Bridging the Reality Gap in Quantum Devices with Physics-Aware Machine Learning

D. L. Craig,[1] H. Moon,[1] F. Fedele,[1] D. T. Lennon,[1] B. van Straaten,[1] F. Vigneau,[1] L. C. Camenzind,[2] D. M. Zumbühl,[2] G. A. D. Briggs,[1] M. A. Osborne,[3] D. Sejdinovic,[4] and N. Ares[3,*]

[1]*Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom*
[2]*Department of Physics, University of Basel, 4056 Basel, Switzerland*
[3]*Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, United Kingdom*
[4]*Department of Statistics, University of Oxford, 24-29 St Giles, Oxford OX1 3LB, United Kingdom*

(Received 19 December 2021; revised 16 May 2023; accepted 29 September 2023; published 4 January 2024)

The discrepancies between reality and simulation impede the optimization and scalability of solid-state quantum devices. Disorder induced by the unpredictable distribution of material defects is one of the major contributions to the reality gap. We bridge this gap using physics-aware machine learning, in particular, using an approach combining a physical model, deep learning, Gaussian random field, and Bayesian inference. This approach enables us to infer the disorder potential of a nanoscale electronic device from electron-transport data. This inference is validated by verifying the algorithm's predictions about the gate-voltage values required for a laterally defined quantum-dot device in AlGaAs/GaAs to produce current features corresponding to a double-quantum-dot regime.

DOI: 10.1103/PhysRevX.14.011001

Subject Areas: Computational Physics, Condensed Matter Physics, Semiconductor Physics

I. INTRODUCTION

Differences between theory and experiment pervade all of science and are one of the driving forces of human discovery. Simulations often require fewer resources than real experiments but rarely capture the full complexity of a system, limiting their practical application. Narrowing the gap between a model and the real world is key for the control of complex systems using machine learning, especially when a machine learning model is trained on a simulation before being applied to real systems [1,2]. The reality gap is widened further when there are quantities which are not directly observable. Such unobservable quantities may be estimated through their influence on other characteristics of the system, for example, indirect observation of black holes [3], observation of the signature of Higgs boson decay [4], or machine learning estimation of human poses from behind walls [5].

Solid-state quantum devices of nominally identical design will often display different characteristics. This variability hinders the scalability of otherwise promising qubit realizations, such as in the spin states of electrons confined in electrostatically defined quantum dots [6–8]. Different devices exhibit different electron-transport features for identical gate-voltage values. This variability is even observed in the same device after being exposed to thermal cycling [9]. In particular, electrostatic disorder induced by randomly located donor ions can be a significant source of variability in delta-doped semiconductor quantum-dot devices [10,11]. Recent theoretical work has used deep learning to reconstruct disorder potentials from partial local density-of-states calculations [12]. Confinement potentials of individual quantum dots have been probed using in-plane magnetic fields [13], but there has been no quantitative experimental study of the disorder present in these devices beyond the observation of its effects [14].

Being able to observe disorder potentials and provide a quantitative measure of the extent to which disorder impacts transport properties could inform the growth of semiconductor platforms for quantum device fabrication. It would also benefit the operation of quantum devices, since particular gate-voltage configurations could be chosen to avoid the negative effects of steep gradients in the electrostatic environment. This is of particular importance in the context of electron shuttling, a promising approach for long-range coupling of spin qubits [15–17]. Once the general characteristics of disorder are revealed, the design of gate architectures can be tailored to allow for more disorder-resilient device design. In this work, we enable all these possibilities by quantifying the disorder potential in a quantum device for an arbitrary gate architecture.

*natalia.ares@eng.ox.ac.uk

The presence of disorder could previously be identified only through its effects on the transport of electrons. To gain this insight into the disorder potential, we use a physics-aware machine learning approach.

Our approach produces disorder potentials through the combination of a physical model accelerated by deep learning, Bayesian inference informed by indirect experimental measurements, and dimensionality reduction of the 2D disorder potential using inducing points and random Fourier features. This novel combination can be applied to observe spatial features of any hidden 2D function that influences observations. Such a 2D function could describe different sources of electrostatic disorder in different material systems [8] or other scalar functions such as temperature on chip [18,19]. In particular, we use transport measurements of an electrostatically defined quantum-dot device in an AlGaAs/GaAs heterostructure to inform and verify our approach.

To infer the disorder potential, we use a combination of transport measurements and predictions from a physical model. The physical model is an electrostatic simulation from which transport features can be estimated. Many simulations with different parameter settings are required to compare this physical model with transport measurements. To accommodate this need without extreme computation times, we develop a fast approximation of the model using deep learning. In this way, we make our approach scalable to large device architectures. Our approach can also be adapted to different material systems by using the relevant electrostatic simulation, such as the models considered in Refs. [8,17,20].

The transport measurements and electrostatic simulations inform the inference algorithm to produce plausible disorder potentials, i.e., posterior samples. The inference mechanism used in this paper follows the philosophy of approximate Bayesian computation [21–24] by utilizing the deep learning approximation of the electrostatic model. Deep generative models could be considered for producing disorder potentials, similar to applications in the design of metasurfaces for integrated photonics [25,26]. In contrast to the abstract latent space of a deep generative model, our physical model retains physical meaning throughout the inference process.

A naive implementation of this inference still leads to unrealistically expensive and wasteful computation. This is because electrons are confined in a two-dimensional electron gas (2DEG), and thus the disorder potential to be inferred is a dense 2D function. We develop a novel reparametrization to greatly reduce the dimensionality of the inference problem, while selecting only the most informative regions of the disorder potential. This maps the nonparametric 2D disorder potential into a parametric model where the parameter space retains physical meaning. This reparametrization approximates the function in the spatial and spectral domains simultaneously using an inducing point approximation of a Gaussian process [27–29] and random Fourier features [30–35]. This approach can be adapted to capture varying sources of disorder present in different device realizations by choosing an appropriate kernel for the Gaussian process.

