Predicting bursting in a complete graph of mixed population through reservoir computing

Suman Saha ,[1,2,*] Arindam Mishra ,[1,*] Subrata Ghosh,[3] Syamal K. Dana ,[1,4] and Chittaranjan Hens [3,†]

[1]*Centre for Mathematical Biology and Ecology, Department of Mathematics, Jadavpur University, Kolkata 700032, India*
[2]*Complex Analysis Group, Translational Health Science and Technology Institute, NCR Biotech Science Cluster, Faridabad 121001, India*
[3]*Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India*
[4]*Division of Dynamics, Faculty of Mechanical Engineering, Lodz University of Technology, 90-924 Lodz, Poland*

We report our investigation and success story, to an extent, on the prediction of spiking and bursting dynamics in globally coupled networks, using echo state network/reservoir computing-based learning procedure. Two exemplary dynamical models, Josephson junctions and Hindmarsh-Rose neurons, are used to construct two separate networks and thereby illustrate the efficacy of our strategy. In the absence of coupling, the networks consist of mixed populations in which few nodes are oscillatory (self-sustained spiking) and the rest of the nodes maintain a quiescent state. When single-input data from one oscillatory node of a network (under stronger interactions between the nodes) is used for learning, the ESN is able to predict the key dynamical features (spiking and bursting) of the other nodes. In comparison, the machine performs with improved predictions if it is fed with two inputs: one from the oscillatory population and another from an excitable population. The machine's leaking parameter plays a crucial role, which can be tuned appropriately to enhance prediction. Furthermore, a cluster synchronization in the mixed population is confirmed from the machine-generated output signals. Our work is expected to be useful as a burst predictor.

I. INTRODUCTION

A recurrent neural network (RNN), a bio-inspired machine learning tool, offers a general solution for complex tasks of predicting signal information. Noticeably, the working principle of RNN does not depend on *a priori* knowledge of the intrinsic dynamical relation of the considered or targeted data [1–3]. Recently another version of RNN called as reservoir computing (RC)/echo state network (ESN) is developed for efficient prediction of complex signals for a considerably longer time [3–8]. In contrast to RNN, the ESN is easier to implement and cost effective since it does not require fine tuning of its inner components except for the readout/output layer which helps matching the target behavior within a close approximation. Having its simplicity of computation and powerful ability of prediction, ESN has been used for calculating Lyapunov exponents, attractor reconstruction [9–11] of dynamical systems and becomes a testing bed for detecting generalized synchronization [12–14] in coupled chaotic systems. Apart from dynamical issues, ESN can identify nonstationarity in steady-state visual evoked

potentials [15], predict stock price in a short time scale [16] and help understand the language processing [17] as well as differentiating speech signals [18]. A reservoir (say, ESN) has three distinct components or layers: an input layer collecting the inputs, a reservoir network with a large number of random elements (analogous to neurons) that expands the input in a high dimensional nonlinear fashion and an output layer to produce the expected target. The readout or output layer is the only part where the weights are trained to produce a desired output which should be closer to the target data. Researchers have also devoted to find the optimal parameters of an ESN for accurate detection of target data [13,19–25].

Here we first employ the ESN to predict different kinds of bursting dynamics originating from a mixed population of Josephson junctions [26], where excitable and oscillatory junction nodes (subpopulations) connect to each other over the top of a complete graph. We test here whether the ESN can, (i) predict the spiking and bursting signals with considerably high accuracy, (ii) how many inputs required for efficient prediction of spiking and bursting dynamics, and (iii) identify cluster synchronization within the subpopulations. From this perspective, we further investigate the role of the leakage parameter of the ESN on the success of prediction. We have also checked how the machine performs in presence of heterogeneity of system parameters and in the desynchronized regime of the dynamical network under learning and prediction (details are given in Ref. [34]). In a broader sense, the paper aims at an optimal and suitable strategy in the machine learning process through ESN for prediction of spiking and bursting dynamics as well as the collective behavior of coupled networks. Note that spiking is a repeated firing state that occurs when neurons
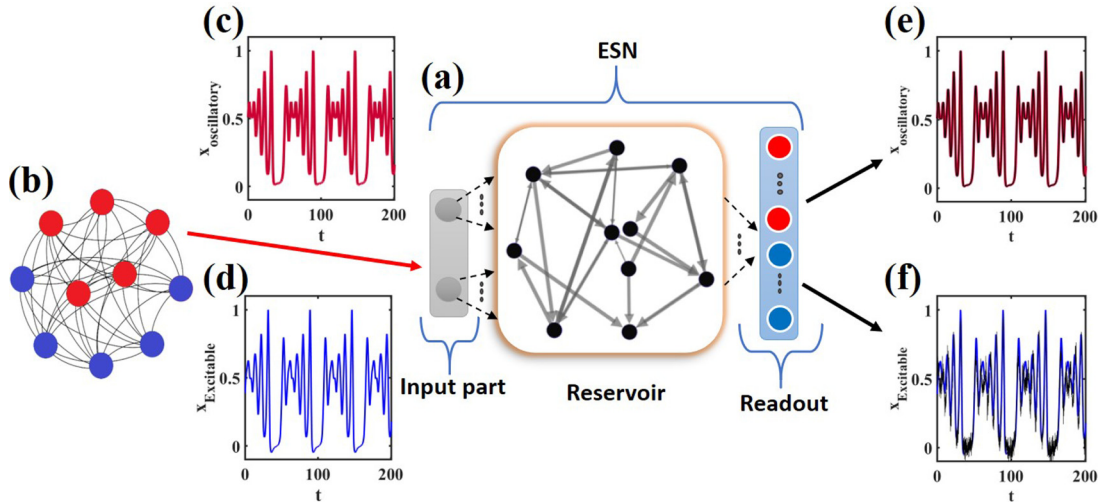
FIG. 1. Bursting prediction by ESN. (a) Components of the machine. (b) A complete graph of spiking (filled red circles) and quiescent oscillators (filled blue circles). [(c) and (d)] For strong mean-field coupling ($\epsilon = 15$), both the subpopulations (oscillatory nodes in red color, excitable nodes in blue color) emerge into bursting oscillation: Six spikes with different amplitudes in each burst. [(e) and (f)] Predicted signals (black lines) by the machine after feeding only one oscillatory input into the machine. Bursting of one oscillatory node (red line) is shown closely matching with machine data (black line). Bursting in the excitable unit (blue line) is also matched with the machine predicted signal (black line). However, there is a small fluctuation in the subthreshold oscillation of the predicted signal near the slow manifold, which shows a mismatch or poor prediction.

