

Adversarial Hamiltonian learning of quantum dots in a minimal Kitaev chain

Koch, Rouven; Van Driel, David; Bordin, Alberto; Lado, Jose L.; Greplova, Eliska

DOI

[10.1103/PhysRevApplied.20.044081](https://doi.org/10.1103/PhysRevApplied.20.044081)

Publication date

2023

Document Version

Final published version

Published in

Physical Review Applied

Citation (APA)

Koch, R., Van Driel, D., Bordin, A., Lado, J. L., & Greplova, E. (2023). Adversarial Hamiltonian learning of quantum dots in a minimal Kitaev chain. *Physical Review Applied*, 20(4), Article 044081. <https://doi.org/10.1103/PhysRevApplied.20.044081>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Adversarial Hamiltonian learning of quantum dots in a minimal Kitaev chain

Rouven Koch^{1,*}, David van Driel^{2,3}, Alberto Bordin^{2,3}, Jose L. Lado¹, and Eliska Greplova³

¹*Department of Applied Physics, Aalto University, Espoo 02150, Finland*

²*QuTech, Delft University of Technology, Delft 2600 GA, Netherlands*

³*Kavli Institute of Nanoscience, Delft University of Technology, Delft 2600 GA, Netherlands*



(Received 26 April 2023; revised 31 May 2023; accepted 6 October 2023; published 31 October 2023)

Determining Hamiltonian parameters from noisy experimental measurements is a key task for the control of experimental quantum systems. An interesting experimental platform where precise knowledge of device parameters is useful is the quantum-dot-based Kitaev chain. In these systems, the fine tuning of Hamiltonian parameters is crucial in order to reach the desired regime with stable midgap modes. In this work, we demonstrate an adversarial machine-learning algorithm to determine the parameters of a quantum-dot-based Kitaev chain. We train a convolutional conditional generative adversarial neural network (CCGAN) with simulated differential conductance data and use the model to predict the parameters at which Majorana bound states are predicted to appear. In particular, the CCGAN model facilitates a rapid, numerically efficient exploration of the phase diagram describing the transition between elastic co-tunneling and crossed Andreev reflection regimes. We verify the theoretical predictions of the model by applying it to experimentally measured conductance obtained from a minimal Kitaev chain consisting of two spin-polarized quantum dots coupled by a superconductor-semiconductor hybrid. Our model accurately predicts, with an average success probability of 97%, whether the measurement was taken in the elastic co-tunneling or crossed Andreev reflection-dominated regime. Our work constitutes a stepping stone towards fast, reliable parameter prediction for tuning quantum dot systems into distinct Hamiltonian regimes. Ultimately, our results yield a strategy to support Kitaev-chain tuning that is scalable to longer chains.

DOI: [10.1103/PhysRevApplied.20.044081](https://doi.org/10.1103/PhysRevApplied.20.044081)

I. INTRODUCTION

The engineering of topological quantum matter represents a critical step for bringing theoretical predictions of exotic physics to the lab. One of the paradigmatic topological modes are Majorana bound states, associated with topological superconducting orders [1–3], and have been studied extensively due to their potential for quantum computing [4,5]. The practical realization of Majorana bound states has proven to be greatly challenging, which motivated the development of multiple experimental platforms for their realization, including semiconductors [6], atomic chains [7], van der Waals heterostructures [8], and heavy-fermion materials [9]. Majorana bound states are predicted to appear at the edges of a spinless p -wave superconductor in one dimension (1D), the so-called Kitaev chain [1–3,10]. A minimal version of such a Kitaev chain has been recently experimentally realized in the system of two quantum dots that are mutually coupled both via crossed Andreev reflection and elastic co-tunneling [11,12]. The relative strengths of these two couplings can be tuned to

drive the system into the so-called sweet spot, where the electron tunneling and Andreev reflection have the same strength [13–15]. An arrival into the sweet-spot regime gives rise to the midgap modes consistent with a “poor man’s Majorana bound state” [11]. Fine tuning a polarized quantum dot chain simultaneously presents a significant experimental challenge and is one of the limiting factors for scaling this type of quantum experiment [12].

Inferring the description of a quantum system from available data, a problem known as Hamiltonian learning, is a central problem in quantum systems [16–21]. Neural network-based algorithms have recently risen to prominence as a tool for Hamiltonian learning and parameter estimation from experimental data [22–30]. While machine-learning algorithms are by far not a universal answer to parameter determination challenges in noisy quantum systems, their generalization properties and fast evaluation make them suitable candidates to address technical questions connected to quantum control and parameter estimation. Neural-network algorithms have been successfully used in supervised settings for both Hamiltonian estimation [20,31,32] as well as tuning challenges [33–36]. More recently, generative models have been shown

*rouven.koch@aalto.fi

to be powerful simulation and parameter estimation tools in the context of generic many-body physics problems [27,37].

Here, we demonstrate that convolutional conditional generative adversarial networks (CCGANs) allow parameter estimation on both simulated and experimental data. The key advantage of the conditional generative model is the ability to extract information from (experimental) data and, what is more, evaluate the ML output as a function of Hamiltonian parameters. We show that conditional generative models constitute an alternative figure of merit of parameter estimation on experimental data. We demonstrate our adversarial methodology in determining the poor man's Majorana sweet spot from differential conductance measurement. We find that our model is able to produce full conductance spectra that are effectively indistinguishable from real measurement data and to estimate the ratio of crossed Andreev reflection and elastic co-tunneling rates with high precision. Incidentally, our algorithm correctly identified measurement-data instances that were later discovered to be mislabeled due to the large error bar in an experimental labeling procedure. Overall, the machine-learning model reached 100% (94%) accuracy on the first (second) measured data set in distinguishing tunneling- and Andreev reflection-dominated regimes. The two measured sets yielded average accuracy of 97% in regime recognition. The ability to quickly recognize the position with respect to the sweet spot is key when driving the Kitaev chain to the fine-tuned point. Our algorithm thus presents a key stepping stone towards Majorana bound state tuning.

The work is organized as follows. In Sec. II we explain key methods used in this work: the physics of a Kitaev chain, the architecture and the training of conditional generative adversarial models, and details about experimental devices and the measurement procedure. In Sec. III we summarize and explain the results we obtained on both simulated and experimental data and quantify the prediction power of our model. In Sec. IV we summarize our results and outline the path of embedding pretrained neural-network models into the Majorana bound state tuning workflow.

II. METHODS

In this section, we review the fundamental methodologies used in our paper. We first elaborate on the theoretical description of the Kitaev model and introduce the key parameters required for sweet-spot tuning. Afterwards, we provide a description of the CCGAN architecture and the training process. We also summarize the key components of our experimental platform and provide details on the measurement process.

A. Kitaev chain

The model we consider to generate the theoretical training data is a minimal one-dimensional Kitaev model of two sites, modeling a two-quantum-dot system with the Hamiltonian

$$H = td_L^\dagger d_R + \Delta d_L^\dagger d_R^\dagger + \epsilon_L d_L^\dagger d_L + \epsilon_R d_R^\dagger d_R + \text{h.c.}, \quad (1)$$

where t is the amplitude of elastic co-tunneling (ECT), Δ is the amplitude of crossed Andreev reflection (CAR), and ϵ_L (ϵ_R) is the on-site energy of the left (right) quantum dot. Here, ECT is the tunneling of a single electron from the left quantum dot to the right one. CAR is the reflection of an electron from the left quantum dot into a hole to the right quantum dot. Both are virtual tunneling processes mediated by the Andreev bound state (ABS) residing in the hybrid semiconducting-superconducting section [13,14,38].

The Hamiltonian can be written as

$$H = \frac{1}{2} \Psi^\dagger h \Psi + \frac{1}{2} (\epsilon_L + \epsilon_R), \quad (2)$$

taking the Nambu spinor with the basis function $\Psi^T = (d_L, d_R, d_L^\dagger, d_R^\dagger)$ and the matrix h has the form

$$h = \begin{pmatrix} \epsilon_L & t & 0 & \Delta \\ t & \epsilon_R & -\Delta & 0 \\ 0 & -\Delta & -\epsilon_L & -t \\ \Delta & 0 & -t & -\epsilon_R \end{pmatrix}. \quad (3)$$

To emulate the experimentally accessible regime, we compute the differential conductance that accounts for the transport measurements through the double-dot system. Specifically, we consider each quantum dot to be tunnel coupled to leads of normal metal with the rates Γ_L and Γ_R for the left and right dot, respectively. We will use the \mathcal{S} -matrix formalism, which describes the transmission and reflection coefficients of electrons and holes from the leads to each quantum dot and we can in turn relate these coefficients to the differential conductance. The \mathcal{S} matrix can generally be written as

$$\mathcal{S}(\omega) = \begin{pmatrix} s_{ee} & s_{eh} \\ s_{he} & s_{hh} \end{pmatrix} = \mathbb{K} - iW^\dagger \left(\omega - H + \frac{1}{2} iWW^\dagger \right)^{-1} W, \quad (4)$$

where $W = \text{diag}\{\sqrt{\Gamma_L}, \sqrt{\Gamma_R}, -\sqrt{\Gamma_L}, -\sqrt{\Gamma_R}\}$ is the tunnel matrix with the tunnel-coupling strengths Γ_L and Γ_R between the quantum dots and the normal leads.

