Abstract—Electrical neuromodulation as a palliative treatment has been increasingly used in the control of epilepsy. However, current neuromodulations commonly implement predetermined actuation strategies and lack the capability of self-adaptively adjusting stimulation inputs. In this work, rooted in optimal control theory, we propose a Koopman-MPC framework for real-time closed-loop electrical neuromodulation in epilepsy, which integrates i) a deep Koopman operator based linear dynamical model to predict the temporal evolution of epileptic electroencephalogram (EEG) with an approximate finite-dimensional linear dynamics and ii) a model predictive control (MPC) module to design optimal seizure suppression strategies. The Koopman operator based linear dynamical model is embedded in the latent state space of the autoencoder neural network, in which we can approximate and update the Koopman operator online. The linear dynamical property of the Koopman operator ensures the convexity of the optimization problem for subsequent MPC control. The proposed deep Koopman operator model shows greater predictive capability than the baseline models (e.g., vector autoregressive model, kernel based method and recurrent neural network (RNN)) in both synthetic and real epileptic EEG data. Moreover, compared with the RNN-MPC framework, our Koopman-MPC framework can suppress seizure dynamics with better computational efficiency in both the Jansen-Rit model and the Epileptor model. Koopman-MPC framework opens a new window for model-based closed-loop neuromodulation and sheds light on nonlinear neurodynamics and feedback control policies.

Index Terms—Closed-loop neuromodulation, deep Koopman operator, electrical neuromodulation, model predictive control, online learning.

I. INTRODUCTION

EPILEPSY is a neurological disorder characterized by the occurrence of spontaneous epileptic seizures, during which the neuronal population fires in an abnormal, excessive and synchronized manner [1], [2]. Clinically, the main biomarkers for the diagnosis of epilepsy with electrophysiological signals are epileptiform spikes and seizure-like waves, which are usually monitored by noninvasive electroencephalogram (EEG). The invasive stereoelectroencephalogram (iEEG) and intracranial electroencephalogram (iEEG) are utilized to better localize the seizure onset zone [3], [4], [5].

A range of treatments have been developed to control seizures, such as anti-epilepsy medication, palliative/radical surgery, and electrical neurostimulation [2], [6], [7], [8], [9]. Compared to surgical resections, electrical neurostimulation, e.g., deep brain stimulation (DBS) and transcranial electrical stimulation (TES), has fewer long-term side effects on brain function, and has been becoming a considerable clinical intervention for intractable epilepsy [8], [10]. For instance, electrical nerve stimulation applies a certain dose of high-frequency current within a safe range (e.g., 50 Hz) to a specific targeting region (i.e., 0.5 ~ 5 mA in the mesial temporal cortex), aiming at modulating the seizure-like neural oscillations to the seizure-free state [8], [11].

However, current neuromodulation mainly implements predetermined open-loop actuation strategies, rather than a closed-loop stimulation strategy that is updated in real time. The open-loop neurostimulation actually lacks feedback signals from the observed real-time seizure waves for adjusting the control inputs during neurostimulation, which would limit its flexibility in practical applications [9]. More recently, closed-loop neuromodulation is increasingly gaining attention in clinical applications such as seizure control [8], [9], [12]. By framing the neuromodulation as a closed-loop state-space control problem, the main goal then is to design an optimal control law to steer the present system state to the desired state. Some pioneering works have been devoted to this area, ranging from designing state-space models to designing control methods. For example, Pequito et al. presented a spectral control method to ensure that the poles of a closed-loop system are within the
Prespecified spectrum range [13]. Ashourvan et al. proposed a pole-placement spectral static output feedback control-theoretic strategy for suppressing seizure dynamics [14]. Recently, model predictive control (MPC), as a model-based and self-tuning control method to maintain a stable and robust control process, has also been brought into neurostimulation applications [15], [16]. For instance, Chang et al. integrated a nonlinear auto-regressive moving-average (NARMA) Volterra model as an identified dynamical system into MPC for seizure suppression [15]. Chatterjee et al. proposed a fractional-order model based MPC framework in neuromodulation [16].

Seizure dynamics is a complex networked nonlinear dynamic process, which we have not yet fully understood [17], [18], [19]. The control of seizures requires modeling the underlying complex seizure dynamics and obtaining optimal control law for seizure suppression, as opposed to the seizure detection or classification tasks with deep learning solutions [20], [21], [22], [23], [24], [25], [26]. Therefore, system identification is crucial for uncovering the system dynamics over the seizure period by building a dynamical model and estimating its parameters based on the observed input-output data. Several traditional system identification methods have been used to model the seizure dynamics with EEG data, such as the auto-regressive moving-average model [15] and the fractional-order system model [16]. However, these two methods have not yet been validated in suppressing network-coupled seizure dynamics. With advances in machine learning, data-driven models with neural networks have shown great potential in system identification [27], [28], [29]. For instance, recurrent neural networks (RNNs) have proven promising in fitting time-series data and modeling dynamical systems in fluid flow control [29] and process industries [28]. However, integrated the highly nonlinear RNN models into a MPC framework for real-time optimization control, the quadratic programming problem becomes a non-convex optimization problem that is computationally inefficient.

When modeling seizure dynamics and designing seizure suppression controllers in the MPC framework, it is necessary to consider the trade-off between model accuracy and complexity. Higher model accuracy improves the predictive performance for approximating the nonlinear seizure system, while lower model complexity guarantees computational efficiency when performing MPC optimizations. For the non-convex nonlinear quadratic programming optimization problem originated from MPC optimization of a nonlinear complex dynamic model, finding its optimal solution is theoretically difficult and numerically inefficient. As an alternative, operator-theoretic approaches, mainly based on the Perron-Frobenius operator or its adjoint Koopman operator [30], aim to approximate a nonlinear dynamical system with linear operators. These operator based methods can help avoid the non-convex MPC optimization problems. However, the predefined Koopman operator model has substantial limitations in modeling time-varying or complex switching dynamical systems. Therefore, a more flexible way to approximate the finite-dimensional Koopman operator in an invariant subspace is needed [31]. Recently, autoencoder architecture has been employed to learn the finite-dimensional Koopman operator instead of the hand-crafted design of the basis function in the prediction of fluid flow dynamics [31], [32]. Considering the linear properties of the Koopman operator and its successful applications in predicting fluid flow dynamics (complex nonlinear system), we aim to integrate the Koopman operator with model predictive control for closed-loop neurostimulation.