send signal to other neurons through a rise of action potential [27–30], whereas bursting is a bunch of spikes, i.e., a repeated firing within a small duration of time intercepted (periodically or irregularly) by a silent or steady state. Finally, we test our scheme with a network of biologically plausible spiking and bursting Hindmarsh-Rose (HR) neuron model [31,32]. For both of our example networks, the ESN efficiently predicts the critical point of emergence of the clustering states.

II. BASICS OF ESN BASED LEARNING

The ESN described, in this article, is a standard discrete-time leaky tanh (tan-hyperbolic function) network. The internal state of each node of the reservoir updates themselves following a recurrent relation:

$$\mathbf{r}(n+1) = (1-\alpha)\mathbf{r}(n) + \alpha \times \tanh(\mathbf{W}_{\text{res}}\mathbf{r}(n) + \mathbf{W}_{\text{in}}\mathbf{s}(n+1)), \tag{1}$$

where $\mathbf{W}_{\text{res}}$ is a $N \times N$ matrix of internal connectivity weights of the reservoir shown in Fig. 1(a). The matrix $\mathbf{W}_{\text{in}}$ (size: $N \times K$) represents input weights and K is determined by the number of external inputs. $\mathbf{r}$ is an N-dimensional vector representing instantaneous internal state dynamics of the reservoir and $\mathbf{s}$ is the K-dimensional input vector.

The tanh function is applied element-wise and α is a leakage constant. In the ESN, the elements of $\mathbf{W}_{\text{res}}$ and $\mathbf{W}_{\text{in}}$ are kept unchanged before or at the instant of post-processing. To train the readout/output weights, the states from the reservoir $\mathbf{r}(\mathbf{n})$ and external input $\mathbf{s}(\mathbf{n})$ are collected in $\mathbf{X}(n) = [\mathbf{r}(n); \mathbf{s}(n)]$ for each instant of time n. The output relation at time n is captured by

$$\mathbf{y}(n+1) = \tanh(\mathbf{W}_{\text{out}}\mathbf{X}(n)). \tag{2}$$

For a long time series $n = 1, .., Q$, we can write the relation in a vector form [5]

$$\mathbf{Y} = \tanh(\mathbf{W}_{\text{out}}\mathbf{X}). \tag{3}$$

Here, $\mathbf{Y} \in R^{K_o \times Q}$ where Q is the length of the time series of each node and K_o is the number of target nodes/outputs. The size of the collecting matrix of the internal dynamics is given by $\mathbf{X} \in R^{(N+K) \times Q}$. The most simplest way to determine $\mathbf{W}_{\text{out}}$ is to take the pseudoinverse of the operator $\mathbf{X}$:

$$\mathbf{W}_{\text{out}} = \tanh^{-1}(\mathbf{Y})\mathbf{X}^{\dagger}. \tag{4}$$

However, the most suitable way of capturing $\mathbf{W}_{\text{out}}$ is the least-square method [33], which looks like

$$\mathbf{W}_{\text{out}} = (\mathbf{X}^T\mathbf{X} + \lambda I)^{-1}\mathbf{X}^T(\tanh^{-1}\mathbf{Y}), \tag{5}$$

where λ is a regularization factor and $\mathbf{Y}$ is the true/target data and I is the identity matrix. In the absence of regularization, the system converges to the previous equation (4) (see Ref. [34], Sec. 1 for details). Note that, the output nonlinearity tanh is optional. To make a random reservoir qualify as an echo state network, it must exhibit certain damping properties. They can be ascertained by ensuring that the weight matrix $\mathbf{W}_{\text{res}}$ has a spectral radius smaller than unity [33] which has been maintained throughout this article for each simulation.

A. Network setup

To construct the ESN, a 1000×1000 reservoir weight matrix $\mathbf{W}_{\text{res}}$ is generated with dense connectivity and random weights are drawn from a uniform distribution over $(-1, 1)$ and re-scaled to a spectral radius of 0.8. Input weights ($\mathbf{W}_{\text{in}}$) are also randomly chosen from a uniform distribution over $(-1, 1)$. One input from the mixed population (or two qualitatively different inputs) will be fed into the reservoir. Therefore, the size of $\mathbf{W}_{\text{in}}$ is either $N \times 1$ or $N \times 2$ (see

Ref. [34], Sec.1). The leakage constant α is fixed at 0.3, however, its effect on prediction is tested and described below for a range of values.

B. Training and testing

The length of the training sequence is taken as 40 000; the first 10 000 steps are discarded to wash out the initial transient and the echo signals $\mathbf{x}(n)$ are sampled from the remaining $Q = 30\,000$ steps. After the training process, 20 000 data points are tested in the backdrop of the learned machine. The machine extracted and target data are represented by $d_i(n)$ and $y_i(n)$ ($n = 1, \ldots, 20\,000, i = 1, \ldots, K_o$), respectively, when the mean squared error (MSE) is calculated by $\text{MSE}_{\text{test}_i} = \frac{1}{20000} \sum_{n=1}^{20\,000} (d_i(n) - y_i(n))^2$.

To illustrate the success of ESN on the prediction of bursting, we employ globally coupled networks of spiking and excitable (in quiescent state) oscillators shown in Fig. 1(b), where spiking oscillators are represented by red circles and quiescent nodes by blue circles.

