Supplementary Material: Sharp Thresholds of the Information Cascade Fragility Under a Mismatched Model

A Proof of Theorem 2

We start with the case where $\liminf_{t\to\infty}\frac{\|\mathcal{Q}_t\|}{\log t}=0$. To lower bound the error probability we need to understand first how the mismatched MAP decisions evolve over time/players. We next establish some notation, which will simplify the analysis. For $i\in\{1,2\}$ and $t\geq 1$, recall our notation in (4)–(8). Also, for $x\in\{1,2\}$, let

$$\phi_i(x)\triangleq\frac{\alpha}{\alpha+\beta}\mathbb{1}\left[x=i\right]+\frac{\beta}{\alpha+\beta}\mathbb{1}\left[x\neq i\right]. \tag{22}$$

Note that $\hat{\mathsf{D}}_i^t=\hat{\mathsf{L}}_i^{t-1}\phi_i(X_t)$. Finally, we define

$$\mathsf{R}_t\triangleq\frac{\hat{\mathsf{L}}_1^t}{\hat{\mathsf{L}}_2^t},\qquad \mathsf{R}'_t\triangleq\frac{\hat{\mathsf{D}}_1^t}{\hat{\mathsf{D}}_2^t}. \tag{23}$$

It is clear that

$$\mathsf{R}'_t=\mathsf{R}_{t-1}\frac{\phi_1(X_t)}{\phi_2(X_t)}. \tag{24}$$

Since the ratio $\phi_1(X_t)/\phi_2(X_t)$ can take values in $\{\beta/\alpha,\alpha/\beta\}$, we have three possible cases only:

- If $\mathsf{R}_{t-1}<\beta/\alpha$ then, clearly $\mathsf{R}'_t<1$, irrespective of the value of X_t. Thus, $Z_t=1$ only if player t is irrational and $X_t=1$. Otherwise, $Z_t=2$.
- If $\mathsf{R}_{t-1}\in[\beta/\alpha,\alpha/\beta]$ then it can be easily shown that $Z_t=X_t$.
- If $\mathsf{R}_{t-1}>\alpha/\beta$ then, clearly $\mathsf{R}'_t>1$, irrespective of the value of X_t. Thus, $Z_t=2$ only if player t is irrational and $X_t=2$. Otherwise, $Z_t=1$.

Let $\{\mathcal{F}_t\}_{t\geq 0}$ denote the filtration spanned by $\{Z_t\}_{t\geq 1}$. Let $\hat{\mathbb{P}}_1(Z_t=i|\mathcal{F}_{t-1})$ be the probability that the guess by player t is i, given the history and $\theta=1$, and the evaluation of this probability is with respect to the mismatched revealers probabilities $\mathcal{Q}$. Then, based on the above, we have

$$\hat{\mathbb{P}}_1(Z_t=1|\mathcal{F}_{t-1}) =\begin{cases}\frac{\alpha}{\alpha+\beta}q_t, & \text{if } \mathsf{R}_{t-1}<\beta/\alpha,\\ \frac{\alpha}{\alpha+\beta}, & \text{if } \mathsf{R}_{t-1}\in[\beta/\alpha,\alpha/\beta],\\ 1-\frac{\beta}{\alpha+\beta}q_t, & \text{if } \mathsf{R}_{t-1}>\alpha/\beta,\end{cases} \tag{25}$$

and

$$\hat{\mathbb{P}}_2(Z_t=1|\mathcal{F}_{t-1}) =\begin{cases}\frac{\beta}{\alpha+\beta}q_t, & \text{if } \mathsf{R}_{t-1}<\beta/\alpha,\\ \frac{\beta}{\alpha+\beta}, & \text{if } \mathsf{R}_{t-1}\in[\beta/\alpha,\alpha/\beta],\\ 1-\frac{\alpha}{\alpha+\beta}q_t, & \text{if } \mathsf{R}_{t-1}>\alpha/\beta.\end{cases} \tag{26}$$

There are two main sources for wrong action: 1) the t^{th} player is irrational, which happens with probability p_t, and his draw is of minority color type, or, 2) the t^{th} player is rational, but his mismatched MAP estimate is wrong. Accordingly, we can write

$$\begin{aligned}\mathsf{P}_{e,t}(\mathcal{P}^\star,\mathcal{Q})&=\mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq\theta\right)\cdot(1-p_t^\star)\\ &\quad+\mathbb{P}(X_t\neq\theta)\cdot p_t^\star\\ &\geq\mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq\theta\right)\cdot(1-p_t^\star)\\ &=\mathbb{P}_1\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)=2\right)\cdot(1-p_t^\star),\end{aligned} \tag{27}$$

where the last equality follows by symmetry. We next lower bound the probability term at the r.h.s. of (27). To this end, we define an event that implies that the output of the mismatched MAP is 2, given that $\theta=1$. Let $\bar{t}_1,\bar{t}_2\leq t^\star$ be three natural numbers, to be defined in the sequel. Let $\mathsf{rev}(t^\star)\triangleq\{i_1,i_2,\ldots,i_{t^\star}\}$ be the set of the *last* $t^\star$ revealers. Define the following event

$$\begin{aligned}\mathcal{E}(\bar{t}_1,\bar{t}_2,t^\star)\triangleq\{&X_i=2,\forall i\in\mathsf{rev}(t^\star)\\ &\cup\{i_{\bar{t}_1},i_{\bar{t}_1}+1\ldots,i_{\bar{t}_1}+\bar{t}_2\}\},\end{aligned} \tag{28}$$

namely, it is the event that the last $t^\star$ revealers are such that their private signal is "2", and all consecutive players $i_{\bar{t}_1},i_{\bar{t}_1}+1,\ldots,i_{\bar{t}_1}+\bar{t}_2$ (either revealers or rationals) are such that their private signal is "2" as well. We claim that by carefully choosing the values of $\bar{t}_1$, $\bar{t}_2$, and $t^\star$, we can show that $\mathcal{E}(\bar{t}_1,\bar{t}_2,t^\star)$ implies that the t^{th} player (mismatched MAP) guess is "2". To this end, note that the first two individuals follow their private signals, that is $Z_1=X_1$ and $Z_2=X_2$. Therefore, if $X_1=X_2=1$, then $\mathsf{R}_2=(\alpha/\beta)^2$, otherwise, $\mathsf{R}_2=(\beta/\alpha)^2$. Depending on either one of the above situations, the proceeding guesses depend on whether a student is a revealer or not, and value of the likelihood R_t at each step t. Accordingly, at step i_1-1, the worst-case (largest) attainable value of the likelihood ratio is

$$\mathsf{R}_{i_1-1}=\left(\frac{\alpha}{\beta}\right)^2\cdot\prod_{i=3}^{i_1-1}\frac{1-\frac{\beta}{\alpha-\beta}q_i}{1-\frac{\alpha}{\alpha-\beta}q_i} \tag{29}$$

$$\leq\left(\frac{\alpha}{\beta}\right)^2 e^{\frac{\alpha-\beta}{\alpha+\beta}\sum_{i=3}^{i_1-1}q_i}, \tag{30}$$

which corresponds to the situation where $X_1=X_2=1$, and the proceeding players decisions up to $t\leq i_1-1$

are $Z_t = 1$. Now, over $\mathcal{E}(\bar{t}_1, \bar{t}_2, t^\star)$, we know that the i_1 player is a revealer and its private information is $X_{i_1} = 2$. Since he is a revealer its decision will be $Z_{i_1} = 2$, which implies that the likelihood ratio is

$$\mathsf{R}_{i_1} = \left(\frac{\alpha}{\beta}\right) \cdot \prod_{i=3}^{i_1-1} \frac{1 - \frac{\beta}{\alpha-\beta} q_i}{1 - \frac{\alpha}{\alpha-\beta} q_i}. \tag{31}$$

The index of the next revealer is i_2. Thus, since $\mathsf{R}_{i_1} > \alpha/\beta$, the decisions of the proceeding players up to player i_2, are "1", which imply that

$$\mathsf{R}_{i_2-1} = \left(\frac{\alpha}{\beta}\right) \cdot \prod_{i=3}^{i_2-1} \frac{1 - \frac{\beta}{\alpha-\beta} q_i}{1 - \frac{\alpha}{\alpha-\beta} q_i}. \tag{32}$$

Then, since i_2 is a revealer and its private information is $X_{i_2} = 2$, we have

$$\mathsf{R}_{i_2} = \prod_{i=3}^{i_1-1} \frac{1 - \frac{\beta}{\alpha-\beta} q_i}{1 - \frac{\alpha}{\alpha-\beta} q_i}, \tag{33}$$

and the above process continues. In particular, assuming that $\mathsf{R}_{i_{j-1}} > \alpha/\beta$, the likelihood ratio after the $i_j{}^{\text{th}}$ decision is

$$\mathsf{R}_{i_j} = \left(\frac{\beta}{\alpha}\right)^{j-2} \prod_{i=3}^{i_j-1} \frac{1 - \frac{\beta}{\alpha-\beta} q_i}{1 - \frac{\alpha}{\alpha-\beta} q_i}. \tag{34}$$

We denote by $\bar{t}_1$ the index at which the likelihood ratio $\mathsf{R}_{i_{\bar{t}_1}}$, after the $i_{\bar{t}_1}{}^{\text{th}}$ decision, is in the interval $[\beta/\alpha, \alpha/\beta]$. Indeed, when this happens, according to (25)–(26), the likelihood ratio $\mathsf{R}_{i_{\bar{t}_1}}$ will be multiplied by either β/α or α/β depending on whether the private information is "2" or "1", respectively, until the likelihood ratio value will be either below β/α or above α/β. Accordingly, over $\mathcal{E}(\bar{t}_1, \bar{t}_2, t^\star)$, we force the private information of players $i_{\bar{t}_1}, i_{\bar{t}_1} + 1, \ldots, i_{\bar{t}_1} + \bar{t}_2$ to be "2", so that the corresponding likelihoods will be multiplied by β/α, until the likelihood ratio value will get bellow β/α, and accordingly, $\bar{t}_2$ is chosen such that

$$\mathsf{R}_{\bar{t}_1+\bar{t}_2} = \left(\frac{\beta}{\alpha}\right)^{\bar{t}_1+\bar{t}_2-2} \prod_{i=3}^{i_{\bar{t}}-1} \frac{1 - \frac{\beta}{\alpha-\beta} q_i}{1 - \frac{\alpha}{\alpha-\beta} q_i}, \tag{35}$$

will be less than β/α. Next, due to the fact that over $\mathcal{E}(\bar{t}_1, \bar{t}_2, t^\star)$ the leftover revealers are such that their private information is "2", and the likelihood ratio is below β/α (and so MAP outputs "2"), it is clear that all the leftover players decisions will be "2". To assure that we take enough revealers at the end, we chose $t^\star$ such that,

$$\left(\frac{\beta}{\alpha}\right)^{t^\star-2} \prod_{i=3}^{t-t^\star} \frac{1 - \frac{\beta}{\alpha-\beta} q_i}{1 - \frac{\alpha}{\alpha-\beta} q_i} < \frac{\beta}{\alpha}, \tag{36}$$

which reflects the case where the players decision are always "1" up to time $t - t^\star$, and then the left over $t^\star$ players are all revealers with "2" being their private information. It is evident that (36) holds if

$$t^\star \geq 3 + \frac{\alpha - \beta}{\alpha + \beta} \cdot \frac{\|\mathcal{Q}_t\|}{\log(\alpha/\beta)}. \tag{37}$$

