Dreaming of electrical waves: Generative modeling of cardiac excitation waves using diffusion models

View Online   Export Citation   CrossMark

Tanish Baranwal,[1,2] (ID) **Jan Lebert,**[1] (ID) **and Jan Christoph**[1,3,a] (ID)

AFFILIATIONS

[1] Cardiovascular Research Institute, University of California, San Francisco, San Francisco, California 94158, USA
[2] Department of Electrical Engineering and Computer Science, University of California, Berkeley, California 94720-1770, USA
[3] Division of Cardiology, University of California, San Francisco, San Francisco, California 94143, USA

[a]Author to whom correspondence should be addressed: jan.christoph@ucsf.edu. **URL:** http://cardiacvision.ucsf.edu

ABSTRACT

Electrical waves in the heart form rotating spiral or scroll waves during life-threatening arrhythmias, such as atrial or ventricular fibrillation. The wave dynamics are typically modeled using coupled partial differential equations, which describe reaction–diffusion dynamics in excitable media. More recently, data-driven generative modeling has emerged as an alternative to generate spatio-temporal patterns in physical and biological systems. Here, we explore denoising diffusion probabilistic models for the generative modeling of electrical wave patterns in cardiac tissue. We trained diffusion models with simulated electrical wave patterns to be able to generate such wave patterns in unconditional and conditional generation tasks. For instance, we explored the diffusion-based (i) parameter-specific generation, (ii) evolution, and (iii) inpainting of spiral wave dynamics, including reconstructing three-dimensional scroll wave dynamics from superficial two-dimensional measurements. Furthermore, we generated arbitrarily shaped bi-ventricular geometries and simultaneously initiated scroll wave patterns inside these geometries using diffusion. We characterized and compared the diffusion-generated solutions to solutions obtained with corresponding biophysical models and found that diffusion models learn to replicate spiral and scroll wave dynamics so well that they could be used for data-driven modeling of excitation waves in cardiac tissue. For instance, an ensemble of diffusion-generated spiral wave dynamics exhibits similar self-termination statistics as the corresponding ensemble simulated with a biophysical model. However, we also found that diffusion models produce artifacts if training data are lacking, e.g., during self-termination, and "hallucinate" wave patterns when insufficiently constrained.

I. INTRODUCTION

Waves in excitable media exhibit complex spatio-temporal dynamics.[1,2] In two-dimensional media, they form linear, focal, or rotating spiral-shaped waves or compositions thereof. In three-dimensional media, they manifest as planar or spherical focal waves or, if perturbed, take on more complicated rotational shapes referred to as scroll waves. Spiral and scroll wave dynamics have been studied for many decades, as they are associated with heart rhythm disorders, such as atrial fibrillation, polymorphic ventricular tachycardia, or ventricular fibrillation.[2–12] In the heart, electrical excitation initiates the contraction of the heart muscle, and it is hypothesized that the abnormal, rapid, and irregular contractions during cardiac tachyarrhythmias are caused by spiral- and scroll-shaped waves of electrical excitation.

The electrical waves can be reproduced and studied in computer simulations using biophysical models.[13–15] These models consist of coupled partial differential equations (PDEs), which describe the electrical excitability u and refractoriness r of cardiac muscle cells and the coupling between them; see Eqs. (1) and (2). The equations model reaction–diffusion dynamics, where the exchange of currents through ion channels between cells are modeled as a diffusive process and the cells as nonlinear oscillators. Integrating these equations in time and over space in a spatially extended system using, for instance, the finite difference or finite element method produces nonlinear waves of electrical excitation mediated via diffusion.

Diffusion, on the other hand, is a term that has recently emerged in the field of artificial intelligence (AI), referring to a class of generative neural networks, which employ a diffusive process to generate data.[16–18] The training procedure consists of a forward noising process where noise is iteratively added to the training data. The neural networks, termed denoising diffusion probabilistic models (DDPMs)[18] or diffusion models, learn to reverse this process, ultimately enabling them to create data from noise, as shown in Fig. 1. Diffusion models are very successful in generating data, such as images,[19–21] videos,[22] and audio,[23] and they are increasingly also used for technical applications in physics, engineering, medicine, and biology.[24–28] Diffusion models likely also have many useful applications in cardiology that have yet to be explored. For example, they could be used in electrophysiological studies to generate synthetic action potential wave patterns and arrhythmia morphologies, either to fill in or reconstruct missing measurement data, or to simulate cardiac dynamics in a data-driven fashion. Diffusion-generated solutions could be particularly useful in situations in which measurements can only be obtained partially or indirectly, or when biophysical model equations or parameters are lacking.

In this numerical study, we explore diffusion models for their application in cardiac electrophysiology and arrhythmia research. We investigated whether they can be used to reconstruct or simulate electrical impulse phenomena in computer simulations of excitable media and simulated electrical spiral and scroll waves in two- and three-dimensional square-, bulk-, and heart-shaped tissues with isotropic and anisotropic diffusive spread of the excitation.

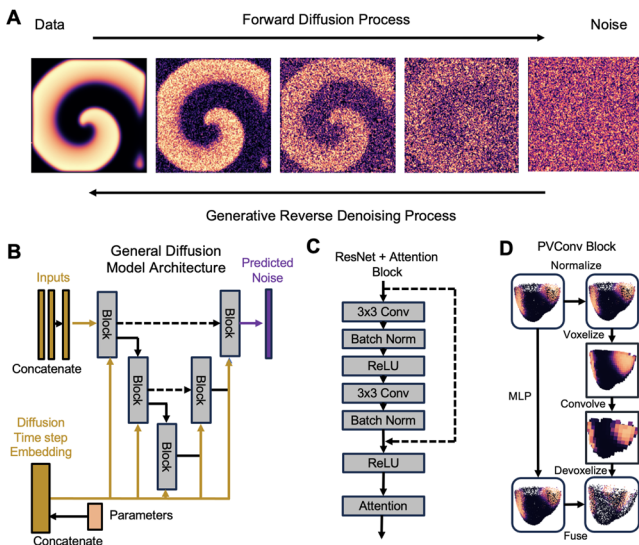More specifically, we used diffusion models for the following tasks.

FIG. 1. Diffusion-based generative modeling of electrical wave dynamics in cardiac tissue. (a) Forward diffusion process and generative reverse denoising process. The training data consist of spiral and scroll wave dynamics in excitable media. (b) General diffusion model architecture for processing image data with underlying U-Net architecture. (c) ResNet Attention block. (d) Diffusion model for generating scroll waves in heart-shaped geometries represented as pointclouds with corresponding scalar-valued data (Point-Voxel Diffusion[29]).

Task 1: Generation of parameter-specific two-dimensional spiral waves; see Sec. III A.
Task 2: Generation of scroll waves in bi-ventricular heart-shapes; see Sec. III B.
Task 3: Evolving spiral wave dynamics over time; see Sec. III C.
Task 4: Reconstruction of three-dimensional scroll waves from two-dimensional surface observations; see Sec. III D.
Task 5: Inpainting of two-dimensional spiral wave dynamics; see Sec. III E.
Task 6: Unconditional generation of two-dimensional spiral wave patterns; see Sec. III F.

We determined how reliable diffusion models are when generating such spatio-temporal physiological dynamics. Generative neural networks, such as diffusion models, generative adversarial networks (GANs), or large language models (LLMs) are known to be capable of producing a continuum of output including false or undesired output, which is often referred to as "hallucination." We show that diffusion models can generate electrical waves in many different ways: out of the blue in an unconstrained generative process or when the generative process is constrained or guided by parameters or boundary conditions such as partial data, or a recent dynamical state of the system. In particular, the latter generative mode corresponds to diffusion-based data-driven modeling of cardiac dynamics. We found that hallucination occurs when the generation task is insufficiently constrained, which raises concerns over the reliability of diffusion models in diagnostic applications.

II. METHODS

A. Simulations of electrical wave dynamics in heart muscle tissue

We simulated nonlinear waves of electrical excitation using coupled partial differential equations (PDEs) in (i) two-dimensional rectangular-shaped, (ii) three-dimensional bulk-shaped, and (iii) three-dimensional heart-shaped geometries, respectively. In all three cases, we used the phenomenological Aliev–Panfilov model,[15]

$$\frac{\partial u}{\partial t} = \nabla \cdot (\boldsymbol{D}\nabla u) - ku(u-a)(u-1) - ur, \tag{1}$$

$$\frac{\partial r}{\partial t} = \left(\varepsilon_0 + \frac{\mu_1 r}{u + \mu_2}\right)(ku(a+1-u) - r). \tag{2}$$

The dynamic variables u and r represent the local electrical excitation and refractoriness in dimensionless, normalized units, respectively. The parameters D, k, a, ε_0, μ_1, and μ_2 determine the properties of the waves (e.g., excitability, wavelength, conduction speed/diffusivity, number of waves, and distance between them). We varied the parameters D and ε_0 to change the properties of the excitation waves and produce different training data for different tasks (Tasks 1–6); see Table I and Secs. III A–III G. The simulations in the simplified (rectangular, bulk) and heart-shaped geometries were performed as described by Lebert, Mittal, and Christoph[30] as well as Lebert *et al.*,[31] respectively. Correspondingly, the system in Eqs. (1) and (2) was integrated using the forward Euler method and the smoothed particle hydrodynamics method,[32,33] respectively. All the simulations were performed in dimensionless units with $\Delta x = 1$

TABLE I. Parameters of the biophysical model (Aliev–Panfilov[15]) used to simulate electrical wave patterns. Task 1: Parameter-specific generation of spiral waves; see Figs. 2 and 3. D' and ε'_0 were varied systematically over a 5×5 grid of values, as shown in Fig. 2(a). Task 2: scroll wave dynamics in a bi-ventricular heart-shaped medium; see Fig. 4. Here, **D** is an anisotropic diffusion tensor; see Lebert et al.[31] The diffusion coefficient $D_\perp$ perpendicular to the fiber orientation was set to 0.1 mm^2/ms; the parallel coefficient $D_\parallel$ was randomly chosen for each simulation from the interval 0.2–0.4 mm^2/ms. Tasks 3, 5, and 6: spiral wave dynamics in 2D isotropic medium; see Figs. 5, 8, and 9. Task 4: scroll waves in anisotropic 3D bulk shown in Fig. 6. Here, **D** is an anisotropic diffusion tensor with $D_\perp = 0.05$ and $D_\parallel = 0.2$; see the work of Lebert, Mittal, and Christoph[30] for details.

Param	Task 1	Task 2	Task 3a/5a	Task 3b/5b	Task 4	Task 6
D	D'	**D**	1	1	**D**	1
K	8	8	8.5	7.5	8	8
A	0.1	0.2	0.1	0.1	0.05	0.05
ε_0	ε'_0	0.002	0.003	0.001	0.002	0.002
μ_1	0.2	0.2	0.16	0.16	0.8	0.2
μ_2	0.3	0.3	0.3	0.3	0.3	0.3

and integration time steps, which proved to be numerically stable. The two-dimensional simulations were isotropic, whereas the three-dimensional simulations were anisotropic with a locally varying fiber direction and faster wave propagation along the fiber direction; see Table I. The fiber architectures were created as described by Lebert, Mittal, and Christoph[30] as well as Lebert et al.[31] The bi-ventricular heart geometries and underlying rule-based fiber architectures were randomly initialized.

The simulation/model parameters were chosen specifically for each task; see Table I. For example, we simulated a range of parameter-specific spiral wave dynamics (Task 1), as shown in Figs. 2(a) and 2(b), by varying the parameters D and ε_0 and by applying a random number of pacing stimuli applied in random locations to cause wave break and create spiral waves. We simulated two different regimes of spiral wave dynamics, as shown in Figs. 5 and 8, using two different parameter sets: one with few (Tasks 3a, 5a) and one with more spiral waves (Tasks 3b, 5b). We simulated scroll wave dynamics in a bulk with $128 \times 128 \times 40$ voxels, as shown in Fig. 6 (Task 4), and in bi-ventricular geometries, as shown in Fig. 4 and described by Lebert et al.[30] (Task 2) using a fixed set of parameters. For each task, we performed hundreds of simulations to generate sufficient training data and separated training data and data used for evaluation. For example, for Task 4, we performed 125 simulations, where 100 simulations were used for training and 25 for evaluation, as described by Lebert, Mittal, and Christoph.[30] The initial conditions u_0, r_0 were randomized and, therefore, different in each simulation; also see the work of Lebert et al.[31] If the spiral or scroll wave dynamics self-terminated prematurely, we restarted the simulation.