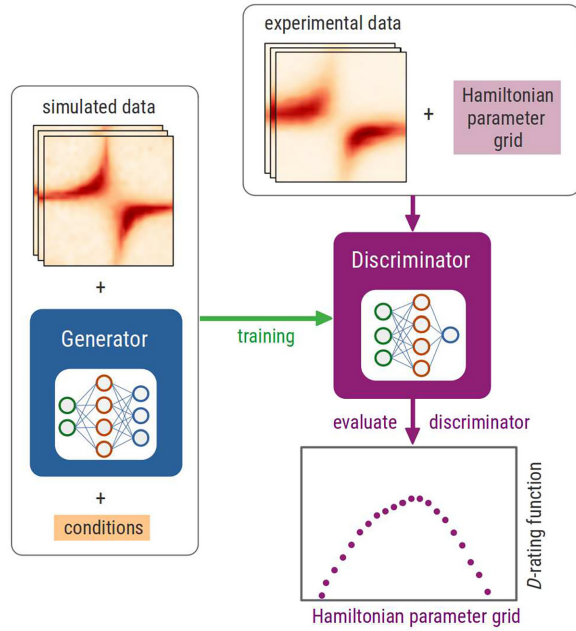


FIG. 1. Workflow of the Hamiltonian learning process with the CCGAN. The discriminator is trained with simulated data, the generator, and conditional parameters. The trained discriminator evaluates the experimental input on a discrete grid of values to obtain the discriminator-rating function.

With the \mathcal{S} matrix at hand, we compute the differential conductance, $G_{\alpha\beta}^0$ in the zero-temperature limit as

$$G_{\alpha\beta}^0 = dI_{\alpha}/dV_{\beta} = \frac{e^2}{h} (\delta_{\alpha\beta} - |s_{ee}^{\alpha\beta}(\omega)|^2 + |s_{he}^{\alpha\beta}|^2), \quad (5)$$

where $\alpha(\beta)$ defines the left or right quantum dot and $s_{ij}^{\alpha\beta}(\omega)$ refers to the matrix elements of the matrices s_{ij} as defined in Eq. (4). Finally, to obtain the finite-temperature differential conductance we convolve the zero-temperature conductance with a Fermi-Dirac distribution. This leads to the expression

$$G^T(\omega) = \int dE \frac{G^0(E)}{4k_B T \cosh^2[(E - \omega)/2k_B T]}. \quad (6)$$

In general, we can divide the parameter space and corresponding differential conductance into three regimes. In the two regimes where either ECT and CAR are the dominating effects, we obtain avoided crossings as depicted in Fig. 1(b). In the sweet spot, where the magnitude of ECT and CAR is equivalent, the avoided crossings vanish. A more in-depth description of the processes behind the avoided crossings can be found in Ref. [12]. Example experimental measurements of each regime can be found in Figs. 5(c), 5(f), and 5(j).

We use the formalism described above to create the data for training the machine-learning (ML) algorithm,

described in the next section. For the data generation, we vary the parameters of Eq. (3) as well as the tunnel rates Γ_L and Γ_R . In particular, we set the ECT magnitude $t = 1$ and include Δ values between 0.3 and 1.7 to include a wide variety of parameters for the ECT-dominated regime ($t > \Delta$), the CAR regime ($t < \Delta$), as well as the sweet spot ($t \approx \Delta$) [39]. The onsite energies $\epsilon_{L/R}$ are chosen randomly in the interval $L_{\epsilon} = [-1.5, 1.5]$ around the center position ϵ_0 , which results in a shift of the center position of the avoided crossings. Additionally, we are adding background noise and a bias with the magnitude of 10^{-3} . The coupling strengths to the leads $\Gamma_{L/R}$ are adjusted to match the experimental measurements.

B. Conditional generative adversarial networks

The key ingredient of our machine-learning approach to Hamiltonian learning from experimental data is a generative model conditioned on the experimentally tunable parameters. By training the generative model to reproduce the experimental measurements, we can study the dependence on successful and faithful data generation on these conditional parameters and thus determine the best match. Using a generative model in place of a feed-forward prediction has the additional advantage of visualizing the complete underlying parameter distribution and thus resulting in a more educated parameter estimate.

The ML algorithm we are using in this work for the Hamiltonian learning is a conditional generative adversarial network (CGAN) [40,41]. The general idea of GANs is to combine two separate deep neural networks (NNs) competing against each other in a min-max game during the training process. In the case of the CGAN, the generator network G learns to map from a random input vector \mathbf{z} , known as latent space, in combination with a conditional parameter \mathbf{y} to the data distribution p_{data} of input images of the training set. The discriminator network D has the task to distinguish between samples of the training set and generated “fake” samples from the generator, given as constraints of the conditional parameters.

The CGAN value function $V(D, G)$ is defined as

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x}|\mathbf{y})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z}|\mathbf{y})))], \quad (7)$$

where \mathbf{x} is the input data, $p_z(\mathbf{z})$ is the noise distribution of the latent space, and $\log[1 - D(G(\mathbf{z}|\mathbf{y}))]$ and $\log[D(\mathbf{x}|\mathbf{y})]$ are the corresponding terms for the generator and discriminator, respectively. In comparison of the original GAN value function [40], the additional conditional labels appear only in the terms corresponding to G and D . During the training, the parameters of both networks get updated simultaneously by minimizing the generator term and maximizing the discriminator term. If trained successfully, the

generator is able to randomly create other samples indistinguishable from the data distribution p_{data} . In contrast, the discriminator becomes very accurate in determining if a given input belongs to p_{data} or is a “fake” sample.

Adding conditional Hamiltonian parameters to the GAN architecture allows us to control the physical parameters of the conductance maps created by the generator, which in the case of the original GAN would be purely random. In the following, we focus on the ratio between Δ and t of the Hamiltonian of Eq. (7) as the continuous conditional label, which allows us to explore the complete phase space of the two-site Kitaev chain. Due to the fact that we are dealing with input data with high spatial correlations among pixels, we choose an architecture for G and D consisting of convolutional neural networks (CNNs) (see Appendix A for more details).

After the training process, both the generator and discriminator networks are trained equally well. For the Hamiltonian learning process, we exploit the trained discriminator to estimate parameters of a given measurement. It is worth noting that, the generation of conductance maps could be carried out with other generative algorithms such as diffusion models or autoencoders [42], but the existence of a discriminator in CGANs allows to part of the trained architecture to be directly used for Hamiltonian learning. In general, the discriminator provides the rating in the range between 0 and 1 of a given input array, with the corresponding label, to be from the real data set or the “fake” one. To use the discriminator for the Hamiltonian learning process and to extract the most-likely parameter, we developed the following workflow, which can be divided into three basic steps, which are also shown in Fig. 1. First, we define a landscape for the parameter range we want to estimate (here $\Delta/t \in I = [0.3, 1.7]$). Second, we feed the input spectrum, simulated or measured, into the discriminator in combination with every possible value of the parameter we want to extract. Finally, the discriminator returns a score characterizing the likeliness that the input data provided corresponds to each possible parameter set. Combining all the outputs, we obtain a *discriminator-rating (D-rating) function*, estimating which Hamiltonian parameter is most likely belonging to the input image. In comparison to feed-forward NNs, this procedure gives us an uncertainty estimation of the maximum predicted value, and can naturally account for situations in which different parameter sets are consistent with the data.

C. Experiments: Two-quantum-dot system

The experimental data was obtained using the device sketched in Fig. 2. An InSb nanowire (NW) is placed on an array of bottom gates, which can change its electrochemical potential locally. Afterwards, the middle section of the NW is contacted by a thin (approximately 8-nm) aluminum shell, which is kept grounded throughout the experiment.

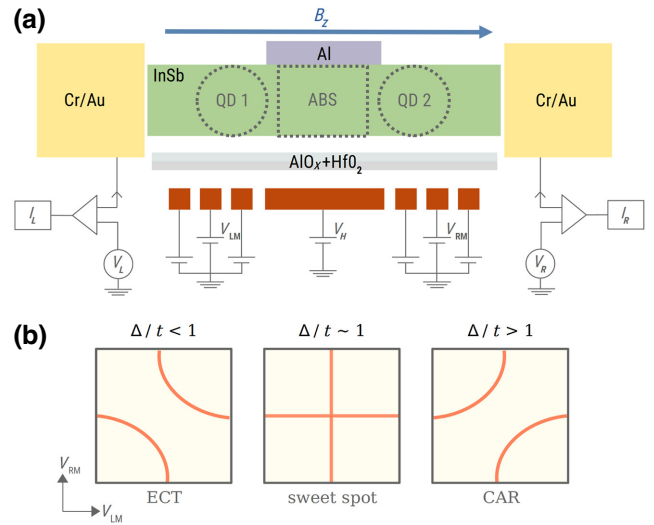


FIG. 2. (a) Sketch of the experimental InSb nanowire setup containing two spin-polarized quantum dots (QDs) and a hybrid hosting Andreev bound states (ABSs). Two Cr/Au normal leads allow for measuring the local and nonlocal differential conductance. We give more details about the device and measurement process in Sec. II C. (b) Sketch of the avoided crossings in the (local) differential conductance for the three different regimes of the Hamiltonian. At the sweet spot (middle panel), the avoided crossings vanish.

