A Proof of Theorem 2

Given any two function $f(t)$ and $g(t)$, we first define a convolution operator $\star$ as follows

$$f(t) \star g(t) = \int_0^t f(t-s) g(s) \mathrm{d}s \tag{12}$$

Therefore, the user activity model for $x_i(t)$ can be expressed as

$$x_i(t) = \underset{\substack{\uparrow \\ \text{base}}}{b_i(t)} + \underbrace{\sum_{j=1}^{U} \alpha_{ij} \kappa_{\omega_2}(t) \star \big(h(x_j(t)) \mathrm{d}N_j(t)\big)}_{\text{social neighbor influence}} \tag{13}$$

Before we apply the differential operator d to (13), we also need the following two properties:

- $\mathrm{d}\kappa_{\omega_2}(t) = -\omega_2 \kappa_{\omega_2}(t)\mathrm{d}t$ for $t \geqslant 0$ and $\kappa_{\omega_2}(0) = 1$.
- The differential of the convolution of two functions is expressed as: $\mathrm{d}(f \star g) = f(0)g + g \star \mathrm{d}f$.

With the above two properties, we set $f = \kappa_{\omega_2}(t)$ and $g = \sum_j \alpha_{ij} h(x_j) \mathrm{d}N_j(t)$, and take the differential of $x_i(t)$ in (13) as follows

$$\mathrm{d}x_i(t) = \mathrm{d}b_i(t) + \mathrm{d}(f \star g) \tag{14}$$

$$= \mathrm{d}b_i(t) + \sum_{j=1}^{U} \alpha_{ij} h(x_j) \mathrm{d}N_j(t) - \omega_2 \Big(\sum_{j=1}^{U} \alpha_{ij} k_{\omega_2}(t) \star (h(x_j) \cdot \mathrm{d}N_j(t)) \Big) \mathrm{d}t \tag{15}$$

$$= \mathrm{d}b_i(t) + \sum_{j=1}^{U} \alpha_{ij} h(x_j) \mathrm{d}N_j(t) - \omega_2 (x_i(t) - b_i(t)) \mathrm{d}t \tag{16}$$

$$= \mathrm{d}b_i(t) + \omega_2 (b_i(t) - x_i(t)) \mathrm{d}t + \sum_{j=1}^{U} \alpha_{ij} h(x_j(t)) \mathrm{d}N_j(t) \tag{17}$$

This completes the proof for the SDE formulation of $x_i(t)$.

Similarly, we can express the intensity function using the convolution operator as follows

$$\lambda_i(t) = \underset{\substack{\uparrow \\ \text{base}}}{\eta_i(t)} + \underbrace{\sum_{j=1}^{U} \beta_{ij} \kappa_{\omega_1}(t) \star \mathrm{d}N_j(t)}_{\text{social neighbor influence}} \tag{18}$$

Then we set $f = \kappa_{\omega_1}(t)$, $g = \sum_j \beta_{ij} \mathrm{d}N_j(t)$, and can show the following equation:

$$\mathrm{d}\lambda_i(t) = \mathrm{d}\eta_i(t) + \omega_1 \big(\eta_i(t) - \lambda_i(t)\big) \mathrm{d}t + \sum_j \beta_{ij} \mathrm{d}N_j(t) \tag{19}$$

This completes the proof.

B Extensions to time-varying networks

Real world social networks can change over time. Users can follow or unfollow each other as time goes by and new users can join the network (Farajtabar et al., 2015). In this section, we extend our framework to networks with time-varying edges and node birth processes.

First, for a fixed network, the expectation in the objective function in (7) is over the stochastic pair $\{\boldsymbol{w}(t), \boldsymbol{N}(t)\}$ for $t \in (t_0, T]$. Since the network is stochastic now, we also need to take the expectation of the adjacency matrix $\boldsymbol{A}(t) = (\alpha_{ij}(t))$ to derive the HJB equation. Hence the input to Algorithm 1 is $\mathbb{E}[\boldsymbol{A}(t)] = (\mathbb{E}[\alpha_{ij}(t)])$ instead of $\boldsymbol{A}$. Specifically, we replace $\boldsymbol{h}_j(\boldsymbol{x})$ in the HJB equation (10) by $\mathbb{E}[\boldsymbol{h}_j(\boldsymbol{x})]$:

$$\sum_j \lambda_j(t)(V(\boldsymbol{x} + \mathbb{E}[\boldsymbol{h}_j(\boldsymbol{x}, t)], t) - V(\boldsymbol{x}, t)) \tag{20}$$

where $\mathbb{E}[\boldsymbol{h}_j(\boldsymbol{x}, t)] = (\mathbb{E}[h_{1j}(t)], \cdots, \mathbb{E}[h_{Uj}(t)])^\top$ and $\mathbb{E}[h_{ij}(t)] = \mathbb{E}[\alpha_{ij}(t)]x_j(t)$. Next, we compute $\mathbb{E}[\alpha_{ij}(t)]$ in two types of networks.

Networks with link creation. We model the creation of link from node $i \to j$ as a survival process $\alpha_{ij}(t)$. If a link is created, $\alpha_{ij}(t) = 1$ and zero otherwise. Its intensity function is defined as

$$\sigma_{ij}(t) = (1 - \alpha_{ij}(t))\gamma_i, \tag{21}$$

where the term $\gamma_i \geqslant 0$ denotes the Poisson intensity, which models the node i's own initiative to create links to others. The coefficient $1 - \alpha_{ij}(t)$ ensures a link is created only once, and intensity is set to 0 after that. Given a sequence of link creation events, we can learn $\{\gamma_i\}$ using maximum likelihood estimation (Aalen et al., 2008) as follows.

Parameter estimation of the link creation process. Given data $e_i = (t_i, u_i, s_i)$, which means at time t_i node u_i is added to the network and connects to s_i, we set $\mathcal{E} = \{e_i\}$ and optimize the concave log-likelihood function to learn the parameters of the Poisson intensity $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_U)^\top$:

$$\max_{\boldsymbol{\gamma} \geqslant 0} \sum_{e_i \in \mathcal{E}} \log(\sigma_{u_i s_i}(t_i)) - \sum_{u,s \in [n]} \int_0^T \sigma_{us}(\tau) \mathrm{d}\tau$$

This objective function can be solved efficiently with many optimization algorithms, such as the Quasi-Newton algorithm.

