Repulsive Markovian Models for Opinion Dynamics

Citation for the original published paper (version of record):
Heiker, C., Gaetan, E., Giarré, L. et al (2024). Repulsive Markovian Models for Opinion Dynamics. Systems and Control Letters, 185. http://dx.doi.org/10.1016/j.sysconle.2024.105720

(article starts on next page)

Repulsive Markovian Models for Opinion Dynamics*

Carl-Johan Heiker[1,†], Elisa Gaetan[2], Laura Giarré[2] and Paolo Falcone[1,2]

***Abstract*— We consider the problem of modeling a decision-making process in a network of stochastic agents, each described as a Markov chain. Two approaches for describing disagreement among agents as social forces are studied. These forces modulate the rates at which agents transition between decisions. We define *similarity* conditions between the two disagreement models and derive a method for obtaining two model instances that fulfill this property. Moreover, we show that a condition for significantly reducing the state-space dimension through *marginalization* can be derived for both models. However, using a counterexample, we also demonstrate that *similarity* is not generally possible for models that can be *marginalized*. Finally, we recommend which disagreement model to use based on the results of our comparison.**

***Index Terms*— Opinion dynamics, Agent-based systems, Markov processes**

I. Introduction

In many areas, from market economy and politics to vehicular traffic, groups of individuals need to make collective decisions through communication. Opinion dynamics [1] can be used to describe and predict the outcome of these processes. However, for complex systems that evolve from human interaction, deriving deterministic models may require more information than what is practically available. Markov chains can be used in stochastic abstractions of such systems. For example, [2] models highway traffic as vehicles that randomly arrive to– and depart from lanes, so that the number of vehicles in each lane is a Continuous Time Markov Chain (CTMC). Based on a similar arrival- and departure process, [3] describes a system of traffic intersections as interconnected CTMCs. Using a different approach, [4] models an intersection as a stochastic decision process to predict road users' decisions to yield or go through the crossing. Each road user switches between the two decision states as a CTMC, according to the Markovian opinion dynamics framework proposed in [5], where CTMC agents change each other's state probabilities by transition rate modulation. In [6], this framework is used to describe how stubborn agents induce rifts in community opinion and how social power develops depending on the interaction topology. Additionally, [7] showed that a marginalized form of this Markovian opinion dynamic model can be obtained analytically. Within this framework, there is a need to develop models of basic interaction forms that can be observed in applications.

*This work is supported by the Swedish Innovation Agency (VINNOVA) under the grant 2018-05005, and partially supported by the EU NextGenerationEU (PNRR) within the Italian National Ph.D. Program in Autonomous Systems (DAuSy).

†Corresponding author.

1Carl-Johan Heiker and Paolo Falcone are with the Mechatronics group, Departement of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden. {heikerc, paolo.falcone}@chalmers.se

2Elisa Gaetan, Laura Giarré and Paolo Falcone are with the Engineering Department 'Enzo Ferrari', University of Modena and Reggio Emilia, Modena, Italy. {elisa.gaetan, laura.giarre}@unimore.it

(a) Indirect repulsion. (b) Direct repulsion.

Fig. 1: Drivers a and b decide between lanes L_1 and L_2. Indirect repulsion increases the rate of different decisions, while direct repulsion decreases the rate of identical decisions.

In [4], an *indirect repulsion force* models disagreement in a network of Markovian agents by increasing the frequency at which agents make conflicting decisions. In this paper, we present the *direct repulsion force* which instead describes disagreement as a decrease in the rates at which agents make identical decisions. The difference between the two principles is illustrated in the following example.

Example 1: Drivers a and b in Fig. 1 choose between highway lanes L_1 and L_2. Indirect repulsion causes a to choose L_2 at a *higher* frequency since b is not in L_2 (Fig. 1a). Conversely, direct repulsion makes a decide L_1 at a *lower* frequency since b is in L_1 (Fig. 1b).

As direct repulsion may seem more intuitive than indirect repulsion, a technical comparison is necessary to determine if the two alternative disagreement models can similarly describe the same decision process. This objective is achieved with the following four contributions: 1) The direct repulsion force is defined; 2) It is shown that an agent network $\mathbb{D}$ constructed using direct repulsion can be reduced to a small-scale model by analytical *marginalization*; 3) The *similarity* property is defined for two general networks, and a method for obtaining $\mathbb{D}$ similar to an agent network $\mathbb{I}$, constructed using indirect repulsion, is formulated; 4) A counterexample is derived to show that, in general, the similarity conditions do not hold for marginalizable networks.

In the following, Section II describes agent- and network models, Section III covers the repulsion forces, Section IV shows how to derive the marginalization of $\mathbb{D}$, and Section V defines similarity between $\mathbb{D}$ and $\mathbb{I}$. In Section VI we construct a counterexample showing that marginalizable networks cannot always be similar, while Section VII demonstrates an example in which similarity is obtained when $\mathbb{I}$ is marginalizable. Finally, conclusions are drawn in Section VIII.

II. Basic Markovian network model

A. Agents as Continuous-Time Markov Chains

Let $\mathcal{N}$ denote a set of N agents. We model the isolated decision process of the n:th member as a time-homogeneous CTMC over a set $\mathcal{S} = \{s_1, \dots, s_M\}$ of M decision states. State transitions $s_i \rightarrow s_j$, where $i \neq j$, occur at non-negative real rates $Q_n[i,j]$ from the $M \times M$ matrix Q_n, in which each diagonal element $Q_n[i,i]$ is the negated sum of the off-diagonal elements in row i. The vector $\Pi_n(t) = \begin{bmatrix}\pi_1^n(t) & \pi_2^n(t) & \dots & \pi_M^n(t)\end{bmatrix}^T$ is a probability distribution over $\mathcal{S}$ describing the probability that the agent is in each state at time t. It is given by solving

$$\dot{\Pi}_n(t) = Q_n^T \Pi_n(t), \tag{1}$$

where $\dot{\Pi}_n(t)$ denotes the first time derivative of $\Pi_n(t)$, from some initial condition $\Pi_n(0)$. The solution $\Pi_n(t)$, including the transient behavior, is fully characterized by Q_n and $\Pi_n(0)$ according to [8].