Note that when $\|\mathcal{Q}_t\|$ is a finite number which happens to be the case when $q_t = o(t^{-1})$ (as opposed to $\|\mathcal{P}_t^\star\|$ which grows logarithmically with t), implies that $t^\star$ is finite, which in turn is the reason for the fact that the error probability is finite. Also, by the same token, it is clear that $\bar{t}_1$ and $\bar{t}_2$ are finite too, with $\bar{t}_2 < t^\star$. The latter implies that $\mathcal{E}(\bar{t}_1, t^\star, t^\star) \subseteq \mathcal{E}(\bar{t}_1, \bar{t}_2, t^\star)$. Finally, we need to make sure that the size of the set of all revealers, denoted by $\mathsf{Rev}_t \triangleq \{i \in [t] :$ player i is revealer$\}$ is bigger than $t^\star$ with high probability. Let $M_t \triangleq \sum_{i=1}^t p_i \sim \log t$. Then, by Chernoff's bound $\mathbb{P}\left(|\mathsf{Rev}_t| \leq t^\star\right) \leq \exp(t^\star \log \frac{M_t}{t^\star} - M_t + t^\star) = O(t^{-1}) \leq 1/2$. Thus, we get that

$$\begin{aligned}
&\mathbb{P}_1\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1}, X_t) = 2\right) \geq \mathbb{P}_1\left[\mathcal{E}(\bar{t}_1, \bar{t}_2, t^\star)\right] \\
&\geq \sum_{\mathcal{A}:|\mathcal{A}|>t^\star} \mathbb{P}\left(\mathsf{Rev}_t = \mathcal{A}\right) \mathbb{P}_1\left[\mathcal{E}(\bar{t}_1, \bar{t}_2, t^\star) \middle| \mathsf{Rev}_t = \mathcal{A}\right] \\
&\geq \sum_{\mathcal{A}:|\mathcal{A}|>t^\star} \mathbb{P}\left(\mathsf{Rev}_t = \mathcal{A}\right) \mathbb{P}_1\left[\mathcal{E}(\bar{t}_1, t^\star, t^\star) \middle| \mathsf{Rev}_t = \mathcal{A}\right] \\
&\geq \left(\frac{\beta}{\alpha+\beta}\right)^{2t^\star} \cdot \mathbb{P}\left(|\mathsf{Rev}_t| > t^\star\right) \\
&\geq \frac{1}{2}\left(\frac{\beta}{\alpha+\beta}\right)^{2t^\star}.
\end{aligned} \tag{38}$$

Therefore, using (27), (38), and the fact that $\liminf_{t\to\infty} \frac{\|\mathcal{Q}_t\|}{\log t} = 0$ we get

$$\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) = \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star, \mathcal{Q})}{\log t} \tag{39}$$

$$\leq 2\log\left(\frac{\alpha+\beta}{\beta}\right) \cdot \liminf_{t\to\infty} \frac{t^\star}{\log t} = 0. \tag{40}$$

Since it is clear that $\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) \geq 0$ we may conclude that in this regime $\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) = 0$. Finally, proving that $\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) = 0$ for the case $\liminf_{t\to\infty} \frac{\|\mathcal{Q}_t\|}{\log t} = \infty$ follows from Theorem 3 (specifically, using the fact that for $\rho > \rho_1$ we have $\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) = 0$), and a monotonicity property of the error probability w.r.t. the revealing probabilities $\mathcal{Q}$. We provide the complete details in Appendix B.2.

B Proof of Theorem 3

We split the proofs into several upper and lower bounds, which together characterize tightly the asymptotic learning rate. Note that by "lower bounds" ("upper bounds") we mean lower- (upper-) bounding the learning rate by upper- (lower-) bounding the error probability.

B.1 Lower Bound: $\rho_0 \leq \rho \leq \rho_1$

We analyze next the probability of wrong action by the t^{th} player. Accordingly, there are two main sources for wrong action: 1) the t^{th} player is irrational, which happens with probability p_t, and his draw is of minority color type, or, 2) the t^{th} player is rational, but his mismatched MAP estimate is wrong. Accordingly, we can write

$$\begin{aligned}\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q}) &= \mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq\theta\right)\cdot(1-p_t)\\ &\quad+\mathbb{P}(X_t\neq\theta)\cdot p_t \qquad (41)\\ &= \mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq\theta\right)\cdot(1-p_t)\\ &\quad+\frac{\beta}{\alpha+\beta}\cdot p_t. \qquad (42)\end{aligned}$$

Therefore, to upper bound $\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q})$ we need to upper bound the probability that the mismatched MAP estimator is incorrect. To this end, we next establish some notation, which will simplify the analysis. For $i\in\{1,2\}$ and $t\geq 1$, recall our notations in (4)–(8), as well as the definitions in (22)–(24). Finally, recall that depending on the value that R_{t-1} takes, there are three modes of operation for the MAP estimator (see, the paragraph following (24)). In particular, the error probability associated with the MAP estimator can be upper bounded as follows,

$$\begin{aligned}\mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq\theta\right) &= \mathbb{P}_1\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq 1\right)\\ &= \mathbb{P}_1\left(\mathsf{R}_t'\leq 1\right)\\ &\leq \mathbb{P}_1\left(\mathsf{R}_{t-1}\leq\frac{\alpha}{\beta}\right). \qquad (43)\end{aligned}$$

Thus, it is suffice to upper bound $\mathbb{P}_1(\mathsf{R}_{t-1}\leq\alpha/\beta)$. To this end, we use similar techniques used in [Peres et al., 2018, Sec. 2.2], but with modifications which handle the mismatch aspect of our model. Let $\{\mathcal{F}_t\}_{t\geq 0}$ denote the filtration spanned by $\{Z_t\}_{t\geq 1}$. Then, for $\lambda\in[0,1]$, we have

$$\begin{aligned}&\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}\middle|\mathcal{F}_{t-1}\right]\\ &= \mathbb{E}_1\left[\left(\frac{\hat{\mathbb{P}}_1(Z_t|\mathcal{F}_{t-1})}{\hat{\mathbb{P}}_2(Z_t|\mathcal{F}_{t-1})}\right)^{-\lambda}\middle|\mathcal{F}_{t-1}\right] \qquad (44)\\ &= \sum_{i\in\{1,2\}}\mathbb{P}_1(Z_t=i|\mathcal{F}_{t-1})\left(\frac{\hat{\mathbb{P}}_1(Z_t=i|\mathcal{F}_{t-1})}{\hat{\mathbb{P}}_2(Z_t=i|\mathcal{F}_{t-1})}\right)^{-\lambda}, \qquad (45)\end{aligned}$$

where $\mathbb{P}_1(Z_t=i|\mathcal{F}_{t-1})$ is the probability that player t guess is i given the history and $\theta=1$, and the evaluation of this probability is with respect to the underlying true revealers probabilities $\mathcal{P}$. On the other hand, $\hat{\mathbb{P}}_1(Z_t=i|\mathcal{F}_{t-1})$ is the probability that player t guess is i given the history and $\theta=1$, and the evaluation of this probability is with respect to the mismatched revealers probabilities $\mathcal{Q}$. Accordingly, the values of these probabilities are given in (25)–(26), and

$$\mathbb{P}_1(Z_t=1|\mathcal{F}_{t-1}) = \begin{cases}\frac{\alpha}{\alpha+\beta}p_t, & \text{if } \mathsf{R}_{t-1}<\beta/\alpha,\\ \frac{\alpha}{\alpha+\beta}, & \text{if } \mathsf{R}_{t-1}\in[\beta/\alpha,\alpha/\beta],\\ 1-\frac{\beta}{\alpha+\beta}p_t, & \text{if } \mathsf{R}_{t-1}>\alpha/\beta.\end{cases} \qquad (46)$$

Therefore, we have,

$$\begin{aligned}&\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}\middle|\mathcal{F}_{t-1},\mathsf{R}_{t-1}<\frac{\beta}{\alpha}\right]\\ &= \frac{\alpha^{1-\lambda}\beta^{\lambda}}{\alpha+\beta}p_t+\left(1-\frac{\alpha}{\alpha+\beta}p_t\right)\left[\frac{1-\frac{\beta}{\alpha+\beta}q_t}{1-\frac{\alpha}{\alpha+\beta}q_t}\right]^{\lambda},\\ &\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}\middle|\mathcal{F}_{t-1},\mathsf{R}_{t-1}\in\left[\frac{\beta}{\alpha},\frac{\alpha}{\beta}\right]\right]\\ &= \frac{\alpha^{1-\lambda}\beta^{\lambda}+\alpha^{\lambda}\beta^{1-\lambda}}{\alpha+\beta},\\ &\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}\middle|\mathcal{F}_{t-1},\mathsf{R}_{t-1}>\frac{\alpha}{\beta}\right]\\ &= \frac{\alpha^{\lambda}\beta^{1-\lambda}}{\alpha+\beta}p_t+\left(1-\frac{\beta}{\alpha+\beta}p_t\right)\left[\frac{1-\frac{\alpha}{\alpha+\beta}q_t}{1-\frac{\beta}{\alpha+\beta}q_t}\right]^{\lambda}. \qquad (47)\end{aligned}$$

Since we assume that both p_t and q_t decay with t, we use the fact that $(1-\delta)^{\lambda}=1-\lambda\cdot\delta+\Theta(\delta^2)$, as $\delta\to 0$. Let

$$f_{\lambda}\equiv f_{\lambda}(\alpha,\beta)\triangleq\frac{\alpha-\alpha^{1-\lambda}\beta^{\lambda}}{\alpha+\beta}, \qquad (48)$$

$$g_{\lambda}\equiv g_{\lambda}(\alpha,\beta)\triangleq\frac{(\alpha-\beta)\lambda}{\alpha+\beta}, \qquad (49)$$

$$h_{\lambda}\equiv h_{\lambda}(\alpha,\beta)\triangleq\frac{\beta-\alpha^{\lambda}\beta^{1-\lambda}}{\alpha+\beta}. \qquad (50)$$

Then, we have

$$\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}\middle|\mathcal{F}_{t-1},\mathsf{R}_{t-1}<\frac{\beta}{\alpha}\right] = 1-f_\lambda\cdot p_t+g_\lambda\cdot q_t+O(p_t^2+q_t^2), \quad (51)$$

$$\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}\middle|\mathcal{F}_{t-1},\mathsf{R}_{t-1}\in\left[\frac{\beta}{\alpha},\frac{\alpha}{\beta}\right]\right] = \frac{\alpha^{1-\lambda}\beta^{\lambda}+\alpha^{\lambda}\beta^{1-\lambda}}{\alpha+\beta}, \quad (52)$$

$$\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}\middle|\mathcal{F}_{t-1},\mathsf{R}_{t-1}>\frac{\alpha}{\beta}\right] = 1-h_\lambda\cdot p_t-g_\lambda\cdot q_t+O(p_t^2+q_t^2). \quad (53)$$

Define the following two sets:

$$\mathcal{A}_t\triangleq\left\{i\in[t]:\mathsf{R}_{i-1}>\frac{\alpha}{\beta}\right\}, \quad (54)$$

and $\mathcal{B}_t\triangleq\mathcal{A}_t^c=[t]\setminus\mathcal{A}_t$. Define also,

$$\mathsf{R}_t^{(1)}\triangleq\prod_{i\in\mathcal{A}_t}\frac{\mathsf{R}_i}{\mathsf{R}_{i-1}} \quad (55)$$

and

$$\mathsf{R}_t^{(2)}\triangleq\prod_{i\in\mathcal{B}_t}\frac{\mathsf{R}_i}{\mathsf{R}_{i-1}}. \quad (56)$$

Note that $\mathsf{R}_t=\mathsf{R}_t^{(1)}\cdot\mathsf{R}_t^{(2)}$. We these definitions, using (51)–(53), we note that there exist some constants C and C' independent of λ, such that for all λ,

$$\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}e^{h_\lambda p_t+g_\lambda q_t}\middle|\mathsf{R}_{t-1},t\in\mathcal{A}_t\right]\leq e^{C'(p_t^2+q_t^2)}, \quad (57)$$

and

$$\mathbb{E}_1\left[\left(\frac{\mathsf{R}_t}{\mathsf{R}_{t-1}}\right)^{-\lambda}e^{f_\lambda p_t-g_\lambda q_t}\middle|\mathsf{R}_{t-1},t\in\mathcal{B}_t\right]\leq e^{C(p_t^2+q_t^2)}. \quad (58)$$