In this study, we propose a deep Koopman operator based model predictive control framework (i.e., Koopman-MPC framework) for closed-loop seizure suppression in a real-time manner. The Koopman-MPC framework can meet the needs of sufficient prediction accuracy and low model complexity for real-time optimization. The major contributions of this study are summarized as follows.

- A Koopman-MPC framework is proposed (Section II), including a Koopman operator based linear dynamical model embedded into a finite-dimensional invariant subspace, and a MPC controller to optimize the neurostimulation policy for seizure control. The design of Koopman-based model leverages the flexibility of online learning of Koopman operator, and enables real-time optimization for MPC controller.
- A deep autoencoder neural network is employed to learn, rather than to hand-crafted design, the embedding linear invariant subspace (Section III). Experimental results demonstrate that the deep Koopman operator model outperforms other baseline models in seizure prediction, for both synthetic data and real EEG data (Section IV-D).
- The seizure suppression performance of the Koopman-MPC framework is validated and compared with the baseline RNN-MPC framework in both Jansen-Rit model and Epileptor model. The results show that Koopman-MPC framework successfully suppresses seizure dynamics with higher computational efficiency (Section IV-E) and allows online learning of the Koopman operator for closed-loop epilepsy control (Section IV-F).

II. KOOPMAN-MPC FRAMEWORK

In this section, we first introduce the Koopman operator theory and its extension to the linear forced dynamical systems, and then we propose the Koopman operator based model predictive control for obtaining the optimal control signals over the Koopman observations.

A. Koopman Operator

Consider a nonlinear discrete-time dynamical system

\[ x_{i+1} = f(x_i), \]

where \( x_i \in \mathcal{M} \subset \mathbb{R}^n \) denotes the vector of state variables at time step \( i \) and the function \( f(\cdot) : \mathcal{M} \rightarrow \mathcal{M} \) governs nonlinear dynamics of the system under a finite low-dimensional manifold \( \mathcal{M} \), in which the future state \( x_{i+1} \) only depends on the current state \( x_i \).

According to the Koopman operator theory, there exists an infinite-dimensional linear operator \( K \) defined on an infinite-dimensional function space such that the observation functions \( g(\cdot) \) satisfy an advancing linear forward relation, shown in Fig. 1.

\[ g(x_{i+1}) = Kg(x_i) = g \circ f(x_i). \]
The Koopman operator based MPC closed-loop neurostimulation framework for seizure suppression is solved by MPC with the approximated Koopman operator based linear dynamical system (Koopman-MPC) that is a transformation from a non-linear manifold space \( \mathcal{M} \subset \mathbb{R}^n \) to a linear manifold space \( \mathbb{R}^m \) that satisfies the linear forward relation \( K \).

The modal decomposition of the Koopman operator \( K \) refers to spectral analysis (Eigendecomposition) [33]. The eigenvalue and corresponding eigenvector of Koopman operator \( K \) have specific physical meanings. The eigenvalues of \( K \) reflect system stability, exponential damping ratio and spectral information, while the eigenvectors represent the independent spatial coordinates in the Koopman invariant space.

One nonnegligible issue about the Koopman operator modeling is how to design (or learn) the nonlinear embedding function \( g(\cdot) \). Some methods have been proposed to design a finite-dimensional embedding function \( \hat{g}(\cdot) \) to approximate the infinite-dimensional space (i.e., the extended dynamic mode decomposition [34] and auto-encoder [31], [32]). In this study, we adopt the auto-encoder solution.

Given a finite sequence of system state measurements \( x_i \ (i = \{0, 1, \ldots, n\}) \), the extended dynamic mode decomposition method constructs a basis function vector \( \Phi = [\Phi_1, \Phi_2, \ldots, \Phi_m]^T \) to construct two feature states \( \Psi_X \) and \( \Psi_X^\dagger \), as follows:
\[
\Psi_X = [\Phi(x_0), \Phi(x_1), \ldots, \Phi(x_{n-1})] \in \mathbb{R}^{m \times n}
\]
\[
\Psi_X^\dagger = [\Phi(x_1), \Phi(x_2), \ldots, \Phi(x_n)] \in \mathbb{R}^{m \times n}
\] (3)

According to the linear least-square estimation, the finite-dimensional approximation of the Koopman operator \( K \in \mathbb{R}^{m \times n} \) can be inferred as
\[
K = \Psi_X \Psi_X^\dagger
\] (4)
where \( \Psi_X^\dagger \) is the pseudo inverse of \( \Psi_X \) and defined as \( \Psi_X^\dagger = \Psi_X^T (\Psi_X \Psi_X^T + \lambda I)^{-1} \), and \( \lambda \) is the regularization factor. The approximated linear dynamical system is rewritten as
\[
\Phi(x_{i+1}) = \hat{K} \Phi(x_i)
\] (5)

Let us consider a nonlinear forced dynamics with input \( u_i \),
\[x_{i+1} = f(x_i, u_i), \]
where \( u_i \in \mathbb{R}^{r \times 1} \) is a \( r \)-dimension control input at time step \( i \), namely there are \( r \) inputs in this system. Then, we extend the approximated Koopman operator based linear dynamical system into a linear forced dynamical system, described as (7). In this case, the objective is simultaneously to learn an approximated finite-dimensional representation space of the state variables and its relationship with control inputs.
\[
\Phi(x_{i+1}) = \hat{K} \Phi(x_i) + \hat{B} u_i
\]
\[
= [\hat{K}, \hat{B}] \begin{bmatrix}
\Phi(x_i) \\
u_i
\end{bmatrix}
\] (7)

where \( \hat{B} \in \mathbb{R}^{m \times r} \) is the control matrix. Then, the approximated Koopman operator and the control matrix can be inferred from
\[
[K, \hat{B}] = \Psi_X \Psi_X^\dagger
\] (8)
where \( \Psi_X = \begin{bmatrix} \Psi_X^T & U \end{bmatrix} = [\Phi(x_0), \Phi(x_1), \ldots, \Phi(x_{n-1})] \) is the augmented state.

In this study, we design a deep autoencoder neural network as the basis function to learn the finite-dimensional linear invariant subspace in a self-supervised manner (Section III) instead of the hand-crafted basis function \( \Phi(\cdot) \).

