This is the accepted manuscript made available via CHORUS. The article has been published as:

Modeling and numerical simulations of the influenced Sznajd model

Farshad Salimi Naneh Karan, Aravinda Ramakrishnan Srinivasan, and Subhadeep Chakraborty

Phys. Rev. E **96**, 022310 — Published 11 August 2017

DOI: 10.1103/PhysRevE.96.022310

Modeling and Numerical Simulations of the Influenced Sznajd Model

Farshad Salimi Naneh Karan, Aravinda Ramakrishnan Srinivasan, and Subhadeep Chakraborty*
Department of Mechanical, Aerospace, and Biomedical Engineering
University of Tennessee, Knoxville

This paper investigates the effects of *independent nonconformists* or influencers on the behavioral dynamic of a population of agents interacting with each other based on the Sznajd model. The system is modeled on a complete graph using the Master Equation. The acquired equation has been numerically solved. Accuracy of the mathematical model and its corresponding assumptions have been validated by numerical simulations. Regions of initial magnetization have been found from where the system converges to one of two unique steady-state PDFs, depending on the distribution of influencers. The scaling property and entropy of the stationary system in presence of varying level of influence have been presented and discussed.

I. INTRODUCTION

Social scientists, for many years have developed theories of group position [1], social identity [2], and system justification [3]. Now, such theories can be validated quantitatively by analyses of 'retweets', 'via', 'hat tip', and 'mention' conventions which have been shown to be analogous to broadcasting one's position, and helps explain how virality, meme propagation, and opinion formation occur on Twitter [4]. Fast diffusion of uncorroborated information through such non-traditional pathways has often exposed our weakness to safeguard against 'perception manipulation' attacks or 'cognitive hacking' [5]. The widespread adoption of tools related to network-based social engagement and its vast implications have generated a renewed interest in scientists to study the nature, topology, and mechanisms of interactions among individuals in a networked society.

The study of macroscopic changes in the outcome of ordering processes in complex networks mostly proceeds along several well-researched avenues. Of relevance to this paper, are the studies **(a)** investigating effects of different interaction mechanisms such as the Voter model [6], the Sznajd model [7] and the Bounded Confidence model [8] on large-scale ordering dynamics, and **(b)** investigating effects of different network topologies such as the complete graph, the Watts-Strogatz small world topology, the Barabási-Albert scale free network, etc. on the mechanism and rate of information diffusion. For such studies, models from statistical physics such as the Ising model have been often adopted, in which a Master equation for global variables extracts basic features of the ordering process. Presence of influences and zealots in networks is another dimension where researchers have made considerable contributions in relatively recent years.

The Voter model is considered one of the simplest interaction mechanisms which models the spin flip of two neighboring nodes; if two neighboring agents have different opinions (binary states, spins), then one of the agents changes its state to match the opinion of the other one. This mechanism has been modeled on different networks with and without zealots [9–11]. However, because only two agents are present in this opinion exchange, the dynamics do not consider the social validation aspect. The concept of social validation in opinion change through interactions is introduced by Sznajd-Weron and Sznajd in [12].

In all the above-mentioned models, opinions are modelized as numbers, integer or real. Each agent is initialized with a random number as their representative opinions. As interactions proceed, the agents rearrange their opinion variables, through mutual discussions. At some stage, the system reaches a configuration which is stable under the chosen dynamics; this final configuration may represent consensus with all agents sharing the same opinion, polarization with two main clusters of opinions ("parties"), or fragmentation where several opinion clusters survive. In all such evolutionary models of societies, the detailed behavior of each human, inherently the complex outcome of many internal processes, is largely overlooked. Decisions can be described in terms of three essential components such as alternatives, anticipated consequences, and uncertainty. Some of the earliest attempts to formalize this theory were by von Neumann & Morgenstern [13] and Savage [14, 15] who popularized the expected utility family of theories. The heart of the theory, sometimes called the rational expectations principle or expectancy-value model [16], proposes that each alternative course of action or choice should be evaluated by weighting its global expected satisfaction-dissatisfaction with the probabilities that the component consequences will occur and be experienced. These rational actor models prove that social validation is surely not the sole reason of an individual's opinion formation in social settings [17]. Effects such as education [18], culture [19], and group identity [20] are also important when an individual makes a decision.

Gross simplification adopted in the Sznajd model is justified by assuming that higher level features, such as symmetries, dimensionality, and topology of the interaction network are strong indicators of the global behavior of the society, along with microscopic details of individual motives, perceptions and judgments, and worthy of studying in isolation.

II. RECENT STUDIES ON SZNAJD MODEL

In recent years nonconformity has been a focal point for a large number of studies on the Sznajd model. In these

* schakrab@utk.edu

studies, two types of nonconformity are usually considered - anti-conformity and independence. In a study by Nyczka et. al., the authors investigate the effect of two types of social responses: conformity and anti-conformity in the dynamics of the Sznajd model on a complete graph [21]. In this dynamic, conformists choose the opinion of the group with probability p_1 (i.e. $\uparrow\uparrow\Downarrow \longrightarrow \uparrow\uparrow\Uparrow$ with probability p_1, no change with probability $1-p_1$), and anticonformists choose the opposite opinion of the group with probability p_2 (i.e. $\uparrow\uparrow\Uparrow \longrightarrow \uparrow\uparrow\Downarrow$ with probability p_2, no change with probability $1-p_2$). In another study by the same group, authors expand this idea to investigate the qualitative and quantitative differences of the two types of nonconformity under the framework of the q-voter model [22] on a complete graph. In this dynamic, anticonformists choose the opposite opinion of the group, and the independent agents do not follow the group; they act independently and choose the opposite opinion of the group with probability $\frac{1}{2}$.

Jędrzejewski [23], expanded these studies with the Pair Approximation methodology to study the behavior of a q-voter model with stochastic noise characterized as Independence on several complex networks. In this model, with probability p, a chosen agent acts independently and adopts the opposite opinion or preserves the old one with equal chances. Otherwise, with probability $1-p$, the agent will be a conformist.

A special class of independent agents are 'zealots' (sometimes called 'inflexibles' or 'influencers'). These are agents who do not change their state throughout the progression of the dynamics. Mobilia uses Mean-Field approximation method to analytically model the nonlinear q-voter model in presence of zealots [24]. In this work, for simplicity, repetition in choosing agents is allowed. The derived transition probabilities are used for bifurcation analysis and finding critical zealotry density. Analyses presented in this work relate to special cases of symmetric and parameterized asymmetric zealotry distributions.

In this paper, we model and investigate the effects of influencers and initial conditions on the dynamics of the Sznajd model. The problem setup is closely linked to [24], but in contrast to Mobilia's work, does not allow repetitive agent choices. Disallowing repetitive agent choices is not only more realistic, but surprisingly, leads to a new master equation which can be solved under specific conditions. Also, we do not observe rapid phase transitions in this dynamic. Here, influencers are nonconformists of the Independent type, who never change their opinions (as opposed to [23]) but affect other agents' decisions. It is worth mentioning that influencers of this study are different from what have been used in [21, 22]. In these studies, the nonconformists are **Anticonformists**, however in our study, we have chosen **Independents** as nonconformists.

In section III, the influenced Sznajd model is formulated on a complete graph. Then, the derived differential equation is used to study the behavior of the society at the steady state without considering any type of symmetry for influencers. Section IV is dedicated to numerical simulations and validation of the mathematical model. Finally, the findings of this paper are summarized and concluded in section V.