Using the simulation data, we generated different training datasets for each task; see Tasks 1–6 presented in Table I and Secs. III A–III G for details. We found that the diffusion models discussed in Secs. III A and III F can already generate spiral wave patterns with as few as 100 training samples; see supplementary material, Fig. 2. However, in order to increase the diversity and quality of the generations, we typically used thousands to tens of thousands of

training samples for all Tasks 1–6; see Table III. Each training dataset consisted of samples randomly chosen from only the different training simulations. Correspondingly, each evaluation dataset consisted of samples randomly chosen from only the evaluation simulations. Training and evaluation datasets were sampled from completely separate datasets. There is no overlap or crosstalk between training and evaluation samples.

B. Denoising diffusion model

We used a denoising diffusion probabilistic modeling[18] neural network architecture, which we refer to as diffusion model for simplicity. Diffusion models consist of a forward diffusion process and a reverse diffusion process, as shown in Fig. 1. During the forward diffusion process, Gaussian noise is added incrementally to an input image until it is indistinguishable from random noise. This produces a sequence of samples $(x_0, \ldots, x_T)$ with increasing noise, starting from the data point x_0 from the real data distribution $q(x)$ and ending with what is indistinguishable from an isotropic Gaussian distribution,

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathbf{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I}). \tag{3}$$

The step sizes are controlled by a variance schedule β_t,

$$q(\mathbf{x}_{1:T} | \mathbf{x}_0) = \prod_{t=1}^{T} q(\mathbf{x}_t | \mathbf{x}_{t-1}). \tag{4}$$

When sampling new data from the data distribution $q(x)$, a model p_θ is learned to estimate $q(x_{t-1}|x_t)$, which is also approximated by a Gaussian distribution. This is referred to as the reverse diffusion process,

$$p_\theta(x_{0:T}) = p(\mathbf{x}_T) \prod_{t=1}^{T} p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t), \tag{5}$$

$$p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) = \mathbf{N}\left(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t)\right). \tag{6}$$

This allows the model p_θ to only have to estimate the two parameters μ and Σ of the estimated denoising step. Commonly, Σ_θ is fixed to a constant variance schedule and is not learnable. This means that in order to estimate p_θ, a model needs to learn $\mu_\theta(x_t, t)$. Electrical wave dynamics can be treated as image-like data, and the U-Net architecture from Dhariwal and Nichol[34] is used to estimate the noise at each step of the reverse diffusion process. The model is trained using pairs taken from the forward diffusion process x_t and x_{t-1} and taking the mean squared error (MSE) between the noise estimated by the model and the true noise at that step.

We implemented six different diffusion models for different tasks (Tasks 1–6); see the results in Secs. III A–III G for additional task-specific details. Throughout this paper, we refer to conditioned and unconditioned diffusion models. Conditioned diffusion models exert a task that is constrained. For instance, the parameter-specific model in Sec. III A is conditioned by the input parameters that guide the diffusion process to produce certain types of wave pattern, and the inpainting model in Sec. III E is conditioned by the surrounding wave pattern as it needs to fill in the missing parts while matching the pattern to the surrounding pattern. While conditioned diffusion models generate wave patterns under certain
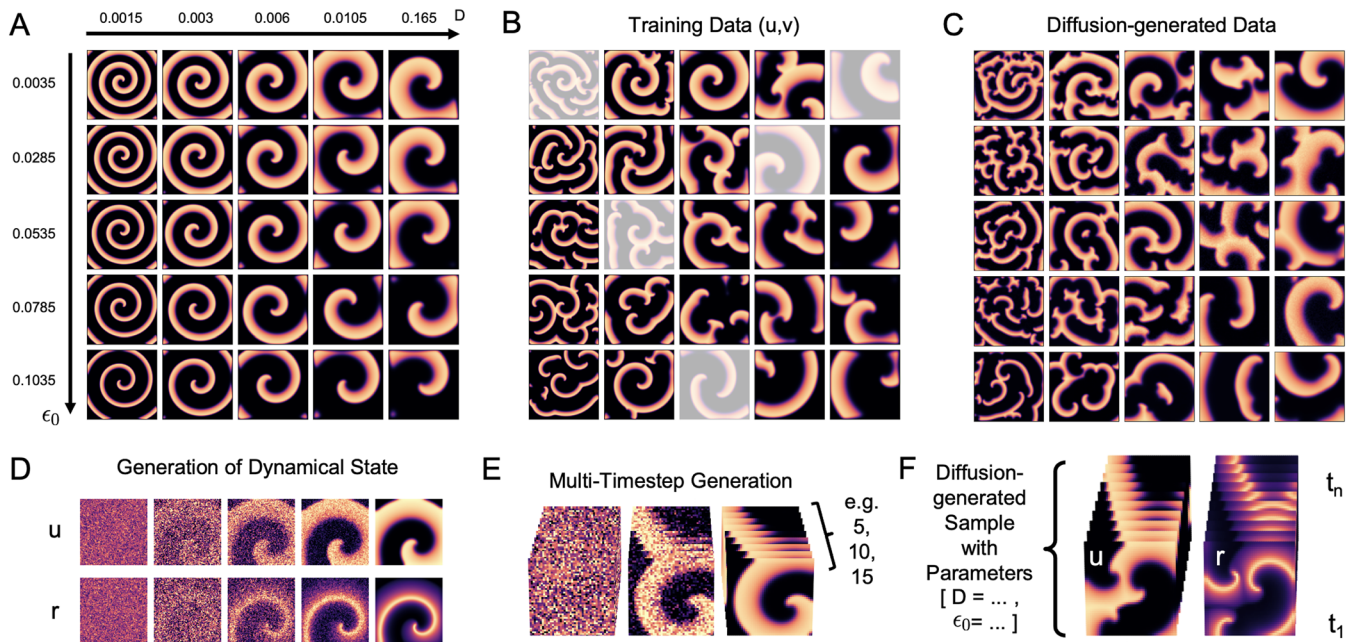
FIG. 2. Parameter-specific generation of spiral wave dynamics using diffusion-based generative modeling (Task 1); see also the supplementary material, Videos 1–5. (a) Different simulations of spiral waves while varying parameters D and ε_0 in Eqs. (1) and (2), or diffusion constant and time scale separation parameter, respectively, which influence the width and distance between the waves, respectively. (b) The diffusion model was trained with data consisting of multi-spiral wave dynamics for the same parameter combinations (D, ε_0) as in panel (a) with 500 simulations per combination. Some parameter combinations (white) were left out during training (5 of 20). (c) The diffusion model generates parameter-specific spiral wave patterns for all parameter combinations, even though it was not trained on all of them. (d) The diffusion model can generate a full dynamical state with both dynamic variables u and r as well as (e) multiple time steps of such states at once: $(u, r)(x, y, t_1, \ldots, t_n)$ with $t_n = 2, 3, \ldots, 15$. (f) Diffusion-generated multi-time-step sample ($t_n = 10$) showing spatio-temporal spiral wave pattern with fast and slow variables (corresponds to $t_s = 100$ simulation time steps; see Sec. III A).

constraints, an unconditional model can dream up any wave pattern without any guidance. All the diffusion models preprocessed the images by down-sampling them to 64×64 in order to reduce memory consumption during training and then up-sampling the generated images to 128×128. We did not see a decrease in performance from doing this down- and up-sampling. The conditioned diffusion models in Secs. III A, III C, and III E were implemented following the work of Saharia et al.[20] using an implementation by Jiang and Belousov.[35] The unconditioned diffusion model in Secs. III F and III G was implemented following the work of Ho et al.[18] using the diffusers library.[36] The diffusion model in Sec. III B was implemented following the work of Zhou et al.[29] using the official codebase. All the diffusion models include a U-Net[37] architecture and were implemented in PyTorch.[38]

C. General training details

The networks were trained using the Adam[39] optimizer with a learning rate of 10^{-4} for the bulk prediction task and 10^{-3} for all other tasks. We used a batch size of 8 for the bulk prediction tasks and a batch size of 32 for all the other tasks. All neural network models were implemented in PyTorch.[38] Training and reconstructions were performed on a NVIDIA RTX A5000 graphics processing unit (GPU); see Table II for an overview of training durations.

TABLE II. Different model sizes (trainable parameters) and training times used in this study. Training was performed on a single NVIDIA RTX A5000 GPU.

Model	Trainable parameters	Training time
Task 1	308, 672, 266	2 days
Task 2	31, 092, 676	0.5 day
Task 3	62, 644, 805	1.5 days
Task 4	965, 266, 792	9 days
Task 5	62, 640, 193	1.5 days
Task 6	113, 673, 219	1 days
Classification	11, 689, 512	5 min

D. Evaluation

We evaluated the diffusion models accuracies using the root mean squared error (RMSE), the mean absolute error (MAE), or the multi-resolution perceptual error (MR)[40] depending on the model and task. We computed the errors per frame, averaging over all the frames of a separate evaluation dataset that was not part of the training dataset. While RMSE and MAE correspond to a measure of the average difference per pixel, MR is a measure for the similarity of two patterns and, in more general terms, for how the waves perceptually look to the human eye. The issue with RMSE and MAE is that

they can produce high errors when images are similar but not perfectly congruent (e.g., a shifted or slightly wider spiral wave pattern, which is otherwise identical). By contrast, the MR calculates the difference between two images over their embedding in feature space and, therefore, provides a much more holistic comparison of two images over multiple spatial scales and feature hierarchies;[40] see also Fig. 8 and Sec. III E. We used MR in addition to RMSE and MAE to overcome their limitations related to a crude pixel-wise comparison. MR captures when two patterns are qualitatively very similar, which RMSE and MAE do not capture per se.

III. RESULTS

A. Parameter-specific generation of spiral wave dynamics

Diffusion models can generate parameter-specific spiral wave patterns when given a set of parameters as input, as shown in Fig. 2. We trained a diffusion model to generate a parameter-specific spatio-temporal spiral wave pattern from 2 input parameters D and ε_0 (Task 1), also shown in Figs. 2(d) and 2(e),

$$(\xi(x,y), D, \varepsilon_0) \rightarrow (\tilde{u}, \tilde{r})(x, y, t_1, \ldots, t_n). \tag{7}$$

Here, $\xi(x,y)$ is the initial noise; D and ε_0 are parameters of the biophysical Aliev–Panfilov model in Eqs. (1) and (2), which influence the spiral wave properties; and $(\tilde{u}, \tilde{r})$ is the generated spatio-temporal spiral wave pattern consisting of multiple time steps t_n, as shown in Figs. 2(d)–2(f). The diffusion model generates both dynamic variables u, r and multiple frames $(t_1, \ldots, t_n)$ at once (here usually $n = 5$). We refer to this generation process as multi-time-step generation 'conditioned by the parameters D and ε_0'. We found that the generative process can be conditioned and consequently guided by parameters to produce spiral wave dynamics with specific properties, which equally arise with a corresponding biophysical model with the same parameters, as shown in Figs. 2(b) and 2(c). Spiral wave shapes can be very different depending on the model parameters; see Figs. 2(a), 2(b), and 9 in Qu et al.[41] and Figs. 5–9 in Bartocci et al.[42] Here, the Aliev–Panfilov model produces spiral waves with wider/thinner arms and longer/shorter diastolic intervals when varying the parameters D and ε_0 in Eqs. (1) and (2), as shown in Fig. 2(a). Our diffusion model reproduces these different parameter-specific regimes when conditioned with the respective parameters, as shown in Figs. 2 and 3. Importantly, the diffusion model can generate parameter-specific spiral wave dynamics, even if the parameter combination was not part of the training data. However, it fails to generate plausible wave patterns outside the distribution of training data, as shown in the supplementary material, Fig. 3, which is generally known to be true for many deep learning models.