Recall that $\|\mathcal{P}_t\|=\sum_{i=1}^t p_i$ and $\|\mathcal{Q}_t\|=\sum_{i=1}^t q_i$. Also, let $\Gamma_t^p\triangleq\sum_{i\in\mathcal{A}_t}^t p_i$ and $\Gamma_t^q\triangleq\sum_{i\in\mathcal{A}_t}^t q_i$. By induction, we can easily see that there exists a constant C such that for any $\lambda_1,\lambda_2\in[0,1]$,

$$\mathbb{E}_1\left[\left(\mathsf{R}_t^{(1)}\right)^{-\lambda_1}e^{h_{\lambda_1}\Gamma_t^p+g_{\lambda_1}\Gamma_t^q}\left(\mathsf{R}_t^{(2)}\right)^{-\lambda_2} e^{f_{\lambda_2}(\|\mathcal{P}_t\|-\Gamma_t^p)-g_{\lambda_2}(\|\mathcal{Q}_t\|-\Gamma_t^q)}\right]\leq e^{C\sum_{i=1}^t(p_i^2+q_i^2)}, \quad (59)$$

and since $\sum_{i=1}^t(p_i^2+q_i^2)$ is finite, we can upper bound the r.h.s. of (59) by a constant C_0. Now, as was shown in [Peres et al., 2018, Appendix A], the condition $\mathsf{R}_t\leq\frac{\alpha}{\beta}$ implies that $\mathsf{R}_t^{(1)}\leq 1$. Thus, we may write

$$\mathbb{P}_1\left(\mathsf{R}_t\leq\frac{\alpha}{\beta}\right)\leq\mathbb{P}_1\left(\mathsf{R}_t\leq\frac{\alpha}{\beta},\ t^{-C_1}\leq\mathsf{R}_t^{(1)}\leq 1\right) + \mathbb{P}_1\left(\mathsf{R}_t^{(1)}\leq t^{-C_1}\right). \quad (60)$$

A simple application of multiplicative Chrenoff's bound shows that the second term on the r.h.s. of the above inequality is upper bounded by t^{-2}, for some constant $C_1<\infty$. We next upper bound the first term on the r.h.s. of the above inequality. Since $\mathsf{R}_t=\mathsf{R}_t^{(1)}\cdot\mathsf{R}_t^{(2)}$, we can write

$$\begin{aligned}\mathbb{P}_1&\left(\mathsf{R}_t\leq\frac{\alpha}{\beta},\ t^{-C_1}\leq\mathsf{R}_t^{(1)}\leq 1\right)\\ &\leq\sum_{x=0}^{C_1\log t}\mathbb{P}_1\left(\mathsf{R}_t^{(1)}\in\left[e^{-(x+1)},e^{-x}\right],\mathsf{R}_t^{(2)}\leq e^{x+C_3}\right)\\ &\leq\sum_{x=0}^{C_1\log t}\mathbb{P}_1\left(\mathsf{R}_t^{(1)}\leq e^{-x},\mathsf{R}_t^{(2)}\leq e^{x+C_3}\right),\end{aligned} \quad (61)$$

where $C_3\triangleq 1+\log\frac{\alpha}{\beta}$. Then, for any $\lambda_1,\lambda_2\in[0,1]$, we have

$$\begin{aligned}\mathbb{P}_1&\left(\mathsf{R}_t^{(1)}\leq e^{-x},\mathsf{R}_t^{(2)}\leq e^{x+C_3}\right)\\ &=\mathbb{P}_1\left[\left(\mathsf{R}_t^{(1)}\right)^{-\lambda_1}\geq e^{\lambda_1 x},\left(\mathsf{R}_t^{(2)}\right)^{-\lambda_2}\geq e^{-\lambda_2(x+C_3)}\right],\end{aligned} \quad (62)$$

and then,

$$\begin{aligned}\mathbb{P}_1&\left(\mathsf{R}_t^{(1)}\leq e^{-x},\mathsf{R}_t^{(2)}\leq e^{x+C_3}\right)\\ &=\mathbb{P}_1\Bigg[\left(\mathsf{R}_t^{(1)}\right)^{-\lambda_1}e^{h_{\lambda_1}\Gamma_t^p+g_{\lambda_1}\Gamma_t^q}\geq e^{\lambda_1 x+h_{\lambda_1}\Gamma_t^p+g_{\lambda_1}\Gamma_t^q},\\ &\qquad\left(\mathsf{R}_t^{(2)}\right)^{-\lambda_2}e^{f_{\lambda_2}(\|\mathcal{P}_t\|-\Gamma_t^p)-g_{\lambda_2}(\|\mathcal{Q}_t\|-\Gamma_t^q)}\\ &\qquad\geq e^{-(x+C_3)\lambda_2+f_{\lambda_2}(\|\mathcal{P}_t\|-\Gamma_t^p)-g_{\lambda_2}(\|\mathcal{Q}_t\|-\Gamma_t^q)}\Bigg].\end{aligned} \quad (63)$$

Now using the facts that $\mathbb{P}[X_1\geq X_2,X_3\geq X_4]\leq\mathbb{P}[X_1\cdot X_3\geq X_2\cdot X_4]$, for non-negative random variables X_1^4, and Markov inequality along with (59), we get

$$\begin{aligned}\mathbb{P}_1&\left(\mathsf{R}_t^{(1)}\leq e^{-x},\mathsf{R}_t^{(2)}\leq e^{x+C_3}\right)\\ &\leq C_0\cdot\mathbb{E}_1\Big[e^{-\lambda_1 x-h_{\lambda_1}\Gamma_t^p-g_{\lambda_1}\Gamma_t^q+\lambda_2(x+C_3)}\\ &\qquad e^{-f_{\lambda_2}(\|\mathcal{P}_t\|-\Gamma_t^p)+g_{\lambda_2}(\|\mathcal{Q}_t\|-\Gamma_t^q)}\Big].\end{aligned} \quad (64)$$

Using the facts that $\|\mathcal{Q}_t\| = \rho \cdot \|\mathcal{P}_t\|$ and $\Gamma_t^q = \rho \cdot \Gamma_t^p$, we obtain,

$$\mathbb{P}_1\left(\mathsf{R}_t^{(1)} \le e^{-x}, \mathsf{R}_t^{(2)} \le e^{x+C_3}\right) \le C_0 \cdot \mathbb{E}_1\left[e^{x(\lambda_2-\lambda_1)+\lambda_2 C_3-(f_{\lambda_2}-\rho g_{\lambda_2})\|\mathcal{P}_t\|} e^{(f_{\lambda_2}-\rho g_{\lambda_2}-h_{\lambda_1}-\rho g_{\lambda_1})\Gamma_t^p}\right]. \quad (65)$$

By symmetry, we can get the above upper bound with $(\lambda_2-\lambda_1)$ replaced by $(\lambda_1-\lambda_2)$. Indeed, to show this we replace (62) with

$$\mathbb{P}_1\left(\mathsf{R}_t^{(1)} \le e^{-x}, \mathsf{R}_t^{(2)} \le e^{x+C_3}\right) = \mathbb{P}_1\left[\left(\mathsf{R}_t^{(1)}\right)^{-(1-\lambda_1)} \ge e^{(1-\lambda_1)x}, \left(\mathsf{R}_t^{(2)}\right)^{-(1-\lambda_2)} \ge e^{-(1-\lambda_2)(x+C_3)}\right], \quad (66)$$

and follow the (63)–(64). Therefore, we may write

$$\mathbb{P}_1\left(\mathsf{R}_t^{(1)} \le e^{-x}, \mathsf{R}_t^{(2)} \le e^{x+C_3}\right) \le C_0 \cdot \mathbb{E}_1\left[e^{-x|\lambda_2-\lambda_1|+\lambda_2 C_3-(f_{\lambda_2}-\rho g_{\lambda_2})\|\mathcal{P}_t\|} e^{(f_{\lambda_2}-\rho g_{\lambda_2}-h_{\lambda_1}-\rho g_{\lambda_1})\Gamma_t^p}\right]. \quad (67)$$

We can now optimize our choices of λ_1 and λ_2 to minimize the above upper bound. We take $\lambda_1+\lambda_2=1$. For such a pair it is easy to check that,

$$f_{\lambda_2}-\rho g_{\lambda_2}-h_{\lambda_1}-\rho g_{\lambda_1} = \frac{\alpha-\beta}{\alpha+\beta}(1-\rho). \quad (68)$$

We next consider the case where $\rho \ge 1$, for which the r.h.s. of (68) is negative, and so,

$$\mathbb{P}_1\left(\mathsf{R}_t^{(1)} \le e^{-x}, \mathsf{R}_t^{(2)} \le e^{x+C_3}\right) \le C_0 \cdot \mathbb{E}_1\left[e^{-x|2\lambda_1-1|+\lambda_2 C_3-(f_{1-\lambda_1}-\rho g_{1-\lambda_1})\|\mathcal{P}_t\|} e^{\frac{\alpha-\beta}{\alpha+\beta}(1-\rho)\Gamma_t^p}\right] \quad (69)$$

$$\le C_0 \cdot e^{-x|2\lambda_1-1|+(1-\lambda_1)C_3-(f_{1-\lambda_1}-\rho g_{1-\lambda_1})\|\mathcal{P}_t\|}. \quad (70)$$

In the interval $\lambda_1 \in [0,1]$, it can be checked that $f_{1-\lambda_1}-\rho g_{1-\lambda_1}$ is maximized at

$$\lambda_1^\star = \min\left(1, \frac{\log\left(\rho \cdot \frac{\alpha/\beta-1}{\log \alpha/\beta}\right)}{\log \alpha/\beta}\right). \quad (71)$$

We mention here that $\lambda_1^\star$ satisfies the following equality

$$\left(\frac{\alpha}{\beta}\right)^{\lambda_1^\star} = \frac{(\alpha-\beta)\rho}{\beta \log \frac{\alpha}{\beta}}, \quad (72)$$

which proves to be useful. It can be shown that $\lambda_1^\star \ge 1/2$. Thus, whenever $\lambda_1^\star < 1$, which happens to be the case exactly when $\rho \le \rho_1$, we have

$$\mathbb{P}_1\left(\mathsf{R}_t^{(1)} \le e^{-x}, \mathsf{R}_t^{(2)} \le e^{x+C_3}\right) \le C_0 \cdot e^{-x|2\lambda_1^\star-1|+(1-\lambda_1^\star)C_3-(f_{1-\lambda_1^\star}-\rho g_{1-\lambda_1^\star})\|\mathcal{P}_t\|}. \quad (73)$$

It can be checked that

$$f_{1-\lambda_1^\star}-\rho g_{1-\lambda_1^\star} = \frac{\alpha \log \frac{\alpha}{\beta} - \rho(\alpha-\beta)}{(\alpha+\beta)\log\frac{\alpha}{\beta}} - \rho\frac{\alpha-\beta}{\alpha+\beta}(1-\lambda_1^\star) \quad (74)$$

$$= \frac{\alpha \log \frac{\alpha}{\beta} - \rho(\alpha-\beta)}{(\alpha+\beta)\log\frac{\alpha}{\beta}} - \frac{\rho(\alpha-\beta)}{\alpha+\beta}\left[1 - \frac{\log\left[\frac{\rho(\alpha/\beta-1)}{\log\frac{\alpha}{\beta}}\right]}{\log\frac{\alpha}{\beta}}\right] \quad (75)$$

$$= \frac{\frac{\alpha}{\beta}\log\frac{\alpha}{\beta} - \rho(\frac{\alpha}{\beta}-1)\left[1+\log\frac{\frac{\alpha}{\beta}\log\frac{\alpha}{\beta}}{\rho(\frac{\alpha}{\beta}-1)}\right]}{(1+\frac{\alpha}{\beta})\log\frac{\alpha}{\beta}} \quad (76)$$

$$= \delta(\alpha/\beta, \rho), \quad (77)$$

where $\delta(\alpha/\beta,\rho)$ is defined (18). Combining the above result with (60), we get

$$\mathbb{P}_1\left(\mathsf{R}_t \le \frac{\alpha}{\beta}\right) \le C_0' e^{(1-\lambda_1^\star)C_3-\delta(\alpha/\beta,\rho)\|\mathcal{P}_t\|} + \frac{1}{t^2}, \quad (78)$$

where we have used the fact that $\sum_{x=0}^{C_1 \log t} e^{-x|2\lambda_1^\star-1|}$ is finite, and absorbed its value in the constant C_0'. Then, substituting the above result in (43) and then in (42), we obtain

$$\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q}) \le \left[C_0' e^{(1-\lambda_1^\star)C_3-\delta(\gamma,\rho)\|\mathcal{P}_t\|} + \frac{1}{t^2}\right]\cdot(1-p_t) + \frac{\beta}{\alpha+\beta}\cdot p_t. \quad (79)$$