As done in [5] and [4], we assume that Q_n is irreducible, which ensures that the associated CTMC is ergodic. This implies that the agent has a nonzero probability to visit all states, and that $\Pi_n(t)$ converges to a unique, stationary state probability $\bar{\Pi}_n$.

B. Agent networks

For a general network $\mathbb{X}$ of N agents with M decisions there are M^N network states, each defined as a tuple $s^{\mathbb{X}} = \langle s^1, \dots, s^n, \dots s^N\rangle$, where s^n denotes the state of the n:th agent. The network states are organized in the set

$$\mathcal{S}_{\mathbb{X}} = \mathcal{S}_1 \times \dots \times \mathcal{S}_n \times \dots \times \mathcal{S}_N, \tag{2}$$

where each $\mathcal{S}_n$ represents the same decisions as $\mathcal{S}$, but is indexed to distinguish between agents. The probability that two agents transition at exactly the same time is zero, so every possible transition between two network states is caused by the state transition of a single agent. Hence, $\mathbb{X}$ can be described as a CTMC over $\mathcal{S}_{\mathbb{X}}$. The probability that $\mathbb{X}$ is in each state at time t is denoted $\Pi_{\mathbb{X}}(t)$ and given by solving

$$\dot{\Pi}_{\mathbb{X}}(t) = Q^T \Pi_{\mathbb{X}}(t), \tag{3}$$

where $\dot{\Pi}_{\mathbb{X}}(t)$ denotes the first time derivative of $\Pi_{\mathbb{X}}(t)$, from some initial condition $\Pi_{\mathbb{X}}(0)$. The network transition rate matrix is

$$Q = \sum_{n=1}^{N} I_{M^{n-1}} \otimes Q_n \otimes I_{M^{N-n}}, \tag{4}$$

where $\otimes$ represents the Kronecker product, Q_n is the rate matrix of the n:th agent, and I denotes identity matrices with dimensions M^{n-1} and M^{N-n}, respectively. Since the $M^N \times M^N$ matrix Q describes a CTMC, it obeys the same definition as in II-A, albeit for a different, larger state space. As explained in [5], this network CTMC inherits the ergodicity property of the individual agents.

III. Indirect and direct repulsion

In the following, we recall the indirect repulsion force presented in [4] and propose an alternative direct formulation for modeling disagreement between stochastic agents.

A. Indirect repulsion

The indirect repulsion force is expressed as

$$^{+}\xi_j^n(\mathcal{R}_\ell, \gamma) = \gamma \sum_k \Gamma^{\mathcal{A}}_{\mathcal{R}_\ell}[n,k]\big(1 - \mathbf{1}_j^k\big) \tag{5}$$

and describes how the n:th agent, a member of the group $\mathcal{A}$, *increases* the rate of transitions to s_j. The increase is directly dependent on which of the agents, indexed k, in the group $\mathcal{R}_\ell$ are in states *different from* s_j, an event denoted by the negated indicator function $(1 - \mathbf{1}_j^k)$. $\mathcal{A}$ and $\mathcal{R}_\ell$ are disjoint subsets of $\mathcal{N}$ that interact according to the directed graph $G^{\mathcal{A}}_{\mathcal{R}_\ell} = \big(\mathcal{N}, \Gamma^{\mathcal{A}}_{\mathcal{R}_\ell}\big)$, where $\Gamma^{\mathcal{A}}_{\mathcal{R}_\ell}$ is an $N \times N$ row-normalized, non-negative real adjacency matrix of agent-to-agent influence strength. A group $\mathcal{A}$ can be repulsed by several disjoint groups $\mathcal{R}_\ell \in \mathcal{R}$ with a unique graph describing each conflict. However, each group can be repulsed by other groups (assuming the role of $\mathcal{A}$ in one conflict), and repulse other groups (assuming the role of $\mathcal{R}_\ell$ in another conflict), simultaneously.

The parameter $\gamma \geq 0$ sets the magnitude of the repulsion force in proportion to the weighted average decisions of $\mathcal{R}_\ell$. It can be a function of both time-independent traits, such as agent index n and the pair $\big(\mathcal{A}, \mathcal{R}_\ell\big)$, but also of the time-dependent network state $s^{\mathbb{X}}$. However, state independence is required for analytical marginalization to be possible. We call this property *marginalizability*.

To express indirect repulsion in the network model (3), we define the $M^N \times M^N$ network transition rate matrix R^+ with the same indexing convention as Q. If a network transition from the a:th to the b:th state in $\mathcal{S}_{\mathbb{X}}$, denoted $s_a^{\mathbb{X}} \rightarrow s_b^{\mathbb{X}}$ where $a \neq b$, is caused by the n:th agent's transition $s_i \rightarrow s_j$, then $R^+[a,b] = \sum_{\mathcal{R}_\ell \in R} {}^{+}\xi_j^n(\mathcal{R}_\ell, \gamma | s_a^{\mathbb{X}}, s_b^{\mathbb{X}})$. While the indicator functions from (5) are random variables, they can be evaluated deterministically given the network states $s_a^{\mathbb{X}}$ and $s_b^{\mathbb{X}}$, such that $R^+[a,b]$ becomes a nonnegative real scalar. Each diagonal element in R^+, however, is the negated sum of off-diagonal elements in the corresponding row.