The leftmost and rightmost three bottom gates are used to define quantum dots (QDs), whose electrochemical potentials are controlled by V_{LM}/V_{RM} , respectively. The NW is contacted by two normal, Cr/Au leads at its ends. These can be biased with a voltage $V_{L/R}$ with respect to the grounded Al. For more details on device fabrication, see Refs. [43,44]. The middle gate V_H is used to tune the electrochemical potential of the hybrid InSb-Al section [45]. This changes the energy of Andreev bound states in the hybrid, which mediate ECT and CAR between the two QDs [46]. As a result, ECT and CAR rates are modulated by V_H through their dependence on the ABS energy [14,38].

The QD excitations are spin polarized by applying an external magnetic field [47–49]. We include in our analysis measurements for both $B_Z = 150$ mT and $B_Z = 100$ mT applied along the NW axis. The measurements to be classified by the CGAN are taken with both normal leads grounded ($V_L = V_R = 0$). The Δ/t ratio is determined using finite bias measurements. We set $V_L = -V_R = 25 \mu\text{eV}$ to ensure only ECT occurs at a finite rate, which is maximum when $\epsilon_L = \epsilon_R$. In turn, we use a symmetric bias $V_L = V_R = 25 \mu\text{eV}$ to have only CAR, which is maximum when $\epsilon_L = -\epsilon_R$. The ratio between CAR and ECT can be estimated from finite bias experimental measurements through $\Delta/t \approx I_{\text{max}}^{\text{CAR}}/I_{\text{max}}^{\text{ECT}}$ [38]. These measurements are taken for different values of V_H , together with $V_L = V_R = 0$ QD-QD scans, to benchmark the CGAN

predictions for different Δ/t ratios. See Fig. 10 for an example determination of I_{\max}^{CAR} , following the method of Wang *et al.* [50].

III. RESULTS AND DISCUSSION

This section is divided into two parts. First, we focus on the predictions of the Δ/t ratio of the CCGAN for the simulated data of the Kitaev model. Second, we analyze the results and predictions for the experimental measurements of the two-quantum-dot system in the InSb nanowire [12].

A. Simulated data

In this section, we apply the CCGAN, in particular, the trained discriminator, to predict the exact value of the Δ/t ratio of the simulated test set, consisting of 10 000 samples. Figure 3 shows the predicted versus real values of the whole test set. The training set consists of 140 000 samples for $\Delta/t \in [0.3, 1.7]$. The computed conductance images mimic the experimental measurements shown in the next section (more information about the training parameters can be found in Appendix A). Taking all data points into account leads to a mean error (δ) of $\delta_{\text{test}} = 0.087$. In general, the predictions of the CCGAN for the test set are very accurate and follow the ideal match (dotted orange line) closely. The deviations from the line are related to the intrinsic randomness of the CGAN algorithm as well as the added noise in the creation of the data set. The missing one-to-one correspondence of the Δ/t parameter for a given conductance image and the added noise leads to uncertainty when inferring the Hamiltonian. We observe in Fig. 3 that the discriminator predicts the outer boundaries, especially $\Delta/t = 0.3$, with higher probability than other values in the vicinity close to the boundaries. This might be related to an undersampling of the boundary regions compared to parameters closer to the center. We also observe a general bias towards smaller ratios in Fig. 3, which can be explained by a large overlap between the distributions within the sweet-spot regime for different parameter values in the vicinity of $\Delta/t \approx 1$. The differential conductance becomes degenerate and it is (visually) difficult to distinguish between the ECT and CAR regime. A similar behavior can be found in Ref. [51] for the Ising model at low temperatures.

We divide the Δ/t ratio into the two regimes, the CAR and ECT regime above and below the sweet spot, and use the CCGAN as a classifier leads to an overall accuracy to predict the regime of 95.3% for the test set data. Taking into account that we predict the absolute values of Δ/t , the parameter values directly around $\Delta/t = 1.0$ lead to most of the classification errors and, therefore, the accuracy of 95% shows the predictive power of the CCGAN. The classification is done by regression into larger or smaller $\Delta/t = 1$ and then used to classify the regime, this leads to some errors around $\Delta/t = 1$. If we add a sweet-spot zone

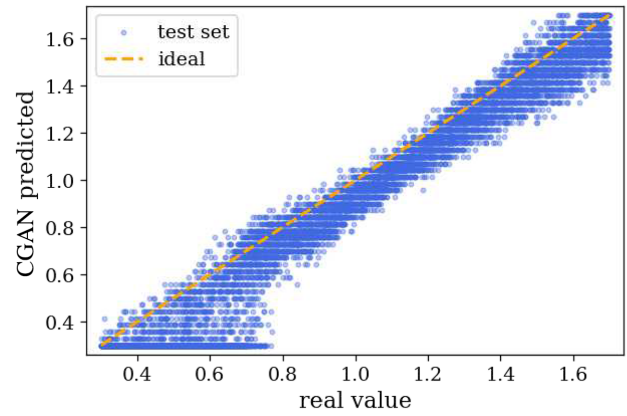


FIG. 3. CCGAN predictions vs. real values of the Δ/t -ratio for the full test set consisting of 10 000 samples in the interval range from 0.3 to 1.7. The orange line shows the ideal prediction case.

between 0.95 and 1.05 we increase the accuracy to 97%. In this case, we consider it as a third regime and do not count it in the accuracy of the previous two-regime classification, i.e., excluding small regression errors around $\Delta/t = 1$.

Figure 4 shows three examples of conductance images, one for each parameter regime, randomly taken from the test set. The panels of Figs. 4(a), 4(c), and 4(e) show the generated conductance image with the corresponding D -rating function [panels Figs. 4(b), 4(d), and 4(f)], as explained in Sec. II B. For all three cases, the D -rating function has a maximum at or in the near vicinity of the real maximum value. However, in the vicinity of the sweet spot [Figs. 4(c) and 4(d)] the “ D -rating function” does not decay as fast as for the other two cases with parameter values further away from the sweet spot. Therefore, the discriminator is more confident in determining the maximum value in Figs. 4(b) and 4(f), showing a higher uncertainty for the predictions around the sweet spot in (d). This uncertainty cannot be determined with supervised NNs using CNNs, which just predict an absolute value. The reason for the increased accuracy can be related to the blurred conductance image without a clear avoided crossing, which is characteristic when approaching the sweet spot from both the ECT and CAR regime.

In addition to the Hamiltonian inference with the discriminator, the generator of the CCGAN can generate high-quality conductance images for a given Δ/t ratio. In Fig. 5, we used the generator to create images for $\Delta/t = 0.7, 0.95, 1.15$, which can be compared with experimental measurements with the same or similar measured parameter shown in Figs. 5(c), 5(f), and 5(j). The generated images for the ECT regime, illustrated in Figs. 5(a) and 5(b), show the expected behavior of the avoided crossings as depicted in Fig. 2(b). The differences between Figs. 5(a) and 5(b) are related to the intrinsic randomness of the CCGAN, which relates to noise and uncertainties in the experimental measurements as well as differences between

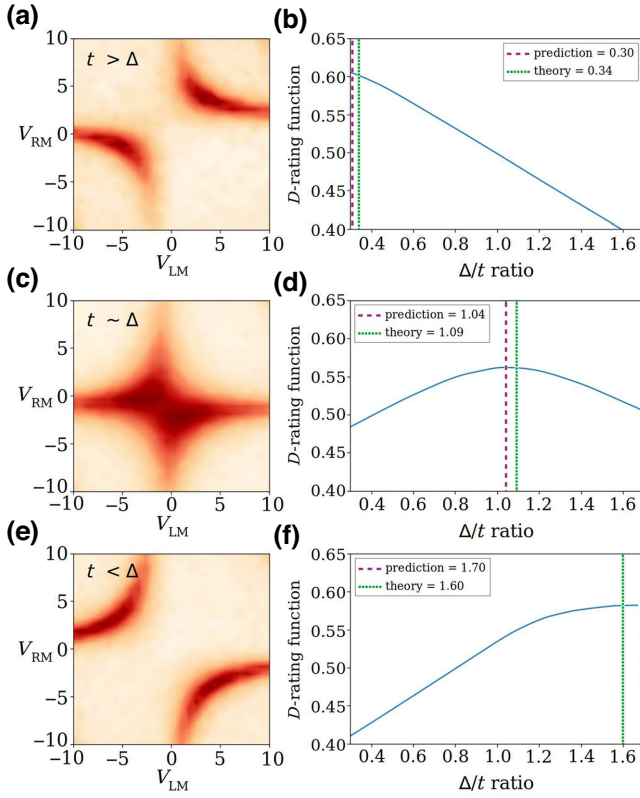


FIG. 4. Predictions for the Kitaev model with the CCGAN for the three regimes of the Hamiltonian. Examples are taken from the test set. (a),(b) show an example for the $t > \Delta$ regime, (c),(d) the “sweet spot” around $t \approx \Delta$, and (e),(f) the $t < \Delta$ regime. (a),(c),(e) are plots of the differential conductance G , (b),(d),(f) show the D -rating function corresponding to each image.

the Kitaev model and real-world data. The same effects can be seen in the sweet-spot regime in Figs. 5(d) and 5(e), as well as in the CAR regime in Fig. 5(h) and Fig. 5(i), where we obtain the expected avoided crossings with a similar magnitude and width.