To assess the performance of inference results, we use the disorder potentials produced by the algorithm to predict the electron-transport regime of new measurements. These predictions provide good agreement with experiment, indicating that our physics-aware method is effective. The physical model can determine the number of quantum dots at a given voltage location. Using posterior disorder samples within this model allows us to predict the voltage locations of double quantum dots and verify these predictions with the experiment in different thermal cycles of the device. Our posterior samples also reproduce ground truth disorders in simulation. Results show that our physics-aware machine learning provides a clear advantage over an uninformed model of the disorder potential when predicting the location of double-quantum-dot features in gate-voltage space.

II. THE DEVICE

A bias is applied to Ohmic contacts to drive current through the device from source to drain, and applying voltages to the gates allows for the control of this current. With appropriate gate voltages, electrons may be confined to form quantum dots. Current peaks as a function of the gate voltage in transport measurements are a signature of Coulomb blockade, indicating the formation of quantum dots. A random distribution of Si donor ions contributes a disordered component to the electrostatic potential experienced by electrons confined in a 2DEG. The distribution of donor ions is thought to freeze at low temperatures with rearrangement only possible significantly above device operating temperature [36].

Our device has eight Ti/Au gate electrodes to which DC voltages can be applied to control electron transport in a 2DEG within a GaAs/AlGaAs heterostructure [9,37]. The gate architecture of the device used in the experiments is depicted in Fig. 1(a), where each of the gate voltages can be set to any value between 0 and −2 V. In our device, gate G6 is held at 0 V to avoid leakage currents. The device is operated at millikelvin temperatures.

III. DEVICE MODEL

A. Electrostatics

As part of our physics-aware machine learning method summarized in Fig. 1(b), we require a model of the quantum-dot device. The effects of gate electrodes and donor ions on the electron density in the 2DEG are calculated self-consistently using the pinned surface model [38,39]. Delta doping results in donor ions being randomly

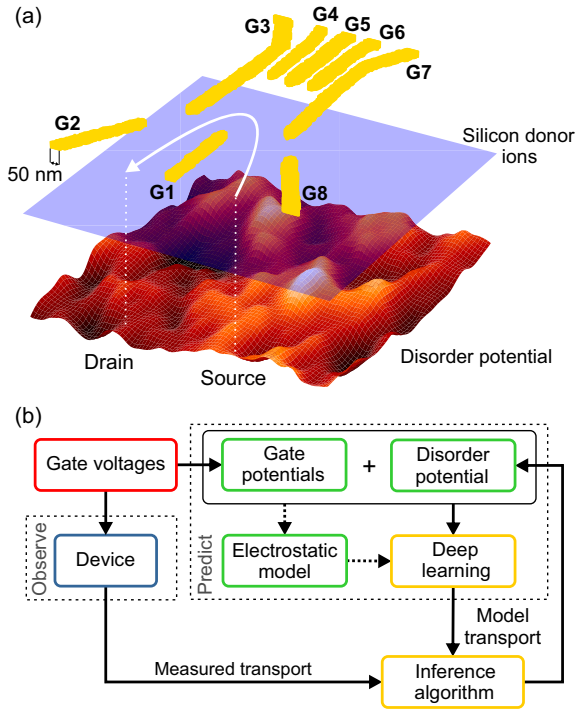

FIG. 1. (a) Device geometry including the gate electrodes (labeled G1–G8), donor ion plane, and an example disorder potential experienced by confined electrons. Typical flow of current from source to drain is indicated by the white arrow. (b) Schematic of the disorder inference process. Colors indicate the following: red for experimentally controllable variables, green for quantities relevant to the electrostatic model, blue for experimental device, and yellow for machine learning methods. Dashed arrows represent the process of generating training data for the deep learning approximation and are not part of the disorder inference process.

located in a plane at a constant height of 45 nm above the 2DEG. With $\mathbf{r} = (x, y)$ denoting a location in the 2DEG plane, the total electrostatic potential is

$$\phi_{\mathrm{tot}}(\mathbf{r}) = \phi_{\mathrm{g}}(\mathbf{r}) + \phi_{\mathrm{d}}(\mathbf{r}) + \phi_{\mathrm{s}}(\mathbf{r}) + \phi_{\mathrm{e}}(\mathbf{r}), \qquad (1)$$

where the electrostatic potential contributions are ϕ_{g} from the gate electrodes, ϕ_{d} from the randomly located donor ions, ϕ_{s} from surface states, and ϕ_{e} from the presence of electrons in the 2DEG.

The potential ϕ_{g} results from the combined effect of the potential from each gate electrode weighted by the applied voltages. It has been shown that this model underestimates the magnitude of ϕ_{g} [38], so we use experimental data to fit an appropriate scale factor for each thermal cycle as discussed in Appendix E. The surface potential is determined by the Schottky barrier with the gates, as discussed by Ref. [40]. Following this work, we set the surface potential to a constant value of $\phi_{\mathrm{s}} = -800$ mV. The potential in the 2DEG from a donor at location $\mathbf{r}_k$ in the donor plane is $\phi_{\mathrm{d}}(\mathbf{r}, \mathbf{r}_k)$. The random potential from all donor ions is then $\phi_{\mathrm{d}}(\mathbf{r}) = \sum_k \phi_{\mathrm{d}}(\mathbf{r}, \mathbf{r}_k)$, summing over the location of each donor. Examples of ϕ_{g} and ϕ_{d} are shown in Figs. 2(a) and 2(b), respectively.

Calculated using the Thomas-Fermi approximation in 2D, the electron density contributes to ϕ_{tot} while also depending on ϕ_{tot}. A self-consistent solution for ϕ_{tot} is computed using an iterative under-relaxation process as in Ref. [39], with an example shown in Fig. 2(c).

B. Modeling the transport regime

To model the transport regime of the device, we consider the transport path of an electron from source to drain. If any point on the transport path has a fully depleted electron density, we say the classical channel for transport is closed (i.e., current does not flow freely). When the channel is closed, the device can be in the quantum-dot regime with transport features from quantum tunneling events, or pinch-off where no current flows at all. When scanning a random combination of all gate voltages, we can approximate the device as an open or closed channel.

