Online Appendix to "Optimal Signaling of Content Accuracy: Engagement vs. Misinformation"

Ozan Candogan and Kimon Drakopoulos

A. Proofs for Section 3

In this section, we first briefly discuss some technical assumptions we impose on signaling mechanisms and agents' strategies. Then, we provide proofs of the claims stated in Section 3.

A.1. Technical Assumptions on Signaling Mechanisms and Agents' Strategies

Denote by $(\Omega, \Sigma, \mathbb{P})$ the underlying probability space, where Ω denotes the set of possible outcomes, Σ denotes the σ-algebra, and $\mathbb{P}$ denotes the probability measure. We assume that $\Omega = \Omega_e \times \Omega_r$, where $\Omega_e = [0, \alpha]$ captures the possible realizations of the inaccuracy level and Ω_r denotes the possible outcomes associated with the randomization device that the platform uses to generate the (randomized) signals. We further assume that Σ is the smallest σ-algebra containing $\Sigma_e \times \Sigma_r$, where Σ_e is the Borel σ-algebra associated with $\Omega_e = [0, \alpha]$ and Σ_r is the σ-algebra associated with Ω_r. The probability measure $\mathbb{P}$ is assumed to be a product measure $\mathbb{P}_e \times \mathbb{P}_r$. Here $\mathbb{P}_r$ is the probability measure associated with the randomization device. The probability measure $\mathbb{P}_e$ is such that $\mathbb{P}_e(A_e) = \frac{1}{\alpha}\bar{m}(A_e)$ for any $A_e \in \Sigma_e$, where $\bar{m}$ denotes the Lebesgue measure. In what follows, with some abuse of notation, we denote by $\int_A f(z)dz$ the Lebesgue integral of a function f over a set A.

The random variables $\{s_i\}_{i \in V}$ and ϵ are defined on this probability space, i.e., they are measurable functions from Ω to reals (where, as usual, we assume that the latter is equipped with the Borel σ-algebra). In particular, for $\omega = (z, \omega_r) \in \Omega$, we have $\epsilon(\omega) = z$ and $s_i(\omega) = s_i(z, \omega_r)$. Note that together with the definition of $\mathbb{P}_e$ this guarantees that ϵ admits a uniform distribution on $[0, \alpha]$.

We emphasize that the common underlying probability space assumption imposes a mild but nontrivial restriction on F. For instance, suppose F is such that it (deterministically) sets s_i to 1 for $\epsilon \in \bar{B}$, and 0 otherwise, for some set $\bar{B}$ that is not a Borel set. Then, it can be seen that $s_i : \Omega \to \mathbb{R}$ is not a measurable function. Hence, such F violates our assumption and does not yield a valid signaling mechanism. Note that for a valid signaling mechanism, the collection $\{F(\cdot;\bar{\epsilon})\}_{\bar{\epsilon} \in [0,\alpha]}$ characterizes the conditional distribution of $\{s_i\}$ given ϵ.[18]

[18] We could alternatively define a signaling mechanism as a collection of measurable functions $\{s_i\}$ defined on the probability space $(\Omega, \Sigma, \mathbb{P})$. The distribution of these random variables conditional on ϵ would yield the distributions $\{F(\cdot;\bar{\epsilon})\}_{\bar{\epsilon} \in [0,\alpha]}$.

We assume that the signals are integrable, i.e., $\mathbb{E}[|s_i|] < \infty$ for all $i \in V$. Furthermore, we assume that agents' strategies $\{\mu_i\}$ are measurable, and in particular $\mu_i^{-1}(0) \subset S_i \subset \mathbb{R}$ and $\mu_i^{-1}(1) \subset S_i \subset \mathbb{R}$ are Borel sets. Observe that under our assumptions, the conditional expectation of each element of the collection $\{s_1, \dots, s_n, \epsilon\}$ with respect to any element of $\{s_1, \dots, s_n\}$ is well-defined (and finite).

A.2. Proofs

We start by providing an auxiliary lemma which establishes that the performance of *any* mechanism (as opposed to only the optimal ones in (3)) can be matched using a straightforward mechanism. This result, analogous to known results in the information design literature (see, e.g., Kamenica and Gentzkow (2011)), implies the optimality of straightforward mechanisms. However, is tailored to our setting which features multiple receivers and a continuum of states.

LEMMA 2. *Consider a signaling mechanism that achieves engagement equal to L and total misinformation equal to M. Then, there exists a straightforward mechanism that achieves total engagement equal to L and total misinformation equal to M at its straightforward equilibrium.*

Proof: Let $F \in \mathcal{F}$ denote a given mechanism, and assume that equilibrium $Q \in \mathcal{Q}(F)$ achieves engagement L and misinformation M. Denote by S_i^k the set of signals that at equilibrium Q induce the action $x_i = k$ for $k \in \{0,1\}$, i.e., $S_i^k = \{s \in S_i : \mu_i(s) = k\}$. Consider the mechanism $F^\star \in \mathcal{F}$, for which the signals $s_i^\star : \Omega \to S_i$ are such that

$$s_i^\star(\omega) = 1, \qquad \text{for } \omega \text{ such that } s_i(\omega) \in S_i^1 \text{ and}$$

$$s_i^\star(\omega) = 0, \qquad \text{for } \omega \text{ such that } s_i(\omega) \in S_i^0.$$

In other words, all signals that induce action 1 (action 0) in mechanism F are collected to a single recommendation $s_i^\star = 1$ ($s_i^\star = 0$) in $F^\star$.

We next prove that the constructed mechanism $F^\star$ is straightforward, i.e., it is an equilibrium for each agent i to use the strategy $\mu_i^\star(s_i^\star) = s_i^\star$. Assume first that $\mathbb{P}(s_i \in S_i^k) > 0$ and hence $\mathbb{P}(s_i^\star = k) > 0$ for all k, i. Suppose that all agents other than i use $\{\mu_j^\star\}_{j \neq i}$. Thus, by construction, we have $\mu_j^\star(s_j^\star(\omega)) = s_j^\star(\omega) = \mu_j(s_j(\omega))$ for all $\omega \in \Omega$, $j \neq i$.

The equilibrium conditions imply that in Q, for equilibrium strategy profile $\{\mu_j\}_j$, we have

$$0 \leq \mathbb{E}\left[v - b\epsilon + \sum_j g_{ij}\mu_j(s_j) \mid s_i \in S_i^1\right] = \mathbb{E}\left[v - b\epsilon + \sum_j g_{ij}\mu_j^\star(s_j^\star) \mid s_i \in S_i^1\right], \tag{13}$$

and similarly

$$0 \geq \mathbb{E}\left[v - b\epsilon + \sum_j g_{ij}\mu_j(s_j) \mid s_i \in S_i^0\right] = \mathbb{E}\left[v - b\epsilon + \sum_j g_{ij}\mu_j^\star(s_j^\star) \mid s_i \in S_i^0\right]. \tag{14}$$

Noting that $\{\omega | s_i(\omega) \in S_i^1\} = \{\omega | s_i^\star(\omega) = 1\}$ and $\{\omega | s_i(\omega) \in S_i^0\} = \{\omega | s_i^\star(\omega) = 0\}$ by construction, these inequalities immediately imply that

$$\mathbb{E}\left[v - b\epsilon + \sum_j g_{ij}\mu_j^\star(s_j^\star) \mid s_i^\star = 0\right] \leq 0 \leq \mathbb{E}\left[v - b\epsilon + \sum_j g_{ij}\mu_j^\star(s_j^\star) \mid s_i^\star = 1\right]. \tag{15}$$

Thus, it follows that in $F^\star$, if agent i receives a signal to engage (not engage), then her payoff is maximized by engaging (not engaging) with the content. Hence, $\{\mu_i^\star\}$ is a straightforward equilibrium and $F^\star$ is a straightforward mechanism.

If $\mathbb{P}(s_i \in S_i^k) = 0$ for some k, i, then, under $F^\star$, the corresponding agents almost surely receive either signal 0 or 1 and, to characterize the equilibrium, only one of the inequalities in (15) suffices. In addition, only one of (13) and (14) hold for equilibrium Q which establishes the optimality of following the aforementioned signal. Thus, straightforwardness of $F^\star$ readily follows.

Finally, by construction, the engagement and the misinformation achieved by the straightforward equilibrium of the new mechanism $F^\star$ are the same as in the equilibrium Q of F since in both mechanisms, all agents engage for precisely the same $\omega \in \Omega$. Hence, the claim follows. □

Proof of Theorem 1. The second part of the theorem follows immediately from the first part, i.e., if a straightforward threshold mechanism yields same level of engagement as a given mechanism and lower misinformation, then whenever a solution to (3) exists, there exists a solution that is a straightforward threshold mechanism. Thus, we focus on establishing the first part.

Given an arbitrary mechanism F_0, by Lemma 2, it follows that there exists a straightforward mechanism F with associated random variables $\{s_i\}$ whose straightforward equilibrium achieves the same level of engagement and misinformation. Note that F need not be a threshold mechanism itself. For each $i \in V$, we construct the following thresholds:

$$\sigma_i = \alpha \mathbb{P}(s_i = 1). \tag{16}$$

We denote by $\hat{F}$ the corresponding threshold mechanism, where $\{\hat{s}_i = \mathbf{1}(\epsilon \leq \sigma_i)\}$. We claim that $\hat{F}$ is a straightforward threshold mechanism whose straightforward equilibrium achieves the same level of engagement as F (and hence F_0), but yields weakly lower misinformation. In what follows, we first assume that $1 > \mathbb{P}(s_i = 1) > 0$ for all i. We then treat the case where $\mathbb{P}(s_i = 1) \in \{0, 1\}$ separately.

Recall that signaling mechanism F is a collection of joint (cumulative) probability distributions $F(\cdot; z)$ of random variables $(s_i, \ldots, s_n)$ for $z \in [0, \alpha]$. For the remainder of this proof, for any $y_i \in S_i = \{0, 1\}$, we define:

$$f_i(y_i; z) := \begin{cases} F(0, \mathbf{1}_{-i}; z) & \text{if } y_i = 0, \\ F(1, \mathbf{1}_{-i}; z) - F(0, \mathbf{1}_{-i}; z) & \text{otherwise,} \end{cases}$$

where $\mathbf{1}_{-i}$ is a vector in $\times_{k \neq i} S_k$ whose entries are all equal to one. Note that $f_i(y_i; z)$ captures the probability that agent i receives signal $y_i \in S_i$ when the inaccuracy level is $\epsilon = z$.

We start by showing that straightforwardness of F implies that $\hat{F}$ is also straightforward. To this end, we first focus on the case when the platform sends a signal $\hat{s}_i = 1$ to agent i, and we establish that when agents $j \neq i$ follow strategies $\hat{\mu}_j(\hat{s}_j) = \hat{s}_j$, the agent i finds it optimal to follow the platform's signal and engage. Then, we show that when the platform sends a signal $\hat{s}_i = 0$, the agent finds it optimal not to engage. Jointly, these claims imply that the strategy profile $\hat{\mu}_j(\hat{s}_j) = \hat{s}_j$ for all $j \in V$ is a straightforward equilibrium.

Assume that $j \neq i$ follow strategies $\hat{\mu}_j(\hat{s}_j) = \hat{s}_j$, and consider the case where agent i receives signal $\hat{s}_i = 1$. We claim that for agent i, the expected value of ϵ conditioned on $\hat{s}_i = 1$ from the mechanism $\hat{F}$ is weakly smaller than that from the mechanism F under the same recommendation (i.e., conditioned on $s_i = 1$). To establish this result, we make use of two auxiliary lemmata, which are stated next:

Lemma 3. *Consider a (Lebesgue) measurable function $g : [0, \alpha] \to [0, 1]$, and let us write*

$$A = \int_{[0,\alpha]} g(x) dx.$$

Then,

$$\int_{[0,A]} x dx \leq \int_{[0,\alpha]} x g(x) dx \leq \int_{[\alpha - A, \alpha]} x dx. \tag{17}$$

Proof: Both inequalities follow by observing that the middle integral in (17) is minimized/maximized when $g(\cdot)$ is a threshold-type function (i.e., equal to zero below/above a threshold and equal to one otherwise). □

Lemma 4. *The functions $f_i(\cdot; z)$ satisfy the following:*

$$\int_{[0,\alpha]} f_i(1; z) dz = \sigma_i, \tag{18}$$

$$\int_{[0,\alpha]} f_i(0;z)dz = \alpha - \sigma_i, \tag{19}$$

$$\int_{[0,\alpha]} z f_i(1;z)dz = \sigma_i \mathbb{E}[\epsilon | s_i = 1], \tag{20}$$

$$\int_{[0,\alpha]} z f_i(0;z)dz = (\alpha - \sigma_i)\mathbb{E}[\epsilon | s_i = 0]. \tag{21}$$

Proof: Recall that $(\Omega, \Sigma, \mathbb{P})$ is such that $\Omega = \Omega_e \times \Omega_r$, Σ is the smallest sigma algebra containing $\Sigma_e \times \Sigma_r$, and $\mathbb{P}$ is a product measure between $\mathbb{P}_e$, and $\mathbb{P}_r$. Denote each $\omega \in \Omega$ more explicitly by the tuple $\omega = (z, \omega_r)$ where $z \in \Omega_e$, $\omega_r \in \Omega_r$, and observe that $\epsilon(\omega) = \epsilon(z, \omega_r) = z$. Using Fubini's theorem for any $A \in \Sigma_e$ we have

$$\begin{aligned}\mathbb{P}(s_i = 1, \epsilon \in A) &= \int \mathbf{1}(s_i(\omega) = 1, \epsilon(\omega) \in A)\mathbb{P}(d\omega) = \int \frac{1}{\alpha} \int \mathbf{1}(s_i(z, \omega_r) = 1, z \in A)\mathbb{P}_r(d\omega_r)dz \\ &= \int_A \frac{1}{\alpha} f_i(1;z)dz.\end{aligned} \tag{22}$$

Rearranging terms and using this identity for $A = [0, \alpha]$ yields

$$\int_{[0,\alpha]} f_i(1;z)dz = \alpha \mathbb{P}(s_i = 1) = \sigma_i. \tag{23}$$

Repeating the same argument, we also obtain

$$\mathbb{P}(s_i = 0, \epsilon \in A) = \int_A \frac{1}{\alpha} f_i(0;z)dz \text{ and} \tag{24}$$

$$\int_{[0,\alpha]} f_i(0;z)dz = \alpha \mathbb{P}(s_i = 0) = \alpha - \sigma_i. \tag{25}$$

From (23) and (25) equalities in (18) and (19) follow.

Next observe that

$$\mathbb{E}[\epsilon \mid s_i = 1] = \frac{\mathbb{E}[\epsilon \mathbf{1}(s_i = 1)]}{\mathbb{P}(s_i = 1)} = \frac{\alpha \mathbb{E}[\epsilon \mathbf{1}(s_i = 1)]}{\sigma_i}. \tag{26}$$

On the other hand, once again using Fubini's theorem, we obtain

$$\mathbb{E}[\epsilon \mathbf{1}(s_i = 1)] = \int \epsilon(\omega)\mathbf{1}(s_i(\omega) = 1)\mathbb{P}(d\omega) = \int \frac{1}{\alpha} \int z \mathbf{1}(s_i(z, \omega_r) = 1)\mathbb{P}_r(d\omega_r)dz = \int_{[0,\alpha]} \frac{1}{\alpha} z f_i(1;z)dz.$$

Together with (26), this implies (20).

Finally, following the same approach, we have

$$\mathbb{E}[\epsilon \mid s_i = 0] = \frac{\mathbb{E}[\epsilon \mathbf{1}(s_i = 0)]}{\mathbb{P}(s_i = 0)} = \frac{\alpha \mathbb{E}[\epsilon \mathbf{1}(s_i = 0)]}{\alpha - \sigma_i}, \tag{27}$$

and also

$$\mathbb{E}[\epsilon\mathbf{1}(s_i=0)] = \int \epsilon(\omega)\mathbf{1}(s_i(\omega)=0)\mathbb{P}(d\omega) = \int \frac{1}{\alpha}\int z\mathbf{1}(s_i(z,\omega_r)=0)\mathbb{P}_r(d\omega_r)dz = \int_{[0,\alpha]} \frac{1}{\alpha} z f_i(0;z)dz.$$

Hence, using the last equality together with (27), we also have (21), and the claim follows. □

Using these auxiliary lemmata, we are ready to characterize the outcome when the platform sends signal $\hat{s}_i = 1$ under mechanism $\hat{F}$. Observe that we have

$$\mathbb{E}[\epsilon \mid \hat{s}_i = 1] = \frac{\sigma_i}{2} = \frac{\int_{[0,\sigma_i]} z dz}{\sigma_i} \leq \frac{\int_{[0,\alpha]} z f_i(1;z)dz}{\sigma_i} = \frac{\sigma_i \mathbb{E}[\epsilon | s_i = 1]}{\sigma_i} = \mathbb{E}[\epsilon \mid s_i = 1]. \tag{28}$$

Here, the inequality follows from Lemma 3 and the fact that $\int_{[0,\alpha]} f_i(1;z)dz = \sigma_i$, as established in Lemma 4. The third equality also follows from Lemma 4.

In addition, for each pair $i, j \in V$, using (16), we write

$$\begin{aligned}\mathbb{P}(s_j = 1 \mid s_i = 1) &= \frac{\mathbb{P}(s_j = 1, s_i = 1)}{\mathbb{P}(s_i = 1)} \leq \frac{\min\{\mathbb{P}(s_i = 1), \mathbb{P}(s_j = 1)\}}{\sigma_i/\alpha} \\ &= \frac{\min\{\sigma_i, \sigma_j\}}{\sigma_i} = \mathbb{P}(\hat{s}_j = 1 \mid \hat{s}_i = 1),\end{aligned} \tag{29}$$

where the inequality follows from the fact that $\mathbb{P}(s_j = 1, s_i = 1) \leq \mathbb{P}(s_j = 1)$ and $\mathbb{P}(s_j = 1, s_i = 1) \leq \mathbb{P}(s_i = 1)$, for all $i, j \in V$.

It follows from (28) and (29) that when the platform sends a signal to engage, the payoff from engaging is higher under $\hat{F}$ than it is under F. This is the case because under $\hat{F}$, both the disutility due to inaccuracy is lower and the network externality due to having connections who engage with the same content is higher.

Next, consider the case where $\hat{s}_i = 0$, i.e., the platform sends a "Do not engage" signal to agent i in mechanism $\hat{F}$. We can express the associated expected inaccuracy level as follows:

$$\mathbb{E}[\epsilon \mid \hat{s}_i = 0] = \frac{\alpha + \sigma_i}{2} = \frac{\int_{[\sigma_i,\alpha]} \sigma d\sigma}{\alpha - \sigma_i} \geq \frac{\int_{[0,\alpha]} z f_i(0;z)dz}{\alpha - \sigma_i} = \mathbb{E}[\epsilon \mid s_i = 0]. \tag{30}$$

Here, the inequality follows from Lemma 3 and the fact that $\int_{[0,\alpha]} f_i(0;z)dz = \alpha - \sigma_i$, as established in Lemma 4. The last equality is also a direct consequence of Lemma 4.

Similarly, for each pair $i,j \in V$, we obtain

$$\begin{aligned}\mathbb{P}(s_j = 1 \mid s_i = 0) &= \frac{\mathbb{P}(s_j = 1, s_i = 0)}{\mathbb{P}(s_i = 0)} = \frac{\mathbb{P}(s_j = 1) - \mathbb{P}(s_j = 1, s_i = 1)}{(\alpha - \sigma_i)/\alpha} \\ &\geq \frac{\sigma_j - \alpha \min\{\mathbb{P}(s_j = 1), \mathbb{P}(s_i = 1)\}}{\alpha - \sigma_i} \\ &= \frac{\sigma_j - \min\{\sigma_i, \sigma_j\}}{\alpha - \sigma_i} = \frac{\max\{\sigma_i, \sigma_j\} - \sigma_i}{\alpha - \sigma_i} = \mathbb{P}(\hat{s}_j = 1 \mid \hat{s}_i = 0).\end{aligned} \tag{31}$$

Here, the inequality follows again from the fact that $\mathbb{P}(s_j = 1, s_i = 1) \leq \mathbb{P}(s_j = 1)$ and $\mathbb{P}(s_j = 1, s_i = 1) \leq \mathbb{P}(s_i = 1)$, for all $i,j \in V$.

It follows from (30) and (31) that when the platform sends a signal to not to engage, the payoff from engaging is lower under $\hat{F}$ than it is under F. This is the case, because under $\hat{F}$ both the disutility due to inaccuracy level is larger and the network externality due to having connections who engage with the same content is smaller.