In Fig. 5, we observe good agreement with our simulations for the CAR regime ($\Delta/t = 1.15$) when comparing Figs. 5(h) and 5(i) with Fig. 5(j). For the ECT-dominated regime (a)–(c), we observe differences in our model with the measurement Fig. 5(c) in the shape and size of the avoided crossings. The differences can be related to the approximation of $t = 1 = \text{const}$, which is usually not the case in the experimental setup where t varies among different measurements, which has an effect on the size and distance of the avoided crossings. For the experimental sweet spot in Fig. 5(f) we picked the conductance image where the avoided crossings vanish and therefore visually can be determined as $\Delta/t \approx 1$. However, the current-inferred ratio is $\Delta/t = 0.85$, which hints at some difficulties in experimentally determining the Δ/t ratio. We find that by picking $\Delta/t = 0.95$, the generator is able to reproduce the image of Fig. 5(f) accurately.

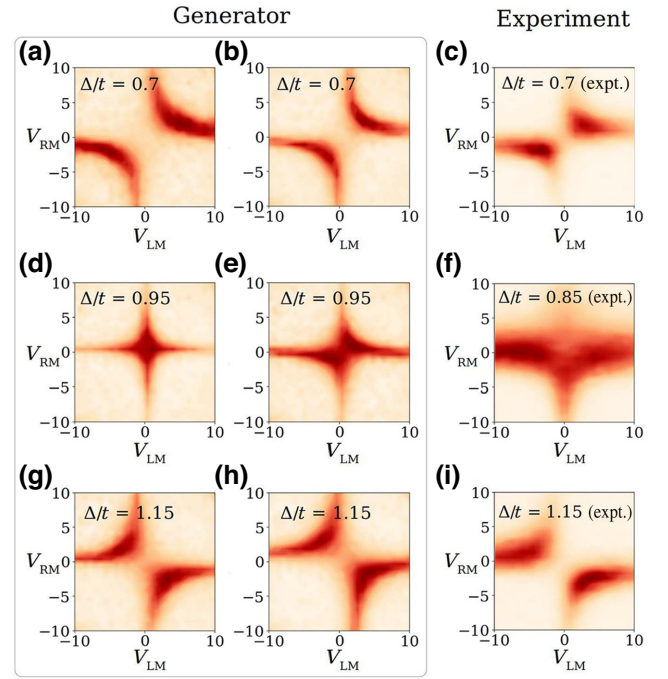


FIG. 5. Generated images with the CCGAN for the three parameter regimes in comparison to measurements. (a)–(c) show the ECT-dominated regime for $\Delta/t = 0.7$, (d)–(f) the vicinity of the sweet spot for $\Delta/t = 0.95$ (0.85 from current measurements), and the CAR regime in (h)–(j) for $\Delta/t = 1.15$. The experimental images have been cut to 28×28 pixels around the center of the avoided crossings.

The individual analysis of the discriminator and generator in Figs. 4 and 5 are a strong indication that the CCGAN is trained properly. Furthermore, we want to highlight at this point again that the overall goal is to create an algorithm that works for both, theory as well as for experimental data, which will be analyzed in the next section.

B. Experimental data

The same CCGAN used in the previous section is now applied to the local differential conductance measurements on the experimental two-quantum-dot system described in Sec. II C. We take the parameter range of $\Delta/t \in [0.3, 1.7]$, covering a wide range of values. As explained in Sec. II C, changing the hybrid electrochemical potential with gate V_H leads to a wide variety of behavior in the conductance measurements of the quantum dot system, ranging from the CAR to the ECT-dominated regime, including the sweet-spot condition of $\Delta \approx t$ where the “poor man’s Majorana” modes are predicted to appear. During the tuning process of the two-quantum-dot system into different Hamiltonian regimes, it is possible to experimentally extract values for the ratio Δ/t (see Sec. II C).

The experimental conductance measurements can be fed into the ML workflow as explained in Appendix C. In this work, we consider two conductance measurements series containing 29 and 39 measurements with the experimentally extracted values for $\Delta/t \in [0.3, 1.7]$, measured at a magnetic field of $B_Z = 150$ mT and $B_Z = 100$ mT, respectively. The results of the 68 measurements are shown in Figs. 7(a) and 7(b) and directly compare the CCGAN predictions with the experimentally current-inferred values. We analyze both measurement series separately and show the (averaged) CCGAN prediction versus the current-inferred Δ/t ratio for 29 and 39 measurements. The standard deviation of the CCGAN predictions are obtained via the workflow described in Appendix C, the error for the current-inferred parameters as described in Sec. II C. Analyzing the 68 measurements, some of the current-inferred parameters in the parameter range between $\Delta/t = 0.8$ and $\Delta/t = 1.2$ show some inconsistencies between the current-inferred value and the appearance of the avoided crossings in the conductance measurements. For the first measurement series [Fig. 7(a)], we detect three experimentally wrongly labeled measurements, where the avoided crossings are visually in a different regime than determined by the current-inferred Δ/t value. These three mislabeled conductance measurements are shown in Figs. 6(a)–6(c), in which the avoided crossings appear in the CAR-dominated regime of $\Delta/t > 1$ even though the measured parameters are smaller than 1, i.e., in the ECT regime. These measurements can visually be classified in the CAR. The predictions of the CCGAN predict the correct regime for these three mislabeled measurements correctly. This is the first hint about the accuracy of the experimentally extracted values and that our algorithm can also be used for outlier detection. The averaged mean error between the measurement values and predictions of the CCGAN, excluding the mislabeled measurements, is $\delta_{\text{total}_1} = 0.164$ ($\delta_{\text{total}_1} = 0.148$) for the first 29 measurements. The classification into the correct regime leads to a classification accuracy of 89.7% taking all 29 measurements into account and 100% when neglecting the wrong labels shown in Fig. 6. The predictions of the CCGAN follow the trend of the “ideal match” with the experimentally current-inferred values well taking into account the huge error bars of the experimental values of 20–30% and the CCGAN uncertainties. Particularly for the ECT-dominated regime of $\Delta/t < 1$ the predictions follow the “ideal-matching” trend well opening the possibility for tuning the system towards the sweet-spot conditions.

Figure 7(b) shows the results for the second measurement series, including 39 measurements at a Zeeman field of $B_Z = 100$ mT. Similar to the first series, some difficulties arise around the sweet spot of $\Delta/t \approx 1$. For five of the measurements, the noise levels are exceptionally high (orange dots) and the classification into the three different regimes becomes difficult since it is not possible

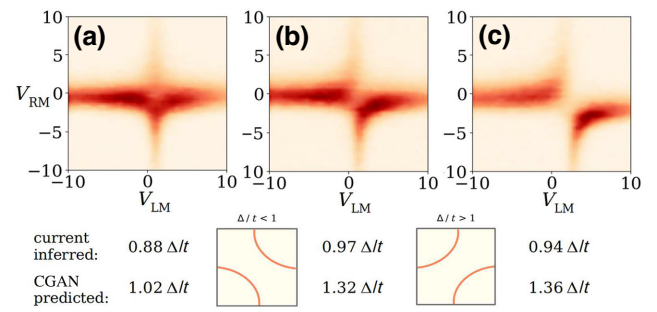


FIG. 6. Shown are the three mislabeled differential conductance images of Fig. 7 with their corresponding Δ/t ratio (a)–(c). The current-inferred parameter ratios of the measurements predict the avoided crossings visually to be in the wrong regime [see Fig. 1(b)].

for us to visually or experimentally verify the assigned label. Moreover, the CCGAN predicts the wrong regime for two measurements (purple dots) and for one sample, the experimental current-inferred value is clearly in the wrong regime (red dot). Excluding the noisy samples, we obtain a mean error of $\delta_{\text{total}_2} = 0.154$, whereas including the noisy samples leads to $\delta_{\text{total}_2} = 0.191$. The accuracy to predict the correct regime is 94% when we consider the 34 measurements, excluding the five noisy ones.