Our motivation is to extract/predict information (time signal) of a large number of nodes by feeding information of the dynamics (time series data) of a small number of nodes as input to the reservoir. Assume that we use the input from one oscillatory node ($x_{\text{oscillatory}}$), then $\mathbf{s}(n)$ is one-dimensional. In each time step, this value is multiplied with the element of input matrix $\mathbf{W}_{\text{in}}$ and then added to the intrinsic dynamics of the reservoir described by Eq. (1). In the same manner, we can add, multiple inputs to the reservoir (see Ref. [34], Sec. 2 for detail description). Before describing details of our main results of our selected networks, we present a glimpse of our strategy and the story of prediction on the dynamics of the Josephson junction network [26]. For a larger coupling strength, all the nodes show bursting behavior in Figs. 1(c)–1(d). However, data of bursting dynamics from only one oscillatory node shown in Fig. 1(c) are fed into the ESN for training and the rest of the nodes' dynamics are predicted from the trained machine. Interestingly, bursting of the oscillatory (red) nodes and the bursting response of the quiescent nodes (blue) are predicted as shown in Figs. 1(e) and 1(f), respectively, with a good approximation. Further details are presented in the next section.

III. PREDICTION OF BURSTING AND CLUSTERING: NETWORK OF JOSEPHSON JUNCTIONS

We consider here a globally coupled network; the local dynamics of each node is represented by resistive-capacitive-shunted junctions (RCSJ) [35–39]. This superconducting Josephson junction model is an analog of the classical pendulum model. Above a critical DC-bias current and for a wide range of damping parameter, an isolated RCSJ shows self-oscillation mimicking the periodic spiking behavior alike neurons' spiking [37]. In presence of a periodic forcing or a shunted inductor, the RCSJ may reveal typical spiking and bursting behaviors [27,36,37,40]. An array of forced RCSJs may also reveal complex collective behavior (chimera, clustering, and extreme events) [41,42] under a global mean field interaction. As suggested above, we consider here a mixed population of RCSJs (total number of junctions is N_d), where

p number of oscillators are in self-oscillatory mode and the rest of ($N_d - p$) oscillators are in a quiescent state [26]. The dynamics of two subpopulations of oscillatory and excitable nodes in the network is captured by two sets of equations,

$$\dot{\theta}_i = x_i, \tag{6}$$

$$\dot{x}_i = I_i - a_i x_i - \sin \theta_i + \frac{\epsilon}{N_d} \sum_{k=1}^{N_d} (x_k - x_i), \tag{7}$$

$$\dot{\theta}_j = x_j, \tag{8}$$

$$\dot{x}_j = I_j - a_j x_j - \sin \theta_j + \frac{\epsilon}{N_d} \sum_{k=1}^{N_d} (x_k - x_j), \tag{9}$$

where $i = 1, 2, \ldots, p$ and $j = p + 1, p + 2, \ldots, N_d$ are the self-oscillatory and excitable nodes, respectively. The variable θ_i of the ith uncoupled node is the phase difference of the junction, $\dot{\theta}_i = x_i$ (dot denotes time derivative) is the voltage across the junction and $a_i = [h/2\pi e I_i R^2 C]^{1/2}$ is the damping parameter, h is the Planck's constant, e is the electronic charge and $I_i (i \neq j)$ is the DC bias current. It is to be mentioned that the isolated RSCJ model has similarities with the Sakaguchi-Kuramoto phase model with inertia [43–45] and also with the working model of a power grid [46].

In the absence of coupling ($\epsilon = 0$) and $a = 1.5$, each node can reveal two types of dynamics: oscillatory spiking behavior ($I_i > 1.0$) and a quiescent state ($I_i < 1.0$). Below $I_i = 1.0$, the system has two equilibrium points: one saddle and a stable node. At $I_i = 1.0$, two fixed points coalesce and the junction generates a stable limit cycle through saddle-node on invariant circle (SNIC) bifurcation for $I_i > 1.0$ [38,39]. Under an external periodic forcing, an isolated junction shows *a parabolic bursting* [36,37,40] typically seen in neuron models [47]. Here we consider a network of junctions with a subpopulation of oscillatory nodes by setting the bias current at $I_1 = I_2 = \cdots = I_p = 1.25$ and another subpopulation as excitable (in quiescent state) by setting $I_{p+1} = \ldots = I_{N_d} = 0.5$ with our choice of fixed $a_i = 1.5$. It has already been established [26] that such a mixed population in a globally coupled graph, originates bursting dynamics whose pattern is determined by the fraction of population p/N_d and the coupling strength ϵ. Now we present the details of how a trained ESN can proceed to predict the bursting behavior of the network nodes. We consider the network size $N_d = 10$, where half ($p = 5$) of the population is excitable and the rest is oscillatory. To understand the spiking-bursting behavior, we have plotted two bifurcation diagrams in Fig. 2 where maxima of the spikes (in a burst) of one oscillatory node (in red dots) in Figs. 2(a) and maxima of one excitable node (in blue dots) in Fig. 2(b) are drawn for varying coupling ϵ.

For weak coupling (left of the dashed vertical lines), the original oscillatory nodes exhibit recurrent spiking oscillation and the excitable nodes show subthreshold oscillation. The number of spikes in a single burst emerges one by one above a critical coupling strength ($\epsilon_c = \epsilon \sim 4.8$). For instance, both the nodes will have four spikes (blue and red dotted lines) in a single burst at a coupling $\epsilon = 9$ (black vertical line). First, we feed time series information from one oscillatory node into the reservoir for training. The rest of the nodes will be
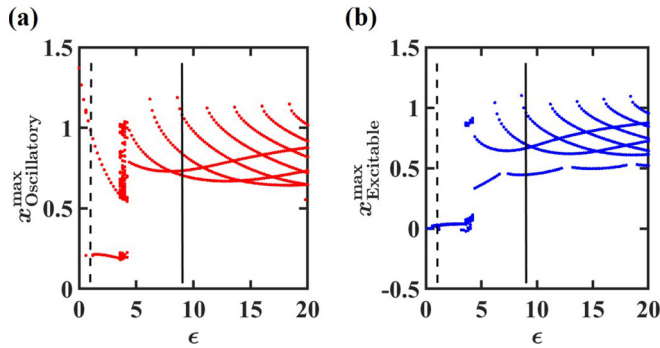
FIG. 2. Bifurcation diagram of oscillatory and quiescent nodes. x_i variable of the RCSJ is used for study. Peak values are plotted as a function of coupling strength (ϵ) for (a) one arbitrarily chosen oscillatory node and (b) one from the excitable subpopulation. For weak coupling ($\epsilon = 1$) indicated by a dashed vertical line, the oscillatory node has single and recurrent spikes (a). The quiescent node shows (b) subthreshold oscillation in this regime (dashed vertical line). For larger coupling ($\epsilon > 4$), both the nodes start bursting oscillations with multiple spikes (the solid vertical line at $\epsilon = 9$). The number of spikes in a burst increases with ϵ.