B. Koopman-MPC Framework for Seizure Suppression

Model predictive control is an optimization-based control framework. It minimizes the objective function in a finite prediction horizon with the control inputs and states transition constraints. The MPC optimization for highly nonlinear dynamical systems is a nonconvex problem of NP-hardness, and we prefer to model it as the MPC convex optimization problem on linear systems with a unique solution. The predictive capability and linearity of the approximated Koopman operator facilitate the design of the model predictive controller.

In this study, we construct the Koopman-MPC framework to control the nonlinear dynamics of seizures following the desired reference signals (seizure-free signals), as shown in Fig. 2. The objective function is defined as follows,
\[
\min_{\Delta u} \sum_{i=1}^{T_p} (\Phi(x_i) - \Phi(x_{ref}))^T Q_X (\Phi(x_i) - \Phi(x_{ref})) + \sum_{i=1}^{T_p} \Delta u_i^T Q_u \Delta u_i
\]
s.t. \[\Phi(x_{i+1}) = \Phi(x_i) + \Delta x_i ,\]
\[u_i = u_{i-1} + \Delta u_i ,\]
\[u_i \in [u_{\min}, u_{\max}] ,\]
\[\Delta u_i \in [\Delta u_{\min}, \Delta u_{\max}] \]

(A) Brain Dynamics Identification

(B) Closed-loop Neurostimulation

Fig. 1. Koopman operator theory \( g(\cdot) \) is a transformation from a non-linear manifold space \( \mathcal{M} \subset \mathbb{R}^n \) to a linear manifold space \( \mathbb{R}^m \) that satisfies the linear forward relation \( K \).

Fig. 2. The Koopman operator based MPC closed-loop neurostimulation framework for seizure suppression. The control plant can be the real brain or a generative model of brain dynamics (e.g., the Jansen-Rit model or the Epilepto model). In the brain dynamics identification (A), the system parameters (i.e., Koopman operator \( K \) and Control matrix \( B \)) are identified by the input-output data pair \((u_i, y_i)\). In the closed-loop neurostimulation (B), the optimal control \( u^* \) is solved by MPC with the identified deep Koopman model.
where $T_p$ and $T_c$ are the length of predictive horizon and control horizon, respectively. $\Phi(x_i)$ and $\Phi(x_{ref}) \in \mathbb{R}^{n \times 1}$ are the finite-dimensional representations of the system state $x_i$ and reference signals $x_{ref}$ in the Koopman invariant subspace, respectively. $\Delta u_i \in \mathbb{R}^{r \times 1}$ ($r$ is the number of system inputs) is the increment of input at step $i$. $Q_X \in \mathbb{R}^{m \times m}$ is the positive definite weighted matrix for penalizing the deviation and $Q_u \in \mathbb{R}^{r \times r}$ is a non-negative matrix for penalizing the incremental amplitudes of control inputs. For real applications, the inputs are constrained between the upper and lower bound. The equation $\Phi(x_{i+1}) = K\Phi(x_i) + Bu_i$ is the linear state transition constraint condition that predicts the future states. In this study, the input signals are constrained from $-30 \text{ mV}$ to $5 \text{ mV}$ in the Jansen-Rit model and $-4 \text{ mV}$ to $0.5 \text{ mV}$ in the Epileptor model. The increment of input signals is constrained from $-20 \text{ mV}$ to $0.5 \text{ mV}$ in the Jansen-Rit model and $-0.25 \text{ mV}$ to $0.05 \text{ mV}$ in the Epileptor model. $Q_X$ is set to be an identity matrix $I$. $Q_u$ equals to 0.011.

To solve the optimization problem in the Koopman-MPC framework in (9), we utilize linear quadratic programming solver at each time step to find a sequence of incremental inputs $\Delta u_i$ ($i = 1, 2, \ldots, T_c$), and only the first incremental input $\Delta u_1$ is applied to the system for seizure suppression.

### III. DEEP KOOPMAN OPERATOR BASED SEIZURE DYNAMICAL MODEL

We design an autoencoder framework with fully-connected layers to learn a finite-dimensional invariant subspace, where we can approximate and update the Koopman operator. The proposed deep Koopman operator model is shown in Fig. 3(A). As a comparison, the RNN-based predictive model is presented in Fig. 3(B). Both the RNN-type models and Koopman-based model are composed of an encoder, a decoder and a dynamical system that predicts the future states (Fig. 3). The main difference is that the dynamics in hidden state space of the RNN-type network is a nonlinear system, while the dynamics in hidden state space of the deep Koopman model is a linear system characterized by the Koopman operator.

A. Autoencoder Architecture to Learn the Linear Embedding of Nonlinear Seizure Dynamics

The Koopman operator based autoencoder consists of an encoder and a decoder neural network, both of which contain three independent fully-connected layers. The first two layers are followed by a hyperbolic tangent activation function (tanh), while the last layer has no activation function and directly generates the output.

As shown in Fig. 3(A), $X$ is the input signal to the model, where $X = [x_0, x_1, \ldots, x_n] \in \mathbb{R}^{k \times (n+1)}$ denotes the past $n$ steps of data $x_{i\text{n-1}} = [x_0, x_1, \ldots, x_{n-1}]$ and the current state $x_n$. $k$ is the dimension of the state space of the original system. The encoder embeds the low-dimensional input data into a high-dimensional latent space and generates $Z \in \mathbb{R}^{m \times (n+1)}$ ($m > k$),

$$Z = encoder(X)$$

The following two cases are considered in the linear dynamical system within the finite-dimensional space.

(1) Without external inputs.

In this case, the prediction process is governed by

$$z_{n+1} = \tilde{K} z_n.$$  

We estimate $\tilde{K}$ by solving the linear least squares problem of $Z_{1:n} = KZ_{0:n-1}$.  

(2) With external inputs.

In this case, we consider the basic linear forced dynamical system for prediction

$$z_{n+1} = \tilde{K} z_n + \hat{B} u_n = [\tilde{K}, \hat{B}] \begin{bmatrix} z_n \\ u_n \end{bmatrix}.$$  

We estimate the Koopman operator $\tilde{K}$ and control matrix $\hat{B}$ with $Z_{0:n-1}, Z_{1:n}$ and input series $u_{0:n-1} = [u_0, u_1, \ldots, u_{n-1}]$.

$$[\tilde{K}, \hat{B}] = Z_{1:n} \left[ Z_{0:n-1} \mid u_{0:n-1} \right].$$

where $[Z_{0:n-1} \mid u_{0:n-1}]$ is the pseudo inverse of the augmented state of $Z_{0:n-1}$ and $u_{0:n-1}$.