A semiclassical electron trajectory between source and drain is calculated by formulating ϕ_{tot} as a graph, where each pixel is a node with nearest-neighbor edges weighted by the mean of connected node values. The minimum spanning tree (MST) [41] of the graph is calculated, and the unique path from source to drain is determined as shown in Fig. 2(e). With the electrostatic potential energy defined as $U(\mathbf{r}) = -e\phi_{\mathrm{tot}}(\mathbf{r})$, the path through the MST has the minimum possible maximum value of U. The location of this point will be called the minimax point, $\mathbf{r}_* = (x_*, y_*)$, with $U^* \equiv U(\mathbf{r}_*)$. If U^* is greater than or equal to the Fermi energy μ_F, the model transport channel is considered closed.

The electron trajectory approximated by the MST path can also be used to determine the number of quantum dots formed by a given ϕ_{tot}. The number of dots defined in the device can be determined using regions of the 1D MST path where $U(\mathbf{r}_*) < \mu_F$, which are delimited by barriers with $U(\mathbf{r}_*) \geq \mu_F$. An example of the electron density and path corresponding to a single dot in our model is shown in Figs. 2(d)–2(f). Since dots in our device are two-dimensional objects in the plane of the 2DEG, the one-dimensional MST path from source to drain is not sufficient to fully determine the number of dots. Additional paths through the MST are calculated to ensure dot labels are robust to all possible configurations of the electron density. Transport features corresponding to quantum dots can be observed only near the closed-channel boundary due to tunnel barriers typically suppressing current far beyond this boundary. The dots identified using our model are not affected by this limitation. An artificial limit on the distance from the closed-channel boundary could be introduced, but this may exclude potentially promising dot locations.

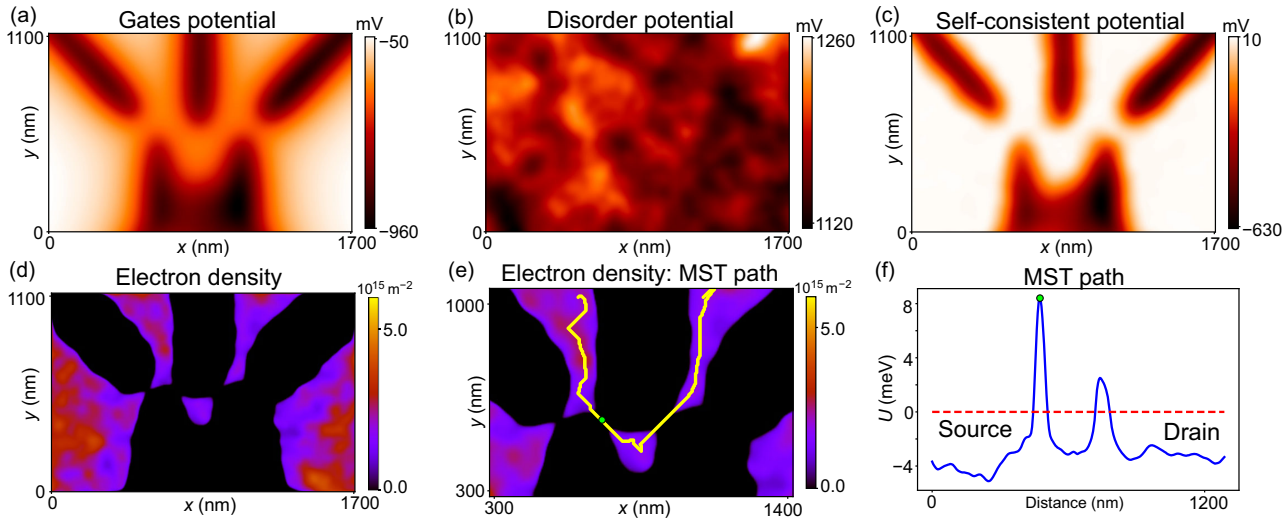

FIG. 2. An example of the device model as discussed in Sec. III. Spatial coordinates x and y are used to indicate the scale of the device. (a) Electrostatic potential from the gate electrodes ϕ_g, (b) a disorder potential ϕ_d, and (c) self-consistent potential ϕ_{tot} given the potentials in (a) and (b). (d) The electron density in the 2DEG given the potential in (c), which shows a single dot. (e) Example of MST path from source to drain in 2D (yellow line) with the location of U^* is marked by a green circle. (f) The potential energy U corresponding to the MST path in (e) with U^* marked by a green circle. The horizontal axis indicates the total distance moved in 2D space. The channel is closed since the value of U^* is above the Fermi level indicated by the red dashed line.

C. Deep learning approximation

For disorder inference, we require fast prediction of the transport regime determined by U^* in our model, given gate voltages and a disorder potential. The self-consistent electrostatic model and MST path require up to 10 sec to calculate U^* in serial computation. This computation time is impractical for the large batches of U^* required for the inference algorithm. Deep learning methods, and their ease of implementation on GPU hardware, allow for a significant acceleration [42,43].

A convolutional neural network (CNN) is trained to calculate $\phi_{tot}(\mathbf{r}_*)$. The architecture of a CNN is particularly suited to data in 2D grids such as the potentials in our electrostatic model. Each input is a 2D potential capturing the influence of the gates, disorder, and Schottky barrier $\phi_{in} = \phi_g + \phi_d + \phi_s$ where ϕ_g and ϕ_d are randomly generated and ϕ_s remains constant. The output training data consist of the self-consistent potential ϕ_{tot} and $\phi_{tot}(\mathbf{r}_*)$ corresponding to each input, with $U^* = -\phi_{tot}(\mathbf{r}_*)$ in units of electron volts. The complete mapping, expressed as $\phi_{in} \to \phi_{tot} \to \phi_{tot}(\mathbf{r}_*)$, is approximated by the CNN $\mathcal{F}_U$. The resolution of ϕ_{in} and ϕ_{tot} is reduced from that used in the electrostatic model to improve the performance of $\mathcal{F}_U$. A series of resolution-preserving convolutions in a residual neural-network architecture [44] learn the nonlinear transformation $\phi_{in} \to \phi_{tot}$, and further layers learn the mapping $\phi_{tot} \to \phi_{tot}(\mathbf{r}_*)$. A recurrent CNN [45] could have been trained to generate the iterative solution; however, we require only the self-consistent potential and not intermediate potentials. Test results achieve a mean absolute error (MAE) of 1.27 meV in U^* estimations, with a 1.2% error when classifying the transport regime using U^*. Batching inputs and using a GPU (GTX 1080 Ti) gives a computation time of 0.6 ms for a single U^* using $\mathcal{F}_U$, a speedup of order 10^4 over the electrostatic model and path-finding algorithm. This evaluation of U^* is also significantly faster than the measurement of current and remains fast for a range of input resolutions, as shown in Appendix B. This is relevant when considering the applicability of our approach to larger and more complex devices such as quantum-dot arrays which would require evaluation of U^* over a wider area of electrostatic potential. The parallel computation of CNN outputs surpasses any acceleration which could be achieved by optimizing the exact computation of transport channel and dot number discussed in Sec. III, which cannot be parallelized.