Next, given the learned parameters, we obtain the following ordinary differential equation (ODE) that describes the time-evolution of $\mathbb{E}[\alpha_{ij}(t)]$:

$$\mathrm{d}\mathbb{E}[\alpha_{ij}(t)] \underset{(a)}{=} \mathbb{E}[\mathrm{d}\alpha_{ij}(t)] \underset{(b)}{=} \sigma_{ij}(t)\mathrm{d}t \underset{(c)}{=} (1 - \mathbb{E}[\alpha_{ij}(t)])\gamma_i \mathrm{d}t, \tag{22}$$

where (a) holds because the operator d and $\mathbb{E}$ are exchangeable, (b) is from the definition of intensity function, and (c) is from (21). The initial condition is $\mathbb{E}[\alpha_{ij}(0)] = 0$ since i and j are not connected initially. We can easily solve this ODE in analytical form:

$$\mathbb{E}[\alpha_{ij}(t)] = 1 - \exp(-\gamma_i t)$$

Networks with node birth. The network's dimension can grow as new users join it. Since the dimension of $\boldsymbol{A}(t)$ changes over time, it is very challenging to control such network, and it remains unknown how to derive the HJB equation for such case. We propose an efficient method by connecting the stochasticity of the node birth process to that of link creation process. More specifically, we have the following observation.

Observation. *The process of adding a new user v to the existing network $\boldsymbol{A} \in \Re^{(N-1)\times(N-1)}$ and connects to user s is equivalent to link creation process of setting $\boldsymbol{A}(t) \in \Re^{N\times N}$ to be the existing network and letting $\alpha_{vs}(t) = 1$.*

With this observation, we can fix the dimension of $\boldsymbol{A}(t)$ beforehand, and add a link whenever a user joins the network. This procedure is memory-efficient since we do not need to maintain a sequence of size-growing matrices. More importantly, we transform the stochasticity of the network's dimension to the stochasticity of link creation process with a fixed network dimension. Finally, the difference between link creation and node birth is: we control each node in the link creation case, but do not control the node until it joins the network in the node birth case.

C Proof of Theorem 4

Theorem 4 (Generalized Ito's Lemma)**.** *Given the SDE in (5), let $V(\boldsymbol{x}, t)$ be twice-differentiable in $\boldsymbol{x}$ and once in t; then we have:*

$$\mathrm{d}V = \Big\{V_t + \frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top) + V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u})\Big\}\mathrm{d}t + V_{\boldsymbol{x}}^\top\boldsymbol{g}\mathrm{d}\boldsymbol{w} + \big(V(\boldsymbol{x}+\boldsymbol{h}, t) - V(\boldsymbol{x}, t)\big)\mathrm{d}\boldsymbol{N}(t) \tag{23}$$

To prove the theorem, we will first provide some background and useful formulas as follows.

$$(\mathrm{d}t)^2 = 0, \mathrm{d}t\mathrm{d}\boldsymbol{N}(t) = 0, \mathrm{d}t\mathrm{d}\boldsymbol{w}(t) = 0, \mathrm{d}\boldsymbol{w}(t)\mathrm{d}\boldsymbol{N}(t) = 0, \mathrm{d}\boldsymbol{w}(t)\mathrm{d}\boldsymbol{w}(t)^\top = \mathrm{d}t\boldsymbol{I} \tag{24}$$

All the above equations hold in the *mean square limit* sense. The mean square limit definition enables us to extend the calculus rules for deterministic functions and properly define stochastic calculus rules such as stochastic differential and stochastic integration for stochastic processes. See (Hanson, 2007) for the proof of these equations.

Proof. We first restate the SDE in (5) as follows

$$\begin{aligned}\mathrm{d}\boldsymbol{x} &= \big(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{u}\big)\mathrm{d}t + \boldsymbol{g}(\boldsymbol{x})\mathrm{d}\boldsymbol{w}(t) + \boldsymbol{h}(\boldsymbol{x})\mathrm{d}\boldsymbol{N}(t)\\ &= \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{h}(\boldsymbol{x})\mathrm{d}\boldsymbol{N}(t),\end{aligned}$$

where $\boldsymbol{F}(\boldsymbol{x})$ denotes the continuous part of the SDE and is define as

$$\boldsymbol{F}(\boldsymbol{x}) = \big(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{u}\big)\mathrm{d}t + \boldsymbol{g}(\boldsymbol{x})\mathrm{d}\boldsymbol{w}(t)$$

Note that the term $\boldsymbol{h}\mathrm{d}\boldsymbol{N}(t)$ denotes the discontinuous part of the SDE. For the simplicity of notation, we set $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}$ and $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{h}$ and omit the dependency on $\boldsymbol{x}$.

Next, we expand $\mathrm{d}V$ according to its definition as follows

$$\mathrm{d}V(\boldsymbol{x}, t) = V(\boldsymbol{x}(t+\mathrm{d}t), t+\mathrm{d}t) - V(\boldsymbol{x}, t)$$

With the definition $\boldsymbol{x}(t+\mathrm{d}t) = \boldsymbol{x}(t) + \mathrm{d}\boldsymbol{x}$, we can expand $V(\boldsymbol{x}(t+\mathrm{d}t), t+\mathrm{d}t)$ using Taylor expansion on variable t as follows

$$\begin{aligned}V(\boldsymbol{x}(t+\mathrm{d}t), t+\mathrm{d}t) &= V(\boldsymbol{x}+\mathrm{d}\boldsymbol{x}, t+\mathrm{d}t) &(25)\\ &= V(\boldsymbol{x}+\mathrm{d}\boldsymbol{x}, t) + V_t(\boldsymbol{x}, t)\mathrm{d}t &(26)\end{aligned}$$

Next, we expand $V(\boldsymbol{x}+\mathrm{d}\boldsymbol{x}, t)$ as follows

$$\begin{aligned}&V(\boldsymbol{x}+\mathrm{d}\boldsymbol{x}, t)\\ &= V(\boldsymbol{x}+\boldsymbol{F}+\boldsymbol{h}\mathrm{d}\boldsymbol{N}(t), t) &(27)\\ &= \Big(V(\boldsymbol{x}+\boldsymbol{F}+\boldsymbol{h}, t) - V(\boldsymbol{x}+\boldsymbol{F}, t)\Big)\mathrm{d}\boldsymbol{N}(t) + V(\boldsymbol{x}+\boldsymbol{F}, t) &(28)\\ &= \Big[\underbrace{V(\boldsymbol{x}+\boldsymbol{h}, t) + V_{\boldsymbol{x}}(\boldsymbol{x}+\boldsymbol{h})^\top\boldsymbol{F} + \frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}(\boldsymbol{x}+\boldsymbol{h})\boldsymbol{F}^\top}_{\text{Taylor expansion 1}} - \Big(\underbrace{V(\boldsymbol{x}, t) + V_{\boldsymbol{x}}^\top\boldsymbol{F} + \frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}\boldsymbol{F}^\top}_{\text{Taylor expansion 2}}\Big)\Big]\mathrm{d}\boldsymbol{N}(t) &(29)\\ &\quad + \underbrace{V(\boldsymbol{x}, t) + V_{\boldsymbol{x}}^\top\boldsymbol{F} + \frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}\boldsymbol{F}^\top}_{\text{Taylor expansion 2}} &(30)\\ &= \Big(V(\boldsymbol{x}+\boldsymbol{h}, t) - V(\boldsymbol{x}, t)\Big)\mathrm{d}\boldsymbol{N}(t) + \Big(V_{\boldsymbol{x}}(\boldsymbol{x}+\boldsymbol{h}) - V_{\boldsymbol{x}}\Big)^\top\boldsymbol{F}\mathrm{d}\boldsymbol{N}(t) &(31)\\ &\quad + V(\boldsymbol{x}, t) + V_{\boldsymbol{x}}^\top\boldsymbol{F} + \frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}\boldsymbol{F}^\top + \Big(\frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}(\boldsymbol{x}+\boldsymbol{h})\boldsymbol{F}^\top - \frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}(\boldsymbol{x})\boldsymbol{F}^\top\Big)\mathrm{d}\boldsymbol{N}(t) &(32)\end{aligned}$$

Next, we show the reasoning from (27) to (30).

