Recommender System Design via Online Feedback Optimization

Anonymous Authors[1]

Abstract

Conventional recommender systems enhance user engagement through personalized content. However, personalization usually induces significant side effects on opinion formation, such as polarization and echo chambers that need to be prevented. With this motivation, we design a recommender system algorithm that addresses user engagement maximization and opinion polarization mitigation by operating in feedback with the social platform. The recommender is agnostic about real-time opinions, network topology, and users' clicking behaviour, all estimated online. We numerically verify the efficacy of the designed recommender on synthetic data. We show that by providing network-aware recommendations to the users as opposed to users' tailored content, we significantly reduce polarization effects without sacrificing user engagement.

1. Motivation

Online social platforms use recommender systems to provide users with tailored content to maximize engagement over the platform. The state-of-the-art algorithms for recommender systems exploit methods, such as content-based filtering (Bansal et al., 2015) and collaborative filtering (Eirinaki et al., 2014), that combine information personalization, popularity and similarity of interests with other users to provide a set of media feed that attracts users' interests. However, it is a well established fact that content personalization leads to undesired effects over users opinions such as echo chambers formation and polarization (Lazer, 2015; Bakshy et al., 2015). By drawing on Online Feedback Optimization (OFO) (Hauswirth et al., 2024), we design a recommender system algorithm in feedback with the social platform whose aim is to maximize users' engagement while penalizing opinion polarization (see Fig.1). We show that, by considering the network topology of social interactions over the platform, something that traditional machine learning algorithms usually do not, we are able to reduce polarization in opinions without sacrificing users' engagement. The dynamics of opinions, the agent's clicking behaviour and the topology of interactions are all assumed to be unknown and inferred online.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the Workshop on Foundations of Reinforcement Learning and Control at the International Conference on Machine Learning (ICML). Do not distribute.

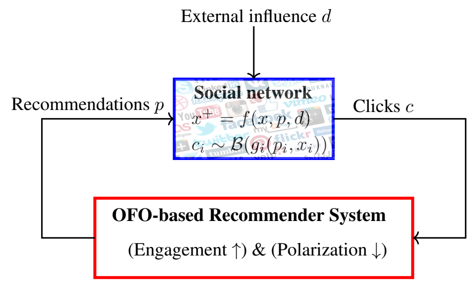

Figure 1. Illustration of the closed-loop between social network and recommender system.

2. Problem Setting

We consider the problem of designing a recommender system for a social network consisting of n users (Fig. 1), indexed by $i \in [n]$. User opinions are collected into a vector $x \in [-1, 1]^n$, with x_i being the opinion of the i-th user. The temporal evolution of the users' opinions is dictated by

$$x^{k+1} = f(x^k, p^k, d), \tag{1}$$

where $p \in [-1,\ 1]^n$ is the position vector of the recommendations, $d \in [-1,\ 1]^n$ represents an external influence to the platform and $f : ([-1,\ 1]^n)^3 \to [-1,\ 1]^n$ encodes the influence of interactions among users. We assume there is a single *polarizing* topic of discussion on the platform. Therefore, a recommendation with position $p = +1$ (-1) can be interpreted as a news strongly in support of (against) the issue. We provide an example of (1) in Appendix A.2.

The following assumption ensures that the dynamics (1) are well-posed and admit a steady-state mapping.

Assumption 1 (Well-posedness)**.** The following hold: (i) The dynamics (1) are forward invariant in $[-1,\ 1]^n$, i.e.,

$$x^0, p^0, d \in [-1,\ 1]^n \ \Rightarrow\ x^k \in [-1,\ 1]^n,\ \forall k \in \mathbb{N}_0.$$

(ii) The dynamics (1) is uniformly exponentially stable and admit a unique steady-state map $h : \left([-1,\ 1]^n\right)^2 \to [-1,\ 1]^n$ satisfying

$$h(p, d) = f(h(p, d), p, d), \quad \forall p, d \in [-1,\ 1]^n.$$

(iii) The map $h(p, d)$ is continuously-differentiable and $L-$Lipschitz (Nesterov, 2014) with respect to p. □

Engagement of the users over the platform is measured in terms of the clicking ratio on the provided news articles. For a user i, we model the probability of clicking on a recommendation as a random variable c_i drawn from a Bernoulli distribution with unknown argument $g_i(p_i, x_i)$, i.e. $\mathcal{B}(g_i(p_i, x_i))$. The argument $g_i(p_i, x_i) \in [0, 1]$, represents the probability that the user i, with opinion x_i clicks on a news expressing the position p_i. We will refer to $g_i(p_i, x_i)$ as the *clicking behaviour* of user i. We provide an example of a clicking behaviour in Appendix A.3

3. Problem Formulation

The goal of the recommender system is to provide recommendations that optimize a specific metric, denoted as $\varphi(p, x)$. We consider a multi-objective cost function that combines engagement maximization and polarization mitigation as

$$\varphi(p, x) = \varphi^{\text{clk}}(p, x) + \varphi^{\text{pol}}(x), \tag{2}$$

The engagement-related term in the cost (2) is defined as

$$\varphi^{\text{clk}}(p, x) = -\sum_{i\in[n]} \mathbb{E}_{c_i\sim\mathcal{B}(g_i(x_i,p_i))}[\, c_i \,],$$

where $\mathbb{E}_{c_i\sim\mathcal{B}(g_i(x_i,p_i))}[\, c_i \,]$ is the expectation of user i's clicking, given their opinion x_i and a recommendation with position p_i. The second term in (2) is given by $\varphi^{\text{pol}}(x) := \sum_{i\in[n]} s_i(x)$, where s_i is a soft penalty function defined as

$$s_i(x_i) = \begin{cases} (x_i - \epsilon_1)^2 & x_i < \epsilon_1 \\ 0 & \epsilon_1 \le x_i \le \epsilon_2 \\ (\epsilon_2 - x_i)^2 & x_i > \epsilon_2 \end{cases}.$$

The parameters $\epsilon_1 \le \epsilon_2$ are used to control the degree of penalty towards extreme opinions. Specifically, a smaller positive (negative) choice for ϵ_2 (ϵ_1) indicates a higher penalty on extreme positive (negative) opinions.

The recommender aims at regulating the system (1) to the solution of the following steady-state optimization problem:

$$\underset{p,\,x}{\text{minimize}} \quad \varphi(p, x) \tag{3a}$$

$$\text{s.t} \quad x = h(p, d) \tag{3b}$$

$$p \in [-1, 1]^n \tag{3c}$$

Problem (3) is, in principle, non-convex, since both the steady-state mapping in (3b) and clicking behaviour in (3a) are, in general, unknown. If the dynamics of opinions (f), the opinions (x_i) in (1) and the clicking behavior (g_i) were known, and an accurate prediction of external influences d were available, then we could solve (3) offline. However, in practice, none of these information is readily available. Instead, the recommender system only has access to the users' online feedback in the forms of clicks $\{c_i^k\}$ on the provided recommendations. Thus, the only measurements we collect from the users is the observed clicking ratio $y^k := \sum_{t=k-T}^{k} c^t/(T+1), \forall\, k \ge T$, for some time window T. Our feedback control problem is formally stated as:

Problem 1. *Design a feedback controller so that* (1) *tracks a solution* (p^*, x^*) *of the optimization problem* (3)*, by assuming only clicks* c_i^k *are available.*

4. Problem Solution

We approach Problem 1 by designing a dynamic feedback controller inspired by the projected-gradient descent algorithm in (Belgioioso et al., 2021). The resulting recommender dynamically generates positions as

$$p^{k+1} = \mathbb{P}_{[-1,1]^n}\left[p^k - \eta\, \Phi(p^k, x^{k+1})\right], \quad \forall k \in \mathbb{N} \tag{4}$$

where $\mathbb{P}_{[-1,1]^n}[z]$ represents the Euclidean projection of some $z \in \mathbb{R}^n$ onto $[-1, 1]^n$, η is the step-size, and

$$\Phi(p, x) := \nabla_p \varphi(p, x) + \nabla_p h(p, d)^\top \nabla_x \varphi(p, x) \tag{5}$$

represents the gradient obtained by applying the chain-rule of differentiation to the cost $\varphi(p, x)$ in (3), with opinions at steady state, i.e. $x = h(p, d)$.

In practice, evaluating the gradient (5) at each sampling instant requires access to: (i) Real-time users' opinions x^k; (ii) Sensitivity mapping $\nabla_p h(p, d)$; (iii) Gradients $\nabla_p \varphi(p, x)$ and $\nabla_x \varphi(p, x)$. None of these information is readily available, making a direct implementation of the recommender design (4) impractical.

To cope with these challenges, we augment the controller (4) with three auxiliary levels that estimate users' opinions, sensitivity, and clicking behaviour online, as illustrated in Fig. 2. Specifically, we structure the design on three levels: Level 1: Real-time opinions and users' clicking behaviour estimation via supervised learning; Level 2: Online sensitivity learning via Kalman filtering; Level 3: Gradient estimation via a forward difference method. In the following three sections, we briefly describe each layer and analyze the stability properties of the feedback interconnection.