The one-step linear approximated Koopman dynamics can be extended to higher-order case, similar to the higher-order autoregressive model, by considering encoded inputs from more previous time steps $z_{n-i}$ ($i = 0, 1, \ldots, r-1$), where $r$ is the number of previous steps.

$$z_{n+1} = \sum_{i=0}^{r-1} K_i z_{n-i} + \hat{B} u_n = [\tilde{K}_0, \ldots, \tilde{K}_{r-1}, \hat{B}] \begin{bmatrix} z_n \\ \vdots \\ z_{n-r+1} \\ u_n \end{bmatrix}. \tag{14}$$

The higher-order linear Koopman dynamics can still be solved by the linear least squares estimator.

The seizure dynamic is predicted in the Koopman invariant subspace by rolling the linear forward equation as (11), (12) or (14) accordingly.
The decoder enables us to transform $Z$ and the prediction of $z_{n+i}$, $(i = 1, 2, \ldots, T_p)$ back into the original state space
\[ \hat{X} = \text{decoder}(Z), \quad \hat{x}_{n+i} = \text{decoder}(z_{n+i}). \]  

(15)

**B. Definition of Explicit Loss Function**

The objective of the deep Koopman operator based autoencoder is to learn a finite invariant subspace that approximates the Koopman operator. The loss is specifically designed with different theoretical considerations for learning the finite-dimensional approximated Koopman operator and building the linear dynamical system. In summary, the explicit loss function for training the deep Koopman operator model covers the following: reconstruction term and prediction-reconstruction term with Mean-Squared Error (MSE).

1. Reconstruction error of $X$:
\[ L_{\text{recon}} = \frac{1}{n * D} \| X - \hat{X} \|_2^2, \]  

(16)

2. Prediction-reconstruction error:
\[ L_{\text{pred}} = \frac{1}{T_p} \| x_{n+i} - \hat{x}_{n+i} \|_2^2, \]  

(17)

where $D$ is the dimension of the original state space. $n$ is the number of steps in both $X$ and $\hat{X}$. $T_p$ is the length of prediction.

The total loss function can be summarized as
\[ L = \lambda_{\text{recon}} L_{\text{recon}} + \lambda_{\text{pred}} L_{\text{pred}} \]  

(18)

where $\lambda_{\text{recon}}$ and $\lambda_{\text{pred}}$ are the hyper-parameters to balance the reconstruction error and prediction-reconstruction error. In our study, we set them equal to 1 by default.

**IV. EXPERIMENTS AND RESULTS**

We first test the performance of the deep Koopman operator model in predicting seizure dynamics using synthetic data and real data and then test the performance of the Koopman-MPC framework on seizure suppression using a simulation platform.

**A. Dataset Description**

The deep Koopman operator model is validated with both synthetic data and real data. We used the well-known Jansen-Rit model [35] and the Epileptor [17] model to generate seizure dynamics. By varying the model parameter (i.e., the excitatory synaptic gain, $A$), Jansen-Rit model can synthesize the seizure-free EEG and seizure EEG signals. In our study, we mimic the seizure propagation process in two distant interconnected cortical columns by a coupled Jansen-Rit model described in [35] and the seizure transition from interictal period to ictal period. The seizure propagates from the epileptogenic foci (Cortex 1) to a seizure-free region (Cortex 2). The seizure transition process is induced by altering the excitatory parameter $A$ in cortex 1. An example of synthetic data with seizure transition and seizure propagation is shown in Fig. 4(A). The ordinary differential equations (ODEs) of the coupled Jansen-Rit model and its parameter settings are summarized in Appendix Section A and Table V.

The Epileptor model integrates the dynamics in seizure interictal and ictal periods using ODEs with three different time scales, as described in Appendix Section B. Changing the excitability parameter $x_0$ in the Epileptor model can generate seizure oscillations or seizure-free dynamics. We also simulate the seizure propagation process in a coupled Epileptor model with two interconnected cortexs, as shown in Fig. 4(B).

We synthesize 4000 s recording with 100 Hz sampling rate in each synthetic model. The first half of each recording was separated as the training and another as testing data. Both models synthesize the cycles of interictal-ictal data and repeat over time. In the Jansen Rit model, the excitatory parameter is sampled from the Gaussian distribution with the varying mean value, which results in different patterns of seizure dynamics.

The real data comes from an open dataset where the raw iEEG data for patients comes from NIH, UMF, UMMC, and JHH in the OpenNeuro repository (https://openneuro.org/datasets/ds003029/versions/1.0.3) [36]. It includes iEEG signals during the interictal period and ictal period with 1000 Hz sampling rate. The recording of the real data with randomly selected channels is shown in Fig. 4(C). We select one of the patients with 84 channels and two different time stamp recordings. One of the recordings is used for training the model, and the other is used for testing.

**B. Baseline Models and Evaluation Metrics**

To evaluate the performance of data prediction, the Koopman operator model is compared with several baseline models, including two vector autoregressive models (i.e., a 5-order VAR, a 10-order VAR), a vector autoregressive moving average model (i.e., VARMA with 5-order AR and 5-order MA), a kernel-based model (i.e., SINDy [37]) and three types of recurrent neural network models (i.e., Vanilla RNN, Gated Recurrent Units, and Long-short term memory). The order of VAR and VARMA means the number of preceding data that are used to predict the value at the present time. Both VAR and VARMA models are estimated on the original space of the raw data, and the kernel-based model is estimated on the spanned space of
C. Hyperparameter Selection in Koopman Model

The dimension of the approximated Koopman operator is a hyperparameter in our study. The larger the dimension of the Koopman operator, the more complex the approximated linear dynamical system. Here, we introduce the Bayesian Information Criterion (BIC) for hyperparameter tuning to tradeoff the model complexity and accuracy. The definition of BIC is as follows,

$$\text{BIC} = n \times LL + k \times \log(n),$$

(19)

where $n$ is the number of samples in the training dataset; $LL$ is the log-likelihood (e.g., log of the mean squared error); and $k$ is the number of parameters in the model. In hyperparameter tuning of our model, we treat $n$ as the batch size of the training set, $LL$ as the log of $MSE$ in model prediction, and $k$ as the dimension of the Koopman Operator (the latent dimension in the latent layer multiplies the order of the linear system). The tuning of the dimension of the Koopman operator in two synthetic models is shown in Fig. 5(A) and (D) respectively.