First, we derive a stochastic calculus rule for the point process. Specifically, since $\mathrm{d}\boldsymbol{N}(t) \in \{0, 1\}$, there are two cases for (27): If a jump happens, *i.e.*, $\mathrm{d}\boldsymbol{N}(t) = 1$, (27) is equivalent to $V(\boldsymbol{x}+\boldsymbol{F}(\boldsymbol{x})+\boldsymbol{h}(\boldsymbol{x}), t)$; otherwise, we have $\mathrm{d}\boldsymbol{N}(t) = 0$ and (27) is equivalent to $V(\boldsymbol{x}+\boldsymbol{F}(\boldsymbol{x}), t)$. Hence (28) is equivalent to (27). This stochastic rule essentially takes $\mathrm{d}\boldsymbol{N}(t)$ from inside the value function V to the outside.

Second, from (28) to (30), we have used the following Taylor expansions.

Taylor expansion 1. For $V(\boldsymbol{x}+\boldsymbol{F}+\boldsymbol{h},t)$, we expand it around $V(\boldsymbol{x}+\boldsymbol{h},t)$ on the $\boldsymbol{x}$-dimension

$$V(\boldsymbol{x}+\boldsymbol{F}+\boldsymbol{h},t)=V(\boldsymbol{x}+\boldsymbol{h},t)+V_{\boldsymbol{x}}(\boldsymbol{x}+\boldsymbol{h})^\top\boldsymbol{F}+\frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}(\boldsymbol{x}+\boldsymbol{h})\boldsymbol{F}^\top$$

Taylor expansion 2. For $V(\boldsymbol{x}+\boldsymbol{F},t)$, we expand it around $V(\boldsymbol{x},t)$ along the $\boldsymbol{x}$ dimension

$$V(\boldsymbol{x}+\boldsymbol{F},t)=V(\boldsymbol{x},t)+V_{\boldsymbol{x}}^\top\boldsymbol{F}+\frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}\boldsymbol{F}^\top$$

Next, we simplify each term in (31) and (32). We keep the first term and expand the second term, $\Big(V_{\boldsymbol{x}}(\boldsymbol{x}+\boldsymbol{h})-V_{\boldsymbol{x}}\Big)^\top\boldsymbol{F}\mathrm{d}\boldsymbol{N}(t)$ as follows

$$\begin{aligned}\Big(V_{\boldsymbol{x}}(\boldsymbol{x}+\boldsymbol{h})-V_{\boldsymbol{x}}\Big)^\top\boldsymbol{F}\mathrm{d}\boldsymbol{N}(t)&=\Big(V_{\boldsymbol{x}}(\boldsymbol{x}+\boldsymbol{h})-V_{\boldsymbol{x}}\Big)^\top((\boldsymbol{f}+\boldsymbol{u})\mathrm{d}t+\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t))\mathrm{d}\boldsymbol{N}(t)\\&=\Big(V_{\boldsymbol{x}}(\boldsymbol{x}+\boldsymbol{h})-V_{\boldsymbol{x}}\Big)^\top\Big((\boldsymbol{f}+\boldsymbol{u})\mathrm{d}t\mathrm{d}\boldsymbol{N}(t)+\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t)\mathrm{d}\boldsymbol{N}(t)\Big),\\&=0\end{aligned}\tag{33}$$

where we have used the equations: $\mathrm{d}t\mathrm{d}\boldsymbol{N}(t)=0$ and $\mathrm{d}\boldsymbol{w}(t)\mathrm{d}\boldsymbol{N}(t)=0$ in the Ito mean square limit sense from (24).

We keep the third term and expand the fourth term $V_{\boldsymbol{x}}^\top\boldsymbol{F}$ as

$$V_{\boldsymbol{x}}^\top\boldsymbol{F}=V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u})\mathrm{d}t+V_{\boldsymbol{x}}^\top\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t)\tag{34}$$

The fifth term $\frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}\boldsymbol{F}^\top$ is expanded as follows

$$\begin{aligned}&\frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}\boldsymbol{F}^\top\\&=\frac{1}{2}\Big((\boldsymbol{f}+\boldsymbol{u})\mathrm{d}t+\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t)\Big)V_{\boldsymbol{xx}}\Big((\boldsymbol{f}+\boldsymbol{u})\mathrm{d}t+\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t)\Big)^\top\\&=\frac{1}{2}\Big((\boldsymbol{f}+\boldsymbol{u})V_{\boldsymbol{xx}}(\boldsymbol{f}+\boldsymbol{u})^\top(\mathrm{d}t)^2+2(\boldsymbol{f}+\boldsymbol{u})\mathrm{d}tV_{\boldsymbol{xx}}(\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t))^\top+(\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t))V_{\boldsymbol{xx}}(\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t))^\top\Big)\\&=\frac{1}{2}\Big(0+0+\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top)\mathrm{d}t\Big)\\&=\frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top)\mathrm{d}t,\end{aligned}\tag{35}$$

where we have used the property that $(\mathrm{d}t)^2=0$, $\mathrm{d}t\mathrm{d}\boldsymbol{w}=0$, and $\mathrm{d}\boldsymbol{w}(t)\mathrm{d}\boldsymbol{w}(t)^\top=\mathrm{d}t\boldsymbol{I}$ from (24).