4.1. Level 1: Opinion & Clicking Behaviour Estimation

The steady-state opinion and clicking behaviour are estimated using an Artificial Neural Network (ANN). Note that

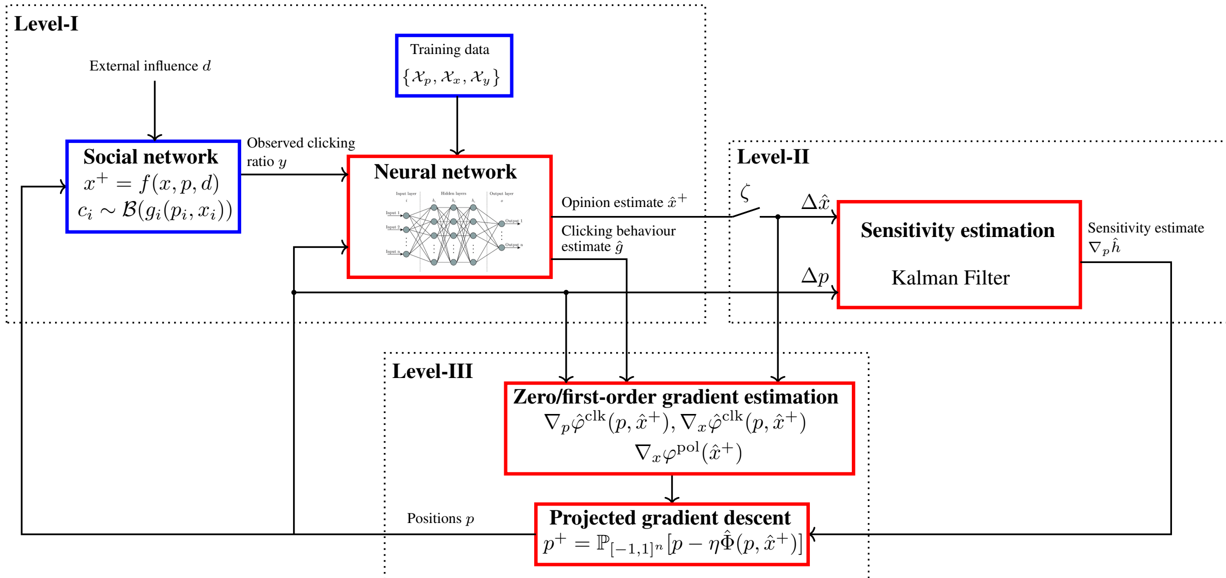

Figure 2. Block diagram of the proposed recommender system design with the three levels. Level 1: Opinion and clicking behaviour estimation, Level 2: Sensitivity estimation, Level 3: Gradient estimation and Optimization.

the use of the ANN allows us take into account for the dynamics of opinions thus circumventing the problem of getting access to online state measurements.

To bound the resulting opinions and clicking behaviour estimation error, we work under the following regularity assumptions (Cothren et al., 2023; Dean & Recht, 2021).

Assumption 2 (Clicking behaviour)**.** The clicking behaviour $g(p, x)$ is M_x-Lipschitz with respect to x, and L_p- and L_x-smooth with respect to p and x, respectively. □

Assumption 3 (Artificial Neural Network)**.** It holds that

1. The image $g(\mathcal{P}, \mathcal{X})$ is compact for any $\mathcal{X}, \mathcal{P} \subseteq [-1, 1]^n$.
2. There exist a continuous mapping $\beta : [0, 1]^n \times [-1, 1]^n \to [-1, 1]^n$ such that $\beta(y, p) = x + \theta(x)$, with $\|\theta(x)\| \le \theta_x$, for all $x \in \mathcal{X}$, and $\theta_x < \infty$. Moreover, the image $\beta(\mathcal{Y}, \mathcal{P})$ is compact for any $\mathcal{Y} \subseteq [0, 1]^n, \mathcal{P} \subseteq [-1, 1]^n$.
3. There exists a constant $\alpha_y < \infty$ such that the composite function $g(p, \beta(y, p)) = y + \nabla_x g(p, x)^\top \theta(x) + \alpha(y)$, with $\|\alpha(y)\| \le \alpha_y$, for all $y \in \mathcal{Y}$. □

It follows by Assumptions 2–3 that the modelling errors of the opinion map $\beta(y, p)$ and clicking behaviour $g(p, x)$ are upper-bounded by θ_x and α_y, respectively.

There are two ANNs involved in this process: One to infer the opinions and another one to infer the clicking behaviour. As far as the opinion estimation is concerned, for every user i, the ANN takes as input the clicking ratio y_i and the position vector p and gives as output the users' opinion estimate $\hat{x}_i$. The reason why we take the whole position vector p as an input is due to the network structure of the problem: The position provided to users connected with agent i will also (indirectly) affect agent i's steady-state opinion x_i. Note that, x_i does not directly depend on the clicking ratios of other users. Thus, we have $n+1$ input neurons (n associated with p and one associated with y_i) and 1 output neuron for each user. Further, we use one intermediate layer comprising of $n + 2$ neurons with a hyperbolic tangent activation function.

For clicking behaviour estimation, for each user i, the ANN takes as input the steady-state opinion x_i and the position provided to the i^{th} user, p_i, and provides as output the clicking behaviour estimate $\hat{g}_i$. The clicking behaviour is specific for each user, therefore, a user's likelihood of clicking on the provided position is not affected by neither other users' opinions nor other users' positions. Thus, we have 2 input neurons (p_i, x_i) and one output (y_i) neuron for each user. Further, we use 3 intermediate layers comprising of 5 neurons with a hyperbolic tangent activation function.

The procedure to acquire the training data is explained in Appendix A.4. We now describe the upper-bounds on the steady-state opinion and clicking behaviour estimation errors. We define $e_x := h(p, d) - \hat{\beta}(y, p)$ and $e_y := g(p, h(p, d)) - \hat{g}(p, x)$ as the opinion and clicking behaviour estimation errors, respectively, where, $\hat{\beta}$ and $\hat{g}$ are the opinion estimate and clicking behaviour maps learned by the ANN.

Lemma 1 (Opinion estimation error)**.** *Under Assumption 3, the steady-state opinion estimation error is upper bounded in the ℓ_2-norm as $\|e_x\| \leq \epsilon_x$, with*

$$\epsilon_x := \sqrt{n}\Big[3\sup_{X\in\mathcal{X}}\|\beta(X)-\hat{\beta}(X)\|_\infty + 2\sup_{i\in[n]}\omega_{\beta_i}(\gamma_x)+ \gamma_x\sup_{i\in[n]}|v_{0,i}|\Big]+\theta_x, \tag{6}$$

where $\sup_{X\in\mathcal{X}}\|\beta(X)-\hat{\beta}(X)\|_\infty$ represents the maximum opinion estimation error during training of the ANN, $\omega_{\beta_i}(\gamma_x)$ denotes the minimum modulus of continuity of β_i on the training set $\mathcal{X}$ (see Definition 4, Appendix A.1) and $v_{0,i}$ represents the bias from the hidden layer to the output layer of the ANN on user i.

Proof. Refer to Appendix A.5. □

Lemma 2 (Clicking behaviour estimation error)**.** *Under Assumptions 2–3, the clicking behaviour estimation error is upper-bounded in the ℓ_2-norm as $\|e_y\| \leq \epsilon_g$, with*

$$\epsilon_g \leq \sqrt{n}\Big[3\sup_{Y\in\mathcal{Y}}\|g(Y)-\hat{g}(Y)\|_\infty + 2\sup_{i\in[n]}\omega_{g_i}(\gamma_y)+ \gamma_y\sup_{i\in[n]}|w_{0,i}|\Big]+M_x\epsilon_x+\alpha_y, \tag{7}$$

where $\sup_{Y\in\mathcal{Y}}\|g(Y)-\hat{g}(Y)\|_\infty$ is the maximum estimation error during ANN training, $\omega_{g_i}(\gamma_y)$ denotes the minimum modulus of continuity of g_i on the training set $\mathcal{Y}$ and $w_{0,i}$ represents the estimated bias weight from the hidden layer to the output layer of the ANN on user i.

Proof. Refer to Appendix A.6 □

4.2. Level 2: Online Sensitivity Learning

In order to estimate the sensitivity $\nabla_p h(p,d)$ in (5) in real time, similar to (Picallo et al., 2022), we adopt a Kalman filter based approach. We denote by $\ell \in \mathbb{R}^{n^2}$ the vectorized sensitivity $\nabla_p h(p,d) \in \mathbb{R}^{n\times n}$, namely, $\ell := \mathrm{vec}(\nabla_p h(p,d))$, and model the sensitivity dynamics as a random walk

$$\ell^k = \ell^{k-1} + w^{k-1},$$

where, $w^k \sim \mathcal{N}(0_{n^2}, Q^k)$ is the process noise, with Q^k as its corresponding covariance matrix. The measurement model is described by

$$\Delta x_{\mathrm{ss}}^{k+1,k} = \Delta\tilde{p}^{k,k-1}\ell^k + v^k,$$

where $\Delta x_{\mathrm{ss}}^{k+1,k} := h(p^k,d) - h(p^{k-1},d)$ is the change in the steady-state opinions for a corresponding change in positions $\Delta p^{k,k-1} := p^k - p^{k-1}$. Further, $\Delta\tilde{p}^{l,m} := \Delta(p^{l,m})^\top \otimes I_n \in \mathbb{R}^{n\times n^2}$, where $\otimes$ indicates the Kronecker product. The measurement noise is described by $v^k \sim \mathcal{N}(0_n, R^k)$, where R^k is its corresponding covariance matrix. This noise accounts for the contribution of the external influence d to the change of opinions Δx.

Similar to (Picallo et al., 2022), the posterior update of estimates $\hat{\ell}^k$ and covariance Σ^k are given by

$$\begin{aligned}\hat{\ell}^k &= \hat{\ell}^{k-1} + \zeta^k K^{k-1}\big(\Delta\hat{x}^{k+1,\tau_i+1} - \Delta\tilde{p}^{k,\tau_i}\hat{\ell}^{k-1}\big)\\ \Sigma^k &= \Sigma^{k-1} + \zeta^k\big(Q^k - K^{k-1}\Delta\tilde{p}^{k,\tau_i}\Sigma^{k-1}\big),\end{aligned} \tag{8}$$

where ζ enforces an auxiliary trigger mechanism, with $\zeta^k = 1$ for k being integer multiples of the time period T. Further, we refer to $\mathcal{T}$ as the set of trigger time instances and τ_i being the latest trigger time instant before k. The trigger mechanism is introduced to enforce time-scale separation between the plant dynamics and the controller updates (Hauswirth et al., 2021) and to allow a sufficient number of clicks to ensure a proper estimate of the clicking ratio, which is needed for opinion estimation.

Note that the sensitivity learning is based on the opinion estimates $\hat{x}$, introduced in Section 4.1, rather than the real opinions. Finally, the Kalman filter gain K^k in (8) is given by

$$K^k = \Sigma^k(\Delta\tilde{p}^{k,\tau_i})^\top(R^k + \Delta\tilde{p}^{k,\tau_i}\Sigma^k(\Delta\tilde{p}^{k,\tau_i})^\top)^{-1}.$$

Next, we postulate some regularity assumptions for the process and measurement models.