To make predictions of voltage locations with a given number of quantum dots using disorder inference results, a fast method for counting dots is required. A second CNN is trained to approximate the number of quantum dots at a given set of gate voltages. The input is ϕ_{in} defined for $\mathcal{F}_U$ in the previous paragraph, with the output being the number of dots, $N_{dot} \in \{0, 1, 2, 3\}$. The network learns the mapping $\mathcal{F}_D : \phi_{in} \to P(N_{dot})$, where $P(N_{dot})$ is the probability for a given N_{dot}, and classification is determined by the maximum $P(N_{dot})$. Because of the sparsity of dots in gate-voltage space, the training set used for $\mathcal{F}_U$ is such that the classifier cannot accurately determine N_{dot}, but only the presence or absence of dots. We thus use an intermediate classifier which produces a new training set that ideally includes only gate voltages for which $N_{dot} > 0$. A mixture of selected (dot-abundant) data and the original (dot-sparse) data are used to train $\mathcal{F}_D$. When determining the maximum

number of dots in the direction of a given vector of gates voltages, $\mathcal{F}_D$ achieves 95.8% classification accuracy. Using a GPU with batched inputs, the computation time for a single classification with $\mathcal{F}_D$ is 0.6 ms. Further details of networks $\mathcal{F}_U$ and $\mathcal{F}_D$ can be found in Appendix B.

IV. INFERENCE ALGORITHM

A. Disorder potential reparametrization

The disorder potential used in the electrostatic model is a dense 2D grid covering the entire 2DEG plane, as displayed in Fig. 2(b). A dense grid is unnecessary for inference since ϕ_d is continuous and values can be interpolated from a sparse grid. Using a dense grid would be unfeasible even with the reduced-resolution CNN inputs.

We propose a novel reparametrization algorithm, with the objective to find a set of n_Z locations $Z = \{\mathbf{r}_k^Z | k = 1, \ldots, n_Z\}$, where the disorder potential values on those locations sufficiently determine the transport regime. Following the literature of Gaussian process regression [27], Z defines a set of inducing points. For the experiments in this paper, Z is parametrized as a 14×14 uniform grid defined by two corner points, with the initial grid shown in Fig. 3(a).

Our reparametrization requires the locations of the inducing points Z as well as the values of the disorder potential on these points, represented by a vector $\boldsymbol{\alpha} = [\phi_d(\mathbf{r}_1^Z), \ldots, \phi_d(\mathbf{r}_{n_Z}^Z)]$. The full dense grid of ϕ_d cannot be exactly recovered from the values on Z because the inducing points are too sparse and random disorder potential variations between the points could influence the transport regime. This variability is encoded in the vector $\boldsymbol{\beta} = [\epsilon_1, \ldots, \epsilon_{2q}]$ containing amplitudes of random Fourier features [30–35], where q is the number of frequencies considered. The parameters contained in $\boldsymbol{\beta}$ are thus dependent on Z.

With optimal inducing points Z_{opt}, the disorder potential values contained in $\boldsymbol{\alpha}$ sufficiently determine the transport regime, while the contribution of random Fourier features from $\boldsymbol{\beta}$ is marginal. A numerical optimizer is used to find Z_{opt}, where the optimization objective is to minimize the effect of $\boldsymbol{\beta}$ on transport regime predictions made by $\mathcal{F}_U$. During optimization, the disorder potential contributing to the input of $\mathcal{F}_U$ is approximately reconstructed from Z, $\boldsymbol{\alpha}$, and $\boldsymbol{\beta}$ using a deterministic function f (see Appendix D),

$$\phi_d \approx \hat{\phi} = f(Z, \boldsymbol{\alpha}, \boldsymbol{\beta}). \tag{2}$$

The optimization of Z can be performed on simulated data, and the optimized inducing points used by our inference algorithm are shown in Fig. 3(b). We observe that the inducing points are located where the transport channel is more likely to be depleted, so that the disorder potential on Z_{opt} can determine the transport regime of the device. For different device designs, Z_{opt} will depend on the gate electrode geometry, and the parameters n_Z and q can be chosen to maintain a tractable inference problem with suitable resolution of Z_{opt} for larger, more complex device architectures. Detailed formulation and implementation of the inducing point optimization algorithm can be found in Appendix E.

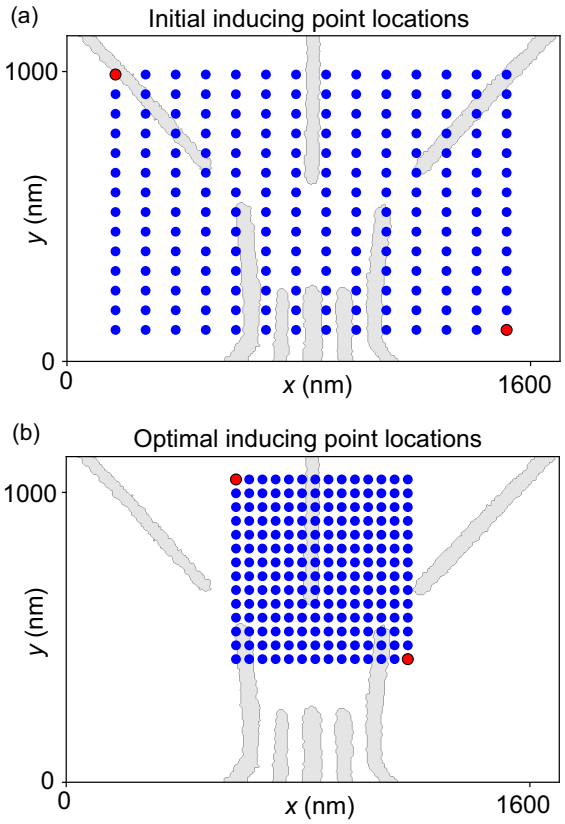

FIG. 3. Gate architecture overlaid with inducing point locations Z indicated by blue dots (a) before optimization and (b) after optimization. Red circles indicate the two corners defining the rectangular array of inducing points. The location of these corners are optimized to ensure that the disorder potential values at the inducing points determine whether the transport channel is open or closed. Hence, the optimized inducing points are located closer to the source and drain reservoirs.