Since under $\hat{F}$ the payoff due to engagement under the engage (do not engage) signal is larger (lower) than F, and since F is straightforward, it follows that $\hat{F}$ is also straightforward. We conclude the proof by establishing that the engagement achieved at a straightforward equilibrium $\hat{Q}$ of $\hat{F}$ is the same as that achieved at a straightforward equilibrium Q of F. However, the misinformation associated with $\hat{Q}$ is weakly lower than that associated with Q.

To see this, first note that the total engagement level for the original mechanism satisfies

$$E(Q) = \sum_{i \in V} \mathbb{P}(s_i = 1) = \sum_{i \in V} \frac{\sigma_i}{\alpha} = \sum_{i \in V} \mathbb{P}(\hat{s}_i = 1) = E(\hat{Q}),$$

where the first and last equalities follow from the definition of engagement and straightforwardness of equilibria Q and $\hat{Q}$, the second equality follows from (16) and the third equality follows from the construction of the thresholds.

Second, observe that

$$M(Q) = \sum_{i \in V} \mathbb{E}[\epsilon | s_i = 1] = \sum_{i \in V} \int_{[0,\alpha]} \frac{1}{\sigma_i} z f_i(1; z) dz \geq \sum_{i \in V} \frac{1}{\sigma_i} \int_{[0,\sigma_i]} z dz = \sum_{i \in V} \mathbb{E}[\epsilon | \hat{s}_i = 1] = M(\hat{Q}),$$

where the first and last equalities follow from the definition of $M(\cdot)$ and the fact that Q and $\hat{Q}$ are straightforward equilibria, the second equality follows from Lemma 4, the third equality follows from the definition of threshold mechanisms, and the inequality follows from Lemma 3.

Thus, we conclude that the straightforward equilibrium of the mechanism $\hat{F}$ yields the same engagement and weakly lower misinformation. Hence, the claim follows.

Finally, observe that if $\mathbb{P}(s_i = 1) \in \{0, 1\}$ for some i, then $\sigma_i \in \{0, \alpha\}$. Hence, almost surely both F and $\hat{F}$ send the same signals to such agents. Since F has a straightforward equilibrium where these agents find it optimal to follow the signal they receive, in $\hat{F}$ aforementioned agents find it optimal to follow their signals as well. Thus, applying the argument above to the remaining agents, it follows that $\hat{F}$ is a straightforward mechanism. In addition, as before, the engagement associated with these agents is the same in both mechanisms and the misinformation is weakly lower. Hence, the claim follows as before. □

Proof of Proposition 1. Suppose that agents other than i follow the recommendation they receive, and consider agent i. Assume that $\sigma_i > 0$ and first focus on the recommendation to "Engage" ($s_i = 1$). The expected utility of agent i from engaging consists of three terms: the satisfaction term v; the inaccuracy disutilty term, which is equal to

$$-b\mathbb{E}[\epsilon \mid s_i = 1] = -b\mathbb{E}[\epsilon \mid \epsilon \leq \sigma_i] = -b\frac{\sigma_i}{2};$$

and the externality term, which is equal to

$$\sum_{j\in V} g_{ij}\mathbb{P}(x_j = 1 \mid s_i = 1) = \sum_{j\in V} g_{ij}\mathbb{P}(s_j = 1 \mid s_i = 1) = \sum_{j\in V} g_{ij}\frac{\min\{\sigma_i, \sigma_j\}}{\sigma_i},$$

where we make use of the fact that agents $j \neq i$ follow the signal of the platform. Therefore, it is incentive compatible for agent i to follow the "Engage" recommendation if and only if $v - b\frac{\sigma_i}{2} + \sum_{j\in V} g_{ij}\frac{\min\{\sigma_i,\sigma_j\}}{\sigma_i} \geq 0$, or equivalently if and only if (4) holds. If $\sigma_i = 0$, then almost surely agent i receives the signal $s_i = 0$. Hence, for equilibrium characterization, her action under the "Engage" recommendation ($s_i = 1$) is irrelevant. In addition, for this case, (4) trivially holds.

Similarly, assume that $\sigma_i < \alpha$, and consider the recommendation "Do not engage" ($s_i = 0$). Once again, agent i's utility from engaging consists of three terms: the satisfaction term v; the penalty term, which is now equal to

$$-b\mathbb{E}[\epsilon \mid s_i = 0] = -b\mathbb{E}[\epsilon \mid \epsilon > \sigma_i] = -b\frac{\alpha + \sigma_i}{2};$$

and the externality term, which is equal to

$$\sum_{j\in V} g_{ij}\mathbb{P}(x_j = 1 \mid s_i = 0) = \sum_{j\in V} g_{ij}\mathbb{P}(s_j = 1 \mid s_i = 0) = \sum_{j\in V} g_{ij}\frac{\max\{\sigma_i, \sigma_j\} - \sigma_i}{\alpha - \sigma_i}.$$

Therefore, it is incentive compatible for agent i to follow the "Do not engage" recommendation if and only if (5) holds. If, on the other hand, $\sigma_i = \alpha$, then agent i surely receives the signal

$s_i = 1$. Hence, in this case, the agent's action under the engage recommendation is not relevant for equilibrium characterization. Moreover, in this case (5) trivially holds.

Thus, we conclude that agent i finds it optimal to follow her recommendation if and only if the thresholds satisfy (4) and (5). Equivalently, a mechanism is straightforward if these conditions hold for any $i \in V$, and the claim follows. □

Proof of Theorem 2.

Before we establish the proof of the theorem, we introduce some necessary definitions, and auxiliary results.

When studying common threshold mechanisms (with $\sigma \in (0,\alpha)$), it is convenient to introduce the following functions, which capture the expected payoffs from different actions conditional on the public signal of the platform:[19]

$$u_i(x_i, x_{-i}, s) \triangleq \begin{cases} \mathbf{1}(x_i = 1) \cdot \mathbb{E}\left[v - b\sigma + \sum_j g_{ij} x_j | \epsilon \leq \sigma\right] & \text{for } s = 1, \\ \mathbf{1}(x_i = 1) \cdot \mathbb{E}\left[v - b\sigma + \sum_j g_{ij} x_j | \epsilon > \sigma\right] & \text{for } s = 0. \end{cases} \tag{32}$$

Note that, by definition we have the following: $u_i(x_i, x_{-i}, s) = \mathbb{E}[\tilde{U}_i(x_i, x_{-i}) \mid \epsilon \leq \sigma]$ if $s = 1$, and $u_i(x_i, x_{-i}, s) = \mathbb{E}[\tilde{U}_i(x_i, x_{-i}) \mid \epsilon > \sigma]$ if $s = 0$.

Observe that in common threshold mechanisms, either all agents receive the "Engage" signal or no agent receives the "Engage" signal. Thus, depending on the signal, two *complete information* games with different payoffs are induced. In particular, let $\mathcal{G}_0$ and $\mathcal{G}_1$ be normal-form games with a set of players V and strategy sets $\{0,1\}$ for each $i \in V$. Assume that the payoffs of agents in $\mathcal{G}_0$ are given by $\{u_i(x_i, x_{-i}, 0)\}$ and that in $\mathcal{G}_1$, they are given by $\{u_i(x_i, x_{-i}, 1)\}$. In our setting, if the public signal is $s = 1$, then agents play $\mathcal{G}_1$, and otherwise they play $\mathcal{G}_0$.

We next establish that $\mathcal{G}_0$ and $\mathcal{G}_1$ are supermodular games. We start with a definition of supermodular games, adapted to our setting.

Definition 6. Let $\mathcal{H} = \langle \mathcal{N}, \{\tilde{S}_i\}, \{\tilde{u}_i\} \rangle$ be a game with set of players $\mathcal{N}$, compact set of strategies $\tilde{S}_i \subset \mathbb{R}$ and continuous payoff functions $\tilde{u}_i : \times_{i \in \mathcal{N}} \tilde{S}_i \to \mathbb{R}$ for each $i \in \mathcal{N}$. We say that $\mathcal{H}$ is a **supermodular game** if $\tilde{u}_i(x_i, x_{-i})$ has increasing differences in (x_i, x_{-i}).[20] That is, $\tilde{u}_i(x'_i, x_{-i}) -$

[19] Here, we restrict attention to $\sigma \in (0,\alpha)$ to avoid conditioning on measure zero events. If $\sigma \in \{0,\alpha\}$, the signals are uninformative as the common threshold mechanism either almost surely recommends engagement or no engagement. In such cases, for our analysis of performance levels induced under common threshold mechanisms, it suffices to restrict attention to signals that have a non-zero probability of being sent by the platform.

[20] Note that by construction, the strategy sets of agents are sublattices of $\mathbb{R}$, and payoff functions are continuous. Both of these assumptions can be relaxed, e.g., it suffices to consider more general sublattices (of $\mathbb{R}^{m_i}$ for $m_i \geq 1$), and semicontinuity of payoffs in an agent's own actions (together with continuity in other agents' actions) suffice.

$\tilde{u}_i(x_i, x_{-i}) \leq \tilde{u}_i(x'_i, x'_{-i}) - \tilde{u}_i(x_i, x'_{-i})$ for $x_i \leq x'_i$ and $x_{-i} \leq x'_{-i}$ (where the last inequality is entrywise).

LEMMA 5. *$\mathcal{G}_0$ and $\mathcal{G}_1$ are supermodular games.*

Proof: Observe that in both games, the strategies of agents belong to $\{0,1\}$, and hence, the strategy sets are compact. Since the domains of the payoff functions are discrete, continuity follows trivially. Moreover, the payoffs exhibit increasing differences, i.e.,

$$u_i(1, x_{-i}, s) - u_i(0, x_{-i}, s) \leq u_i(1, x'_{-i}, s) - u_i(0, x'_{-i}, s),$$

where $x'_{-i} \geq x_{-i}$ and $s \in \{0,1\}$. The fact that the inequality holds can trivially be seen by noting that when $x'_{-i} \geq x_{-i}$, the network externality term is weakly larger. Thus, it follows that for both $s=0$ and $s=1$, the induced game is supermodular. □

We now proceed with the proof of Theorem 2. To prove the result, we first characterize the equilibrium strategies under mechanism F in terms of the pure strategy equilibria of the associated supermodular games. Then we use this characterization to study the performance levels of F at different equilibria and their convex hull.

If the platform chooses a threshold σ, when $\epsilon \leq \sigma$, agents play $\mathcal{G}_1$, and otherwise they play $\mathcal{G}_0$. Let $\mathcal{X}_0$ denote the set of pure equilibria of $\mathcal{G}_0$ and $\mathcal{X}_1$ denote the set of pure equilibria of $\mathcal{G}_1$. With some abuse of notation we denote any equilibrium X of $\mathcal{X}_i$ ($i \in \{0,1\}$) with a vector in $\mathbb{R}^n$, where the j-th entry of the vector, X_j, denotes the equilibrium action of agent j. By Lemma 5 it follows that both of these sets admit a largest and a smallest equilibrium (with respect to the usual order in $\mathbb{R}^n$). We denote the largest/smallest equilibria in $\mathcal{X}_i$ respectively by $\overline{X}^i$ and $\underline{X}^i$ for $i \in \{0,1\}$. Using the vector representation of equilibria, for any $X \in \mathcal{X}_i$, we have $\underline{X}^i \leq X \leq \overline{X}_i$, where the inequalities are entrywise. We start by expressing any equilibrium $Q \in \mathcal{Q}(F)$ in terms of $X^0 \in \mathcal{X}_0$ and $X^1 \in \mathcal{X}_1$.

LEMMA 6. *Suppose that the platform uses a common threshold mechanism F with threshold $\sigma \in (0, \alpha)$. Denote the equilibrium strategy profile for any $Q \in \mathcal{Q}(F)$ by $\mu^Q := \{\mu_i^Q\}$, and note that agents' equilibrium actions are given by $Y_0 = \mu^Q(0)$ and $Y_1 = \mu^Q(1)$ for signals $s=0$ and $s=1$, respectively. Let $\mathcal{Y} = \{(Y_0, Y_1) \mid Y_0 = \mu^Q(0), Y_1 = \mu^Q(1), Q \in \mathcal{Q}(F)\}$. Then, we have:*

$$\mathcal{Y} = \mathcal{X} \triangleq \{(X^0, X^1) | X^0 \in \mathcal{X}_0, X^1 \in \mathcal{X}_1\}.$$

Proof: Consider any $Q \in \mathcal{Q}(F)$. We claim that $\mu^Q(0)$ is an equilibrium in $\mathcal{G}_0$. Suppose that this is not the case and that agent i can unilaterally deviate to x_i' and strictly improve her payoff in $\mathcal{G}_0$, i.e., $u_i(x_i', \mu_{-i}^Q(0), 0) > u_i(\mu_i^Q(0), \mu_{-i}^Q(0), 0)$. Let μ_i' denote a strategy of agent i in the original game, such that $\mu_i'(0) = x_i'$ and $\mu_i'(1) = \mu_i^Q(1)$. We have

$$u_i(\mu_i'(0), \mu_{-i}^Q(0), 0) > u_i(\mu_i^Q(0), \mu_{-i}^Q(0), 0),$$

and $u_i(\mu_i^Q(1), \mu_{-i}^Q(1), 1) = u_i(\mu_i'(1), \mu_{-i}^Q(1), 1)$. Thus, we conclude that $\{\mu_i^Q\}_i$ does not satisfy (2), and we obtain a contradiction to the fact that Q is an equilibrium under mechanism F. Hence, $\mu^Q(0)$ is an equilibrium in $\mathcal{G}_0$. Following the same approach, it also follows that $\mu^Q(1)$ is an equilibrium in $\mathcal{G}_1$. Hence, we conclude that $\mathcal{Y} \subset \mathcal{X}$.

Conversely, consider any $X^0 \in \mathcal{X}_0, X^1 \in \mathcal{X}_1$. Let μ be the strategy profile that satisfies $\mu(0) = X^0$, and $\mu(1) = X^1$. Suppose that this is not an equilibrium of the original game under mechanism F and that agent i strictly improves her payoff by deviating to strategy μ_i'. Then, (2) implies that agent i strictly improves her expected payoff conditional on signal realization $s = 0$ or $s = 1$ by using μ_i'. Consider the case $s = 0$; then deviating from X^0, using $x_i' = \mu_i'(s)$ strictly improves the payoff of agent i in $\mathcal{G}_0$, contradicting the assumption that X^0 is an equilibrium. The case of $s = 1$ similarly contradicts the assumption that X^1 is an equilibrium. Thus, it follows that $\mathcal{Y} \supset \mathcal{X}$. Since we also have $\mathcal{Y} \subset \mathcal{X}$, the claim follows. □

Having characterized the structure of the set of equilibria that are induced by common threshold mechanisms, we now proceed to analyze the structure of the performance levels that are achievable by these mechanisms. Consider a signaling mechanism F and an equilibrium $Q \in \mathcal{Q}(F)$. Recall that the performance of Q is given by $P(Q) = (E(Q), M(Q))$, and performance levels achievable at an equilibrium of F are denoted by $\mathcal{P}(F) = \{P(Q) | Q \in \mathcal{Q}(F)\}$.

Suppose that $\sigma \in (0, \alpha)$. As established in Lemma 6, the equilibrium actions of agents in a common threshold mechanism F are given as product sets of equilibria of corresponding supermodular games $\mathcal{G}_0$ and $\mathcal{G}_1$. We next establish that $\mathcal{P}(F)$ is contained in the convex hull of four equilibria associated with these supermodular games. Before proving our result, we define these equilibria:

– $Q^{(B,B)}$ with strategy profile $\mu^{(B,B)}$: $\mu^{(B,B)}(0) = \overline{X}^0$ and $\mu^{(B,B)}(1) = \overline{X}^1$: this is the strategy profile where for each signal realization, agents' actions constitute the best (largest) equilibrium of the corresponding supermodular game.

– $Q^{(B,W)}$ with strategy profile $\mu^{(B,W)}$: $\mu^{(B,W)}(0) = \overline{X}^0$ and $\mu^{(B,W)}(1) = \underline{X}^1$: this is the strategy profile where for $s = 0$, agents' actions constitute the best (largest) equilibrium of the corresponding supermodular game, while for $s = 1$, agents' actions constitute the worst (smallest) equilibrium.

– $Q^{(W,B)}$ with strategy profile $\mu^{(W,B)}$: $\mu^{(W,B)}(0)=\underline{X}^0$ and $\mu^{(W,B)}(1)=\overline{X}^1$: this is the strategy profile where for $s=0$, agents' actions constitute the worst (smallest) equilibrium of the corresponding supermodular game, while for $s=1$, agents' actions constitute the best (largest) equilibrium.

– $Q^{(W,W)}$ with strategy profile $\mu^{(W,W)}$: $\mu^{(W,W)}(0)=\underline{X}^0$ and $\mu^{(W,W)}(1)=\underline{X}^1$: this is the strategy profile where for each signal realization, agents' actions constitute the worst (smallest) equilibrium of the corresponding supermodular game.

Lemma 7. *Suppose that* $\sigma\in(0,\alpha)$, *then*

$$\mathcal{P}(F)\subset\mathcal{C}(F)=\mathcal{C}'(F)\triangleq \text{conv}(\{P(Q^{(B,B)}),P(Q^{(B,W)}),P(Q^{(W,B)}),P(Q^{(W,W)})\}).$$

Proof: To prove the claim, it suffices to establish that $\mathcal{C}(F)=\mathcal{C}'(F)$. From the definition of $\mathcal{C}(F)$, it can be seen that this is equivalent to having $P(Q)\in\mathcal{C}'(F)$ for all $Q\in\mathcal{Q}(F)$.

Consider $Q\in\mathcal{Q}(F)$, and using Lemma 6, let $(X^0,X^1)\in\mathcal{X}_0\times\mathcal{X}_1$ denote the corresponding actions of agents under signals $s=0$ and $s=1$, respectively. Using these sets, engagement and misinformation associated with Q can be explicitly given as follows:

$$\begin{aligned} E(Q)&=p_1|X^1|+p_0|X^0|\\ M(Q)&=p_1|X^1|\frac{\sigma}{2}+p_0|X^0|\frac{\alpha+\sigma}{2}, \end{aligned}$$

where $p_1\triangleq\mathbb{P}(s=1)$, $p_0\triangleq 1-p_1$, and with abuse of notation, we denote by $|X^k|$ the number of nonzero entries of $X^k\in\{0,1\}^n$ for $k\in\{0,1\}$.

Suppose that $P(Q)$ does not belong to $\mathcal{C}'(F)$. Then, by the separating hyperplane theorem, there exist scalars $\gamma_e,\gamma_m\in\mathbb{R}$ such that

$$\gamma_e E(Q)-\gamma_m M(Q)>\gamma_e E'-\gamma_m M' \tag{33}$$

for any $(E',M')\in\mathcal{C}'(F)$. Observe that

$$\gamma_e E(Q)-\gamma_m M(Q)=|X^1|(\gamma_e p_1-\gamma_m p_1\frac{\sigma}{2})+|X^0|(\gamma_e p_0-\gamma_m p_0\frac{\alpha+\sigma}{2}).$$

Let $Y^1=\overline{X}^1$ if $(\gamma_e p_1-\gamma_m p_1\frac{\sigma}{2})\geq 0$, and $Y^1=\underline{X}^1$ otherwise. Similarly, let $Y^0=\overline{X}^0$ if $(\gamma_e p_0-\gamma_m p_0\frac{\alpha+\sigma}{2})\geq 0$, and $Y^0=\underline{X}^0$ otherwise. Observe that by Lemma 6, we have $\underline{X}^1\leq X^1\leq\overline{X}^1$ and, similarly, $\underline{X}^0\leq X^0\leq\overline{X}^0$. Thus, it follows that

$$\gamma_e E(Q)-\gamma_m M(Q)\leq|Y^1|(\gamma_e p_1-\gamma_m p_1\frac{\sigma}{2})+|Y^0|(\gamma_e p_0-\gamma_m p_0\frac{\alpha+\sigma}{2}).$$

On the other hand, by construction, (Y^0, Y^1) is the equilibrium actions for one of the equilibria $Q^{(B,B)}$, $Q^{(B,W)}$, $Q^{(W,B)}$, $Q^{(W,W)}$. Thus, we obtain a contradiction to (33), and the claim follows. □

Lemma 7 characterizes equilibrium performance levels for $\sigma \in (0,\alpha)$. Note that if $\sigma \in \{0,\alpha\}$ agents almost surely receive the signal $s=0$, or $s=1$. Suppose $\sigma=0$, and agents almost surely receive signal $s=0$. In this case, equilibrium play almost surely coincides with an equilibrium of $\mathcal{G}_0$. Thus, it can be seen that $Q^{(W,W)}, Q^{(W,B)}, Q^{(B,W)}, Q^{(B,B)}$ remain to be equilibria for common threshold mechanisms with $\sigma=0$, and we have $P(Q^{(W,W)}) = P(Q^{(W,B)})$, and $P(Q^{(B,W)}) = P(Q^{(B,B)})$. Moreover, since equilibrium play almost surely corresponds to an equilibrium of $\mathcal{G}_0$, we conclude that equilibrium performance levels belong to the convex hull of the performance levels associated with the best/worst equilibria of $\mathcal{G}_0$. These observations imply that equilibrium performance levels once again belong to $\mathcal{C}'(F)$, i.e., $\mathcal{C}(F)=\mathcal{C}'(F)$. Recalling that when $\sigma=\alpha$ agents almost surely receive signal $s=1$, and following a similar argument, it follows that in this case as well $\mathcal{C}(F)=\mathcal{C}'(F)$.