Assumption 4 (Gaussian noise)**.** The process and measurement noise w, v are white Gaussian. Moreover, the steady-state opinion estimation error $e_x = h(p,d) - \hat{\beta}(y,p)$ is uncorrelated with w and v. □

The assumption of white noise for the process model is standard in the context of sensitivity learning in feedback optimization (Picallo et al., 2022). Intuitively, the process perturbation determines the degree of trust one puts on the sensitivity estimates. It is also reasonable to assume that the external influence is uncorrelated among users in certain cases, for example the extended FJ model in Appendix A.2 where $d \equiv x_0$. We also note that since $\hat{x}$ is estimated using the ANN, the steady-state estimation error e_x is not correlated with the process or measurement noise in (4.2). Under Assumption 4, the covariances simplify as $Q^k = (\sigma_q^k)^2 I_{n^2}$ and $R^k = (\sigma_r^k)^2 I_n$, for some $\sigma_q, \sigma_r > 0$.

To ensure that the sensitivity matrix is correctly inferred, we must guarantee that the input positions Δp are persistently exciting (Willems et al., 2005) (see Definition 3). This is carried out by introducing a dither signal in (4).

4.3. Level 3: Gradient Estimation & Optimization

In order to estimate the gradient of the engagement maximization cost φ^{clk}, we use the *finite forward difference* method as in (Scheinberg, 2022), yielding

$$\nabla_x \hat{\varphi}_i^{\mathrm{clk}}(p,x) = \frac{\hat{\varphi}^{\mathrm{clk}}(p, x+\mu e_i) - \hat{\varphi}^{\mathrm{clk}}(p,x)}{\mu} \tag{9}$$

$$\nabla_p \hat{\varphi}_i^{\mathrm{clk}}(p,x) = \frac{\hat{\varphi}^{\mathrm{clk}}(p+\mu e_i, x) - \hat{\varphi}^{\mathrm{clk}}(p,x)}{\mu}, \tag{10}$$

where $\nabla_j \hat{\varphi}_i^{\mathrm{clk}}, j \in \{p, x\}$ denotes the i^{th} entry of the gradient, $e_i \in \mathbb{R}^n$ refers to the i^{th} vector of the canonical basis of $\mathbb{R}^n$ and μ is a smoothing parameter. The cost $\hat{\varphi}^{\mathrm{clk}}(p,x) = -1_n^\top \hat{g}(p,x)$ is evaluated as described in Section 4.1. The smoothing parameter μ is chosen small enough so that the gradient estimate provides a good approximation of the true value. However, having μ in (9)-(10) too small can lead to numerical instability, as formalized in the following:

Lemma 3 (Gradient estimation error)**.** *Under Assumptions 2–3, the gradient estimation error is upper bounded as*

$$\|\nabla_j \hat{\varphi}^{\mathrm{clk}} - \nabla_j \varphi^{\mathrm{clk}}\| \leq 2n^{3/4}\sqrt{\epsilon_g L_j}, \quad j \in \{x, p\}.$$

Moreover, the upper bound is tight and reached in correspondence of $\mu^* = 2n^{1/4}\sqrt{\epsilon_g / L_j}$, with $j \in \{x, p\}$.

Proof. Refer to Appendix A.7. □

We now present the projected-gradient update rule in (4) augmented with sensitivity, state, and gradient estimations, and state stationarity bounds with respect to the positions p^k. The augmented update reads compactly as

$$p^{k+1} = \mathbb{P}_{[-1,1]^n}\left[p^k - \zeta^k \eta\, \hat{\Phi}(p^k)\right], \tag{11}$$

where the gradient surrogate $\hat{\Phi}(p^k)$ is constructed by combining sensitivity, opinion, and gradient estimates as

$$\begin{aligned}\hat{\Phi}(p^k) =& \nabla_p \hat{\varphi}^{\mathrm{clk}}(p^k, \hat{x}^{k+1}) + (\hat{H}^k)^\top \nabla_x \hat{\varphi}^{\mathrm{clk}}(p^k, \hat{x}^{k+1}) \\ &+ \gamma (\hat{H}^k)^\top \nabla_x \varphi^{\mathrm{pol}}(\hat{x}^{k+1}) - w_{pe}^k/\eta,\end{aligned}$$

where, $\hat{x}$ is the opinion estimate, $\hat{H}^k := \nabla_p \hat{h}(p^k, d)$ is the sensitivity estimate at time k. The additional term $w_{pe}^k \sim \mathcal{N}(0_n, \sigma_{pe}^2 I_n)$ is a dither signal that ensures persistency of excitation of the inputs.

In Algorithm 1, we provide the pseudo-code of the proposed recommender system design. In the offline phase, training of the neural network is carried out for opinion and clicking behaviour estimation. In the online phase a new sensitivity estimate is generated and new positions are provided to the users periodically, every T time steps. The recommendations are provided for N time instances in total.

Algorithm 1 $[y^*, p^*]$ = Recommender$[N, T, n]$

> **Initialization**
> Collect training data
> Opinion and clicking behaviour map estimation $(\hat{\beta}, \hat{g})$
> **Optimization phase**
> **for** $k \in [0, N]$ **do**
>   Collect clicks c^k from users
>   Clicking ratio $y^k \leftarrow \frac{\sum_{t=\tau_i}^k c^t}{k-\tau_i+1}, \tau_i = (i-1)T < k$
>   **if** $\zeta^k = 1$ **then**
> $\mathcal{T} \leftarrow$ append$[k]$
>   Opinion estimate: $\hat{x}_i^{k+1} \leftarrow \hat{\beta}_i(y_i^k, p^k)$
>   Sensitivity estimate: $\hat{H}^k$ using (8)
>   Gradient estimate: Using $\hat{g}, \hat{x}$ for (9)-(10)
>   Obtain p^{k+1} with OFO using (11)
>   **else**
> $\hat{H}^k \leftarrow \hat{H}^{k-1}$
> $p^{k+1} \leftarrow p^k$
>   **end if**
> **end for**$[y^*, p^*] \leftarrow [y^k, p^{k+1}]$

4.4. Closed-loop Convergence Guarantees

In this section, we provide convergence guarantees for the sensitivity estimation process (8) and for the closed-loop interconnection between the opinions and recommendations.

Now, we state the main convergence result with respect to the sensitivity estimation error.

Theorem 1 (Sensitivity estimate convergence)**.** *Under Assumptions 1, 4 and persistently exciting inputs Δp (Definition 3), the sensitivity estimation error $e^k := \ell^k - \hat{\ell}^k$ has its variance bounded in norm, i.e., there exist positive constants $c_1, c_2, C_f < \infty$ and $\xi \in (0,1)$ such that*

$$\mathbb{E}\left[\|e^k\|^2\right] \leq C_f + (c_1 \xi^{c_2 T})^{2|\mathcal{T}|}\mathbb{E}\left[\|e^0\|^2\right]$$

where $C_f = \frac{1}{1-(c_1\xi^{c_2 T})^2}\left[(T-1)\overline{\sigma}_q^2 + K_m^2(\overline{\sigma}_r^2 + 2\epsilon_x^2)\right]$ with $\overline{\sigma}_q = \sup_{k\in\mathbb{N}_0} \sigma_q^k$, $\overline{\sigma}_r = \sup_{k\in\mathbb{N}_0} \sigma_r^k$, $K_m = \sup_{k\in\mathbb{N}_0} \|K^k\|$, with $c_1\xi^{c_2 T} < 1$.

Proof. Refer to Appendix A.8. □

Note that the variance upper bounds depend on the opinion estimation error ϵ_x. Further, note that increasing the sampling period T reduces the bias and variance error through the term $1 - c_1\xi^{c_2 T}$. This is expected as increasing T guarantees greater time-scale separation between opinion dynamics and recommendations. Finally, the upper bound on the sensitivity error variance is proportional to the noise variance of the opinions through the term $\overline{\sigma}_r^2$.

To quantify performance on the recommendations, we use the *fixed-point residual mapping* (Eq. (5), (J. Reddi et al., 2016)):

$$\mathcal{G}(p) := \frac{1}{\eta}\Big(p - \mathbb{P}_{[-1,1]^n}\left[p - \eta\Phi(p, h(p, d))\right]\Big), \quad (12)$$

The fixed-point residual mapping is zero at a critical point of (3), and is a common metric to quantify convergence of iterative algorithms in non-convex regimes (Nesterov, 2014). We are now ready to state the main convergence result. Given that opinions are directly estimated at steady-state, we circumvent the need to separately prove convergence on the opinion dynamics. The following theorem proves convergence of the closed-loop system.

Theorem 2 (Closed-loop scheme convergence). *Let Assumptions 1-4 hold. For* $\eta \in (0, \frac{1}{2L'})$, $\mu^* = 2n^{1/4}\sqrt{\epsilon_g/L_j}, j \in \{p, x\}$, *the sequence* $\{p^k\}_{k\in\mathbb{N}}$ *generated by* (11) *satisfies*

$$\frac{1}{|\mathcal{T}|}\sum_{\substack{l\in\mathcal{T}\\ l\leq k}} \mathbb{E}\left[\|\mathcal{G}(p^l)\|^2\right] \leq K_1, \quad \forall k \geq T \quad (13)$$

where K_1 *is given by*

$$K_1 = \frac{6}{1-2\eta L'}\Bigg\{\underbrace{\frac{\varphi(p^0, h(p^0, d)) - \varphi^*}{3\eta|\mathcal{T}|}}_{\kappa_1} + \frac{\sigma^2_{\text{pe}}}{\eta^2} + 4\Bigg[\underbrace{\gamma^2 L^2\epsilon_x^2 +}_{\kappa_2}$$

$$\underbrace{n^{3/2}\epsilon_g(L_p + L^2 L_x)}_{\kappa_3}\Bigg] + 2(2n\gamma^2 + M_x^2 + 4n^{3/2}\epsilon_g L_x)\frac{\sum_{\substack{l\in\mathcal{T}\\ l\leq k}}\mathbb{E}[\|e^l\|^2]}{|\mathcal{T}|}\Bigg\},$$

with $L' = L_p + L^2(L_x + 2)$ *denoting the Lipschitz constant of the gradient* $\Phi(\cdot, h(\cdot, d))$, *and* $\varphi^* = \inf_{p\in[-1,1]^n} \varphi(p, h(p, d)) > -n$ *the cost function value at a locally optimal solution.*

Proof. Refer to Appendix A.9. □

The bound K_1 shows that achieving an accurate solution of (3), i.e., a local minima $(p^*, h(p^*, d))$, is limited by the deviation of the initial cost at $(p^0, h(p^0, d))$ from the optimal one φ^* (i.e. term κ_1), the polarization and engagement gradient estimation errors (i.e. terms κ_2, κ_3, respectively) including both opinion and clicking behaviour estimation errors, the variance of the dither signal, and the sensitivity estimation error variance $\mathbb{E}[\|e^l\|^2]$.