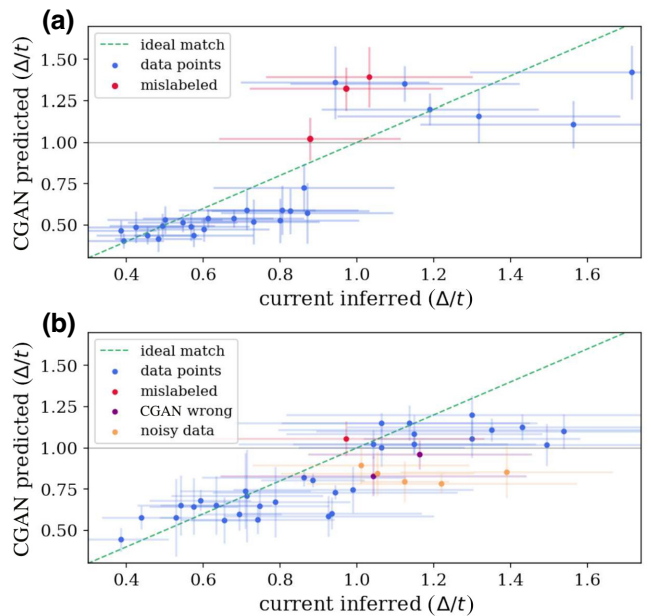


FIG. 7. Predictions of the CCGAN versus current-inferred labels of the Δ/t ratio for the two measurement series with a magnetic field of $B_Z = 150$ mT in (a) and $B_Z = 100$ mT in (b). The dotted green line shows the ideal match and the gray line marks the phase classification of the CCGAN. The standard deviation of each measurement is given as an error bar in the x - y direction and the prediction workflow is described in Appendix C.

For the second measurement series, we observe that the CCGAN predicted values are following the “ideal match” line with the current-inferred parameter and we obtain similar results for the mean error and for the classification accuracy. Particularly for the ECT-dominated regime ($\Delta/t < 1$) we clearly observe the linear trend towards the sweet spot. In the vicinity of the sweet spot ($\Delta/t \approx 1$) the magnitude of the errors increases, and therefore it is more difficult to make predictions or infer parameters close to the sweet spot. In the CAR-dominated regime ($\Delta/t > 1$) the upwards trend is not as clear as for the ECT regime and almost constant and the CCGAN-predicted values are generally of smaller magnitude as the current-inferred values.

However, these results have to take into account different potential error sources: (i) general prediction errors of the CCGAN discriminator due to the limited amount of training data and intrinsic randomness as discussed for Fig. 4, (ii) differences between the Kitaev model approach and real-world data considering the simplicity of the model as well as the approximations we used, and (iii) errors in determining the values experimentally by current measurements considering the huge error bars for these values. As mentioned earlier, in the Kitaev Hamiltonian, we set $t = 1$ and vary only Δ , which is a simplified view compared to the experimental conditions where t can vary during a measurement series. Therefore, by considering only the relation between t and Δ , our model does not consider differences in the avoided crossings related to the absolute magnitude of t . This could be a reason for the higher differences in the CAR-dominated regime in this measurement series. Furthermore, in our analysis, we do not consider the “quality” of each individual measurement, which is related to the noise and visibility of the avoided crossings. Images with very high noise magnitudes deviate from the theoretical model and the training data too much and, therefore, the CCGAN cannot make as accurate predictions about them as for “perfect” measurements. Considering the experimentally wrong-labeled measurements and relatively high error bars, the current-inferred values cannot be taken as ground truth parameter, but rather be used to give us an idea about the accuracy of the CCGAN predictions. For most of the mislabeled measurements, the CCGAN predicts the correct parameter regime.

Another reason why the vicinity of the sweet spot is harder to predict is that there are only very small visual differences on each side of the sweet spot in the local conductance. The absence of clear avoided crossings and higher noise levels lead to higher uncertainties in the prediction. Therefore, to predict the sweet-spot regime accurately, very clean or additional measurements such as the nonlocal conductance are required.

Following the previous discussion, we do not expect that the CCGAN predictions with the current-inferred extraction will yield an ideal match line since the experimental

values can not be seen as perfect values. Furthermore, corrections beyond the minimal model considered can lead to deviations when applying the algorithm to real-world data. Nonetheless, we observe an overall good agreement between our ML algorithm trained on the Kitaev model and the experimental data, the strength of our approach to classify the correct regime and detect outliers in the current-based parameter extraction.

The procedure of predicting the Δ/t value with the CCGAN is illustrated in Fig. 8 for two example measurements. The measurement data, shown in Figs. 8(a) and 8(b), is cropped to 28×28 pixels Figs. 8(c) and 8(d), which is the input size of the CCGAN. The cropped and normalized image is then fed into the discriminator leading to the D -rating function Figs. 8(e) and 8(f) for each measurement. The cropping of the image can be done arbitrarily around the center, which, however, leads to slightly different results. The averaging process to extract the Hamiltonian parameter for the results in Fig. 7 is explained in more detail in Appendix C. Due to the mentioned variation in the training data around the imaginary crossing of the differential conductance, the cropping of the images can also vary up to five pixels away from the center in x and y direction, as, e.g., seen in Figs. 8(c) and 8(d). We use this variation around the center with arbitrary chosen crossing to give a standard deviation for our prediction by averaging over 100 configurations. Furthermore, this variation allows us to automate the centering process with a nonperfect center-finding algorithm.

The fact that our method is not affected by small variations of the actual center is a possible advantage over classical fitting methods, which might lead to very different results when choosing different center positions. Considering that in reality the symmetry of the (avoided) crossings in the differential conductance does not exist due to, e.g., negligible capacitive coupling between the plunger gates of the quantum dots. If this is non-negligible, the avoided crossing lines become tilted and cross over to other quadrants as seen in Ref. [52]. These issues can be overcome by our CCGAN algorithm.

Let us conclude this section with a benchmark comparison to a standard feed-forward model. We used the same Kitaev chain model, described in Sec. II, and training data used for the CCGAN to train a supervised CNN network and make predictions for the same 68 experimental data points. In Fig. 9, we show the CNN-predicted Δ/t value compared with the current-inferred value for the two measurement series at $B_Z = 150$ mT in Fig. 9(a) and $B_Z = 100$ mT in Fig. 9(b). In general, the classification accuracy of the CNN is 92 and 94% and, therefore, slightly lower compared to the CCGAN value. The mean error for both measurement series is $\delta_{\text{CNN-1}} = 0.150$ and $\delta_{\text{CNN-2}} = 0.170$. Furthermore, the predictions for the ECT-dominated regime, are almost constant and not following the trend of the “ideal match” as it was the case before.

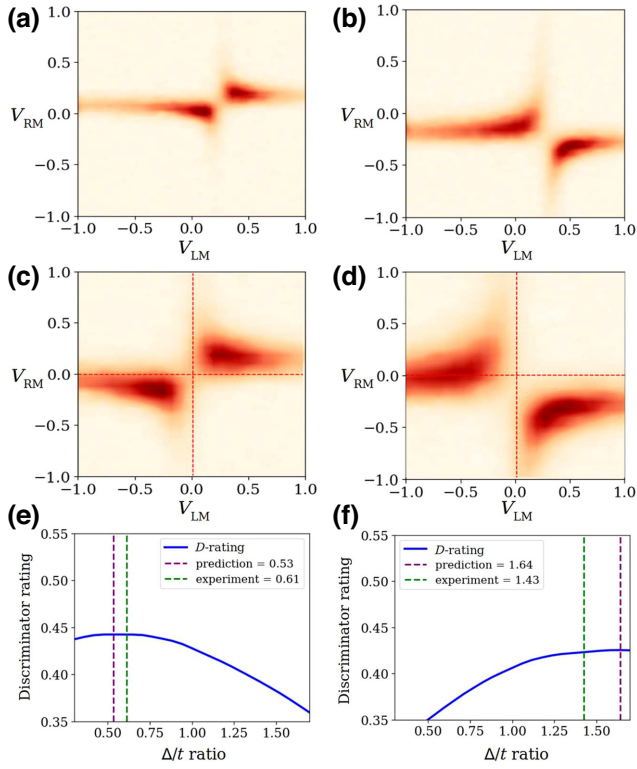


FIG. 8. Predictions for the experimental measurements of the two-quantum-dot system with the CCGAN trained on the full data set for two examples of the measurement series. The original measurements are shown in (a),(b). (c),(d) are the corresponding cropped images, which are the inputs to the CGAN. (e),(f) show the discriminator-rating function predicted by the discriminator of the CCGAN.

More details can be found in Appendix B. In general, we find that the CCGAN yields better results than the simple supervised-learning approach with CNNs when estimating the ECT-CAR amplitude ratio as well as for the classification of the regimes.