We performed 12 500 unique simulations in total for $5 \times 5 = 25$ different parameter pairs (ε_0, D) or 500 simulations per parameter pair. The simulations were performed with combinations of (ε_0, D), as shown in Fig. 2(a). In each simulation, we initialized the spiral waves shown in Fig. 2(a) and then applied a random number of pacing stimuli (between 10 and 40) in random locations to cause wave break and create single- or multi-spiral wave dynamics, as shown in Fig. 2(b). Only every tenth simulation time step over a period of 10 000 simulation time steps was written out, resulting

in 1000 frames showing about 2–3 spiral wave rotations. Per simulation, we extracted 5 multi-time-step training samples showing each a unique spatio-temporal spiral wave pattern, yielding 2500 samples in total per parameter pair or 62 500 samples in total over the grid of 5×5 parameter pairs. Each sample consists of $n = 5$ frames $\{\{u(x,y,t_1), r(x,y,t_1)\}, \ldots, \{u(x,y,t_n), r(x,y,t_n)\}\}$, which are 3 frames apart, effectively covering a period of $t_s = 150$ simulation time steps. Each sample shows a unique spiral wave pattern, and there is no overlap between training samples. We augmented the data by randomly flipping or rotating all frames in a sample by multiples of 90°, effectively increasing the training dataset size by a factor of 8 and ensuring rotational invariance. It should be noted that most spirals shown in Fig. 2(b) are clock-wise rotating as they are simulated data before augmentation. In principle, the number of frames in a training sample can vary (e.g., 5, 10, and 15), but we resorted to 5 for simplicity. The parameters were first encoded using sinusoidal embeddings.[43] We then conditioned the diffusion model by concatenating these sinusoidal parameter encodings to the diffusion time step embedding that is passed into each residual connection in the underlying U-Net,[34] as shown in Fig. 1. Aside from the parameter conditioning, the generation was unconditioned, allowing the diffusion model to dream up any spiral wave pattern.

We verified the parameter-specificity of the diffusion model by initializing the biophysical model from Eqs. (1) and (2) with the first diffusion-generated frame $\{\tilde{u}(x,y,t_1), \tilde{r}(x,y,t_1)\}$, as shown in Fig. 3(a), and integrating the biophysical model for $t_s = 150$ time steps $t_1 \rightarrow t_2 \rightarrow t_3 \rightarrow \cdots \rightarrow t_s$ using either the same parameter combination (ε_0^*, D^*) or a mismatching parameter combination (ε_0, D) to see if the PDE-evolved solutions co-evolve with the spatio-temporal spiral wave pattern generated by the diffusion model; see Fig. 3(b). It should be noted that the diffusion sample times $\{t_n = 1, 2, 3, 4, 5\}$ correspond to $\{t_s = 1, 31, 61, 91, 121, 151\}$ in simulation time steps because of the subsampling during the training data creation (every tenth frame stored from simulation) and training procedure (every third frame used for one training sample). We can compare the two solutions because the diffusion-generated spatio-temporal spiral wave pattern shows a plausible spatio-temporal progression of the wave pattern. Comparing the simulated state $\{u(x,y,t_s), r(x,y,t_s)\}$ at time $t_s = 151$ to the corresponding last state $\{u(x,y,t_n), r(x,y,t_n)\}$ with $t_n = 5$ in the diffusion-generated sample, as shown in Fig. 3(d), we found that the average pixel-wise error (MAE) is smallest with matching parameters $(\varepsilon_0, D) = (\varepsilon_0^*, D^*)$, as shown in Fig. 3(e), regardless of whether they were part of the training data or not. In other words, the diffusion model generates spiral wave dynamics that the biophysical model also produces with the same parameters. More precisely, we found that the error was minimal with matching parameters for only 21 of the $5 \times 5 = 25$ parameter combinations. In the other four cases, it was a nearby combination and the difference in the error was very small. The four ambiguous cases occurred in the central lower left area of the 5×5 parameter grid with medium to thin waves. The ambiguity could result from the waves being more similar to each other or, vice versa, harder to distinguish when comparing the divergence of the wave patterns in our measurements. Correspondingly, we believe the issue will resolve when including longer trajectories or integration times in our measurements. The plots shown in Fig. 3(e) were derived from averaging over 100 simulations initialized with the first
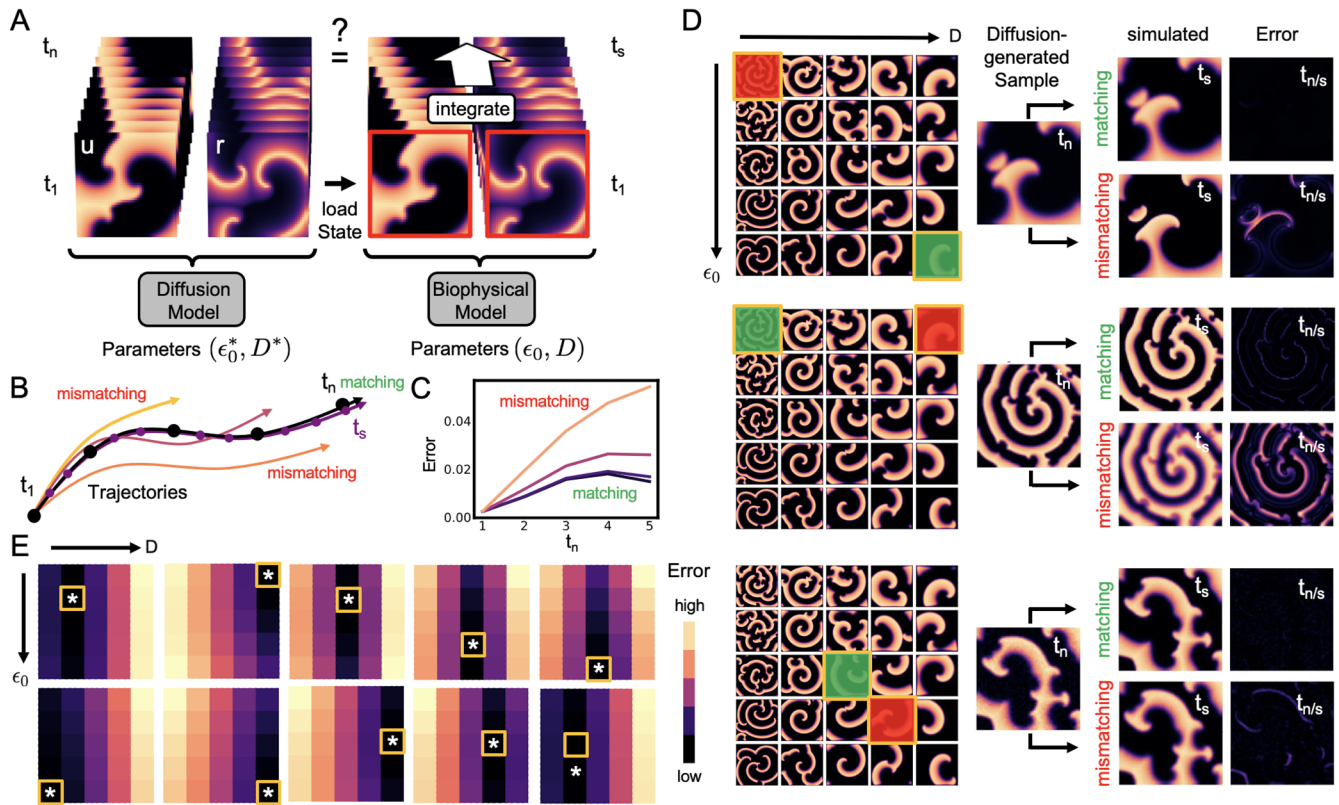
FIG. 3. Scheme to verify whether the diffusion-generated spiral wave patterns shown in Fig. 2 are parameter-specific: (a) The first state (u, r) at t_1 of the diffusion-generated multi-time-step sample generated with parameters (ϵ_0^*, D^*) was loaded into the corresponding biophysical model with either identical or mismatching parameters (ϵ_0, D). The biophysical model was then integrated for t_s time steps, and the solutions were then compared to the last state of the diffusion-generated sample at t_n. This was repeated for all parameter combinations (ϵ_0, D); see Sec. III A for details. (b) Trajectories in phase space starting from state t_1, co-evolving with matching parameters and diverging with mismatching parameters. (c) Error between diffusion-generated and simulated states over time ($t_n = 5$ corresponds to $t_s = 150$) with matching (low error: black/purple) and mismatching (high error: orange/yellow) parameters for one diffusion-generated sample. (d) Examples of diffusion-generated and simulated states at $t_n = t_s$ with matching parameters $((\epsilon_0^*, D^*) = (\epsilon_0, D))$ and mismatching parameters [5×5 grid with the same parameter combinations as shown in Figs. 2(a) and 2(b)]. The diffusion-generated sample was generated with the parameter combination indicated by the green square. The matching biophysical simulation was performed with the same parameters, while the mismatching simulation was performed with the combination indicated by the red square. The simulations deviate from the diffusion-generated samples when the parameters do not match, which causes an error (pixel-wise absolute difference). (e) Confirmation that generations are parameter-specific: the average pixel-wise error (MAE) between simulated and diffusion-generated patterns at t_n (averaged over 100 samples/simulations) is the lowest (black/purple: small, orange/yellow: large) for matching parameter combinations $(\epsilon_0, D) = (\epsilon_0^*, D^*)$. □ parameter combination (ϵ_0^*, D^*) used to generate diffusion-generated sample; ∗ parameter combination of simulation (ϵ_0, D) with the lowest error; □ and ∗ match in 21 of 25 cases; in the other cases, the minimum is nearby with marginal difference in the error. 5×5 grid corresponds to same parameter combinations as shown in Fig. 2(a).

frames of 100 different diffusion-generated samples per parameter combination. It should be noted that the diffusion model was not trained on all parameter combinations. The findings suggest that diffusion models do not need to be trained meticulously on all possible parameter combinations, can interpolate in parameter space, and generate wave dynamics for many more parameter combinations than just the ones they were trained on.

B. Generation of re-entrant scroll waves in heart-shaped geometries

We trained a diffusion model to generate scroll waves in bi-ventricular-shaped geometries (Task 2), as shown in Fig. 4. Panel D shows the denoising diffusion process used to generate