5. Numerical Results

We briefly discuss the results obtained in simulation with the proposed recommendation algorithm. We then make a performance comparison between our algorithm and other OFO algorithms that benefit of more information. We also discuss the effects of personalization and show the benefits of network awareness as opposed to decoupled recommendations.

Table 1. Methods for comparison

Method	Sensitivity	Opinions	Clicking behaviour
M_1 (Oracle)	√	√	√
M_2	×	√	√
M_3	×	×	√
M_4 (Alg. 1)	×	×	×

5.1. Experimental Setting

For the simulations, we consider synthetic data with $n = 15$ users in a social network graph with the user's opinions evolving based on an extended Friedkin-Johnsen model, representing the networked version of the one in (Rossi et al., 2022):

$$x^{k+1} = (I_n - \Gamma_p - \Gamma_d)Ax^k + \Gamma_p p^k + \Gamma_d d^k,$$

where Γ_d, Γ_p are positive diagonal matrices such that $\Gamma_p + \Gamma_d \preceq I_n$, and describe the impact of d and p over the opinions, respectively, and A is a row-stochastic adjacency matrix encoding the social interconnections of the users. The external influence is modelled as a prejudice term, the initial opinion x^0.

We consider the following two different clicking behaviours:

$$C_A := c_i^k \sim \mathcal{B}\left(\frac{1}{2} + \frac{1}{2}x_i^k p_i^k\right),$$

$$C_B := c_i^k \sim \mathcal{B}\Big(\frac{1}{2} + \frac{1}{2}e^{-4(x_i^k - p_i^k)^2}\Big).$$

Clicking behaviour C_A represents confirmation bias over extreme positions (Rossi et al., 2022), and C_B models confirmation bias towards any position $p_i \in [-1, 1]$. To incorporate diversity in clicking behaviours, we randomly assign 8 users to follow clicking behaviour C_A and the remaining ones are attributed C_B.

For the polarization cost $\varphi^{\text{pol}}(x)$ in (2), we set $\epsilon_1 = -0.5$ and $\epsilon_2 = 0.5$. Thus, opinions lying outside the region $[-0.5, 0.5]^n$ are penalized. Further, we set $\gamma = 1$ in (2), thus giving equal importance to engagement maximization and polarization reduction.

5.2. Performance Comparisons

We make a comparison of our algorithm, referred in Table 1 as M_4 with other OFO approaches ($M_1 - M_3$) benefiting from more information. Table 1 summarizes the methods used for comparison and their attributes. For the oracle (method M_1), we employ the standard projected-gradient controller (4), wherein opinions can be directly measured, and the sensitivity and users' clicking behavior is known. Method M_2 requires online sensitivity estimation using a Kalman filter. Method M_3 requires opinion estimation using supervised learning in addition to sensitivity estimation. Finally, method M_4 is our algorithm, combining sensitivity, opinion, clicking behaviour and gradient estimation. It must be noted that the comparisons are not carried over a level-playing field. The methods $M_1 - M_3$ carry a significant advantage over M_4. In fact, $\varphi^{\text{clk}}(p, x)$ can be perfectly computed in $M_1 - M_3$, given that an analytic expression for the users' clicking behaviours is available.

To analyze the convergence of our algorithm and compare it with other OFO methods, we consider the fixed-point residual mapping norm $\|\mathcal{G}(p^k)\|^2$ (Fig. 3). It can be observed that convergence is empirically established in all the proposed methods with this metric. Although our proposed method manifests slightly inferior performances, we emphasize that the clicking behaviour estimation we perform is completely data-driven and online. The availability of an analytical form for the clicking behaviour $g(p, x)$, from which the other methods benefit, would be too ideal for real-life settings.

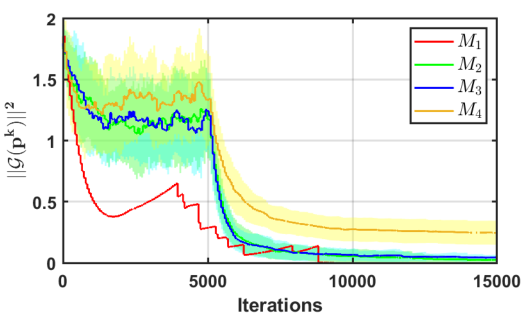

Figure 3. Evolution of the fixed-point residuals $\|\mathcal{G}(p^k)\|^2$ for the algorithms in Table 1. The bold lines represent the mean and the shaded region are the ± 1 standard deviation across the 50 Monte-Carlo trials.

5.3. Benefits of Network Awareness

We show the advantages of providing network-aware recommendations to the users, as opposed to individual decoupled recommendations. We compare our approach with two methods: the recommender from (Rossi et al., 2022) and a naive OFO algorithm both accounting for each user as isolated from the network. For the naive method, we make two significant changes from our proposed OFO approach in Algorithm 1. Since the social network is not considered, for

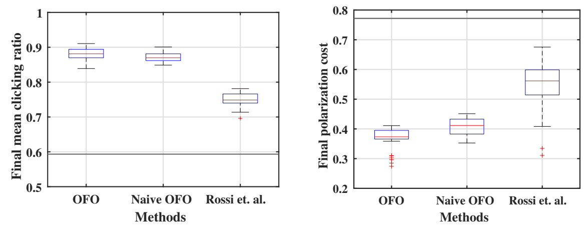

Figure 4. Steady-state mean clicking ratio (left) and polarization (right) obtained by Algorithm 1 (OFO), its network-agnostic version (Naive OFO), and the recommender design in (Rossi et al., 2022). The black lines represent the initial ideal mean clicking ratio (left) and the initial polarization (right).

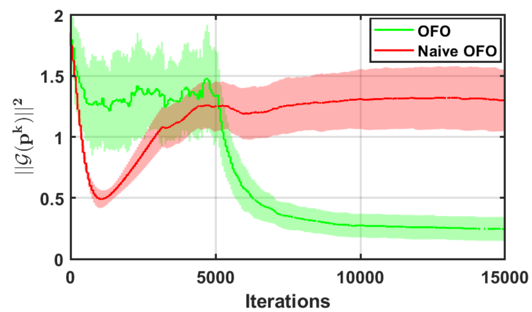

Figure 5. Evolution of the *fixed-point residuals* $\|\mathcal{G}(p^k)\|^2$ obtained by applying Algorithm 1 (OFO) and its network-agnostic version (Naive OFO). The solid lines represent the mean and the shaded region (± 1 standard deviation) the range of changes across the 50 Monte Carlo trials.

each user $i \in [n]$, we train the neural network for opinion estimation using only their own positions and acceptance ratio, i.e. $\hat{x}_i = \hat{\beta}_i(y_i, p_i)$ rather than $\hat{\beta}_i(y_i, p)$. Further, we do not carry out sensitivity estimation and instead provide a random constant diagonal sensitivity with entries in $[0, 0.5]$ in each Monte-Carlo simulation, thus considering each user as isolated.

The intuition behind the proposed naive method is that if the network did not contribute to engagement maximization and/ or polarization minimization, an educated guess for the sensitivity estimate $\hat{H}$ would lead to the same performance as our approach. An advantage of the naive OFO would be a massive reduction in the computational burden, as the online sensitivity estimation using the Kalman filter, which is computationally expensive, would no longer be required.

We observe from Fig. 4 that although our proposed algorithm returns a similar engagement when compared to the naive OFO algorithm, it returns a much lower polarization cost. Therefore, we conclude that network-aware recommendations lead to a significant reduction in polarization without sacrificing much in terms of users' engagement. Moreover, the choice of a random diagonal sensitivity matrix in the naive OFO method lowers the performance of the overall optimization problem (3), as seen from Fig. 5. We also observe a larger overshoot in $\|\mathcal{G}(p^k)\|^2$ with the naive OFO method.

6. Conclusion

Our aim is to improve the understanding of real-world phenomena using a simplified yet insightful model. We designed a recommender system that simultaneously maximizes user engagement and mitigates polarization. Our recommender system solely relies on clicks and does not require any prior knowledge about opinion dynamics and users' clicking behavior. We provided theoretical optimality and closed-loop guarantees for the resulting recommender-social network interconnection. Finally, our simulations demonstrated that our recommender performs favorably against other approaches that do not leverage information about users' interconnections. We provide evidence that polarization risk should be considered at the recommendation level.

Future research directions include relaxing the smoothness assumptions on the clicking behaviour and incorporate other interest attractors towards recommendation other than confirmation bias, e.g. repulsion. Finally, we aim for our network-centered perspective to enhance the existing literature on opinion polarization caused by algorithmic systems and inspire effective mitigation strategies.

References

Bakshy, E., Messing, S., and Adamic, L. A. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.

Bansal, T., Das, M., and Bhattacharyya, C. Content driven user profiling for comment-worthy recommendations of news and blog articles. In *Proceedings of the 9th ACM Conference on Recommender Systems*, RecSys '15, pp. 195–202, New York, NY, USA, 2015. Association for Computing Machinery.

Belgioioso, G., Liao-McPherson, D., de Badyn, M. H., Bolognani, S., Lygeros, J., and Dörfler, F. Sampled-data online feedback equilibrium seeking: Stability and tracking. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 2702–2708. IEEE, 2021.

Breneis, S. On variation functions and their moduli of continuity. *Journal of Mathematical Analysis and Applications*, 491(2):124349, 2020. ISSN 0022-247X.

Cothren, L., Bianchin, G., Dean, S., and Dall'Anese, E. Perception-based sampled-data optimization of dynamical systems. *IFAC-PapersOnLine*, 56(2):5083–5088, 2023. 22nd IFAC World Congress.