B. Bayesian inference

To reconstruct the disorder potential, in addition to determining Z_{opt}, we must infer suitable values of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. To do this, the inference algorithm requires measurements of current in gate-voltage space. We generate random directions in the seven-dimensional gate-voltage space, each defined by a unit vector $\mathbf{u}_j$ normalized such that $\max_i |u_j^i| = 1$ and $u_j^i \leq 0$ for $u_j^i \in \mathbf{u}_j$. A specific voltage location is defined as $\mathbf{v} = R\mathbf{u}_j$, where R is the voltage distance along $\mathbf{u}_j$. In particular, the inference algorithm requires information about the location of the boundary between open- and closed-channel transport. To obtain this information, stored in a dataset D, a current trace is conducted along a given $\mathbf{u}_j$ from the origin at $R = 0$ mV to the device voltage limit at $R = 2000$ mV. Each

current trace contributes two entries in D; the voltages immediately before and after current drop to half the open-channel current, paired with $y = 1$ and $y = 0$, respectively. The resulting dataset can be defined by $D = \{(\mathbf{v}_i, y_i)|i = 1, \ldots, 2n_u\}$, where n_u is the number of unit vectors considered. We use $n_u = 200$ in this paper, which is well below the typical data requirements of deep learning methods used to predict features of quantum devices [46,47].

To infer $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ using D, we define a prior distribution $p(\boldsymbol{\alpha},\boldsymbol{\beta})$ and a likelihood of data $p(D|\boldsymbol{\alpha},\boldsymbol{\beta})$. The posterior distribution then follows the Bayes rule $p(\boldsymbol{\alpha},\boldsymbol{\beta}|D) \propto p(D|\boldsymbol{\alpha},\boldsymbol{\beta})p(\boldsymbol{\alpha},\boldsymbol{\beta})$. In our formulation, $p(\boldsymbol{\alpha},\boldsymbol{\beta})$ follows the multivariate normal distribution having the zero mean vector and diagonal covariance matrix. The likelihood function utilizes the estimated U^* from the CNN $\mathcal{F}_U$ for each data point $(\mathbf{v}_i, y_i)$ by calculating ϕ_g from $\mathbf{v}$, and approximating ϕ_d from $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

A set of n_s posterior samples $\{(\boldsymbol{\alpha}_i,\boldsymbol{\beta}_i)|i = 1, \ldots, n_s\}$ can be drawn from Markov-chain Monte Carlo (MCMC) methods. Using Eq. (2), the posterior samples of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ generate a set of 2D disorder potentials $S_{\hat{\phi}} = \{\hat{\phi}_i|i = 1, \ldots, n_s\}$, which can be used for CNN inputs. The CNN computation is differentiable, unlike the electrostatic model and path-finding algorithm, allowing us to use Hamiltonian Monte Carlo [48] with TENSORFLOW PROBABILITY [49].

V. RESULTS

A. Transport channel prediction

From the Bayesian inference process, we obtain a set of posterior samples of the disorder potential $S_{\hat{\phi}}$. The standard deviation of posterior inducing point values used to generate $S_{\hat{\phi}}$ is shown in Fig. 4 for three thermal cycles of the same device. A low posterior standard deviation on an inducing point means the inference algorithm has learned more about the disorder potential at that location. In each case, the posterior standard deviation is lowest in regions surrounding gate G1 (the "nose"). This reflects the possible locations of U^* existing most frequently in these locations, due to the primary role of G1 in depleting the transport path from source to drain.

After performing inference of the disorder potential, the set of posterior samples is used in the electrostatic model approximated by $\mathcal{F}_U$. We verify the posterior prediction of the distance R_C required to close the transport channel in a simulated and experimental device for a set of unit vectors, $\{\mathbf{u}_i|i = 1, \ldots, n_v\}$. We set R_C to be the point at which the current drops below 50% of the open-channel current. For a given $\mathbf{u} \in \{\mathbf{u}_i|i = 1, \ldots, n_v\}$, we calculate the mean value of R_C predicted using each posterior sample in $S_{\hat{\phi}}$. To probe the generality of inference results, we evaluate predictions using the measurements which inform our inference (training data) and measurements which the inference algorithm does not encounter (test data). The training and test datasets use $n_v = n_u = 200$ and $n_v = 300$ unit vectors, respectively.

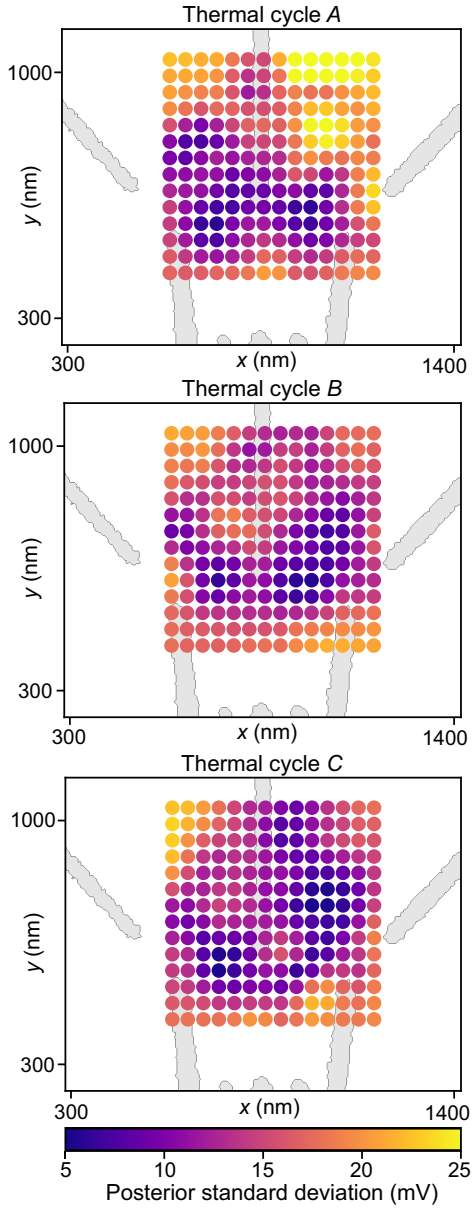

FIG. 4. Disorder inference results using experimental data for three thermal cycles (A, B, (C) of the same device. The inducing point locations Z_{opt} are indicated by circles with the gate structure in the background. The color of each inducing point represents the standard deviation over posterior samples of the disorder potential value at that point.