Finally, the last term is expressed as

$$\begin{aligned}&\Big(\frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}(\boldsymbol{x}+\boldsymbol{h})\boldsymbol{F}^\top-\frac{1}{2}\boldsymbol{F}V_{\boldsymbol{xx}}(\boldsymbol{x})\boldsymbol{F}^\top\Big)\mathrm{d}\boldsymbol{N}(t)\\&=\frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}(\boldsymbol{x}+\boldsymbol{h})\boldsymbol{g}\boldsymbol{g}^\top)\mathrm{d}t\mathrm{d}\boldsymbol{N}(t)-\frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top)\mathrm{d}t\mathrm{d}\boldsymbol{N}(t)=0-0=0\end{aligned}\tag{36}$$

Substituting equation (33), (34), (35), and (36) to equation (31) and (32), we have:

$$\begin{aligned}V(\boldsymbol{x}+\mathrm{d}\boldsymbol{x},t)=&\Big(V(\boldsymbol{x}+\boldsymbol{h},t)-V(\boldsymbol{x},t)\Big)\mathrm{d}\boldsymbol{N}(t)+V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u})\mathrm{d}t+V_{\boldsymbol{x}}^\top\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t)\\&+V(\boldsymbol{x},t)+\frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top)\mathrm{d}t\end{aligned}\tag{37}$$

Plugging (37) to (26), we have:

$$\begin{aligned}V(\boldsymbol{x}(t+\mathrm{d}t),t+\mathrm{d}t)=&\Big(V(\boldsymbol{x}+\boldsymbol{h},t)-V(\boldsymbol{x},t)\Big)\mathrm{d}\boldsymbol{N}(t)+V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u})\mathrm{d}t+V_{\boldsymbol{x}}^\top\boldsymbol{g}\mathrm{d}\boldsymbol{w}(t)\\&+V(\boldsymbol{x},t)+\frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top)\mathrm{d}t+V_t(\boldsymbol{x},t)\mathrm{d}t\end{aligned}$$

Hence after simplification, we obtain (23) and finishes the proof:

$$\begin{aligned}&\mathrm{d}V=V(\boldsymbol{x}(t+\mathrm{d}t),t+\mathrm{d}t)-V(\boldsymbol{x}(t),t)\\&=\Big\{V_t+\frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top)+V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u})\Big\}\mathrm{d}t+V_{\boldsymbol{x}}^\top\boldsymbol{g}\mathrm{d}\boldsymbol{w}+\big(V(\boldsymbol{x}+\boldsymbol{h},t)-V(\boldsymbol{x},t)\big)\mathrm{d}\boldsymbol{N}(t)\end{aligned}$$

D Proof of Theorem 5

Theorem 5. *The HJB equation for the user activity guiding problem in (7) is*

$$-V_t = \min_{\boldsymbol{u}} \Big[\mathcal{L} + \frac{1}{2}\mathrm{tr}\big(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top\big) + V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u}) + \sum\nolimits_{j=1}^{U} \lambda_j(t)\big(V(\boldsymbol{x}+\boldsymbol{h}_j(\boldsymbol{x}),t) - V(\boldsymbol{x},t)\big)\Big]$$

where $\boldsymbol{h}_j(\boldsymbol{x})$ is the j-th column of $\boldsymbol{h}(\boldsymbol{x})$.

Proof. First we express the value function V as follows

$$\begin{aligned}
V(\boldsymbol{x},t) &= \min_{\boldsymbol{u}} \mathbb{E}\Big[V(\boldsymbol{x}(t+\mathrm{d}t), t+\mathrm{d}t) + \int_t^{t+\mathrm{d}t} \mathcal{L}\,\mathrm{d}\tau\Big] && (38)\\
&= \min_{\boldsymbol{u}} \mathbb{E}\Big[V(\boldsymbol{x},t) + \mathrm{d}V + \mathcal{L}\,\mathrm{d}t\Big] && (39)\\
&= \min_{\boldsymbol{u}} \mathbb{E}\Big[V(\boldsymbol{x},t) + \Big\{V_t + \frac{1}{2}\mathrm{tr}(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top) + V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u})\Big\}\mathrm{d}t \\
&\qquad + V_{\boldsymbol{x}}^\top \boldsymbol{g}\mathrm{d}\boldsymbol{w} + \big(V(\boldsymbol{x}+\boldsymbol{h},t) - V(\boldsymbol{x},t)\big)\mathrm{d}\boldsymbol{N}(t) + \mathcal{L}\,\mathrm{d}t\Big] && (40)\\
&= \min_{\boldsymbol{u}} \Big[V(\boldsymbol{x},t) + \Big\{V_t + \mathcal{L} + \frac{1}{2}\mathrm{tr}\big(V_{\boldsymbol{xx}}\boldsymbol{g}\boldsymbol{g}^\top\big) + V_{\boldsymbol{x}}^\top(\boldsymbol{f}+\boldsymbol{u})\Big\}\mathrm{d}t \\
&\qquad + \sum\nolimits_{j=1}^{U} \lambda_j(t)\big(V(\boldsymbol{x}+\boldsymbol{h}_j(\boldsymbol{x}),t) - V(\boldsymbol{x},t)\big)\mathrm{d}t,\Big] && (41)
\end{aligned}$$

where (39) to (40) follows from Theorem 4, and (40) to (41) follows from the properties of Wiener processes and point processes, *i.e.*, $\mathbb{E}[\mathrm{d}\boldsymbol{w}] = 0$ and $\mathbb{E}[\mathrm{d}\boldsymbol{N}(t)] = \boldsymbol{\lambda}(t)\mathrm{d}t$.

Finally, cancelling $V(\boldsymbol{x},t)$ on both sides of (41) and dividing both sides by $\mathrm{d}t$ yields the HJB equation.

E Proof of Proposition 6

For the quadratic cost case (the opinion least square guiding problem), we have: $\phi = \frac{1}{2}\|\boldsymbol{x}(T) - \boldsymbol{a}\|^2$, $\mathcal{L} = \frac{1}{2}\|\boldsymbol{x}(t) - \boldsymbol{a}\|^2 + \frac{\rho}{2}\|\boldsymbol{u}(t)\|^2$. Since the instantaneous cost $\mathcal{L}$ is quadratic in $\boldsymbol{x}$ and $\boldsymbol{u}$, and terminal cost ϕ is quadratic in $\boldsymbol{x}$, if the control $\boldsymbol{u}$ is a linear function of $\boldsymbol{x}$, then the value function V must be quadratic in $\boldsymbol{x}$, since it is the optimal value of the summation of quadratic functions.

Moreover, the fact that $\boldsymbol{u}$ is linear in $\boldsymbol{x}$ is because our SDE model for user activities is linear in both $\boldsymbol{x}$ and $\boldsymbol{u}$. Since $V(T) = \phi(T)$ is quadratic, as illustrated in (Hanson, 2007), one can show by induction that when computing the value of V backward in time, $\boldsymbol{u}$ is always linear in $\boldsymbol{x}$.

Similarly, one can show that the value function V is linear in the state $\boldsymbol{x}$ for the linear cost case (opinion maximization problem), where $\phi = -\sum_u x_u(T)$, $\mathcal{L} = -\sum_u x_u(t) + \frac{\rho}{2}\|\boldsymbol{u}(t)\|^2$.