Dean, S. and Recht, B. Certainty equivalent perception-based control. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144 of *Proceedings of Machine Learning Research*, pp. 399–411. PMLR, 07 – 08 June 2021. URL https://proceedings.mlr.press/v144/dean21a.html.

Eirinaki, M., Louta, M. D., and Varlamis, I. A trust-aware system for personalized user recommendations in social networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(4):409–421, 2014.

Friedkin, N. E. and Johnsen, E. C. Social influence and opinions. *The Journal of Mathematical Sociology*, 15(3-4):193–206, 1990. doi: 10.1080/0022250X.1990.9990069.

Hauswirth, A., Bolognani, S., Hug, G., and Dörfler, F. Timescale separation in autonomous optimization. *IEEE Transactions on Automatic Control*, 66(2):611–624, 2021.

Hauswirth, A., He, Z., Bolognani, S., Hug, G., and Dörfler, F. Optimization algorithms as robust feedback controllers. *Annual Reviews in Control*, 57, 2024. ISSN 1367-5788.

J. Reddi, S., Sra, S., Poczos, B., and Smola, A. J. Proximal stochastic methods for nonsmooth nonconvex finite-sum optimization. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Jazwinski, A. H. Stochastic processes and filtering theory. *Mathematics in Science and Engineering*, 64, 1970.

Lazer, D. The rise of the social algorithm. *Science*, 348 (6239):1090–1091, 2015.

Marchi, M., Gharesifard, B., and Tabuada, P. Training deep residual networks for uniform approximation guarantees. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144 of *Proceedings of Machine Learning Research*, pp. 677–688. PMLR, 07 – 08 June 2021.

Nesterov, Y. *Introductory Lectures on Convex Optimization: A Basic Course*. Springer Publishing Company, Incorporated, first edition, 2014.

Picallo, M., Ortmann, L., Bolognani, S., and Dörfler, F. Adaptive real-time grid operation via online feedback optimization with sensitivity estimation. *Electric Power Systems Research*, 212:108405, 2022.

Reddi, S. J., Sra, S., Poczos, B., and Smola, A. Fast stochastic methods for nonsmooth nonconvex optimization. *arXiv preprint arXiv:1605.06900*, 2016.

Rossi, W. S., Polderman, J. W., and Frasca, P. The closed loop between opinion formation and personalized recommendations. *IEEE Transactions on Control of Network Systems*, 9(3):1092–1103, 2022.

Scheinberg, K. Finite difference gradient approximation: To randomize or not? *INFORMS Journal on Computing*, 34:2384–2388, 2022.

Tabuada, P. and Gharesifard, B. Universal approximation power of deep residual neural networks through the lens of control. *IEEE Transactions on Automatic Control*, 68 (5):2715–2728, 2023. doi: 10.1109/TAC.2022.3190051.

Willems, J. C., Rapisarda, P., Markovsky, I., and De Moor, B. L. A note on persistency of excitation. *Systems & Control Letters*, 54(4):325–329, 2005.

A. Appendix

A.1. Notations & Preliminaries

We let the symbols $\mathbb{R}(\mathbb{R}_+)$, $\mathbb{N}_0$ denote the set of (positive) real numbers and non-negative integers, respectively. The set of integers $\{1, 2, \ldots, n\}$ is denoted by $[n]$. For a vector $y \in \mathbb{R}^n$, we let the symbol $|y|$ denote the vector whose i-th entry, $|y|_i$, is the i-th component of y in modulus, $|y_i|$. The symbol $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ denotes the standard inner product on $\mathbb{R}^n$. The symbol $1_n, (0_n)$, denotes the all ones (zeros) vector of size n. The symbols I_n and O_n, denote the n-dimensional identity and zero matrix, respectively. Given a matrix $A \in \mathbb{R}^{m\times n}$, the matrix is *positive* (*non-negative*) if all its elements are greater (or equal) than zero and we denote it as $A > 0$ ($A \geq 0$). We let $\mathrm{vec}(A)$ denote the vectorized version of A, i.e., if $A = [a_1\ a_2\ \ldots\ a_n]$, with $a_j \in \mathbb{R}^m, j \in [n]$, then $\mathrm{vec}(A) = [a_1^\top a_2^\top \ldots a_n^\top]^\top$. Let matrix $A \in \mathbb{R}^{n\times n}$, we denote by $\det[A]$ its determinant; A is row-stochastic if $A \geq 0$ and $A1_n = 1_n$. The symbol $\mathcal{D}_\mathcal{U} \subseteq \mathbb{R}$ denotes the set of n-dimensional diagonal matrix with entries in $\mathcal{U}$. The symbol $\mathrm{diag}[x]$ denotes the diagonal matrix with x_i on its i-th diagonal entry. Given $\mathcal{C} \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$, we let $\mathbb{P}_\mathcal{C}[x]$ denote the euclidean projection of x over $\mathcal{C}$. The normal distribution is denoted by $\mathcal{N}(\mu, \Sigma)$, with μ, Σ representing the mean and variance, and the uniform distribution in the interval $[a, b]$, with $a, b, \in \mathbb{R}$ is represented by $\mathcal{U}[a, b]$.

We now state some preliminary lemmas and definitions that are recurrently used in this work.

Definition 1 (L-Lipschitzness (Nesterov, 2014))**.** A function $h : \mathbb{R}^n \to \mathbb{R}^n$ is *L-Lipschitz* if for all $x_1, x_2 \in \mathbb{R}^n$, it holds $\|h(x_1) - h(x_2)\| \leq L\|x_1 - x_2\|$.

Lemma 4 (Gradient boundedness (Nesterov, 2014))**.** *Consider a continuously differentiable, L-Lipschitz map $h : \mathbb{R}^n \to \mathbb{R}$, then the Lipschitz constant provides a bound for the 2-norm of its gradient, i.e. $\|\nabla h\| \leq L$.*

Definition 2 (Global β-smoothness (Nesterov, 2014))**.** A continuously differentiable map $\Phi : \mathbb{R}^n \to \mathbb{R}$ is globally β-smooth if its gradient $\nabla\Phi$ is β-Lipschitz.

Lemma 5 (Properties of β-smooth functions (Nesterov, 2014))**.** *Given a globally β-smooth map $\Phi : \mathbb{R}^n \to \mathbb{R}$, then $\Phi(x_1) - \Phi(x_2) - \nabla\Phi^\top(x_2)(x_1 - x_2) \leq \frac{1}{2}\beta\|x_1 - x_2\|^2$. Moreover, if Φ is twice continuously-differentiable, then $\|\nabla^2\Phi\| \leq \beta$.*

Definition 3 (Persistently exciting input (Willems et al., 2005))**.** Given an input/output system, an input $u \in \mathbb{R}^n$ is *persistently exciting* for the system if the corresponding Henkel matrix has full rank, i.e., there exists $S \in \mathbb{N}$ such that

$$\mathrm{Rank}[\Delta u^{t-S-1,t-S-2}, \Delta u^{t-S,t-S-1}, \ldots, \Delta u^{t,t-1}] = n,$$

where $\Delta u^{t_1,t_2} := u^{t_1} - u^{t_2}$.

Definition 4 (Minimal modulus of continuity(Breneis, 2020))**.** Let $h : \mathcal{X} \to \mathbb{R}$ be a continuous function over a bounded set $\mathcal{X}$. The *minimal modulus of continuity* of h on $\mathcal{X}$ is defined as

$$\omega_h(\gamma) := \sup\Big\{|h(x) - h(y)| : x, y \in \mathcal{X}, \|x - y\|_\infty \leq \gamma\Big\}.$$

A.2. Opinion dynamics example: Friedkin-Johnsen model

The opinions in a Friedkin-Johnsen model evolve according to

$$x^{k+1} = (I_n - \Gamma_p - \Gamma_d)Ax^k + \Gamma_p p^k + \Gamma_d d, \tag{14}$$

where Γ_d, Γ_p are positive diagonal matrices such that $\Gamma_p + \Gamma_d \preceq I_n$, and describe the impact of d and p over the opinions, and A is a row-stochastic adjacency matrix encoding the social interconnections of the users. The dynamics (14) is forward-invariant in $[-1, 1]^n$, since the opinions are a convex combination of $x, p, d \in [-1, 1]^n$.

Further, the steady-state mapping is single-valued, affine (hence, continuously differentiable) and reads as

$$h(p, d) = (I_n - (I_n - \Gamma_p - \Gamma_d)A)^{-1}(\Gamma_p p + \Gamma_d d).$$

Finally, we note that by taking $d \equiv x^0$ and $\Gamma_p = O_n$, the dynamics (14) boil down to the standard FJ model (Friedkin & Johnsen, 1990). △

A.3. Clicking behaviour example: Extremity confirmation bias

A user $i \in [n]$, holding opinion x_i, affected by *extremity confirmation bias* (Rossi et al., 2022) clicks on a recommendation p_i with probability

$$c_i \sim \mathcal{B}\left(\frac{1}{2} + \frac{1}{2}x_i p_i\right),$$

where the clicking behaviour $g_i(p_i, x_i) = \frac{1}{2} + \frac{1}{2}x_i p_i$ models confirmation bias towards extreme recommendations. In fact, if the opinion $x_i \approx \pm 1$ and the position $p_i \approx \pm 1$ (∓ 1), then the probability of clicking is almost 1 (0). The clicking probability becomes random (0.5) as the user position approaches the neutral stance $p_i \approx 0$, highlighting diminished engagement for less polarized content or when recommendations directly counter the user's stance on the issue.

A.4. Acquiring training data

The training of the neural network is carried out offline via feed-forward and back-propagation. Algorithm 2, provides the pseudo-code to acquire the training data.