Thus, for any $\sigma \in [0,\alpha]$, equilibrium performance levels for mechanism F, denoted by $\mathcal{P}(F)$, belong to the convex hull of performance levels associated with equilibria: $Q^{(B,B)}, Q^{(B,W)}, Q^{(W,B)}, Q^{(W,W)}$. To finish the proof of Theorem 2, we explicitly characterize the performances associated with $Q^{(W,W)}, Q^{(W,B)}, Q^{(B,W)}, Q^{(B,B)}$.

(i) First, we focus on the case where $v \le b\sigma/2$. Consider $\mathcal{G}_0$, and assume that no neighbors of some agent i engage with the content. Then, agent i's payoff from engaging is given by $v - b\frac{\alpha+\sigma}{2} < v - b\frac{\sigma}{2} \le 0$, and she finds it optimal not to engage. Since i is arbitrary, having no agents engage is an equilibrium. Given that the actions of each agent belong to $\{0,1\}$, we conclude that the worst equilibrium $\underline{X}^0$ of $\mathcal{G}_0$ is such that no agent engages.

Similarly, consider $\mathcal{G}_1$, and assume that no neighbor of agent i engages with the content. Then, agent i's payoff from engaging is given by $v - b\frac{\sigma}{2} \le 0$. In this case as well, the agent finds it optimal not to engage. As before, having no agents engage is an equilibrium. This constitutes the worst equilibrium $\underline{X}^1$ of $\mathcal{G}_1$.

Having characterized the worst equilibria of $\mathcal{G}_0$ and $\mathcal{G}_1$, we next focus on the best equilibria. First consider $\mathcal{G}_1$. We claim that it is an equilibrium to have all agents in A_1 engage and agents in A_1^c not engage. To see this, note that for the described strategy profile for $i \in A_1$, the payoff from engaging is given by $v - b\frac{\sigma}{2} + \sum_{j\in A_1} g_{ij} \ge v - b\frac{\sigma}{2} + k_1 \ge 0$, where the first equality follows from the definition of A_1, and the last one follows from the definition of k_1. Thus, agent i finds it optimal to engage. Conversely, for $i \notin A_1$ the payoff from engaging is given by $v - b\frac{\sigma}{2} + \sum_{j\in A_1} g_{ij} \le v - b\frac{\sigma}{2} + (k_1 - 1) < 0$, where the first inequality follows since $i \notin A_1$ implies that i has $k_1 - 1$ or fewer neighbors in A_1. The second inequality follows from the definition of k_1. Thus, we conclude that agent $i \notin A_1$ finds it optimal not to engage. Hence, it indeed is an equilibrium to have only the agents in A_1 engage.

We next argue that this is the best equilibrium of $\mathcal{G}_1$. Suppose that this is not true, in which case there is an equilibrium where agents in $A \supsetneq A_1$ engage. Let $j \in A \setminus A_1$ be an agent for whom $\sum_{k\in A} g_{jk}$ is minimized among all agents in $A \setminus A_1$. Observe that $\sum_{k\in A} g_{jk} \le k_1 - 1$, as otherwise we obtain a contradiction to the fact that A_1 is the maximal induced subgraph where agents have at least k_1 neighbors. Observe that the payoff of agent j is given by $v - b\frac{\sigma}{2} + \sum_{k\in A} g_{jk} \le v - b\frac{\sigma}{2} + (k_1 - 1) < 0$. Thus, agent j is better off not engaging and hence, we obtain a contradiction. We conclude that at the best equilibrium of $\mathcal{G}_1$, only agents in A_1 engage.

The best equilibrium of $\mathcal{G}_0$ admits an identical characterization after replacing $-b\frac{\sigma}{2}$ with $-b\frac{\sigma+\alpha}{2}$ in the above argument (and using A_0 in place of A_1), accounting for the fact that in $\mathcal{G}_0$, the misinformation penalty term in the payoff is $-b\frac{\sigma+\alpha}{2}$, as opposed to $-b\frac{\sigma}{2}$. We conclude that the best equilibrium of $\mathcal{G}_0$ is such that only agents in A_0 engage.

Given best/worst equilibria of $\mathcal{G}_0$ and $\mathcal{G}_1$, we next characterize the engagement and misinformation associated with $Q^{(B,B)}, Q^{(B,W)}, Q^{(W,B)}, Q^{(W,W)}$. Observe that the platform sends the signal $s=1$ with probability σ/α and the signal $s=0$ with probability $(\alpha-\sigma)/\alpha$. Using this, together with the characterization of best/worst equilibria given above, we obtain the following:

– $Q^{(W,W)}$: Associated equilibrium actions are $\underline{X}_0 = 0$ and $\underline{X}_1 = 0$ for $s=0$ and $s=1$, respectively. Thus, we have $E(Q^{(W,W)}) = 0$ and $M(Q^{(W,W)}) = 0$.

– $Q^{(W,B)}$: Associated equilibrium actions are $\underline{X}_0 = 0$ and $\overline{X}_1 = e_{A_1}$ for $s=0$ and $s=1$, where e_{A_1} is a vector that has one for entries $i \in A_1$ and zero otherwise. Thus, we have $E(Q^{(W,B)}) = \frac{\sigma}{\alpha}|A_1|$ and $M(Q^{(W,B)}) = \frac{\sigma^2}{2\alpha}|A_1|$.

– $Q^{(B,W)}$: Associated equilibrium actions are $\overline{X}_0 = e_{A_0}$ and $\underline{X}_1 = 0$ for $s=0$ and $s=1$, where e_{A_0} has one for entries $i \in A_0$ and zero otherwise. Thus, we have $E(Q^{(B,W)}) = \frac{\alpha-\sigma}{\alpha}|A_0|$ and $M(Q^{(B,W)}) = \frac{\alpha^2-\sigma^2}{2\alpha}|A_0|$.

– $Q^{(B,B)}$: Associated equilibrium actions are $\overline{X}_0 = e_{A_0}$ and $\overline{X}_1 = e_{A_1}$. Thus, we have $E(Q^{(B,B)}) = \frac{\sigma}{\alpha}|A_1| + \frac{\alpha-\sigma}{\alpha}|A_0|$ and $M(Q^{(B,B)}) = \frac{\sigma^2}{2\alpha}|A_1| + \frac{\alpha^2-\sigma^2}{2\alpha}|A_0|$.

Observe that when $0 < \sigma < \alpha$, these points are not collinear provided that $A_0, A_1 \neq \emptyset$. This can be seen by noting that $\frac{M(Q^{(W,B)})}{E(Q^{(W,B)})} = \frac{\sigma}{2}$, whereas $\frac{M(Q^{(B,W)})}{E(Q^{(B,W)})} = \frac{\alpha+\sigma}{2}$. Thus, the region covered by this equilibria has nontrivial area.

(ii) Next, we consider the case with $\frac{b\sigma}{2} < v \le \frac{b(\sigma+\alpha)}{2}$. Note that $v > \frac{b\sigma}{2}$ immediately implies that in $\mathcal{G}_1$, agents always have nonnegative payoff from engaging, even in the absence of network externalities. Thus, $\overline{X}^1 = \underline{X}^1 = e$, where e denotes the vector of ones.

On the other hand, in $\mathcal{G}_0$, the payoffs from engaging are given by $v - b\frac{\sigma+\alpha}{2} + \sum_j g_{ij}x_j$. Thus, in the absence of network effects, we have $v - b\frac{\sigma+\alpha}{2} \le 0$. Hence, in the worst equilibrium, nobody engages and we obtain $\underline{X}_0 = 0$. Also, following the same argument as before, in the best equilibrium, only agents in A_0 engage, i.e., $\overline{X}_0 = e_{A_0}$. Summarizing, when $\frac{b\sigma}{2} < v \le \frac{b(\sigma+\alpha)}{2}$ we have the following:

– $Q^{(W,W)} = Q^{(W,B)}$: Associated equilibrium actions are $\underline{X}^0 = 0$ and $\overline{X}^1 = e$ respectively for $s = 0$ and $s = 1$. Thus, we have $E(Q^{(W,W)}) = E(Q^{(W,B)}) = |V|\frac{\sigma}{\alpha}$ and $M(Q^{(W,W)}) = M(Q^{(W,B)}) = |V|\frac{\sigma^2}{2\alpha}$.

– $Q^{(B,W)} = Q^{(B,B)}$: Associated equilibrium actions are $\underline{X}^0 = e_{A_0}$ and $\overline{X}^1 = e$ respectively for $s = 0$ and $s = 1$. Thus, we have $E(Q^{(B,W)}) = E(Q^{(B,B)}) = |V|\frac{\sigma}{\alpha} + |A_0|\frac{\alpha-\sigma}{\alpha}$ and $M(Q^{(W,W)}) = M(Q^{(W,B)}) = |V|\frac{\sigma^2}{2\alpha} + |A_0|\frac{\alpha^2-\sigma^2}{2\alpha}$.

(iii) Suppose next that $v > \frac{b(\sigma+\alpha)}{2}$. As in case (ii), since $v > \frac{b\sigma}{2}$, in $\mathcal{G}_1$, agents always have incentive to engage and hence we have $\overline{X}^1 = \underline{X}^1 = e$. Consider $\mathcal{G}_0$. Since $v > \frac{b(\sigma+\alpha)}{2}$, similar to $\mathcal{G}_1$, all agents always have incentive to engage even in the absence of any neighbors who engage. Thus, we also have $\overline{X}^0 = \underline{X}^0 = e$. Hence, in this case, we have $Q = Q^{(W,W)} = Q^{(W,B)} = Q^{(B,W)} = Q^{(B,B)}$ and $E(Q) = |V|$, $M(Q) = |V|\frac{\sigma^2}{2\alpha} + |V|\frac{\alpha^2-\sigma^2}{2\alpha} = |V|\frac{\alpha}{2}$. □

A.3. Example: Achievable performance levels for common threshold mechanisms

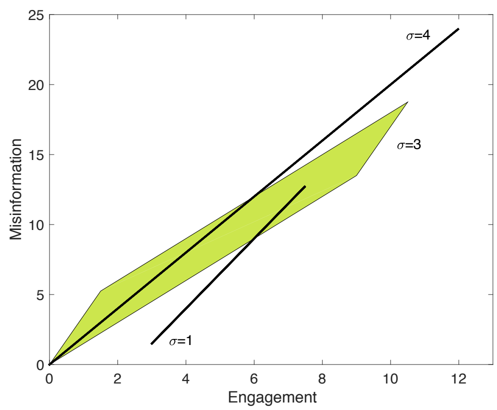
In this section, we illustrate the achievable performance levels for common threshold mechanisms over a simple example.

Consider the network ($\mathcal{G}$) given in Figure 1. This network consists of 12 nodes: half of the nodes are organized in a complete graph (hereafter subgraph $\mathcal{G}_L$), the other half constitute a cycle (hereafter $\mathcal{G}_R$), and two (dashed) edges connect the two halves. Denote the set of nodes of $\mathcal{G}_L$ and $\mathcal{G}_R$ respectively by V_L and V_R.

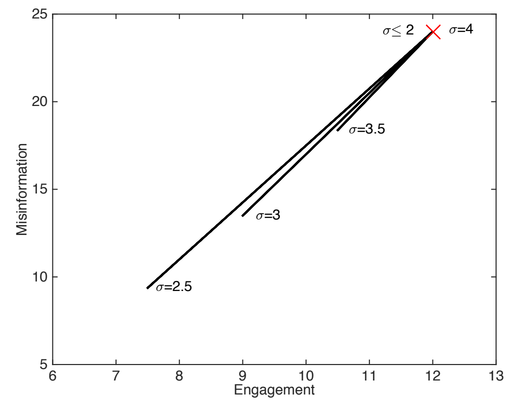
It can be easily checked that given a network $\mathcal{G}$, there do not exist parameters α, b, v such that the cases (i) and (iii) in Theorem 2 jointly emerge. Thus, we analyze the network in Figure 1 for two sets of parameters, where respectively cases (i), (ii) and (ii), (iii) jointly emerge.

In Figure 7(a), we focus on parameters $\alpha = 4, b = 2, v = 2$. Theorem 2 suggests that for these parameters, when $\sigma \geq 2$, equilibrium characterization is given as in case (i), whereas when $\sigma < 2$, case (ii) provides a characterization. Suppose $\sigma = 3$, and note that we have $k_0 = 5$, $k_1 = 1$, and thus, $A_0 = V_L$, whereas $A_1 = V_L \cup V_R$. As given in Theorem 2, in this case, the convex hull of equilibrium performance levels ($\mathcal{C}(F)$) has nontrivial area. This corresponds to the shaded region in the figure. On the other hand, for $\sigma = 4$, while the characterization is still given by Theorem 2(i), we have $k_0 = 6, k_1 = 2$. Thus, $A_0 = \emptyset$. Hence, $\mathcal{C}(F)$ collapses to the line segment in the figure associated with $\sigma = 4$. Similarly, for $\sigma = 1$, we have another line segment in Figure 7(a). This one is characterized by Theorem 2 (ii).

Figure 7(b), on the other hand, considers $v = 6$, while keeping α, b the same. In this case, for $\sigma \geq 2$ and $\sigma < 2$, respectively Theorem 2(ii) and (iii) provide the equilibrium characterization. In the latter case, the set of equilibrium performance levels becomes a singleton, depicted by 'x' in the figure. Interestingly, even in case (ii) (e.g., for $\sigma = 2$ or $\sigma = 4$), the equilibrium performance levels may collapse to a singleton. Otherwise, they correspond to the line segments in Figure 7(b). □



(a) Cases (i) and (ii)



(b) Cases (ii) and (iii)

Figure 7 $\mathcal{C}(F)$ for different sets of parameters: (a) $\alpha = 4, b = 2, v = 2$; (b) v is updated to 6 and other parameters remain unchanged. For $\sigma \leq 2$ and $\sigma = 4$, the equilibrium performance set collapses to a point, designated by "x" in the figure.

B. Proofs and Examples of Section 4

B.1. Proofs

B.1.1. Proofs of Lemma 1 and Proposition 2

Proof of Lemma 1. (i) We use the identity

$$\max\{\sigma_i, \sigma_j\} = \sigma_i + \sigma_j - \min\{\sigma_i, \sigma_j\}$$

to rewrite (5) as

$$\sigma_i^2 \leq \frac{2}{b} v\sigma_i + \frac{2}{b}\sum_{j\in V} g_{ij} \min\{\sigma_i, \sigma_j\} + \alpha^2 - \frac{2}{b} v\alpha - \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_j.$$

Note that $\sigma_j \leq \alpha$, and therefore,

$$\alpha^2 - \frac{2}{b} v\alpha - \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_j \geq \alpha\left[\alpha - \frac{2}{b} v - \frac{2}{b}\sum_{j\in V} g_{ij}\right] \geq \alpha\left[\alpha - \frac{2}{b} v - \frac{2}{b} d_{max}\right] > 0,$$

where the last inequality follows from the assumptions of the lemma. Therefore, by (4)

$$\sigma_i^2 \leq \frac{2}{b} v\sigma_i + \frac{2}{b}\sum_{j\in V} g_{ij} \min\{\sigma_i, \sigma_j\} < \frac{2}{b} v\sigma_i + \frac{2}{b}\sum_{j\in V} g_{ij} \min\{\sigma_i, \sigma_j\} + \alpha^2 - \frac{2}{b} v\alpha - \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_j,$$

and hence, constraint (5) is not binding.

(ii) Assume that the optimal thresholds $\{\sigma_i\}$ are such that $\sigma_i = \alpha$ for some $i \in V$. Then, (4) implies that

$$\alpha^2 \le \frac{2}{b} v\alpha + \frac{2}{b} \sum_{j \in V} g_{ij}\alpha \le \frac{2}{b} v\alpha + \frac{2}{b} d_{max}\alpha,$$

which contradicts with the assumption of the lemma. Hence, $\sigma_j < \alpha$ for all j.

Assume next that $\sigma_{i^*} = 0$ for some $i^* \in V$. We claim that for small enough $\delta > 0$, the solution $\tilde{\sigma}_i = \sigma_i + \mathbf{1}(i = i^*)\delta$ is feasible and yields strictly larger engagement, contradicting the optimality of $\{\sigma_i\}_i$.

From part (i) the constraint (5) is not binding for any $i \in V$, and therefore, we do not need to check them for feasibility for small enough δ. The constraint $\sigma_i \in [0, \alpha]$ is trivially satisfied for all $i \in V$ by the feasibility of $\{\sigma_i\}$ and the fact that $\sigma_{i^*} = 0$. Furthermore, the constraint (4) is satisfied for all $i \in V \setminus (\{i^*\} \cup N_{i^*})$. This is because for such i and any $k \in \{i\} \cup N_i$, we have $\tilde{\sigma}_k = \sigma_k$. Hence, the constraint (4) corresponding to such i remain intact after the update of thresholds to $\{\tilde{\sigma}_j\}_j$. For all $i \in N_{i^*}$, the right hand side of the constraint (4) (weakly) increases, and therefore, the constraint is satisfied. We conclude by verifying that (4) holds for i^*. Indeed, for small enough δ, we have

$$\delta^2 \le \frac{2}{b} v\delta \le \frac{2}{b} v\delta + \frac{2}{b} \sum_{j \in V} g_{i^*j} \min\{\delta, \tilde{\sigma}_j\},$$

and hence, the constraint continues to hold under $\{\tilde{\sigma}_j\}_j$. Thus, we conclude that the constructed solution is feasible. It can be readily seen that it leads to a strictly higher objective than $\{\sigma_j\}_j$, contradicting its optimality. Hence, we have $\sigma_j > 0$ for all $j \in V$.

Thus, we conclude that only (4) can be binding at the optimal solution. Assume that for some agent $i \in V$, this constraint is not binding, i.e.,

$$\sigma_i^2 < \frac{2}{b} v\sigma_i + \frac{2}{b} \sum_{j \in V} g_{ij} \min\{\sigma_i, \sigma_j\}.$$

Then, σ_i can be increased by some $\delta > 0$ without violating any constraint in the family (4), as the relevant constraint is not binding for i, and for $i \ne j$, increasing σ_i relaxes the aforementioned constraint. Since the remaining constraints are also not binding at the optimal solution, it follows that by increasing σ_i, a new feasible solution with higher engagement is obtained. This contradicts optimality of $\{\sigma_i\}$; hence, we conclude that $\sigma_i^2 = \frac{2}{b} v\sigma_i + \frac{2}{b} \sum_{j \in V} g_{ij} \min\{\sigma_i, \sigma_j\}$ for all $i \in V$. □

Proof of Proposition 2. (i) Suppose (7) has a solution where $\sigma_j > 0$ for all $j \in V$, and let $\sigma_i \le \sigma_j$ for all j. Since $\min\{\sigma_i, \sigma_j\} = \sigma_i$ for all j and $\sigma_i > 0$, (7) implies that $\sigma_i \ge 2(v + d_{min})/b$. Hence, to study solutions of (7) where $\sigma_j \in (0,\alpha]$ for all $j \in V$, it suffices to restrict attention to the set

$$X = \left[\frac{2(v+d_{min})}{b}, \alpha\right]^n.$$

We next show that (7) indeed has a unique solution in X, which in turn is its unique solution in $(0,\alpha]^n$. Using simple algebra and restricting attention to positive solutions, we rewrite (7) as follows:

$$\sigma_i = \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_{j\in V} g_{ij}\min\{\sigma_i,\sigma_j\}}. \tag{34}$$

Therefore, all positive solutions of (7) are fixed points of the operator, $T(\sigma) = (T_i(\sigma))_{i\in V}$, where

$$T_i(\sigma) = \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_{j\in V} g_{ij}\min\{\sigma_i,\sigma_j\}}. \tag{35}$$

We next establish that for $\sigma \in X$, we have $T(\sigma) \in X$. Note that this is equivalent to establishing $T_i(\sigma) \le \alpha$ and $T_i(\sigma) \ge \frac{2(v+d_{min})}{b}$ for any $\sigma \in X$ and $i \in V$.