III. INFLUENCED SZNAJD MODEL

In this paper, a complete graph is considered. A complete graph can be characterized by a "magnetization parameter" defined as $m = \frac{N_+ - N_-}{N}$, where, N_+ represents the number of agents in state $+1$, and N_- represents the number of agents in state -1 at any instant. $N = N_+ + N_-$ nodes make up the vertices of the graph [25]. The system also includes $I = I_+ + I_-$ influences, in which I_+ represents the number of influences in state $+1$, and I_- represents the number of influences in state -1.

In the original Sznajd Model [12], the idea of *social validation* is used to introduce a spin dynamic with ± 1 alignments:

- In each time step a pair of spins S_i and S_{i+1} is chosen to change spins of their nearest neighbors, namely S_{i-1} and S_{i+2}.
- If $S_i = S_{i+1}$ then $S_{i-1} = S_i$ and $S_{i+2} = S_i$ (*social validation*).
- If $S_i = -S_{i+1}$ then $S_{i-1} = S_{i+1}$ and $S_{i+2} = S_i$.

In this study, a simplified Sznajd dynamics, modified for a complete graph, rather than a one-dimensional lattice, introduced in [7], is used which inherits the *social validation* concept, but has slightly different update rules:

- In each time step a pair of nodes (i and j), respectively in states S_i and S_j are chosen at random and attempt to change the state S_k of a randomly-chosen common nearest neighbor, $k \in \mathcal{N}_i \cap \mathcal{N}_j$, where $\mathcal{N}_i$ is the set of neighbor nodes of node i.
- If $S_i = S_j$ then $S_k = S_i$ (*social validation*).
- If $S_i = -S_j$ then nothing happens.

Due to the stochastic process of spin flips, Master Equation (ME) is used to formulate this model. ME represents the time evolution of the probability of a system having any configuration of ± 1 spins defined by m. Eqn. 1 represents the general form of the ME which includes in/out-flow rates calculated as probabilities [25].

$$\dot{P}_m = r_{m+\frac{2}{N}} P_{m+\frac{2}{N}} + g_{m-\frac{2}{N}} P_{m-\frac{2}{N}} - (r_m + g_m) P_m \tag{1}$$

where, in/out-flow rates are derived based on the Sznajd model as:

$$
\begin{aligned}
r_m &= P(m \to m - \frac{2}{N}) = \left(\frac{N_- + I_-}{N+I}\right)\left(\frac{N_- + I_- - 1}{N+I-1}\right)\left(\frac{N_+}{N+I-2}\right) \\
g_m &= P(m \to m + \frac{2}{N}) = \left(\frac{N_+ + I_+}{N+I}\right)\left(\frac{N_+ + I_+ - 1}{N+I-1}\right)\left(\frac{N_-}{N+I-2}\right) \\
r_{m+\frac{2}{N}} &= P(m + \frac{2}{N} \to m) = \left(\frac{N_- + I_- - 1}{N+I}\right)\left(\frac{N_- + I_- - 2}{N+I-1}\right)\left(\frac{N_+ + 1}{N+I-2}\right) \\
g_{m-\frac{2}{N}} &= P(m - \frac{2}{N} \to m) = \left(\frac{N_+ + I_+ - 1}{N+I}\right)\left(\frac{N_+ + I_+ - 2}{N+I-1}\right)\left(\frac{N_- + 1}{N+I-2}\right)
\end{aligned}
\tag{2}
$$

By substituting Eqn. 2 in Eqn. 1, after some manipulations, and by defining the control variable as $u = \frac{I_+ - I_-}{I}$, for large but finite population sizes N, the master equation can be expressed as:

$$
\begin{aligned}
\dot{P}_m = \frac{1}{N^4}\Bigg[&\left(Nm^2 - N + \frac{N^2}{2} - \frac{N^2m^2}{2} + \frac{I^2}{2} + \frac{I^2u^2}{2} + IN - I^2mu - Im^2N\right)\frac{\partial^2 P_m}{\partial m^2} \\
&+ \Bigg(10mN + INu - INm - \frac{N^3m}{2} - I^2uN - N^2Iu - 2N^2m + \frac{N^3m^3}{2} \\
&+ \frac{I^2mN}{2} + \frac{I^2u^2mN}{2} + N^2Ium^2 - 4ImN - 2I^2u + 4Iu\Bigg)\frac{\partial P_m}{\partial m} \\
&+ \left(4N - 3IN - N^2 - \frac{N^3}{2} + \frac{3N^3m^2}{2} + 2N^2Imu + \frac{I^2N}{2} + \frac{I^2u^2N}{2}\right)P_m\Bigg]
\end{aligned}
\tag{3}
$$

Assuming that a steady state density function exists, $\lim_{t\to\infty}\frac{dP_m}{dt} = 0$ [26], the steady state density $P_{ss}(m)$ has to satisfy:

$$
\begin{aligned}
&\left(Nm^2 - N + \frac{N^2}{2} - \frac{N^2m^2}{2} + \frac{I^2}{2} + \frac{I^2u^2}{2} + IN - I^2mu - Im^2N\right)\frac{d^2P_{ss}(m)}{dm^2} + \\
&\Bigg(10mN + INu - INm - \frac{N^3m}{2} - I^2uN - N^2Iu - 2N^2m + \frac{N^3m^3}{2} + \\
&\frac{I^2mN}{2} + \frac{I^2u^2mN}{2} + N^2Ium^2 - 4ImN - 2I^2u + 4Iu\Bigg)\frac{dP_{ss}(m)}{dm} + \\
&\left(4N - 3IN - N^2 - \frac{N^3}{2} + \frac{3N^3m^2}{2} + 2N^2Imu + \frac{I^2N}{2} + \frac{I^2u^2N}{2}\right)P_{ss}(m) = 0
\end{aligned}
\tag{4}
$$

which is a homogeneous second order linear differential equation. The terms can be rearranged to a more compact form:

$$
\begin{aligned}
\frac{d}{dm}\left[\left(\Sigma_{i=0}^{3}a_im^i\right)P_{ss}(m) + \left(\Sigma_{i=0}^{2}b_im^i\right)\frac{dP_{ss}(m)}{dm}\right] \\
= -4mN\frac{dP_{ss}(m)}{dm}
\end{aligned}
\tag{5}
$$

Integrating the same equation and using the identity $\int_{-1}^{1}P_{ss}(m)dm = 1$, the general form of a first order ODE is obtained:

$$
B(m)\frac{dP_{ss}(m)}{dm} + A(m)P_{ss}(m) = 4N \tag{6}
$$

The coefficients of this equation are presented in Table I

The analytical solution of such equations has the general form of:

$$
P_{ss}(m) = \frac{1}{v(m)}\left(\int v(m)\frac{4N}{B(m)}dm\right) + C, \tag{7}
$$

where $v(m) = e^{\int\frac{A(m)}{B(m)}dm}$. $P_{ss}(m)$ after normalization, pro-

a_0	$4Iu - I^2u - N^2Iu - NI^2u + NIu$
a_1	$8N - 3NI - N^2 - \frac{N^3}{2} + \frac{NI^2}{2} + \frac{NI^2u^2}{2}$
a_2	N^2Iu
a_3	$\frac{N^3}{2}$
b_0	$-N + \frac{N^2}{2} + \frac{I^2}{2} + \frac{I^2u^2}{2} + IN$
b_1	$-I^2u$
b_2	$N - NI - \frac{N^2}{2}$
A	$a_3m^3 + a_2m^2 + a_1m + a_0$
B	$b_2m^2 + b_1m + b_0$

TABLE I

vides the steady state distribution of the population starting from an initial distribution m_0. In the simplified setting of $I = 0$ and $N \to \infty$, the population almost surely converges to either all $+1$ or all -1 states, the two attractors for that dynamics, i.e. $P_{ss}(m) = \delta(m \pm 1)$ where δ is the Dirac delta function. Which direction the population converges to depends on the initial starting condition, i.e. whether m_0 is greater or less than 0. However, for all these cases, it is observed that $P_{ss}(m = 0) = 0$, i.e. at steady state the probability of the population being equally divided between $+1$ and -1 states is almost surely 0, which stems from $m = 0$ being an unstable equilibrium and the stochastic nature of the process.