Algorithm 2 $[\mathcal{X}_p, \mathcal{X}_x, \mathcal{X}_y]$ = Training$[N, T, m]$

```
for j = 1 to m do
    p̄ ~ U[−1, 1]^n
    x^{k+1} = f(x^k, p̄, d^k), k ∈ {0, 1, ..., N − 1},
    c_i^k = B(g_i(p̄_i, x_i^k)), ∀i ∈ [n], k ∈ {0, 1, ..., N − 1}
    Obtain final opinion from users x^N
    Clicking ratio y^N = 1/(T+1) Σ_{k=N−T}^{N} c^k
    Positions set X_p ← append[p̄]
    Steady-state opinions set X_x ← append[x^N]
    Clicking ratio set X_y ← append[y^N]
end for
```

For each training sample $k \in [m]$, we provide news expressing a fixed random position $\overline{p}$ to the users for a period of N time steps. The term T is a design parameter, with $(N - T)$ representing a time instant at which opinions have reached a steady-state. After N steps, we obtain the users' opinion x^N and compute their clicking ratio for the position $\overline{p}$ based on the last T time steps. This process is carried out over m Monte Carlo trials, thus collecting m training data points.

A.5. Proof of Lemma 1

The proof follows from (Corollary 5.2, (Tabuada & Gharesifard, 2023)) with the continuous function β playing the same role as f in (Tabuada & Gharesifard, 2023).

In order to provide an explicit upper-bound for e_x we point out that the neural network makes use of the injection layer, whose map is represented by $u : \mathbb{R}^{n+1} \to \mathbb{R}^{n+2}$, the hidden layer whose map is represented by $z : \mathbb{R}^{n+2} \to \mathbb{R}^{n+2}$ and the output layer, whose map is represented by $v : \mathbb{R}^{n+2} \to \mathbb{R}$, with $v(z) = v^\top z + v_0$. Given that the input and output layers are linear maps, we can directly make use of (Theorem 7, (Marchi et al., 2021)) to state the following upper-bound $\forall i \in [n]$:

$$|\beta_i(X) - \hat{\beta}_i(X)| \leq 3 \sup_{X \in \mathcal{X}} |\beta_i(X) - \hat{\beta}_i(X)| + 2\omega_{\beta_i}(\gamma_x) + |v_{0,i}|\gamma_x,$$

where $\sup_{X \in \mathcal{X}} |\beta_i(X) - \hat{\beta}_i(X)|$ refers to the maximum training error on each user i, $\omega_{\beta_i}(\gamma_x)$ refers to the minimum modulus of continuity of β_i on $\mathcal{X}$, $v_{0,i}$ represents the estimated bias weight v_0 from the hidden layer to the output layer of the neural network for user i. By replacing the modelling error θ_x of β from Assumption 3, it is now possible to state the following:

$$\|e_x\| \leq \sqrt{n}\Big[3 \sup_{X \in \mathcal{X}} \|\beta(X) - \hat{\beta}(X)\|_\infty + 2 \sup_{i \in [n]} \omega_{\beta_i}(\gamma_x) + \gamma_x \sup_{i \in [n]} |v_{0,i}|\Big] + \theta_x.$$

A.6. Proof of Lemma 2

Before making use of (Corollary 5.2, (Tabuada & Gharesifard, 2023)) to prove the upper-bound on e_y, it is to be noted that the arguments of g and $\hat{g}$ in the definition of e_y, are different. Thus, we re-write e_y as $e_y = \hat{g}(p, \hat{x}) - g(p, \hat{x}) + g(p, \hat{x}) - g(p, h(p, d))$ and note the following:

$$\begin{aligned}\|e_y\| &\underset{(a)}{\leq} \|g(p, h(p, d) + e_x) - g(p, h(p, d))\| + \\ &\quad \|\hat{g}(p, \hat{x}) - g(p, \hat{x})\| \\ &\underset{(b)}{\leq} \|\hat{g}(p, \hat{x}) - g(p, \hat{x})\| + \|\nabla_x g^\top(p, h(p, d)) e_x\| + \alpha_y \\ &\underset{(c)}{\leq} \|\hat{g}(p, \hat{x}) - g(p, \hat{x})\| + M_x \epsilon_x + \alpha_y. \end{aligned} \tag{15}$$

In (a), we make use of the identity $\hat{x} = x + e_x$, with the state estimation error e_x defined in Lemma 1. In (b), we use the Taylor series expansion on $g(p, x + e_x)$ around x. In addition, the higher-order terms in the expansion of $g(p, x + e_x)$ are upper-bounded by the modelling error on the clicking behaviour, i.e. $\|\mathcal{O}(g(p, x))\| \leq \alpha_y$. In (c), we use Assumption 2, i.e. $g(p, x)$ is M_x-Lipshitz with respect to x and the fact that the opinion estimation error is upper-bounded with $\|e_x\| \leq \epsilon_x$ from Lemma 1.

The proof for the existence of an upper-bound on $\|\hat{g}(p, \hat{x}) - g(p, \hat{x})\|$ is similar to the proof in Appendix A.5. The bound can now be stated as follows, for every $i \in [n]$:

$$|g_i(Y) - \hat{g}_i(Y)| \leq 3 \sup_{Y \in \mathcal{Y}} |g_i(Y) - \hat{g}_i(Y)| + 2\omega_{g_i}(\gamma_y) + |w_{0,i}|\gamma_y, \tag{16}$$

where $\sup_{Y \in \mathcal{Y}} |g_i(Y) - \hat{g}_i(Y)|$ refers to the maximum training error on each user i, $\omega_{g_i}(\gamma_y)$ refers to the modulus of continuity of g_i on $\mathcal{Y}$, $w_{0,i}$ represents the estimated bias weight

w_0 from the hidden layer to the output layer of the neural network for user i. Using (16) in (15), it is now possible to state the following:

$$\|e_y\| \le \sqrt{n}\Big[3\sup_{Y\in\mathcal{Y}}\|g(Y)-\hat{g}(Y)\|_\infty + 2\sup_{i\in[n]}\omega_{g_i}(\gamma_y)+ \gamma_y \sup_{i\in[n]}|w_{0,i}|\Big] + M_x\epsilon_x + \alpha_y.$$

A.7. Proof of Lemma 3

Due to Assumption 2 and Lemma 5, for any $p, x_1, x_2 \in \mathbb{R}^n$ it holds that

$$\varphi^{\mathrm{clk}}(p,x_1)-\varphi^{\mathrm{clk}}(p,x_2)-(\nabla_x\varphi^{\mathrm{clk}})^\top(p,x_2)(x_1-x_2) \le \frac{1}{2}L_x\|x_1-x_2\|^2.$$

In the above equation, we replace x_1, x_2 with $x+\mu e_i$ and x, respectively, so that

$$\varphi^{\mathrm{clk}}(p,x+\mu e_i)-\varphi^{\mathrm{clk}}(p,x)-\mu(\nabla_x\varphi^{\mathrm{clk}})^\top(p,x)e_i \le \frac{1}{2}L_x\mu^2. \tag{17}$$

We now add and subtract $\hat{\varphi}^{\mathrm{clk}}(p,\hat{x}+\mu e_i)$ and $\hat{\varphi}^{\mathrm{clk}}(p,\hat{x})$ in the above equation and we note that $\hat{\varphi}^{\mathrm{clk}}(p,\hat{x}) := -1_n^\top\hat{g}(p,\hat{x})$ and $\varphi^{\mathrm{clk}}(p,x) := -1_n^\top g(p,x)$. Using these definitions in (17), we obtain:

$$\hat{\varphi}^{\mathrm{clk}}(p,\hat{x}+\mu e_i)-\hat{\varphi}^{\mathrm{clk}}(p,\hat{x}) \le 1_n^\top[e_y(p,x)+e_y(p,x+\mu e_i)] + \mu(\nabla_x\varphi^{\mathrm{clk}}(p,x))^\top e_i + \frac{1}{2}L_x\mu^2, \tag{18}$$

where $e_y(p,x)$ is the clicking behaviour estimation error as defined in (7). Dividing both sides by μ in (18) and using the gradient estimate definition in (9), we obtain

$$\Big(\nabla_x\hat{\varphi}^{\mathrm{clk}}\Big)_i(p,\hat{x}) - \Big(\nabla_x\varphi^{\mathrm{clk}}\Big)_i(p,x) \le \frac{L_x\mu}{2} + \frac{1}{\mu}1_n^\top(e_y(p,x)+e_y(p,x+\mu e_i)).$$

for all $i\in[n]$. Taking the modulus on both sides, we obtain $|1_n^\top(e_y(p,x)+e_y(p,x+\mu e_i))| \le \sqrt{n}\epsilon_g$ using the Cauchy-Schwartz inequality and the upper-bound on the clicking behaviour estimation error $\|e_y\|\le\epsilon_g$ as in (7). We now write

$$\Big|\Big(\nabla_x\hat{\varphi}^{\mathrm{clk}}\Big)_i(p,\hat{x}) - \Big(\nabla_x\varphi^{\mathrm{clk}}\Big)_i(p,x)\Big| \le \frac{1}{2}L_x\mu + 2\frac{\sqrt{n}\epsilon_g}{\mu},$$

from which (9) follows. Analogous reasoning is followed for the gradient estimation error with respect to p. To obtain the tight upper-bound, we use first-order optimality conditions with respect to μ on the term $\frac{1}{2}L_x\mu + 2\frac{\sqrt{n}\epsilon_g}{\mu}$, thus obtaining $\mu^* = 2n^{1/4}\sqrt{\epsilon_g/L_x}$. Since the second-order derivative of the term $\frac{1}{2}L_x\mu + 2\frac{\sqrt{n}\epsilon_g}{\mu}$ is strictly positive, μ^* is the smoothing parameter that provides the lowest upper-bound on the gradient estimation error.

A.8. Proof of Theorem 1

The Kalman filter is uniformly asymptotically stable provided the pairs $(I_{n^2}, \sqrt{Q^k}), (I_{n^2}, \Delta\tilde{p}^{k,\tau_i})$ are uniformly completely controllable and observable, respectively (Theorem 7.4, (Jazwinski, 1970)). Since Q^k is a design parameter used to control the degree of trust in the process model, it is possible to make Q^k positive definite for all $k\in\mathbb{N}_0$. Further, persistently exciting inputs Δp that satisfy the Henkel matrix condition in Definition 3 guarantees uniform complete observability (see (Picallo et al., 2022)).