IV. CONCLUSION

In this work, we show a demonstration of CGANs for Hamiltonian learning of simulated and real experimental measurements in quantum-dot-based Kitaev chain systems that host poor man's Majoranas in a parameter sweet spot. We show that our algorithm is capable of predicting both simulated and experimental measurements with good accuracy as well as giving an estimation about the uncertainty of each prediction via the discriminator-rating function. This is a clear advantage over supervised neural networks predicting just a single value for each input. This work can be seen as the first proof of principle that CGANs can be a powerful tool to support experimental measurements and extract underlying Hamiltonians from real-world problems. It is worthwhile to note that, taking

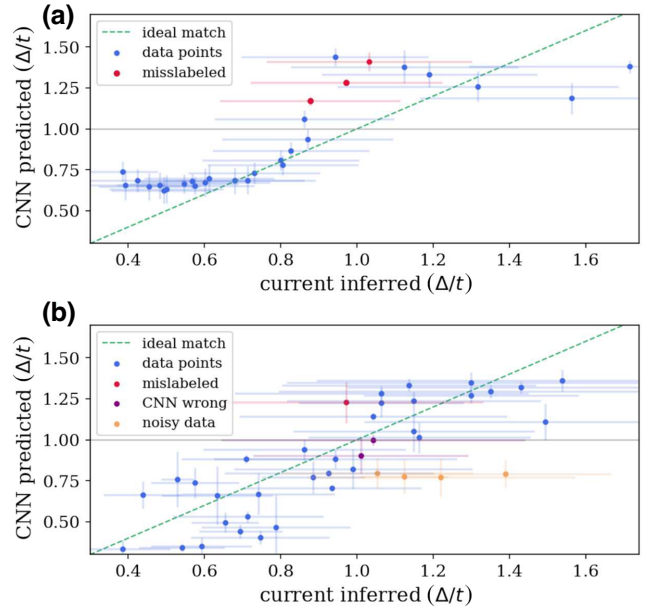


FIG. 9. Predictions of the CNN (supervised) versus experimental labels of the Δ/t -ratio for the two experimental measurement series. The standard deviation of the prediction is given as an error bar for each data point. Panel (a) shows the 29 measurements at $B_Z = 150$ mT and (b) 39 measurements at $B_Z = 100$ mT.

into account all sources of error and the still existing difficulties in training GANs, there is room for improvements in the training of the GAN as well as pre- and postprocessing of especially the experimental data. Interestingly, we show that even when training the CGAN on the relatively simple theoretical model yields robust Hamiltonian prediction and in the case presented here outperforms previously known labeling methods for the experimental measurements.

Our ML tuning approach can relatively easily be embedded into an experimental tuning workflow, run directly in the laboratory, and is generally independent of the experimental setup. We also give a possible prospect of our approach for the application to longer chains. This is interesting because a long chain of quantum dots is required for a Kitaev chain to be considered topological, which constitutes a much more complex system. To observe Majorana zero modes at the ends of a Kitaev chain, $t = \Delta$ has to be satisfied locally for neighboring quantum dots. In other words, an N -site chain can be tuned to host Majoranas at its ends by tuning one pair of quantum dots at a time, as shown in Ref. [53]. Therefore, we believe that our present CCGAN can be applied to Kitaev chains of arbitrary lengths, provided that the density of states can be probed locally. We thus put forward an alternative figure of merit for experimental tuning up Kitaev-chain systems into the $t = \Delta$ sweet spot.

CODE AND EXPERIMENTAL DATA AVAILABILITY

This paper is supplemented by a GitLab repository with all the code and data necessary to reproduce our results <https://gitlab.com/QMAI/papers/adversarialkitaevchain>. The raw experimental data, as well as code used for labeling, are available at <https://zenodo.org/record/7798010>.

ACKNOWLEDGMENTS

We acknowledge useful discussions with Frank Schäfer and Julian Arnold. We acknowledge Francesco Zatelli for contributing the scattering matrix code used to generate the training data. R.K. and J.L.L. acknowledge the computational resources provided by the Aalto Science-IT project, and the financial support from the Academy of Finland Projects No. 331342, No. 336243, and No. 358088, and the Jane and Aatos Erkkö Foundation. E.G. acknowledges project Engineered Topological Quantum Networks (with Project No. VI.Veni.212.278) of the research programme NWO Talent Programme Veni Science domain 2021, which is financed by the Dutch Research Council (NWO). E.G. and R.K. acknowledge support of the Kavli Institute of Nanoscience Delft. D.v.D. and A.B. acknowledge funding by the Dutch Organization for Scientific Research (NWO).

APPENDIX A: ARCHITECTURE AND TRAINING OF THE CCGAN

1. Architecture

The code for the CCGAN and for the CNN can be found in Ref. [54]. The code uses Tensorflow [55] and Keras [56]. The generator architecture of the CCGAN is shown in Table I, consisting of eight layers and in total 995 329 (trainable) parameters. The layers include an input concatenation of the latent vector (dimension 100) with the continuous conditional parameter, a dense layer followed by deconvolutional layers (C2DT) with dropout of 25% between deconvolutional layers. The output dimension is (28, 28, 1) since we are using “gray-scale” images with the size of 28×28 pixels.

The discriminator network is shown in Table II and consists of 11 layers and 878 721 total parameters. The input layer with dimension (28, 28, 1) is followed by 2 times the combination of a convolutional layer (C2D), a max pooling operation (i.e., down-sampling along its spatial dimensions), and a dropout layer. The result is flattened to a one-dimensional array and concatenated with the conditional parameter, before being fed into two dense layers. We are using the ReLU activation function [57] for hidden layers and a sigmoid function for the last layers. We are using the Adam optimizer [58] for the training of the total CCGAN with a learning rate of 0.00001. For the generator,

TABLE I. Architecture of the generator network of the CCGAN.

Generator Layer	Output shape
Concatenation	101
Dense	6272
Reshape	(7, 7, 128)
C2DT	(14, 14, 128)
Dropout	(14, 14, 28)
C2DT	(28, 28, 64)
Dropout	(28, 28, 64)
C2DT	(28, 28, 1)
Total parameters	995 329

we use the MSE as a loss metric, for discriminator binary cross-entropy.

2. Training

The training set consists of 140 000 conductance images, generated with the Kitaev model introduced in Sec. II. The batch size is set to 32 and we trained the network for a total number of 40 epochs. The training was performed on a Nvidia GeForce GTX 1070 GPU, which takes approximately 5 h.

APPENDIX B: RESULTS: SUPERVISED CNN BENCHMARK

Figure 9 shows the results of the (supervised) CNN for the Hamiltonian inference process. Similar to Figs. 7 and 9 shows the predictions of the CNN versus current-inferred values of Δ/t for the first measurement series at $B_z = 150$ mT in panel Fig. 9(a) and for the second at $B_z = 100$ mT in panel Fig. 9(b).

For the first measurement series, we obtain with the CNN approach a mean error of $\delta_{\text{CNN-1}} = 0.172$ and

TABLE II. Architecture of the discriminator network of the CCGAN.

Discriminator Layer	Output shape
Input	(28, 28, 1)
C2D	(28, 28, 64)
MaxPooling2D	(14, 14, 64)
Dropout	(14, 14, 64)
C2D	(14, 14, 128)
MaxPooling2D	(7, 7, 128)
Dropout	(7, 7, 128)
Flatten	6272
Concatenate	6273
Dense	100
Dense	1
Total parameters	878 721

$\delta_{\text{CNN-1}} = 0.150$ when neglecting the three mislabeled values. The classification accuracy, neglecting these three measurements is 92% (24/26 measurements). Compared to the predictions of the CCGAN (see Fig. 7), the CNN predictions show a step-function-like behavior and a plateau in the ECT and CAR phase, not following the “ideal match.”

The CNN predictions of the second series are shown in Fig. 9(b). Excluding the noisy measurements (orange dots) we obtain a mean error of $\delta_{\text{CNN-2}} = 0.170$ and a classification accuracy of 94%. As for the CCGAN, the CNN predicts for two out of 35 measurements the wrong regime. The deviation from the “ideal match” is larger compared to the CCGAN in Fig. 7(b), especially for the ECT-dominated regime. However, we observe the general upwards trend, following the “ideal match” line mostly within the range of the experimental error bars.

APPENDIX C: EXPERIMENTAL DATA PREPROCESSING AND WORKFLOW

The workflow of the predictions for experimental data with the CCGAN consists of five steps and is also shown in Fig. 8 for two example measurements. First, the original image [e.g., Figs. 8(a) and 8(b)] of the differential conductance has to be cropped from 52×52 to 28×28 pixels [see Figs. 8(c) and 8(d)] to match the dimension the CCGAN is trained on. Second, the cropped image has to be scaled according to the training data, which is normalized to values between 0 and 1 with an average maximum value of 0.88. Therefore, we normalize all experimental measurements to 0.88 to be consistent. Then, we define a parameter grid for the Δ/t ratio by splitting up the interval between 0.3 and 1.7 in 100 steps. Afterward, the discriminator evaluates the cropped image with each point on the parameter grid, giving a “discriminator rating” for each Δ/t ratio yielding a probability distribution as shown in Figs. 8(e) and 8(f) for two example measurements. The maximum value corresponds to the Δ/t -ratio with the highest probability to belong to the corresponding differential conductance. Finally, the first four steps are repeated 100 times for arbitrary croppings around the center position to obtain a standard deviation for the predictions of the discriminator network [59].