Therefore, taking $p_t = \frac{(1+\gamma)\kappa(\gamma)}{t} \wedge 1 = p_t^\star$, we obtain that for $1 \le \rho \le \rho_1$,

$$\mathsf{E}(\mathcal{P}^\star,\mathcal{Q}) = \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star,\mathcal{Q})}{\log t} \quad (80)$$

$$\ge \delta(\alpha/\beta,\rho)\cdot\liminf_{t\to\infty}\frac{\|\mathcal{P}_t\|}{\log t} \quad (81)$$

$$= \delta(\gamma,\rho)\left[(1+\gamma)\kappa(\gamma)\right], \quad (82)$$

as claimed. Next, we consider the case where $\rho < 1$. In this case, the the r.h.s. of (68) is positive, and so,

$$\begin{aligned}\mathbb{P}_1\left(\mathsf{R}_t^{(1)} \leq e^{-x}, \mathsf{R}_t^{(2)} \leq e^{x+C_3}\right) \\ \leq C_0 \cdot \mathbb{E}_1\left[e^{-x|2\lambda_1-1|+\lambda_2 C_3-(f_{1-\lambda_1}-\rho g_{1-\lambda_1})\|\mathcal{P}_t\|}\right. \\ \left. e^{\frac{\alpha-\beta}{\alpha+\beta}(1-\rho)\Gamma_t^p}\right] \qquad (83) \\ \leq C_0 \cdot e^{-x|2\lambda_1-1|+(1-\lambda_1)C_3-(f_{1-\lambda_1}-\rho g_{1-\lambda_1})\|\mathcal{P}_t\|} \\ \cdot e^{\frac{\alpha-\beta}{\alpha+\beta}(1-\rho)\|\mathcal{P}_t\|}. \qquad (84)\end{aligned}$$

Again, $f_{1-\lambda_1} - \rho g_{1-\lambda_1}$ is maximized at

$$\lambda_1^\star = \frac{\log\left(\rho \cdot \frac{\alpha/\beta-1}{\log \alpha/\beta}\right)}{\log \alpha/\beta}, \qquad (85)$$

and note that for $\rho < 1$ it is always the case that $\lambda_1^\star < 1$. Also, for $\rho_0 < \rho \leq 1$, we have that $\lambda_1^\star \in [0,1]$. Hence, we may write

$$\begin{aligned}\mathbb{P}_1\left(\mathsf{R}_t^{(1)} \leq e^{-x}, \mathsf{R}_t^{(2)} \leq e^{x+C_3}\right) \\ \leq C_0 e^{-x|2\lambda_1^\star-1|+(1-\lambda_1^\star)C_3-\left[\delta(\alpha/\beta,\rho)-\frac{\alpha-\beta}{\alpha+\beta}(1-\rho)\right]\|\mathcal{P}_t\|}. \qquad (86)\end{aligned}$$

Combining the above result with (60), we get

$$\begin{aligned}\mathbb{P}_1\left(\mathsf{R}_t \leq \frac{\alpha}{\beta}\right) \leq C_0' e^{(1-\lambda_1^\star)C_3-\left[\delta(\gamma,\rho)-\frac{\alpha-\beta}{\alpha+\beta}(1-\rho)\right]\|\mathcal{P}_t\|} \\ + \frac{1}{t^2}. \qquad (87)\end{aligned}$$

Then, substituting the above result in (43) and then in (42), we obtain

$$\begin{aligned}\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q}) \leq \frac{\beta}{\alpha+\beta}\cdot p_t + \\ \left[C_0' e^{(1-\lambda_1^\star)C_3-\left[\delta(\gamma,\rho)-\frac{\alpha-\beta}{\alpha+\beta}(1-\rho)\right]\|\mathcal{P}_t\|} + \frac{1}{t^2}\right]\cdot(1-p_t). \qquad (88)\end{aligned}$$

Therefore, taking $p_t = \frac{\alpha+\beta}{\beta}\frac{\kappa(\alpha,\beta)}{t} \wedge 1 = p_t^\star$, we obtain that for $\rho_0 \leq \rho \leq 1$,

$$\begin{aligned}\mathsf{E}(\mathcal{P}^\star,\mathcal{Q}) &= \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star,\mathcal{Q})}{\log t} \\ &\geq \left[\delta(\gamma,\rho) - (1-\rho)\frac{\gamma-1}{\gamma+1}\right]\cdot \liminf_{t\to\infty}\frac{\|\mathcal{P}_t^\star\|}{\log t} \\ &= \left[\delta(\gamma,\rho) - (1-\rho)\frac{\gamma-1}{\gamma+1}\right](1+\gamma)\kappa(\gamma), \qquad (89)\end{aligned}$$

as claimed.

B.2 Upper Bound: $\rho \geq \rho_1$ and $\|\mathcal{Q}_t\| \gg \log t$

We prove that for $\rho \geq \rho_1$ we have $\mathsf{E}(\mathcal{P}^\star,\mathcal{Q}) = 0$. We show that this is correct also when $\liminf_{t\to\infty}\frac{\|\mathcal{Q}_t\|}{\log t} = \infty$ as stated in Theorem 2. To this end, first note that from (8), we have

$$\begin{aligned}\mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq\theta\right) &= \mathbb{P}_1\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1},X_t)\neq 1\right) \\ &\geq \mathbb{P}_1\left(\mathsf{R}_t' < 1\right) \\ &\geq \mathbb{P}_1\left(\mathsf{R}_{t-1} < \frac{\beta}{\alpha}\right), \qquad (90)\end{aligned}$$

and so it is suffice to lower bound the r.h.s. of (90). It is clear that

$$\mathbb{P}_1\left(\mathsf{R}_t < \frac{\beta}{\alpha}\right) \geq \mathbb{P}_1\left(\mathsf{R}_i < \frac{\beta}{\alpha},\ \forall i\in[t]\right). \qquad (91)$$

Accordingly, in order to obtain a lower bound on (91) we define three events that together imply that $\mathsf{R}_t < \beta/\alpha$. We note that the derivations bellow follow [Peres et al., 2018, Sec. 2.2], with modifications which handle the mismatch aspect of our model. We need a few definitions. Let $\tau(s) \triangleq \min\{t \geq 1 : \|\mathcal{Q}_t\| \geq s\}$, and $t_0 \triangleq \tau(2\frac{\alpha-\beta}{\alpha+\beta}\log\frac{\alpha}{\beta}+2)$. Define

$$\mathcal{E}_0 \triangleq \left\{\mathsf{R}_{t_0} < (\beta/\alpha)^4\right\}. \qquad (92)$$

The above initial event takes the mismatched likelihood ratio below β/α, and the events we define below ensure that it always stays below this bar. Let $\mathsf{J}_t \triangleq \log \mathsf{R}_t$, and define the stopping time $T \triangleq \min\{s \geq t_0 : J_s \notin [-\log t, 2\log(\beta/\alpha)]\}$. We define the events

$$\mathcal{E}_1 \triangleq \{\mathsf{J}_T \leq -\log t\}, \qquad (93)$$

and

$$\mathcal{E}_2 \triangleq \left\{\min_{s\in[t]} \mathsf{J}_s \geq -\log^{3/4} t\right\}. \qquad (94)$$

We observe that $\mathcal{E}_0\cap\mathcal{E}_1\cap\mathcal{E}_2$ imply together that $\mathsf{J}_s \in \left[-\log^{3/4} t, 2\log\frac{\beta}{\alpha}\right]$, for all $s \in [t_0, t]$, which in turn implies that $\mathsf{R}_t < \beta/\alpha$. Thus,

$$\mathbb{P}_1\left(\mathsf{R}_t \leq \frac{\beta}{\alpha}\right) \geq \mathbb{P}_1\left(\mathcal{E}_0\cap\mathcal{E}_1\cap\mathcal{E}_2\right). \qquad (95)$$

We next lower bound the probability of the event $\mathcal{E}_0$ which is easier to handle. Note that according to our setting the first two individuals follow their private signal, and hence if $X_1 = X_2 = 2$, we have $Z_1 = Z_2 = 2$. This in turn implies that $\mathsf{R}_2 = (\beta/\alpha)^2$. Now, if $X_i = 2$ for all $i \in \{3,4,\ldots,t_0\}$, then it is clear that $Z_i = 2$, for all $i \in \{3,4,\ldots,t_0\}$ as well. Accordingly, using (25)–(26), this implies that the mismatched likelihood ratio at time t_0 is given by

$$\mathsf{R}_{t_0} = \left(\frac{\beta}{\alpha}\right)^2 \prod_{i=3}^{t_0} \frac{1-\frac{\alpha}{\alpha+\beta}q_i}{1-\frac{\beta}{\alpha+\beta}q_i} \tag{96}$$

$$\leq \left(\frac{\beta}{\alpha}\right)^2 \exp\left(-\frac{\alpha-\beta}{\alpha+\beta}\sum_{i=3}^{t_0} q_i\right). \tag{97}$$

However, by the definition of t_0, we know that

$$\sum_{i=3}^{t_0} q_i \geq \|\mathcal{Q}_{t_0}\| - 2 \geq 2\frac{\alpha+\beta}{\alpha-\beta}\log\frac{\alpha}{\beta}, \tag{98}$$

which together with (97) implies that $\mathsf{R}_{t_0} \leq (\beta/\alpha)^4$. Thus,

$$\mathbb{P}_1(\mathcal{E}_0) \geq \mathbb{P}_1(X_i = 2\ \forall i \in [t_0]) = \left(\frac{\beta}{\alpha+\beta}\right)^{t_0}. \tag{99}$$

Therefore, because t_0 is a constant it is suffice to lower bound the probability $\mathbb{P}_1(\mathcal{E}_1 \cap \mathcal{E}_2|\mathcal{E}_0)$. Given $\mathcal{E}_0$, the log-likelihood ratio J_t performs a random walk from time t_0 until the stopping time T. Specifically, using (25)–(26), and (46), for $s \geq t_0$, we may write

$$\mathsf{J}_{s\wedge T} = \mathsf{J}_{t_0} + \sum_{i=t_0+1}^{s\wedge T} \xi_i \tag{100}$$

where $\{\xi_i\}$ are statistically independent random variables such that

$$\mathbb{P}_1\left(\xi_i = \log\frac{\alpha}{\beta}\right) = \frac{\alpha}{\alpha+\beta}p_i, \tag{101}$$

and

$$\mathbb{P}_1\left(\xi_i = \log\frac{1-\frac{\alpha}{\alpha+\beta}q_i}{1-\frac{\beta}{\alpha+\beta}q_i}\right) = 1 - \frac{\alpha}{\alpha+\beta}p_i. \tag{102}$$