We can now derive an explicit analytical expression for the upper-bound on the variance $\mathbb{E}[\|e^k\|^2]$. By making use of (4.2) and (8) on ℓ^k and $\hat{\ell}^k$, by adding and subtracting $\Delta x_{\mathrm{ss}}^{k+1,\tau_i+1}$and by making use of (4.2) and (4.2) on $\Delta x_{\mathrm{ss}}^{k+1,\tau_i+1}$, one gets

$$e^k = \big(I_{n^2}-\zeta^kK^{k-1}\Delta\tilde{p}^{k,\tau_i}\big)(e^{k-1}+w^{k-1}) - \zeta^kK^{k-1}v^k + \zeta^kK^{k-1}\Delta e_x^{\tau_i+1,k+1}. \tag{19}$$

Taking the expectation of the norm squared on both sides in (19), we obtain:

$$\begin{aligned}\mathbb{E}[\|e^k\|^2] &\underset{(a)}{=} \mathbb{E}[\|(I_{n^2}-\zeta^kK^{k-1}\Delta\tilde{p}^{k,\tau_i})(e^{k-1}+w^{k-1})- \zeta^kK^{k-1}\big(v^k+\Delta e_x^{k+1,\tau_i+1}\big)\|^2]\\ &\underset{(b)}{\le} \|I_{n^2}-\zeta^kK^{k-1}\Delta\tilde{p}^{k,\tau_i}\|^2(\mathbb{E}[\|e^{k-1}\|^2]+(\sigma_q^{k-1})^2) +\|\zeta^kK^{k-1}\|^2\big((\sigma_r^k)^2+2\epsilon_x^2\big)\end{aligned}$$

In (b), we expand the norm and use Assumption 4 to state that the expectation on the cross-coupled terms are all zero. We then use $\mathbb{E}[\|w^{k-1}\|^2] = (\sigma_q^{k-1})^2$ and $\mathbb{E}[\|v^k\|^2] = (\sigma_r^k)^2$. Further, Lemma 1 allows us to state that $\mathbb{E}[\|\Delta e_x^{k+1,\tau_i+1}\|^2] \le 2\epsilon_x^2$. Considering a trigger at time k, we have:

$$\mathbb{E}[\|e^{\tau_{i+1}}\|^2] \le \|I_{n^2}-K^{\tau_{i+1}-1}\Delta\tilde{p}^{\tau_{i+1},\tau_i}\|^2\mathbb{E}[\|e^{\tau_i}\|^2] + (T-1)\overline{\sigma}_q^2 + \|K^{\tau_{i+1}-1}\|^2\big(\overline{\sigma}_r^2+2\epsilon_x^2\big), \tag{20}$$

where $\overline{\sigma}_q^2 = \sup_{t\in\mathbb{N}_0}(\sigma_q^t)^2$ and $\overline{\sigma}_r^2 = \sup_{t\in\mathbb{N}_0}(\sigma_r^t)^2$. Tracing back (20) recursively to $k=0$, we obtain the following relation:

$$\mathbb{E}[\|e^{\tau_{i+1}}\|^2] \le (c_1\xi^{c_2T})^{2|\mathcal{T}|}\mathbb{E}[\|e^0\|^2] + \frac{1-(c_1\xi^{c_2T})^{2|\mathcal{T}|}}{1-c_1\xi^{c_2T}}\big[(T-1)\overline{\sigma}_q^2 + K_m^2\big(\overline{\sigma}_r^2+2\epsilon_x^2\big)\big],$$

We now have the following asymptotic result on the variance:

$$\lim_{|\mathcal{T}|\to\infty}\mathbb{E}[\|e^{\tau_i}\|^2] \le \frac{1}{1-c_1\xi^{c_2T}}\big[(T-1)\overline{\sigma}_q^2+K_m^2\big(\overline{\sigma}_r^2+2\epsilon_x^2\big)\big],$$

A.9. Proof of Theorem 2

Before describing the proof, we state the following supporting lemma and remark.

Lemma 6 (Projections with smooth functions (Reddi et al., 2016)). *Let $y = \mathbb{P}_{[-1,1]^n}[x - \eta u]$ with $y, x, u \in \mathbb{R}^n$. Then, the following inequality holds:*

$$\varphi(y) \leq \varphi(z) + \langle y - z, \Phi(x) - u\rangle + \Big[\frac{L'}{2} - \frac{1}{2\eta}\Big]\|y - x\|^2 + \Big[\frac{L'}{2} + \frac{1}{2\eta}\Big]\|z - x\|^2 - \frac{1}{2\eta}\|y - z\|^2, \quad \forall z \in \mathbb{R}^n$$

where φ is the cost function to be minimized and Φ its gradient. Further, L' and η are the smoothness factor of φ and the step-size of the gradient descent algorithm, respectively.

Proof. The proof is given in (Lemma 2,(Reddi et al., 2016)). □

Remark 1 (Composite gradient lipschitzness). We observe that the Lipschitz and smoothness constant for $\varphi^{\text{pol}}(x)$ are $2\sqrt{n}$ and 2, respectively. Using Assumptions 1(iii), 2, the composite gradient $\Phi(p)$ is thus $L'-$Lipschitz with respect to p, where $L' = L_p + L^2(L_x + 2)$. Therefore, the composite cost function $\varphi(p, h(p, d))$ is $L'-$smooth with respect to p.

We now state the following gradient update in the case where the gradients, sensitivity and opinions are known:

$$\overline{p}^{k+1} = \mathbb{P}_{[-1,1]^n}[p^k - \eta\Phi(p^k, h(p^k, d))],$$

where $\Phi(p, h(p, d))$ is the composite gradient. The above will serve as a benchmark to investigate the stationarity of our algorithm. We are now in the position to prove the inequalities in (13). To do so, we use Lemma 6 with $y = \overline{p}^{k+1}$, $x = p^k$ and $u = \Phi(p^k, h(p^k, d))$ is the composite gradient. Choosing $z = p^k$ and taking the expectation on both sides, we obtain:

$$\mathbb{E}\Big[\varphi(\overline{p}^{k+1})\Big] \leq \mathbb{E}\Big[\varphi(p^k) + \Big(\frac{L'}{2} - \frac{1}{\eta}\Big)\|\overline{p}^{k+1} - p^k\|^2\Big]. \quad (21)$$

We now define the update step with our algorithm:

$$p^{k+1} = \mathbb{P}_{[-1,1]^n}[p^k - \eta\zeta^k\hat{\Phi}^k].$$

We use Lemma 6 with $y = p^{k+1}$, $x = p^k$ and $u = \zeta^k\hat{\Phi}^k$. Choosing $z = \overline{p}^{k+1}$ and taking the expectation on both sides of the inequality, we obtain:

$$\begin{aligned}\mathbb{E}\Big[\varphi(p^{k+1})\Big] \leq \mathbb{E}\Big[&\varphi(\overline{p}^{k+1}) + \Big(\frac{L'}{2} - \frac{1}{2\eta}\Big)\|p^{k+1} - p^k\|^2 + \\ &\langle p^{k+1} - \overline{p}^{k+1}, \Phi(p^k, h(p^k, d)) - \zeta^k\hat{\Phi}^k\rangle + \quad (22)\\ &\Big(\frac{L'}{2} + \frac{1}{2\eta}\Big)\|\overline{p}^{k+1} - p^k\|^2 - \frac{1}{2\eta}\|p^{k+1} - \overline{p}^{k+1}\|^2\Big]\end{aligned}$$

We now add inequalities (21) and (22) to obtain:

$$\begin{aligned}\mathbb{E}\Big[\varphi(p^{k+1})\Big] \leq \mathbb{E}\Big[&\varphi(p^k) + \Big(\frac{L'}{2} - \frac{1}{2\eta}\Big)\|p^{k+1} - p^k\|^2 + \\ &\underbrace{\langle p^{k+1} - \overline{p}^{k+1}, \Phi(p^k, h(p^k, d)) - \zeta^k\hat{\Phi}^k\rangle}_{T_1} + \quad (23)\\ &\Big(L' - \frac{1}{2\eta}\Big)\|\overline{p}^{k+1} - p^k\|^2 - \frac{1}{2\eta}\|p^{k+1} - \overline{p}^{k+1}\|^2\Big].\end{aligned}$$

We now focus on the term T_1. Using Cauchy-Schwartz relation and the fact that the geometric mean of two non-negative real numbers is always less than its arithmetic mean, we obtain the following: $T_1 \leq \|p^{k+1} - \overline{p}^{k+1}\|\;\|\Phi(p^k, h(p^k, d)) - \zeta^k\hat{\Phi}^k\| \leq \frac{1}{2\eta}\|p^{k+1} - \overline{p}^{k+1}\|^2 + \frac{\eta}{2}\|\Phi(p^k, h(p^k, d)) - \zeta^k\hat{\Phi}^k\|^2$. Using the above inequality in (23), we obtain:

$$\begin{aligned}\mathbb{E}\Big[\varphi(p^{k+1})\Big] \leq \mathbb{E}\Big[&\varphi(p^k) + \underbrace{\Big(\frac{L'}{2} - \frac{1}{2\eta}\Big)\|p^{k+1} - p^k\|^2}_{T_2} + \quad (24)\\ &\frac{\eta}{2}\|\Phi(p^k, h(p^k, d)) - \zeta^k\hat{\Phi}^k\|^2 + \Big(L' - \frac{1}{2\eta}\Big)\eta^2\|\mathcal{G}(p^k)\|^2\Big].\end{aligned}$$

In the above inequality, we used the definition of fixed-point residual mapping according to (12).