For a simulated device in which the true disorder can be chosen but is hidden from the algorithm, we compare the performance of random and posterior disorder potentials when predicting the value of R_C over five independent iterations of the inference algorithm. Random disorder potentials generated using the electrostatic model with randomly located donor ions predict R_C with a mean absolute percentage error (MAPE) of 7.0% across training

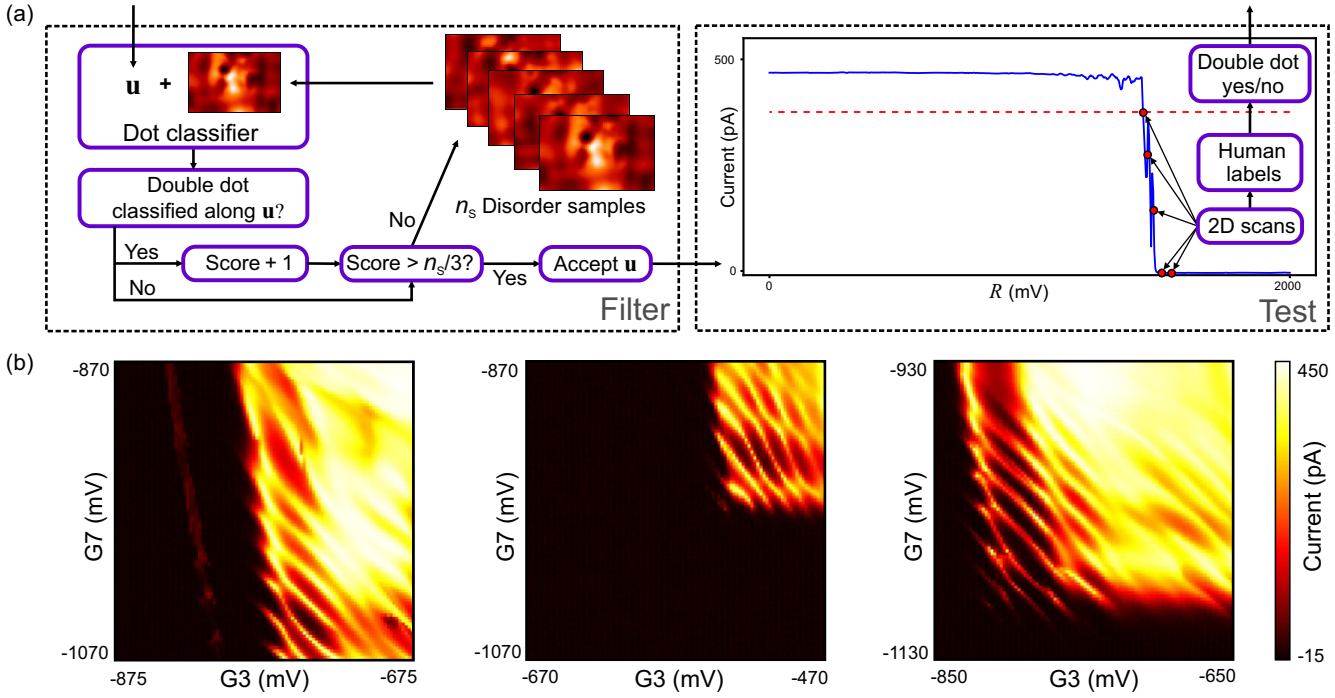

FIG. 5. (a) Predicting double-dot locations. A unit vector $\mathbf{u}$ is passed to the filter which determines whether the vector is considered for the test device (which can be a real or simulated device). The filter uses the dot classifier $\mathcal{F}_D$ to scan along $\mathbf{v} = R\mathbf{u}$ for each of the n_s disorder samples, and the score is increased for each disorder sample which produces a double dot in the scan. Posterior disorder samples from the inference algorithm are shown. Vectors with a score greater than $n_s/3$ are accepted to be tested. An example current trace (solid blue line) of a voltage vector from the origin to the limit of device operation is shown. The dashed red line indicates the 80% threshold used to begin 2D current scans over gates G3 and G7. The 2D scans are taken at intervals along the original current trace (indicated by red circles). The resulting 2D current scans are passed to multiple human experts to label the presence of double quantum dots. (b) Example current scans over G3 and G7 which score highly for double dots when labeled by six human experts for three different unit vectors. Each scan is a 200 mV × 200 mV window with the voltages associated with the direction of $\mathbf{u}$ at the center.

and test data. In contrast, posterior samples predict the value of R_C with a MAPE of 0.3% on training data and 0.5% on test data. These results show that the inference algorithm is successful in finding disorder potentials which effectively describe features of a simulated device.

We then verify the posterior prediction of R_C in a real device. Thermal cycling the device four times, we run a total of five iterations of the inference algorithm. The value of R_C is predicted with a MAPE of 1.5% for training data and 2.0% for test data. The MAE of R_C predictions is 24.3 and 31.9 mV for training and test data, respectively. Random disorder potentials predict R_C with a MAPE of 7.5% across training and test data. The inference results are thus effective in predicting the gate voltages which close the transport channel in a real device. The added complexities of fabrication and materials defects, as well as a quantum-mechanical treatment of transport features such as Coulomb peaks are not considered by our electrostatic model, and hence not directly captured by the Gaussian process, which may contribute to the reduced performance of the inference algorithm compared to the simulated device. Still, this extra complexity may be partially captured by our inference algorithm, since a Gaussian process contains greater flexibility than a direct fit of a physical model. This flexibility, combined with an appropriate Gaussian process kernel, is imperative to the applicability of this algorithm to different device realizations.