We first prove that $T_i(\sigma) \le \alpha$ for any $\sigma \in X$ and $i \in V$. To that end, first observe that

$$T_i(\sigma) \le \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_{j\in V} g_{ij}\alpha} \le \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\alpha d_{max}}.$$

Thus, to show that $T_i(\sigma) \le \alpha$, it suffices to establish that $\frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\alpha d_{max}} \le \alpha$ or equivalently $\sqrt{\frac{v^2}{b^2} + \frac{2}{b}\alpha d_{max}} \le \alpha - \frac{v}{b}$. Since $\alpha \ge v/b$, after taking the squares the latter inequality can be written as follows:

$$\frac{v^2}{b^2} + \frac{2}{b}\alpha d_{max} \le \alpha^2 + \frac{v^2}{b^2} - 2\alpha\frac{v}{b}.$$

Cancelling out common terms and rearranging, this is equivalent to

$$\frac{2}{b}\alpha(d_{max}+v) \le \alpha^2.$$

Note that by the assumptions of the proposition, this inequality trivially holds, and hence, we conclude that $T_i(\sigma) \le \alpha$ for any $\sigma \in X$ and $i \in V$.

Next, we prove that $T_i(\sigma) \ge \frac{2(v+d_{min})}{b}$ for any $\sigma \in X$ and $i \in V$. Observe that

$$\begin{aligned} T_i(\sigma) &\ge \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_{j\in V} g_{ij}\frac{2(v+d_{min})}{b}} \ge \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\frac{2(v+d_{min})}{b}d_{min}} \\ &= \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + 4\frac{v}{b}\frac{d_{min}}{b} + 4\frac{d_{min}^2}{b^2}} = \frac{v}{b} + \left(\frac{v}{b} + \frac{2d_{min}}{b}\right) = \frac{2(v+d_{min})}{b}. \end{aligned} \tag{36}$$

Hence, we obtain $T_i(\sigma) \geq \frac{2(v+d_{min})}{b}$, as claimed. Thus, after restricting the domain of $T(\cdot)$ to X, we see that its image is also contained in X, i.e., $T: X \to X$.

We use S to denote the set of fixed points of this operator:

$$S = \{\sigma \in X : \ T(\sigma) = \sigma\}.$$

Let $\geq$ denote the natural partial order in X, i.e., for $\hat{\sigma}, \sigma' \in X$, we say $\hat{\sigma} \geq \sigma'$ if $\hat{\sigma}_i \geq \sigma'_i$ for all $i \in V$. Moreover, we denote by $\hat{\sigma} > \sigma'$ the relations $\hat{\sigma} \geq \sigma'$ and $\hat{\sigma} \neq \sigma'$.

Since X is a compact subset of $\mathbb{R}^n$, it follows that $(X, \geq)$ constitutes a complete lattice. It follows from (35) that when $\hat{\sigma} \geq \sigma'$, we have $T(\hat{\sigma}) \geq T(\sigma')$, i.e., $T(\cdot)$ is an increasing operator on $(X, \geq)$. Using Tarski's fixed point theorem (see, e.g., Topkis (2011)), we conclude that S is a (nonempty) complete lattice. We next establish that S is a singleton.

For the purposes of contradiction, assume that $\tilde{\sigma}, \hat{\sigma} \in S$, where $\hat{\sigma}$ is the smallest point of the complete lattice S, and $\tilde{\sigma} > \hat{\sigma}$. Without loss of generality, we index nodes in increasing order of $\hat{\sigma}_i$, i.e., $\hat{\sigma}_1 \leq \hat{\sigma}_2 \leq \ldots \leq \hat{\sigma}_n$, and prove by induction that $\hat{\sigma}_i = \tilde{\sigma}_i$ for $i = 1, \ldots, n$. Let us start with $i = 1$ and observe that

$$\begin{aligned}\left(\tilde{\sigma}_1 - \frac{v}{b}\right)^2 - \left(\hat{\sigma}_1 - \frac{v}{b}\right)^2 &= \left(T_1(\tilde{\sigma}) - \frac{v}{b}\right)^2 - \left(T_1(\hat{\sigma}) - \frac{v}{b}\right)^2 \\ &= \frac{2}{b}\sum_{j \in V} g_{1j}(\min\{\tilde{\sigma}_1, \tilde{\sigma}_j\} - \min\{\hat{\sigma}_1, \hat{\sigma}_j\}) \leq \frac{2}{b}\sum_{j \in V} g_{1j}(\tilde{\sigma}_1 - \hat{\sigma}_1),\end{aligned} \tag{37}$$

where the inequality follows from the fact that $\tilde{\sigma}_1 \geq \min\{\tilde{\sigma}_1, \tilde{\sigma}_j\}$ and $\hat{\sigma}_1 = \min\{\hat{\sigma}_1, \hat{\sigma}_j\}$. Therefore,

$$(\tilde{\sigma}_1 - \hat{\sigma}_1)\left(\tilde{\sigma}_1 + \hat{\sigma}_1 - 2\frac{v}{b}\right) \leq \frac{2}{b}\sum_{j \in V} g_{1j}(\tilde{\sigma}_1 - \hat{\sigma}_1).$$

If $\tilde{\sigma}_1 \neq \hat{\sigma}_1$, it has to be the case that $\tilde{\sigma}_1 > \hat{\sigma}_1$ (since, by construction, $\tilde{\sigma}$ is larger), and therefore,

$$\tilde{\sigma}_1 + \hat{\sigma}_1 - 2\frac{v}{b} \leq \frac{2d_1}{b}.$$

Observe that $\hat{\sigma}_1$ is the smallest element of $\hat{\sigma}$, and therefore, by (35) and the fact that $\hat{\sigma} \in S$ is a fixed point of $T(\cdot)$, we obtain

$$\hat{\sigma}_1^2 = \frac{2}{b}v\hat{\sigma}_1 + \frac{2}{b}\sum_{j \in V} g_{ij}\hat{\sigma}_1.$$

Hence, $\hat{\sigma}_1 = 2(v + d_1)/b$, and consequently, $\tilde{\sigma}_1 \leq 0$, which is a contradiction since $\tilde{\sigma} \in S \subseteq X$.

Let us proceed with the induction by assuming that for some $k>1$ it is the case that $\tilde{\sigma}_i=\hat{\sigma}_i$ for all $i\le k-1$. Using $T_k(\tilde{\sigma})=\tilde{\sigma}_k$, we have

$$\begin{aligned}\left(\tilde{\sigma}_k-\frac{v}{b}\right)^2&=\frac{v^2}{b^2}+\frac{2}{b}\sum_{j<k}g_{kj}\min\{\tilde{\sigma}_k,\hat{\sigma}_j\}+\frac{2}{b}\sum_{j\ge k}g_{kj}\min\{\tilde{\sigma}_k,\tilde{\sigma}_j\}\\&\le\frac{v^2}{b^2}+\frac{2}{b}\sum_{j<k}g_{kj}\hat{\sigma}_j+\frac{2}{b}\sum_{j\ge k}g_{kj}\min\{\tilde{\sigma}_k,\tilde{\sigma}_j\}.\end{aligned}\tag{38}$$

Using $T_k(\hat{\sigma})=\hat{\sigma}_k$ we also obtain

$$\begin{aligned}\left(\hat{\sigma}_k-\frac{v}{b}\right)^2&=\frac{v^2}{b^2}+\frac{2}{b}\sum_{j<k}g_{kj}\min\{\hat{\sigma}_k,\hat{\sigma}_j\}+\frac{2}{b}\sum_{j\ge k}g_{kj}\min\{\hat{\sigma}_k,\hat{\sigma}_j\}\\&=\frac{v^2}{b^2}+\frac{2}{b}\sum_{j<k}g_{kj}\hat{\sigma}_j+\frac{2}{b}\sum_{j\ge k}g_{kj}\hat{\sigma}_k,\end{aligned}\tag{39}$$

where the last equality uses the fact that $\hat{\sigma}_1\le\hat{\sigma}_2\le\cdots\le\hat{\sigma}_n$.

Using these observations, we get

$$\begin{aligned}(\tilde{\sigma}_k-\hat{\sigma}_k)\left(\tilde{\sigma}_k+\hat{\sigma}_k-\frac{2v}{b}\right)&=\left(\tilde{\sigma}_k-\frac{v}{b}\right)^2-\left(\hat{\sigma}_k-\frac{v}{b}\right)^2\\&\le\frac{2}{b}\sum_{j\ge k}g_{kj}\min\{\tilde{\sigma}_k,\tilde{\sigma}_j\}-\frac{2}{b}\sum_{j\ge k}g_{kj}\hat{\sigma}_k\\&\le\frac{2}{b}\sum_{j\ge k}g_{kj}\tilde{\sigma}_k-\frac{2}{b}\sum_{j\ge k}g_{kj}\hat{\sigma}_k\\&=(\tilde{\sigma}_k-\hat{\sigma}_k)\frac{2}{b}\sum_{j\ge k}g_{kj}.\end{aligned}\tag{40}$$

Here, the first inequality follows from (38) and (39), whereas the second inequality uses the fact that $\min\{\tilde{\sigma}_k,\tilde{\sigma}_j\}\le\tilde{\sigma}_k$. Assume that $\hat{\sigma}_k\ne\tilde{\sigma}_k$. Then, since $\tilde{\sigma}>\hat{\sigma}$ it has to be the case that $\tilde{\sigma}_k>\hat{\sigma}_k$. This in turn implies that the above inequality can be written as

$$\tilde{\sigma}_k+\hat{\sigma}_k-2\frac{v}{b}-\frac{2}{b}d_k^*\le 0,\tag{41}$$

where $d_k^*=\sum_{j\ge k}g_{kj}$.

We claim that

$$\hat{\sigma}_k\ge 2\frac{v}{b}+\frac{2}{b}d_k^*.$$

To see this, note that by equation (39), we have

$$\left(\hat{\sigma}_k-\frac{v}{b}\right)^2=\frac{v^2}{b^2}+\frac{2}{b}\sum_{j<k}g_{kj}\hat{\sigma}_j+\frac{2}{b}d_k^*\hat{\sigma}_k,$$

which can be equivalently written as

$$f(\hat{\sigma}_k) = \frac{2}{b}\sum_{j<k} g_{kj}\hat{\sigma}_j, \tag{42}$$

where

$$f(x) = \left(x - \frac{v}{b}\right)^2 - \frac{v^2}{b^2} - \frac{2}{b}d_k^* x.$$

It is straightforward to verify that $f(x)$ has a root at 0 and a root at $2\frac{v}{b} + \frac{2}{b}d_k^*$. By convexity, in between these points, $f(x) \leq 0$. Therefore, the unique positive solution to (42) is larger than $2\frac{v}{b} + \frac{2}{b}d_k^*$, i.e., $\hat{\sigma}_k \geq \frac{2v}{b} + \frac{2}{b}d_k^*$. Thus, (41) implies that $\tilde{\sigma}_k \leq 0$. Hence, we obtain a contradiction to the fact that $\tilde{\sigma} \in S$, and it follows that $\tilde{\sigma}_k = \hat{\sigma}_k$.

By induction, we conclude that $\tilde{\sigma} = \hat{\sigma}$, and hence, S is a singleton. This in turn implies that (7) has a single solution where $\sigma_k \in (0, \alpha]$ for all k. The optimality of this solution follows directly from Lemma 1.

(ii) Observe that (8) always admits a unique solution. This is because its left hand side is linear, and right hand side is strictly concave in y_i, and for $y_i = 0$ the left hand side is strictly smaller. Moreover, at every visit to S3, the algorithm decreases the cardinality of set V by one. Thus, it follows that the algorithm terminates in finite time. Without loss of generality, assume that the algorithm determines $\sigma = \{\sigma_i\}_{i\in V}$ in the following order: $\sigma_1, \ldots, \sigma_n$.

Let a_i^k and a_{i+1}^k respectively denote the value of a_i and a_{i+1} in kth visit of the algorithm to Step S2 (or prior to their update in Step S3). Similarly, let V^k denote the set V in kth visit of the algorithm to Step S2.

Observe that by definition, we have

$$\sigma_i = \frac{v}{b} + \sqrt{a_i^i + \frac{2}{b}\sigma_i \sum_{j\in V^i} g_{ij}}. \tag{43}$$

Moreover, using the fact $a_{i+1}^{i+1} = a_{i+1}^i + \frac{2}{b}\sigma_i g_{i+1,i}$, we obtain

$$\sigma_{i+1} = \frac{v}{b} + \sqrt{a_{i+1}^{i+1} + \frac{2}{b}\sigma_{i+1}\sum_{j\in V^{i+1}} g_{i+1,j}} = \frac{v}{b} + \sqrt{a_{i+1}^i + \frac{2}{b}\sigma_i g_{i+1,i} + \frac{2}{b}\sigma_{i+1}\sum_{j\in V^{i+1}} g_{i+1,j}}. \tag{44}$$

Suppose for contradiction that $\sigma_{i+1} < \sigma_i$. Consider the equation

$$y = \frac{v}{b} + \sqrt{a_{i+1}^i + \frac{2}{b}y\sum_{j\in V^i} g_{i+1,j}}.$$

Let $y^\star > 0$ denote the unique solution of this equation. Observe that $y^\star \geq \sigma_i$, as otherwise, in the ith visit of the algorithm to Step S2, we get $y^\star = y_{i+1}^\star < y_i^\star$ and obtain a contradiction to the fact that at step i, we have $i \in \arg\min_{k\in V^i} y_k^\star$. Since we also assumed $\sigma_{i+1} < \sigma_i$, we have the following relation:

$$y^\star \geq \sigma_i > \sigma_{i+1}. \tag{45}$$

Define the following functions with domains $\mathbb{R}_+$:

$$f(x) := \frac{v}{b} + \sqrt{a_{i+1}^i + \frac{2}{b} x \sum_{j\in V^i} g_{i+1,j}} - x,$$

$$g(x) := \frac{v}{b} + \sqrt{a_{i+1}^i + \frac{2}{b}\sigma_i g_{i+1,i} + \frac{2}{b} x \sum_{j\in V^{i+1}} g_{i+1,j}} - x.$$

Observe that $f(0) > 0$, $f(y^\star) = 0$, and $f(x) = 0$ has a unique solution since it is strictly concave and $f(0) > 0$. Using these observations, together with the fact that $y^\star \geq \sigma_i$, also implies that $f(y^\star) \leq f(\sigma_i)$. Similarly, observe that by (44), we have $g(\sigma_{i+1}) = 0$. Moreover, $g(0) > 0$ and similar reasoning as before yields $g(x) = 0$ has a unique solution. Based on these observations, together with $\sigma_i > \sigma_{i+1}$, we conclude that $0 = g(\sigma_{i+1}) > g(\sigma_i)$. Finally, observe that by construction, $f(\sigma_i) = g(\sigma_i)$. Collecting all these observations, we conclude that

$$0 = f(y^\star) \leq f(\sigma_i) = g(\sigma_i) < g(\sigma_{i+1}) = 0.$$

Thus, we obtain a contradiction and conclude that $\sigma_i \leq \sigma_{i+1}$. The claim follows since i is arbitrary.

(iii) As argued in part (ii), the algorithm terminates in finite time. Without loss of generality, again assume that the algorithm determines $\sigma = \{\sigma_i\}_{i\in V}$ in the following order: $\sigma_1, \dots, \sigma_n$. By part (ii), we have $\sigma_1 \leq \dots \leq \sigma_n$.

We first establish that $\sigma \in X$. Assume not. Then, there are two cases possible: $\sigma_i > \alpha$ or $\sigma_i < \frac{2(v+d_{min})}{b}$ for some i. Denote the smallest i that satisfies these inequalities with i_1 and i_2, respectively. It follows from the algorithm that

$$\sigma_{i_1} = \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_{j<i_1} g_{i_1 j}\sigma_j + \frac{2}{b}\sigma_{i_1}\sum_{j\geq i_1} g_{i_1 j}} \leq \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sigma_{i_1}\sum_{j\in V} g_{i_1 j}} \leq \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sigma_{i_1} d_{max}},$$

where we use the fact that when the algorithm sets i_1, $a_{i_1} = \frac{v^2}{b^2} + \frac{2}{b}\sum_{j<i_1} g_{i_1 j}\sigma_j$ and $\sigma_j \leq \alpha < \sigma_{i_1}$ for $j < i_1$ by definition of i_1. On the other hand, rearranging terms, the previous inequality yields

$$\sigma_{i_1}^2 + \frac{v^2}{b^2} - 2\sigma_{i_1}\frac{v}{b} \leq \frac{v^2}{b^2} + \frac{2}{b}\sigma_{i_1} d_{max}.$$

After canceling out the common terms and rearranging, this can be equivalently written as follows:

$$\sigma_{i_1}(\sigma_{i_1} - 2\frac{v}{b}) \leq \frac{2}{b}\sigma_{i_1} d_{max}.$$

This in turn implies $\sigma_{i_1} \leq 2\frac{v}{b} + \frac{2}{b}d_{max} < \alpha$, where the last inequality follows by the assumptions of the proposition. Hence, we reach a contradiction to $\sigma_{i_1} > \alpha$ and conclude $\sigma_i \leq \alpha$ for all i.

Next, we show that $\sigma_i \geq \frac{2(v+d_{min})}{b}$ for all $i \in V$. Observe that we have

$$\sigma_{i_2} = \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_{j<i_2} g_{i_2 j}\sigma_j + \frac{2}{b}\sigma_{i_2}\sum_{j\geq i_2} g_{i_2 j}} \geq \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sigma_{i_2}\sum_{j\in V} g_{i_2 j}} \geq \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sigma_{i_2} d_{min}},$$

where we use the fact that when the algorithm sets i_2, $a_{i_2} = \frac{v^2}{b^2} + \frac{2}{b}\sum_{j<i_2} g_{i_2 j}\sigma_j$ and $\sigma_j \geq 2(v + d_{min})/b > \sigma_{i_2}$ for $j < i_2$ by definition of i_2. Rearranging terms, the previous inequality yields

$$\sigma_{i_2}^2 + \frac{v^2}{b^2} - 2\sigma_{i_2}\frac{v}{b} \geq \frac{v^2}{b^2} + \frac{2}{b}\sigma_{i_2} d_{min}.$$

After canceling out the common terms and rearranging, we obtain

$$\sigma_{i_2} - 2\frac{v}{b} \geq \frac{2}{b} d_{min}.$$

Hence, we obtain a contradiction to $\sigma_{i_2} < \frac{2}{b}(v + d_{min})$. Thus, it follows that $\sigma_i \geq \frac{2}{b}(v + d_{min})$ for all i. Hence, $\sigma \in X$, as claimed.

We conclude the proof by showing that the constructed $\sigma \in X$ corresponds to the unique solution of (7) where $\sigma_i > 0$ for all i. Let a_i^k denote the value of a_i in the kth visit of the algorithm to Step S2 (or prior to the updates in Step S3). Consider the σ constructed by the algorithm; we argue that it is a fixed point of $T(\cdot)$, i.e.,

$$\sigma_i = T_i(\sigma) = \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_j g_{ij}\min\{\sigma_i, \sigma_j\}}.$$

Since $\sigma_1 \leq \cdots \leq \sigma_n$, this condition can be written as follows:

$$\sigma_i = \frac{v}{b} + \sqrt{\frac{v^2}{b^2} + \frac{2}{b}\sum_{j<i} g_{ij}\sigma_j + \frac{2}{b}\sum_{j\geq i} g_{ij}\sigma_i}. \tag{46}$$

Consider the ith visit of the algorithm to step S3 (where σ_i is determined). It can be seen that before the update in Step S3, $a_i^i = \frac{v^2}{b^2} + \frac{2}{b}\sum_{k<i} g_{ik}\sigma_k$. Using this observation (46), can be restated

as

$$\sigma_i = \frac{v}{b} + \sqrt{a_i^i + \frac{2}{b}\sum_{j\geq i} g_{ij}\sigma_i}.$$

On the other hand, this condition is trivially satisfied since by construction, $y_i^\star$ solves this equation in Step S2 (in the ith visit of the algorithm to this step) and $\sigma_i = y_i^\star$. Thus, we conclude that $T_i(\sigma) = \sigma_i$. Since i is arbitrary, we conclude that the σ obtained by the algorithm is a fixed point of $T(\cdot)$. Moreover, as established before, $\sigma \in X$. Hence, using part (i) we conclude that the algorithm terminates with the unique fixed point of $T(\cdot)$ in X, which also corresponds to the unique solution of (7) where $\sigma_i > 0$ for all $i \in V$. □

B.1.2. Auxiliary Results on Optimal thresholds We next present structural results that relate the network position of each agent to her optimal threshold in the engagement maximization problem. Our first such result focuses on degrees and provides bounds on thresholds in terms of degrees.