We now assume that there is a trigger at time instant k, i.e. $\zeta^k = 1$. To obtain a feasible upper-bound on $\mathbb{E}[\|\mathcal{G}(p^k)\|^2]$, we need $L' - 1/2\eta < 0$. Thus, the step-size is constrained with $\eta \in (0, \frac{1}{2L'})$. With this constraint, we have $T_2 \leq 0$. This leads to the following:

$$\begin{aligned}\mathbb{E}\Big[\|\mathcal{G}(p^k)\|^2\Big] \leq \frac{2}{\eta(1 - 2\eta L')}\Big\{&\mathbb{E}\Big[\varphi(p^k)\Big] - \mathbb{E}\Big[\varphi(p^{k+1})\Big] + \\ &\frac{\eta}{2}\underbrace{\mathbb{E}\Big[\|\Phi(p^k, h(p^k, d)) - \hat{\Phi}^k\|^2\Big]}_{T_3}\Big\}. \quad (25)\end{aligned}$$

We now analyze the term T_3. For the sake of convenience, we drop the arguments of the gradient p, x and time argument k in the gradient terms. Thus, we denote $\nabla_p\varphi^{\text{clk}} = \nabla_p\varphi^{\text{clk}}(p^k, h(p^k, d))$, $\nabla_x\varphi^{\text{clk}} = \nabla_x\varphi^{\text{clk}}(p^k, h(p^k, d))$, $\nabla_p\hat{\varphi}^{\text{clk}} = \nabla_p\varphi^{\text{clk}}(p^k, \hat{x}^{k+1})$ and $\nabla_x\hat{\varphi}^{\text{clk}} = \nabla_x\varphi^{\text{clk}}(p^k, \hat{x}^{k+1})$. Using the definitions of $\Phi(p^k, h(p^k, d))$, $\hat{\Phi}^k$, we have:

$$\begin{aligned}T_3 = \mathbb{E}\Big[\|&\nabla_p\varphi^{\text{clk}} - \nabla_p\hat{\varphi}^{\text{clk}} + H^{k\top}\nabla_x\varphi^{\text{clk}} - \hat{H}^{k\top}\nabla_x\hat{\varphi}^{\text{clk}} + \quad (26)\\ &\gamma H^{k\top}\nabla_x\varphi^{\text{pol}}(h(p^k, d)) - \gamma\hat{H}^{k\top}\nabla_x\varphi^{\text{pol}}(\hat{x}^{k+1}) + \frac{w^k_{\text{pe}}}{\eta}\|^2\Big],\end{aligned}$$

where $H^k = \nabla_p h(p^k, d)$ is the true sensitivity and $\hat{H}^k$ is its estimate at time k. We now analyze the upper-bound on

T_3:

$$
\begin{aligned}
T_3 &\overset{(a)}{\leq} 6\Big\{\|\nabla_p \varphi^{\text{clk}} - \nabla_p \hat{\varphi}^{\text{clk}}\|^2 + \|(H^k)^\top(\nabla_x \varphi^{\text{clk}} - \nabla_x \hat{\varphi}^{\text{clk}})\|^2 + \\
&\quad \frac{\sigma_{\text{pe}}^2}{\eta^2} + \gamma^2\|(H^k)^\top(\nabla_x \varphi^{\text{pol}}(h(p^k,d)) - \nabla_x \varphi^{\text{pol}}(\hat{x}^{k+1}))\|^2 + \\
&\quad (\gamma^2\|\nabla_x \varphi^{\text{pol}}(\hat{x}^{k+1})\|^2 + \|\nabla_x \hat{\varphi}^{\text{clk}}\|^2)\mathbb{E}\Big[\|H^k - \hat{H}^k\|^2\Big]\Big\} \\
&\overset{(b)}{\leq} 6\Big\{4n^{3/2}\epsilon_g(L_p + L^2 L_x) + \frac{\sigma_{\text{pe}}^2}{\eta^2} + \\
&\quad \gamma^2\|(H^k)^\top(\nabla_x \varphi^{\text{pol}}(h(p^k,d)) - \nabla_x \varphi^{\text{pol}}(\hat{x}^{k+1}))\|^2 + \\
&\quad (\gamma^2\|\nabla_x \varphi^{\text{pol}}(\hat{x}^{k+1})\|^2 + \|\nabla_x \hat{\varphi}^{\text{clk}}\|^2)\mathbb{E}\Big[\|H^k - \hat{H}^k\|^2\Big]\Big\} \\
&\overset{(c)}{\leq} 6\Big\{4n^{3/2}\epsilon_g(L_p + L^2 L_x) + \frac{\sigma_{\text{pe}}^2}{\eta^2} + 4\gamma^2 L^2\|h(p^k,d) - \hat{x}^{k+1}\|^2 + \\
&\quad (\gamma^2\|\nabla_x \varphi^{\text{pol}}(\hat{x}^{k+1})\|^2 + \|\nabla_x \hat{\varphi}^{\text{clk}}\|^2)\mathbb{E}\Big[\|H^k - \hat{H}^k\|^2\Big]\Big\} \\
&\overset{(d)}{\leq} 6\Big\{4n^{3/2}\epsilon_g(L_p + L^2 L_x) + \frac{\sigma_{\text{pe}}^2}{\eta^2} + 4\gamma^2 L^2 \epsilon_x^2 + \\
&\quad (\gamma^2\|\nabla_x \varphi^{\text{pol}}(\hat{x}^{k+1})\|^2 + \|\nabla_x \hat{\varphi}^{\text{clk}}\|^2)\mathbb{E}\Big[\|H^k - \hat{H}^k\|^2\Big]\Big\} \\
&\overset{(e)}{\leq} 6\Big\{4n^{3/2}\epsilon_g(L_p + L^2 L_x) + \frac{\sigma_{\text{pe}}^2}{\eta^2} + 4\gamma^2 L^2 \epsilon_x^2 + \\
&\quad (4n\gamma^2 + \|\nabla_x \hat{\varphi}^{\text{clk}} \pm \nabla_x \varphi^{\text{clk}}\|^2)\mathbb{E}\Big[\|e^k\|^2\Big]\Big\} \\
&\overset{(f)}{\leq} 6\Big\{4n^{3/2}\epsilon_g(L_p + L^2 L_x) + \frac{\sigma_{\text{pe}}^2}{\eta^2} + 4\gamma^2 L^2 \epsilon_x^2 + \\
&\quad (4n\gamma^2 + 2(M_x^2 + 4n^{3/2}\epsilon_g L_x))\mathbb{E}\Big[\|e^k\|^2\Big]\Big\}
\end{aligned} \tag{27}
$$

To obtain (a), we added and subtracted $(H^k)^\top \nabla_x \hat{\varphi}^{\text{clk}}$ and $\gamma(H^k)^\top \nabla_x \varphi^{\text{pol}}(\hat{x}^{k+1})$ inside the norm in (26). We then used the fact that $\mathbb{E}\Big[\|\sum_{j=1}^m r_j\|^2\Big] \leq m\sum_{j=1}^m \mathbb{E}\Big[\|r_j\|^2\Big]$. We also make use of the fact that $\mathbb{E}[\|w_{\text{pe}}^k\|^2] = \sigma_{\text{pe}}^2$. In (b), we made use of Lemma 3 for the gradient estimate accuracy on φ^{clk} and Assumption 1(iii) for Lipschitz condition on $h(p,d)$ to arrive at the term $4n^{3/2}\epsilon_g(L_p + L^2 L_x)$. In (c), we made use of Remark 1 for the smoothness condition on φ^{pol} and Assumption 1(iii) for Lipschitz condition on $h(p,d)$ to arrive at the term $4L^2\|x - \hat{x}\|^2$. In (d), we made use of the fact that the norm of the steady-state opinion estimation error is upper-bounded by ϵ_x. In (e), we made use of Remark 1 for the Lipschitz condition on φ^{pol}, thus arriving at the term $4n$. We also add and subtract the term $\nabla_x \varphi^{\text{clk}}$. Further, we use inequality $\|H^k - \hat{H}^k\| \leq \|H^k - \hat{H}^k\|_F = \|e^k\|$, where $\|\cdot\|_F$ refers to the Frobenius norm and e^k is the sensitivity estimation error. In (f), we made use of the fact that $\|a+b\|^2 \leq 2(\|a\|^2 + \|b\|^2)$. We then use Assumption 2 to state that $\|\nabla_x \varphi^{\text{clk}}\| \leq M_x$ and Lemma 3 to state the upper-bound on $\|\nabla_x \varphi^{\text{clk}} - \nabla_x \hat{\varphi}^{\text{clk}}\|$.

We now use the inequality (27) in (25) and add the inequalities over the trigger time instances up to k, leading to telescopic cancellation. Thus, we have:

$$
\begin{aligned}
\sum_{\substack{l\in\mathcal{T}\\ l\leq k}} \mathbb{E}[\|\mathcal{G}(p^l)\|^2] \leq \frac{2}{\eta(1-2\eta L')}\Big\{\mathbb{E}\Big[\varphi(p^0)\Big] - \mathbb{E}\Big[\varphi(p^{k+1})\Big]\Big\} + \\
\frac{6|\mathcal{T}|}{1-2\eta L'}\Big\{4n^{3/2}\epsilon_g(L_p + L^2 L_x) + \frac{\sigma_{\text{pe}}^2}{\eta^2} + 4\gamma^2 L^2 \epsilon_x^2\Big\} + \\
\frac{12\big[2n\gamma^2 + (M_x^2 + 4n^{3/2}\epsilon_g L_x)\big]}{1-2\eta L'} \sum_{\substack{l\in\mathcal{T}\\ l\leq k}} \mathbb{E}\Big[\|e^l\|^2\Big].
\end{aligned}
$$

Since the positions do not change between two consecutive trigger time instances, it is sufficient to investigate convergence guarantees at the trigger time instances.

Using Theorem 1 for the upper-bound on $\mathbb{E}[\|e^k\|^2]$, the summation of this term over the trigger instances leads to the following:

$$
\begin{aligned}
\sum_{\substack{l\in\mathcal{T}\\ l\leq k}} \mathbb{E}\Big[\|e^l\|^2\Big] &\leq |\mathcal{T}|C_f + \mathbb{E}\Big[\|e^0\|^2\Big] \sum_{\substack{l\in\mathcal{T}\\ l\leq k}} (c_1 \xi^{c_2 T})^{2|\mathcal{T}|} \\
&|\mathcal{T}|C_f + \mathbb{E}\Big[\|e^0\|^2\Big]\Big(\frac{1-(c_1\xi^{c_2 T})^{2|\mathcal{T}|}}{1-(c_1\xi^{c_2 T})^2}\Big).
\end{aligned} \tag{28}
$$

We start the algorithm with $p^0 = 0_n$, thus $\mathbb{E}\Big[\varphi(p^0)\Big] = \varphi(0)$. Further, $\exists\ \varphi^* \leq \mathbb{E}\Big[\varphi(p^k)\Big], \forall k \in \mathbb{N}_0$, i.e. φ^* is a local optimal value. Thus, with the above formulations and (28), we obtain (13).