Now, note that

$$\mathbb{E}_1[\xi_i] = \frac{\alpha}{\alpha+\beta}\log\frac{\alpha}{\beta}p_i - \frac{\alpha-\beta}{\alpha+\beta}q_i + \Theta(q_i^2+p_i^2)$$
$$= \frac{\rho+\gamma(\log\gamma-\rho)}{1+\gamma}p_i + \Theta(p_i^2) \tag{103}$$

where $\gamma = \alpha/\beta$. The important observation here is that $\rho \geq \rho_1$ is equivalent $\rho+\gamma(\log\gamma-\rho) \leq 0$, which implies that the log-likelihood ratio has a downward (non-positive) drift. More precisely, it can be seen that the expectation can be written as $\mathbb{E}_1[\xi_i] = -\eta \cdot p_i + \Theta(p_i^2)$, for some $\eta > 0$. Since p_i is decaying with i, it is clear that there exists a *finite* index $i_0 \in \mathbb{N}$, such that $\mathbb{E}_1[\xi_i] \leq 0$, for all $i \geq i_0$. Accordingly, letting $\bar{t}_0 \triangleq t_0 \vee i_0$, we obtain that under $\mathbb{P}_1$, the random walk $\{\mathsf{J}_{s\wedge T}\}_{s\geq \bar{t}_0}$ is a supermartingale. For simplicity of notation, for the rest of the proof we use t_0 in place of $\bar{t}_0$. Therefore, by the optional stopping theorem we have

$$\mathbb{E}_1[\mathsf{J}_T|\mathcal{E}_0] \leq \mathbb{E}_1[\mathsf{J}_{t_0}] \leq 4\log\frac{\beta}{\alpha}. \tag{104}$$

On the other hand, by the definition of T, it is either the case that $\mathsf{J}_T > 2\log\frac{\beta}{\alpha}$, in which case $\mathsf{J}_T \in (2\log\frac{\beta}{\alpha}, \log\frac{\beta}{\alpha}]$, or $\mathsf{J}_T < -\log t$, and then $\mathsf{J}_T \in [-\log t - \log\frac{\beta}{\alpha}, -\log t)$. Thus, we can write

$$\mathbb{E}_1[\mathsf{J}_T|\mathcal{E}_0] = \mathbb{E}_1\left[\mathsf{J}_T\mathbb{1}\left[\mathsf{J}_T > 2\log\frac{\beta}{\alpha}\right]|\mathcal{E}_0\right]$$
$$+ \mathbb{E}_1\left[\mathsf{J}_T\mathbb{1}\left[\mathsf{J}_T < -\log t\right]|\mathcal{E}_0\right] \tag{105}$$

$$\geq [1-\mathbb{P}_1(\mathcal{E}_1|\mathcal{E}_0)]\, 2\log\frac{\beta}{\alpha}$$
$$- \mathbb{P}_1(\mathcal{E}_1|\mathcal{E}_0)\cdot\left(\log t + \log\frac{\beta}{\alpha}\right) \tag{106}$$

$$\geq 2\log\frac{\beta}{\alpha} - \mathbb{P}_1(\mathcal{E}_1|\mathcal{E}_0)\cdot\log t, \tag{107}$$

which together with (104) implies that

$$\mathbb{P}_1(\mathcal{E}_1|\mathcal{E}_0) \geq \frac{2\log\gamma}{\log t}. \tag{108}$$

Finally, using classical results on the tails of supermartingales (see, e.g., [Freedman, 1975, Fan et al., 2015]), we have

$$\mathbb{P}\left(\min_{s\in[t]} \mathsf{J}_s < -\log^{3/4} t \,\middle|\, \mathcal{E}_0\right) \leq e^{-c\sqrt{\log t}}, \tag{109}$$

and thus

$$\mathbb{P}_1(\mathcal{E}_1 \cap \mathcal{E}_2|\mathcal{E}_0) \geq \frac{\log\gamma}{\log t}, \tag{110}$$

for t large enough. Combining (27), (90), (95), (99), and (110), we obtain

$$\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) = \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star, \mathcal{Q})}{\log t} \tag{111}$$

$$\leq \liminf_{t\to\infty} -\frac{\log \mathbb{P}_1(\mathcal{E}_1 \cap \mathcal{E}_2|\mathcal{E}_0)}{\log t} = 0, \tag{112}$$

which concludes the proof. Finally, using the above arguments we prove that when $\bar{\mathcal{Q}}$ satisfies $\liminf_{t\to\infty} \frac{\|\bar{\mathcal{Q}}_t\|}{\log t} = \infty$, then $\mathsf{E}(\mathcal{P}^\star, \bar{\mathcal{Q}}) = 0$, as stated in Theorem 2. Specifically, let $\mathcal{Q}$ be any sequence of assumed revealing probabilities such that $q_t = \rho \cdot p_t^\star$, with $\rho > \rho_1$. Let $\mathsf{R}_t^{\bar{\mathcal{Q}}}$ and $\mathsf{R}_t^{\mathcal{Q}}$ designate the likelihoods corresponding to the revealing probabilities $\bar{\mathcal{Q}}$ and $\mathcal{Q}$, respectively. Then, from (91) it is clear that

$$\mathbb{P}\left(\mathsf{MAP}_{\bar{\mathcal{Q}}}(Z_1^t, X_{t+1}) \neq \theta\right) \geq \mathbb{P}_1\left(\mathsf{R}_t^{\bar{\mathcal{Q}}} < \frac{\beta}{\alpha}\right) \tag{113}$$

$$\geq \mathbb{P}_1\left(\mathsf{R}_i^{\bar{\mathcal{Q}}} < \frac{\beta}{\alpha},\ \forall i \in [t]\right). \tag{114}$$

Now, note that for large enough t it must be the case that $\bar{q}_t > q_t$. The log-likelihood ratio $\mathsf{J}_t^{\bar{\mathcal{Q}}}$ performs a random walk with probabilities given in (101)–(102), with q_t replaced by $\bar{q}_t$. Accordingly, it is clear that the random variables $\{\xi_i\}$ can take only smaller values under $\bar{\mathcal{Q}}$ compared to $\mathcal{Q}$. This in turn implies that $\mathsf{R}^{\bar{\mathcal{Q}}} \leq \mathsf{R}^{\mathcal{Q}}$, and thus,

$$\begin{aligned}\mathbb{P}\left(\mathsf{MAP}_{\bar{\mathcal{Q}}}(Z_1^t, X_{t+1}) \neq \theta\right) &\geq \mathbb{P}_1\left(\mathsf{R}_i^{\bar{\mathcal{Q}}} < \frac{\beta}{\alpha},\ \forall i \in [t]\right)\\ &\geq \mathbb{P}_1\left(\mathsf{R}_i^{\mathcal{Q}} < \frac{\beta}{\alpha},\ \forall i \in [t]\right)\\ &\geq \mathbb{P}_1(\mathcal{E}_0 \cap \mathcal{E}_1 \cap \mathcal{E}_2), \quad (115)\end{aligned}$$

which is the same lower bound we started with for $\mathcal{Q}$.

B.3 Upper Bound: $1 \leq \rho \leq \rho_1$

We consider the case where $1 \leq \rho \leq \rho_1$, and continue from (103). Indeed, in this regime, the expectation in (103) is non-negative and thus the log-likelihood ratio J_s has an upward drift. We remove this drift by defining a new measure $\tilde{\mathbb{P}}_1$, such that for $i > t_0$,

$$\tilde{\mathbb{P}}_1\left(\xi_i = \log\frac{\alpha}{\beta}\right) = \nu_i, \quad (116)$$

$$\tilde{\mathbb{P}}_1\left(\xi_i = \log\frac{1-\frac{\alpha}{\alpha+\beta}q_i}{1-\frac{\beta}{\alpha+\beta}q_i}\right) = 1-\nu_i, \quad (117)$$

where

$$\nu_i \triangleq \frac{\log\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}}{\log\left(\frac{\alpha}{\beta}\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}\right)}. \quad (118)$$

Now, under $\tilde{\mathbb{P}}_1$, the random walk $\{\mathsf{J}_{s\wedge T}\}_{s\geq t_0}$ is a martingale, and thus using the same steps we used in (104)–(110), we obtain that

$$\tilde{\mathbb{P}}_1\left(\mathcal{E}_1 \cap \mathcal{E}_2 | \mathcal{E}_0\right) \geq \frac{\log\gamma}{\log t}, \quad (119)$$

for t large enough. Next, performing a change of measure we may write

$$\mathbb{P}_1\left(\mathcal{E}_1 \cap \mathcal{E}_2 | \mathcal{E}_0\right) = \tilde{\mathbb{E}}_1\left[\frac{\mathrm{d}\mathbb{P}_1(\cdot|\mathcal{E}_0)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot|\mathcal{E}_0)}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right]\middle|\, \mathcal{E}_0\right], \quad (120)$$

so we need to understand how the Radon-Nikodym derivative of $\mathbb{P}_1(\cdot|\mathcal{E}_0)$ w.r.t. $\tilde{\mathbb{P}}_1(\cdot|\mathcal{E}_0)$ behaves. Note that

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}_1(\cdot|\mathcal{E}_0)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot|\mathcal{E}_0)} = \prod_{i=t_0+1}^{t}\Bigg\{&\frac{\frac{\alpha}{\alpha+\beta}p_i^\star}{\nu_i}\mathbb{1}\left[\xi_i = \log\frac{\alpha}{\beta}\right]\\ &+\frac{1-\frac{\alpha}{\alpha+\beta}p_i^\star}{1-\nu_i}\mathbb{1}\left[\xi_i = \log\frac{1-\frac{\alpha}{\alpha+\beta}q_i}{1-\frac{\beta}{\alpha+\beta}q_i}\right]\Bigg\}. \quad (121)\end{aligned}$$

We claim that each factor in the product can be lower bounded for some $C = C(\alpha,\beta)$ as follows

$$\begin{aligned}&\frac{\frac{\alpha}{\alpha+\beta}p_i^\star}{\nu_i}\mathbb{1}\left[\xi_i = \log\frac{\alpha}{\beta}\right]\\ &\quad+\frac{1-\frac{\alpha}{\alpha+\beta}p_i^\star}{1-\nu_i}\mathbb{1}\left[\xi_i = \log\frac{1-\frac{\alpha}{\alpha+\beta}q_i}{1-\frac{\beta}{\alpha+\beta}q_i}\right]\\ &\quad\geq e^{(1-\lambda^\star)\xi_i}\cdot K_i(\xi_i) \quad (122)\end{aligned}$$

where $\lambda^\star = \lambda_1^\star$ is defined in (71), and

$$\begin{aligned}K_i(\xi_i) \triangleq\ & e^{-\delta(\gamma,\rho)p_i^\star - C(p_i^\star)^2}\cdot\mathbb{1}\left[\xi_i = \log\frac{1-\frac{\alpha}{\alpha+\beta}q_i}{1-\frac{\beta}{\alpha+\beta}q_i}\right]\\ &+e^{-\rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)p_i^\star - C(p_i^\star)^2}\cdot\mathbb{1}\left[\xi_i = \log\frac{\alpha}{\beta}\right]. \quad (123)\end{aligned}$$