Keeping these in mind, the ODE is solved separately over two ranges, P_{ss_l} solved over $-1 \leq m \leq 0$ and P_{ss_h} solved over $0 \leq m \leq 1$ with $P_{ss}(m = 0) = 0$, as the boundary condition for both. It will be shown next, that the entire span $[-1, 1]$ of possible starting values for m_0 can be divided into two distinct regions,

- a *convergence zone (Z_C), starting from which the population consistently converges toward a particular attractive state - consequently, the steady state PDF matches either P_{ss_l} or P_{ss_h}, and*
- *an indeterminate zone (Z_I), starting from which the population can, by chance converge toward any of the two attracting states, thus $P_{ss}(m)$ has a bimodal distribution which is a combination of P_{ss_l} and P_{ss_h}.*

It may be noted, that for the simple case of $I = 0$ and $N \to \infty$, Z_I is a singleton set, with $Z_I = \{0\}$.

IV. RESULTS

Using a variable-step Runge-Kutta method, Eqn. 6 is numerically solved and plotted for different population sizes. Then, results from the Monte Carlo (MC) simulations are overlayed on it.

Fig. 1 includes numerical solution of Eqn. 6 (solid lines) as well as data points from MC simulations for different population sizes (N). Larger populations are characterized by higher peak values and lower variances. Additionally, the assumption of large but finite N lower bounds the population size. This can be seen in the $N = 50$ plot, where results

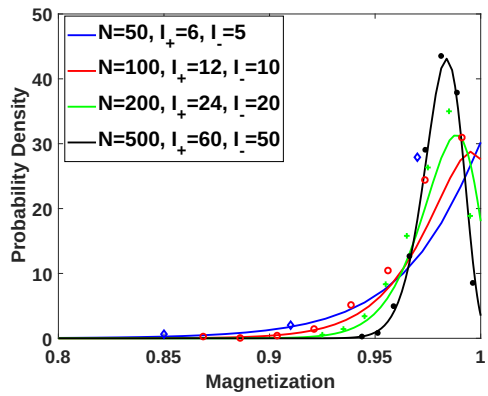

FIG. 1: *Effect of the population size N on the steady state density $P_{ss}(m)$*

from numerical simulations are not in accordance with the numerical solution of the mathematical model. High degree of fidelity between MC simulations data and numerical solutions of the model is observed for $N > 100$.

A caveat on the assumption of large population size should be mentioned here. Realistically, N cannot be unbounded, since there is a cognitive limit to the number of people with whom one can maintain social relationship - this is known as the Dunbar number [27]. It is true that influence in the world of social network is not completely dictated by relationships in which an individual knows the identity of each person and how he/she relates to every other person. For example, the Pew Research Center puts the average number of friends for a Facebook user at 338 [28]. However, these kinds of social networks rarely ever assume a complete graph topology, but can be more realistically modeled as a scale-free network. Keeping this in mind, and since the derivations in this paper rely on the complete graph assumption, all subsequent results are for an intermediate population size ($N = 100$) which is large enough for numerical accuracy, but smaller than the permissible Dunbar number limit ($N_{limit} = 150$).

A. Convergence and Indeterminate Zones

This section aims to elaborate on the effect of initial magnetization (m_0) on the steady state of the system. Each result is from 1000 runs of Monte Carlo simulations in which the magnetization parameter m is treated as a random variable. The rescaled histogram of the random variable is used to estimate the probability density function of the system at steady state. Different population sizes are considered and simulated with all the possible initial conditions (m_0 varies with steps equal to $\frac{2}{N}$).

For ease of representation, the mean magnetization $\langle m_{ss} \rangle$ estimated from the steady state PDF, $P_{ss}(m)$, for each simulation is plotted. Although, not the most comprehensive representation, Fig. 2 by representing the behavior of the mean steady state magnetization of systems as initial magnetization varies, provides a visual representation of the conver-

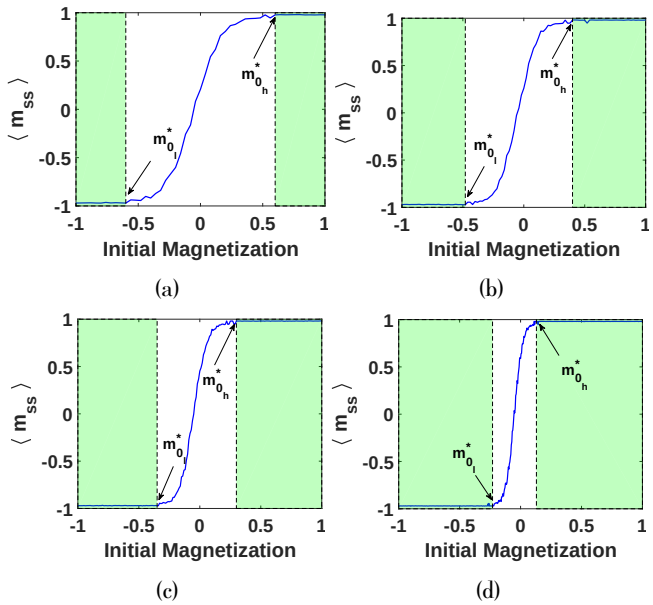

FIG. 2: *Dependency of mean steady state magnetization on initial magnetization,*
(a) $N = 50, I_+ = 6, I_- = 5$, *(b)* $N = 100, I_+ = 12, I_- = 10$, *(c)* $N = 200, I_+ = 24, I_- = 20$, *(d)* $N = 500, I_+ = 60, I_- = 50$

gence and indeterminate zones.

Fig. 2 depicts the two regions, convergence and indeterminate zones. The convergence zone is again composed of two regions - $Z_C = Z_{C_l} \cup Z_{C_h}$, where

$$\begin{aligned} Z_{C_l} &= \{m_0 \leq m^*_{0_l} | P_{ss}(m) = P_{ss_l}(m)\} \\ Z_{C_h} &= \{m_0 \geq m^*_{0_h} | P_{ss}(m) = P_{ss_h}(m)\} \end{aligned} \tag{8}$$

The indeterminate zone

$$Z_I = \{m^*_{0_l} < m_0 < m^*_{0_h} | P_{ss}(m) = C_l P_{ss_l}(m) + C_h P_{ss_h}(m)\}$$

where $C_l > 0$, $C_h > 0$. For different configurations of population size and influences, Z_{C_l} and Z_{C_h} are represented by green with boundaries denoted by $m^*_{0_l}$ and $m^*_{0_h}$.