After selecting the dimension of the approximated Koopman operator, we then tune the time lengths of states for updating the Koopman operator. We quantify the $MSE$ and $R^2$ given different time lengths of states. The results are shown in Fig. 5(B), (C) and (E), (F) respectively. In our setting, the optimal hyperparameters of the Koopman model in predicting the synthetic data generated by Jansen-Rit model and the Epileptor model are latent dimension=18, order=1, time length=100 and latent dimension=10, order=2, time length=70 respectively.

D. Prediction of Seizure Dynamics

The Koopman operator model and the RNN-type predictive models are trained to predict seizure dynamics. The VAR, VARMA and Kernel-based predictive models are online learning methods that do not require a training procedure. In the training procedure, the network parameters (e.g., the encoder and the decoder), the Koopman operator, and the control matrix of the Koopman operator model are learned, and the network parameters and the RNN-cell parameters of RNN-type predictive models are learned. In the test procedure, the Koopman operator and the control matrix can be updated in real-time explicitly thanks to its linear property.

1) Prediction of the Synthetic Data: Prediction results from the synthetic data with the coupled Jansen-Rit model and Epileptor model are shown in Fig. 6 & 7, Table I & II respectively.
TABLE I
THE PREDICTION PERFORMANCE IN THE SYNTHETIC DATA (COUPLED JR MODEL)

| Model          | # of param. | $MSE \downarrow$ | $MAE \downarrow$ | $M_{AE} \downarrow$ | $EV \uparrow$ | $R^2 \uparrow$ |
|----------------|-------------|-------------------|------------------|---------------------|-------------|-------------|
| VAR (5-order)  | 0           | 31.35±0.7436      | 2.84±0.2059      | 0.85±0.370         | -1.19±0.821 | -1.20±0.823 |
| VAR (10-order) | 0           | 36.64±1.0335      | 3.28±0.0777      | 1.29±0.535         | -0.47±0.267 | -0.49±0.269 |
| VARMA (5-order AR, 5-order MA) | 0  | 23.82±0.6330      | 2.54±0.1635      | 0.82±0.407         | -3.11±1.732 | -3.12±1.733 |
| SINDy (Poly+Four. Kernel) | 0 | 73.91±1.8388     | 4.94±1.154       | 1.63±0.762         | -0.03±0.060 | -0.08±0.059 |
| Vanilla RNN-Cell (14-dim) | 4125 | 4.41±1.358       | 0.89±0.234       | 0.19±0.074         | 0.81±0.055 | 0.81±0.056 |
| LSTM-Cell (20-dim) | 6492 | 4.76±1.462       | 0.91±0.285       | 0.16±0.084         | 0.80±0.048 | 0.80±0.048 |
| GRU-Cell (18-dim) | 5231 | 4.21±1.322       | 0.80±0.248       | 0.17±0.085         | 0.81±0.027 | 0.81±0.027 |
| Koopman (18-dim,1-order) | 1460 | 1.57±0.796       | 0.55±0.132       | 0.17±0.067         | 0.93±0.059 | 0.93±0.060 |

TABLE II
THE PREDICTION PERFORMANCE IN THE SYNTHETIC DATA (THE EPILEPTOR MODEL)

| Model          | # of param. | $MSE \downarrow$ | $MAE \downarrow$ | $M_{AE} \downarrow$ | $EV \uparrow$ | $R^2 \uparrow$ |
|----------------|-------------|-------------------|------------------|---------------------|-------------|-------------|
| VAR (5-order)  | 0           | 0.05±0.021        | 0.01±0.021       | 0.03±0.010         | 0.92±0.039  | 0.92±0.039  |
| VAR (10-order) | 0           | 0.05±0.023        | 0.10±0.022       | 0.03±0.011         | 0.96±0.040  | 0.96±0.040  |
| VARMA (5-order AR, 5-order MA) | 0 | 0.10±0.044       | 0.15±0.030       | 0.07±0.014         | 0.86±0.057  | 0.86±0.058  |
| SINDy (Poly+Four. Kernel) | 0 | 0.37±0.092       | 0.41±0.088       | 0.29±0.118         | 0.46±0.217  | 0.46±0.275  |
| Vanilla RNN-Cell (10-dim & 10-step) | 1007 | 0.02±0.007       | 0.00±0.010       | 0.04±0.007         | 0.96±0.013  | 0.96±0.013  |
| Vanilla RNN-Cell (18-dim) | 5555 | 0.01±0.005       | 0.06±0.010       | 0.03±0.005         | 0.98±0.009  | 0.98±0.009  |
| GRU-Cell (10-dim & 10-step) | 1267 | 0.01±0.006       | 0.05±0.010       | 0.03±0.005         | 0.97±0.011  | 0.97±0.012  |
| GRU-Cell (20-dim & 60-step) | 9032 | 0.01±0.004       | 0.05±0.009       | 0.03±0.005         | 0.98±0.006  | 0.98±0.006  |
| LSTM-Cell (10-dim & 10-step) | 1397 | 0.01±0.006       | 0.06±0.010       | 0.03±0.006         | 0.97±0.011  | 0.97±0.011  |
| LSTM-Cell (20-dim & 60-step) | 10092 | 0.01±0.004       | 0.05±0.010       | 0.03±0.006         | 0.98±0.007  | 0.98±0.007  |
| Koopman (10-dim,2-order) | 712  | 0.03±0.010       | 0.09±0.013       | 0.03±0.006         | 0.94±0.021  | 0.94±0.021  |

Specifically, Fig. 6 & 7 illustrate the prediction of temporal evolution when using the Koopman operator based model, GRU-cell model, 5-order VAR and SINDy model. The results show two deep learning models (i.e., Koopman operator based model and the GRU model) can better predict future dynamics. Table I shows six quantification metrics (i.e., # of parameters, $MSE$, $MAE$, $M_{AE}$, $EV$, and $R^2$) for the coupled JR model, and Table II for the Epileptor model. Table I & II demonstrate that the Koopman model outperforms or approximates to the baseline models. Our results suggest that the deep Koopman model can learn the invariant subspace and predict the future time series.