An agent network modeled with indirect repulsion is denoted $\mathbb{I}$, and its decision probabilities $\Pi_{\mathbb{I}}(t)$ are found by solving

$$\dot{\Pi}_{\mathbb{I}}(t) = (Q + R^+)^T \Pi_{\mathbb{I}}(t), \tag{6}$$

where $\dot{\Pi}_{\mathbb{I}}(t)$ is the first time derivative of $\Pi_{\mathbb{I}}(t)$, from some initial condition $\Pi_{\mathbb{I}}(0)$.

B. Direct repulsion

As an alternative form of repulsion, we now describe the *direct* repulsion force as

$$^{-}\xi_j^n(\mathcal{R}_\ell, \gamma) = \gamma \sum_k \Gamma^{\mathcal{A}}_{\mathcal{R}_\ell}[n,k]\mathbf{1}_j^k. \tag{7}$$

This force is denoted "−" as it, in contrast to (5) which is denoted "+", directly *reduces* the n:th agent's transition rate to s_j depending on which agents, indexed k, are in s_j. However, to prevent negative transition rates, γ must be limited, as shown next.

Given state-independent influence parameters γ, assume that the n:th agent is maximally repulsed. That is, every group $\mathcal{R}_\ell$ has all its agents in s_j so that $\mathbf{1}_j^k = 1$ for all k and $\mathcal{R}_\ell$. Since each $\Gamma_{\mathcal{R}_\ell}^{\mathcal{A}}$ is row normalized, the total rate reduction is $\sum_{\mathcal{R}_\ell \in \mathcal{R}} {}^{-}\xi_j^n(\mathcal{R}_\ell, \gamma) = \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma$. If this is less than or equal to the minimal isolated transition rate $q_{\text{min}}^n = \min_{i \neq j} Q_n[i,j]$, none of the n:th agent's transition rates are negative. Assuming that γ is dependent on n and $(\mathcal{A}, \mathcal{R}_\ell)$, we require

$$\sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma \le q_{\text{min}}^n, \quad \forall n. \tag{8}$$

As for indirect repulsion, we define a $M^N \times M^N$ transition rate matrix R^- to describe direct repulsion in the network model (3). A network transition from the a:th to the b:th state in $\mathcal{S}_{\mathbb{X}}$, where $a \neq b$, that is caused by the n:th agent's transition $s_i \to s_j$ obtains the rate reduction $R^-[a,b] = \sum_{\mathcal{R}_\ell \in R} {}^{-}\xi_j^n(\mathcal{R}_\ell, \gamma, |s_a^{\mathbb{X}}, s_b^{\mathbb{X}})$. Each diagonal element in R^- is the negative sum of the off-diagonal elements in the corresponding row. A network described with direct repulsion is denoted $\mathbb{D}$, and its decision probabilities $\Pi_{\mathbb{D}}(t)$ are found by solving

$$\dot{\Pi}_{\mathbb{D}}(t) = (Q - R^-)^T \Pi_{\mathbb{D}}(t), \tag{9}$$

where $\dot{\Pi}_{\mathbb{D}}(t)$ is the first time derivative of $\Pi_{\mathbb{D}}(t)$, from some initial condition $\Pi_{\mathbb{D}}(0)$.

Remark 1: Unlike indirect repulsion, direct repulsion needs to be limited, requiring knowledge of q_{min}^n. Additionally, its transition rate reductions may have unwanted effects on the transient phase of the network decision probabilities. These disadvantages are important to consider when choosing repulsion model.

IV. Marginalization with Direct Repulsion

Through *marginalization*, the probability distribution of a single variable can be derived by summation over a joint probability distribution. This way, MN individual agent decision probabilities can be derived from an M^N network model. However, [4] showed that a marginalization of $\mathbb{I}$ can be expressed analytically, without using the network model. This is much more computationally efficient. In the following, we show how to analytically derive a marginalization of $\mathbb{D}$ under the same conditions.

Theorem 1: Each row in the marginalized model describes the probability that the n:th agent is in s_j as

$$\dot{\pi}_j^n = \sum_{i=1}^{M} Q^n[i,j]\pi_i^n + \pi_j^n \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma - \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma \sum_k \Gamma_{\mathcal{R}_\ell}^{\mathcal{A}}[n,k]\pi_j^k. \tag{10}$$

Proof: We let $\mathbb{E}[\mathbf{1}(t)] = \pi(t)$, where $\mathbb{E}[\cdot]$ denotes the expected value of a random variable, and express the probability that the n:th agent is in s_j after δt time using the infinitesimal definition of the CTMC. Then,

$$\mathbb{E}\big[\mathbf{1}_j^n(t+\delta t)|s^{\mathbb{X}}\big] = \mathbf{1}_j^n\big(1 - \delta t \mathbb{Q}_O\big) + \big(1 - \mathbf{1}_j^n\big)\delta t \mathbb{Q}_I, \tag{11}$$

where $\mathbb{Q}_O$ and $\mathbb{Q}_I$ represent the total in- and outgoing rates from s_j. Taking the expected value of the RHS,

$$\frac{\mathbb{E}\big[\mathbf{1}_j^n(t+\delta t)|s^{\mathbb{X}}\big] - \mathbb{E}\big[\mathbf{1}_j^n\big]}{\delta t} = \mathbb{E}\big[-\mathbf{1}_j^n\big(\mathbb{Q}_O + \mathbb{Q}_I\big) + \mathbb{Q}_I\big], \tag{12}$$

and the LHS approaches $\dot{\pi}_j^n$ as $\delta t \to 0$. Next, we split $\mathbb{Q}_I$ and $\mathbb{Q}_O$ into two parts: isolated and repulsive rates. Then, we derive the contributions to $\dot{\pi}_j^n$ for each part.