predicted after the training procedure is finished: 30 000 data points from one node are used to train the readout weights. After that, one oscillatory input of 20 000 data length (points) is used to predict 20 000 data points of rest of the nodes. We examine that the ESN predicts the signal with a low accuracy when the network is under weak interactions ($\epsilon \sim 1$). The original signals extracted from the complete graph are shown

in Fig. 3(a). The time series of four oscillatory nodes (in solid, dashed, dotted, and dotted-dashed red lines) are shown (other than the input node), which are asynchronous. However, the subthreshold oscillations of low amplitude (time series in blue) of the excitable five nodes are synchronized, which is further confirmed by a spatiotemporal plot in Fig. 3(b) (five nodes, 6–10, at right). The readout of the ESN is tuned at the time of the training process by information from one oscillatory unit only. In the post-process, the predicted signals (of 9 nodes leaving the input node) in Figs. 3(c) and 3(d) do not match the original signals [cf. Figs. 3(a) and 3(b)]. It explains a weak correlation between the original and the predicted signals. This happens because the original oscillatory signals are not in a synchronous state and, therefore, the machine fails to predict the signals of the other nodes. By contrast, for a higher coupling ($\epsilon = 15$), the oscillatory and excitable units maintain a correlated evolution in time, basically, showing up an almost synchronized bursting state; time series and spatiotemporal plots in Figs. 3(e) and 3(f) provide a proof. The synchronized bursting signals (red lines) of the oscillatory nodes and the excitable nodes (blue line) are shown in the time series plot. Both the subpopulations are separately synchronized and the spikes within each burst are synchronized for the entire population (red and blue lines almost merging to each other). After the readout weights are trained, the ESN efficiently predicts the signals of the other nodes. Note that, time series data from only one oscillatory node is fed into the reservoir and the rest of the synchronized oscillatory nodes (red lines) of excitable nodes (blue lines) are predicted in Fig. 3(g) as already described in Figs. 1(e)
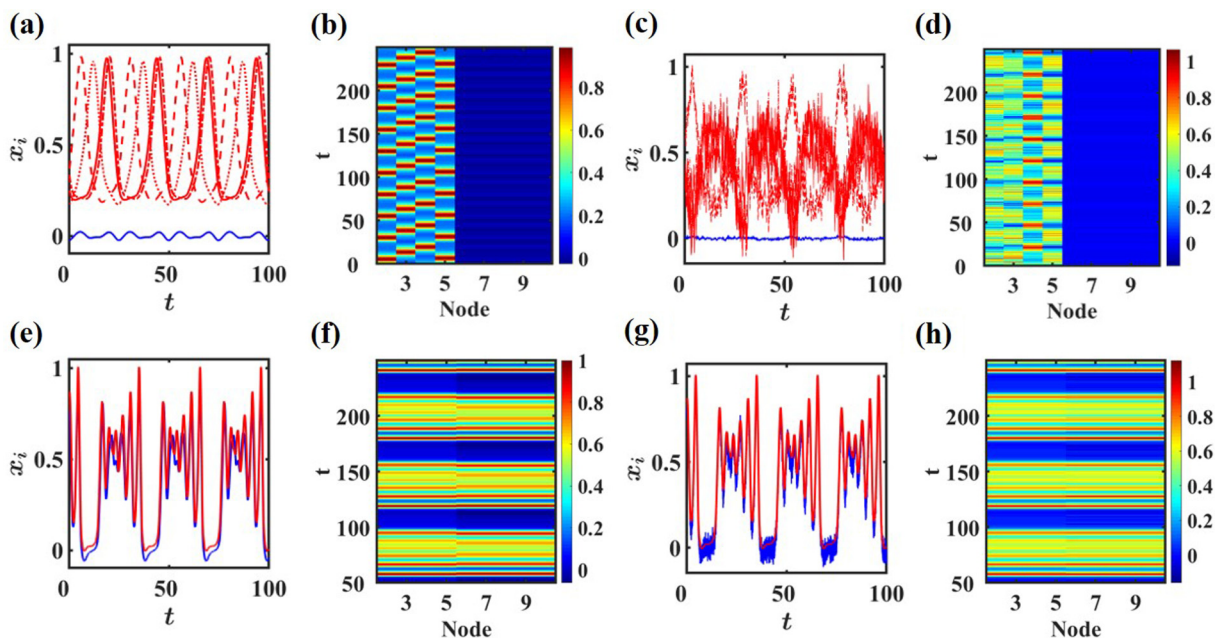