It can be observed that as population size increases, the convergence zone covers larger intervals of m_0. In addition, with linear scaling of simulation parameters, systems starting from their respective convergence zones converge to the same mean steady state magnetization value. In other words, for two generic systems α and β, with a scale factor of η, if $(m_{0_\alpha} = m_{0_\beta}) \leq m^*_{0_l}$ or if $(m_{0_\alpha} = m_{0_\beta}) \geq m^*_{0_h}$, then $\langle m_{ss}(N, I_+, I_-)\rangle = \langle m_{ss}(\eta N, \eta I_+, \eta I_-)\rangle$, $\eta = 2, 3, 4, \ldots$.

It may be noted that although the results presented so far are illustrated with the help of the mean of the stationary distribution, in fact, the hypothesis is that for each α, the sequence, $P^\alpha_n(m)$ converges in distribution to $P_{ss}(m)$, i.e., $\lim_{n\to\infty} P^\alpha_n(m) = P_{ss}(m)$.

To prove this, denoting a system configuration as $sys = \{N, I_+, I_-, m_0\}$, the following null and alternative hypotheses are constructed:

- **Null Hypothesis** (H_0)**:** For sys_α and sys_β with $N_\alpha = N_\beta$, $I_{+_\alpha} = I_{+_\beta}$, $I_{-_\alpha} = I_{-_\beta}$, if $(m_{0_\alpha} \neq m_{0_\beta}) \leq m^*_{0_l}$ or if $(m_{0_\alpha} \neq m_{0_\beta}) \geq m^*_{0_h}$ then $P^\alpha_{ss}(m) \sim P^\beta_{ss}(m)$.
- **Alternative Hypothesis** (H_1)**:** For sys_α and sys_β with $N_\alpha = N_\beta$, $I_{+_\alpha} = I_{+_\beta}$, $I_{-_\alpha} = I_{-_\beta}$, if $(m_{0_\alpha} \neq m_{0_\beta}) \leq m^*_{0_l}$ or if $(m_{0_\alpha} \neq m_{0_\beta}) \geq m^*_{0_h}$ then $P^\alpha_{ss}(m) \nsim P^\beta_{ss}(m)$.

Simulation data with different m_0's are tested at $\alpha = 0.01$ significance level using the two-sample Kolmogorov-Smirnov (KS) test. Results of the testing do not reject the null hypothesis; thus, the null hypothesis stands true with 99% certainty.

Fig. 3 illustrates steady state PDF's of a system with different initial conditions (marked by red stars in Fig. 3a) for visual verification. It may be noted that when initial conditions are chosen from Z_{C_h}(Figs. 3f, g, h), or from Z_{C_l} (Figs. 3c, d, e), the steady state PDFs each time converge respectively to $P_{ss_l}(m)$ and $P_{ss_h}(m)$.

For $m^*_{0_l} \leq m_0 \leq m^*_{0_h}$ (Fig. 3b), the PDF is a mixture of two PDFs. K-S tests reveal that the rescaled PDFs of these modes are the same as P_{ssl} and P_{ssh} proving that the solution of Eqn. 4 contains both specific solutions; or equivalently, $C_l \neq 0$ and $C_h \neq 0$.

B. Convergence Zone with Varying Control Inputs

In section IV A, a fixed control input is used to find the convergence zones related to different population sizes. The objective of this section is to investigate whether there exist convergence zones for other values of u. Also, how such convergence zones depend on the control input is studied.

To do this, control input u is treated as an independent variable in Fig. 4. Similar to section IV A, MC simulations are performed for all possible combinations of m_0 and u. Then, mean magnetization of the system at the steady state is calculated and plotted.

A specific control input can correspond to different combinations of I_+ and I_-; for e.g., $I_+ = 8, I_- = 2$ and $I_+ = 16, I_- = 4$ both represent $u = 0.6$. To fully specify the problem, we define *influence ratio*, $\lambda = \frac{I}{N}$, as a simulation parameter, where $|I| = |I_+| + |I_-|$. Somewhat arbitrarily, we focus on the range $0 \leq \lambda \leq 1$ for λ, which regulates the total number of influences in the system to be at most the same size as the population. In Fig. 4, the control input changes with steps equal to $\frac{2}{I}$ to cover all possible discrete values in $[-1, 1]$.

Existence of convergence zones: Fig. 4 depicts that for different combinations of λ and u, region(s) ($m_0 \leq m^*_{0_l}$ or $m_0 \geq m^*_{0_h}$) exist, where the mean steady state magnetization $\langle m_{ss}\rangle$ loses its dependency on the initial magnetization, and all systems converge to a unique PDF.

To determine boundaries of convergence zones, a new set of figures are presented. Fig. 5a represents the case with $\lambda = 0.2$ with markers for $m^*_{0_l}$ and $m^*_{0_h}$. Fig. 5b is the two

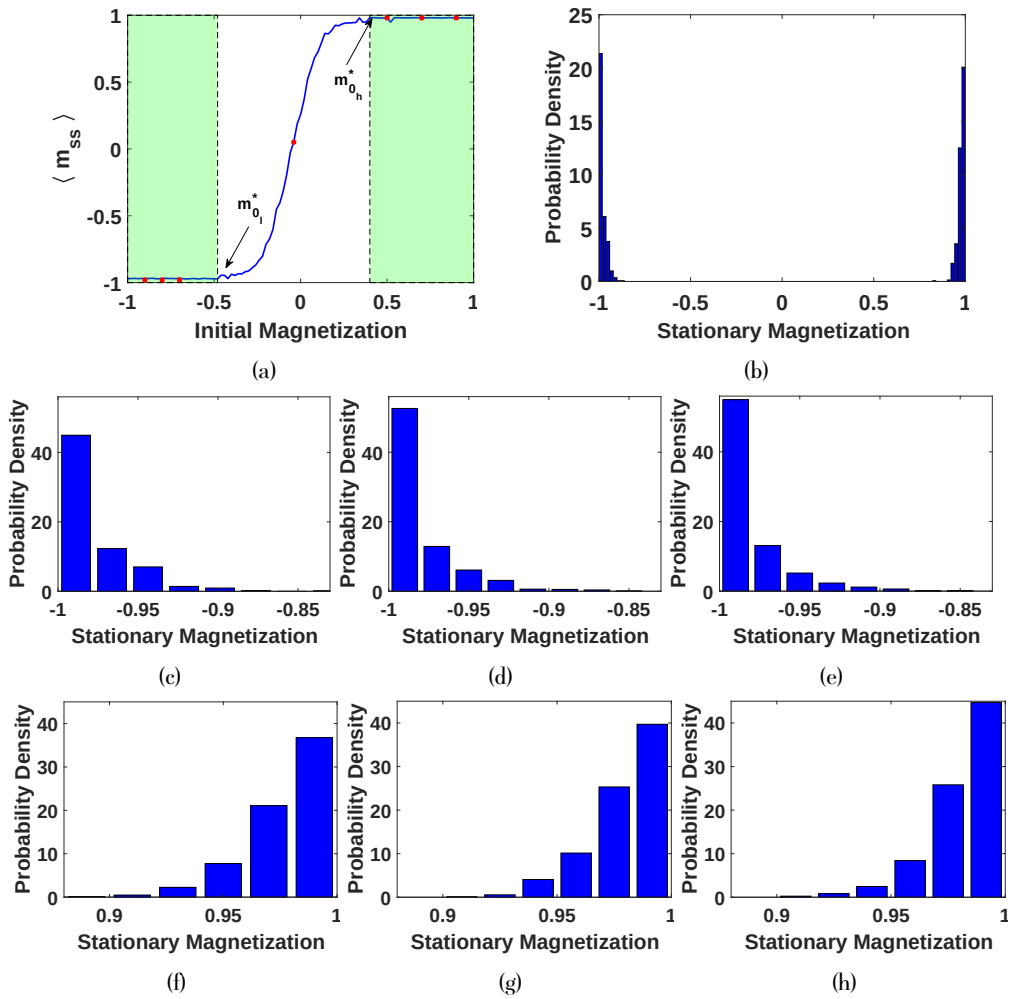

FIG. 3: *Convergence zone and stationary PDF's of a system with* $N = 100, I_+ = 12, I_- = 10$, *(a) chosen initial values* (m_0) *for hypothesis testing, (b)* $m_0 = -0.04$, *(c)* $m_0 = -0.7$, *(d)* $m_0 = -0.8$, *(e)* $m_0 = -0.9$, *(f)* $m_0 = 0.5$, *(g)* $m_0 = 0.7$, *(h)* $m_0 = 0.9$