bi-ventricular-shaped point clouds with corresponding excitation values per point, where both the shape and the electrical wave pattern are generated simultaneously by using the diffusion model (two representative examples). Figure 4(f) shows further examples of diffusion-generated scroll waves, which are visually indistinguishable from the scroll wave patterns shown in panel E, which were simulated using the biophysical model in Eqs. (1) and (2). The generative process was completely unconstrained and not explicitly conditioned. Correspondingly, the model generates any scroll wave pattern in any bi-ventricular shape that it can come up with given what it has learned from the training data. The scroll wave patterns are anisotropic because the training data were simulated with anisotropic ventricular fiber architecture. The training data, therefore, implicitly warrant anisotropy during the generations. Even
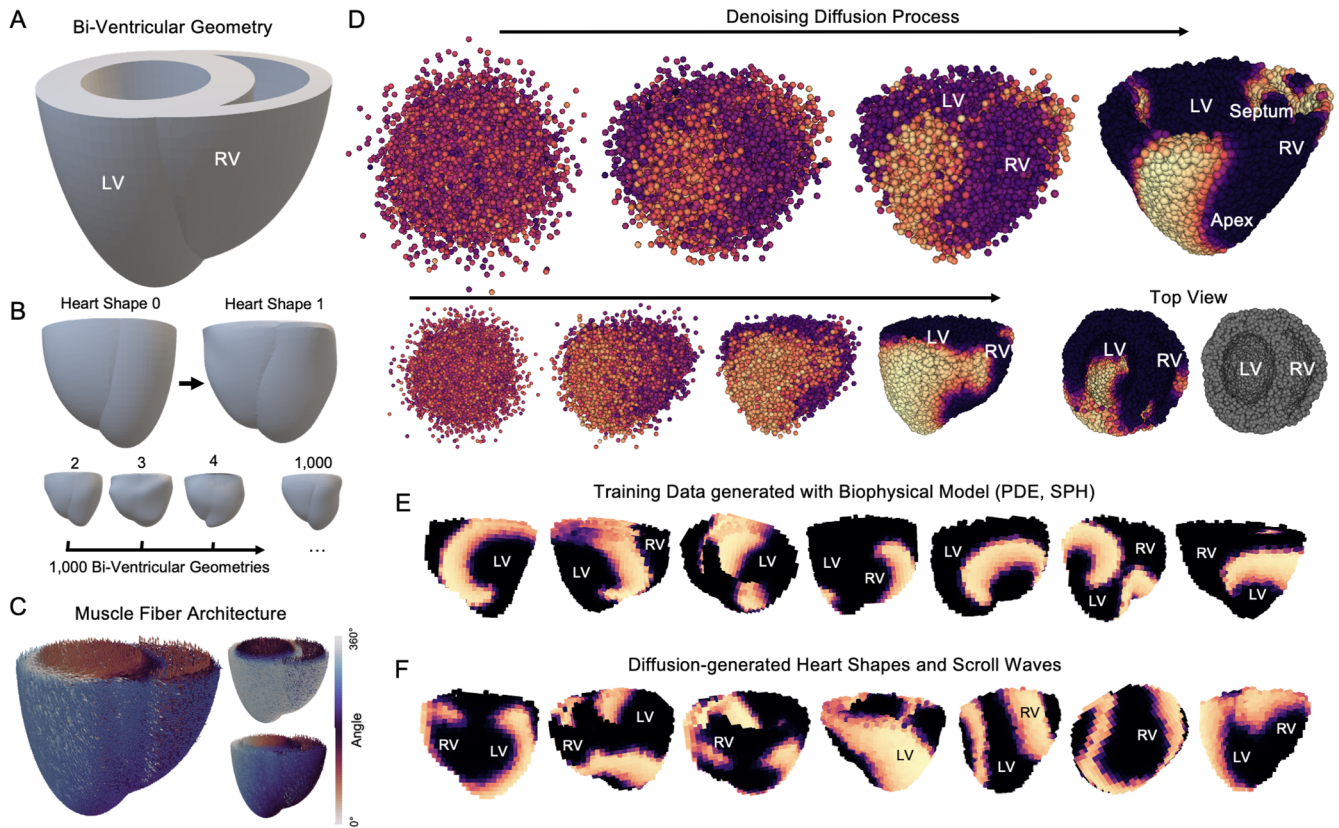
FIG. 4. Diffusion-based modeling of re-entrant electrical waves in heart-shaped bi-ventricular geometries (Task 2); see also the supplementary material, Video 6. (a) Template geometry used in simulations to generate training data. (b) 1000 randomized, unique variations of template geometry to create unique training samples, as described in Ref. 30. (c) Geometry-dependent bi-ventricular muscle fiber architecture initiated in each simulation. (d) Denoising process during diffusion-based generation of electrical scroll wave pattern in bi-ventricular heart shapes. Both the tissue geometry and wave pattern are generated simultaneously. (e) Training data used to train diffusion model consisting of 5000 training samples showing electrical scroll wave patterns. Each simulation consists of 32 000 particles, subsampled to 16 000 particles for training, here voxelized and volume-rendered for visualization. The training data were simulated using a biophysical model (Aliev–Panfilov), see Eqs. (1) and (2), and integrated using the SPH-method,[32,33] see Sec. II A. (f) Additional examples of diffusion-generated electrical scroll waves in bi-ventricular heart shapes. The diffusion model generates a bi-ventricular shape (each shape different) as well as a full dynamical state with both dynamic variables $(u, r)(\vec{x})$. The scroll wave patterns are anisotropic due to the specific ventricular muscle fiber organization in the training data.

though the diffusion model generates only a single scroll wave pattern $u(\vec{x}, t)$, and only the excitatory variable u, it is easy to imagine how this pattern could also be evolved over time $u(\vec{x}, t_1, t_2, \ldots)$ as described in the next Sec. III C and shown in Fig. 5.

The simulations were performed as described by Lebert *et al.*[30] Accordingly, the simulated training data consist of point clouds of $i \sim 32\,000$ vertices, $p(\vec{x})_i$ representing bi-ventricular heart shapes. The simulations were performed with 1000 unique bi-ventricular shapes created from a template geometry; see Figs. 4(a) and 4(b). Accordingly, the diffusion model comes up with similar shapes during the generative process. The excitatory variable u_i is defined per vertex i. We down-sampled the data to 16 000 points and used Point-Voxel Diffusion[29] trained on 5000 training samples obtained from the simulations where each training sample consists of a single point cloud of excitation values $u(\vec{x}, t)$ at a particular time t. We trained the model to output 16 384 points (with a latent dimension of 512) for 400 epochs.

C. Generative diffusion-based simulation of electrical wave dynamics

Electrical impulse propagation in the heart is usually simulated by integrating partial differential equations in space and over time by using, for example, the finite difference or finite element methods. We trained a diffusion model to calculate an immediate future time step of a given spatio-temporal excitation wave pattern (Task 3),

$$(u, r)(\vec{x}, t) \rightarrow (\tilde{u}, \tilde{r})(\vec{x}, t + \tau), \tag{8}$$

where (u, r) are the dynamic variables from Eqs. (1) and (2) and τ is an infinitesimal temporal increment or the integration time step. In other words, we employed diffusion-based data-driven modeling to evolve electrical wave dynamics over time rather than using a biophysical model to simulate the dynamics. More precisely, we trained a diffusion model to predict the next five time steps from
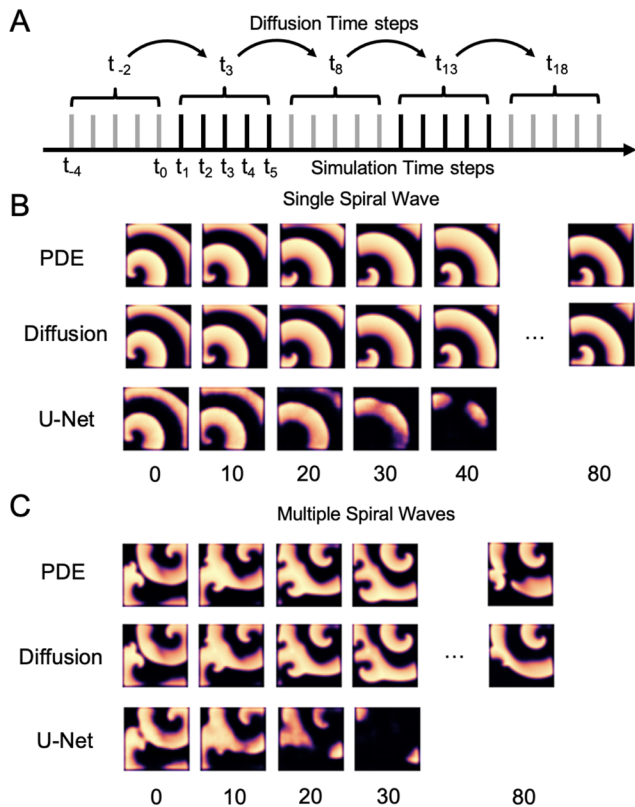
FIG. 5. Data-driven modeling of spiral wave dynamics using diffusion models. (a) Spatio-temporal prediction (Task 3) of five frames from previous five frames; see Sec. III C. Panels (b) and (c): comparison of ground-truth data (GT) simulated with the biophysical model (finite differences) and data-driven methods (Diffusion vs U-Net) to evolve the wave pattern. With a single spiral wave, the output of the diffusion model is visually indistinguishable from the ground-truth for many rotations, while U-Net quickly fails to sustain the wave pattern. With more complicated wave patterns, the diffusion models begins to deviate from the biophysical model after 2–3 spiral rotations (80 time steps).

the previous five time steps of the dynamics, resulting in an integration scheme that updates a brief spatio-temporal pattern instead of a static spatial pattern,

$$(u, r)(\vec{x}, t_{-4}, \ldots, t_0) \rightarrow (\tilde{u}, \tilde{r})(\vec{x}, t_1, \ldots, t_5). \quad (9)$$

Here, $\{t_{-4}, t_{-3}, t_{-2}, t_{-1}, t_0\}$ are the four previous time steps, and the current time step t_0 and $\{t_1, \ldots, t_5\}$ are the next five time steps predicted by the model, as shown in Fig. 5(a). We found this multi-time-step prediction scheme more stable than updating the dynamics one time step at a time or predicting the next time step from the previous n time steps in an auto-regressive manner. We found empirically that using five time steps to predict the next five time steps was a good compromise between performance and training time. We tested using ten time steps to predict the next ten time steps, which worked as well (and presumably better), but the model was consequently bigger and the training time much larger. In all cases, we used no temporal subsampling. We conditioned the diffusion model

by concatenating $(u, r)(\vec{x}, t_{-4}, \ldots, t_0)$ to the initial noisy distribution $\xi(x, y)$, adding five channels to the input of the underlying U-Net ($10 \times 128 \times 128$ pixel³). We trained and evaluated the model with 15 000 and 5000 samples, respectively. The training samples show either simple or complex two-dimensional spiral wave dynamics simulated with two parameter sets; see Table I (Task 3a/b) and Figs. 5(b) and 5(c). The same parameters were also used in Task 5 and shown in Fig. 8. For each of the two parameter sets, we ran 100 simulations and sampled the training samples from 75 simulations and the evaluation samples from 25 simulations, respectively.

Figures 5(b) and 5(c) show the ground-truth (PDE) spiral wave patterns for up to 80 simulation time steps and the corresponding evolved spiral wave patterns predicted either with our diffusion model or a correspondingly trained U-Net model. While the dynamics quickly degenerate with U-Net, the diffusion model successfully sustains and evolves the dynamics over a very long time. The ability to sustain the wave pattern is likely related to diffusion models being able to learn and mimic shapes. The diffusion model produces spiral waves with either stable or meandering cores, which exhibit breakup and (self-) interactions. With the single spiral wave shown in Fig. 5(b), the diffusion model's output matches the biophysical model's output for many rotations; see the supplementary material, Video 7. Eventually, the original (ground-truth) dynamics diverge from the diffusion-generated dynamics, which, to some extent, is to be expected as the dynamics would also diverge with, for instance, two different classical integration methods. With the more complex spiral wave dynamics shown in Fig. 5(c), the diffusion output diverges rapidly within less than two rotations of the spiral wave pattern; see the supplementary material, Video 8. Interestingly, the diffusion model appears to favor more stable wave dynamics (less wave break); see the right panel in Fig. 5(c). We can only speculate that this could be related to a bias in the training data, e.g., an underrepresentation of finer spatial scales; also see Sec. III G for further details regarding the ensemble behavior of the dynamics.

The diffusion-based time stepping appears to only work well with spatio-temporal data, which suggests that spatio-temporal data are unique enough so that the model is sufficiently constrained (or conditioned), which, in turn, enables it to predict the next spatio-temporal segment reasonably well. However, how these findings generalize to various dynamical regimes with different Lyapunov times warrants further research.

Updating the dynamics in a 128×128 pixel² simulation domain takes 1.1 ms on a NVIDIA A5000 GPU per multi-frame prediction. Together with the results in Sec. III B, our findings suggest that diffusion-based modeling could be used to simulate spatio-temporal dynamics in the heart.

D. Reconstruction of three-dimensional scroll wave dynamics from surface observations

Measuring electrophysiological wave phenomena beneath the heart surface is a long-standing challenge in cardiovascular research and diagnostics. Catheter electrodes or optical mapping provide only superficial data from the heart surface, and intramural measurements from within the heart muscle with electrodes are sparse. To address this challenge, various numerical methods were introduced, which aim at reconstructing transmural wave patterns from observations of the dynamics on the tissue's surface.[30,44–46] The

numerical reconstructions are particularly relevant in the context of tachyarrhythmias, such as ventricular or atrial fibrillation, as they may provide a better understanding of the underlying three-dimensional spatio-temporal organization of the electrical waves within the heart muscle. Recently, Lebert *et al.*[30] and Stenger *et al.*[44] demonstrated that convolutional encoding–decoding neural networks (different U-Net-types) can be used to reconstruct three-dimensional scroll wave dynamics inside a thick bulk-shaped excitable medium from two-dimensional observations of the dynamics on the top and/or bottom surfaces (representing the epi- and endocardium). At the same time, Stenger *et al.* also demonstrated this briefly with a diffusion model.[44] However, several aspects of the deep learning-based reconstructions remain underexplored, in particular with the diffusion-based approach.