First, let $\mathbb{Q}_O = \sum_{i \neq j} Q_n[j,i]$ and $\mathbb{Q}_I = \sum_{i \neq j} Q_n[i,j]\mathbf{1}_i^n$ in the RHS of (12), which becomes

$$\mathbb{E}\Big[-\mathbf{1}_j^n\Big(\sum_{i \neq j} Q_n[j,i] + \sum_{i \neq j} Q_n[i,j]\mathbf{1}_i^n\Big) + \sum_{i \neq j} Q_n[i,j]\mathbf{1}_i^n\Big]. \tag{13}$$

Since $\sum_{i \neq j} Q_n[j,i] = -Q_n[j,j]$ and $\mathbf{1}_j^n \mathbf{1}_i^n = 0$ for all t, this is reduced to

$$\mathbb{E}\Big[Q_n[j,j]\mathbf{1}_j^n + \sum_{i \neq j} Q_n[i,j]\mathbf{1}_i^n\Big] = \mathbb{E}\Big[\sum_{i=1}^{M} Q_n[i,j]\mathbf{1}_i^n\Big]. \tag{14}$$

The contribution to $\dot{\pi}_j^n$ from the isolated rates is thus

$$\sum_{i=1}^{M} Q^n[i,j]\pi_i^n, \tag{15}$$

which corresponds to the first term in (10).

Second, by setting $\mathbb{Q}_O = -\sum_{\mathcal{R}_\ell \in \mathcal{R}} \sum_{i \neq j} {}^{-}\xi_i^n(\mathcal{R}_\ell, \gamma)$ and $\mathbb{Q}_I = -\sum_{\mathcal{R}_\ell \in \mathcal{R}} {}^{-}\xi_j^n(\mathcal{R}_\ell, \gamma)$, we evaluate the rate contributions from the direct repulsion force (7). Assuming that γ is state-independent, but possibly dependent on n, $\mathcal{A}$ and $\mathcal{R}_\ell$, the RHS of (12) becomes

$$\mathbb{E}\Big[\mathbf{1}_j^n \sum_{\mathcal{R}_\ell \in \mathcal{R}} \sum_{i=1}^{M} {}^{-}\xi_i^n(\mathcal{R}_\ell, \gamma) - \sum_{\mathcal{R}_\ell \in \mathcal{R}} {}^{-}\xi_j^n(\mathcal{R}_\ell, \gamma)\Big]. \tag{16}$$

As γ is independent on s_i and s_j, we obtain

$$\mathbb{E}\Big[\mathbf{1}_j^n \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma \sum_{i=1}^{M} \sum_k \Gamma_{\mathcal{R}_\ell}^{\mathcal{A}}[n,k]\mathbf{1}_i^k - \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma \sum_k \Gamma_{\mathcal{R}_\ell}^{\mathcal{A}}[n,k]\mathbf{1}_j^k\Big]. \tag{17}$$

If we for each agent in $\mathcal{R}_\ell$, indexed k, sum over the indicator functions for every possible state $i = 1 \dots M$, $\mathbf{1}_i^k = 1$ for exactly one i. Additionally, $\Gamma_{\mathcal{R}_\ell}^{\mathcal{A}}$ is row normalized. Thus,

$$\mathbb{E}\Big[\mathbf{1}_j^n \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma - \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma \sum_k \Gamma_{\mathcal{R}_\ell}^{\mathcal{A}}[n,k]\mathbf{1}_j^k\Big], \tag{18}$$

which rewritten in terms of π is

$$\pi_j^n \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma - \sum_{\mathcal{R}_\ell \in \mathcal{R}} \gamma \sum_k \Gamma_{\mathcal{R}_\ell}^{\mathcal{A}}[n,k]\pi_j^k, \tag{19}$$

the second and third term in (10). ■

We use (10) to construct the marginalized model

$$\dot{\Pi}_m(t) = (Q_m - R_m)^T \Pi_m(t), \tag{20}$$

where Q_m and R_m are $MN \times MN$ matrices constructed from the terms (15) and (19), respectively.

A. Effects of marginalizability on R^+ and R^-

Assume that two different network transitions $s_a^{\mathbb{X}} \to s_b^{\mathbb{X}}$ and $s_c^{\mathbb{X}} \to s_d^{\mathbb{X}}$ are caused by the same agent making the transitions $s_i \to s_j$ and $s_k \to s_l$, respectively. If the same repulsive agents who are in s_j when $s_a^{\mathbb{X}} \to s_b^{\mathbb{X}}$, are also in s_l for $s_c^{\mathbb{X}} \to s_d^{\mathbb{X}}$, the transitioning agent faces an identical opposition in both scenarios. Then, the summation term in (5) (or (7) for direct repulsion) will be identical for both network transitions. Moreover, if all parameters are state-independent (a necessary condition for marginalizability) the influence strength is identical in both scenarios. Consequentially, the force and thereby the corresponding element in R^+ (or R^-) is the same for $s_a^{\mathbb{X}} \to s_b^{\mathbb{X}}$ and $s_c^{\mathbb{X}} \to s_d^{\mathbb{X}}$.