LEMMA 8. *Let $\{\sigma_i\}_{i\in V}$ be the solution to (7) with $\sigma_i > 0$ for all $i \in V$. Then,*

(i) for $i \in \arg\min_{j\in N_i} \sigma_j$, we have

$$\sigma_i = \frac{2}{b}(v + d_i),$$

(ii) $\sigma_i \geq \frac{2}{b}(v + d_{min})$,

(iii) $\sigma_i \leq \frac{2}{b}(v + d_i)$.

Proof:

(i) Consider i such that $\sigma_i \leq \sigma_j$ for all $j \in N_i$. Writing (7) for such an i (and recalling that $\sigma = \{\sigma_j\}_{j\in V}$ is a solution to (7)), we obtain:

$$\sigma_i^2 = \frac{2}{b} v\sigma_i + \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_i.$$

Since $\sigma_i > 0$, this implies

$$\sigma_i = \frac{2}{b}(v + d_i),$$

as claimed.

(ii) Follows directly from (i).

(iii) By equation (7), we have

$$\sigma_i^2 \leq \frac{2v}{b}\sigma_i + \frac{2}{b} d_i\sigma_i,$$

where the inequality follows after observing $\min\{\sigma_i, \sigma_j\} \leq \sigma_i$ for all $i, j \in V$. Since $\sigma_j > 0$ for all j, the claim follows from this inequality. □

For agents who have the smallest thresholds, the fixed-point equation (7) simplifies to a linear equation, allowing for the closed-form characterization in this lemma. As we will shortly see agents with the smallest thresholds are those with the smallest degree. For the remaining agents, (7) can be used to obtain an explicit bound on the thresholds. Both types of characterizations provided in Lemma 8 feature agents' degrees, suggesting that network structure can significantly impact the optimal thresholds, especially in settings where agents have vastly different degrees.

Intuitively, agents who have the smallest degrees benefit the least from network externalities. As such, they have the least incentive to engage with the content. Hence, unless their thresholds and, consequently, the expected inaccuracy level from engaging with the content (conditional on an "Engage" recommendation) are small, they find it optimal not to engage. This observation suggests that the platform may find it optimal to set the thresholds of these agents lower than the rest. Using the characterization in Lemma 8, our next corollary establishes that this indeed is the case.

COROLLARY 1. *Let F be the optimal threshold mechanism, with thresholds $\{\sigma_i\}_{i\in V}$. Suppose that $d_k \leq d_i$ for all $i \in V$. Then, $\sigma_k \leq \sigma_i$ for all $i \in V$.*

Proof: Observe that by Lemma 8(iii), $\sigma_k \leq \frac{2}{b}(v+d_k) = \frac{2}{b}(v+d_{min})$. On the other hand, by Lemma 8(ii), we also have $\sigma_k \geq \frac{2}{b}(v+d_{min})$. These inequalities imply that $\sigma_k = \frac{2}{b}(v+d_{min})$. On the other hand, Lemma 8(ii) implies that $\sigma_i \geq \frac{2}{b}(v+d_{min}) = \sigma_k$ for all $i \in V$, and the claim follows. □

In order to further explore the impact of the network structure on the optimal thresholds, we next focus on a particular class of networks where all agents have the same degree k, also known as k-regular graphs. As the next lemma suggests, for such networks, sending private signals to agents does not benefit the platform for maximizing engagement, and common threshold mechanisms are optimal.

LEMMA 9. *If $\mathcal{G}$ is k-regular and $2(v+k)/b < \alpha$, then optimal thresholds are given by $\sigma_i = 2(v+k)/b$ for all $i \in V$.*

Proof: It can readily checked that $\sigma_i = 2(v+k)/b$ for all $i \in V$ is a solution to (7). Proposition 2 implies that the aforementioned $\{\sigma_i\}$ constitute thresholds of an engagement maximizing mechanism. Hence, the claim follows. □

Intuitively, in k-regular graphs, nodes are "symmetric" in terms of their network position. Thus, for such graphs, choosing thresholds differently for each node should not help the platform increase engagement. Our lemma formalizes this intuition.

We proceed by characterizing assumptions on the primitives that ensure that optimal thresholds are identical and equal to α for all agents.

LEMMA 10. *If $b \leq \frac{2}{\alpha}(v+d_{min})$, then setting $\sigma_i = \alpha$ induces an optimal mechanism, and the corresponding engagement is equal to n.*

Proof: Suppose $\sigma_i = \alpha$ for all $i \in V$. Observe that for this choice of $\{\sigma_i\}$, (5) and $\sigma_i \in [0,\alpha]$ trivially hold, whereas (4) holds since the lemma assumes $b \leq \frac{2}{\alpha}(v+d_{min})$. Thus, the constructed solution is feasible in the engagement maximization problem. The corresponding engagement is $\frac{1}{\alpha}\sum_{i\in V}\sigma_i = n$. The optimality of the constructed mechanism follows since n is also the largest achievable engagement. □

Lemma 10 establishes that whenever agents do not incur too much disutility from engaging with inaccurate content (i.e., b is small), the problem of maximizing engagement is trivial: the platform always recommends engagement. In other words, the corresponding optimal mechanism is a no-intervention mechanism.

In contrast, when this disutility parameter is nontrivial (i.e., b is large), in the absence of any intervention by the platform, agents may cease to engage with the content, which in turn reduces the overall engagement. As we establish in our next lemma, in such cases, it is optimal for the platform to intervene by setting (some) thresholds to levels strictly below α.

LEMMA 11. *Assume that $b > \frac{2}{\alpha}(v+d_{min})$. Let $\{\sigma_i\}_{i\in V}$ be an optimal solution of* (OPT-E).

(i) There exists $i \in V$ such that $\sigma_i < \alpha$.

(ii) The engagement of the corresponding straightforward equilibrium Q is strictly smaller than n, i.e., $E(Q) < n$.

(iii) Consider a relaxation of (OPT-E)*, where the constraint $\sigma_i \in [0,\alpha]$ is omitted. The optimal solution of this problem is such that $\sigma_i \in [0,\alpha]$ for all $i \in V$.*

(iv) The optimal solution of (OPT-E) *is unique.*

Proof: (i) It is straightforward to verify that if $\sigma_i = \alpha$ for all $i \in V$, constraint (4) is violated under the assumptions of the lemma for any minimum degree node.

(ii) Follows immediately from (i).

(iii) Assume that $(\hat{\sigma}_i)_{i\in V}$ is an optimal solution of the relaxed problem (where $\sigma_i \in [0,\alpha]$ is not imposed). To prove the claim, it suffices to establish that this solution always satisfies $\sigma_i \in [0,\alpha]$. For the purposes of contradiction, assume that this is not the case. Then, either there exists some $i \in V$ such that $\hat{\sigma}_i > \alpha$ or there exists some $i^* \in V$ such that $\hat{\sigma}_{i^*} < 0$.

First suppose that there exists some $i \in V$ for which $\hat{\sigma}_i > \alpha$. Then, from constraint (5), we obtain

$$\alpha^2 - \hat{\sigma}_i^2 \geq \frac{2}{b}v(\alpha - \hat{\sigma}_i) + \frac{2}{b}\sum_{j\in V} g_{ij}\max\{\hat{\sigma}_i, \hat{\sigma}_j\} - \frac{2}{b}d_i\hat{\sigma}_i \geq \frac{2}{b}v(\alpha - \hat{\sigma}_i),$$

where the last inequality follows from $\max\{\hat{\sigma}_i, \hat{\sigma}_j\} \geq \hat{\sigma}_i$. Therefore, since $\hat{\sigma}_i > \alpha$, we get

$$\hat{\sigma}_i \leq \frac{2v}{b} - \alpha \leq 0,$$

where the last inequality follows from the assumptions of the lemma. Thus, it follows that $\hat{\sigma}_i \leq \alpha$ for all $i \in V$.

Next, assume that there exists $i^* \in V$ for which $\hat{\sigma}_{i^*} < 0$. Then, using the relation $\min\{\hat{\sigma}_{i^*}, \hat{\sigma}_j\} \leq \hat{\sigma}_{i^*}$ in (4) and canceling out common terms, we obtain $\hat{\sigma}_{i^*} \geq \frac{2v}{b} + \frac{2}{b}\sum_j g_{i^*j} \geq \frac{2v}{b} + \frac{2}{b} d_{i^*} \geq 0$. Hence, we reach a contradiction, and it follows that $\hat{\sigma}_i \geq 0$ for all $i \in V$.

We conclude that even after $\sigma_i \in [0, \alpha]$ is relaxed, the optimal solution satisfies this constraint. Hence, the set of optimal solutions of the engagement maximization problem and the following problem (obtained after relaxing $\sigma_i \in [0, \alpha]$) coincide:

$$\begin{aligned} \max_{\sigma_1, \ldots, \sigma_n} \quad & \frac{1}{\alpha} \sum_{i \in V} \sigma_i \\ s.t. \quad & \bar{f}_i(\sigma) \leq 0 \quad \forall i \in V, \\ & \bar{g}_i(\sigma) \leq 0 \quad \forall i \in V, \end{aligned} \tag{R}$$

where

$$\begin{aligned} \bar{f}_i(\sigma) &= \sigma_i^2 - \frac{2}{b} v \sigma_i - \frac{2}{b} \sum_{j \in V} g_{ij} \min\{\sigma_i, \sigma_j\}, \\ \bar{g}_i(\sigma) &= -\alpha^2 + \sigma_i^2 + \frac{2}{b} v(\alpha - \sigma_i) + \frac{2}{b} \sum_{j \in V} g_{ij} \max\{\sigma_i, \sigma_j\} - \frac{2}{b} d_i \sigma_i. \end{aligned}$$

(iv) We first establish that (R) has a unique solution. For the purposes of contradiction, assume that there are two solutions, σ and $\hat{\sigma}$ such that $\sigma_i \neq \hat{\sigma}_i$ for some $i \in V$. Note that since (R) is a concave maximization problem, its optimal set is convex. Therefore, the point

$$\sigma^m = \frac{\sigma + \hat{\sigma}}{2}$$

is also optimal. Note that the functions $\bar{f}_i(\sigma)$ and $\bar{g}_i(\sigma)$ can both be expressed as the sum of a strictly convex term (σ_i^2) and convex functions of $\{\sigma_j\}$. Therefore, using the feasibility conditions

$$\bar{f}_i(\sigma) \leq 0, \quad \bar{f}_i(\hat{\sigma}) \leq 0, \quad \bar{g}_i(\sigma) \leq 0, \quad \bar{g}_i(\hat{\sigma}) \leq 0,$$

we obtain

$$\bar{f}_i(\sigma^m) < 0, \quad \bar{g}_i(\sigma^m) < 0.$$

Hence, σ^m is an optimal solution for which the constraints are not binding. Therefore, for small enough δ, $(\sigma_i^m + \delta)_{i\in V}$ is also a feasible solution for which the objective is strictly larger than that of σ^m, contradicting the optimality of σ^m. Hence, it follows that (R) has a unique solution. Together with (iii), this implies that the engagement maximization problem also has a unique solution. □

B.1.3. Proofs of Proposition 3, 4 and Theorem 3

Proof of Proposition 3. Recall that $\sigma_i^e = \sigma_i \frac{b}{2v}$ for all $i \in V$, where $\{\sigma_i\}_{i\in V}$ is the solution of (7) where each σ_i is positive. Equivalently, by Proposition 2, these thresholds constitute an optimal solution of the engagement maximization problem. Before showing the connection between the engagement centrality and Bonacich centrality, we present an algebraic lemma that will be of use.

LEMMA 12. *Let* $\{\sigma_i\}_{i\in V}$ *denote the optimal solution of* (OPT-E). *We have* $\frac{\sigma_i\sigma_j}{\beta} \geq \min\{\sigma_i, \sigma_j\} \geq \frac{\sigma_i\sigma_j}{\alpha}$, *where* $\beta = 2(v + d_{min})/b$.

Proof: Without loss of generality, assume that $\sigma_i \leq \sigma_j$. Then, $\min\{\sigma_i, \sigma_j\} = \sigma_i$. Note that σ_j is trivially no larger than α, i.e., $\sigma_j \leq \alpha$. Therefore,

$$\frac{\sigma_i\sigma_j}{\alpha} \leq \sigma_i = \min\{\sigma_i, \sigma_j\}.$$

Similarly, by Lemma 8, $\sigma_j \geq \beta$, from which we obtain

$$\frac{\sigma_i\sigma_j}{\beta} \geq \sigma_i = \min\{\sigma_i, \sigma_j\}.$$

Together with the previous inequality this establishes the claim. □

From Equation (7), for any $i \in V$, we have

$$\sigma_i^2 = \frac{2}{b} v\sigma_i + \frac{2}{b} \sum_{j\in V} g_{ij} \min\{\sigma_i, \sigma_j\}.$$

By Lemma 12, we obtain

$$\frac{2}{b} v\sigma_i + \frac{2}{b} \sum_j g_{ij} \frac{\sigma_i\sigma_j}{\beta} \geq \sigma_i^2 \geq \frac{2}{b} v\sigma_i + \frac{2}{b} \sum_j g_{ij} \frac{\sigma_i\sigma_j}{\alpha}.$$

Since by Lemma 8 we have $\sigma_j \geq \beta > 0$ for all j, the last inequality can be equivalently written as follows:

$$\frac{2}{b} v + \frac{2}{b} \sum_j g_{ij} \frac{\sigma_j}{\beta} \geq \sigma_i \geq \frac{2}{b} v + \frac{2}{b} \sum_j g_{ij} \frac{\sigma_j}{\alpha}.$$

Denoting by σ the vector of $\{\sigma_i\}$, in vector form, this inequality be expressed as

$$\frac{2}{b}v\mathbf{1}+\frac{2}{b\beta}G\sigma \geq \sigma \geq \frac{2}{b}v\mathbf{1}+\frac{2}{b\alpha}G\sigma, \tag{47}$$

where the inequality is entrywise. Since $b>\frac{2}{\alpha}(v+d_{max})$, it follows that $(I-\frac{2}{\alpha b}G)$ is diagonally dominant, and its inverse is given by $\sum_{k=0}^{\infty}(\frac{2}{\alpha b}G)^k$, which has nonnegative entries. Moreover, by the assumptions of the proposition, we also have that $(I-\frac{2}{\beta b}G)^{-1}$ has nonnegative entries. Using these observations to rearrange terms in (47) (and keeping in mind that vector inequalities are preserved when they are multiplied by matrices that have nonnegative entries), we obtain

$$\frac{2v}{b}\left(I-\frac{2}{\beta b}G\right)^{-1}\mathbf{1}\geq\sigma\geq\frac{2v}{b}\left(I-\frac{2}{\alpha b}G\right)^{-1}\mathbf{1}.$$

The claim follows after observing that the quantities on the left and right correspond to Bonacich centrality vectors with different parameters. □

Proof of Proposition 4. Observe that if $2(v+d_{min})/\alpha\geq b$, Lemma 10 implies that engagement can be maximized using a common threshold mechanism with $\sigma=\alpha$. Observe that this mechanism sends the same signal, regardless of the realization of ϵ, and is uninformative. Hence, agents' strategies remain optimal, if the platform were to use instead $\sigma=0$ – another mechanism with uninformative signals. Thus, the first part of the claim follows.

Suppose next that $2(v+d_{min})/\alpha<b$, and define:

$$f_L(\sigma):=|V|\frac{\sigma}{\alpha}+|A_0(\sigma)|\frac{\alpha-\sigma}{\alpha}\text{ and }f_R(\sigma):=|A_0(\sigma)|\frac{\alpha-\sigma}{\alpha}+|A_1(\sigma)|\frac{\sigma}{\alpha}.$$

Consider maximal equilibria in cases (i) and (ii) of Theorem 2. For $\alpha\geq\sigma\geq 2v/b$, case (i) of the theorem applies, and it follows that for any given σ satisfying these inequalities, the maximum corresponding engagement is given by $f_R(\sigma)$. Since $2v/b-\alpha<0$, Theorem 2 implies that for σ satisfying $0\leq\sigma<2v/b$, the maximum corresponding engagement is given by $f_L(\sigma)$. Note that when $\sigma\leq 2v/b$, $A_1(\sigma)=V$, and therefore $f_L(\sigma)=f_R(\sigma)$. Hence, we conclude that for any σ the corresponding expected engagement is given by $f_R(\sigma)$. Moreover, if $f_R(\sigma)$ admits a maximizer in $[0,\alpha]$ it achieves the maximum engagement. Note that $f_R(\cdot)$ is left-continuous by definition, and increasing in the intervals where it is continuous (i.e., where $|A_0(\cdot)|$ and $|A_1(\cdot)|$ are constant). Thus, $\arg\max_\sigma f_R(\sigma)$ is well-defined and the claim follows.

□

Proof of Theorem 3. First note that if $\alpha \geq \frac{2}{b}v$, then setting $\sigma_i = 0$ for all i is feasible in (OPT-M). This leads to an objective value of zero, which, after observing the nonnegativity of the objective, implies that the constructed solution is optimal. Thus, we conduct our analysis restricting attention to $\alpha < \frac{2}{b}v$. We claim that in this regime, the optimal solution is given by $\sigma_i = \min\{\alpha, \frac{2v}{b} - \alpha\}$ for all i. We prove the claim by first relaxing (OPT-M) and showing that the aforementioned solution is optimal for this relaxation. Then, we argue that this solution is also feasible, and hence optimal, in (OPT-M).

We start by relaxing the constraint (4) of (OPT-M) and considering the following problem:

$$\begin{aligned} \min_{\sigma_1,\dots,\sigma_n} \quad & \frac{1}{\alpha}\sum_{i\in V}\frac{\sigma_i^2}{2} \\ s.t. \quad & \alpha^2 - \sigma_i^2 \geq \frac{2}{b}v(\alpha-\sigma_i) + \frac{2}{b}\sum_{j\in V} g_{ij}\max\{\sigma_i,\sigma_j\} - \frac{2}{b}d_i\sigma_i \quad \text{for } i \in V \\ & 0 \leq \sigma_i \leq \alpha \quad \text{for } i \in V. \end{aligned} \tag{M2}$$

Observe that

$$\frac{2}{b}v(\alpha-\sigma_i) + \frac{2}{b}\sum_{j\in V} g_{ij}\max\{\sigma_i,\sigma_j\} - \frac{2}{b}d_i\sigma_i \geq \frac{2}{b}v(\alpha-\sigma_i) + \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_i - \frac{2}{b}d_i\sigma_i = \frac{2}{b}v(\alpha-\sigma_i).$$

We use this observation to further relax the first constraint of (M2) by replacing the right-hand side of the first constraint with $\frac{2}{b}v(\alpha-\sigma_i)$:

$$\begin{aligned} \min_{\sigma_1,\dots,\sigma_n} \quad & \frac{1}{\alpha}\sum_{i\in V}\frac{\sigma_i^2}{2} \\ s.t. \quad & \alpha^2 - \sigma_i^2 \geq \frac{2}{b}v(\alpha-\sigma_i) \quad \text{for } i \in V \\ & 0 \leq \sigma_i \leq \alpha \quad \text{for } i \in V. \end{aligned} \tag{M3}$$

Note that this problem decouples over i. Thus, solving the following problem for each i yields an optimal solution to (M3):

$$\begin{aligned} \min_{\sigma_i} \quad & \frac{1}{\alpha}\frac{\sigma_i^2}{2} \\ s.t. \quad & \alpha^2 - \sigma_i^2 \geq \frac{2}{b}v(\alpha-\sigma_i), \\ & 0 \leq \sigma_i \leq \alpha. \end{aligned} \tag{M4}$$

Suppose that the optimal solution is such that $\sigma_i^\star < \alpha$. Then, feasibility implies that $\alpha^2 - (\sigma_i^\star)^2 \geq \frac{2}{b}v(\alpha - \sigma_i^\star)$, or equivalently

$$\alpha + \sigma_i^\star \geq \frac{2}{b}v.$$

This implies that $\sigma_i^\star \geq \frac{2}{b}v - \alpha$. Since the objective is strictly increasing in σ_i (for $\sigma_i \geq 0$) and $\frac{2}{b}v - \alpha > 0$, we conclude that if the optimal solution is not $\sigma_i^\star = \alpha$, then it is given by $\sigma_i^\star = \frac{2}{b}v - \alpha$. Moreover, since the objective is increasing in σ_i, we conclude that if $\frac{2}{b}v - \alpha \leq \alpha$, then setting $\sigma_i^\star = \frac{2}{b}v - \alpha$ yields a feasible and optimal solution. Otherwise, $\sigma_i^\star = \alpha$ yields an optimal solution. These observations imply that the optimal solution to (M3) is such that

$$\sigma_i^\star = \min\{\alpha, \frac{2v}{b} - \alpha\}, \tag{48}$$

for all $i \in V$. Since (M3) is obtained by relaxing (OPT-M), if (48) is feasible in this problem, then it is optimal.