We trained a diffusion model to predict three-dimensional scroll wave dynamics inside a bulk from two-dimensional observations of the dynamics on the surface of the bulk (Task 4), as described by Lebert *et al.*[30] and shown in Fig. 6. More specifically, the model was trained to predict a single three-dimensional snapshot of the excitatory variable $u_t(x, y, z)$ at a given time t at every voxel in a bulk with $128 \times 128 \times 40$ voxels from the five previous two-dimensional snapshots of the dynamics on the bulk's surface,

$$(u_1(x, y), \ldots, u_5(x, y)) \rightarrow \tilde{u}_5(x, y, z), \tag{10}$$

where $\tilde{u}$ is a prediction of the true dynamics u. The snapshots were measured either (i) on the top surface only (single-surface mode), resulting in a spatio-temporal measurement consisting of five snapshots $[u_1(x, y, 1), \ldots, u_5(x, y, 1)]$ or (ii) on the top and bottom surface (dual-surface mode), resulting in $2 \cdot 5$ snapshots $[u_1(x, y, 1), u_1(x, y, 40), u_2(x, y, 1), \ldots, u_5(x, y, 40)]$, as described by Lebert *et al.*[30] The five snapshots were sampled at equidistant times at $t_{-4\tau}, t_{-3\tau}, t_{-2\tau}, t_{-\tau}, t_0$ with $u_i = u(t_i)$ over about one rotational period T of the scroll wave dynamics ($\tau \approx T/5$), which we found to provide sufficient information to reconstruct the dynamics, as described in Refs. 30, 47, and 48. Accordingly, we conditioned the diffusion model by concatenating these sequences as additional channels in the U-Net inputs (interleaved in the dual-surface mode, odd indices for the top layer and even indices for the bottom layer).

To explore an alternative extension of our reconstruction approach, we conditioned the diffusion model using the output of a generic U-Net model, which was trained and applied as described in Refs. 30, to create a combined model that can potentially take advantage of the strengths of both the U-Net and diffusion models; also see Fig. 7. Accordingly, we conditioned the combined model
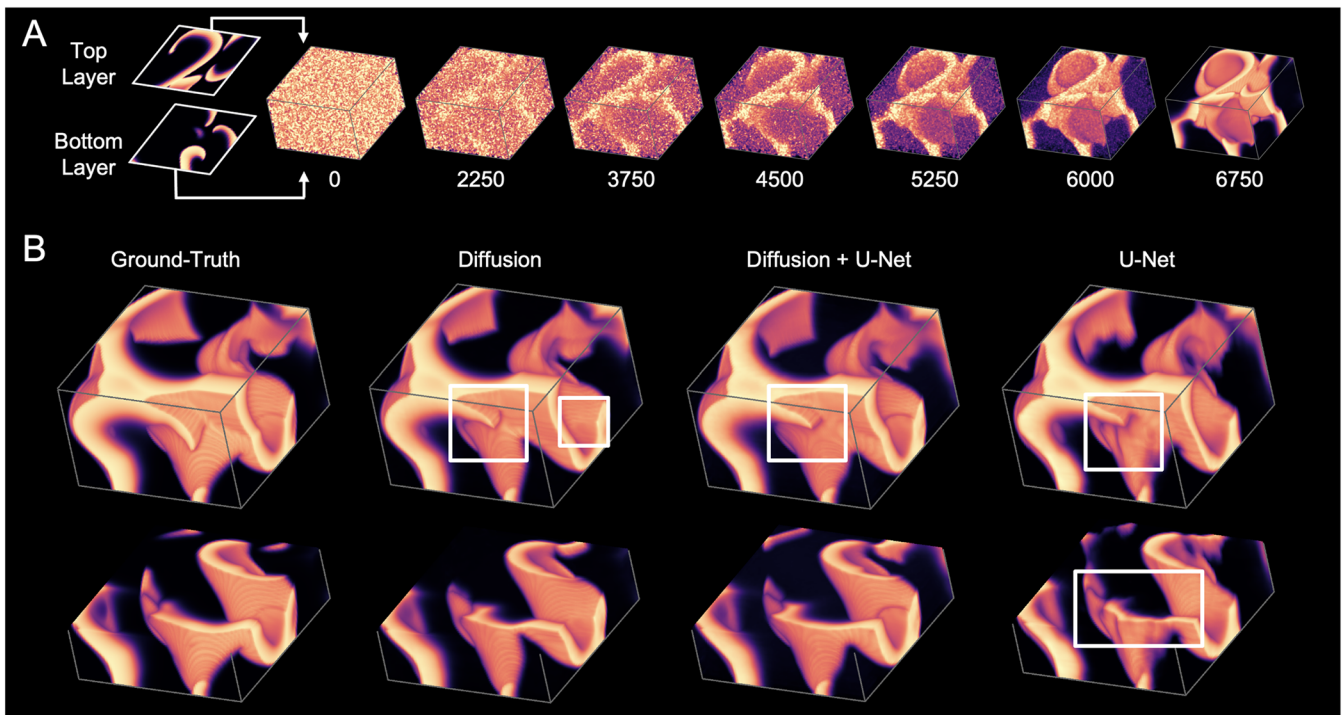

FIG. 6. Diffusion-based reconstruction of scroll wave dynamics inside a three-dimensional bulk from two-dimensional observations of the dynamics on the bulk's top and bottom surface (Task 4); also see the supplementary material, Video 9. The bulk is fully opaque and measurements can only be obtained from its surface. (a) Illustration of the diffusion process over denoising iterations. (b) Scroll wave dynamics (left: ground-truth) and reconstructed scroll waves with diffusion (center left), U-Net (right), and U-Net refined with diffusion (center right); also see Refs. 30 and 44. The U-Net is from our previous study.[30] The top row shows the full bulk, while the bottom row shows the lower half of the bulk to highlight the central midwall layer of the bulk, which exhibits the highest reconstruction error. While diffusion produces smoother wave patterns than U-Net, particularly at deeper layers, the overall reconstruction accuracies are not significantly different across the three approaches; also see Fig. 7. The white squares highlight slight differences between reconstructions and ground-truth. The bulk is $128 \times 128 \times 40$ voxels (aspect ratio was altered to emphasize transmural wave morphology); also see Ref. 30. The simulation parameters are presented in Table I (Task 4).

with the sequences of 5 (or 2 · 5) two-dimensional snapshots and 1 three-dimensional prediction $\tilde{u}(x, y, z)$ of the U-Net model, which, in turn, also analyzed 5 snapshots as input. The two- and three-dimensional inputs were concatenated to obtain $(128 \times 128 \times 45)$ or $(128 \times 128 \times 50)$ samples in single- vs dual-surface mode as conditions, respectively. This leads to a total of four conditioning modes that we tested (single- vs dual-surface, diffusion vs combination of diffusion + U-Net). Generally, the different model versions required three-dimensional input samples, e.g., $(128 \times 128 \times 5)$ or $(128 \times 128 \times 45)$. To denoise a $128 \times 128 \times 40$ volume image with $5 \cdot 128 \times 128$ snapshots as conditioning, the model corresponds to an $R(128 \times 128 \times 45) \rightarrow R(128 \times 128 \times 40)$ function. However, internally, because the denoising diffusion process works on the intermediate noisy bulk data, the overall data consist of the conditioning data concatenated to the noisy data, which then results in an array size of, for example, $128 \times 128 \times 80$.

Training was performed with 20 000 training samples, which were generated in 100 simulations; also see Sec. II A, and the model was evaluated on 5000 separate samples. The simulated bulk is completely opaque (only surface voxels can be observed) and thick enough to sustain three-dimensional scroll wave dynamics (128 × 128 × 40 voxels, 1–2 scroll wavelengths), as shown in Fig. 6. We used the same simulation data and parameters as Lebert et al.;[30] also see also Table I (Task 4).

Figure 6(a) shows the denoising process during the scroll wave prediction task in the bulk using the diffusion model. The scroll wave pattern is reconstructed from the top and bottom surface layers.
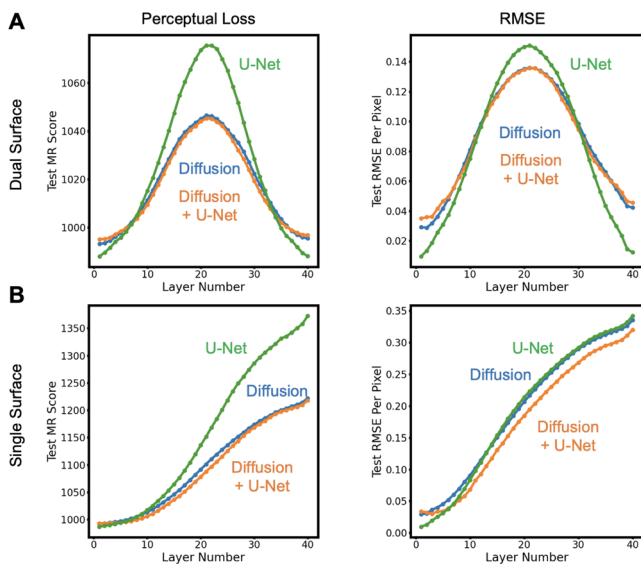
FIG. 7. Transmural reconstruction error per layer number or depth with diffusion (blue), U-Net (green, from the work of Lebert, Mittal, and Christoph[30]), or diffusion + U-Net (orange); also see Fig. 6. The reconstruction error was quantified using either the perceptual error[40] (left) or absolute difference (RMSE, right). (a) Dual-surface reconstruction analyzing the top and bottom layers: U-Net performs worse at midwall, but slightly better closer to the surfaces than diffusion. Perceptual error and RMSE produce very different error profiles. (b) Single-surface prediction: the perceptual error increases more steeply with U-Net as with diffusion, while all models perform similar to RMSE.

Interestingly, the rough shape of the scroll wave pattern is already captured early in the denoising process, while later stages enhance finer structures. Figure 6(b) shows a comparison of the predictions obtained with diffusion vs U-Net vs the combination of the two with diffusion refining the U-Net output; also see Fig. 7. All the models are able to predict three-dimensional scroll waves from two-dimensional observations using either only the top (single-surface mode) or both the top and bottom surface layers (dual-surface mode); also see Fig. 7. The diffusion model slightly outperforms U-Net, but their combination does not significantly increase the reconstruction accuracy beyond the accuracy of the diffusion model. Stenger et al.[44] found that diffusion performs substantially better than U-Net with long observations (32 snapshots). Here, we used fewer observations (only five snapshots), which likely causes this discrepancy.

Most importantly, the diffusion-based reconstructions exhibit one striking feature: while U-Net reconstructions become fuzzier with increasing depth, diffusion maintains the shape, smoothness, and overall look of the scroll waves much better throughout the bulk. This is also reflected by the perceptual error, as shown in Fig. 7 and also Sec. III E. However, even though the visual impression suggests otherwise, we find, on average, no dramatic improvement of the overall reconstruction accuracy (RMSE) with diffusion over U-Net; see Fig. 7. Upon closer inspection, one notices that diffusion produces minor mismatches at deeper layers [white boxes shown in Fig. 6(b)], suggesting that its output looks better than it is and is not necessarily more accurate than with U-Net; also see Sec. III E. We hoped that guiding the diffusion model with the output from the U-Net model could mitigate these issues, but, on average, the error remained the same, as shown in Fig. 7. Unlike in the work of Stenger et al.,[44] our model produces smooth scroll wave patterns without residual noise. Our diffusion model predicts the bulk at once and not layer by layer, which could cause the smoother appearance of the waves.