dimensional location of the same markers on the $m_0 - u$ plane. As it can be seen, for most values of the control input, both $m^*_{0_l}$ and $m^*_{0_h}$ are present. However, in extreme cases of control input one of them is absent. For example, $u = -1$ and $u = 1$ lack $m^*_{0_h}$ and $m^*_{0_l}$ respectively. Same analysis can be applied to systems with other values of λ.

Exit probability, i.e. the probability that the system ultimately reaches consensus as a function of the initial composition of the population, is a very important first-passage property [29]. In this work, the concept of exit probability is not completely applicable since the influenced system never reaches consensus. However, the proof of our Null Hypothesis does show some similarities with the concept of exit probability. Here, we have shown that two similar systems with initial conditions belonging to the same convergence zone will converge to the same probability distribution in their stationary states. The limits of convergence zones depend on N, λ, and u.

Effect of influence ratio (λ)**:** Fig. 4a represents an uninfluenced society. $\langle m_{ss} \rangle$ is only a function of m_0, and boundaries of convergence zone are clear on both sides. However, as soon as influences are added to the society ($\lambda = 0.2$), the behavior of the $\langle m_{ss} \rangle$ starts to change, and dependency on u is noticeable. For instance, on plane $u = -1$, the $\langle m_{ss} \rangle$s for positive m_0s decrease since all of the influences are focused on the negative side. Also, on plane $u = 1$, the $\langle m_{ss} \rangle$s for negative m_0s increase because all the influences are focused on the positive side.

More importantly, at $\lambda = 0.2$, number of influences is not high enough to make the $\langle m_{ss} \rangle$ flatten as m_0 changes on planes $u = \pm 1$. This is the reason why there is neither a $m^*_{0_h}$ on plane $u = -1$ nor a $m^*_{0_l}$ on $u = 1$. However, for higher values of λ, number of influences is high enough to overcome the effect of m_0 and make the $\langle m_{ss} \rangle$ flatten (Figs. 4c-f).

In Figs. 4c-e, as λ increases, the control input becomes the dominant parameter in changing the behavior of $\langle m_{ss} \rangle$, and $\langle m_{ss} \rangle$ completely loses its dependency on m_0 as $u \longrightarrow \pm 1$. In addition, when $u \longrightarrow 0$, m_0 is the dominant parameter in the behavior of $\langle m_{ss} \rangle$. Interestingly, Fig. 4f, $\langle m_{ss} \rangle$ is a function of u only. It

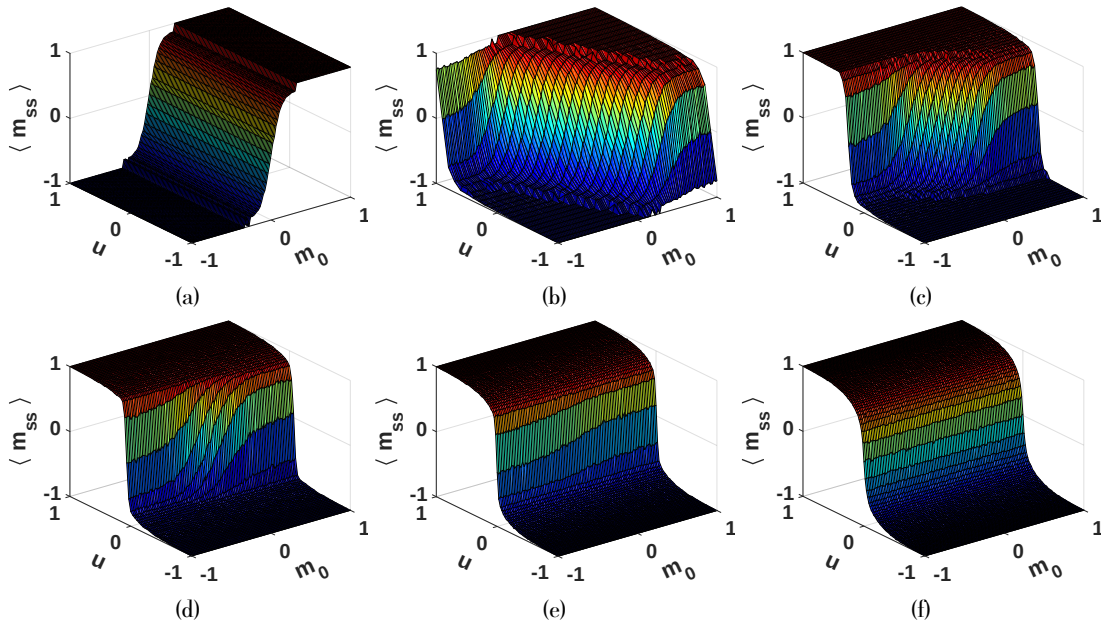

FIG. 4: *Dependency of mean steady state magnetization on initial magnetization and control input for* $N = 100$, *(a)* $\lambda = 0$, *(b)* $\lambda = 0.2$, *(c)* $\lambda = 0.4$, *(d)* $\lambda = 0.6$, *(e)* $\lambda = 0.8$, *(f)* $\lambda = 1$

is interesting to note that when the influence size (I) exactly matches the population size (N) the steady state PDF completely loses its dependency on the initial condition.

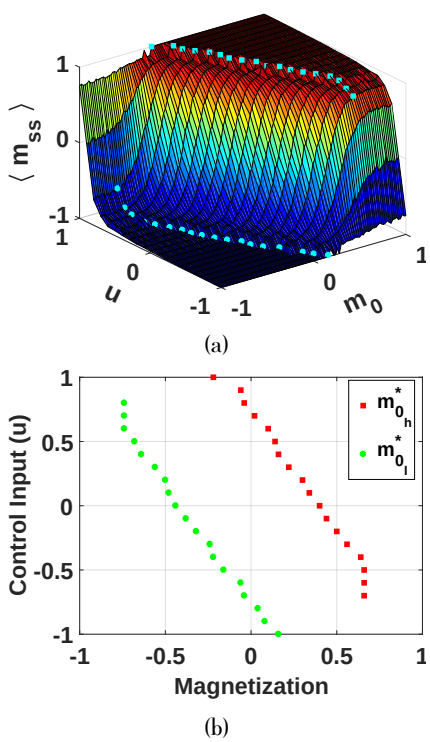

FIG. 5: *(a)* $m^*_{0_l}$ *and* $m^*_{0_h}$ *marked on the 3D graph related to* $\lambda = 0.2$ *for all possible values of* u. *(b) 2D representation of the boundaries on the* $m_0 - u$ *plane.*

Comparison to previous studies: In the study by Slanina and Lavička [7], it is proven that for a simplified Sznajd model dynamic on a complete graph, there exists a phase transition at $m_0 = 0$. However, in our simulations this phase transition is absent in Fig. 4. In the case of $\lambda = 0$, this is due to the fact that, in their analyses, Slanina and Lavička assume the population size is infinite, $N \to \infty$; however, in our simulations we assume the population is large but finite. In cases where $\lambda \neq 0$, the absence of a phase transition is partly because of the large but finite N assumption, and in part because of the presence of influences.