Observe that under the solution (48), we have $\sigma_i \in [0, \alpha]$ for all $i \in V$ by construction. In addition, we have $\alpha^2 - (\sigma_i^\star)^2 \geq \frac{2}{b}v(\alpha - \sigma_i^\star)$. This inequality trivially follows if $\sigma_i^\star = \alpha$. Otherwise, observe that $\sigma_i^\star = \frac{2}{b}v - \alpha$, and hence,

$$\alpha^2 - (\sigma_i^\star)^2 = (\alpha - \sigma_i^\star)(\alpha + \sigma_i^\star) = \frac{2}{b}v(\alpha - \sigma_i^\star).$$

On the other hand, we also have

$$\frac{2}{b}v(\alpha - \sigma_i^\star) + \frac{2}{b}\sum_{j\in V} g_{ij}\max\{\sigma_i^\star, \sigma_j^\star\} - \frac{2}{b}d_i\sigma_i^\star = \frac{2}{b}v(\alpha - \sigma_i^\star) + \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_i^\star - \frac{2}{b}d_i\sigma_i^\star = \frac{2}{b}v(\alpha - \sigma_i^\star).$$

Using these observations, it follows that the constraint (5) of (OPT-M) trivially holds under the constructed solution. Finally, observe that

$$\frac{2}{b}v\sigma_i^\star + \frac{2}{b}\sum_{j\in V} g_{ij}\min\{\sigma_i^\star, \sigma_j^\star\} \geq \frac{2}{b}v\sigma_i^\star > \alpha\sigma_i^\star \geq (\sigma_i^\star)^2.$$

Thus, the constraint (4) of (OPT-M) also trivially holds for the solution (48). Since this solution is optimal for a relaxation of (OPT-M) and also satisfies all the constraints of (OPT-M), we conclude that it is optimal in this problem.

Summarizing, if $\alpha \geq \frac{2}{b}v$, then the optimal solution to (OPT-M) is given by $\sigma_i = 0$ for all i. Otherwise, it is given by the solution in (48). Hence, the claim follows. □

B.2. Additional Results and Examples

We next provide additional results and examples that complement our findings in Section 4.

B.2.1. Example: Optimal thresholds vs. Network Structure In this section, we provide an example that illustrates that in the engagement maximization problem nodes that have larger degree may have smaller or larger thresholds than the rest.

EXAMPLE 1. Assume that $\mathcal{G}$ is a star graph with k leaves. Assume further that $\frac{2}{b}(v+k)<\alpha$. The leaves have the minimum degree, and therefore, by Corollary 1, the optimal thresholds corresponding to the leaves are the smallest. By part (i) of Lemma 8, it follows that for a leaf node l the threshold is given by $\sigma_l = \frac{2}{b}(v+1)$. For the center, the unique positive solution of (7) yields

$$\sigma_c = \frac{v}{b} + \frac{\sqrt{4k(v+1)+v^2}}{b}.$$

Thus, in this example the agent with a larger degree has a larger threshold.

Next, consider a graph that consists of two different connected components: a star with k leaves and a $(k-1)$-regular graph. Here, each node of the $(k-1)$-regular graph has a smaller degree than the center of the star. It can be readily checked that (OPT-E) decouples over disconnected components of the graph. Lemma 9 implies that the optimal threshold for each node of the $(k-1)$-regular graph is equal to

$$\sigma_r = \frac{2}{b}(v+k-1).$$

Thus, when k is large enough, $\sigma_r > \sigma_c$ even though the degrees of the nodes in the regular graph are smaller than the degree of the center of the star. □

B.2.2. Engagement Maximization for Common Threshold Mechanisms In this section we provide an algorithm for finding the engagement-maximizing common threshold mechanism. Our algorithm relies on the following proposition.

PROPOSITION 6. *(i) Let $B_0 = \{\sigma | (b\frac{\sigma+\alpha}{2} - v) \in \{0,1,\dots,n\}\}$, $B_1 = \{\sigma | (b\frac{\sigma}{2} - v) \in \{0,1,\dots,n\}\}$, and $B = B_1 \cup B_2$. Consider $\sigma_1, \sigma_2 \in B$ such that $(\sigma_1,\sigma_2)\cap B = \emptyset$. The maximum achievable engagement is an increasing function of σ in $(\sigma_1,\sigma_2]$.*

(ii) If engagement is discontinuous at some $\sigma \in B$, then engagement achievable at $\sigma+\delta$ is strictly lower than that at σ for any $\delta > 0$.

(iii) Moreover, maximum engagement (among all common threshold mechanisms) is achieved by some $\sigma \in B$.

Proof: Consider the characterization of common threshold mechanisms in Theorem 2. Observe that B_0 contains the σ where k_0 changes, i.e., k_0 associated with $\sigma+\delta$ is different than the one associated with $\sigma \in B_0$ for any $\delta > 0$. Similarly, B_1 contains the σ where k_1 changes. Thus, it follows that $|A_0(\sigma)|$ and $|A_1(\sigma)|$ are constant for $\sigma \in (\sigma_1,\sigma_2]$, where σ_1,σ_2 are as defined in the statement of the proposition. As can be seen from the proof of Proposition 4, engagement for a given σ is given by $|A_0(\sigma)|\frac{\alpha-\sigma}{\alpha}+|A_1(\sigma)|\frac{\sigma}{\alpha}$. These observations imply that the maximum engagement that corresponds to a given σ changes continuously in $(\sigma_1,\sigma_2]$. Moreover, since $|V|\geq|A_1(\sigma)|\geq|A_0(\sigma)|$, it follows that the corresponding engagement is weakly increasing in $\sigma\in(\sigma_1,\sigma_2]$. Hence, we conclude that the optimal engagement in this interval is achieved for $\sigma=\sigma_2$. Since this is true for any such interval, we conclude that optimal engagement is obtained for $\sigma\in B$.

Suppose that maximum engagement changes discontinuously at $\sigma\in B$. Then, it must be the case that $|A_i(\sigma)|\neq|A_i(\sigma+\delta)|$ for arbitrarily small $\delta>0$ for $i=0$ or $i=1$. However, since $A_i(\sigma)\subset A_i(\sigma')$ for $\sigma'\leq\sigma$, we conclude that $|A_i(\sigma)|>|A_i(\sigma+\delta)|$. Hence, if engagement changes discontinuously at σ, then it has to be the case that engagement achievable at $\sigma+\delta$ is strictly lower than that at σ, as claimed. □

This finding leads to two important conclusions. First, if the platform resorts to using a common threshold mechanism, the corresponding threshold must be designed with extreme care, as the performance is not robust to perturbations (as formalized in part (i)). Second, in general, non-motonocity suggests that it may be difficult to find the optimal σ. However, Proposition 6 suggests that when searching over optimal σ, it suffices to restrict attention to finitely many common thresholds that belong to B. As can be seen from the proof of Proposition 4, engagement for a given σ is given by $|A_0(\sigma)|\frac{\alpha-\sigma}{\alpha}+|A_1(\sigma)|\frac{\sigma}{\alpha}$. Since $|B|\leq 2n+2$, this immediately implies a simple algorithm (Algorithm 2) for characterizing the thresholds that maximize engagement among all common threshold mechanisms.

Algorithm 2 An algorithm for computing common threshold $\sigma^\star$ that maximizes engagement.

S1. If $2v/b>\alpha$, set $\sigma^\star=0$, and go to Step S3. Otherwise, compute B_0,B_1, and set $B=B_0\cup B_1$.

S2. Compute $\sigma^\star=\arg\max_{\sigma\in B}|A_0(\sigma)|\frac{\alpha-\sigma}{\alpha}+|A_1(\sigma)|\frac{\sigma}{\alpha}$.

S3. Return $\sigma^\star$.

B.2.3. Performance Comparison: Optimal vs. Common Threshold vs. No-Intervention

Engagement. Our findings in Sections 4.2.1 and 4.2.2 naturally bring up the question of comparing the engagement achieved by the optimal signaling mechanism, the optimal common threshold mechanism, and the no-intervention mechanisms.

In this section, we first focus on k-regular graphs. As discussed earlier (and established in Lemma 9), for these graphs, common threshold mechanisms are optimal among all mechanisms for the engagement maximization problem. However, it is not clear if, for these graphs, not intervening at all already guarantees high engagement. Our first corollary argues that for k-regular graphs, the benefit of intervention can range from being extremely significant to being non-existent.

Corollary 2. *Let E_O, E_P, and E_N respectively denote the maximum engagement that can be obtained by the optimal (private) signaling mechanism, a common threshold mechanism, and a no-intervention mechanism. Moreover, assume that the underlying network is k-regular. Then*

(i) if $b \leq 2(v+k)/\alpha$, then $E_N = E_P = E_O = |V|$,

(ii) if $b > 2(v+k)/\alpha$, then

$$E_N = 0 < E_P = E_O = |V|\frac{2(v+k)}{\alpha b}.$$

Proof: (i) Follows immediately from Lemma 10.

(ii) In this case, $k < \lceil b\frac{\alpha}{2} - v \rceil$. Hence, $A_0(0) = \emptyset$, and by Theorem 2 we conclude that under common threshold $\sigma = 0$ the maximum achievable engagement is 0. Common threshold mechanisms with $\sigma = 0$ and $\sigma = \alpha$ share the same equilibria, since in both cases the signaling mechanism almost surely sends a single signal, regardless of ϵ. These observations imply that for the no-intervention mechanism, we obtain an engagement of 0.

Observe that Lemma 9 implies that setting $\sigma_i = 2(v+k)/b$ for all $i \in V$ maximizes engagement under the assumption of part (ii) of the corollary. Hence, common threshold mechanisms are optimal, and the maximum engagement achieved for the common threshold and the optimal signaling mechanisms is given as $2\frac{n}{\alpha}\frac{v+k}{b}$. □

This corollary implies that no intervention may lead to a drastic reduction in the amount of engagement on the platform depending on the connectivity (k) of the underlying network and the parameters of agents' payoffs. For example, when agents are not as well connected (small k) and $b > 2(v+k)/\alpha$, in the absence of intervention, the platform observes no engagement, whereas with an appropriately designed common threshold mechanism, an amount of engagement proportional to the size of the network can be obtained.

In Corollary 2, since we focus on k-regular graphs, there is no gap between common threshold and private signaling mechanisms in terms of engagement. We next analyze more general graphs and explore whether these mechanisms induce very different engagement levels.

Corollary 3. *Fix integers $m, n_0 > 1$ such that $n_0 > 2m$. Let nodes be labeled as $1, \ldots, n = 2n_0$. Consider a graph that consists of two disjoint graphs with $n/2 = n_0$ nodes each. The first graph consists of nodes $V_1 = \{1, \ldots, n_0\}$, where each node $i \in V_1$ is connected to nodes $j \in V_1$ such that $j \equiv j' \pmod{n_0}$ for some $j' \in \{i-1, i+1\}$. The second graph consists of nodes $V_2 = \{n_0+1, \ldots, 2n_0\}$, where each node $i \in V_2$ is connected to nodes $j \in V_2$ such that $j \equiv j' \pmod{n_0}$ for some $j' \in \{i-m, \ldots, i-1, i+1, i+m\}$.*

Let E_O and E_P respectively denote the maximum engagement that can be obtained by the optimal (private) signaling mechanism and a common threshold mechanism. Assume that $b > 2(v+2m)/\alpha$. We have:

$$E_O = 2n\frac{v+m+1}{\alpha b}, \text{ while } E_P = n \cdot \max\left\{\frac{2(v+2)}{\alpha b}, \frac{v+2m}{\alpha b}\right\}.$$

Proof: Observe that for the optimal mechanism, the problem decouples over the two disjoint components of the graph. By construction, one component is 2-regular, whereas the other is $2m$-regular. Lemma 9 suggests that the optimal threshold is such that for any i in the first component, we have $\sigma_i = 2(v+2)/b$, and for any i in the second component, we have $\sigma_i = 2(v+2m)/b$. Thus, corresponding overall engagement is given by

$$E_O = \frac{n}{2\alpha}\frac{2(v+2)}{b} + \frac{n}{2\alpha}\frac{2(v+2m)}{b} = 2\frac{n}{\alpha b}(v+m+1).$$

By the assumptions of the corollary, for all $\sigma \geq 0$, we have

$$b\frac{\sigma+\alpha}{2} - v > 2m.$$

Hence, $A_0(\sigma) = \emptyset$ for all $\sigma \geq 0$. On the other hand, we have $b\frac{\sigma}{2} - v \leq 2$ and $A_1(\sigma) = V$ for $\sigma \leq 2(v+2)/b$. Similarly, $A_1(\sigma) = V_2$ for $2(v+2)/b < \sigma \leq 2(v+2m)/b$, and $A_1(\sigma) = \emptyset$ for $\sigma > 2(v+2m)/b$. As can be seen from the proof of Proposition 4, the expected engagement for a given σ is given by $|A_0(\sigma)|\frac{\alpha-\sigma}{\alpha} + |A_1(\sigma)|\frac{\sigma}{\alpha}$. Using these observations, the engagement levels for different values of σ are given as follows:

(i) if $\sigma \leq 2(v+2)/b$, then $A_0(\sigma) = \emptyset$ and $A_1(\sigma) = V$, and hence, $|A_0(\sigma)| = 0$, while $|A_1(\sigma)| = n$. In this regime, the expected engagement is $\frac{n\sigma}{\alpha}$.

(ii) if $2(v+2)/b < \sigma \leq 2(v+2m)/b$, then $A_0(\sigma) = \emptyset$ and $A_1(\sigma) = V_2$, and hence, $|A_0(\sigma)| = 0$ and $|A_1(\sigma)| = n/2$. The corresponding engagement level is given by $\frac{n\sigma}{2\alpha}$.

(iii) if $2(v+2m)/b < \sigma$, then $A_1(\sigma) = \emptyset$, and therefore, $|A_0(\sigma)| = 0$ and $|A_1(\sigma)| = 0$. The corresponding engagement level is 0.

The largest engagement in case (i) is achieved for $\sigma = 2(v+2)/b$ and is given by $\frac{2n(v+2)}{\alpha b}$, and the largest engagement in case (ii) is achieved for $\sigma = 2(v+2m)/b$ and is given by $\frac{n(v+2m)}{\alpha b}$. In case (iii), the engagement is always equal to zero. This observation readily implies that

$$E_P = n \cdot \max\left\{\frac{2(v+2)}{\alpha b}, \frac{v+2m}{\alpha b}\right\},$$

and the claim follows. □

This corollary readily implies that when agents in the network are heterogeneous in terms of their network positions, the engagement achieved under the best common threshold mechanism can be substantially lower than that under the optimal private signaling mechanism. To see this, note that in the corollary by choosing $m = 2$ and v arbitrarily small, we immediately obtain examples where $E_O/E_P \approx 3/2$, i.e., the optimal private signaling mechanism leads to 50% more engagement.

In conclusion, we have established that an engagement-maximizing platform greatly benefits from agent-specific signaling of content accuracy, which exposes more central agents to content with higher expected inaccuracy level.

Misinformation. We next compare (optimal) common threshold mechanisms with no-intervention mechanisms and quantify the improvement due to intervention.

COROLLARY 4. *If $\frac{2v}{b} \leq \alpha$ or $\frac{2v}{b} \geq 2\alpha$, then no-intervention mechanisms are optimal. Otherwise, let M_P, M_N respectively denote the misinformation associated with the optimal common threshold mechanism and no-intervention mechanism respectively. Then,*

$$M_N = |V|\frac{\alpha}{2}, \text{ and } M_P = |V|\frac{(\frac{2v}{b}-\alpha)^2}{2\alpha}.$$

This corollary implies that for moderate levels of disutility parameter b (so that $\alpha < 2v/b < 2\alpha$), no-intervention mechanisms may lead to substantially more misinformation than common threshold mechanisms. In particular, in this regime, we have

$$\frac{M_N}{M_P} = \left(\frac{\alpha}{2v/b-\alpha}\right)^2.$$

Thus, for $2v/b \approx \alpha$, the misinformation ratio becomes unbounded, clearly highlighting the substantial effect of intervention on misinformation levels.

Proof of Corollary 4. The fact that no intervention mechanisms are optimal when $\frac{2v}{b} \leq \alpha$ and $\frac{2v}{b} \geq 2\alpha$ readily follows from Theorem 3.

When $\alpha < \frac{2v}{b} < 2\alpha$, the misinformation under the no intervention mechanism is given by Theorem 2(iii) and equals $|V|\frac{\alpha}{2}$, as claimed. By Theorem 3, the optimal common threshold mechanism features a threshold of $\sigma = \frac{2v}{b} - \alpha$. The corresponding minimal misinformation is given by $|V|\frac{\sigma^2}{2\alpha} = |V|\frac{(\frac{2v}{b}-\alpha)^2}{2\alpha}$. □

B.2.4. Revealing the Inaccuracy Level Publicly In this section we consider an alternative benchmark where the platform reveals the true inaccuracy level ϵ of the content publicly to all the agents (for any realization of ϵ). We refer to such mechanisms as *true inaccuracy level revelation* mechanisms, and explore the expected engagement and misinformation under these mechanisms.

Observe that under true inaccuracy level revelation the agents play a game where their payoffs are as given in (1). It can be readily checked that for any given ϵ the induced game $\mathcal{G}_\epsilon$ is a supermodular game. Thus, for any given ϵ, the induced game has best and worst equilibria, which we respectively denote by $\overline{X}^\epsilon$ and $\underline{X}^\epsilon$. These equilibria can be characterized following the approach in the proof of Theorem 2. In this theorem, for a given threshold σ, when agents receive the public signal $s_i = 1$, they play a supermodular game where the expected inaccuracy level is given by $\sigma/2$, and the inaccuracy disutility is given by $-b\sigma/2$ (similarly when $s_i = 0$, they play a supermodular game where the inaccuracy disutility is $-b(\sigma+\alpha)/2$). The corresponding best and worst equilibria are characterized in terms of the k-cores of the underlying network. When ϵ is publicly revealed, as opposed to focusing on the expected inaccuracy level $\sigma/2$, agents base their decision on the inaccuracy level ϵ of the content. Thus, repeating the steps in the proof of Theorem 2 (and replacing $\sigma/2$ with ϵ in the payoffs), the following result can be established:

LEMMA 13. *Let k_ϵ be the smallest integer satisfying $k_\epsilon \geq b\epsilon - v$, and A_{k_ϵ} denote the k_ϵ-core of the underlying network.*

- *If $v \leq b\epsilon$, in $\overline{X}^\epsilon$ agents in A_{k_ϵ} engage and in $\underline{X}^\epsilon$ no agent engages.*
- *If $v > b\epsilon$ then both in $\overline{X}^\epsilon$ and $\underline{X}^\epsilon$ all agents engage.*

Observe that this result allows for characterizing the maximum engagement/minimum misinformation that can be obtained when the platform uses true inaccuracy level revelation mechanisms. In particular, to characterize maximum expected engagement (minimum expected misinformation) induced under such mechanisms, it suffices to focus on the best (worst) equilibrium $\overline{X}^\epsilon$ ($\underline{X}^\epsilon$) for every ϵ and take an expectation over ϵ.

Let $\epsilon_0 := 0$, and $\epsilon_k := (k-1+v)/b$ for $k \geq 1$. As before, let A_k denote the set of nodes in the k-core of the network. Observe that by the lemma given above for $\epsilon \in (\epsilon_k, \epsilon_{k+1}]$ at the best equilibrium agents in A_k engage. Thus, the expected maximum engagement under true inaccuracy level revelation mechanism can be given as follows:

$$\sum_{k=0}^{\infty} \frac{\epsilon_{k+1} - \epsilon_k}{\alpha} |A_k|,$$

and the corresponding expected misinformation is given by:

$$\sum_{k=0}^{\infty} \frac{\epsilon_{k+1}^2 - \epsilon_k^2}{2\alpha} |A_k|.$$

Similarly, expected minimum engagement under this mechanism is $|V|\frac{\epsilon_1}{\alpha}$, and the corresponding minimum expected misinformation is given by $|V|\frac{\epsilon_1^2}{2\alpha}$. Note that in these expressions, the terms $\{\epsilon_k\}$ are functions of the disutility parameter b, as can be seen from their definition.

Thus, we conclude that if the platform were to reveal the true inaccuracy level, the maximum expected engagement achievable under this mechanism can be expressed in terms of the k-cores of the underlying network and $\{\epsilon_k\}$. Similarly, the minimum misinformation can be characterized in terms of the number of nodes in the underlying network and ϵ_1. For the numerical examples of Section 4.4, following this observation we first find the k-cores of the network and then use them to compute the induced maximum expected engagement and minimum expected misinformation under true inaccuracy level revelation mechanisms, as outlined above.