It is important to highlight that our deep-learning-based scroll wave reconstruction approach was only trained on the Aliev–Panfilov scroll wave dynamics and, therefore, assumes scroll waves inside the tissue. All three types of neural networks were trained with thousands of corresponding pairs of three- and two-dimensional data of scroll waves and observations thereof. Therefore, the training data implicitly restrict the approach to a particular distribution of data and its characteristics (specific electrophysiological model that produces waves with a particular shape, isotropic vs anisotropic wave patterns, wavelength relative to medium thickness), and the approach is task-specific (single- vs dual-surface observations). It would be interesting to test how the reconstructions perform and what type of waves they produce with significantly different data.
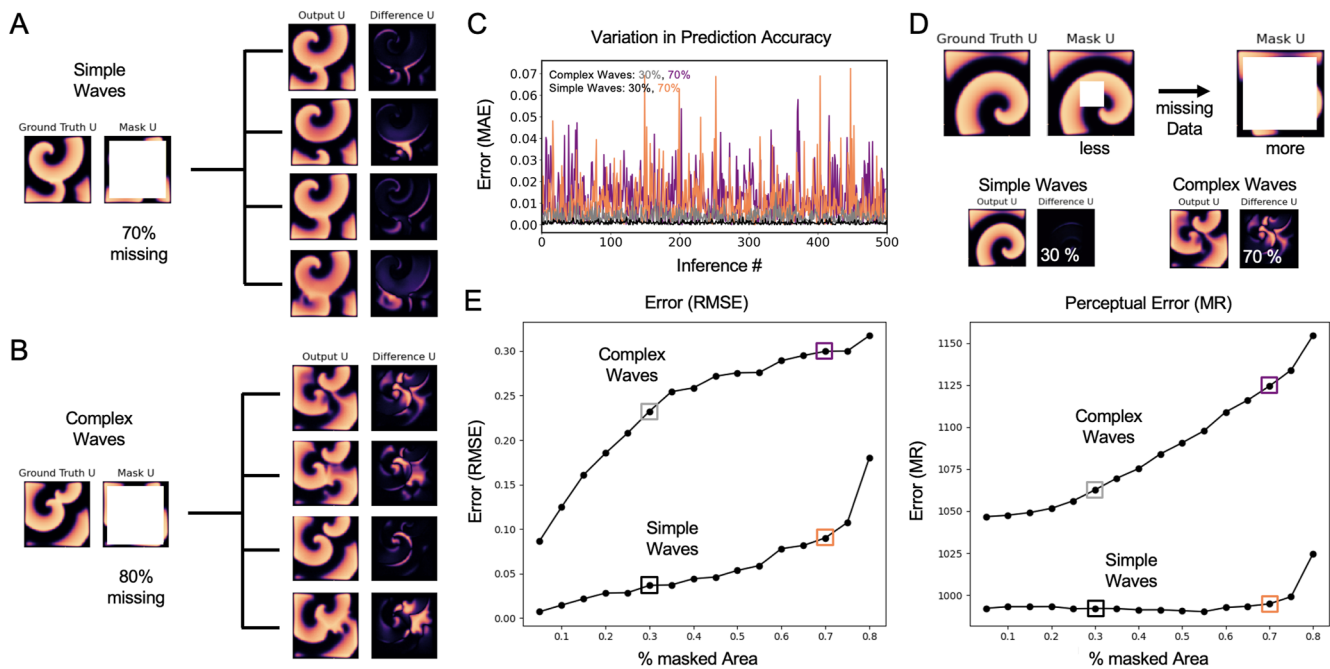
E. Hallucination during inpainting of 2D spiral wave dynamics

Generative models are known to hallucinate, which means that they may generate output that looks, sounds, or reads convincing but is inaccurate.[49,50] For example, recent large language models (LLMs) are known to confidently present made-up knowledge as if it was factual. In computer vision, diffusion models may generate unexpected scenes, which are abstract and not part of human day-to-day

experience. While it is easy to identify hallucination in diffusion-generated visual scenes, it is not necessarily obvious with spiral or scroll wave patterns when they include hallucinations; also see Sec. III F. In Fig. 6, the diffusion-generated reconstructed scroll wave patterns at midwall look convincing and can be misinterpreted as accurate solutions, but they are just as inaccurate as the output from the U-Net model.

We further explored this hallucinating behavior in a two-dimensional inpainting task of spiral wave dynamics (Task 5), as shown in Fig. 8, and can confirm that hallucination occurs, particularly when the task is insufficiently constrained; also see the supplementary material, Video 10. We varied the size of a square region at the center of the medium, within which the diffusion model was tasked to interpolate the missing spiral wave pattern from the surrounding data. Hallucination is minimal with a small square, which is reflected by the error (RMSE) on the left sides of the graphs as shown in Fig. 8(e). However, we observed that the diffusion model comes up with many different spiral wave patterns when the square region is large or more data are missing; see Figs. 8(a) and 8(b) and the supplementary material, Video 10. Figure 8(c) shows the variation of the diffusion model output when the same task is repeated 500 times with simple vs complex waves and 30% vs 70% missing

data, respectively. The variation in the output as quantified by the error (MAE) between the ground-truth and the individual predicted spiral wave pattern is particularly large, with 70% missing data. Hallucination also becomes stronger when the wave dynamics are more complicated (we tested two parameter sets; see Table I, Task 5a/b); compare panels A and B and the upper and lower curves in panel E (average error calculated over 500 unique samples per data point) in Fig. 8. In Fig. 8(e), we compared the root mean squared error (RMSE), which reflects the pixel-wise congruency of the ground-truth and the predicted pattern, with a perceptual error (MR), which reflects similarities or differences in the patterns independently from spatial mismatches (as it is calculated on the embedding of the pattern). The perceptual error indicates that with simple waves, the variations in the output of the diffusion model are small regardless of the size of the masked area. In other words, differences in waves only correspond to slight spatial mismatches (which cause high RMSE), while the wave shapes are very similar qualitatively. The spikes in panel C (with 70%), on the other hand, correspond to large qualitative changes of the wave pattern, which occur occasionally with both complex and simple waves. Our data suggest that (i) the diffusion model hallucinates if it has the freedom to generate many potential solutions to a problem; (ii)

FIG. 8. Hallucination of the diffusion model during inpainting of electrical spiral wave dynamics (Task 5). (a) The diffusion model predicts missing data inside a square region (white) at the center of the medium. With most of the data missing (70%), the diffusion model dreams up wave patterns that look convincingly like spiral wave patterns but deviate substantially from the ground-truth: four repetitive predictions for the same interpolation task. (b) The effect becomes more severe with more complex waves and fewer data (80% missing); also see panel D and the supplementary material, Video 10. (c) The model generates significantly different output when the same prediction task is performed repeatedly (500 times, black: simple waves with 30%, gray: simple waves with 30%, orange: complex waves with 70%, pink: complex waves with 70% missing data, respectively). (d) Combinations of less vs more missing data (30% vs 70%) and simple and complex waves. (e) Average prediction error (RMSE vs perceptual MR) with increasing percentage of missing data. Hallucination is minimal with simpler wave patterns and 10%–30% missing data and increases steeply with complex wave patterns. The perceptual error indicates a higher qualitative agreement between inpainted and ground-truth patterns, which is not reflected by RMSE, as also shown in Fig. 7.

hallucination can be associated with qualitative changes in the topology of the wave pattern; and (iii) hallucination can be mitigated by sufficiently constraining the task the diffusion model is supposed to perform.

The inpainting diffusion model (Task 5) was trained as follows: we masked or left out data from a square region at the center of the 128×128 pixel^2 simulation domain and trained the network with corresponding pairs of masked $u_m(x, y)$ and ground-truth data $u(x, y)$ to reconstruct the missing parts of the spiral wave pattern,

$$(u_{m,1}(x, y), \ldots, u_{m,5}(x, y)) \rightarrow \tilde{u}_5(x, y), \qquad (11)$$

where $\tilde{u}$ is a prediction of the ground-truth dynamics u and the model reads a short spatio-temporal sequence of five two-dimensional snapshots. Masked pixels were replaced by zeros. We conditioned the diffusion model by concatenating $(u_{m,1}(x, y), \ldots, u_{m,5}(x, y))$ to the initial noisy distribution $\xi(x, y)$, adding five channels to the input of the underlying U-Net $(6 \times 128 \times 128 \text{ pixel}^3)$. We simulated two spiral wave regimes and trained two separate models: (i) one with largely only one spiral wave (Task 5a) and (ii) one with multiple, more chaotic spiral waves (Task 5b); see Table I for the corresponding simulation parameters. Both the models were trained equally with a range of masks with uniform distribution (of the percentage of masked area vs total area). We varied mask sizes $m \in [0.05, \ldots, 0.8]$ (percentage masked area vs total area from 5% to 80% in increments of 5%). Training and evaluation was performed with 27 500 and 6000 samples, respectively. The training samples were drawn randomly from 100 simulations performed with the biophysical model defined in Eqs. (1) and (2) for each regime, and the evaluation samples were drawn from 25 separate simulations for each regime.

F. Visual similarity of diffusion-generated vs PDE-generated spiral waves

Diffusion-generated "fake" spiral waves are visually indistinguishable from real spiral wave patterns simulated with a biophysical model, as shown in Figs. 2(b), 2(c), 9(a), and 9(b). Each panel in Figs. 2(b), 2(c), 9(a), and 9(b) shows randomly chosen, representative examples of spiral wave patterns simulated with a biophysical model or generated with diffusion, respectively. The biophysical model integrates partial differential equations (PDEs), whereas the diffusion model mimics these solutions. We found it impossible to distinguish diffusion-generated from PDE-simulated spiral wave patterns visually (we tested this systematically with different lab members). However, despite the visual similarity, a ResNet18[51] classifier fine-tuned on the two classes of images shown in Figs. 9(a) and 9(b) is able to distinguish the two groups of spiral wave patterns with an accuracy of 99.7% (separate training and validation/test datasets). This may be due to the invisible artifacts from the denoising process or the capability of CNNs to learn minute differences between classes. This is well explored[52,53] and was in part the motivation behind the joint generator-discriminator training process of GANs. An analysis of the fine-tuned ResNet18 classifier using Grad-CAM[54] provides some insights into the classification mechanism but is difficult to interpret overall and remains inconclusive; see Fig. 9(c) and supplementary material, Fig. 3.
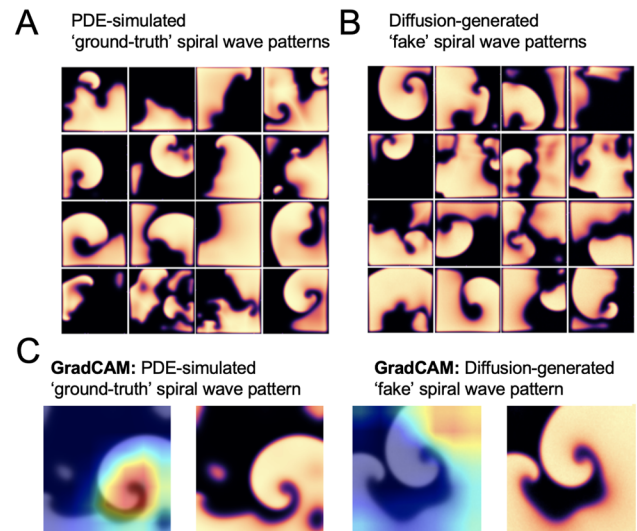
FIG. 9. Comparison of (a) spiral wave patterns simulated in computer simulations using a biophysical model (PDE, "ground-truth") and (b) "fake" spiral wave patterns generated with diffusion model from Eq. (12) (unconditional generation, Task 6). The two datasets are visually indistinguishable from each other. (c) ResNet18[51] classifier can distinguish real from 'fake' spiral waves with 99% accuracy. Grad-CAM[54] highlights the features to which the classification is sensitive.

It is only possible to visually distinguish real from "fake" spiral waves when the model was not trained for long enough. In that case, the generations often include noisy spiral wave images; see supplementary material, Fig. 1. The training dataset size does not seem to impact the image quality: both models that trained with small (100–1, 000 samples) and large training datasets (more than 10 000 samples) exhibit noisy images if they are not trained for long enough. Otherwise, training our diffusion models with sufficiently large training datasets is required to cover a wider range of the many possible, highly chaotic, and diverse spiral wave patterns. We found that diffusion models can generate plausible-looking spiral wave patterns with as few as 100 training samples; see supplementary material, Fig. 2. Interestingly, even the parameter-specific model can generate spiral wave patterns when the training samples are distributed over the 25 parameter pairs (4 samples/pair). Nevertheless, when calculating the "Fréchet inception distance" or the "FID

TABLE III. Training dataset size evaluated using "Fréchet inception distance" or the "FID score."[55] The lower the score, the better the similarity between the real and diffusion-generated images measured over the entire distribution of images. The training dataset is from Fig. 9.

Training samples	FID score
100	27.54
500	12.22
1 000	8.77
5 000	8.34
10 000	7.01
50 000	7.05

Score,"[55] which measures how visually similar the generated and real images are and how well the generated images capture the entire distribution of real images, we found that at least tens of thousands of training samples are necessary for training, as presented in Table III.