2) Prediction of the Real iEEG Data: We spilled the real 84-channel iEEG dataset into the training data and testing data. All the deep learning models are trained with the training data and tested with the testing data. Fig. 8 visualizes the prediction of seizure dynamics from the real iEEG test data using Koopman-based model, compared with the 10-order VAR model and a GRU-based model. The quantification metrics on the training and testing data are shown in Table III. Table III shows that although the RNN-type models (RNN-Cell, LSTM-Cell and GRU-Cell model) have better performance in training data, the performance is largely reduced in testing data. Online learning models (5-order VAR model, 10-order VAR model and Koopman model) have better generalizability.

Results from synthetic data and real data indicate that the deep Koopman operator model can learn a finite-dimensional invariant subspace to transform the original nonlinear dynamical system to a linear system. The dynamics of the entire system are embedded into the approximated Koopman operator $K$ and...
TABLE III
THE PREDICTION PERFORMANCE IN THE REAL SEIZURE DYNAMICS DATA (EPILEPSY-IEEG-MULTICENTER-DATASET)

| Model                        | # of param. | \(MSE\) ↓ | \(MAE\) ↓ | \(McAE\) ↓ | \(EV\) ↑ | \(R^2\) ↑ |
|------------------------------|-------------|-----------|-----------|-----------|---------|---------|
| VAR (5-order)                | 0           | 0.084     | 0.111     | 0.108     | 0.123   | 0.051   | 0.090   | 0.807   | 0.812   |
| VAR (10-order)               | 0           | 0.073     | 0.100     | 0.104     | 0.122   | 0.050   | 0.060   | 0.829   | 0.829   |
| SYDy (Poly+Four Kernel)      | 0           | 0.961     | 0.235     | 0.329     | 0.229   | 0.164   | 0.157   | \(-6.437\) | \(-1.687\) |
| Vanilla RNN-Cell (100-dim)   | 379k        | 0.054     | 0.150     | 0.151     | 0.232   | 0.109   | 0.154   | 0.737   | 0.624   | 0.735   | 0.622   |
| LSTM-Cell (100-dim)          | 410k        | 0.028     | 0.127     | 0.111     | 0.207   | 0.081   | 0.132   | 0.850   | 0.691   | 0.856   | 0.699   |
| GRU-Cell (100-dim)           | 400k        | 0.028     | 0.125     | 0.109     | 0.203   | 0.079   | 0.129   | 0.857   | 0.700   | 0.858   | 0.658   |
| Koopman (140-dim, 1-order)   | 102k        | 0.049     | 0.072     | 0.112     | 0.139   | 0.070   | 0.086   | 0.851   | 0.858   | 0.851   | 0.858   |

Fig. 9. Seizure suppression performance of MPC controllers with Jansen-Rit model. (A) Koopman-MPC controller; (B) GRU-MPC controller; (C) LSTM-MPC controller. The experiments are conducted on the simulation platform using coupled Jansen-Rit Model. From top to bottom, we show the EEG waveform in cortex 1, cortex 2, and the control signal applied to cortex 1, respectively. The red/gray line indicates the waveform with/without control.

In LSTM-MPC and GRU-MPC frameworks, the hidden states lie in a highly nonlinear manifold space, rather than the linear transition space in the Koopman-MPC framework (9). Thus, the nonlinear dynamical system in RNN-type models (e.g., LSTM-Cell and GRU-Cell model) constructs a highly nonlinear quadratic programming optimization problem in MPC, which cannot be solved simply by a quadratic programming solver to obtain the optimal solution. To this end, we use a deep learning solution with gradient descent for optimization. Specifically, we transfer the LSTM-MPC and GRU-MPC optimization problem into a pyTorch framework and solve it with the Adam solver (with a fixed learning rate of 0.01 and 10-step iterations).

We conducted numerical experiments on the Jansen-Rit simulation platform and the Epileptor platform to test the effects of the Koopman-MPC framework on seizure control. The stimulation signal is only applied to the epileptogenic foci (cortex 1). We set a 10-step predictive horizon and a 10-step control horizon to obtain the optimal control signal for seizure suppression. In each experiment, we first collect some input-output data pairs for updating the model parameters in the deep Koopman operator model, as well as initializing the hidden state of the GRU and LSTM model. Then, the optimal control signal is designed by MPC controller and inserted into the Jansen-Rit model and the Epileptor model respectively. Fig. 9 and Fig. 10 show the suppressed seizure signals (top and middle) and the corresponding control signals (bottom), which are provided by

---

Fig. 10. Seizure suppression performance of MPC controllers with the Epileptor model. (A) Koopman-MPC controller; (B) GRU-MPC controller; (C) LSTM-MPC controller. The experiments are conducted on the simulation platform using a coupled Epileptor Model. From top to bottom, we show the EEG waveform in cortex 1, cortex 2, and the control signal applied to cortex 1, respectively. The red/gray line indicates the waveform with/without control.

---

the control matrix \( \hat{B} \), which can be re-calibrated at each step, underlying the flexibility of online learning. In contrast, the RNN-type models learn the hidden nonlinear dynamics and evolve under a pre-trained dynamical law. Its system properties cannot be updated online. The vector autoregressive model fails to capture the highly nonlinear dynamics with large amplitudes in the predictions.

E. MPC-Based Seizure Suppression

We test the seizure suppression performance in the Jansen-Rit model and the Epileptor model. We examine the control ability and the computational efficiency of MPC optimization in seizure suppression with the Koopman-based model, the GRU-based model and the LSTM-based model. Since seizure dynamics are highly nonlinear, it is difficult for a predictive model to track the dynamics of seizures over a long time. The Koopman-MPC model therefore periodically re-estimates the Koopman operator \( \hat{K} \) and control matrix \( \hat{B} \). To be noted, the network parameters in the Koopman-based model and RNN-type models (LSTM-Cell and GRU-Cell model) are not updated once the models have been trained. Only the linear dynamical system parameters (i.e., \( \hat{K} \) and \( \hat{B} \)) can be updated online. The optimal neurostimulation strategy under the MPC framework is solved by the linear quadratic programming solver.
Koopman-MPC framework (Fig. 9(A) & Fig. 10(A)), GRU-MPC framework (Fig. 9(B) & Fig. 10(B)) and LSTM-MPC framework (Fig. 9(C) & Fig. 10(B)). In Fig. 9, although all models successfully suppress the seizures, the Koopman-MPC controller responds faster and with fewer seizure waves after triggering the stimulator. In Fig. 10, all three frameworks can suppress the transition into the ictal period (the fast oscillation with negative amplitude). But suffering in the oscillation in the interictal period (the oscillation with positive amplitude). The main reason might be related to the stimulation artifacts. The performance is further confirmed by the quantification of the computational efficiency in Table IV. The Koopman operator MPC framework only takes 0.02 s in the Jansen-Rit model and takes 0.015 s in the Epileptor model to obtain the optimal control signal at each time step, which is sufficiently fast for the closed-loop neurostimulation applications.