APPENDIX D: EXTRACTION OF THE EXPERIMENTAL LABELS

Figure 10 depicts a three-terminal measurement of correlated current as a function of the plunger gate voltages of two quantum dots arranged in a Kitaev chain configuration. The bias is set antisymmetrically $V_L = -V_R$ to ensure that only elastic co-tunneling (ECT) can occur. The graph shows a positive slope indicating that current arises only when both dots are on resonance, indicative of ECT. By taking the maximum of I_{corr} for each value

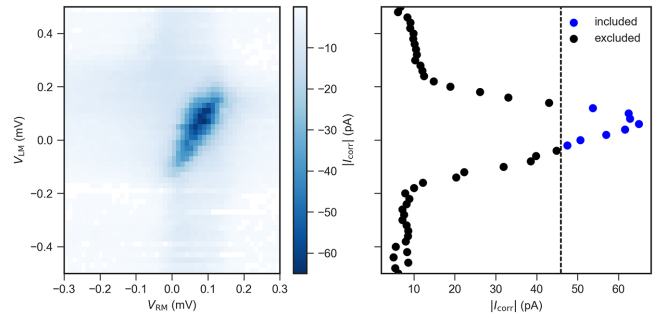


FIG. 10. (Left) The correlated current $I_{\text{corr}} = \sqrt{-\min(I_L \times I_R)}$ as measured on the left and right lead respectively, for varying QD plunger gates V_{LM} and V_{RM} at an antisymmetric bias configuration $V_L = -V_R = 25 \mu\text{eV}$. (Right) The maximum of I_{corr} for every linecut of V_{RM} . The mean $\langle I_{\text{corr}} \rangle$ and standard deviation $\sqrt{\langle I_{\text{corr}}^2 \rangle}$ are calculated from the points above the FWHM indicated by the black dashed line. $\langle I_{\text{corr}} \rangle$ is taken to be $I_{\text{max}}^{\text{ECT}}$ due to the bias configuration allowing only ECT processes.

of V_{RM} , we can estimate $I_{\text{max}}^{\text{ECT}}$, which is proportional to t [38]. Repeating this for a symmetric bias configuration $V_L = V_R$ yields $I_{\text{max}}^{\text{CAR}}$, which is proportional to Δ . Then, a charge-stability measurement is done for $V_L = V_R = 0$, which shows the (un)avoided crossings associated with CAR and ECT. These are labeled according to the Δ/t ratio obtained from the finite bias measurements.

- [1] Masatoshi Sato and Yoichi Ando, Topological superconductors: A review, *Rep. Progr. Phys.* **80**, 076501 (2017).
- [2] C. W. J. Beenakker, Search for Majorana fermions in superconductors, *Annu. Rev. Condens. Matter Phys.* **4**, 113 (2013).
- [3] Steven R. Elliott and Marcel Franz, *Colloquium: Majorana fermions in nuclear, particle, and solid-state physics*, *Rev. Mod. Phys.* **87**, 137 (2015).
- [4] Jason Alicea, Yuval Oreg, Gil Refael, Felix von Oppen, and Matthew P. A. Fisher, Non-Abelian statistics and topological quantum information processing in 1D wire networks, *Nat. Phys.* **7**, 412 (2011).
- [5] David Aasen, Michael Hell, Ryan V. Mishmash, Andrew Higginbotham, Jeroen Danon, Martin Leijnse, Thomas S. Jespersen, Joshua A. Folk, Charles M. Marcus, Karsten Flensberg, and Jason Alicea, Milestones toward Majorana-based quantum computing, *Phys. Rev. X* **6**, 031016 (2016).
- [6] Elsa Prada, Pablo San-Jose, Michiel W. A. de Moor, Attila Geresdi, Eduardo J. H. Lee, Jelena Klinovaja, Daniel Loss, Jesper Nygård, Ramón Aguado, and Leo P. Kouwenhoven, From Andreev to Majorana bound states in hybrid superconductor–semiconductor nanowires, *Nat. Rev. Phys.* **2**, 575 (2020).
- [7] Benjamin E. Feldman, Mallika T. Randeria, Jian Li, Sangjun Jeon, Yonglong Xie, Zhijun Wang, Ilya K. Drozdov, B. Andrei Bernevig, and Ali Yazdani, High-resolution studies of the Majorana atomic chain platform, *Nat. Phys.* **13**, 286 (2016).

- [8] Shawlienu Kezilebieke, Md Nurul Huda, Viliam Vaňo, Markus Aapro, Somesh C. Ganguli, Orlando J. Silveira, Szczepan Głodzik, Adam S. Foster, Teemu Ojanen, and Peter Liljeroth, Topological superconductivity in a van der Waals heterostructure, *Nature* **588**, 424 (2020).
- [9] Lin Jiao, Sean Howard, Sheng Ran, Zhenyu Wang, Jorge Olivares Rodriguez, Manfred Sigrist, Ziqiang Wang, Nicholas P. Butch, and Vidya Madhavan, Chiral superconductivity in heavy-fermion metal UTe₂, *Nature* **579**, 523 (2020).
- [10] A. Yu Kitaev, Unpaired Majorana fermions in quantum wires, *Phys.-Usp.* **44**, 131 (2001).
- [11] Martin Leijnse and Karsten Flensberg, Parity qubits and poor man’s Majorana bound states in double quantum dots, *Phys. Rev. B* **86**, 134528 (2012).
- [12] Tom Dvir, Guanzhong Wang, Nick van Loo, Chun-Xiao Liu, Grzegorz P. Mazur, Alberto Bordin, Sebastiaan L. D. ten Haaf, Ji-Yin Wang, David van Driel, Francesco Zatelli, Xiang Li, Filip K. Malinowski, Sasa Gazibegovic, Ghada Badawy, Erik P. A. M. Bakkers, Michael Wimmer, and Leo P. Kouwenhoven, Realization of a minimal Kitaev chain in coupled quantum dots, *Nature* **614**, 445 (2023).
- [13] Guanzhong Wang, Tom Dvir, Grzegorz P. Mazur, Chun-Xiao Liu, Nick van Loo, Sebastiaan L. D. ten Haaf, Alberto Bordin, Sasa Gazibegovic, Ghada Badawy, Erik P. A. M. Bakkers, Michael Wimmer, and Leo P. Kouwenhoven, Singlet and triplet cooper pair splitting in hybrid superconducting nanowires, *Nature* **612**, 448 (2022).
- [14] Alberto Bordin, Guanzhong Wang, Chun-Xiao Liu, Sebastiaan L. D. ten Haaf, Grzegorz P. Mazur, Nick van Loo, Di Xu, David van Driel, Francesco Zatelli, Sasa Gazibegovic, Ghada Badawy, Erik P. A. M. Bakkers, Michael Wimmer, Leo P. Kouwenhoven, and Tom Dvir, Controlled crossed Andreev reflection and elastic co-tunneling mediated by Andreev bound states, *ArXiv:2212.02274* (2022).
- [15] Jay D. Sau and S. Das Sarma, Realizing a robust practical Majorana chain in a quantum-dot-superconductor linear array, *Nat. Commun.* **3**, 964 (2012).
- [16] C. Di Franco, M. Paternostro, and M. S. Kim, Hamiltonian tomography in an access-limited setting without state initialization, *Phys. Rev. Lett.* **102**, 187203 (2009).
- [17] Sonia G. Schirmer and Daniel K. L. Oi, Two-qubit Hamiltonian tomography by Bayesian analysis of noisy data, *Phys. Rev. A* **80**, 022333 (2009).
- [18] Luigi Seveso and Matteo G. A. Paris, Estimation of general Hamiltonian parameters via controlled energy measurements, *Phys. Rev. A* **98**, 032114 (2018).
- [19] Ian Hincks, Christopher Granade, and David G. Cory, Statistical inference with quantum measurements: Methodologies for nitrogen vacancy centers in diamond, *New J. Phys.* **20**, 013022 (2018).
- [20] Jianwei Wang, Stefano Paesani, Raffaele Santagati, Sebastian Knauer, Antonio A. Gentile, Nathan Wiebe, Maurangelo Petruzzella, Jeremy L. O’Brien, John G. Rarity, Anthony Laing, and Mark G. Thompson, Experimental quantum Hamiltonian learning, *Nat. Phys.* **13**, 551 (2017).
- [21] Valentin Gebhart, Raffaele Santagati, Antonio Andrea Gentile, Erik M. Gauger, David Craig, Natalia Ares, Leonardo Banchi, Florian Marquardt, Luca Pezzè, and Cristian Bonato, Learning quantum systems, *Nat. Rev. Phys.* **5**, 141 (2023).
- [22] James R. Garrison and Tarun Grover, Does a single eigenstate encode the full Hamiltonian?, *Phys. Rev. X* **8**, 021026 (2018).
- [23] Eli Chertkov and Bryan K. Clark, Computational inverse method for constructing spaces of quantum models from wave functions, *Phys. Rev. X* **8**, 031029 (2018).
- [24] X. Turkeshi, T. Mendes-Santos, G. Giudici, and M. Dalmonte, Entanglement-guided search for parent Hamiltonians, *Phys. Rev. Lett.* **122**, 150606 (2019).
- [25] Eyal Bairey, Itai Arad, and Netanel H. Lindner, Learning a local Hamiltonian from local measurements, *Phys. Rev. Lett.* **122**, 020504 (2019).
- [26] Agnes Valenti, Guliuxin Jin, Julian Léonard, Sebastian D. Huber, and Eliska Greplova, Scalable Hamiltonian learning for large-scale out-of-equilibrium quantum dynamics, *Phys. Rev. A* **105**, 023302 (2022).
- [27] Rouven Koch and Jose L. Lado, Designing quantum many-body matter with conditional generative adversarial networks, *Phys. Rev. Res.* **4**, 033223 (2022).
- [28] Agnes Valenti, Evert van Nieuwenburg, Sebastian Huber, and Eliska Greplova, Hamiltonian learning for quantum error correction, *Phys. Rev. Res.* **1**, 033092 (2019).
- [29] Netta Karjalainen, Zina Lippo, Guangze Chen, Rouven Koch, Adolfo O. Fumega, and Jose L. Lado, Hamiltonian inference from dynamical excitations in confined quantum magnets, *Phys. Rev. Appl.* **20**, 024054 (2023).
- [30] Liangyu Che, Chao Wei, Yulei Huang, Dafa Zhao, Shunzhong Xue, Xinfang Nie, Jun Li, Dawei Lu, and Tao Xin, Learning quantum Hamiltonians from single-qubit measurements, *Phys. Rev. Res.* **3**, 023246 (2021).
- [31] Lukas J. Fiderer, Jonas Schuff, and Daniel Braun, Neural-network heuristics for adaptive Bayesian quantum estimation, *PRX Quantum* **2**, 020303 (2021).
- [32] Antonio A. Gentile, Brian Flynn, Sebastian Knauer, Nathan Wiebe, Stefano Paesani, Christopher E. Granade, John G. Rarity, Raffaele Santagati, and Anthony Laing, Learning models of quantum systems from experiments, *Nat. Phys.* **17**, 837 (2021).
- [33] R. Durrer, B. Kratochwil, J. V. Koski, A. J. Landig, C. Reichl, W. Wegscheider, T. Ihn, and E. Greplova, Automated tuning of double quantum dots into specific charge states using neural networks, *Phys. Rev. Appl.* **13**, 054019 (2020).
- [34] J. Darulová, S. J. Pauka, N. Wiebe, K. W. Chan, G. C. Gardener, M. J. Manfra, M. C. Cassidy, and M. Troyer, Autonomous tuning and charge-state detection of gate-defined quantum dots, *Phys. Rev. Appl.* **13**, 054005 (2020).
- [35] Markku P. V. Stenberg, Oliver Köhn, and Frank K. Wilhelm, Characterization of decohering quantum systems: Machine learning approach, *Phys. Rev. A* **93**, 012122 (2016).
- [36] Joshua Ziegler, Thomas McJunkin, E. S. Joseph, Sandesh S. Kalantre, Benjamin Harpt, D. E. Savage, M. G. Lagally, M. A. Eriksson, Jacob M. Taylor, and Justyna P. Zwolak, Toward robust autotuning of noisy quantum dot devices, *Phys. Rev. Appl.* **17**, 024069 (2022).
- [37] Dries Sels and Eugene Demler, Quantum generative model for sampling many-body spectral functions, *Phys. Rev. B* **103**, 014301 (2021).
- [38] Chun-Xiao Liu, Guanzhong Wang, Tom Dvir, and Michael Wimmer, Tunable superconducting coupling of