V. Similarity between direct and indirect repulsion

To establish if a network $\mathbb{D}$ defined with direct repulsion and a network $\mathbb{I}$ formulated with indirect repulsion can be used to similarly model a decision process, we define two criteria for *similarity* between two general networks: 1) the difference in the networks' expected decision state holding time is minimized w.r.t. some criterion (e.g. (22) in this paper), and 2) the networks have identical stationary decision probabilities. Next, we define this property formally to derive a method that finds $\mathbb{D}$ similar to $\mathbb{I}$.

A. Minimal difference in expected decision state holding time

Our first criterion for similarity is that the time that $\mathbb{D}$ spends in each decision state should be as close as possible to that of $\mathbb{I}$. As described in [8], the state holding time V_i of a state s_i in a CTMC with rate matrix Q is exponentially distributed with parameter $-Q[i,i]$. It follows that the expectation $\mathbb{E}[V_i] = -1/Q[i,i]$. For a state $s_i^{\mathbb{X}}$ in the networks $\mathbb{I}$ and $\mathbb{D}$, the cumulative probability distribution functions of state holding time are

$$P[V_i^+ \leq t] = 1 - e^{\left(Q[i,i]+R^+[i,i]\right)t} \quad \text{and} \tag{21a}$$

$$P[V_i^- \leq t] = 1 - e^{\left(Q[i,i]-R^-[i,i]\right)t}, \tag{21b}$$

respectively. Thus, $\mathbb{I}$ decreases the expected state holding time, while $\mathbb{D}$ increases it. To keep the decision state holding times of both networks as close as possible, the increase of $\mathbb{E}[V_i^-]$ should be minimized, equivalent to minimizing the diagonals of R^-. For this purpose, we formulate the objective function

$$f(r,H) = r^T A^T H A r, \tag{22}$$

where r is a column vector such that r_i represents the i:th positive R^- element found by nested row-wise and left-to-right iteration. A is a $(0,1)$-matrix such that $Ar = -\text{diag}(R^-)$, and H is a tuning matrix for emphasizing the minimization of a subset of diagonal elements.

B. Equal stationary decision probabilities

Our second criterion for similarity between $\mathbb{D}$ and $\mathbb{I}$ is that they should reach identical stationary state probabilities. While the transition rate matrix of an ergodic CTMC maps to a unique stationary state, several matrices can produce the same stationary state. We need to find the off-diagonal elements of R^- such that the stationary state equations of (6) and (9), expressed as

$$(Q + R^+)^T \bar{\Pi} = 0 \text{ and} \tag{23a}$$

$$(Q - R^-)^T \bar{\Pi} = 0, \tag{23b}$$

hold given Q, R^+ and the shared stationary distribution $\bar{\Pi}$. Since Q becomes redundant in finding R^-, (23) reduces to

$$(R^-)^T \bar{\Pi} = -(R^+)^T \bar{\Pi}. \tag{24}$$

We can rearrange the linear combination $(R^-)^T \bar{\Pi}$ into a product between a known matrix $M_{\bar{\Pi}}$ constructed from elements of $\bar{\Pi}$ and the vector r of unknown R^- off-diagonals. The rate reductions in r must be positive, but cannot exceed the transition rates in Q. In total, we require

$$M_{\bar{\Pi}} r = -(R^+)^T \bar{\Pi} \tag{25a}$$

$$0 < r < q, \tag{25b}$$

where q is a vector of Q elements and indexed like r.

All solutions to (25a) are given by

$$r = -M_{\bar{\Pi}}^{\dagger} (R^+)^T \bar{\Pi} + (I - M_{\bar{\Pi}}^{\dagger} M_{\bar{\Pi}}) w, \tag{26}$$

where $\dagger$ denotes the *Moore-Penrose inverse* and w is an arbitrary vector, see e.g. [9]. Solutions exist iff $M_{\bar{\Pi}} M_{\bar{\Pi}}^{\dagger} = -(R^+)^T \bar{\Pi}$. If the latter holds, the solution is unique iff $M_{\bar{\Pi}}$ has full column rank, and $I - M_{\bar{\Pi}}^{\dagger} M_{\bar{\Pi}}$ is a zero matrix. While this condition can be used to check the existence of any solution r, we also require $r < q$, which in practice is easily enforced directly in a numerical solver.

C. Obtaining $\mathbb{D}$ similar to $\mathbb{I}$ by constrained optimization

Minimizing the objective function (22) under the constraints in (25) defines the constrained optimization problem

$$\underset{r}{\text{minimize}} \quad f(r,H) \tag{27a}$$

$$\text{subject to} \quad M_{\bar{\Pi}} r = -(R^+)^T \bar{\Pi}, \tag{27b}$$

$$0 < r < q, \tag{27c}$$

which, for any number of agents and decisions, finds the off-diagonal elements of the matrix R^- such that $\mathbb{D}$ is *similar* to $\mathbb{I}$. In the following, we investigate how the feasibility of (27) is affected by network structure and additional constraints for obtaining marginalizable networks.

VI. Feasibility under similarity- and marginalizability constraints

To show that similarity is not generally feasible for marginalizable networks, we return to Example 1 in the introduction, and investigate (27) in three cases:

1) No additional marginalization constraints are imposed,
2) $\mathbb{D}$ is marginalizable, and
3) $\mathbb{D}$ and $\mathbb{I}$ are both marginalizable.