Furthermore, we found that the parameter-specific model with which we generated the patterns shown in Fig. 2(c) required more training samples and training iterations than the unconditional model with which we generated the patterns shown in Fig. 9(b). Further research is needed on the data-efficiency of diffusion models in these different applications. Data-efficiency is an important consideration when looking at the possibility of fine-tuning or training diffusion models from scratch on experimental data. It is possible that diffusion models require a large amount of data to perform well in complex applications, which could be a major limitation.

The diffusion model used to generate the wave patterns shown in Fig. 9(b) was trained to generate two-dimensional spiral wave patterns, including both dynamic variables from noise in an unconditional fashion (Task 6),

$$\xi(x,y) \rightarrow (\tilde{u}, \tilde{r})(x,y). \tag{12}$$

The model was trained on spiral wave patterns obtained with the Aliev–Panfilov model with a fixed parameter set; see Table I (Task 6), which we adapted from the work of Lilienkamp, Christoph, and Parlitz.[56] Consequently, the model is not conditioned with input parameters, as in Sec. III A, but can dream up any spiral wave pattern it can come up with given its training; see Fig. 9(b) and also supplementary material, Videos 1 and 2. Moreover, the model is completely unrestricted in that it is not trained to perform certain tasks, such as inpainting, nor constrained by certain boundary conditions or parameters that guide the generative process. The unconditional model was trained with 50 000 training samples of spiral wave patterns $(u,r)(x,y)$ simulated in an isotropic excitable medium with size 128×128 pixel^2, as shown in Fig. 9(a). All the simulations ran for a fixed simulation time, until the end of phase 1 shown in Fig. 10(b). They were initialized with a random pulse protocol[56] to cause wave break and induce spiral wave dynamics, and they were stopped shortly after the spiral wave dynamics had fully developed; next, also see Sec. III G. The training samples (T.S.) were sampled from the last 300 time steps at the end of phase 1 of each simulation from an ensemble of 5000 simulations, as shown in Fig. 10(b). The images shown in Figs. 9(a) and 9(b) are a random selection from the simulated training samples and diffusion-generated data, respectively (16 images from each class).

G. Self-termination behavior of diffusion-generated spiral waves

Spiral wave dynamics eventually self-terminate if one waits long enough. Figure 10(a) shows examples of short-, medium-, and long-lived spiral wave dynamics, which were simulated with the same parameters as in Sec. III F using the biophysical model from Eqs. (1) and (2). Two of the three examples survive for only about 100–300 simulation time steps before self-termination, while one survives for longer than 500 simulation time steps. When performing many simulations, most spiral wave dynamics self-terminate rather quickly, while only few episodes survive for long times; also see Figs. 10(b) and 10(c). The overall distribution of self-termination or survival

times of spiral wave dynamics was previously found to be exponential.[56] Here, we found a similar behavior with diffusion-generated spiral wave dynamics.

We performed the same simulations as in the previous Sec. III F, but let the simulations run until they eventually self-terminated; see phase 2 in Fig. 10(b). We also initiated simulations using the unconditional diffusion model (Task 6) from Eq. (12) from Sec. III F, since this model generates a full dynamical state $(u,r)(x,y)$ and also let those simulations run until they eventually self-terminated. Both types of simulations were evolved using the biophysical model (PDE-evolved). The histogram in Fig. 10(c)
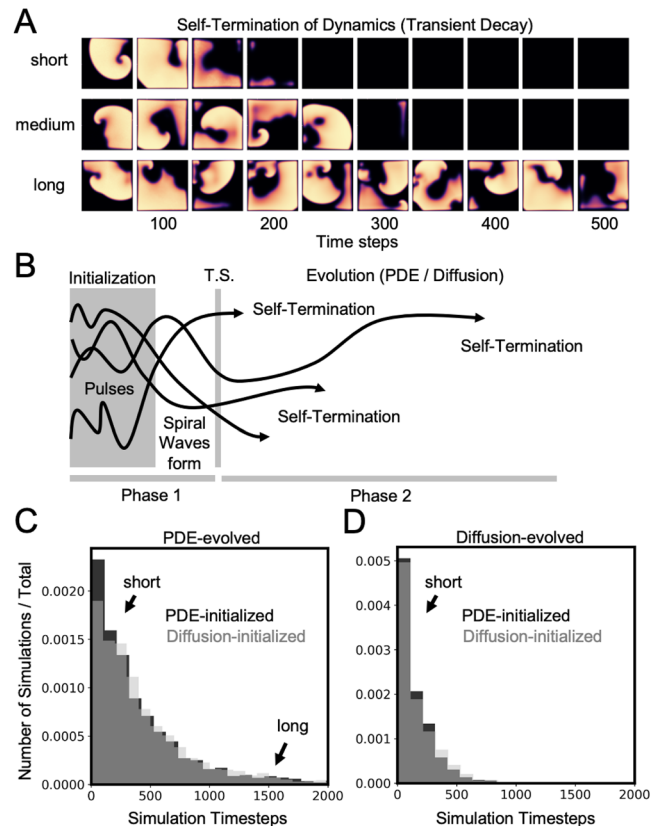
FIG. 10. Self-termination characteristics of real and "fake" spiral wave dynamics. (a) Short-, medium-, and long-lived spiral wave dynamics simulated with the biophysical model. (b) Trajectories of different spiral wave episodes. In phase 1, the spiral waves are induced with a random pulse protocol using the biophysical model. The training samples (T.S.) were sampled over a period of 300 frames right after the spiral waves had formed; also see Fig. 9(b). In phase 2, the spiral wave dynamics were further evolved using either the biophysical model or diffusion model from Eq. (8). At the beginning of phase 2, the simulations were either initialized or continued with the real dynamical states (u,r) from phase 1 or the "fake" dynamical states $(\tilde{u}, \tilde{r})$ generated with the unconditional diffusion model from Eq. (12). Self-termination times were measured for all 4 cases with respect to the beginning of phase 2. Panels (c) and (d): distributions of self-termination (or survival) times measured over an ensemble of 5000 spiral wave dynamics showing an exponential decay. (c) The dynamics were evolved with the biophysical model (PDE) and initialized with either the biophysical or diffusion model. (d) The dynamics were evolved with the diffusion model and initialized with either the biophysical or diffusion model.

shows the distributions of self-termination times (or survival times) for ensembles of 5000 simulations of diffusion-initialized spiral wave dynamics (gray) vs conventional spiral wave dynamics (black). The self-termination times were calculated with respect to the beginning of phase 2. Both distributions match, demonstrating that the diffusion-generated spiral wave patterns adhere to the same self-termination statistics with a similar exponential decay rate as their biophysical counterparts. In other words, it appears that the unconditional diffusion-generated spiral wave patterns shown in Fig. 9(b) do not just look like the real spiral wave patterns shown in Fig. 9(a), but correspond to real physical solutions.

In another experiment, we compared the self-termination statistics when both spiral wave dynamics in phase 2 are evolved with the time-stepping diffusion model from Eq. (8) in Sec. III C. The histogram in Fig. 10(d) shows that both distributions match (calculated for ensembles of 5000 simulations each). However, both diffusion-evolved dynamics self-terminate much sooner than their PDE-evolved counterparts, highlighting that the time-stepping diffusion model from Sec. III C behaves differently than the finite differences time-stepping scheme of the biophysical model. This correlates with the observation in Sec. III C that diffusion-evolved spiral wave dynamics appear to exhibit less wave break than the corresponding PDE-evolved dynamics. Less wave break could contribute to a shortening of the survival times of the dynamics. In addition, while measuring the self-termination times with the diffusion-evolved dynamics, we encountered one curious phenomenon: shortly before self-termination, the dynamics would be abruptly taken over by severe noise. This issue could be solved by adding more training data of self-termination events. Taken together, these findings suggest that diffusion-generated spiral wave dynamics adhere to the same laws as their biophysical counterparts, but there are some reservations that require further exploration.

IV. DISCUSSION

Generative AI provides the potential for many promising applications in the biological and biomedical sciences. Here, we demonstrated that denoising diffusion probabilistic models (DDPMs) can be used to model waves of electrical excitation in cardiac tissue. Diffusion models can be used to reconstruct or create parameter-specific wave patterns and, most importantly, simulate electrical wave propagation in a data-driven manner. In other words, diffusion models can learn to evolve cardiac wave dynamics from previously seen data without knowledge about the underlying physics. Therefore, they could potentially be used to create a data-driven model of the heart's electrophysiological system from measurement data. We found that diffusion models not only generate electrical wave dynamics that look like and are visually indistinguishable from simulated wave dynamics, but the diffusion-generated dynamics also preserve some of the inherent characteristics of the original dynamics. For instance, we found that diffusion-generated spiral wave dynamics exhibit similar self-termination statistics as their counterparts in excitable media; see Sec. III G.

While we have some confidence that the diffusion-generated waves are indeed legitimate solutions, we also remain cautious and further research is needed to confirm whether diffusion models provide a valid alternative to conventional biophysical modeling. At this point, we cannot rule out that diffusion models merely

emulate rather than simulate spiral wave dynamics. This concern is particularly critical when the dynamics are chaotic and sensitive to slight physical perturbations or differences in the numerical integration. We evolved both simpler and more complicated spiral wave dynamics, and while the diffusion model generated plausible-looking simpler wave dynamics for very long times (in contrast to U-Net), which co-evolved over a reasonable period of time with the ground-truth dynamics (keep in mind that even different solvers would lead to diverging results), the more complicated spiral wave dynamics diverged very quickly from their ground-truth counterpart and exhibited less wavebreak. The latter observation could be an indication that bias in the training data influences the behavior of the dynamics in yet inexplicable ways. In addition, we observed unfamiliar artifacts, such as a sudden onset of noise shortly before the self-termination of the spiral wave dynamics. This particular artifact could be mitigated by including more training data of self-termination events, which hints at fundamental issues with selective and insufficient training data.

A major concern with diffusion models is their ability to hallucinate. Hallucination is an inherent property of generative modeling and a feature and bug at the same time. Diffusion models can generate a continuum of outputs, of which some are made up and false. The main issue is that the false output is hard to identify as diffusion models excel at learning the data distribution and generating realistically looking data points from this distribution. This raises concerns over the applicability of diffusion models in healthcare, where they could produce misleading output, which could lead to an incorrect diagnosis or treatment. Here, we found that the extent to which diffusion models hallucinate is related to how much the task that the network is supposed to perform is constrained. If the problem is more constrained, then the space of possible solutions becomes smaller and there is less potential for hallucination (e.g., when evolving dynamics). Therefore, sufficiently constraining diffusion models is essential as well as developing methods to quantify and mitigate hallucination. Nevertheless, the perceived weaknesses with regard to hallucination can also be a major advantage in other situations: diffusion models excel when tasked to generate a starting point in underconstrained tasks, and therefore, they could serve as a powerful prior for difficult cardiac modeling or reconstruction tasks.

Overall, despite the potential drawbacks, diffusion models are a promising tool with many potential applications in cardiac research and diagnostics. Diffusion models can, in principle, generate parameter- or even model-specific (scroll) wave dynamics in three-dimensional heart-shaped geometries, suggesting that they could be used to simulate atrial or ventricular fibrillation in an individualized segmentation of a patient's heart while also integrating patient- or disease-specific information (e.g., ion channel abnormalities) and/or measurement data (e.g., catheter mapping or electrocardiographic data). In particular, diffusion models offer the possibility to learn different integration time scales, perform simulations at arbitrary resolutions, and skip the tedious part of finding the right initial conditions for simulating such dynamics as they can readily generate spiral and scroll wave patterns instantaneously.

V. CONCLUSIONS

We demonstrated that denoising diffusion probabilistic models (DDPMs) can be used for generating electrophysiological wave

patterns in cardiac tissue. They can be used for recovering missing data, evolving spatio-temporal dynamics, generating electrophysiological wave patterns in arbitrary geometries, or generating parameter-specific dynamics, among other tasks. The diffusion-generated waves are visually indistinguishable from and behave very similar to waves simulated with biophysical models. However, diffusion models tend to hallucinate with insufficient constraining, produce artifacts in situations in which training data are lacking, and produce high computational upfront costs for training. In the future, diffusion models could be used for data-driven modeling of various physiological phenomena in the heart.