F Additional synthetic experiments

Synthetic experimental setup. We consider a network with 1000 users, where the network topology matrix is randomly generated with a sparsity of 0.001. We simulate the opinion SDE on the observation window $[0, 10]$ by applying Euler forward method to compute the difference form of (6) with $\omega = 1$:

$$x_i(t_{k+1}) = x_i(t_k) + \big(b_i + u_i(t_k) - x_i(t_k)\big)\Delta t + \theta \Delta w_i(t_k) + \sum\nolimits_{j=1}^{U} \alpha_{ij} x_j(t_k) \Delta N_j(t_k),$$

where the observation window is divided into 100 time stamps $\{t_k\}$, with interval length $\Delta t = 0.1$. The Wiener increments Δw_i is sampled from the normal distribution $\mathcal{N}(0, \sqrt{\Delta t})$ and the Hawkes increments $\Delta N_j(t_k)$ is computed by counting the number of events on $[t_k, t_{k+1})$ for user j. The events for each user is simulated by the Otaga's thinning algorithm (Ogata, 1981). The thinning algorithm is essentially a rejection sampling algorithm where samples are first proposed from a homogeneous Poisson process and then samples are kept according to the ratio between the actual intensity and that of the Poisson process.

We set the baseline opinion uniformly at random, $b_i \sim \mathcal{U}[-1, 1]$, noise level $\theta = 0.2$, $\alpha_{ij} \sim \mathcal{U}[0, 0.01]$, initial opinion $x_i(0) = -10$, and $\omega = 1$ for the exponential triggering kernel κ_ω. We repeat simulation of the SDE for ten times and report average performance. We set the tradeoff (budget level) parameter $\rho = 10$, and our results generalize beyond this value.

Network visualization. We conduct control over this 1000-user network with four different initial and target states. Figure 6 shows that our framework works efficiently.

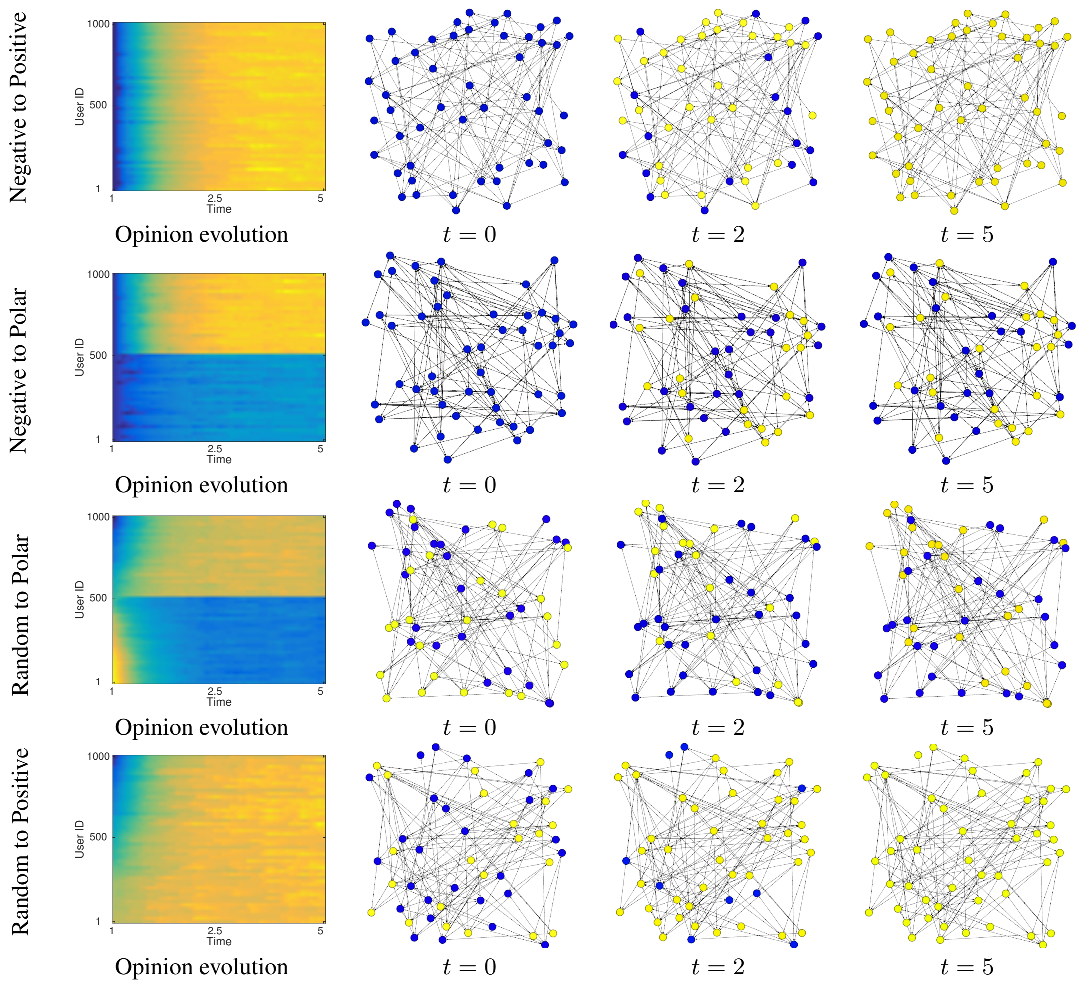

Figure 6: Controlled opinion of four networks with 1,000 users. The first column is the description of opinion change. The second column shows the opinion *value* per user over time. The three right columns show three snapshots of the opinion *polarity* in the network with 50 sub-users at different times. Yellow means positive and blue means negative polarity. Since the controlled trajectory converges fast, we use time range of $[0, 5]$. Parameters are same except for different initial and target state: Set index I to denote user 1-500 and II to denote the rest. 1st row: $\boldsymbol{x}_0 = -10, \boldsymbol{a} = 10$. 2nd row: $\boldsymbol{x}_0 = -10, \boldsymbol{a}(\text{I}) = -5, \boldsymbol{a}(\text{II}) = 10$. 3rd row: $\boldsymbol{x}_0$ sampled uniformly from $[-10, 10]$ and sorted in decreasing order, $\boldsymbol{a}(\text{I}) = -10, \boldsymbol{a}(\text{II}) = 5$. 4th row: $\boldsymbol{x}_0$ is same as (c), $\boldsymbol{a} = 10$.