B. Double-dot prediction

Having demonstrated the success of the inference algorithm in determining the values of R_C, we use the posterior disorder samples to predict transport features beyond the training domain of the inducing point optimization and inference algorithm. We specifically consider features corresponding to the double-quantum-dot regime. We implement a method requiring minimal knowledge of the transport characteristics of a particular device. Three pieces of information are required: (i) quantum dots form near the closed-channel boundary, (ii) gates G3 and G7 in our device couple most strongly to dot energy levels, and (iii) double quantum dots form features with periodicity in two gate-voltage directions in transport measurements.

The method of finding double quantum dots using posterior disorder samples is summarized in Fig. 5(a). Random unit vectors $\mathbf{u}$ are generated and scanned from $R = 0$ mV to $R = 2000$ mV in the simulated device. A randomly chosen unit vector is unlikely to lead to

double-dot transport features given the sparsity of double dots in voltage space. Based on predictions made by $\mathcal{F}_D$ for each posterior disorder, we select candidate voltage vectors. If $\mathcal{F}_D$ detects a double dot along a vector for a given posterior disorder, the vector's score is increased by 1. Vectors with a score greater than a selected threshold (taken to be $n_s/3$) are accepted to be investigated further in a test device.

Accepted vectors indicate only a direction in gate-voltage space in which double-quantum-dot features could be observed in a test device. As these features are expected to be found near the closed-channel boundary in transport measurements, we investigate multiple voltage locations near this boundary along each accepted vector. To investigate each accepted $\mathbf{u}$, an automated protocol performs a current trace along $\mathbf{u}$ from the origin to the device voltage limit. The gate voltages are then set to the boundary between open- and closed-channel regimes along $\mathbf{u}$, identified by a drop of 20% from open-channel current. To allow for the identification of double-quantum-dot transport features, gates G3 and G7 are scanned in a 200 mV $\times$ 200 mV window centered at this gate-voltage location. Such 2D scans are subsequently performed at intervals of 13.3 mV in R along the direction of $\mathbf{u}$ until the maximum current value in a 2D scan drops below 100 pA. The 2D scans are labeled by six human experts to determine the presence of double-quantum-dot features along $\mathbf{u}$. Multiple human experts are required as double-quantum-dot features are often subjective to human labelers and difficult to identify computationally [9]. Further details of the vector filtering can be found in Appendix G with example labeling of 2D scans shown in the Supplemental Material [50].

Similar to our R_C predictions, we first test the predictive power of posterior disorders in a simulated device in which the true disorder potential is known. By selecting random unit vectors, we find a mean double-dot occurrence rate of 0.83% using several random disorder potentials generated using the electrostatic model with randomly located donor ions. We thus perform disorder inference followed by vector filtering. We do not scan gates G3 and G7 as in the real device since $\mathcal{F}_D$ can determine the number of dots at a point in voltage space. After performing vector filtering, double dots are correctly identified in 28% of instances using random disorders and in 67% of instances using posterior samples. This demonstrates that posterior disorder samples have greater predictive power than random disorder potentials.

In the real device, we produce two sets of posterior samples from independent iterations of disorder inference. Accepted vectors and the associated labels from both iterations are combined to provide larger sets of results using posterior and random disorders. Examples of 2D current scans which score highly for double-quantum-dot features are shown in Fig. 5(b). To assess the success of our dot prediction method, a binomial distribution is fitted to posterior and random results, where the probability of finding a double dot along an accepted voltage vector is $P(\mathrm{DD}) = p$, with $P(\overline{\mathrm{DD}}) = 1 - p$. The fit results in 95% confidence intervals of $0.235 < p_{\mathrm{post}} < 0.449$ using posterior samples, and $0.004 < p_{\mathrm{rand}} < 0.149$ using random disorders.

These values, with a separation of the 95% confidence intervals, demonstrate that using posterior disorders results in a higher rate of success than random disorders in finding experimental double quantum dots. Our results show that the inference algorithm produces disorder potentials with predictive power beyond the original domain of training data and can reduce the human expertise required to tune a double quantum dot.

In addition to the comparison of random and posterior disorders, we also perform the filtering process with featureless (i.e., constant-valued) disorder potentials. Fewer vectors are accepted to be tested than when using posterior or random disorders, and vectors which produce the highest scoring 2D scans, identified in results of posterior disorder predictions, are not found. Further details can be found in Appendix G.

VI. CONCLUSION

We demonstrate that hidden disorder in a nanoscale electronic device can be inferred with indirect measurements and physics-aware machine learning. The reparametrization of the disorder potential proves effective in reducing the dimensionality of the problem, and the successful acceleration of an electrostatic model with a differentiable convolutional neural network allows for Bayesian inference. The entire inference process, from inducing point location optimization to selecting posterior samples, is general and applicable to any gate structure. The device specifics are contained in the electrostatic model and can easily be adapted to other gate architectures and materials. For example, in Ref. [20] we adapt the electrostatic model for a linear dot device laterally defined in a SiGe heterostructure obtaining good agreement with experiment for a device confining holes instead of electrons and with a different gate geometry. Different device realizations can also be considered by using an appropriate physical model and updating the Gaussian process kernel to capture the disorder characteristics.

The prediction of double-dot locations using both random and posterior disorders shows the benefits of model-assisted tuning, and results indicate that the posterior disorders perform better in this task for various thermal cycles of the device. We therefore conclude that the use of physics-aware machine learning narrows the reality gap. The remaining gap between simulation and experiment can be attributed to further unknowns such as fabrication defects in gate design, material defects, and quantum transport properties which are not considered in our physical model. Even with the added complexity of real

devices, the predictive power of posterior disorders indicates that the inference algorithm has the flexibility to effectively capture the electrostatic landscape, and is thus applicable to different types of devices. The generality of this method and the minimal data required for inference are promising qualities for future utility in understanding nanoscale quantum devices. The inference algorithm can be easily scaled to larger, more complex devices. In this way, our approach could be used to study the variability of devices in large arrays, extending over micro- and millimeter areas [51–53]. While our results demonstrate a way to bridge the reality gap introduced by disorder, future devices may be suitably designed such that the gap induced by disorder is no longer crucial to understanding device operation. An alternative combination of deep learning and Bayesian inference could assist in the design of gate architectures which are robust to the impact of disorder. Machine-learning-assisted device design is a promising path forward, and may pave the way for truly scalable quantum devices.