In Example 1, $N = 2$ drivers, a and b, choose between $M = 2$ lanes at rates in

$$Q_1 = \begin{bmatrix} -q^a_{12} & q^a_{12} \\ q^a_{21} & -q^a_{21} \end{bmatrix}, \quad Q_2 = \begin{bmatrix} -q^b_{12} & q^b_{12} \\ q^b_{21} & -q^b_{21} \end{bmatrix}, \tag{28}$$

respectively, so that $\mathbb{X}$ has a state space $\mathcal{S}_{\mathbb{X}} = \{s^1_1, s^1_2\} \times \{s^2_1, s^2_2\}$ according to (2). Its isolated transition rate matrix is

$$Q = \begin{bmatrix} -q^a_{12} - q^b_{12} & q^b_{12} & q^a_{12} & 0 \\ q^b_{21} & -q^a_{12} - q^b_{21} & 0 & q^a_{12} \\ q^a_{21} & 0 & -q^a_{21} - q^b_{12} & q^b_{12} \\ 0 & q^a_{21} & q^b_{21} & -q^a_{21} - q^b_{21} \end{bmatrix}. \tag{29}$$

Assuming mutual repulsion, the structure of the network transition rate matrices for indirect and direct repulsion are

$$R^+ = \begin{bmatrix} -(r^{a+}_{12} + r^{b+}_{12}) & r^{b+}_{12} & r^{a+}_{12} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & r^{a+}_{21} & r^{b+}_{21} & -(r^{a+}_{21} + r^{b+}_{21}) \end{bmatrix}, \tag{30a}$$

$$R^- = \begin{bmatrix} 0 & 0 & 0 & 0 \\ r^{b-}_{21} & -(r^{a-}_{12} + r^{b-}_{21}) & 0 & r^{a-}_{12} \\ r^{a-}_{21} & 0 & -(r^{a-}_{21} + r^{b-}_{12}) & r^{b-}_{12} \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{30b}$$

Here, $r^{a/b+}_{ij}$ and $r^{a/b-}_{ij}$ denote the rate increase and decrease for each agent's transition $s_i \to s_j$ as a result of indirect and direct repulsion, respectively. To formulate the constraints in (27b) and (27c), we construct

$$M_{\bar{\Pi}} = \begin{bmatrix} \bar{\pi}_2 & 0 & \bar{\pi}_3 & 0 \\ -\bar{\pi}_2 & -\bar{\pi}_2 & 0 & 0 \\ 0 & 0 & -\bar{\pi}_3 & -\bar{\pi}_3 \\ 0 & \bar{\pi}_2 & 0 & \bar{\pi}_3 \end{bmatrix}, \tag{31a}$$

$$q = \begin{bmatrix} q^b_{21} & q^a_{12} & q^a_{21} & q^b_{12} \end{bmatrix}^T, \tag{31b}$$

$$r = \begin{bmatrix} r^{b-}_{21} & r^{a-}_{12} & r^{a-}_{21} & r^{b-}_{12} \end{bmatrix}^T, \tag{31c}$$

$$\bar{\Pi} = \begin{bmatrix} \bar{\pi}_1 & \bar{\pi}_2 & \bar{\pi}_3 & \bar{\pi}_4 \end{bmatrix}^T. \tag{31d}$$

Fig. 2 shows the CTMC representations of a, b and $\mathbb{X}$. For the sake of demonstration, $\mathbb{X}$ has both the indirect and direct repulsion rates added to its isolated transition rates, thus depicting $\mathbb{I}$ and $\mathbb{D}$ simultaneously. In this special case, indirect and direct repulsion never affects the same network transition, which can also be seen in the non-overlapping sparsity patterns of (30a) and (30b).

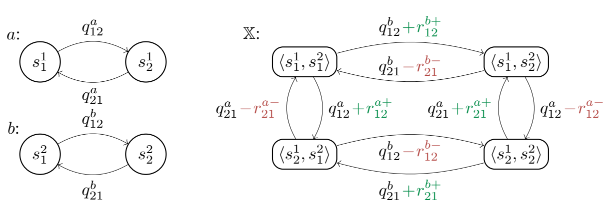

Fig. 2: CTMCs a and b and their network CTMC $\mathbb{X}$.

A. No marginalization constraints

Evaluating the constraints (27b) and (27c) for Example network 1 yields

$$0 < r^{b-}_{21} = \frac{\bar{\pi}_1 \, r^{b+}_{12} - \bar{\pi}_4 \, r^{b+}_{21} + \bar{\pi}_3 \, x}{\bar{\pi}_2} < q^b_{21}, \tag{32a}$$

$$0 < r^{a-}_{12} = \frac{\bar{\pi}_4 \, r^{a+}_{21} + \bar{\pi}_4 \, r^{b+}_{21} - \bar{\pi}_3 \, x}{\bar{\pi}_2} < q^a_{12}, \tag{32b}$$

$$0 < r^{a-}_{21} = \frac{\bar{\pi}_1 \, r^{a+}_{12} + \bar{\pi}_4 \, r^{b+}_{21} - \bar{\pi}_3 \, x}{\bar{\pi}_3} < q^a_{21}, \tag{32c}$$

$$0 < r^{b-}_{12} = x < q^b_{12}. \tag{32d}$$

Thus, the rate x is a free parameter on an interval $(x_{\min}, x_{\max})$, such that (32) hold, showing that similarity can be achieved without marginalization of the networks.