G Optimal control policy for least square opinion guiding

In this section, we derive the optimal control policy for the opinion SDE defined in (6) with the least square opinion guiding cost. First, we restate the controlled SDE in (6) as follows.

$$\mathrm{d}x_i(t) = \big(b_i + u_i(\boldsymbol{x}, t) - x_i(t)\big)\mathrm{d}t + \theta \mathrm{d}w_i(t) + \sum\nolimits_{j=1}^{U} \alpha_{ij} x_j(t) \mathrm{d}N_j(t)$$

Putting it in the vector form, we have:

$$\mathrm{d}\boldsymbol{x}(t) = \big(\boldsymbol{b} - \boldsymbol{x} + \boldsymbol{u}\big)\mathrm{d}t + \theta \mathrm{d}\boldsymbol{w}(t) + \boldsymbol{h}(\boldsymbol{x})\mathrm{d}\boldsymbol{N}(t)$$

where the j-th column of $\boldsymbol{h}(\boldsymbol{x})$ captures how much influence that x_j has on all other users and is defined as $\boldsymbol{h}_j(\boldsymbol{x}) = \boldsymbol{B}^j \boldsymbol{x}$, where the matrix $\boldsymbol{B}^j \in \Re^{U\times U}$ and has the j-th column to be $(\alpha_{1j}, \cdots, \alpha_{Uj})^\top$ and zero elsewhere.

We substitute $f = \boldsymbol{b} - \boldsymbol{x}(t) + \boldsymbol{u}(,t)$, $\boldsymbol{g} = \theta$ and $\boldsymbol{h}$ to (10) and obtain the HJB equation as

$$-\frac{\partial V}{\partial t} = \min_{\boldsymbol{u}} \Big\{ \mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, t) + \frac{\theta^2}{2} \mathrm{tr}\big(V_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x}, t)\big) + V_{\boldsymbol{x}}(\boldsymbol{x}, t)^\top (\boldsymbol{b} - \boldsymbol{x}(t) + \boldsymbol{u}(t)) + \sum_{j=1}^{U} \lambda_j(t) \big(V(\boldsymbol{x} + \boldsymbol{h}_j(\boldsymbol{x}), t) - V(\boldsymbol{x}, t)\big) \Big\} \tag{42}$$

For the least square guiding problem, the instantaneous cost and terminal cost are defined as

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, t) = \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{a}\|^2 + \frac{1}{2}\rho \|\boldsymbol{u}\|^2, \qquad \phi(T) = \frac{1}{2}\|\boldsymbol{x}(T) - \boldsymbol{a}\|^2$$

Hence we assume that value function V is quadratic in $\boldsymbol{x}$ with unknown coefficients $\boldsymbol{v}_1(t) \in \Re^U$, $\boldsymbol{v}_{11}(t) \in \Re^{U\times U}$ and $v_0(t) \in \Re$:

$$V(\boldsymbol{x}, t) = v_0(t) + \boldsymbol{v}_1(t)^\top \boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{v}_{11}(t) \boldsymbol{x} \tag{43}$$

To find the optimal control, we substitute (43) to HJB equation and take the gradient of the right-hand side of the HJB equation (42) with respect to $\boldsymbol{u}$ and set it to $\mathbf{0}$. This yields the optimal feedback control policy:

$$\boldsymbol{u}^*(\boldsymbol{x}, t) = -\frac{1}{\rho} V_{\boldsymbol{x}} = -\frac{1}{\rho}\Big(\boldsymbol{v}_1(t) + \boldsymbol{v}_{11}(t)\boldsymbol{x}\Big) \tag{44}$$

Substitute $\boldsymbol{u}^*$ in (44) to the HJB equation, we first compute the four terms on the right side of the HJB equation. Note that the minimization is reached when $\boldsymbol{u} = \boldsymbol{u}^*$.

In the following derivations, we will use the property that $\boldsymbol{v}_{11} = \boldsymbol{v}_{11}^\top$ and $\boldsymbol{a}^\top \boldsymbol{b} = \boldsymbol{b}^\top \boldsymbol{a}$ for any vector $\boldsymbol{a}$ and $\boldsymbol{b}$.

The first term is:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}^*, t) &= \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{a} + \frac{1}{2}\rho \boldsymbol{u}^{*\top} \boldsymbol{u}^* \\
&= \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{a} + \frac{1}{2\rho}(\boldsymbol{v}_1 + \boldsymbol{v}_{11}\boldsymbol{x})^\top (\boldsymbol{v}_1 + \boldsymbol{v}_{11}\boldsymbol{x}) \\
&= \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{a} + \frac{1}{2\rho}\boldsymbol{v}_1^\top \boldsymbol{v}_1 + \frac{1}{\rho}\boldsymbol{v}_1^\top \boldsymbol{v}_{11}\boldsymbol{x} + \frac{1}{2\rho}\boldsymbol{x}^\top \boldsymbol{v}_{11}\boldsymbol{v}_{11}\boldsymbol{x} \\
&= \underbrace{\frac{1}{2\rho}\boldsymbol{v}_1^\top \boldsymbol{v}_1}_{\text{scalar}} + \underbrace{\boldsymbol{x}^\top (\frac{1}{\rho}\boldsymbol{v}_{11}\boldsymbol{v}_1 - \boldsymbol{a})}_{\text{linear}} + \underbrace{\frac{1}{2}\boldsymbol{x}^\top (\frac{1}{\rho}\boldsymbol{v}_{11}\boldsymbol{v}_{11} + \boldsymbol{I})\boldsymbol{x}}_{\text{quadratic}}
\end{aligned}$$

Note that in line 1 of the expansion of $\mathcal{L}$, we dropped the constant term $\frac{1}{2}\boldsymbol{a}^\top \boldsymbol{a}$.

The second term is a scalar: $\mathrm{tr}\big(V_{\boldsymbol{x}\boldsymbol{x}}(\boldsymbol{x},t) = \frac{\theta^2}{2}\mathrm{tr}(\boldsymbol{v}_{11})$. The third term is

$$
\begin{aligned}
&V_{\boldsymbol{x}}^\top(\boldsymbol{b}-\boldsymbol{x}+\boldsymbol{u}^*)\\
&=(\boldsymbol{v}_1+\boldsymbol{v}_{11}\boldsymbol{x})^\top(\boldsymbol{b}-\boldsymbol{x}-\boldsymbol{u}^*) = (\boldsymbol{v}_1+\boldsymbol{v}_{11}\boldsymbol{x})^\top(\boldsymbol{b}-\boldsymbol{x}-\frac{1}{\rho}(\boldsymbol{v}_1+\boldsymbol{v}_{11}\boldsymbol{x}))\\
&=(\boldsymbol{v}_1^\top\boldsymbol{b}-\frac{1}{\rho}\boldsymbol{v}_1^\top\boldsymbol{v}_1)-(\boldsymbol{v}_1^\top\boldsymbol{x}+\frac{1}{\rho}\boldsymbol{v}_1^\top\boldsymbol{v}_{11}\boldsymbol{x}+\frac{1}{\rho}\boldsymbol{v}_1^\top\boldsymbol{v}_{11}\boldsymbol{x}-\boldsymbol{b}^\top\boldsymbol{v}_{11}\boldsymbol{x})-\boldsymbol{x}^\top\boldsymbol{v}_{11}^\top\boldsymbol{x}-\frac{1}{\rho}\boldsymbol{x}^\top\boldsymbol{v}_{11}^\top\boldsymbol{v}_{11}\boldsymbol{x}\\
&=\underbrace{(\boldsymbol{v}_1^\top\boldsymbol{b}-\frac{1}{\rho}\boldsymbol{v}_1^\top\boldsymbol{v}_1)}_{\text{scalar}}-\underbrace{\boldsymbol{x}^\top(\boldsymbol{v}_1+\frac{2}{\rho}\boldsymbol{v}_{11}\boldsymbol{v}_1-\boldsymbol{v}_{11}\boldsymbol{b})}_{\text{linear}}-\underbrace{\frac{1}{2}\boldsymbol{x}^\top(2\boldsymbol{v}_{11}+\frac{2}{\rho}\boldsymbol{v}_{11}\boldsymbol{v}_{11})\boldsymbol{x}}_{\text{quadratic}}
\end{aligned}
$$