SUPPLEMENTARY MATERIAL

The supplementary material provides additional figures (Supplementary Figs. 1–4) and descriptions of the supplementary videos, which can also be found at https://cardiacvision.ucsf.edu/videos/diffusion/ and https://www.youtube.com/@cardiacvision.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

T.B. developed the deep learning methodology and performed the data analysis. T.B. and J.C. performed the two-dimensional simulations. J.C. aided with the data analysis. J.L. performed the SPH and bulk simulations. T.B. and J.C. designed the figures. J.C. wrote the manuscript. T.B. aided in writing the manuscript. All authors discussed and interpreted the results and read and approved the final version of the manuscript. J.L. and J.C. conceived the research.

Tanish Baranwal: Data curation (equal); Formal analysis (lead); Investigation (equal); Methodology (lead); Software (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Jan Lebert**: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Visualization (equal). **Jan Christoph**: Conceptualization (equal); Formal analysis (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are openly available at https://github.com/cardiacvision/diffusion.

REFERENCES

[1] V. I. Krinsky and H. L. Swinney, *Waves and Patterns in Chemical and Biological Media* (MIT Press, 1991).

[2] A. Winfree, "Electrical turbulence in three-dimensional heart muscle," Science **266**, 1003–1006 (1994).

[3] S. Alonso, M. Bär, and B. Echebarria, "Nonlinear physics of electrical wave propagation in the heart: A review," Rep. Prog. Phys. **79**, 096601 (2016).

[4] W.-J. Rappel, "The physics of heart rhythm disorders," Phys. Rep. **978**, 1–45 (2022).

[5] A. M. Pertsov, J. M. Davidenko, R. Salomonsz, W. T. Baxter, and J. Jalife, "Spiral waves of excitation underlie reentrant activity in isolated cardiac muscle," Circ. Res. **72**, 631–650 (1993).

[6] R. A. Gray, J. Jalife, A. Panfilov, W. T. Baxter, C. Cabo, J. M. Davidenko, and A. M. Pertsov, "Nonstationary vortexlike reentrant activity as a mechanism of polymorphic ventricular tachycardia in the isolated rabbit heart," Circulation **91**, 2454–2469 (1995).

[7] P. Pathmanathan and R. A. Gray, "Filament dynamics during simulated ventricular fibrillation in a high-resolution rabbit heart," BioMed Research BioMed Res. Int. **2015**, 720575.

[8] J. Christoph, M. Chebbok, C. Richter, J. Schröder-Schetelig, P. Bittihn, S. Stein, I. Uzelac, F. H. Fenton, G. Hasenfuß, R. F. Gilmour, and S. Luther, "Electromechanical vortex filaments during cardiac fibrillation," Nature **555**, 667–672 (2018).

[9] I. Uzelac, S. Iravanian, N. K. Bhatia, and F. H. Fenton, "Spiral wave breakup: Optical mapping in an explanted human heart shows the transition from ventricular tachycardia to ventricular fibrillation and self-termination," Heart Rhythm **19**, 1914 (2022).

[10] F. Fenton and A. Karma, "Vortex dynamics in three-dimensional continuous myocardium with fiber rotation: Filament instability and fibrillation," Chaos **8**, 20–47 (1998).

[11] O. Berenfeld and A. M. Pertsov, "Dynamics of intramural scroll waves in three-dimensional continuous myocardium with rotational anisotropy," J. Theor. Biol. **199**, 383–394 (1999).

[12] Z. Qu, J. Kil, F. Xie, A. Garfinkel, and J. N. Weiss, "Scroll wave dynamics in a three-dimensional cardiac tissue model: Roles of restitution, thickness, and fiber rotation," Biophys. J. **78**, 2761–2775 (2000).

[13] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," Biophys. J. **1**, 445–466 (1961).

[14] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," Proc. IRE **50**, 2061–2070 (1962).

[15] R. R. Aliev and A. V. Panfilov, "A simple two-variable model of cardiac excitation," Chaos, Solitons Fractals **7**, 293–301 (1996).

[16] J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," arXiv:1503.03585 (2015).

[17] Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," arXiv:1907.05600 (2019).

[18] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," arXiv:2006.11239 [cs.LG] (2020).

[19] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," arXiv:2112.10752 (2021).

[20] C. Saharia, W. Chan, H. Chang, C. A. Lee, J. Ho, T. Salimans, D. J. Fleet, and M. Norouzi, "Palette: Image-to-image diffusion models," arXiv:2111.05826 (2021).

[21] C. Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E. Denton, S. K. S. Ghasemipour, B. K. Ayan, S. S. Mahdavi, R. G. Lopes, T. Salimans, J. Ho, D. J. Fleet, and M. Norouzi, "Photorealistic text-to-image diffusion models with deep language understanding," arXiv:2205.11487 [cs.CV] (2022).

[22] U. Singer, A. Polyak, T. Hayes, X. Yin, J. An, S. Zhang, Q. Hu, H. Yang, O. Ashual, O. Gafni, D. Parikh, S. Gupta, and Y. Taigman, "Make-a-video: Text-to-video generation without text-video data," arXiv:2209.14792 [cs.CV] (2022).

[23] Z. Kong, W. Ping, J. Huang, K. Zhao, and B. Catanzaro, "DiffWave: A versatile diffusion model for audio synthesis," arXiv:2009.09761 [eess.AS] (2021).

[24]W. H. L. Pinaya, P.-D. Tudosiu, J. Dafflon, P. F. da Costa, V. Fernandez, P. Nachev, S. Ourselin, and M. J. Cardoso, "Brain imaging generation with latent diffusion models," arXiv:2209.07162 [eess.IV] (2022).

[25]J. L. Watson, D. Juergens, N. R. Bennett et al., "De novo design of protein structure and function with RF diffusion," Nature 620, 1089–1100 (2023).

[26]Z. Guo, J. Liu, Y. Wang, M. Chen, D. Wang, D. Xu, and J. Cheng, "Diffusion models in bioinformatics and computational biology," Nat. Rev. Bioeng. 2, 136 (2023).

[27]A. Kazerouni, E. K. Aghdam, M. Heidari, R. Azad, M. Fayyaz, I. Hacihaliloglu, and D. Merhof, "Diffusion models in medical imaging: A comprehensive survey," Med. Image Anal. 88, 102846 (2023).

[28]P.-D. Tursio et al., "Realistic morphology-preserving generative modelling of the brain," Nature Mach. Intell. 6, 811–819 (2024).

[29]L. Zhou, Y. Du, and J. Wu, "3D shape generation and completion through point-voxel diffusion," arXiv:2104.03670 (2021).

[30]J. Lebert, M. Mittal, and J. Christoph, "Reconstruction of three-dimensional scroll waves in excitable media from two-dimensional observations using deep neural networks," Phys. Rev. E 107, 014221 (2023).

[31]J. Lebert, D. Deng, L. Fan, L. C. Lee, and J. Christoph, "Deep learning-based prediction of electrical arrhythmia circuits from cardiac motion: An in-silico study," arXiv:2305.07822v1 [physics.med-ph] (2023).

[32]C. Zhang, M. Rezavand, Y. Zhu, Y. Yu, D. Wu, W. Zhang, J. Wang, and X. Hu, "SPHinXsys: An open-source multi-physics and multi-resolution library based on smoothed particle hydrodynamics," Comput. Phys. Commun. 267, 108066 (2021).

[33]C. Zhang, J. Wang, M. Rezavand, D. Wu, and X. Hu, "An integrative smoothed particle hydrodynamics method for modeling cardiac function," Comput. Methods Appl. Mech. Eng. 381, 113847 (2021).

[34]P. Dhariwal and A. Nichol, "Diffusion models beat GANs on image synthesis," arXiv:2105.05233 (2021).

[35]L. Jiang and Y. Belousov, "Unofficial implementation of Palette: Image-to-image diffusion models," https://github.com/Janspiry/Palette-Image-to-Image-Diffusion-Models, 2022.

[36]P. von Platen, S. Patil, A. Lozhkov, P. Cuenca, N. Lambert, K. Rasul, M. Davaadorj, and T. Wolf, Diffusers: State-of-the-art diffusion models.

[37]O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Lecture Notes in Computer Science (Springer International Publishing, 2015), pp. 234–241.

[38]A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An imperative style, high-performance deep learning library," in Advances in Neural Information Processing Systems, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019), Vol. 32, pp. 8024–8035.

[39]D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Conference Track Proceedings of 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, May 7–9, 2015, edited by Y. Bengio and Y. LeCun, arXiv:1412.6980 (2015).

[40]R. Kazmierczak, G. Franchi, N. Belkhir, A. Manzanera, and D. Filliat, "A study of deep perceptual metrics for image quality assessment," arXiv:2202.08692 [cs.CV] (2022).

[41]Z. Qu, F. Xie, A. Garfinkel, and J. N. Weiss, "Origins of spiral wave meander and breakup in a two-dimensional cardiac tissue model," Ann. Biomed. Eng. 28, 755–771 (2000).

[42]E. Bartocci, R. Singh, F. B. von Stein, A. Amedome, A. J. J. Caceres, J. Castillo, E. Closser, G. Deards, A. Goltsev, R. S. Ines, C. Isbilir, J. K. Marc, D. Moore, D. Pardi, S. Sadhu, S. Sanchez, P. Sharma, A. Singh, J. Rogers, A. Wolinetz, T. Grosso-Applewhite, K. Zhao, A. B. Filipski, R. F. Gilmour, R. Grosu, J. Glimm, S. A. Smolka, E. M. Cherry, E. M. Clarke, N. Griffeth, and F. H. Fenton, "Teaching cardiac electrophysiology modeling to undergraduate students: Laboratory exercises and GPU programming for the study of arrhythmias and spiral wave dynamics," Adv. Physiol. Educ. 35, 427–437 (2011).

[43]A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, edited by I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, and R. Garnett, arXiv:1706.03762 (2017), pp. 5998–6008.

[44]R. Stenger, S. Herzog, I. Kottlarz, B. Rüchardt, S. Luther, F. Wörgötter, and U. Parlitz, "Reconstructing in-depth activity for chaotic 3D spatiotemporal excitable media models based on surface data," Chaos 33, 013134 (2023).

[45]M. J. Hoffman, N. S. LaVigne, S. T. Scorse, F. H. Fenton, and E. M. Cherry, "Reconstructing three-dimensional reentrant cardiac electrical wave dynamics using data assimilation," Chaos 26, 013107 (2016).

[46]M. J. Hoffman and E. M. Cherry, "Sensitivity of a data-assimilation system for reconstructing three-dimensional cardiac electrical dynamics," Philos. Trans. R. Soc., A 378, 20190388 (2020).

[47]J. Christoph and J. Lebert, "Inverse mechano-electrical reconstruction of cardiac excitation wave patterns from mechanical deformation using deep learning," Chaos 30, 123134 (2020).

[48]J. Lebert, N. Ravi, F. H. Fenton, and J. Christoph, "Rotor localization and phase mapping of cardiac excitation waves using deep neural networks," Front. Physiol. 12, 782176 (2021).

[49]R. Cohen, I. Kligvasser, E. Rivlin, and D. Freedman, "Looks too good to be true: An information-theoretic analysis of hallucinations in generative restoration models," arXiv:2405.16475 [cs.LG] (2024).

[50]A. T. Kalai and S. S. Vempala, "Calibrated language models must hallucinate," arXiv:2311.14648 [cs.CL] (2024).

[51]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv:1512.03385 (2015).

[52]S.-Y. Wang, O. Wang, R. Zhang, A. Owens, and A. A. Efros, "CNN-generated images are surprisingly easy to spot. . . for now," arXiv:1912.11035 [cs.CV] (2020).

[53]S. Hou, Y. Ju, C. Sun, S. Jia, L. Ke, R. Zhou, A. Nikolich, and S. Lyu, "DeepFake-O-Meter v2.0: An open platform for DeepFake detection," arXiv:2404.13146 [cs.CR] (2024).

[54]R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," Int. J. Comput. Vision 128, 336–359 (2019).

[55]M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," arXiv:1706.08500 [cs.LG] (2018).

[56]T. Lilienkamp, J. Christoph, and U. Parlitz, "Features of chaotic transients in excitable media governed by spiral and scroll waves," Phys. Rev. Lett. 119, 054101 (2017).