FIG. 3. Comparison of original signals with the machine predicted signals from RCSJ network. One input is used from one oscillatory node. Original time signal x_i (a) and spatiotemporal plot (b) of all nodes for weak coupling $\epsilon = 1$. Time series (c) and spatiotemporal plot (d) of predicted signal from the machine. No correlation is found between the machine predicted signals and original signals extracted from dynamical equations. For higher coupling $\epsilon = 15$, bursting appears in the entire network: time series x_i (e) and all the nodes are synchronized as shown in the spatiotemporal plot (f). Machine-based predictions of time evolution in blue (g) is more accurate than the weak coupling setting except for small fluctuations near the slow manifold. Cluster synchronization is also visible from the predicted signal in a spatiotemporal plot (h).
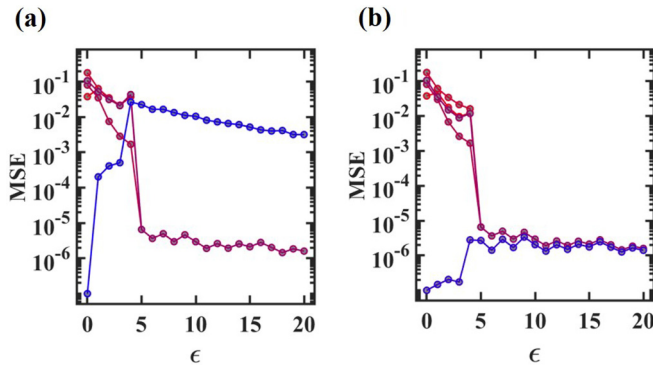
(a)   **(b)**

FIG. 4. MSE between the original and predicted signals in the RCSJ network. (a) One oscillatory node is fed into the machine. The red circled lines represent the oscillatory nodes. They merged near the coupling $\epsilon \sim 5$ and MSE of each oscillatory node sharply dropped to a value $\sim 10^{-5}$. The blue circled lines represent the excitable nodes. MSEs of excitable nodes are almost saturated at $\sim 10^{-2}$ at a higher coupling. (b) For better prediction, two nodes' information is fed into the machine: one oscillatory and another from one node in the excitable subpopulation. MSE for all the nodes sharply dropped near a coupling $\epsilon \sim 5$ confirming the efficiency of prediction of the ESN.

and 1(f). The quiescent nodes are showing some fluctuations at the subthreshold level. However, the bursting pattern (5 or 6 peaks in a single burst and timing of spiking) of each node in the entire graph is closely mimicked in the predicted output. To quantify the accuracy of the data set generated from the machine/reservoir, we calculate the MSE of all the predicted signals with respect to their original counterparts. MSE is plotted in Fig. 4(a) as a function of ϵ. All the red lines (with solid circles) converge around $\epsilon \sim 5$ and maintain synchrony within the population and MSE confirms that the RC-based signals of the oscillatory nodes are highly correlated with their original counterparts as it saturates at $\sim 10^{-5}$ beyond $\epsilon > 5$. At a lower coupling, the blue nodes (solid circles) are in steady state or show small oscillation; the ESN-detected signals also show subthreshold oscillation and hence the prediction for the excitable/quiescent nodes at a weaker coupling is highly accurate and accordingly, the MSE of all the excitable nodes are significantly small. However, when we increase the coupling, the amplitude of subthreshold oscillation of each excitable node becomes larger and larger with ϵ and hence the MSE of all the excitable nodes fluctuates for a range of ϵ, and finally converges to a relatively larger value, $\sim 10^{-2}$ at $\epsilon \sim 5$. Compared to the oscillatory nodes (red lines with filled circles), all the predicted signals of the excitable nodes (blue lines with filled circles) are less accurate for higher coupling (MSE $\sim 10^{-2}$, $\epsilon > 5$). The reason is as follows: Feeding data from one oscillatory node into the reservoir, we expect the ESN to predict the rest of the oscillatory and excitable nodes (total nine nodes). Since the input is from one oscillatory node, the machine will have a better sharing of information with the rest of the oscillatory nodes, at a larger coupling, hence it predicts the oscillatory nodes's behavior in a better way. However, with no direct information from the excitable nodes being shared with the ESN, it cannot accurately identify the signature of subthreshold oscillations of the excitable nodes and hence

prediction about their behavior is poor. For improving the performance of the ESN, we have increased the number of inputs to two when one input is taken from the oscillatory nodes and the other from the excitable nodes as shown in Fig. 4(b) to make a balance of information input. Now, it is seen that above a certain coupling $\epsilon > 5$, MSE of all the nodes (blue and red) coalesce together nicely and saturated at $\sim 10^{-4}$, reflecting the cluster states of the entire networks. For improved performance, we play with the leakage parameter and discuss in Sec. III A. Note that the integration has been performed in MATLAB using the ODE45 routine. At the time of starting the simulations, the initial conditions are chosen randomly between 0.1 to 2 and adiabatically updated for the next coupling. This strategy is kept unchanged for all data generated on the Josephson junction network, in the main text, as well as Ref. [34].

A. Prediction: Impact of leakage constant

The performance of the reservoir is now checked from the perspective of the machine parameter α, the leakage constant. So far we have derived our results using a leakage parameter $\alpha = 0.3$ for the ESN. We now play with this parameter α if it can improve the prediction both for single- and two-input learning. For comparison, we use the MSE measure of one input signal (separately, for excitable and oscillatory nodes) and its machine-generated signal both for the single and the double inputs, and plotted separately, as a function of ϵ and α in Fig. 5.

First we feed data from the time signal of one oscillatory node ($x_{\mathrm{Oscillatory}_1}$) into the reservoir. MSE of one excitable node ($x_{\mathrm{Excitable}_2}$) and its machine-generated counterpart is only shown in Fig. 5(a), which clearly shows that it decreases for larger $\alpha > 0.02$. The machine performance does not change significantly for larger α (including our selected specific value of 0.3) and ϵ. However, compared to the excitable node, the machine performs very well for the oscillatory node ($x_{\mathrm{Oscillatory}_2}$), particularly, in the larger coupling range ($\epsilon > 4.0$) even for low α values, as shown in Fig. 5(b). The leakage parameter does not play any significant change in prediction by the machine except for a range of low (ϵ, α) values.