The fourth term is

$$
\begin{aligned}
&\sum_{j=1}^U \boldsymbol{\lambda}_j(t)(V(\boldsymbol{x}+\boldsymbol{h}_j(\boldsymbol{x}),t)-V(\boldsymbol{x},t))\\
&=\sum_{j=1}^U \boldsymbol{\lambda}_j(t)(\boldsymbol{v}_1^\top\boldsymbol{B}^j\boldsymbol{x}+\frac{1}{2}\boldsymbol{x}^\top\boldsymbol{B}^{j\top}\boldsymbol{v}_{11}\boldsymbol{B}^j\boldsymbol{x}+\frac{1}{2}\boldsymbol{x}^\top 2\boldsymbol{v}_{11}\boldsymbol{B}^j\boldsymbol{x})\\
&=\underbrace{\boldsymbol{x}^\top\boldsymbol{\Lambda}^\top\boldsymbol{v}_1}_{\text{linear}}+\underbrace{\frac{1}{2}\boldsymbol{x}^\top(\sum_{j=1}^U\lambda_j\boldsymbol{B}^{j\top}\boldsymbol{v}_{11}\boldsymbol{B}^j+2\boldsymbol{v}_{11}\boldsymbol{\Lambda})\boldsymbol{x}}_{\text{quadratic}}
\end{aligned}
$$

where $\boldsymbol{\Lambda}(t)=\sum_{j=1}^U\lambda_j(t)\boldsymbol{B}^j$. Next, we compute the left side of HJB equation as:

$$
-V_t = -v_0'(t)-\boldsymbol{x}^\top\boldsymbol{v}_1'(t)-\frac{1}{2}\boldsymbol{x}^\top\boldsymbol{v}_{11}'(t)\boldsymbol{x}
$$

By comparing the coefficients for the scalar, linear and quadratic terms in both left-hand-side and right-hand-side of the HJB equation, we obtain three ODEs as follows.

First, only consider all the coefficients quadratic in $\boldsymbol{x}$:

$$
-\boldsymbol{v}_{11}'(t)=\boldsymbol{I}+2\boldsymbol{v}_{11}(t)(-1+\boldsymbol{\Lambda}(t))+\sum_{j=1}^U\lambda_j(t)\boldsymbol{B}^{j\top}\boldsymbol{v}_{11}(t)\boldsymbol{B}^j-\frac{1}{\rho}\boldsymbol{v}_{11}(t)\boldsymbol{v}_{11}(t)
$$

Second, consider the linear term:

$$
-\boldsymbol{v}_1'(t)=-\boldsymbol{a}+(-1+\boldsymbol{\Lambda}^\top(t)-\frac{1}{\rho}\boldsymbol{v}_{11}(t))\boldsymbol{v}_1(t)+\boldsymbol{v}_{11}(t)\boldsymbol{b}
$$

Third, consider the scalar term:

$$
-v_0'(t)=\boldsymbol{b}^\top\boldsymbol{v}_1(t)+\frac{\theta^2}{2}\mathrm{tr}(\boldsymbol{v}_{11}(t))-\frac{1}{2\rho}\boldsymbol{v}_1^\top(t)\boldsymbol{v}_1(t)
$$

Finally, we compute the terminal condition for the three ODEs by $V(\boldsymbol{x}(T),T)=\phi(\boldsymbol{x}(T),T)$:

$$
\begin{aligned}
V(X(T),T)&=v_0(T)+\boldsymbol{x}(T)^\top\boldsymbol{v}_1(T)+\frac{1}{2}\boldsymbol{x}(T)^\top\boldsymbol{v}_{11}(T)\boldsymbol{x}(T)\\
\phi(\boldsymbol{x}(T),T)&=-\boldsymbol{x}(T)^\top\boldsymbol{a}+\frac{1}{2}\boldsymbol{x}(T)^\top\boldsymbol{x}(T)
\end{aligned}
$$

Hence $v_0(T)=0$, $\boldsymbol{v}_1(T)=-\boldsymbol{a}$ and $\boldsymbol{v}_{11}=\boldsymbol{I}$. Note here we drop the constant term $\frac{1}{2}\boldsymbol{a}^\top\boldsymbol{a}$ in terminal cost ϕ.

Finally, we just need to use Algorithm 1 to solve these ODEs to obtain $\boldsymbol{v}_{11}(t)$ and $\boldsymbol{v}_1(t)$. Substituting $\boldsymbol{v}_{11},\boldsymbol{v}_1$ to (44) leads to the optimal control policy.

H Optimal control policy for opinion influence maximization

In this section, we solve the opinion influence maximization problem. The solving scheme is similar to the least square opinion shaping cost, but the derivation is different due to different cost functions.

First, we choose $\omega = 1$ and restate the controlled opinion SDE in (6) as

$$\mathrm{d}x_i(t) = \big(b_i + u_i(\boldsymbol{x}, t) - x_i(t)\big)\mathrm{d}t + \theta \mathrm{d}w_i(t) + \sum\nolimits_{j=1}^{U} \alpha_{ij} x_j(t) \mathrm{d}N_j(t)$$