F. Online Learning of Koopman Operator for Closed-Loop Neurostimulation

We further test whether the online learning Koopman-MPC framework can predict and control seizures during the state transition process in the Jansen-Rit model (e.g., the state transition among seizure-free, preictal, and ictal period) and in the Epileptor model (e.g., the state transition between interictal and ictal period). In Jansen-Rit model, we manipulate the state transitions by varying the excitatory gain parameter \( A \): from the seizure-free state, to the preictal period (slow oscillation), then to the seizure state (fast oscillation). Specifically, EEG signals are in the non-ictal state for the first 2 seconds (\( A = 7.0 \)), then shift to the preictal period for 3 seconds (\( A = 7.2 \)), and eventually turn to the ictal state for the last 3 seconds (\( A = 7.8 \)). The Epileptor model simulates the interictal and ictal periods with three different time scales. The fastest scale reflects the fast oscillations, while the lowest scale governs the transition between the interictal and ictal periods. The gray lines in Fig. 11(A&D) show the synthetic EEG signals without seizure control.

We conduct eigendecomposition on the estimated Koopman operator \( \hat{K} \) to measure the maximum frequency of the seizure dynamics. Fig. 11(C) and (F) shows the maximum complex part of the eigenvalues of \( \hat{K} \) in spectral analysis for the data generated by the Jansen-Rit model and the Epileptor model, respectively. We find that the maximum spectral of the system is largely increased when the seizure-like wave occurred in both synthetic data, suggesting that spectral analysis of the Koopman-based model might offer a potential biomarker of seizure onset and need to be validated with more real dataset.

Inspired by the results of spectral analysis, we design a closed-loop neurostimulation system using the maximum frequency of \( \hat{K} \) (i.e., \( f \)) as a biomarker. According to Fig. 11(C), we set a threshold \( f_{\text{thres}} = 0.1 \). When \( f > f_{\text{thres}} \), the controller is actuated and triggers the neurostimulation to suppress seizure-like waves. We firstly conduct experiments to validate our closed-loop neurostimulation system in Jansen-Rit model. As shown in Fig. 11(A-C), the neurostimulation is triggered when \( f > f_{\text{thres}} \). The seizures are immediately suppressed after the controller is actuated. When the Koopman-MPC framework detects the seizure dynamics, it triggers the stimulator again.

The closed-loop neurostimulation system is also tested in the Epileptor model (Fig. 11(D)-(E)). We set the threshold of the maximum frequency \( f_{\text{thres}} = 0.1 \) and the threshold of EEG amplitude \( V_{\text{cortex}} = 0 \) as biomarker in the Koopman-MPC framework. The detector starts detecting at 25 s. Once triggering the neurostimulator, the seizure wave with fast oscillation and negative amplitude is controlled, as the black line (with Koopman-MPC control) vs the gray line (without control) in Fig. 11(D). Despite considerable performance to suppress huge negative bursts, the small interictal spikes (e.g., positive oscillations) cannot be completely suppressed. This may be related to the effect of stimulation artifacts or the interaction of internal states in the Epileptor model.

V. DISCUSSION

In this work, we propose a novel deep Koopman-MPC framework for closed-loop electrical neurostimulation. A tailored autoencoder is employed to learn the invariant subspace of the Koopman operator, in which the nonlinear dynamics is mapped into finite-dimensional linear dynamics. The approximation of
the Koopman operator provides sufficient prediction horizon for seizure prediction, which can be integrated into the MPC controller for seizure suppression.

A. Prediction and Control for Seizure Dynamics

The epileptic seizures can be modeled as a dynamical system [15], [16], [17]. System identification is an essential step for modeling and prediction of seizure dynamics. However, the accuracy and the complexity of the identified model is a trade-off in modeling seizure dynamics, especially when the model is integrated into the MPC framework for designing optimal neurostimulation in real-time. In our work, the proposed Koopman operator based model has higher prediction accuracy compared with VAR model, Kernel-based model and nonlinear RNN-type models (Table I, II and III), while maintaining a considerable size of network parameters. One possible reason is that the Koopman model learns an invariant subspace that captures sufficient nonlinear patterns in the approximated Koopman operator for seizure prediction [38]. A great merit of the Koopman-based model is its ability to online update the Koopman operator, whereas the RNN-type models only learn the dynamics of the training data, and they cannot update its governing function online. Another merit of the Koopman model is its linear property, which greatly facilitates the convex optimization in MPC control and provides a unique optimal solution with high computational efficiency (Table IV). In contrast, the RNN-type MPC framework contains a highly nonlinear optimization problem, which is typically computationally expensive. In fact, there is a trade-off between model complexity and computational efficiency when designing models for real-time closed-loop control. A more complex nonlinear model might better identify the system dynamics, but it would suffer from the high computational cost in optimization. Importantly, the proposed Koopman-MPC framework, to some extent, breaks the speed-accuracy tradeoff by combining a finite-dimensional approximated Koopman operator model and a linear MPC optimal controller. Beyond the invasive neurostimulation investigated in this study, the Koopman-MPC framework might be extended to broader neurostimulation protocols involving other types of stimulation (e.g., transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS) and Vagus nerve stimulation (VNS)) in the future.

B. Seizure Simulation Platform

For ethical reasons, it is difficult to directly collect input-output data of patients with epilepsy. Therefore, using a computational model to generate neural dynamics is an alternative way to validate control strategies. In this study, we applied the Jansen-Rit model and the Epileptor model to validate the Koopman model to simulate the EEG data and conduct MPC control. The Jansen-Rit model describes the dynamical interaction of cortical column-level neural mass model, the Jansen-Rit model accounts for the dynamical interactions among three populations of cells (i.e., excitatory, inhibitory, and pyramidal neuron populations) [35]. By varying the bifurcation parameter, the average excitatory synaptic gain $A$, the model can generate neural activity in healthy or seizure conditions [35]. The Epileptor model is also a neural mass model that simulates the state transitions between the interictal and ictal period with three different time scales, the fast time scale controls the high-frequency oscillation of spike-wave while the slow time scale mimics the state transitions [17]. Also, many other dynamical models have been proposed to simulate neurodynamics with physics laws for exploring the underlying neural mechanisms. For instance, the Kuramoto model, as a phase oscillator, is used to synthesize the fMRI phase dynamics series [39] and to study the cognitive process [40]. It would be of high interest to build a more realistic seizure simulation platform for testing different control policies and even for performing virtual surgery.