B. $\mathbb{D}$ is marginalizable

Following the reasoning in Section IV-A, choosing state-independent parameters when constructing R^- requires the additional equality constraints $r^{a-}_{21} = r^{a-}_{12}$ and $r^{b-}_{21} = r^{b-}_{12}$, halving the number of unknowns in (27). The lower bound of (32) reduces into the constraints

$$0 < r^{b-}_{21} = \frac{\bar{\pi}_1 \, r^{b+}_{12}}{\bar{\pi}_2} + \frac{\bar{\pi}_4 \, r^{a+}_{21}}{\bar{\pi}_2} + \frac{\bar{\pi}_1 \, r^{a+}_{12}}{\bar{\pi}_2 - \bar{\pi}_3} - \frac{\bar{\pi}_4 \, r^{a+}_{21}}{\bar{\pi}_2 - \bar{\pi}_3}, \tag{33a}$$

$$0 < r^{a-}_{21} = \frac{\bar{\pi}_4 \, r^{a+}_{21}}{\bar{\pi}_2 - \bar{\pi}_3} - \frac{\bar{\pi}_1 \, r^{a+}_{12}}{\bar{\pi}_2 - \bar{\pi}_3}, \tag{33b}$$

$$0 = \frac{\bar{\pi}_1 \, r^{a+}_{12}}{\bar{\pi}_3} - \frac{\bar{\pi}_1 \, r^{b+}_{12}}{\bar{\pi}_2} - \frac{\bar{\pi}_4 \, r^{a+}_{21}}{\bar{\pi}_2} + \frac{\bar{\pi}_4 \, r^{b+}_{21}}{\bar{\pi}_3}. \tag{33c}$$

The satisfaction of (33c) is determined only by R^+ and $\bar{\Pi}$ and cannot be controlled by choosing R^-. Hence, this counterexample shows that *similarity between $\mathbb{I}$ and $\mathbb{D}$ is not always determined by $\mathbb{D}$, if $\mathbb{D}$ is marginalizable.*

C. $\mathbb{D}$ and $\mathbb{I}$ are marginalizable

Marginalizability of $\mathbb{I}$ requires that $r^{a+}_{12} = r^{a+}_{21}$ and $r^{b+}_{12} = r^{b+}_{21}$ hold, in addition to (33). We can then express r^{b+}_{12} as a function of r^{a+}_{12} in (33c), and substitute r^{b+}_{12} in (33a). This, in combination with (33b), first allows us to derive two sets of conditions

$$\frac{\bar{\pi}_2}{\bar{\pi}_3} > \frac{\bar{\pi}_4}{\bar{\pi}_1} > 1 \quad \text{or} \quad \frac{\bar{\pi}_2}{\bar{\pi}_3} < \frac{\bar{\pi}_4}{\bar{\pi}_1} < 1, \tag{34}$$

that account for the existence and positivity of r.

Second, we must verify that the conditions (34) do not affect the positivity of the entries r^{a+}_{12} and r^{b+}_{12}. From (33c), we assume that $r^{a+}_{12} > 0$ by construction. Therefore, $r^{b+}_{12} > 0$ iff

$$\frac{\bar{\pi}_3 \bar{\pi}_1 - \bar{\pi}_4 \bar{\pi}_2}{\bar{\pi}_2 \bar{\pi}_1 - \bar{\pi}_4 \bar{\pi}_3} > 0, \tag{35}$$

which has two possible sets of solutions,

$$\begin{cases} \frac{\bar{\pi}_3}{\bar{\pi}_2} > \frac{\bar{\pi}_4}{\bar{\pi}_1} \\ \frac{\bar{\pi}_2}{\bar{\pi}_3} > \frac{\bar{\pi}_4}{\bar{\pi}_1} \end{cases} \quad \text{or} \quad \begin{cases} \frac{\bar{\pi}_3}{\bar{\pi}_2} < \frac{\bar{\pi}_4}{\bar{\pi}_1} \\ \frac{\bar{\pi}_2}{\bar{\pi}_3} < \frac{\bar{\pi}_4}{\bar{\pi}_1}. \end{cases} \tag{36}$$

Finally, the constraints (34) for the existence and positivity of the solution must be merged with those for the positivity of the entries of matrix R^+, (36). However, none of these merged sets has a solution, since for both cases there is a conflict between inequalities. This means that it is not possible to find any combinations of $\bar{\Pi}$ entries such that all the constraints hold. Thus, $N = M = 2$ is a counterexample showing that *similarity according to* (27) *is not generally feasible when both* $\mathbb{D}$ *and* $\mathbb{I}$ *are marginalizable.*

VII. Numerical results

In Example 1, the similarity objective function (27a) cannot be minimized as the diagonal elements of (30b) become independent on the free rate r_{12}^{b-} due to (32). However, in the following large-scale example, we show that (27) can be solved when $\mathbb{I}$ is marginalizable, and that minimizing (27a) can lead to a shorter probability transient in $\mathbb{D}$.

Example 2: Assume that $N = 8$ driver agents $c, d \dots j$ choose between $M = 2$ lanes at rates $Q_1[1{,}2]=1$, $Q_2[1{,}2]=2$, $\dots Q_8[1{,}2]=8$ and $Q_{1\dots 8}[2{,}1]=5$, producing a 256×256 Q according to (4). Groups are disregarded, assuming fully connected repulsion between all agents so that $\Gamma^{\mathcal{A}}_{\mathcal{R}_\ell}$ in (5) and (7) always describes repulsion between two agents. For $\mathbb{I}$, we choose state-independent parameters $0.14 \le \gamma \le 0.71$ so that (8) holds for all agents and $\mathbb{I}$ is marginalizable. The objective is to find a $\mathbb{D}$ similar to $\mathbb{I}$ by solving (27) for the vector r containing the 2032 unknown rates in R^-. We assume that R^- is the result of state-dependent γ-parameters so that $\mathbb{D}$ is not marginalizable. The matrix H in (27a) is set to identity for an unweighted minimization of all diagonal elements in R^-. Moreover, to show that the diagonal elements of R^- can indeed be minimized, we find and compare two networks, $\mathbb{D}_1$ and $\mathbb{D}_2$. For $\mathbb{D}_1$, we require $r \ge 0.4$, which is unnecessarily high. For $\mathbb{D}_2$, $r \ge 0.001$, which is closer to the necessary positivity constraint.