Linear scaling property: In section IV A, the linear scaling property with a fixed control input is mentioned. In this section, we investigate if this property holds true for other values of control input. Fig. 6 compares the behavior of $\langle m_{ss} \rangle$ for two population sizes $N = 50$ and $N = 100$ ($\eta = 2$). From visual inspection, linear scaling of simulation parameters preserves the general trend of $\langle m_{ss} \rangle$.

However, point-to-point comparison of $\langle m_{ss} \rangle$ data shows that the relation $\langle m_{ss}(N, I_+, I_-) \rangle = \langle m_{ss}(\eta N, \eta I_+, \eta I_-) \rangle$ does not hold true for all values of u. To determine the interval of control input for which the scalability property stands, we study the point-to-point difference in $\langle m_{ss} \rangle$ data values. Fig. 6c represents $\delta = \langle m_{ss} \rangle|_{N=50} - \langle m_{ss} \rangle|_{N=100}$ calculated on planes $m_0 = -1$ (blue line) and $m_0 = +1$ (red line).

It can be seen that for some values of u, $\langle m_{ss} \rangle$ of the scaled system ($N = 100$) and the base system ($N = 50$) are exactly equal ($\delta = 0$); thus, for the corresponding m_0 scalability property stands. For instance, if $\lambda = 0.2$, for

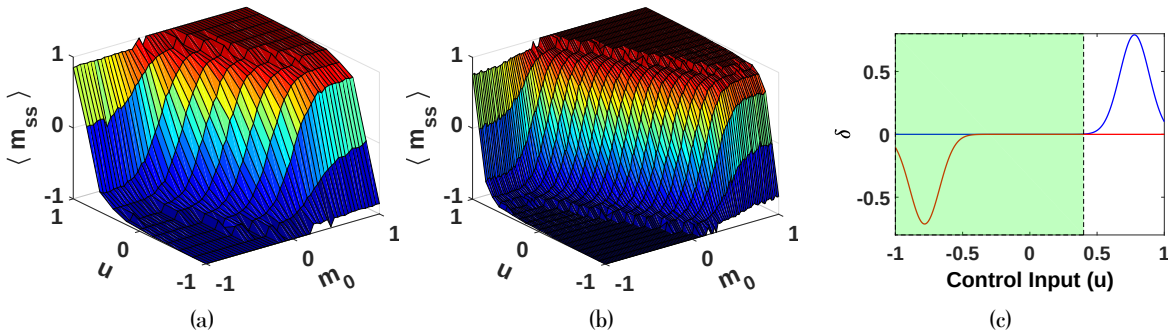

FIG. 6: *Three dimensional linear scaling property, (a) the base system ($N = 50, \lambda = 0.2$), (b) the scaled system ($N = 100, \lambda = 0.2$), (c) Scalability zone for $m_0 = \pm 1$ with $\lambda = 0.2$*

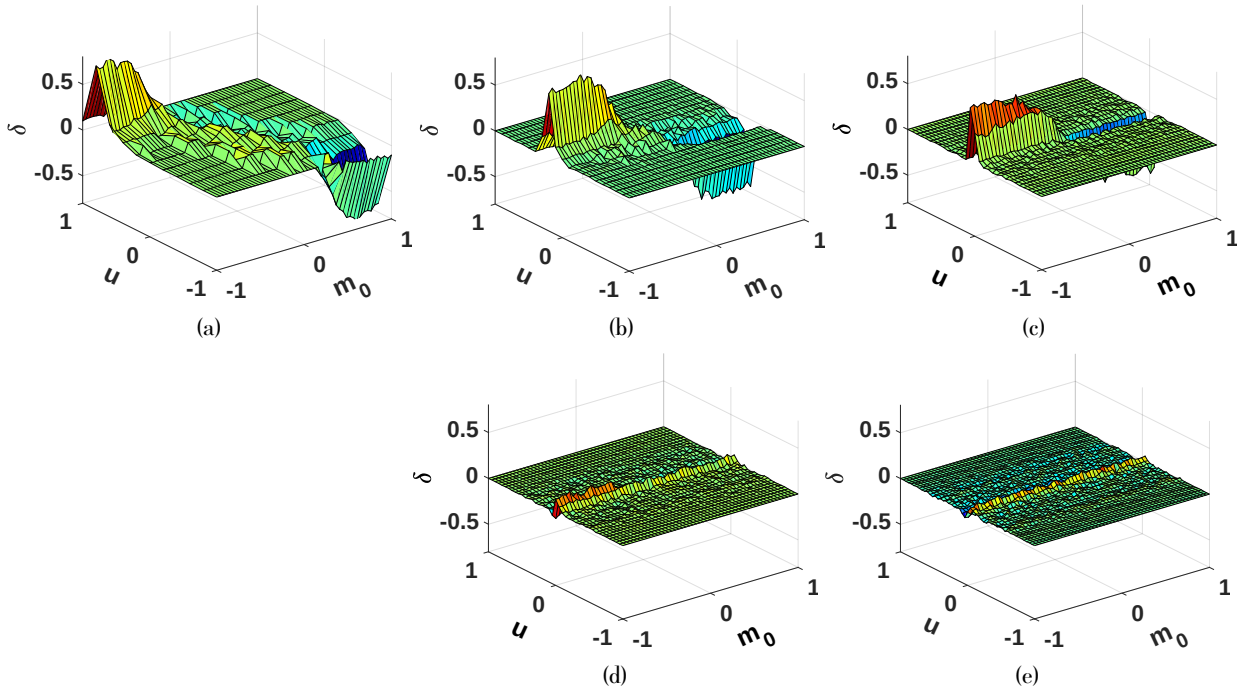

FIG. 7: *Scalability zone for all m_0s for $N = 100$, (a) $\lambda = 0.2$, (b) $\lambda = 0.4$, (c) $\lambda = 0.6$, (d) $\lambda = 0.8$, (e) $\lambda = 1$*

$\forall u \in [-1, 0.4], \delta = 0$ on plane $m_0 = -1$ (the green area). For a unique m_0, intervals of control input for which $\delta = 0$ will be referred to as the **Scalability Zone** of the base system for that m_0.

To study the existence of scalability zone for different values of initial magnetization, the same process has been applied to $\langle m_{ss} \rangle$ data on other available m_0 planes. Fig. 7 presents $\delta(m_0, u)$ for different values of influence ratio. Interesting characteristics are observed for different values of λ regarding existence of scalability zone and its expansion, maximum difference values, and sign of δ.

It can be seen in Fig. 7a that for some values of m_0 close to zero $\delta \neq 0$. However, scalability zone appears and expands in the first and third quadrant of the $m_0 - u$ plane (where $sgn(m_0) = sgn(u)$) towards outer borders. In this figure, dependency of δ on m_0 is clear in a large area. Also, maximum difference happens in the second and fourth quadrants near the outer borders (where $sgn(m_0) = -sgn(u)$).