- quantum dots via Andreev bound states in semiconductor-superconductor nanowires, *Phys. Rev. Lett.* **129**, 267701 (2022).
- [39] Setting the ECT magnitude $t = 1$ is an approximation and one reason for differences compared to the experimental setup where t varies across different measurements. We discuss this in more detail in the results section.
- [40] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, Generative adversarial nets, *Adv. Neural. Inf. Process. Syst.* **27**, 699 (2014).
- [41] Mehdi Mirza and Simon Osindero, Conditional generative adversarial nets, [ArXiv:1411.1784](https://arxiv.org/abs/1411.1784) (2014).
- [42] Daniel Flam-Shepherd, Tony C. Wu, Xuemei Gu, Alba Cervera-Lierta, Mario Krenn, and Alán Aspuru-Guzik, Learning interpretable representations of entanglement in quantum optics experiments using deep generative models, *Nat. Mach. Intell.* **4**, 544 (2022).
- [43] Sebastian Heedt, Marina Quintero-Pérez, Francesco Borsoi, Alexandra Fursina, Nick van Loo, Grzegorz P. Mazur, Michał P. Nowak, Mark Ammerlaan, Kongyi Li, and Svetlana Korneychuk, *et al.*, Shadow-wall lithography of ballistic semiconductor–semiconductor quantum devices, *Nat. Commun.* **12**, 4914 (2021).
- [44] Grzegorz P. Mazur, Nick van Loo, Ji-Yin Wang, Tom Dvir, Guanzhong Wang, Aleksei Khindanov, Svetlana Korneychuk, Francesco Borsoi, Robin C. Dekker, Ghada Badawy, Peter Vinke, Sasa Gazibegovic, Erik P. A. M. Bakkers, Marina Quintero-Pérez, Sebastian Heedt, and Leo P. Kouwenhoven, Spin-mixing enhanced proximity effect in aluminum-based semiconductor–semiconductor hybrids, *Adv. Mater.* **34**, 2202034 (2022).
- [45] Nick van Loo, Grzegorz P. Mazur, Tom Dvir, Guanzhong Wang, Robin C. Dekker, Ji-Yin Wang, Mathilde Lemang, Cristina Sfiligoj, Alberto Bordin, David van Driel, Ghada Badawy, Sasa Gazibegovic, Erik P. A. M. Bakkers, and Leo P. Kouwenhoven, Electrostatic control of the proximity effect in the bulk of semiconductor-superconductor hybrids, *Nat. Commun.* **14**, 3325 (2023).
- [46] Athanasios Tsintzis, Rubén Seoane Souto, and Martin Leijnse, Creating and detecting poor man’s Majorana bound states in interacting quantum dots, *Phys. Rev. B* **106**, L201404 (2022).
- [47] R. Hanson, L. M. K. Vandersypen, L. H. Willems van Beveren, J. M. Elzerman, I. T. Vink, and L. P. Kouwenhoven, Semiconductor few-electron quantum dot operated as a bipolar spin filter, *Phys. Rev. B* **70**, 241304 (2004).
- [48] David van Driel, Guanzhong Wang, Alberto Bordin, Nick van Loo, Francesco Zatelli, Grzegorz P. Mazur, Di Xu, Sasa Gazibegovic, Ghada Badawi, Erik P. A. M. Bakkers, Leo P. Kouwenhoven, and Tom Dvir, Spin-filtered measurements of Andreev bound states, [ArXiv:2212.10241](https://arxiv.org/abs/2212.10241) (2022).
- [49] Alisa Danilenko, Andreas Pöschl, Deividas Sabonis, Vasileios Vlachodimitropoulos, Candice Thomas, Michael J. Manfra, and Charles M. Marcus, Spin-resolved spectroscopy using a quantum dot defined in InAs 2DEG, [ArXiv:2212.10175](https://arxiv.org/abs/2212.10175) (2022).
- [50] Qingzhen Wang, Sebastiaan L. D. ten Haaf, Ivan Kulesh, Di Xiao, Candice Thomas, Michael J. Manfra, and Srijit Goswami, Triplet Cooper pair splitting in a two-dimensional electron gas, [ArXiv:2211.05763](https://arxiv.org/abs/2211.05763) (2022).
- [51] Julian Arnold and Frank Schäfer, Replacing neural networks by optimal analytical predictors for the detection of phase transitions, *Phys. Rev. X* **12**, 031044 (2022).
- [52] W. G. van der Wiel, S. De Franceschi, J. M. Elzerman, T. Fujisawa, S. Tarucha, and L. P. Kouwenhoven, Electron transport through double quantum dots, *Rev. Mod. Phys.* **75**, 1 (2002).
- [53] Ion C. Fulga, Arbel Haim, Anton R. Akhmerov, and Yuval Oreg, Adaptive tuning of Majorana fermions in a quantum dot chain, *New J. Phys.* **15**, 045020 (2013).
- [54] Rouven Koch, David van Driel, Alberto Bordin, Jose L. Lado, and Eliska Greplova, Adversarial Hamiltonian learning of quantum dots in a minimal Kitaev chain, <https://gitlab.com/QMAI/papers/adversarialkitaevchain> (2023).
- [55] Martín Abadi, *et al.*, TensorFlow: Large-scale machine learning on heterogeneous systems, Software available from <https://www.tensorflow.org/>, (2015).
- [56] Francois Chollet, *et al.*, Keras, <https://github.com/fchollet/keras> (2015).
- [57] Abien Fred Agarap, Deep learning using rectified linear units (ReLU), [ArXiv:1803.08375](https://arxiv.org/abs/1803.08375) (2018).
- [58] Diederik P. Kingma and Jimmy Ba, Adam: A method for stochastic optimization, [ArXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
- [59] For some cases, the CCGAN over- or underestimates the probability to exactly 0.3 or 1.7, respectively. This might be related to the undersampling of the boundaries. Therefore, we exclude predictions of exactly 0.3 and 1.7 to obtain the standard deviation.