C. Limitations and Future Works

It is worthy to mention the limitations and future directions of our work. First, we did not collect input-output data from epilepsy patients in a biological experiment. Instead, we ran numerical experiments with the Jansen-Rit model and Epileptor model to simulate the EEG data and conduct MPC control. Second, we did not study the controllability of the identified system, which can be characterized by the approximated Koopman operator and control matrix [41]. The controllability of a linear dynamical system reflects whether the system can reach any state in finite control steps. Studying the controllability of neural dynamical systems will broadly advance neuroscience applications (e.g., cognitive control, neurofeedback control). Lastly, we did not consider the effects of noise in our model. The EEG denoising techniques [42], [43] will contribute to the robustness of Koopman-based models in real-world applications.

The brain is a complex coupled dynamical system [44], [45]. Modeling and controlling the brain network while balancing speed and accuracy is challenging. The development of neuromodulation techniques and cybernetics will promote each other. In our work, we only consider two cortical regions dynamics and their neuromodulation, controlling the complex brain network dynamics requires hierarchical network models and more advanced control theory [46], [47]. The development of control theory would provide new tools for neurostimulation, including the selection of optimal targeted stimulation regions, the robustness and safety of stimulation protocols.

ACKNOWLEDGMENT

The authors gratefully acknowledge the anonymous reviewers for their insightful comments, Mr. Rongwei Liang for discussions and editing, Dr. Yi Yao and Prof. Haiyan Wu for their useful suggestions. We thank the support from the Swarma-TCCI Scholarship to Zhichao Liang and Zixiang Luo and the Neuromagnetic Model Reading Group jointly sponsored by Swarma Club and Tianqiao and Chrissy Chen Institute. All authors declare to have no conflict of interests.

APPENDIX

A. The Coupled Jansen-Rit Model

The Jansen-Rit model describes the dynamical interaction of three subpopulations, the pyramidal cells, the excitatory feedback interneurons and the inhibitory feedback interneurons. It
can be extended into a coupled Jansen-Rit model to simulate neural activity in two interconnected cortical areas in (20). Parameters of the coupled Jansen-Rit Model are summarized in Table V.

\[
\begin{align*}
\dot{y}_0 &= y_0 \\
\dot{y}_1 &= A_1 y_1 (y_1 - y_2) - 2a y_3 - a^2 y_1 \\
\dot{y}_2 &= y_2 + Input \tag{20}
\end{align*}
\]

where the Sigmoid function \( S(v) = \frac{2v_0}{1+e^{v_0 v}} \), Input is the external input applied in cortex 1.

**B. Epileptor Model**

The parameterized epileptor neural field model is composed of five state variables in three different time scales: fast, slow and intermediate time scales. The epileptor model can be summarized as

\[
\begin{align*}
\dot{x}_{1,i} &= \frac{1}{\tau_1} \left[ (y_{1,i} - f_1(x_{1,i},x_{2,i}) - z_i + 3.1) + Input_{i} \right] \\
\dot{y}_{1,i} &= \frac{1}{\tau_1} \left[ 1 - 5x_{1,i}^2 - y_{1,i} \right] \\
\dot{z}_i &= \frac{1}{\tau_0} \left[ 4(x_{1,i} - x_0) - z_i - \sum_{j=1}^{N} K_{i,j} \cdot (x_{1,j} - x_{1,i}) \right] \\
\dot{x}_{2,i} &= \frac{1}{\tau_1} \left[ -y_{2,i} + x_{2,i} - x_{2,i}^3 + 0.45 + 0.002g(x_{1,i}) - 0.3(z_i - 3.5) \right] \\
\dot{y}_{2,i} &= \frac{1}{\tau_2} \left[ -y_{2,i} + f_2(x_{2,i}) \right] \tag{21}
\end{align*}
\]

**C. RNN-Type Model for System Identification**

The RNN-type models have been validated in fluid flow dynamics [29] and process industries [28]. Here, we apply it as baseline model to predict the seizure-like dynamics. We instantiate RNN-type predictive model with three types of recurrent cells (i.e., RNN-Cell, LSTM-Cell and GRU-Cell). The previous states and inputs are used to infer the initial hidden state of the recurrent cell, and the hidden dimension reflects model complexity. The dimension in the output layer of the encoder and the input of the decoder is determined by the hidden dimension of the RNN-type cell. The input dimension of the encoder is determined by number of state and time steps of previous state-input pair. We choose the optimal hyperparameter set for each RNN-type model with the lowest \( MSE \) and the highest \( R^2 \) (Fig. 12 and 13).

---

**TABLE V**

| Parameter | Description | Value |
|-----------|-------------|-------|
| \( A \)   | Average excitatory synaptic gain | 3.4 mV |
| \( B \)   | Average inhibitory synaptic gain | 23 mV |
| \( C_1 \) | Average synaptic connectivity | 108 |
| \( C_2 \) | Average excitatory connectivity | 93.7 |
| \( C_3 \) | Average inhibitory connectivity | 93.7 |
| \( \mu' \) | Average excitatory synaptic gain | 7 mV |
| \( \nu' \) | Average inhibitory synaptic gain | 23 mV |
| \( C_1 \) | Average synaptic connectivity | 108 |
| \( C_2 \) | Average excitatory connectivity | 93.7 |
| \( C_3 \) | Average inhibitory connectivity | 93.7 |

---

**Fig. 12.** Hyperparameter tuning in training RNN-type model in Jansen-Rit Model. The quantification of \( MSE \) (top) and \( R^2 \) (bottom) from (A) GRU-Cell model, (B) LSTM-Cell model, (C) RNN-Cell model.

**Fig. 13.** Hyperparameter tuning in training RNN-type model in Epileptor Model. The quantification of \( MSE \) (top) and \( R^2 \) (bottom) from (A) GRU-Cell model, (B) LSTM-Cell model, (C) RNN-Cell model.