As λ increases, the dependency of δ on m_0 decreases; maximum differences decrease, and they happen close to $u = 0$ plane. The scalability zone is available for all values of m_0. When $\lambda = 1$, δ is a function of u only. Interestingly, $sgn(\delta) = sgn(u) = -sgn(m_0)$ for all values of λ. This means that for $m_0 \geq 0, \delta < 0 \Rightarrow \langle m_{ss} \rangle|_{N=50} < \langle m_{ss} \rangle|_{N=100}$ if $\delta(m_0, u) \neq 0$, or for $m_0 \leq 0, \delta > 0 \Rightarrow \langle m_{ss} \rangle|_{N=50} > \langle m_{ss} \rangle|_{N=100}$ if $\delta(m_0, u) \neq 0$.

So far, our findings are dominantly based on the mean of the probability distribution function of the system at its steady state; however, while calculating the mean, some information of the PDF is lost. For instance, although a family of Gaussian distributions defined as $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2}{2\sigma^2}}$ have a mean value of zero, their disorder or uncertainty cannot be explained by their mean. In information theory the

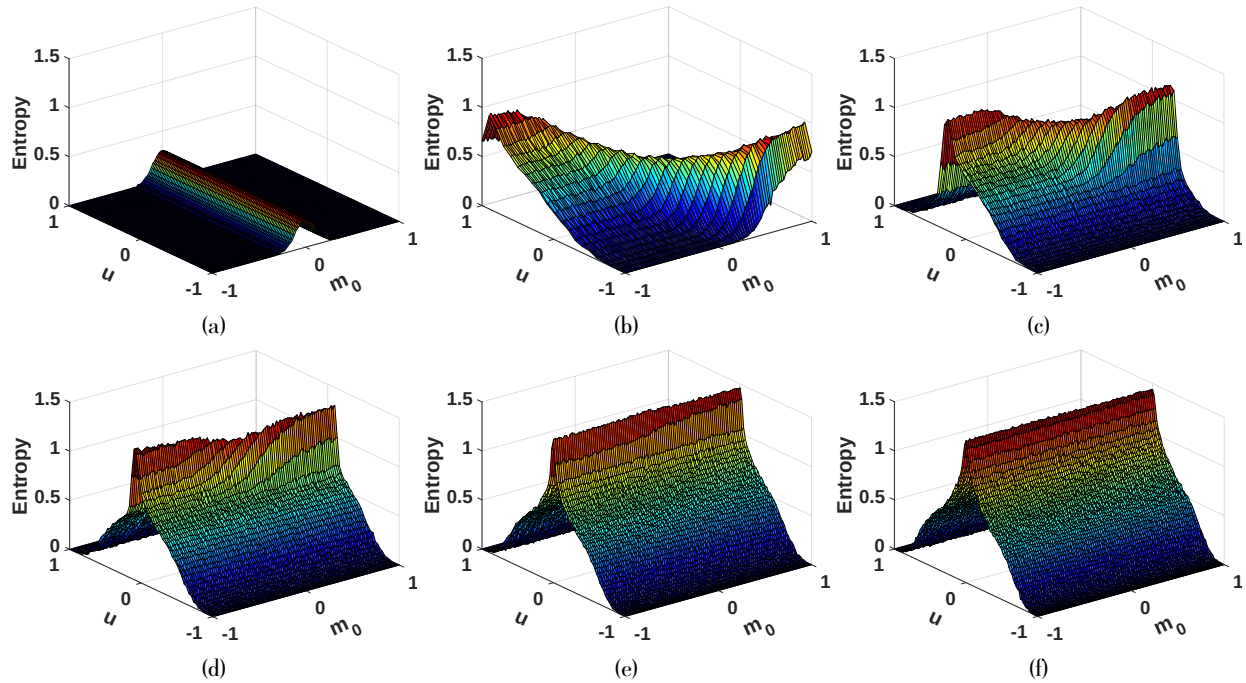

FIG. 8: *Entropy of distributions for different values of* λ *for* $N = 100$, *(a)* $\lambda = 0$, *(b)* $\lambda = 0.2$, *(c)* $\lambda = 0.4$, *(d)* $\lambda = 0.6$, *(e)* $\lambda = 0.8$, *(f)* $\lambda = 1$

concept of *Entropy* is used to study the uncertainty of PDFs. Next section is devoted to the study of system entropy at steady state.

C. Entropy Analysis

Entropy of a random variable X is defined as the expectation of the random variable $-\log P(X)$ with respect to the probability measure P [30, 31]:

$$\begin{aligned} H(X) =& E_{P_X}\{-\log P_X(X)\} \\ =& -\sum_{x\in X} P_X(x)\log P_X(x) \end{aligned} \tag{9}$$

where, E_P denotes the expectation with respect to probability distribution P. We use the steady state magnetization PDF of the systems in Fig. 4 to calculate entropy of each data point. Fig. 8 represents the entropy of the system for different values of influence ratio. It is worth noting that low entropy corresponds to higher certainty and high entropy represents lower certainty in the steady state.

For $\lambda = 0$, there are no influences in the population, and entropy of the system is a function of m_0 only. Entropy is zero for $m_0 \leq m^*_{0_l}$ or $m_0 \geq m^*_{0_h}$ because in these regions there are enough agents at the corresponding state to push the system PDF towards a Dirac delta function at either -1 for $m_0 \leq m^*_{0_l}$ or at $+1$ when $m_0 \geq m^*_{0_h}$. Also, entropy reaches its maximum at $m_0 = 0$, because in this region neither of the groups in ± 1 states are strong enough to completely attract the population towards themselves; this phenomenon results in a more distributed density function over the magnetization axis, and consequently higher entropy.

Figures 8b-f depict that adding influences to the population increases the entropy of the system. The reason is that although small in number, influences are capable of attracting agents to their state; as a result, the steady state magnetization of the system will have a different value from that of the uninfluenced population. So, the density function of the system becomes more distributed over the magnetization axis leading to higher entropy values.

For instance, $u = 0|_{\lambda=0.2}$ represents a system with influences equally divided between the two groups; this might imply that their effects on the steady state PDF will be eliminated by each other, and the resulted PDF will be the same as that of the uninfluenced system ($\lambda = 0$). However, by point-to-point comparison of data points we find that $\forall m_0$ on the $u = 0$ plane, the PDF of the influenced system is more distributed, and $H(m_0, u = 0)|_{\lambda=0} < H(m_0, u = 0)|_{\lambda=0.2}$.

In Fig. 8b, entropy is at its lowest levels when $sgn(m_0) = sgn(u)$. In this situation higher number of influences support the initially more populated state resulting in a sharp less-distributed magnetization density function close to one of the ends of the magnetization axis. The opposite happens when $sgn(m_0) = -sgn(u)$ since more influences are in the favor of the initially less populated state preventing the population from clustering at one of the edges of the magnetization axis; the resulting density function is nicely distributed causing higher entropy. This reasoning explains why the point $(m_0 = 0, u = 0)$ is similar to a saddle point.

It is observed in Fig. 8b that entropy is more sensitive to the control input rather than the initial condition. Consider point $(m_0 = -1, u = -1)$ on the graph. Entropy starts

to propagate from zero faster if $\{m_0 = const.$ and $u \uparrow\}$ compared to the situation where $\{u = const.$ and $m_0 \uparrow\}$. The reason is that influences (even in small numbers) are always active in the process of changing agents' states to theirs. However, an agent is active in attracting other agents to its initial state till s/he changes to the opposite state; i.e. its effect is not ever-lasting. As a result, more agents are needed to have the same effect of a few number of influences on the PDF of the system and the entropy.