Putting it in the vector form, we have:

$$\mathrm{d}\boldsymbol{x}(t) = \big(\boldsymbol{b} - \boldsymbol{x} + \boldsymbol{u}\big)\mathrm{d}t + \theta \mathrm{d}\boldsymbol{w}(t) + \boldsymbol{h}(\boldsymbol{x})\mathrm{d}\boldsymbol{N}(t)$$

where the j-th column of $\boldsymbol{h}(\boldsymbol{x})$ captures how much influence that x_j has on all other users and is defined as $\boldsymbol{h}_j(\boldsymbol{x}) = \boldsymbol{B}^j \boldsymbol{x}$, where the matrix $\boldsymbol{B}^j \in \Re^{U \times U}$ and has the j-th column to be $(\alpha_{1j}, \cdots, \alpha_{Uj})^\top$ and zero elsewhere. We substitute $f = \boldsymbol{b} - \boldsymbol{x}$, $\boldsymbol{g} = \theta$ and $\boldsymbol{h}$ to (10) and obtain the HJB equation as follows

$$-\frac{\partial V}{\partial t} = \min_{\boldsymbol{u}} \Big\{ \mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, t) + \frac{\theta^2}{2} \mathrm{tr}\big(V_{\boldsymbol{xx}}(\boldsymbol{x}, t)\big) + V_{\boldsymbol{x}}(\boldsymbol{x}, t)^\top (\boldsymbol{b} - \boldsymbol{x}(t) + \boldsymbol{u}(t)) + \sum_{j=1}^{U} \lambda_j(t) \big(V(\boldsymbol{x} + \boldsymbol{h}_j(\boldsymbol{x}), t) - V(\boldsymbol{x}, t) \big) \Big\} \quad (45)$$

For opinion influence maximization, we define the cost as follows. Suppose the goal is to maximize the opinion influence at each time on $[0, T]$, the instantaneous cost $\mathcal{L}$ is defined as:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, t) = -\sum_{j=1}^{U} x_i(t) + \frac{1}{2}\|\boldsymbol{u}(t)\|^2 = -\boldsymbol{x}(t)^\top \mathbf{1} + \frac{1}{2}\|\boldsymbol{u}(t)\|^2$$

where $\mathbf{1}$ is the column vector with each entry to be one. For the terminal cost, we have: $\phi(T) = -\boldsymbol{x}(T)^\top \mathbf{1}$.

Following the similar reasoning as the least square opinion guiding problem. Since the terminal cost ϕ is linear in the state $\boldsymbol{x}$, the value function must be linear in $\boldsymbol{x}$, since it is the optimal value of a linear function. Hence we set the value function $V(\boldsymbol{x}, t)$ to be a linear function in $\boldsymbol{x}$ with *unknown coefficients* $\boldsymbol{v}_1(t) \in \Re^U$ and $v_0(t) \in \Re$:

$$V(\boldsymbol{x}, t) = v_0(t) + \boldsymbol{v}_1(t)^\top \boldsymbol{x} \quad (46)$$

To find the optimal control, we substitute (46) to (45) and take the gradient of the right-hand-side of (45) with respect to $\boldsymbol{u}$ and set it to $\mathbf{0}$. This yields the optimal control policy:

$$\boldsymbol{u}^*(t) = -\frac{1}{\rho} \boldsymbol{V}_{\boldsymbol{x}} = -\frac{1}{\rho} \boldsymbol{v}_1(t) \quad (47)$$

Next, we just need to compute $\boldsymbol{v}_1(t)$ to find $\boldsymbol{u}^*$. Substitute $\boldsymbol{u}^*$ in (47) to the HJB equation, we will compute the four terms on the right side of the HJB equation and derive the ODEs by comparing the coefficients. Note that the minimization is reached when $\boldsymbol{u} = \boldsymbol{u}^*$.

First, $\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}^*, t)$ is expanded as:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}^*, t) = -\boldsymbol{x}^\top \mathbf{1} + \frac{1}{2}\|\boldsymbol{u}^*\|^2 = \underbrace{\frac{1}{2\rho} \boldsymbol{v}_1^\top \boldsymbol{v}_1}_{\text{scalar}} - \underbrace{\boldsymbol{x}^\top \mathbf{1}}_{\text{linear}}$$

Since V is linear in $\boldsymbol{x}$, $V_{\boldsymbol{xx}} = 0$. The third term is:

$$V_{\boldsymbol{x}}^\top (\boldsymbol{b} - \boldsymbol{x} + \boldsymbol{u}^*) = \boldsymbol{v}_1^\top (\boldsymbol{b} - \boldsymbol{x} - \frac{1}{\rho}\boldsymbol{v}_1) = \underbrace{\boldsymbol{v}^\top \boldsymbol{b} - \frac{1}{\rho} \boldsymbol{v}^\top \boldsymbol{v}_1}_{\text{scalar}} - \underbrace{\boldsymbol{x}^\top \boldsymbol{v}_1}_{\text{linear}}$$

The fourth term is:

$$\sum_{j=1}^{U} \lambda_j(t) (V(\boldsymbol{x} + \boldsymbol{h}_j(\boldsymbol{x}), t) - V(\boldsymbol{x}, t)) = \sum_{j=1}^{U} \lambda_j(t) \boldsymbol{v}_1^\top \boldsymbol{h}_j(\boldsymbol{x}) = \underbrace{\boldsymbol{x}^\top \boldsymbol{\Lambda}^\top \boldsymbol{v}_1}_{\text{linear}}$$

where $\boldsymbol{\Lambda}(t) = \sum_{j=1}^{U} \lambda_j(t) \boldsymbol{B}^j$. Next, we compute the left-hand-side of HJB equation as:

$$-V_t = -v_0'(t) - \boldsymbol{x}^\top \boldsymbol{v}_1'(t) \quad (48)$$

Then by comparing the coefficients for the scalar and linear terms in both left side and right side of the HJB equation, we obtain two ODEs.

First, only consider all the coefficients linear in $\boldsymbol{x}$:

$$\boldsymbol{v}_1'(t) = \mathbf{1} + \boldsymbol{v}_1(t) - \boldsymbol{\Lambda}^\top \boldsymbol{v}_1(t) \tag{49}$$

Second, consider the linear term:

$$v_0'(t) = -\frac{1}{2\rho}\boldsymbol{v}_1^\top \boldsymbol{v}_1 - \boldsymbol{v}_1^\top \boldsymbol{b} + \frac{1}{\rho}\boldsymbol{v}_1^\top \boldsymbol{v}_1 = -\boldsymbol{v}_1(t)^\top \boldsymbol{b} + \frac{1}{2\rho}\boldsymbol{v}_1(t)^\top \boldsymbol{v}_1(t)$$

Hence we just need to solve the ODEs (48) to obtain $\boldsymbol{v}_1$ and then compute the optimal control $\boldsymbol{u}^*(t)$ from (47).

Finally we derive the terminal conditions for the above two ordinary differential equations. First, $V(T) = \phi(T) = -\boldsymbol{x}(T)^\top \mathbf{1}$ holds from the definition of the value function. Moreover, from the function form of V, we have $= v_0(T) + \boldsymbol{x}^\top \boldsymbol{v}_1(T)$. Hence by comparing the coefficients, we have $v_0(T) = 0$ and $\boldsymbol{v}_1(T) = -\mathbf{1}$.

With the above terminal condition and (49), we will use Algorithm 1 to solve for $\boldsymbol{v}_1(t)$ and obtain the optimal control policy.