Indeed, this inequality can be checked for both potential values of ξ by expanding the expressions in $p_i^\star$. Then, multiplying (122) over all $i \in \{t_0+1,\ldots,t\}$, using the fact that $\mathsf{J}_t = \mathsf{J}_{t_0} + \sum_{i=t_0+1}^t \xi_i$ on the event $\mathcal{E}_1 \cap \mathcal{E}_2$, we obtain that

$$\begin{aligned}&\frac{\mathrm{d}\mathbb{P}_1(\cdot|\mathcal{E}_0)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot|\mathcal{E}_0)}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right] \geq e^{(1-\lambda^\star)(\mathsf{J}_t-\mathsf{J}_{t_0}) - C\sum_{i=1}^t (p_i^\star)^2}\\ &\quad\cdot e^{-\delta(\gamma,\rho)\sum_{i\in\mathcal{V}^c}p_i^\star - \rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}}p_i^\star}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right]\\ &\geq e^{(1-\lambda^\star)\mathsf{J}_t - C'}\\ &\quad\cdot e^{-\delta(\gamma,\rho)\|\mathcal{P}_t^\star\| - \rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}}p_i^\star}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right], \quad (124)\end{aligned}$$

where $\mathcal{V} \triangleq \{i \geq t_0 : \xi_i = \log(\alpha/\beta)\}$, $C' \triangleq C\sum_{i=1}^t (p_i^\star)^2$ is finite, and the second inequality follows because conditioned on $\mathcal{E}_0$ we know that $\mathsf{J}_{t_0} < 0$. Also, recall that on the event $\mathcal{E}_1 \cap \mathcal{E}_2$ we have that $\mathsf{J}_t \geq -\log^{3/4} t$, and thus

$$\begin{aligned}&\frac{\mathrm{d}\mathbb{P}_1(\cdot|\mathcal{E}_0)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot|\mathcal{E}_0)}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right] \geq e^{-(1-\lambda^\star)\log^{3/4} t - C' - \delta(\gamma,\rho)\|\mathcal{P}_t^\star\|}\\ &\quad\cdot e^{-\rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}}p_i^\star}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right]. \quad (125)\end{aligned}$$

We next show that with high probability $|\mathcal{V}|$ is at most logarithmic in t and thus $\sum_{i\in\mathcal{V}} p_i^\star$ is negligible compared to other contributions in the exponent of the r.h.s. of (125). Let $Z \triangleq \sum_{i=t_0+1}^t \mathbb{1}\left[\xi_i = \log\frac{\alpha}{\beta}\right]$. Then, we already saw that under $\tilde{\mathbb{P}}_1$ the random variables $\{\xi\}_{i>t_0}$ are statistically independent. Specifically, Z follows a Poisson-Binomial distribution with success probabilities ν_i given in (118). Using Chernoff's inequality, for a Poisson-Binomial random variable Z with mean μ, and any $s > \mu$, it can be shown that

$$\mathbb{P}\left[\mathsf{Z} \geq s\right] \leq \exp\left(s - \mu - s\log\frac{s}{\mu}\right). \tag{126}$$

Accordingly, in our case it is clear that $\mu = \sum_{i=t_0+1}^{t} \nu_i = C_1(1+o(1)) \cdot \log t$, as $t \to \infty$, for some $C_1(\alpha,\beta)$, due to the fact that $\nu_i \propto q_i = \Theta(t^{-1})$. Taking $s = \ell \cdot \mu$, such that $(\ell - 1 - \ell\log\ell) \leq -\frac{2}{C_1}$, we obtain from (126) that $\tilde{\mathbb{P}}_1\left[\mathsf{Z} \geq s\right] \leq t^{-2}$. Thus, with probability at least $1 - O(t^{-2})$ we have that $|\mathcal{V}| \leq C'' \log t$, for some constant C''. This in turn implies that with the same probability

$$\sum_{i\in\mathcal{V}} p_i^\star \leq C_2(1+o(1)) \cdot \log(\log t) = o(\log t). \tag{127}$$

Therefore, combining (120), (125), and (127), we obtain

$$\begin{aligned}
\mathbb{P}_1\left(\mathcal{E}_1 \cap \mathcal{E}_2 | \mathcal{E}_0\right) &= \tilde{\mathbb{E}}_1\left[\frac{\mathrm{d}\mathbb{P}_1(\cdot|\mathcal{E}_0)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot|\mathcal{E}_0)}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right]\middle| \mathcal{E}_0\right] \qquad (128)\\
&\geq \tilde{\mathbb{E}}_1\left[e^{-(1-\lambda^\star)\log^{3/4} t - C' - \delta(\gamma,\rho)\|\mathcal{P}_t^\star\|}\right.\\
&\qquad \left. \cdot e^{-\rho\left(\frac{1}{2} - \frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}} p_i^\star}\mathbb{1}\left[\mathcal{E}_1 \cap \mathcal{E}_2\right]\middle| \mathcal{E}_0\right]\\
&\geq \left[1 - O(t^{-2})\right] e^{-o(\log t) - \delta(\gamma,\rho)\|\mathcal{P}_t^\star\|}\tilde{\mathbb{P}}_1(\mathcal{E}_1 \cap \mathcal{E}_2|\mathcal{E}_0)\\
&\geq \frac{\left[1 - O(t^{-2})\right]\log\frac{\alpha}{\beta}}{\log t} e^{-o(\log t) - \delta(\gamma,\rho)\|\mathcal{P}_t^\star\|}. \qquad (129)
\end{aligned}$$

Combining (27), (90), (95), (99), and (129), we obtain

$$\begin{aligned}
\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) &= \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star, \mathcal{Q})}{\log t} \qquad (130)\\
&\leq \liminf_{t\to\infty} -\frac{\log \mathbb{P}_1\left(\mathcal{E}_1 \cap \mathcal{E}_2 | \mathcal{E}_0\right)}{\log t} \qquad (131)\\
&\leq \delta(\gamma,\rho) \liminf_{t\to\infty} \frac{\|\mathcal{P}_t^\star\|}{\log t} \qquad (132)\\
&= \delta(\gamma,\rho)[(1+\gamma)\kappa(\gamma)], \qquad (133)
\end{aligned}$$

as claimed.

B.4 Upper Bound: $\rho \leq \rho_0$

For $\rho \leq \rho_0$ we use the fact that when the log-likelihood ratio is above $\log\alpha/\beta$, it has a downward drift. This implies that above $\log\alpha/\beta$, the walk cannot go beyond a certain value. Recall (90). As before, in order to obtain a lower bound on (90) we define an event that implies that $\mathsf{R}_t < \beta/\alpha$. Now, when the log-likelihood ratio $\mathsf{J}_t \triangleq \log \mathsf{R}_t$ is above the line $\log\alpha/\beta$, using (25)–(26), and (46), we may write

$$\mathsf{J}_s = \sum_{i=1}^{s} \xi_i, \tag{134}$$

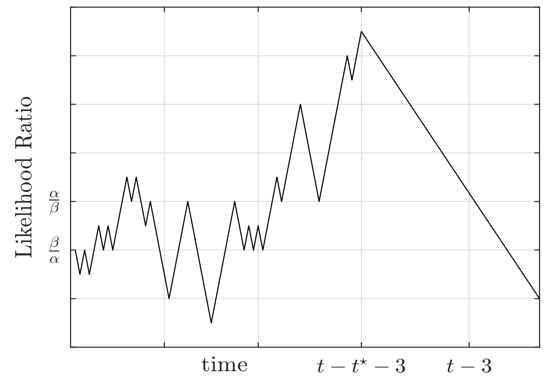

Figure 3: Illustration of a sample path of the random walk constructed in the lower bound.

for $s \geq 0$, where ξ_i's are statistically independent random variables, and

$$\mathbb{P}_1\left(\xi_i = \log\frac{\beta}{\alpha}\right) = \frac{\beta}{\alpha+\beta}p_i, \tag{135}$$

and

$$\mathbb{P}_1\left(\xi_i = \log\frac{1 - \frac{\beta}{\alpha+\beta}q_i}{1 - \frac{\alpha}{\alpha+\beta}q_i}\right) = 1 - \frac{\beta}{\alpha+\beta}p_i. \tag{136}$$

Note that

$$\mathbb{E}_1\left[\xi_i\right] = \frac{(\gamma-1)\rho - \log\gamma}{1+\gamma}p_i + \Theta(p_i^2). \tag{137}$$

Thus, we see that for $\rho \leq \rho_0$, we have $\mathbb{E}_1\left[\xi_i\right] \leq 0$. We next show that with high probability $\max_{1\leq i\leq t} \mathsf{J}_i < \tau_0$, namely, the maximal value that the log-likelihood ratio can achieve is bounded by a certain constant τ_0. Accordingly, using the same arguments as in the proof of Theorem 2 this implies that only a finite number of timestamps are needed in order to drive log-likelihood ratio bellow $\log\alpha/\beta$. Specifically, as in the proof of Theorem 2 it is suffice to assume that the last $t^\star + 3$ revealers are such that their private information is $X_i = 2$. Indeed, if for example, at time $\ell = t - (t^\star + 3)$ the log-likelihood ratio J_ℓ attained its maximal possible value τ_0 (or e^{τ_0} for R_ℓ). Then, after $t^\star$ timestamps, i.e., at time $\ell = t - 3$, the likelihood value is at most $(\beta/\alpha)^{t^\star} e^{\tau_0}$. Accordingly, if we set $t^\star = 1 \vee (\frac{\tau_0}{\log\alpha/\beta} - 1)$, then we get that the likelihood value is $(\beta/\alpha)^{t^\star} e^{\tau_0} \leq \alpha/\beta$, namely, bellow α/β. Thus, in the worst case, at time $\ell = t - 3$, the likelihood ratio value is in the interval $[\beta/\alpha, \alpha/\beta]$. In this interval, the MAP estimator outputs the private signal, namely, $Z_i = X_i = 2$, and accordingly, the likelihood ratio is multiplied by β/α. Therefore, the remaining 3 timestamps simply insure that at time t the likelihood ratio value is below β/α, as required. To wit, if the likelihood value at time $t-3$ is α/β, then at time t it value will be $(\beta/\alpha)^2 < \beta/\alpha$. Thus, by the above arguments, it is clear that we can lower bound the error probability as follows

$$\mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^t, X_{t+1}) \neq \theta\right) \geq \mathbb{P}_1\left(\mathsf{R}_t < \frac{\beta}{\alpha}\right) \tag{138}$$

$$\geq \left(\frac{\beta}{\alpha+\beta}\right)^{3+t^\star} \mathbb{P}_1\left[\max_{1\leq s\leq t} \mathsf{J}_s \leq \tau_0\right]. \tag{139}$$

We next show that there exists a finite value of τ_0 such that the probability term at the r.h.s. of (139) is lower bounded by 1/2. Thus, since $t^\star$ is finite, we obtain

$$\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) = \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star, \mathcal{Q})}{\log t} \tag{140}$$

$$\leq \liminf_{t\to\infty} -\frac{\log\left[\frac{1}{2}\left(\frac{\beta}{\alpha+\beta}\right)^{3+t^\star}\right]}{\log t} \tag{141}$$

$$= 0, \tag{142}$$

as claimed. It is only left to prove that the probability term at the r.h.s. of (139) is lower bounded by 1/2. To this end, for any $\lambda \geq 0$, we have

$$\begin{aligned}
\mathbb{E}_1\left[e^{\lambda \xi_i}\right] &= \frac{\beta}{\alpha+\beta} e^{\lambda \log\frac{\beta}{\alpha}} p_i \\
&\quad + e^{\lambda \log \frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}} \left(1 - \frac{\beta}{\alpha+\beta} p_i\right) \\
&= \frac{\beta}{\alpha+\beta}\left[1 + \left(e^{\lambda\log\frac{\beta}{\alpha}} - 1\right)\right] p_i \\
&\quad + \left[1 + \lambda\rho\frac{\alpha-\beta}{\alpha+\beta}p_i + \Theta(p_i^2)\right]\left(1 - \frac{\beta}{\alpha+\beta}p_i\right) \\
&= 1 + \left[\left(e^{\lambda\log\frac{\beta}{\alpha}} - 1\right)\frac{\beta}{\alpha+\beta} + \lambda\rho\frac{\alpha-\beta}{\alpha+\beta}\right] p_i + \Theta(p_i^2) \\
&= 1 + \left[\left(e^{-\lambda\log\gamma} - 1\right)\frac{1}{\gamma+1} + \lambda\rho\frac{\gamma-1}{\gamma+1}\right] p_i + \Theta(p_i^2).
\end{aligned} \tag{143}$$