C. Additional Results on the E/M Frontier and the Proof of Proposition 5

We start our discussion by first analyzing the optimal mechanisms that solve (12) and characterizing the E/M frontier. Then, we proceed with the proof of the main result of Section 5.

C.1. Optimal Mechanism

Once again, to obtain optimal mechanisms solving (12), we restrict attention to straightforward threshold mechanisms and focus on the following problem:[21]

$$\begin{aligned} M^*(E) = \max_{\sigma_1,\dots,\sigma_n \in [0,\alpha]} \quad & \frac{1}{\alpha}\sum_{i\in V}\frac{\sigma_i^2}{2} \\ s.t \quad & \sum_{i\in V}\sigma_i \geq \alpha E \\ & (4),(5). \end{aligned} \tag{OPT-F}$$

The constraints other than the first one, as before, guarantee that the optimal solution yields thresholds of a straightforward threshold mechanism (see Proposition 1). The first constraint is the engagement requirement constraint from (12). For the remainder of this section, we denote the solution to (OPT-F) for a given $E \geq 0$ by $\sigma_i^*(E)$.

If we remove the two families of incentive compatibility constraints (i.e., the last two constraints), the resulting problem is fairly easy to solve: we set all the thresholds equal to $\alpha E/n$ (whenever $E \leq n$). This provides a lower bound to $M^*(E)$, as the next lemma formalizes.

Lemma 14. $M^*(E) \geq \frac{\alpha}{2n}E^2$.

Proof: Observe that if $E > n$, then (OPT-F) is infeasible; hence $M^*(E) = \infty$, and the claim trivially holds. Assume that $E \leq n$.

Consider the following relaxation of (OPT-F) obtained by omitting the last two constraints:

$$\begin{aligned} \min_{\sigma_1,\dots,\sigma_n} \quad & \frac{1}{\alpha}\sum_{i\in V}\frac{\sigma_i^2}{2} \\ s.t. \quad & \sum_{i\in V}\sigma_i \geq \alpha E \\ & 0 \leq \sigma_i \leq \alpha \quad \text{for all } i \in V. \end{aligned} \tag{49}$$

[21] Recall that Theorem 1 implies that given a signaling mechanism, a straightforward threshold mechanism with weakly lower misinformation and the same amount of engagement can always be obtained. This immediately implies that it still suffices to search for optimal mechanisms by restricting attention to this class of mechanisms.

Since this is an optimization problem with a continuous objective over a compact domain, it admits an optimal solution. Observe that this problem is symmetric in $\{\sigma_i\}$, i.e., if $\{\sigma_1, \dots, \sigma_i, \dots \sigma_j \dots \sigma_n\}$ is an optimal solution, so is $\{\sigma_1, \dots, \sigma_j, \dots \sigma_i \dots \sigma_n\}$. Furthermore, by convexity of the problem (and hence the set of optimal solutions), it admits an optimal solution such that $\sigma_i = \sigma$ for all i. Setting $\sigma_i = \sigma$ for all i in (49), we obtain the following problem:

$$\begin{aligned} \min_{\sigma} \quad & \frac{n}{2\alpha}\sigma^2 \\ s.t. \quad & \sigma \geq \frac{\alpha E}{n}, \\ & 0 \leq \sigma \leq \alpha. \end{aligned}$$

Observe that since the objective is increasing in σ, it follows that the optimal solution to this problem is $\sigma = \frac{\alpha E}{n} \leq \alpha$. Thus, we conclude that (49) admits a solution $\sigma_i = \sigma = \frac{\alpha E}{n}$ for all i. The corresponding objective value is $\frac{n}{2\alpha}\sigma^2 = \frac{\alpha E^2}{2n}$. Since (49) is obtained by relaxing (OPT-F), it follows that $M^*(E) \geq \frac{\alpha E^2}{2n}$, as claimed. □

We proceed by providing a particular parameter regime where the optimal signaling mechanism can achieve the lower bound of Lemma 14. To obtain this result, we show that for this regime, incentive compatibility constraints in (OPT-F) are irrelevant and can be relaxed.

LEMMA 15. *If* $E \in \left[\max\left\{0, \frac{n}{\alpha}\left(\frac{2v}{b} - \alpha\right)\right\}, \min\left\{n, \frac{2n}{\alpha}\frac{d_{min}+v}{b}\right\}\right]$, *then*

$$M^*(E) = \frac{\alpha}{2n}E^2, \text{ and } \sigma_i^*(E) = \frac{\alpha E}{n} \qquad \text{for all } i \in V.$$

Proof: Fix some E that satisfies the constraints of the lemma. Set $\sigma_i = \sigma = \alpha E/n$ for all $i \in V$. We claim that this is a feasible solution in (OPT-F). To see this, note that when $\sigma_i = \sigma$ for all $i \in V$, the constraint (4) of this problem reduces to

$$\sigma^2 \leq \frac{2}{b}v\sigma + d_i\sigma$$

for any $i \in V$. On the other hand, by construction, $0 \leq \sigma = \alpha E/n \leq \frac{2}{b}(d_{min} + v)$, and hence, this inequality trivially holds. Similarly, when $\sigma_i = \sigma$ for all $i \in V$, the constraint (5) of (OPT-F) reduces to

$$\alpha^2 - \sigma^2 \geq \frac{2}{b}v(\alpha - \sigma). \tag{50}$$

On the other hand, the constraints on E imply that $\sigma = \alpha E/n \leq \alpha$ and $\sigma \geq \frac{2v}{b} - \alpha$, and hence (50) holds. The first constraint of (OPT-F) readily follows from the construction. Finally, as discussed before, $\sigma \leq \alpha$, and due to the fact that $E \geq 0$, we have $\sigma \geq 0$.

Observe that the constructed solution leads to a total misinformation of $\frac{1}{2\alpha}\sum_{i\in V}\sigma_i^2 = \alpha E^2/2n$. Therefore, by Lemma 14 it is optimal, i.e., $M^*(E) = \alpha E^2/2n$ and $\sigma_i^*(E) = \sigma = \alpha E/n$ for all $i \in V$.

□

Interestingly, this lemma implies that when the engagement requirement belongs to the given interval, there is a corresponding common threshold mechanism with appropriately chosen thresholds that solves (OPT-F). Hence, the lemma identifies a regime where the E/M frontier is achieved via common threshold mechanisms. This is due to the fact that in this regime, the platform wants to ensure a moderate engagement level while minimizing misinformation. Consistent with the results of Section 4.3, common threshold mechanisms facilitate coordination among agents on the no-engagement action for high levels of inaccuracy level and achieve low misinformation and moderate engagement.

In general, the minimum misinformation that can be achieved on the platform (i.e., the optimal objective of (OPT-M)) is strictly greater than zero. This can be seen from Theorem 3 by noting that the misinformation-minimizing thresholds can be strictly positive and hence lead to nontrivial engagement and misinformation. Thus, as the engagement requirement in (OPT-F) decreases, eventually, the corresponding constraint becomes irrelevant, and the platform's problem becomes equivalent to (OPT-M). In this case, the platform finds it optimal to implement misinformation-minimizing mechanisms. Based on the results of Section 4.3, common threshold mechanisms are optimal for this purpose, and hence they belong to the E/M frontier. Our next lemma characterizes this regime.

LEMMA 16. *Suppose $E \in [0,n]$ is such that $E \le \max\left\{0, \frac{n}{\alpha}\left(\frac{2v}{b} - \alpha\right)\right\}$.*

- *If $\frac{2v}{b} \le \alpha$, then $M^*(E) = \sigma_i^*(E) = 0$ for all $i \in V$.*
- *Otherwise,*

$$M^*(E) = \frac{n}{2\alpha}\left(\min\left\{\alpha, \frac{2v}{b} - \alpha\right\}\right)^2, \text{ and } \sigma_i^*(E) = \min\left\{\alpha, \frac{2v}{b} - \alpha\right\} \text{ for all } i \in V.$$

Proof:

If $2v/b \le \alpha$, then $E = 0$. Setting $\sigma_i = 0$ for all $i \in V$, we obtain a feasible solution of (OPT-F). Since the objective is lower bounded by zero, it follows that this is an optimal solution with $M^*(E) = 0$, and the claim trivially follows.

Suppose instead $2v/b > \alpha$. Let

$$\underline{E} = \frac{n}{\alpha}\min\left\{\alpha, \frac{2v}{b} - \alpha\right\},$$

and observe that for all $E \le \underline{E}$,

$$M^*(E) \le M^*(\underline{E}), \tag{51}$$

since the optimization problem (OPT-F) corresponding to engagement requirement E is a relaxed version of that corresponding to $\underline{E}$. Observe that by setting $\sigma_i = \frac{\alpha}{n}\underline{E} = \min\left\{\alpha, \frac{2v}{b} - \alpha\right\}$ for all $i \in V$, we obtain a feasible solution of an instance of (OPT-F) with $E \leq \underline{E}$. The fact that the constructed solution satisfies the engagement requirement as well as constraints (4) and (5) of (OPT-F) can be readily checked by substituting $\sigma_i = \min\left\{\alpha, \frac{2v}{b} - \alpha\right\}$ in these constraints. The constraint $\sigma_i \in [0, \alpha]$ holds since by construction $\sigma_i \leq \alpha$, and the assumptions of the lemma guarantee that $2v/b - \alpha \geq 0$, which in turn ensures $\sigma_i \geq 0$ for all $i \in V$.

Since the optimal solution of (OPT-F) induces a weakly lower objective value than the constructed solution, for $E \leq \underline{E}$, we have

$$M^*(E) \leq \frac{1}{2\alpha}\sum_{i \in V}\sigma_i^2 = \frac{\alpha}{2n}\underline{E}^2. \tag{52}$$

On the other hand, by Theorem 3, the mechanism with $\sigma_i = \alpha\underline{E}/n$ for all i minimizes misinformation (even in the absence of engagement constraints), and hence, for any $E \leq \underline{E}$, we have

$$M^*(E) \geq \frac{\alpha}{2n}\underline{E}^2.$$

Combining this inequality with (52), we obtain $M^*(E) = M^*(\underline{E}) = \frac{\alpha}{2n}\underline{E}^2$. Moreover, since $\sigma_i = \alpha\underline{E}/n$ is feasible in (OPT-F) for $E \leq \underline{E}$ and achieves this objective value, we conclude that $\sigma_i^*(E) = \sigma_i = \alpha\underline{E}/n$ for $E \leq \underline{E}$. □

Lemmata 15 and 16 jointly characterize the E/M frontier for $E \in \left[0, \min\left\{n, \frac{2n}{\alpha}\frac{d_{min}+v}{b}\right\}\right]$. Note that if $b \leq 2(d_{min} + v)/\alpha$, then the aforementioned interval contains $[0, n]$, and hence, the lemmata above fully characterize the E/M frontier. In other words, this implies that when the disutility parameter b is small, common threshold mechanisms are sufficient to obtain all performance levels on the E/M frontier and to achieve optimal tradeoffs between engagement and misinformation.

We next focus on complementary settings where $b > 2(d_{min} + v)\alpha$ and engagement levels higher than $2n(d_{min} + v)/(\alpha b)$ could be feasible. For $E > 2n(d_{min} + v)/(\alpha b)$, it is not possible to solve (OPT-F) in closed form and obtain the E/M frontier explicitly. Still, in the remainder of this section, we explore the structure of the E/M frontier in this regime and provide a characterization.

LEMMA 17. *Suppose $b > 2(d_{min} + v)/\alpha$ and the network is not regular. Let m denote the number of minimum-degree nodes, i.e., $m = |\{i \mid d_i = d_{min}\}|$. Then, there exists $\delta > 0$ such that, for all*

$$E \in \left(\frac{2n}{\alpha}\frac{d_{min}+v}{b}, \frac{2n}{\alpha}\frac{d_{min}+v}{b} + \delta\right),$$

we have

$$M^*(E) = \frac{2m}{\alpha}\left(\frac{v + d_{min}}{b}\right)^2 + \frac{1}{2\alpha(n-m)}\left(\alpha E - m\frac{2(v + d_{min})}{b}\right)^2.$$

Proof: Let us denote by K the set of minimum degree nodes

$$K = \{i \in V : d_i = d_{min}\}.$$

Observe that since the network is not regular, we have $K \neq V$ and $m < n$.

For notational convenience, we write

$$\underline{E} = \frac{2n(d_{min} + v)}{b\alpha}.$$

Note that for any feasible solution of the program (OPT-F) and any $i \in K$, we have

$$\sigma_i^2 \leq \frac{2}{b} v\sigma_i + \frac{2}{b} \sum_{j \in V} g_{ij} \min\{\sigma_i, \sigma_j\} \leq \frac{2}{b} v\sigma_i + \frac{2}{b} \sum_{j \in V} g_{ij}\sigma_i,$$

which implies

$$\sigma_i \leq \frac{2}{b} v + \frac{2}{b} d_{min}.$$

Therefore, the following program

$$\begin{aligned} \min_{\sigma_1,\dots,\sigma_n} \quad & \frac{1}{\alpha} \sum_{i \in V} \frac{\sigma_i^2}{2} \\ s.t. \quad & \sigma_i \leq \frac{2}{b} v + \frac{2}{b} d_{min}, \quad \text{for all } i \in K \\ & \sum_{i \in V} \sigma_i \geq \alpha E, \end{aligned} \tag{53}$$

is a relaxation of (OPT-F). Since the objective is increasing in $\{\sigma_i\}$, it can be seen that at the optimal solution of (53), the engagement requirement will be binding. Hence, explicitly imposing this constraint, we obtain the following equivalent program:

$$\begin{aligned} \min_{\sigma_1,\dots,\sigma_n} \quad & \frac{1}{\alpha} \sum_{i \in V} \frac{\sigma_i^2}{2} \\ s.t. \quad & \sigma_i \leq \frac{2}{b} v + \frac{2}{b} d_{min}, \quad \text{for all } i \in K, \\ & \sum_{i \in V} \sigma_i = \alpha E. \end{aligned} \tag{54}$$

We next obtain an optimal solution to the above program and establish that it is feasible in (OPT-F). Note that since the above program is a relaxation of (OPT-F), this readily implies the optimality of the constructed solution in the latter program.

Note that (54) is symmetric among $\{\sigma_i\}_{i\in K}$ as well as $\{\sigma_i\}_{i\in V\setminus K}$. Since the problem is convex, we conclude that there exists an optimal solution where $\sigma_i = \sigma_L$ for all $i \in K$ and $\sigma_i = \sigma_H$ for all $i \in V \setminus K$. Feasibility implies that for these solutions we have $m\,\sigma_L + (n-m)\,\sigma_H = \alpha E$. Using these observations, (54) can be written as the following single parameter optimization problem:

$$\min_{\sigma_L \in [0, 2(v+d_{min})/b]} \frac{1}{2\alpha}\left(m\sigma_L^2 + \frac{(\alpha E - m\sigma_L)^2}{n-m}\right).$$

Using the first order optimality conditions, it follows that the optimal solution of the above for a given E is given by

$$\sigma_L(E) := \min\left\{\frac{2(v+d_{min})}{b}, \frac{\alpha E}{n}\right\}.$$

For $E \geq \underline{E}$,

$$\frac{2(v+d_{min})}{b} = \alpha\frac{\underline{E}}{n} \leq \alpha\frac{E}{n}, \tag{55}$$

and therefore

$$\sigma_L(E) = \frac{2(v+d_{min})}{b} \quad \text{and} \quad \sigma_H(E) = \frac{1}{n-m}\left(\alpha E - m\frac{2(v+d_{min})}{b}\right).$$

Furthermore, we note that by (55), $\sigma_L(E) \leq \alpha E/n$, which yields

$$\sigma_L(E) \leq \sigma_H(E), \tag{56}$$

and

$$\sigma_H(\underline{E}) = \sigma_L(\underline{E}). \tag{57}$$

We conclude the proof by showing that $\sigma_i = \sigma_L(E)$, $i \in K$ and $\sigma_i = \sigma_H(E)$, $i \in V \setminus K$ is actually feasible in (OPT-F), when $E \in [\underline{E}, \underline{E} + \delta)$ for $\delta > 0$ small enough. We start by checking constraint (4) of (OPT-F). By (56), $\sigma_i \leq \sigma_j$ for all $i \in K$, $j \in V$, and therefore, for all $i \in K$, this constraint can be written as

$$\sigma_L(E)^2 \leq \frac{2}{b}v\sigma_L(E) + \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_L(E) = \frac{2(v+d_{min})}{b}\sigma_L(E),$$

which trivially holds, by the definition of $\sigma_L(E)$. Note that when $E = \underline{E}$, by (57), we get

$$\sigma_H(\underline{E})^2 = \frac{2}{b}v\sigma_H(\underline{E}) + \frac{2}{b}d_{min}\sigma_H(\underline{E}) < \frac{2}{b}v\sigma_H(\underline{E}) + \frac{2}{b}\sum_{j\in V} g_{ij}\sigma_H(\underline{E}).$$

This observation implies that constraint (4) strictly holds for $i \in V \setminus K$ when $E = \underline{E}$. Furthermore, by continuity, the constraint continues to hold for $E \in [\underline{E}, \underline{E} + \delta)$ and $\delta > 0$ small enough.

Observe that for all $i \in V$ and $E = \underline{E}$, by using (57), the constraint (5) of (OPT-F) reduces to

$$\alpha^2 - \sigma_L(\underline{E})^2 > \frac{2}{b} v(\alpha - \sigma_L(\underline{E})),$$

where the inequality is strict since the lemma assumes $\alpha > 2(d_{min} + v)/b$. Therefore, by continuity, the constraint continues to hold for $E \in [\underline{E}, \underline{E} + \delta)$ and δ small enough. The engagement requirement constraint (i.e., the first constraint of (OPT-F)) follows from the definition of $\sigma_L(E)$ and $\sigma_H(E)$. Finally, the fact that constraint $\sigma_i \in [0, \alpha]$ holds for all $i \in V$, follows from the assumption of the lemma and the observation that

$$\sigma_H(\underline{E}) = \sigma_L(\underline{E}) < \alpha,$$

which, by continuity of $\sigma_L(E), \sigma_H(E)$, implies that the constraint continues to hold for $E \in [\underline{E}, \underline{E} + \delta)$ and δ small enough.

Thus, we conclude that the constructed solution is optimal in (OPT-F) or $E \in [\underline{E}, \underline{E} + \delta)$ and $\delta > 0$ small enough. The objective associated with this solution can be given by

$$\begin{aligned} M^*(E) &= m\frac{\sigma_L(E)^2}{2\alpha} + (n-m)\frac{\sigma_H(E)^2}{2\alpha} \\ &= \frac{m}{2\alpha}\left(\frac{2(v + d_{min})}{b}\right)^2 + \frac{n-m}{2\alpha}\left(\frac{1}{n-m}\left(\alpha E - m\frac{2(v + d_{min})}{b}\right)\right)^2 \\ &= \frac{2m}{\alpha}\left(\frac{v + d_{min}}{b}\right)^2 + \frac{1}{2\alpha(n-m)}\left(\alpha E - m\frac{2(v + d_{min})}{b}\right)^2. \end{aligned}$$

□

Clearly, the engagement cannot be increased indefinitely. Specifically, there exists some $\bar{E}$ for which there does not exist a mechanism that supports engagement higher than $\bar{E}$, i.e.,

$$\bar{E} = \sup\{E : \text{(OPT-F) is feasible}\}.$$

This quantity corresponds to the maximum engagement that can be achieved by a mechanism in $\mathcal{F}$, and can be readily characterized by our results in Section 4.2.1. Our next corollary provides an expression for this point and follows immediately from Proposition 2.

Corollary 5. *Assume that $b > \frac{2}{\alpha}(v + d_{max})$. Then, $\bar{E} = \sum_{i \in V} \sigma_i/\alpha$, where $\{\sigma_i\}_{i \in V}$ is the unique solution to (7), where $\sigma_i \in (0, \alpha)$ for all $i \in V$.*

C.2. Proof of Proposition 5

In order to prove Proposition 5, we will make use of the following auxiliary results.