A sharp contrast appears in prediction once we increase the number of inputs. We feed two inputs into the reservoir: Time signals of one oscillatory ($x_{\mathrm{Oscillatory}_1}$) node and one excitable ($x_{\mathrm{Excitable}_1}$) node are fed and the machine predicts the signals of the rest of the eight nodes ($x_{\mathrm{Oscillatory}_{2,...,5}}$ and $x_{\mathrm{Excitable}_{2,...,5}}$). In this case, the machine performance is significantly enhanced for the whole range of our selected ϵ and α values in the prediction of an excitable node ($x_{\mathrm{Excitable}_2}$) as shown in Fig. 5(c), although the performance of the ESN seems better in a lower range of ϵ. Compared to Fig. 5(a), MSE is sharply decreased (MSE $\sim 10^{-6}$) here. In the case of the oscillatory node ($x_{\mathrm{Oscillatory}_2}$) and its machine-generated signal, the performance does not change significantly compared to the single input case. For the whole range of α and a range of larger $\epsilon > 4.8$, the performance remains almost similar to the single input case. However, MSE appears to decrease in a range of lower coupling ($\epsilon < 4.8$) and higher α values as shown in Fig. 5(d) that, encouragingly, indicates a desired improvement in the performance of the machine in prediction, in this range. Yet remains a range of low $\epsilon - \alpha$ values where
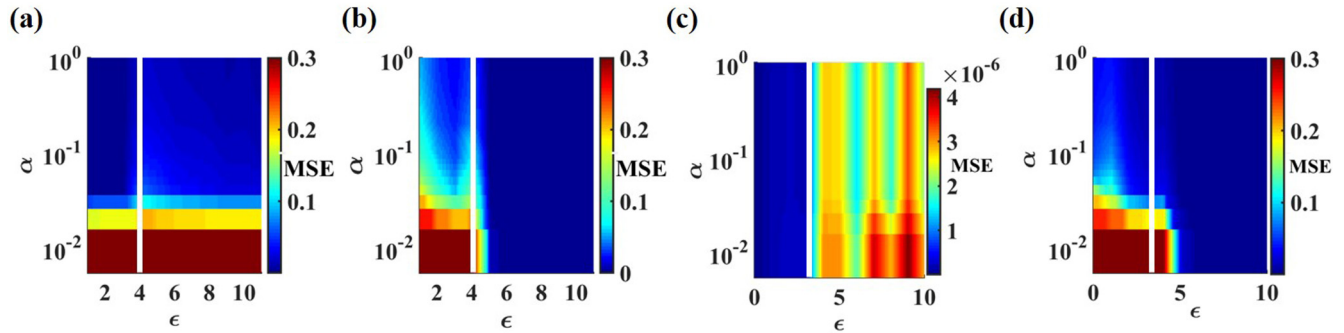
FIG. 5. Impact of machine parameters in RCSJ network. MSEs are plotted in a plane α-ϵ plane. [(a) and (b)] One oscillatory node input is fed to the reservoir. (a) MSE is calculated between real data and machine-generated data of the excitable node. (b) MSE of an oscillatory node's original data and machine-generated data. [(c) and (d)] Two input data: One oscillatory and one excitable nodes' data are fed into the machine. (c) MSE of one excitable node and (d) MSE of one oscillatory node are is plotted. All the plots [(a)–(d)] project MSE for varying α and ϵ. The desynchronized time signals (white vertical bar) and their output from the machine at a coupling strength $\epsilon = 4$ is discussed in detail in Sec. 2 of Ref. [34].

the performance remains poor due to the desynchronization of the network, which is discussed, in detail, in Ref. [34]. The vertical bar (white bar) marks the selected parameters to show the effect of desynchronization in [34]. It is shown that prediction by the ESN can be improved for desynchronized signals using the two-node signal feeding strategy, but one has to make a suitable choice of the leakage parameter. We discussed these issues in Ref. (Secs. 3 and 4) [34]. In the next section, we consider another network of a mixed population of neurons and test the performance of the ESN in one node as well as the two-node training approach.

IV. PREDICTION OF BURSTING AND CLUSTERING: COUPLED HR NEURONS

In a single RCSJ used above to build a network, in our first example, spiking oscillation starts due to tunneling of electron Cooper pairs through the superconducting junction for an applied DC bias current above a critical value [35]. It could generate bursting similar to a neuron [37] when a slow dynamics controls the spiking as induced here by the excitable nodes [26].

To check the performance of the ESN on the prediction of busting dynamics, we present here a biologically plausible and phenomenological neuron model, namely, the HR model [31,32], which responds to a constant external triggering above a threshold to generate recurrent action potential or spiking oscillation that basically represents fast movements of Na$^+$ and K$^+$ ions across the membrane of a neuronal cell. Additionally, the model incorporated [31] a slow Ca^{++} ion movement across the membrane that controlled the spiking oscillation resulting in a type of bursting dynamics. We employ our ESN technique next on a globally coupled network of the HR system. The dynamics of the nodes in the network is captured by

$$\dot{x}_i = y_i + bx_i^2 - ex_i^3 - z_i + I_i + \frac{\epsilon}{N_d} \sum_{k=1}^{N_d} (x_k - x_i) \quad (10)$$

$$\dot{y}_i = c - dx_i^2 - y_i \quad (11)$$

$$\dot{z}_i = rs(x_i - x_R) - rz_i, \quad (12)$$

where $e = c = 1$, $b = 3$, $d = 5$, $s = 4$, $r = 0.01$, $x_R = -1.6$, and $N_d = 10$. Once again, one half of the population is kept in steady state ($I = 1.3$) and the other half in oscillatory state ($I = 4$). Notation i indicates the node index.