In Fig. 8b, as u increases on $m_0 = -1$ plane, entropy ascends till $u = 0.9$. This phenomenon is easy to comprehend since more influences in state $+1$ attract more agents from state -1, and the PDF of the system is more distributed. However, at $u = 1$, since all the influences are at the positive side, they are able to dramatically shift the PDF towards themselves and make it narrower resulting in lower entropy value. Here, the number of influences is not high enough to change the PDF into a Dirac delta function, so entropy does not reach zero.

On the contrary, in Figures 8c-f, number of influences is high enough to completely attract the population towards themselves on planes $u = \pm 1$; so, the entropy is zero. As λ increases, similar to $\langle m_{ss} \rangle$, for larger intervals of u entropy loses its dependency on m_0 and $sgn(m_0)$, and the saddle point disappears; for $\lambda = 1$, entropy depends on the control input only.

V. CONCLUSION

In this paper, effects of *external influences* on the behavioral dynamic of a group of agents who interact with each other based on the Sznajd model are studied. The Sznajd model is formulated on a complete graph in presence of influences, and the governing differential equation for the population behavior at steady state is derived. The resulted ODE shows the dependency of $p_e(m)$ on population size N, control input $u = \frac{I_+ - I_-}{I}$, and total influence size I.

The mathematical model is numerically solved, and the results are compared with data from numerical simulations. Higher peak values and lower variances are observed for larger population sizes. We find good agreement between the results from numerical experiments and numerical solution of the mathematical model for network sizes exceeding $N = 100$.

Based on numerical simulations, regions (called convergence zone) are available where the steady state loses its dependency on the initial condition. By adopting Kolmogorov-Smirnov hypothesis testing method, it is proven with 99% certainty that steady state PDF's of systems with initial conditions belonging to the same convergence zone are equivalent. Different graphs are provided to display this phenomenon. A relationship based on these findings and the general solution of the stationary ODE is drawn for different ranges of m_0. Furthermore, results show that by increasing the population size, the convergence zone covers a larger area. Interestingly, it is observed that linear scaling of simulation parameters causes the same convergence value for the mean steady state magnetization.

Effect of the control input on convergence zones is studied on three dimensional graphs by defining a new parameter $\lambda = \frac{I}{N}$ named *influence ratio*. Results show that for different combinations of u and λ, convergence zone(s) exist. Figures are provided to show the boundaries of the convergence zones. It is shown that increase in the number of influences results in larger convergence zones. When $\lambda = 1$, the whole $m_0 - u$ plane is independent of m_0. In addition, the absence of a phase transition is explained by the large-but-finite population size assumption and presence of influences.

Linear scaling property is also investigated for all combinations of m_0 and u. It is shown that not for all values of the control input the mentioned property stands. Three dimensional figures are provided which show the dependency of the difference between the $\langle m_{ss} \rangle$'s of two linearly scaled systems for different λ's. In general, as influence ratio increases, δ decreases; and δ loses dependency on m_0. The reason behind high values of δ on the second and fourth quadrants of the $m_0 - u$ plane is discussed in detail. It was also shown that for negative initial conditions, the $\langle m_{ss} \rangle$ of the base system is higher than that of the scaled system.

Entropy study of the system reveals that higher influence (although small in number) equals higher entropy. Generally, entropy is a function of the initial magnetization and the control input, but as λ increases, the system loses its dependency on m_0. When $\lambda = 1$ the entropy is only a function of the control input.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support from NSF CIS grant number 1538139.

[1] H. Blumer, American sociological review **19**, 3 (1954).
[2] M. A. Hogg, in *Understanding Peace and Conflict Through Social Identity Theory* (Springer, 2016) pp. 3–17.
[3] J. T. Jost, M. R. Banaji, and B. A. Nosek, Political psychology **25**, 881 (2004).
[4] L. K. Hansen, A. Arvidsson, F. Å. Nielsen, E. Colleoni, and M. Etter, Future information technology , 34 (2011).
[5] G. Cybenko, A. Giani, and P. Thompson, Computer **35**, 50 (2002).
[6] F. Vazquez *et al.*, New Journal of Physics **10**, 063011 (2008).
[7] F. Slanina and H. Lavicka, The European Physical Journal B-Condensed Matter and Complex Systems **35**, 279 (2003).
[8] J. Lorenz, International Journal of Modern Physics C **18**, 1819 (2007).
[9] M. Mobilia, Physical review letters **91**, 028701 (2003).

[10] M. Mobilia, A. Petersen, and S. Redner, Journal of Statistical Mechanics: Theory and Experiment **2007**, P08029 (2007).

[11] A. R. Srinivasan and S. Chakraborty, in *2014 American Control Conference* (IEEE, 2014) pp. 2096–2101.

[12] K. Sznajd-Weron and J. Sznajd, International Journal of Modern Physics C **11**, 1157 (2000).

[13] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior* (Princeton university press, 2007).

[14] M. Friedman and L. J. Savage, Journal of political Economy **56**, 279 (1948).

[15] L. J. Savage, *The foundations of statistics* (Courier Corporation, 1972).

[16] M. Fishbein and I. Ajzen, (1977).

[17] F. S. N. Karan and S. Chakraborty, in *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation* (Springer, Cham, 2017) pp. 182–192.

[18] W. E. Huffman, American Journal of Agricultural Economics **56**, 85 (1974).

[19] D. A. Briley, M. W. Morris, and I. Simonson, Journal of consumer research **27**, 157 (2000).

[20] M. B. Brewer and R. M. Kramer, Journal of personality and social psychology **50**, 543 (1986).

[21] P. Nyczka, J. Cisło, and K. Sznajd-Weron, Physica A: Statistical Mechanics and its Applications **391**, 317 (2012).

[22] P. Nyczka, K. Sznajd-Weron, and J. Cisło, Physical Review E **86**, 011105 (2012).

[23] A. Jędrzejewski, Physical Review E **95**, 012307 (2017).

[24] M. Mobilia, Physical Review E **92**, 012803 (2015).

[25] F. S. N. Karan and S. Chakraborty, IEEE Transactions on Computational Social Systems **3**, 13 (2016).

[26] F. S. N. Karan and S. Chakraborty, "Dynamics of a repulsive voter model," in *Social Computing, Behavioral-Cultural Modeling, and Prediction: 8th International Conference, SBP 2015, Washington, DC, USA, March 31-April 3, 2015. Proceedings*, edited by N. Agarwal, K. Xu, and N. Osgood (Springer International Publishing, Cham, 2015) pp. 428–433.

[27] R. Dunbar, *How many friends does one person need?: Dunbar's number and other evolutionary quirks* (Faber & Faber, 2010).

[28] A. Smith, "6 new facts about facebook," Pew Research Center, Washington, D.C., http://pewrsr.ch/1dm5NmJ (2013).

[29] R. Lambiotte and S. Redner, EPL (Europhysics Letters) **82**, 18007 (2008).

[30] K. Kobayashi *et al.*, *Mathematics of information and coding*, Vol. 203 (American Mathematical Soc., 2007).

[31] W. Wang, A. Taalimi, K. Duan, R. Guo, and H. Qi, in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)* (2016) pp. 1–9.