Let us define the map,

$$\varphi : \lambda \mapsto \left(e^{-\lambda\log\gamma} - 1\right)\frac{1}{\gamma+1} + \lambda\rho\frac{\gamma-1}{\gamma+1}. \tag{144}$$

For $\lambda \ll 1$, we have $\varphi(\lambda) = \frac{(\rho-\rho_0)(\gamma-1)}{\gamma+1}\lambda + O(\lambda^2)$, and since $\rho - \rho_0 \leq 0$, we may conclude that $\varphi(\cdot)$ has a negative derivative at 0, hence its minimum, attained at $\lambda_0 > 0$, is strictly negative, namely, $\varphi(\lambda_0) < 0$. Accordingly, due to statistical independence we may write,

$$\mathbb{E}_1\left[e^{\lambda_0 \mathsf{J}_s}\right] = \mathsf{C}_s e^{\varphi(\lambda_0)\|\mathcal{P}_s^\star\|}, \tag{145}$$

for a certain converging/bounded sequence $\{\mathsf{C}_s\}_{s\geq 1}$. Next, define the random process

$$\mathsf{M}_s \triangleq \frac{\exp(\lambda_0 \mathsf{J}_s)}{\mathbb{E}_1\left[\exp(\lambda_0 \mathsf{J}_s)\right]}, \tag{146}$$

for $s \geq 1$. It is clear that $\{\mathsf{M}_s\}_{s\geq 1}$ is a positive martingale. Thus, using Doob's martingale maximal inequality, we have for $\tau > 0$,

$$\mathbb{P}_1\left[\max_{1\leq s\leq t} \mathsf{M}_s \geq \tau\right] \leq \frac{\mathbb{E}_1(\mathsf{M}_t)}{\tau} = \frac{1}{\tau}, \tag{147}$$

which is equivalent to

$$\mathbb{P}_1\left[\max_{1\leq s\leq t} \frac{e^{\lambda_0 \mathsf{J}_s}}{\mathsf{C}_s e^{\varphi(\lambda_0)\|\mathcal{P}_s^\star\|}} \geq \tau\right] \leq \frac{1}{\tau}. \tag{148}$$

In particular, using the fact that $\varphi(\lambda_0) < 0$, it is clear that the above implies that

$$\mathbb{P}_1\left[\max_{1\leq s\leq t} e^{\lambda_0 \mathsf{J}_s} \geq \tau \cdot \max_{1\leq s\leq t} \mathsf{C}_s\right] \leq \frac{1}{\tau}, \tag{149}$$

or,

$$\mathbb{P}_1\left[\max_{1\leq s\leq t} \mathsf{J}_s \geq \frac{\log\left[\tau \cdot \max_{1\leq s\leq t} \mathsf{C}_s\right]}{\lambda_0}\right] \leq \frac{1}{\tau}. \tag{150}$$

The above can be written also as follows

$$\mathbb{P}_1\left[\max_{1\leq s\leq t} \mathsf{J}_s \geq \tau\right] \leq e^{-\lambda_0 \tau} \cdot \max_{1\leq s\leq t} \mathsf{C}_s. \tag{151}$$

Therefore, taking $\tau > \tau_0 \triangleq \frac{2\log \max_{1\leq s\leq t} \mathsf{C}_s}{\lambda_0}$, we have $\mathbb{P}_1\left[\max_{1\leq s\leq t} \mathsf{J}_s \geq \tau\right] < 1/2$, as claimed.

B.5 Upper Bound: $\rho_0 \leq \rho \leq 1$

For this regime we use similar arguments as in the previous subsection. The main difference here is that the log-likelihood ratio random walk now has a positive drift, and thus is unbounded. However, we claim that this unbounded value is small compared to $\log t$, and thus, the number of timestamps $t^\star$ needed to bring the random walk bellow $\log\frac{\beta}{\alpha}$ is small compared to $\log t$, and more importantly will not affect the learning rate. Specifically, recall (139). Taking $\tau_0 = \log^{3/4} t$, we have,

$$\mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^t, X_{t+1}) \neq \theta\right) \geq \mathbb{P}_1\left(\mathsf{R}_t < \frac{\beta}{\alpha}\right) \tag{152}$$

$$\geq \left(\frac{\beta}{\alpha+\beta}\right)^{3+t^\star} \mathbb{P}_1\left[\max_{1\leq s\leq t} \mathsf{J}_s \leq \log^{3/4} t\right], \tag{153}$$

and since $t^\star = \Theta(\tau_0)$, we have

$$\mathsf{E}(\mathcal{P}^\star, \mathcal{Q}) = \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star, \mathcal{Q})}{\log t} \tag{154}$$

$$\leq \liminf_{t\to\infty} -\frac{\log \mathbb{P}_1\left[\max_{1\leq s\leq t} \mathsf{J}_s \leq \log^{3/4} t\right]}{\log t}. \tag{155}$$

We next upper bound the r.h.s. of the above inequality. To this end, note that above $\log\alpha/\beta$, the log-likelihood ratio process J_s forms a random walk as in (134)–(136), now with a positive drift since $\rho \geq \rho_0$. We remove this drift by defining a new measure $\tilde{\mathbb{P}}_1$, such that,

$$\tilde{\mathbb{P}}_1\left(\xi_i = \log\frac{\beta}{\alpha}\right) = \nu_i, \tag{156}$$

$$\tilde{\mathbb{P}}_1\left(\xi_i = \log\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}\right) = 1-\nu_i, \tag{157}$$

where

$$\nu_i \triangleq \frac{\log\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}}{\log\left(\frac{\alpha}{\beta}\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}\right)}. \tag{158}$$

Now, under $\tilde{\mathbb{P}}_1$, the random walk $\{\mathsf{J}_s\}_s$ is a martingale, and thus, using classical results on the tails of martingales (see, e.g., [Freedman, 1975, Fan et al., 2015])

$$\tilde{\mathbb{P}}_1\left(\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right) \geq 1-e^{-c\sqrt{\log t}}, \tag{159}$$

for t large enough. Next, performing a change of measure we may write

$$\begin{aligned}\mathbb{P}_1&\left(\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right)\\ &= \tilde{\mathbb{E}}_1\left[\frac{\mathrm{d}\mathbb{P}_1(\cdot)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot)}\mathbb{1}\left[\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right]\right],\end{aligned} \tag{160}$$

so we need to understand how the Radon-Nikodym derivative of $\mathbb{P}_1(\cdot)$ w.r.t. $\tilde{\mathbb{P}}_1(\cdot)$ behaves. Note that

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}_1(\cdot)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot)} = \prod_{i=1}^{t}&\left\{\frac{\frac{\beta}{\alpha+\beta}p_i^\star}{\nu_i}\mathbb{1}\left[\xi_i = \log\frac{\beta}{\alpha}\right]\right.\\ &\left.+\frac{1-\frac{\beta}{\alpha+\beta}p_i^\star}{1-\nu_i}\mathbb{1}\left[\xi_i = \log\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}\right]\right\}.\end{aligned} \tag{161}$$

We claim that each factor in the product can be lower bounded for some $C = C(\alpha,\beta)$ as follows

$$\begin{aligned}&\frac{\frac{\beta}{\alpha+\beta}p_i^\star}{\nu_i}\mathbb{1}\left[\xi_i = \log\frac{\beta}{\alpha}\right]\\ &+\frac{1-\frac{\beta}{\alpha+\beta}p_i^\star}{1-\nu_i}\mathbb{1}\left[\xi_i = \log\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}\right] \geq e^{\lambda^\star\xi_i}\tilde{K}_i(\xi_i),\end{aligned} \tag{162}$$

where $\lambda^\star = \lambda_1^\star$ is defined in (71), and

$$\begin{aligned}\tilde{K}_i(\xi_i) &\triangleq e^{-\left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]p_i^\star - C(p_i^\star)^2}\\ &\quad\cdot\mathbb{1}\left[\xi_i = \log\frac{1-\frac{\beta}{\alpha+\beta}q_i}{1-\frac{\alpha}{\alpha+\beta}q_i}\right]\\ &+e^{-\rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)p_i^\star - C(p_i^\star)^2}\cdot\mathbb{1}\left[\xi_i = \log\frac{\beta}{\alpha}\right].\end{aligned} \tag{163}$$

Indeed, this inequality can be checked for both potential values of ξ by expanding the expressions in $p_i^\star$. Then, multiplying (162) over all $i\in\{1,\ldots,t\}$, we obtain that

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}_1(\cdot)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot)} &\geq e^{\lambda^\star\mathsf{J}_t - C\sum_{i=1}^t(p_i^\star)^2}e^{-\left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]\sum_{i\in\mathcal{V}^c}p_i^\star}\\ &\quad\cdot e^{-\rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}}p_i^\star}\\ &\geq e^{\lambda^\star\mathsf{J}_t - C'}e^{-\left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]\|\mathcal{P}_t^\star\|}\\ &\quad\cdot e^{-\rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}}p_i^\star}\\ &\geq e^{\lambda^\star\log\frac{\alpha}{\beta} - C'}e^{-\left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]\|\mathcal{P}_t^\star\|}\\ &\quad\cdot e^{-\rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}}p_i^\star},\end{aligned} \tag{164}$$

where $\mathcal{V}\triangleq\{i\geq 1:\xi_i = \log(\beta/\alpha)\}$, $C'\triangleq C\sum_{i=1}^t(p_i^\star)^2$ is finite. As in (126)–(127), with probability at least $1-O(t^{-2})$, we have $\sum_{i\in\mathcal{V}}p_i^\star = o(\log t)$. Therefore, combining this fact with (159), (160), and (164), we obtain

$$\begin{aligned}\mathbb{P}_1&\left(\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right)\\ &= \tilde{\mathbb{E}}_1\left[\frac{\mathrm{d}\mathbb{P}_1(\cdot)}{\mathrm{d}\tilde{\mathbb{P}}_1(\cdot)}\mathbb{1}\left[\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right]\right]\\ &\geq \tilde{\mathbb{E}}_1\left[e^{\lambda^\star\log\frac{\alpha}{\beta} - C'}e^{-\left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]\|\mathcal{P}_t^\star\|}\right.\\ &\quad\left.\cdot e^{-\rho\left(\frac{1}{2}-\frac{\alpha-\beta}{(\alpha+\beta)\log\frac{\alpha}{\beta}}\right)\sum_{i\in\mathcal{V}}p_i^\star}\mathbb{1}\left[\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right]\right]\\ &\geq [1-O(t^{-2})]e^{-o(\log t)-\left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]\|\mathcal{P}_t^\star\|}\\ &\quad\cdot\tilde{\mathbb{P}}_1\left[\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right]\\ &\geq [1-O(t^{-2})](1-e^{-c\sqrt{\log t}})\\ &\quad\cdot e^{-o(\log t)-\left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]\|\mathcal{P}_t^\star\|}.\end{aligned} \tag{165}$$

Finally, substituting (165) in (155), we finally obtain

$$\begin{aligned}\mathsf{E}(\mathcal{P}^\star,\mathcal{Q}) &= \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star,\mathcal{Q})}{\log t}\\ &\leq \liminf_{t\to\infty} -\frac{\log \mathbb{P}_1\left[\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right]}{\log t}\\ &\leq \left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right]\liminf_{t\to\infty}\frac{\|\mathcal{P}_t^\star\|}{\log t}\\ &= \left[\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right](1+\gamma)\kappa(\gamma), \end{aligned} \tag{166}$$

as claimed.