Time evolution (x_i) of two arbitrarily selected nodes ($i = 1, 2$) from two subpopulations, one oscillatory (red line) and one excitable nodes (blue line), are picked up and compared with their respective machine-generated signals (black lines) as shown in Fig. 6. Time signal of two subpopulations are desynchronized and irregular for $\epsilon = 0.4$. For this coupling, the bursting is weakly predicted (black lines) by the
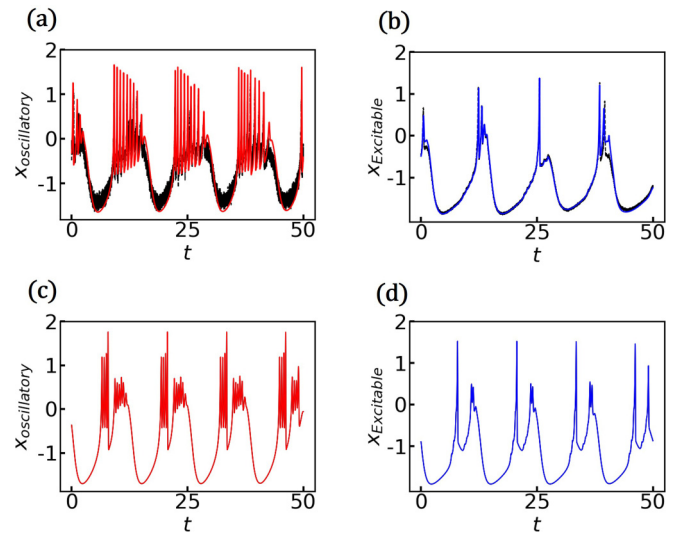

FIG. 6. Comparison of the time signal of two arbitrarily chosen nodes from two different subpopulations of the HR network. Results of one node data feeding to the reservoir are presented. (a) Signal of one oscillatory node (red) and the signal predicted by the machine (black). (b) Signal of one excitable node and machine-based predicted signal are shown with blue and black lines. For (a) and (b), the coupling strength is fixed at $\epsilon = 0.4$. [(c) and (d)] When the coupling ($\epsilon = 1.2$) is increased, the ESN generated signals closely match the original signals of the oscillatory (c) and excitable nodes (d).
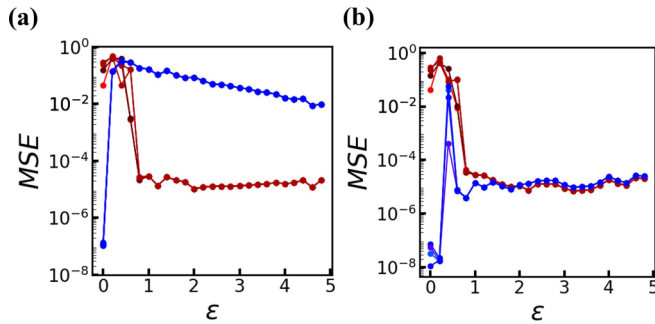
(a) **(b)**

FIG. 7. MSE between the original signal and predicted signals in coupled HR Neuronal networks. (a) One oscillatory node is fed into the machine. (b) Two nodes into the machine: one oscillatory and the other is chosen from the excitable population. MSE for all nodes are sharply dropped to a value $\sim 10^{-5}$ near a coupling $\epsilon \sim 1$.

machine as shown in Figs. 6(a) and 6(b). In particular, for the oscillatory node in Fig. 6(a), the prediction is very poor (cf. no match of red and black colored bursting), which is consistent with our results of the RCSJ network as expected in the weakly desynchronized regime. However, for a larger $\epsilon = 1.2$, the machine performance improves enormously. This is confirmed in Fig. 6(c) from a comparison of one oscillatory node with its predicted signal, which shows a close match (dominant red lines are seen, no trace of black line). The excitable node's signal overlaps (blue line) on its machine predicted signal (black line) in Fig. 6(d) very closely; no separate identity of the predicted signal (black line) is seen. Once again, results of the RCSJ example are confirmed. For higher coupling, the ESN performs well since clustering states emerge within the subpopulations. The machine is able to predict the bursting and the clustering behavior. For confirmation, first we feed one node information into the machine. MSE of the original and machine-generated signals of all the nodes are shown in Fig. 7(a). The predictions for the excitable/quiescent nodes (blue lines with circle) are not good compared to the oscillatory nodes (red line with filled circles). The ESN performance is poor, as usual, for training with data from one oscillatory node. The ESN performance improves for obvious reason, as explained for the RCSJ network, when we feed two nodes' information to the reservoir (one oscillatory node and another in quiescent state) for the training purpose. MSEs as a function of ϵ of all the nodes are shown in Fig. 7(b) for

two-node training strategy. Note that, in both cases, there is a sharp drop of MSE around a coupling strength $\epsilon \sim 1$, which is a signature of clustering in the network.

V. CONCLUSION

A reservoir computing based prediction of spiking and bursting is proposed for a group of globally connected oscillators in a backdrop of mixed parameter setup. We tested our strategy using two bursting models representing the local dynamics of the networks, one superconducting Josephson junction and one biological neuron model. A mixed population of quiescent and oscillatory nodes can generate bursting dynamics in the entire graph of superconducting junctions. We have shown that ESN can predict regular and irregular bursting of the entire graph over a range of coupling strength. The ESN can perform better if we feed the reservoir with information from two nodes from two different subpopulations. We are able to find a wide range of a machine parameters to check the performance of the reservoir computing-based prediction. This reveals a possibility of tuning the machine parameter for improved performance, in general. By using two inputs, the reservoir computing (ESN) even can predict the shape of desynchronized signals in the weak coupling regime. The success of the ESN based prediction was tested for homogeneous and heterogeneous networks in terms of system parameter distribution [34]. The work can be useful for efficient prediction of bursting and detection of cluster synchronization within a population and chimeralike states in the weaker coupling range of a network. These feature extractions are intriguing since otherwise, one cannot predict the onset of cluster and chimera states without using the detailed information of the dynamical models. In the near future, questions remain if more complex signals such as future arrival time of extreme events can be predicted from available time series data, and about extracting network structure information from a suitable setup of the ESN.

Codes to reproduce the results presented here are freely accessible in [48]. Additional information is available from the corresponding author upon reasonable request.

[1] J. Schmidhuber, Neural Networks **61**, 85 (2015).

[2] B. Lusch, J. N. Kutz, and S. L. Brunton, Nat. Commun. **9**, 4950 (2018).

[3] H. Jaeger and H. Haas, Science **304**, 78 (2004).

[4] LF. Seoane, Philos. Trans. R. Soc. London B **374**, 20180377 (2019).

[5] M. Lukoševičius and H. Jaeger, Comput. Sci. Rev. **3**, 127 (2009).

[6] J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, Phys. Rev. Lett. **120**, 024102 (2018).

[7] G. Tanaka *et al.*, Neural Networks **115**, 100 (2019).