LEMMA 18. *(i) Suppose* $\bar{\sigma}_i > 0$*, and* ${\bar{\sigma}_i}^2 = \frac{2}{b}v\bar{\sigma}_i + \frac{2}{b}\sum_j g_{ij}\min\{\bar{\sigma}_i, \bar{\sigma}_j\}$*. Then for* $\sigma_i \in [0, \bar{\sigma}_i]$*, we have* ${\sigma_i}^2 \leq \frac{2}{b}v\sigma_i + \frac{2}{b}\sum_j g_{ij}\min\{\sigma_i, \bar{\sigma}_j\}$*.*

(ii) Suppose $\bar{\sigma}_i < \alpha$*, and* $\alpha^2 - \bar{\sigma}_i^2 = \frac{2}{b}v(\alpha - \bar{\sigma}_i) + \frac{2}{b}\sum_{j\in V} g_{ij}\max\{\bar{\sigma}_i, \bar{\sigma}_j\} - \frac{2}{b}d_i\bar{\sigma}_i$*. Then, for* $\sigma_i \in [\bar{\sigma}_i, \alpha]$*, we have* $\alpha^2 - \sigma_i^2 \geq \frac{2}{b}v(\alpha - \sigma_i) + \frac{2}{b}\sum_{j\in V} g_{ij}\max\{\sigma_i, \bar{\sigma}_j\} - \frac{2}{b}d_i\sigma_i$*.*

Proof: (i) Let $f_L(\sigma_i) = {\sigma_i}^2$, and $f_R(\sigma_i) = \frac{2}{b}v\sigma_i + \frac{2}{b}\sum_j g_{ij}\min\{\sigma_i, \bar{\sigma}_j\}$. Observe that $f_L(0) = f_R(0)$ and $f_L(\bar{\sigma}_i) = f_R(\bar{\sigma}_i)$. Since $f_L(\cdot)$ is strictly convex, and $f_R(\cdot)$ is a concave function, this immediately implies that $f_L(\sigma_i) \leq f_R(\sigma_i)$ for $\sigma_i \in [0, \bar{\sigma}_i]$, and the claim follows.

(ii) Let $f_L(\sigma_i) = \alpha^2 - \sigma_i^2$, and $f_R(\sigma_i) = \frac{2}{b}v(\alpha - \sigma_i) + \frac{2}{b}\sum_{j\in V} g_{ij}\max\{\sigma_i, \bar{\sigma}_j\} - \frac{2}{b}d_i\sigma_i$. Observe that $f_L(\alpha) = f_R(\alpha)$ and $f_L(\bar{\sigma}_i) = f_R(\bar{\sigma}_i)$. Since $f_L(\cdot)$ is strictly concave, and $f_R(\cdot)$ is a convex function, this immediately implies that $f_L(\sigma_i) \geq f_R(\sigma_i)$ for $\sigma_i \in [\bar{\sigma}_i, \alpha]$, and the claim follows. □

We next characterize how given a feasible solution of (OPT-M), another feasible solution can be obtained by perturbing the original solution. Since the constraints of (OPT-F) are a superset of those of (OPT-M), our result also allows us to reason about the feasible solutions of (OPT-F).

LEMMA 19. *Suppose* $\{\sigma_i\}$ *is feasible in* (OPT-M)*. Let* $\overline{S} = \{i \mid \sigma_i \geq \sigma_j,\ \forall j \in V\}$*, and* $\underline{S} = \{i \mid \sigma_i \leq \sigma_j,\ \forall j \in V\}$*. Assume that* $\overline{S} \neq \underline{S}$*.*

(i) Let $\{\hat{\sigma}_i\}$ *be such that* $\hat{\sigma}_i = \sigma_i - \delta$ *for* $i \in \overline{S}$ *and* $\hat{\sigma}_i = \sigma_i$ *for* $i \notin \overline{S}$*, for an arbitrarily small* $\delta > 0$*. Then* $\{\hat{\sigma}_i\}$ *is also feasible in* (OPT-M)*.*

(ii) Let $\{\hat{\sigma}_i\}$ *be such that* $\hat{\sigma}_i = \sigma_i + \delta$ *for* $i \in \underline{S}$ *and* $\hat{\sigma}_i = \sigma_i$ *for* $i \notin \underline{S}$*, for an arbitrarily small* $\delta > 0$*. If* $\sigma_j = 0$ *for* $j \in \underline{S}$*, then* $\{\hat{\sigma}_i\}$ *is also feasible in* (OPT-M)*.*

Proof: First note that updating the $\{\sigma_i\}$ as in (i) or (ii) does not change the ranking of agents in terms of their $\{\sigma_i\}$, i.e., if $\sigma_i \geq \sigma_j$ then $\hat{\sigma}_i \geq \hat{\sigma}_j$. Consequently the agents who achieve minima/maxima in the constraints (4) and (5) of (OPT-M) do not change. Since $\overline{S} \neq \underline{S}$, we have that $\sigma_i > 0$ for $i \in \overline{S}$ and $\sigma_i < \alpha$ for $i \in \underline{S}$. Thus the updates still guarantee that $\hat{\sigma}_i \in [0, \alpha]$ for arbitrarily small δ. To prove the claim we only focus on the constraints (4) and (5) of (OPT-M). Note that for each agent $i \in V$ we have one constraint in each family, where the left hand side is only a function of σ_i, and the right hand side involves either $\min\{\sigma_i, \sigma_j\}$ or $\max\{\sigma_i, \sigma_j\}$ terms. We refer to these respectively as the first/second constraint for agent i.

(i) First consider $j \notin \overline{S}$. Observe that updating $\{\sigma_i\}$ as stated relaxes the second constraint of (OPT-M) for agent $j \notin \overline{S}$. On the other hand for such j, we have $\sigma_j < \max_i \sigma_i$. Thus, the first constraint is not impacted for $j \notin \overline{S}$. Hence, the update does not violate first/second constraint for $j \notin \overline{S}$.

Consider $j \in \overline{S}$. Observe that if the first and second constraint were not binding for agent j under $\{\sigma_i\}$, then the update does not violate feasibility. Suppose the first constraint was binding, i.e., $\sigma_j^2 = \frac{2v}{b}\sigma_j + \frac{2}{b}\sum_{k\in V} g_{jk}\min\{\sigma_j, \sigma_k\}$. By Lemma 18, it follows that keeping the σ_k of remaining agents the same, the first constraint is satisfied for any $\tilde{\sigma}_j \in [0, \sigma_j]$, and in particular it is satisfied for $\tilde{\sigma}_j = \hat{\sigma}_j$. Note that this result continues to hold if σ_j is updated for all agents in $\overline{S}$ jointly, since $\sigma_{j_1} = \sigma_{j_2}$ and $\hat{\sigma}_{j_1} = \hat{\sigma}_{j_2}$ for $j_1, j_2 \in \overline{S}$.

Similarly, suppose that the second constraint was binding, i.e.,

$$\alpha^2 - \sigma_j^2 = \frac{2v}{b}(\alpha - \sigma_j) + \frac{2}{b}\sum_{k\in V} g_{jl}\max\{\sigma_j, \sigma_l\} - \frac{2}{b}d_j\sigma_j = \frac{2v}{b}(\alpha - \sigma_j). \tag{58}$$

Observe that for $i \notin \overline{S}$, we have

$$\alpha^2 - \sigma_i^2 \geq \frac{2v}{b}(\alpha - \sigma_i) + \frac{2}{b}\sum_{k\in V} g_{ik}\max\{\sigma_i, \sigma_k\} - \frac{2}{b}d_i\sigma_i \geq \frac{2v}{b}(\alpha - \sigma_i). \tag{59}$$

Let $h(\sigma) = \alpha^2 - \sigma^2 - \frac{2v}{b}(\alpha - \sigma)$. Observe that $h(\cdot)$ is strictly concave. We have from (58) and (59) that $h(\sigma_j) = 0$, and $h(\sigma_i) \geq 0$. By concavity, we have for any $\sigma \in [\sigma_i, \sigma_j]$ that $h(\sigma) \geq \min\{h(\sigma_j), h(\sigma_i)\} \geq 0$. Thus, it follows that the second constraint is satisfied for any $\tilde{\sigma}_j \in [\sigma_i, \sigma_j]$, and in particular it is satisfied for $\tilde{\sigma}_j = \hat{\sigma}_j$. Thus, we conclude that the constructed solution is feasible in (OPT-M).

(ii) First consider $j \notin \underline{S}$. Observe that updating $\{\sigma_i\}$ as stated relaxes the first constraint in (OPT-M) for $j \notin \underline{S}$. On the other hand for such j, we have $\sigma_j > \min_i \sigma_i$. Thus, the second constraint for $j \notin \underline{S}$ is not impacted. Hence, the update does not impact the first/second constraint for $j \notin \underline{S}$.

Consider $j \in \underline{S}$. Observe that if the first/second constraint was not binding for agent j under $\{\sigma_i\}$, then the update does not violate feasibility. Suppose the second constraint was binding, i.e., $\alpha^2 - \sigma_j^2 = \frac{2v}{b}(\alpha - \sigma_j) + \frac{2}{b}\sum_{k\in V} g_{jl}\max\{\sigma_j, \sigma_l\} - \frac{2}{b}d_j\sigma_j$. By Lemma 18, it follows that the second constraint is satisfied for any $\tilde{\sigma}_j \in [\sigma_j, \alpha]$ when the threshold of only agent j is updated, and in particular it is satisfied for $\tilde{\sigma}_j = \hat{\sigma}_j$. As before, this result continues to hold if σ_j is jointly updated for all $j \in \underline{S}$. Note that since $\sigma_j = 0$, the first constraint was binding for $j \in \underline{S}$, i.e.,

$$\sigma_j^2 = \frac{2v}{b}\sigma_j + \frac{2}{b}\sum_{k\in V} g_{jk}\min\{\sigma_j, \sigma_k\} = 0. \tag{60}$$

On the other hand, as a result of the aforementioned update the left hand side becomes $\delta^2 < \frac{2v}{b}\delta$. Thus, it follows that the minimum constraint is satisfied for $\hat{\sigma}_j$. Hence, we conclude that the constructed solution is feasible in (OPT-M). $\square$

LEMMA 20. *Suppose $E > 0$, and let $\{\sigma_i\}$ be an optimal solution to* (OPT-F)*. Then $\sigma_i > 0$ for all $i \in V$.*

Proof: Let $S=\{j \mid \sigma_j=0\}$. Since $E>0$, by feasibility in (OPT-F), it follows that $V\setminus S\neq 0$.

Consider another threshold mechanism, where $\hat{\sigma}_i=\sigma_i+\delta$ for $i\in S$ and arbitrarily small $\delta>0$. Assume that $\hat{\sigma}_j=\sigma_j$ for $j\notin S$. Note that the constraints of (OPT-F) consist of those of (OPT-M) and an additional constraint $\sum_i \sigma_i\geq E\alpha$. The latter constraint is trivially satisfied by $\{\hat{\sigma}_i\}$ since $\hat{\sigma}_i\geq\sigma_i$, whereas Lemma 19 implies that the rest of the constraints also hold under $\{\hat{\sigma}_i\}$. Thus, we conclude that $\{\hat{\sigma}_i\}$ is feasible in (OPT-F).

Let $\bar{S}=\{i \mid \hat{\sigma}_i\geq\hat{\sigma}_k, \forall k\in V\}$, and observe that since $V\setminus S\neq\emptyset$, we have $\bar{S}\cap S=\emptyset$, and $\bar{S}\neq\emptyset$. We next construct a third threshold mechanism with thresholds $\bar{\sigma}_i=\hat{\sigma}_i$ for $i\notin\bar{S}$, and $\bar{\sigma}_j=\hat{\sigma}_j-\bar{\delta}$ for $j\in\bar{S}$, where $\bar{\delta}|\bar{S}|=\delta|S|$. Lemma 19 implies that (for $\delta,\bar{\delta}$ small enough) these thresholds also satisfy the constraints of (OPT-F) other than $\sum_i\sigma_i\geq E\alpha$. On the other hand, by construction, we have $\sum_i\bar{\sigma}_i=\sum_i\sigma_i\geq E\alpha$. Thus, we conclude that $\{\bar{\sigma}_i\}$ is feasible in (OPT-F).

Observe that

$$\begin{aligned}\sum_i\sigma_i^2-\sum_i\bar{\sigma}_i^2&=\sum_{i\in S}(\sigma_i^2-\bar{\sigma}_i^2)+\sum_{i\in\bar{S}}(\sigma_i^2-\bar{\sigma}_i^2)\\&=\sum_{i\in S}(\sigma_i^2-(\sigma_i+\delta)^2)+\sum_{i\in\bar{S}}(\sigma_i^2-(\sigma_i-\bar{\delta})^2)\\&=|S|\delta^2-2\delta\sum_{i\in S}\sigma_i-|\bar{S}|\bar{\delta}^2+2\bar{\delta}\sum_{i\in\bar{S}}\sigma_i.\end{aligned} \tag{61}$$

Observe that $\sigma_i=\underline{\sigma}$ for all $i\in S$ and similarly $\sigma_i=\overline{\sigma}$ for all $i\in\bar{S}$, for some $\overline{\sigma}>\underline{\sigma}$. Thus, (61) implies that

$$\sum_i\sigma_i^2-\sum_i\bar{\sigma}_i^2=|S|\delta^2-|\bar{S}|\bar{\delta}^2+2(|\bar{S}|\bar{\delta}\overline{\sigma}-|S|\delta\underline{\sigma})>0, \tag{62}$$

for arbitrarily small $\delta,\bar{\delta}$. Here, the last inequality follows by noting that

$$|\bar{S}|\bar{\delta}\overline{\sigma}-|S|\delta\underline{\sigma}=|\bar{S}|\bar{\delta}(\overline{\sigma}-\underline{\sigma})+\underline{\sigma}(|\bar{S}|\bar{\delta}-|S|\delta)=|\bar{S}|\bar{\delta}(\overline{\sigma}-\underline{\sigma})>0.$$

Feasibility of $\{\bar{\sigma}_i\}$ in (OPT-F) together with (62) implies that $\{\sigma_i\}$ is not optimal in (OPT-F). Thus, we obtain a contradiction. Hence $S=\emptyset$, and the claim follows. □

LEMMA 21. *Let F be a common threshold mechanism with threshold $\sigma\in(0,\alpha)$, and $Q\in\mathcal{Q}(F)$ be a corresponding equilibrium such that (F,Q) tuple solves (12) for a given $E\geq 0$. Denote by s the public signal of F. The set of agents who take action 1 when $s=0$ and action 0 when $s=1$ is empty.*

Proof: Let $\{\mu_i\}_{i\in V}$ denote agents' strategies at equilibrium Q, and define

$$S:=\{i\in V \mid \mu_i(0)=1 \quad \textit{and} \quad \mu_i(1)=0\}. \tag{63}$$

Suppose towards a contradiction that $S \neq \emptyset$. Consider a new mechanism $\hat{F}$ with associated random variables $\{\hat{s}_i\}_{i\in V}$ such that for all $\omega \in \Omega$ and $i \in V$, we have:

$$\hat{s}_i(\omega) = x \iff \mu_i(s(\omega)) = x,$$

where $x \in \{0,1\}$. Note that $\hat{F}$ is a straightforward mechanism whose straightforward equilibrium achieves the same engagement/misinformation level as Q, as argued in the proof of Lemma 2.

Note that, by (63) and the definition of $\hat{s}_i$, for all $i \in S$, we have

$$\hat{s}_i(\omega) = 1 \iff \epsilon(\omega) > \sigma.$$

We now construct a third mechanism $\tilde{F}$, whose signals are such that

$$\tilde{s}_i(\omega) = \hat{s}_i(\omega)$$

for all $\omega \in \Omega$ and $i \in V \setminus S$, and

$$\tilde{s}_i(\omega) = 1 \iff \epsilon(\omega) \leq \alpha - \sigma,$$

for all $\omega \in \Omega$ and $i \in S$. We claim that $\tilde{F}$ has a straightforward equilibrium $\tilde{Q}$. Indeed, since under the new mechanism engagement signals ($\tilde{s}_i = 1$) are associated with lower inaccuracy levels, we have

$$\mathbb{E}[\epsilon \mid \tilde{s}_i = 1] \leq \mathbb{E}[\epsilon \mid \hat{s}_i = 1],$$
$$\mathbb{E}[\epsilon \mid \tilde{s}_i = 0] \geq \mathbb{E}[\epsilon \mid \hat{s}_i = 0],$$

for all agents $i \in V$. In addition, since agents in S now receive engagement signal when $\epsilon(\omega) \leq \alpha - \sigma$, for all $i \in V$ we have

$$\sum_{j\in V} g_{ij}\mathbb{E}[\tilde{s}_j \mid \tilde{s}_i = 1] \geq \sum_{j\in V} g_{ij}\mathbb{E}[\hat{s}_j \mid \hat{s}_i = 1],$$

while

$$\sum_{j\in V} g_{ij}\mathbb{E}[\tilde{s}_j \mid \tilde{s}_i = 0] \leq \sum_{j\in V} g_{ij}\mathbb{E}[\hat{s}_j \mid \hat{s}_i = 0].$$

Therefore, since it is incentive compatible for all agents in V to follow the recommendation of $\hat{F}$, it is also incentive compatible to follow the recommendation of $\tilde{F}$. Hence, $\tilde{Q}$ is a straightforward equilibrium. By construction

$$E(\tilde{Q}) = E(Q),$$

but

$$M(Q) - M(\tilde{Q}) = \sum_{i \in S} \frac{1}{\alpha} \left(\int_{[\sigma,\alpha]} z dz - \frac{(\alpha-\sigma)^2}{2} \right) = \sum_{i \in S} \frac{1}{\alpha}(\alpha-\sigma)\sigma > 0.$$

On the other hand,

$$M(Q) \leq M(\tilde{Q}),$$

since the performance level of Q is on the E/M frontier. Thus, we obtain a contradiction and $S = \emptyset$.

□

We are now ready to proceed to the proof of Proposition 5. One direction follows readily from Lemmata 15 and 16. We now prove that if $E > \bar{E}$, where

$$\bar{E} = \frac{2n}{b\alpha}(v + d_{min}) < n,$$

then common thresholds cannot induce performance that is on the E/M frontier. Consider some $E' > \bar{E}$. Assume, for the purposes of contradiction that there exists some common threshold mechanism $\hat{F}$ with threshold $\hat{\sigma} \in [0, \alpha]$ that admits an equilibrium $Q = \{\mu_i\}$ such that $E(Q) = E'$ and its performance level $(E(Q), M(Q))$ is on the E/M Frontier. Note that if $\hat{\sigma} = 0$ or $\hat{\sigma} = \alpha$ the mechanism almost surely sends the same signal. Thus, it follows that the performance levels achievable by mechanisms with these thresholds are identical. Thus, we restrict attention to $\hat{\sigma} > 0$.

If $\hat{\sigma} \in (0, \alpha)$, by Lemma 21, the set of agents who take action 1 when $s = 0$ and action 0 when $s = 0$ is empty. Therefore, there are three sets of agents:

$$S_1 = \{i : \mu_i(0) = 1, \mu_i(1) = 1\},$$
$$S_2 = \{i : \mu_i(0) = 0, \mu_i(1) = 0\},$$
$$S_3 = \{i : \mu_i(0) = 0, \mu_i(1) = 1\}.$$

On the other hand if $\hat{\sigma} = \alpha$ agents trivially belong in one of the aforementioned sets.

Following the same approach as in the proof of Lemma 2, the threshold mechanism $\tilde{F}$ with thresholds $\tilde{\sigma}_i = \alpha$ for $i \in S_1$, $\tilde{\sigma}_i = 0$ for $i \in S_2$ and $\tilde{\sigma}_i = \hat{\sigma}$ for all $i \in S_3$ is a straightforward threshold mechanism whose straightforward equilibrium achieves the same performance level. Since performance level associated with Q is on the E/M Frontier, $\{\tilde{\sigma}_i\}$ solves (OPT-F). By Lemma 20, $S_2 = \emptyset$. By the incentive compatibility constraint (4), for all agents $i \in S_1$ we get

$$\alpha^2 \leq \frac{2}{b} v\alpha + \frac{2}{b} d_i \alpha \leq \frac{2}{b} v\alpha + \frac{2}{b} d_{max}\alpha,$$

which contradicts with the assumption of the proposition, and hence $S_1 = \emptyset$. Therefore, $S_3 = V$. Observe that by feasibility, we have $\sum_{i\in V} \tilde{\sigma}_i = n\hat{\sigma} \geq \alpha E'$. On the other hand, from the incentive compatibility constraint (4) and the fact that $\tilde{\sigma}_i = \hat{\sigma}$ for all $i \in S_3 = V$, we get

$$\hat{\sigma}^2 \leq \frac{2}{b}\hat{\sigma}v + \frac{2}{b}d_i\hat{\sigma},$$

for all $i \in V$, and hence

$$\hat{\sigma}^2 \leq \frac{2}{b}\hat{\sigma}v + \frac{2}{b}d_{min}\hat{\sigma}.$$

Rearranging terms, this yields $\hat{\sigma} \leq \frac{2}{b}(v + d_{min}) = \alpha\frac{\bar{E}}{n} < \alpha\frac{E'}{n}$. Thus, we obtain a contradiction to $n\hat{\sigma} \geq \alpha E'$, and conclude that for $E = E' > \bar{E}$ common thresholds cannot induce performance that is on the E/M frontier.

The second part of the proposition follows from Lemma 17. □