C Additional Proofs

C.1 Proof of Theorem 4

Note that the first two individuals follow their private signal, that is, $Z_i = X_i$, for $i = 1, 2$. Therefore, if if $X_1 = X_2 = 2$, then it is clear that $\mathsf{R}_2 = (\beta/\alpha)^2$, which implies that the MAP estimator outputs 2 as its decision. Accordingly, it should be clear that if all future irrational players draw 2 as their private information, then the MAP estimator continues to output 2. The above scenario gives a lower bound on the error probability. Specifically, let Rev_t denote the set of revealers up to time t. It is clear that $|\mathsf{Rev}_t|$ follows a Poisson-Binomial distribution with mean $\mu = \|\mathcal{P}_t\| = o(\log t)$. Thus, for any $c > 1$, using (126) we get

$$\mathbb{P}\left[|\mathsf{Rev}_t| \geq c\,\|\mathcal{P}_t\|\right] \leq e^{-(c\log c - c+1)\|\mathcal{P}_t\|}. \tag{167}$$

Taking any c such that $c\log c - c + 1 > 0$, it is clear that the r.h.s. of (167) is less than half, and

$$\begin{aligned}&\mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^t, X_{t+1}) \neq \theta\right)\\ &\geq \mathbb{P}_1(X_1=2, X_2=2, X_i=2, \forall i\in\mathsf{Rev}_t)\\ &\geq \frac{1}{2}\left(\frac{\beta}{\alpha+\beta}\right)^2 \mathbb{P}_1\left(\bigcap_{i\in\mathsf{Rev}_t}\{X_i=2\}\middle| |\mathsf{Rev}_t|\leq c\,\|\mathcal{P}_t\|\right)\\ &\geq \frac{1}{2}\left(\frac{\beta}{\alpha+\beta}\right)^{2+c\|\mathcal{P}_t\|}. \end{aligned} \tag{168}$$

Thus,

$$\begin{aligned}\mathsf{E}(\mathcal{P},\mathcal{Q}) &= \liminf_{t\to\infty} -\frac{\log \mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^t, X_{t+1})\neq\theta\right)}{\log t}\\ &\leq [c\log(1+\gamma)]\cdot\liminf_{t\to\infty}\frac{\|\mathcal{P}_t\|}{\log t} = 0, \end{aligned} \tag{169}$$

as claimed.

C.2 Proof of Theorem 5

Since the proof of Theorem 5 follows the steps of the proof of Theorem 3 almost exactly, in this subsection we highlight the few technical differences only. Starting with the lower bounds, using the same steps as in Subsection B.1, one obtains the same upper bounds on the error probability as in (79) and (88), for $1 \leq \rho \leq \rho_1$ and $\rho_0 \leq \rho \leq 1$, respectively. In particular, for $1 \leq \rho \leq \rho_1$ recall that

$$\begin{aligned}\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q}) &\leq \left[C_0' e^{(1-\lambda_1^\star)C_3 - \delta(\gamma,\rho)\|\mathcal{P}_t\|} + \frac{1}{t^2}\right]\cdot(1-p_t)\\ &\quad + \frac{\beta}{\alpha+\beta}\cdot p_t. \end{aligned} \tag{170}$$

Therefore, we get

$$\mathsf{E}(\mathcal{P},\mathcal{Q}) = \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P}^\star,\mathcal{Q})}{\log t} \tag{171}$$

$$\geq 1 \wedge \left[\delta(\alpha/\beta,\rho)\cdot\liminf_{t\to\infty}\frac{\|\mathcal{P}_t\|}{\log t}\right] \tag{172}$$

$$= 1 \wedge \left[\mathsf{C}_\mathsf{p}\cdot\delta(\gamma,\rho)\right], \tag{173}$$

as claimed. Similarly, using (88), we get that

$$\mathsf{E}(\mathcal{P},\mathcal{Q}) \geq 1 \wedge \left[\mathsf{C}_\mathsf{p}\cdot\left(\delta(\gamma,\rho)-\frac{\gamma-1}{\gamma+1}(1-\rho)\right)\right], \tag{174}$$

for $\rho_0 \leq \rho_1$, as stated in Theorem 5.

The upper bounds in Subsections B.2–B.5 remain the same as well. In fact, the only differences are in Subsections B.3 and B.5. Specifically, for $1 \leq \rho \leq \rho_1$ the lower bound in (129) still holds true. Then, recall that

$$\begin{aligned}\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q}) &= \mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1}, X_t)\neq\theta\right)\cdot(1-p_t)\\ &\quad + \mathbb{P}(X_t\neq\theta)\cdot p_t\\ &\geq \mathbb{P}\left(\mathcal{E}_0\cap\mathcal{E}_1\cap\mathcal{E}_2\right)\cdot(1-p_t) + \frac{\beta}{\alpha+\beta}p_t, \end{aligned} \tag{175}$$

and thus combined with (129), we obtain

$$\mathsf{E}(\mathcal{P},\mathcal{Q}) = \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q})}{\log t} \tag{176}$$

$$\leq 1 \wedge \liminf_{t\to\infty} -\frac{\log \mathbb{P}_1\left(\mathcal{E}_1\cap\mathcal{E}_2|\mathcal{E}_0\right)}{\log t} \tag{177}$$

$$\leq 1 \wedge \left[\delta(\gamma,\rho)\cdot\liminf_{t\to\infty}\frac{\|\mathcal{P}_t^\star\|}{\log t}\right] \tag{178}$$

$$= 1 \wedge \left[\mathsf{C}_\mathsf{p}\cdot\delta(\gamma,\rho)\right], \tag{179}$$

as claimed. For $\rho_0 \leq \rho \leq 1$ we have a similar situation. Specifically, the lower bound in (165) still hold true with $\mathcal{P}^\star$ replaced by $\mathcal{P}$. Also, recalling (153) we have

$$\begin{aligned}\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q}) &= \mathbb{P}\left(\mathsf{MAP}_{\mathcal{Q}}(Z_1^{t-1}, X_t)\neq\theta\right)\cdot(1-p_t)\\ &\quad + \mathbb{P}(X_t\neq\theta)\cdot p_t\\ &\geq \left(\frac{\beta}{\alpha+\beta}\right)^{3+t^\star}\mathbb{P}_1\left[\max_{1\leq s\leq t}\mathsf{J}_s \leq \log^{3/4} t\right]\\ &\quad \cdot(1-p_t) + \frac{\beta}{\alpha+\beta}p_t. \end{aligned} \tag{180}$$

Thus, combining (165) and (180), we obtain

$$\begin{aligned}\mathsf{E}(\mathcal{P},\mathcal{Q}) &= \liminf_{t\to\infty} -\frac{\log \mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q})}{\log t} \\ &\leq 1 \wedge \left[\mathsf{C}_{\mathsf{p}} \cdot \left(\delta(\gamma,\rho) - \frac{\gamma-1}{\gamma+1}(1-\rho)\right)\right],\end{aligned}$$

as stated in Theorem 5.

C.3 Proof of Theorem 6

The proof of Theorem 6 follows from two facts. First, recall that a trivial lower bound on the error probability is $\mathsf{P}_{e,t}(\mathcal{P},\mathcal{Q}) \geq \frac{\beta}{\alpha+\beta}p_t$, which implies that $\mathsf{E}(\mathcal{P},\mathcal{Q}) \leq -\lim_{t\to\infty}\log p_t/\log t$. We next show that if $\mathcal{Q}$ is such that $\|\mathcal{Q}_t\|\,/\,\|\mathcal{P}_t\| \to \rho$, and $\rho_0 < \rho < \rho_1$, then the above also lower bounds the learning rate. Indeed, as before, using the same steps as in Appendix B.1, we get the upper bounds in (79) and (88), for $1 \leq \rho \leq \rho_1$ and $\rho_0 \leq \rho \leq 1$, respectively. However, since in this case $\|\mathcal{P}_t\| = \omega(\log t)$ the terms in the squared brackets at the r.h.s. of (79) and (88) are negligible compared to the other $\frac{\beta}{\alpha+\beta}p_t$. This implies that (79) and (88) are dominated by $\frac{\beta}{\alpha+\beta}p_t$ and thus $\mathsf{E}(\mathcal{P},\mathcal{Q}) \geq -\lim_{t\to\infty}\log p_t/\log t$, as well. Finally, it is left to show that $\mathsf{E}(\mathcal{P},\mathcal{Q}) = 0$ in the leftover cases, which follows from the same arguments as in Appendices B.2 and B.4, and therefore omitted.

C.4 Adversarial Model is Too Stringent

In this section we show that the error probability in (1) associated with any estimator is lower bounded by a constant, and accordingly, the total number of errors in (3) is proportional to the number of players N.

To this end, consider the set of revealers $\Pi_{\mathsf{N}} = [\mathsf{N}-\mathsf{V}_{\mathsf{N}}+1 : \mathsf{N}]$ in (3). This choice of Π_{N} corresponds to the case where all revealers appear at the end. Assume that $\mathsf{V}_{\mathsf{N}} = o(\mathsf{N})$, otherwise, $\mathsf{TE}(\mathsf{V}_{\mathsf{N}})$ is trivially proportional to $\Theta(\mathsf{N})$. Then, since all first $\mathsf{N}-\mathsf{V}_{\mathsf{N}}$ players are rational, with a positive probability a wrong cascade will occur. Indeed, this is just the classical herding experiment, proposed and studied in [Anderson and Holt, 1996, Anderson and Holt, 1997] (see also [David and Jon, 2010, Ch. 16]). In fact, each player $t \in [\mathsf{N}-\mathsf{V}_{\mathsf{N}}]$ is wrong with probability at least $\frac{\beta^2}{(\alpha+\beta)^2} = (1+\gamma)^{-2}$, which is the probability that the decisions of the first two players are wrong (both draw marbles of minority type). Therefore, the number of errors in (3) satisfies

$$\begin{aligned}\mathsf{TE}(\mathsf{V}_{\mathsf{N}}) &= \inf_{\hat{\theta}\in\hat{\Theta}} \sup_{\Pi_{\mathsf{N}}\subset[\mathsf{N}]:\ |\Pi_{\mathsf{N}}|=\mathsf{V}_{\mathsf{N}}} \sum_{t=1}^{\mathsf{N}} \mathsf{P}_{e,t}(\hat{\theta}_t,\Pi_{\mathsf{N}}) \\ &\geq \frac{\mathsf{N}-\mathsf{V}_{\mathsf{N}}}{(1+\gamma)^2},\end{aligned}$$

namely, of order $\Theta(\mathsf{N}-\mathsf{V}_{\mathsf{N}})$, which concludes the proof.

D Conclusion and Outlook

In this paper we have studied the effect of mismatch between players on information cascade, contrary to related works where full/partial mismatch was taken for granted. For the mismatch model considered in this paper we have identified when learning is possible and when it is not. Consequently, we demonstrated that the learning rate exhibits several surprising phase transitions.

We hope our work has opened more doors than it closes. There are many questions for future work:

1. It would be interesting to generalize our results to the case where more than two states are possible, each corresponding to multiple private signals.
2. In this paper we focus on . Studying the asymptotic learning rate and the total number of wrong errors of information cascades over random graphs (e.g., Erdős-Rényi random graph, stochastic block models, etc.) is very interesting and of practical importance.
3. Following our negative result on the worst-case model, studying minimax learning rates in adversarial models, by assuming a more *structured* geometry for the set of revealers in order to avoid trivial rates is quite challenging and interesting.
4. It is important to check whether rational players that do not know $\mathcal{P}$ can do better then just assuming some $\mathcal{Q}$. In particular, devising a universal scheme that attains (or at least does not lose too much) the optimal learning rate for $\mathcal{P}$, without knowing $\mathcal{P}$, is an important question. A reasonable approach would be using the same $q_t = \Theta(t^{-1})$, and adapt the leading constant in some way.
5. As discussed in the introduction, it is well-documented in social learning literature that a fully rational model often places unreasonable computational demands on Bayesian players (e.g., [Mossel and Tamuz, 2017]), hence understanding the impact of simpler more efficient strategies is desirable. This situation can be partially captured by our model, since a sub-optimal mismatched MAP, e.g., a majority rule, can be employed by the players intentionally to reduce computational complexity. There are of course other computationally efficient strategies that cannot be covered by our mismatch MAP framework, but we hope that the results and techniques developed in

our paper will prove useful in the analysis of other these strategies as well.