[8] R. S. Zimmermann and U. Parlitz, Chaos **28**, 043118 (2018).

[9] J. Pathak, Z. Lu, B. Hunt, M. Girvan, and E. Ott, Chaos **27**, 121102 (2017).

[10] T. L. Carroll, Phys. Rev. E **98**, 052209 (2018).

[11] Z. Lu, B. R. Hunt, and E. Ott, Chaos **28**, 061104 (2018).

[12] T. Weng, H. Yang, C. Gu, J. Zhang, and M. Small, Phys. Rev. E **99**, 042203 (2019).

[13] T. Lymburn, D. M. Walker, M. Small, and T. Jüngling, Chaos **29**, 093133 (2019).

[14] N. Frolov, V. Maksimenko, A. Lüttjohann, A. Koronovskii, and A. Hramov, Chaos **29**, 091101 (2019).

[15] D. Ibáñez-Soria, A. Soria-Frisch, J. Garcia-Ojalvo, and G. Ruffini, PLoS One **14**, e0218771 (2019).

[16] X. Lin, Z. Yang, and Y. Song, Expert Systems Appl. **36**, 7313 (2009).

[17] X. Hinaut and P. F. Dominey, PLoS One **8**, e52946 (2013).

[18] D. Verstraeten, B. Schrauwen, D. Stroobandt, and J. Van Campenhout, Inform. Process. Lett. **95**, 521 (2005).

[19] O. V. Maslennikova and V. I. Nekorkin, Chaos **29**, 103126 (2019).

[20] A. Haluszczynski and C. R'ath, Chaos **29**, 103143 (2019).

[21] T. Lymburn, A. Khor, T. Stemler, D. C. Corrêa, M. Small, and T. Jüngling, Chaos **29**, 023118 (2019).

[22] T. L. Carroll and L. M. Pecora, Chaos **29**, 083130 (2019).

[23] A. Shirin, I. S. Klickstein, and F. Sorrentino, Chaos **29**, 103147 (2019).

[24] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, Nat. Commun. **2**, 468 (2011).

[25] T. Pyragiene and K. Pyragas, Phys. Lett. A **383**, 3088 (2019).

[26] C. Hens, P. Pal, and S. K. Dana, Phys. Rev. E **92**, 022915 (2015).

[27] J. Borresen and S. Lynch, PLoS One **7**, e48498 (2012).

[28] E. M. Izhikevich, Int. J. Bifur. Chaos **10**, 1171 (2000); A. Sherman and J. Rinzel, Proc. Natl. Acad. Sci. USA **89**, 2471 (1992).

[29] S. Ghosh, A. Mondal, P. Ji, A. Mishra, S. K. Dana, C. G. Antonopoulos, and C. Hens, Front. Comput. Neurosci. **14**, 49 (2020).

[30] S. K. Sharma, A. Mondal, A. Mondal, R. K. Upadhyay, and C. Hens, J. R. Soc., Interface **17**, 20190859 (2020).

[31] J. L. Hindmarsh, R. M. Rose, Proc. R. Soc. London B **221**, 87 (1984).

[32] A. Mishra, S. Saha, M. Vigneshwaran, P. Pal, T. Kapitaniak, and S. K. Dana, Phys. Rev. E **97**, 062311 (2018).

[33] S. Scardapane, D. Wang, and M. Panella, Neural Networks **78**, 65 (2016).

[34] See Supplemental Material at http://link.aps.org/supplemental/ 10.1103/PhysRevResearch.2.033338 for a detail mathematical description of the ESN, and for more illustrations on the impact of leakage parameter and the number of inputs on prediction are reported in Fig. S1, effect of heterogeneity in Figs. S2–S3. Improving prediction in the desynchronization regime is demonstrated in Figs. S4–S5.

[35] B. D. Josephson, Rev. Mod. Phys. **36**, 216 (1964).

[36] S. K. Dana, D. Sengupta, and K. Edoh, IEEE Trans. Circuits Syst. I **48**, 990 (2001).

[37] S. K. Dana, D. Sengupta, and C. Hu, IEEE Trans. Circuits Syst. II, **53**, 1031 (2006).

[38] M. Levi, F. C. Hoppensteadt, and W. L. Miranker, Q. Appl. Math. **36**, 167 (1978).

[39] S. H. Strogatz, *Nonlinear Dynamics and Chaos* (Westview Press, USA, 1998).

[40] P. Crotty, D. Schult, and K. Segall, Phys. Rev. E **82**, 011914 (2010).

[41] A. Mishra, S. Saha, C. Hens, P. K. Roy, M. Bose, P. Louodop, H. A. Cerdeira, and S. K. Dana, Phys. Rev. E **95**, 010201(R) (2017); A. Mishra, S. Saha, P. K. Roy, T. Kapitaniak, and S. K. Dana, Chaos, **27**, 023110 (2017).

[42] A. Ray, A. Mishra, D. Ghosh, T. Kapitaniak, S. K. Dana, and C. Hens, Phys. Rev. E **101**, 032209 (2020).

[43] H. Sakaguchi and Y. Kuramoto, Prog. Theor. Phys. **76**, 576 (1986).

[44] P. Kundu, P. Khanra, C. Hens, and P. Pal, Phys. Rev. E **96**, 052216 (2017).

[45] P. Kundu, C. Hens, B. Barzel, and P. Pal, Europhys. Lett. **120**, 40002 (2018).

[46] F. Hellmann, P. Schultz, P. Jaros, R. Levchenko, T. Kapitaniak, J. Kurths, and Y. Maistrenko, Nat. Commun. **11**, 592 (2020).

[47] G. B. Enmentrout and N. Kopell, SIAM J. Appl. Math. **46**, 233 (1986).

[48] S. Saha, A. Mishra, S. Ghosh, S. K. Dana, and C. Hens, Data cite for "Predicting bursting in a complete graph of mixed population through reservoir computing", https://github.com/ CRHENS/Predicting-bursting-in-a-complete-graph-of-mixed-population-through-reservoir-computing (2020).