We simulate $\mathbb{I}$, $\mathbb{D}_1$ and $\mathbb{D}_2$ to obtain the network probability trajectories from $t = 0$ to $t = 1$. By post-summation over network decision probabilities, we extract the decision probabilities of each agent $c \dots j$ and compare them in Fig. 3. In the first eight plots, it can be seen that $\mathbb{I}$, $\mathbb{D}_1$, and $\mathbb{D}_2$ all produce the same stationary decision probability for each agent, confirming that the second similarity criterion in Section V holds. In the plot of agent c, we also show the network *rise time* until 95% of the stationary probability. Importantly, $\mathbb{I}$ has the fastest rise time $t_1 \approx 0.35$, followed by $\mathbb{D}_2$ with $t_2 \approx 0.55$ and lastly $\mathbb{D}_1$ with $t_3 \approx 0.87$. The eigenvalue plot shows that a higher spectral gap, obtained by minimizing the difference in expected state holding time, corresponds to a shorter transient time in this particular case.

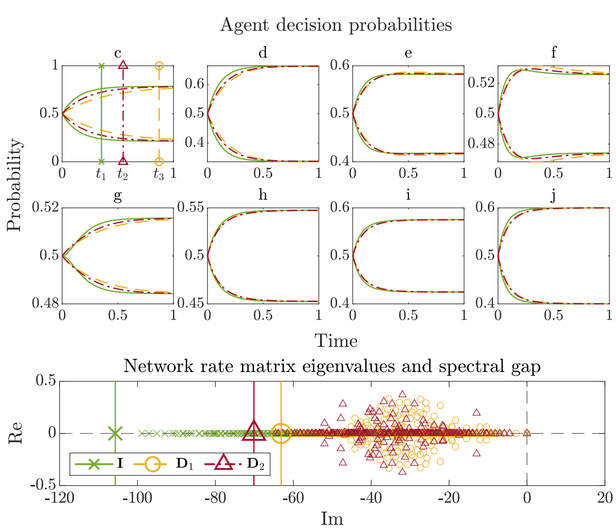

Fig. 3: Agent decision probabilities and eigenvalues.

Remark 2: In the above, we show a case in which we find $\mathbb{D}$ similar to a marginalizable $\mathbb{I}$. When marginalization constraints are enforced also on $\mathbb{D}$, no solution is found.

VIII. Conclusions

We introduce the *direct repulsion force* as an alternative to the *indirect repulsion force* for modeling group-wise disagreement in a decision process among Markovian agents in a network. To see if both methods can similarly describe a decision process, *similarity* is defined as a property between two network models, requiring equality between their respective stationary decision probabilities and a minimal (w.r.t. the cost (27a)) difference between their expected decision state holding times. We show that the state-space dimension of a network defined with direct repulsion also can be significantly reduced using *marginalization*, and for two networks $\mathbb{D}$ and $\mathbb{I}$ defined with direct and indirect repulsion respectively, we demonstrate an example of when $\mathbb{D}$ similar to a marginalizable $\mathbb{I}$ can be obtained. However, we also derive a counterexample to show that the existence of $\mathbb{D}$ similar to $\mathbb{I}$ cannot generally be guaranteed when $\mathbb{D}$ and/or $\mathbb{I}$ are marginalizable. Moreover, direct repulsion requires more information and may increase the time until the decision process becomes stationary. Therefore, we recommend indirect repulsion as a model for disagreement in decision processes among Markovian agents. Future work includes learning indirect repulsion parameters from decision processes in traffic scenarios, such as lane changing.

References

[1] H. Noorazar, K. R. Vixie, A. Talebanpour, and Y. Hu, "From classical to modern opinion dynamics," *Int. J. Mod. Phys. C*, vol. 31, no. 7, 2020.

[2] O. E. Joubari, J. B. Othman, and V. Vèque, "Markov chain mobility model for multi-lane highways," *Mobile Networks and Applications*, vol. 27, no. 3, pp. 1286–1298, 2022.

[3] B. J. Othman, V. Vèque *et al.*, "Continuous time markov chain traffic model for urban environments," in *GLOBECOM*, 2020, pp. 1–6.

[4] C.-J. Heiker and P. Falcone, "Decision modeling in markovian multi-agent systems," in *61st CDC*, 2022, pp. 7235–7240.

[5] P. Bolzern, P. Colaneri, and G. De Nicolao, "Opinion influence and evolution in social networks: A markovian agents model," *Automatica*, vol. 100, pp. 219–230, 2019.

[6] P. Bolzern, P. Colaneri, and G. De Nicolao, "Opinion dynamics in social networks: The effect of centralized interaction tuning on emerging behaviors," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 362–372, 2020.

[7] P. Bolzern, P. Colaneri, and G. De Nicolao, "Opinion dynamics in social networks with heterogeneous markovian agents," in *IEEE CDC*, Miami, USA, Dec. 2018, pp. 6180–6185.

[8] C. G. Cassandras and S. Lafortune, "Markov chains," in *Introduction to discrete event systems*, 2nd ed. Springer, 2008, ch. 7, pp. 368–428.

[9] M. James, "The generalised inverse," *The Mathematical Gazette*, vol. 62, no. 420, pp. 109–114, 1978. [Online]. Available: http://www.jstor.org/stable/3617665