ACKNOWLEDGMENTS

We acknowledge J. Zimmerman and A. C. Gossard for the growth of the AlGaAs/GaAs heterostructure. D. C. would like to thank E. M. Gauger for support and useful discussions on this manuscript. This work was supported by the Royal Society (Grant No. URF\R1\191150), the EPSRC National Quantum Technology Hub in Networked Quantum Information Technology (Grant No. EP/M013243/1), Quantum Technology Capital (Grant No. EP/N014995/1), EPSRC Platform Grant (Grant No. EP/R029229/1), the European Research Council (Grant Agreement No. 948932), FQXi Grant No. FQXI-IAF19-01, the Swiss NSF Project No. 179024, the Swiss Nanoscience Institute, the NCCR SPIN, and the EU H2020 European Microkelvin Platform EMP Grant No. 824109. We acknowledge the use of the University of Oxford Advanced Research Computing facility in carrying out this work.

D. L. C. and H. M. contributed equally to this work.

APPENDIX A: PHYSICAL MODEL

The gate electrodes exist on the surface of the device, and the potential from each gate at a depth $d = 115$ nm beneath the gates in the plane of the 2DEG is determined using analytic expressions. A further 5 nm is added to the depth of the AlGaAs/GaAs heterojunction (110 nm) to account for the extent of the electron density beyond this junction. A representation of the gates is taken from an SEM image of a device of identical design. The image of each gate is used, and the potential from each pixel is calculated individually and summed to give the total potential from each gate $\phi_{g_i}(\mathbf{r})$. The total gate potential is

$$\phi_g(\mathbf{r}) = G_{SF}\sum_i v_i \phi_{g_i}(\mathbf{r}), \tag{A1}$$

where the sum is over all gates and v_i is the voltage applied to the ith gate. The gate scale factor G_{SF} is used as the pinned surface model underestimates the magnitude of the gate potential. This underestimation is observed when gate voltages stop current in the experimental device but do not deplete the transport path from source to drain in the simulated device.

Donor ions exist in a plane at a constant height $h = 45$ nm above the 2DEG. The potential in the 2DEG from a single donor ion at location $\mathbf{r}_{ij} = (x_i, y_j)$ in the donor plane is

$$\phi_d(\mathbf{r}, \mathbf{r}_{ij}) = \frac{e}{4\pi\epsilon\epsilon_0}[(|\mathbf{r} - \mathbf{r}_{ij}|^2 + h^2)^{-1/2} - (|\mathbf{r} - \mathbf{r}_{ij}|^2 + (2d - h)^2)^{-1/2}]. \tag{A2}$$

The potential from each donor ion is summed to give the disorder potential $\phi_d(\mathbf{r})$.

The electron number density is calculated using the Thomas-Fermi approximation in 2D,

$$n(\mathbf{r}) = 2\frac{m^*}{\pi\hbar^2}(\mu_F - U(\mathbf{r}))\Theta(\mu_F - U(\mathbf{r})),$$

where m^* is the effective mass of an electron in GaAs, μ_F is the chemical potential or Fermi level of the 2DEG which is set to zero, and $\Theta(\cdot)$ is the Heaviside step function. The factor of 2 accounts for spin degeneracy, and the Heaviside step function approximates the Fermi distribution at low temperatures. The electrostatic potential associated with the electron density is

$$\phi_e(\mathbf{r}) = -\frac{e}{4\pi\epsilon\epsilon_0}\int \mathbf{dr}'\frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

A self-consistent solution is computed using an iterative under-relaxation process. The device is fabricated from a wafer (Gossard-060926C) with 2DEG density $n = 2.64 \times 10^{15}$ m^{-2} and delta-doping density $n_\delta \approx 6 \times 10^{16}$ m^{-2}. As the 2DEG density can be accurately measured, and there is no guarantee that all Si donors become effective dopants, we fit n_δ such that the electrostatic model produces the known 2DEG density. The fitted value is $n_\delta = 1.25 \times 10^{16}$ m^{-2} giving a calculated mean electron density of $\langle n\rangle = (2.64 \pm 0.04) \times 10^{15}$ m^{-2}, which is the mean and standard deviation uncertainty of 100 calculations. This value is in agreement with the experimental value of 2.64×10^{15} m^{-2}.

A semiclassical trajectory of electrons between source and drain is calculated by formulating the 2D potential ϕ_{tot} as a graph G_ϕ, where each pixel is a node. G_ϕ is defined as

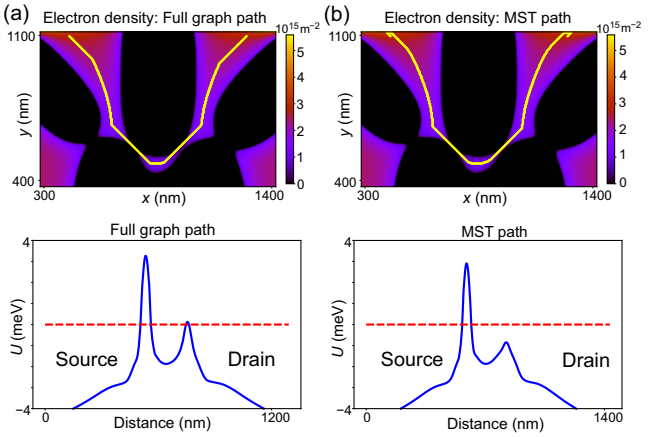

FIG. 6. Comparison of the semiclassical electron trajectory for (a) the full graph G_ϕ and (b) the MST of G_ϕ at identical gate voltages. The electron density with the relevant transport path (yellow line) is shown in the upper plot of each panel. The potential along each path (blue line) is shown in the lower plot of each panel, where the Fermi energy μ_F is indicated by a red dashed line. The "Distance" axis indicates the total distance moved in 2D space. The transport path using the full graph overestimates the potential barrier height as it shows two regions where $U > \mu_F$. The electron density demonstrates that there is only one region where $U > \mu_F$, matching the prediction of the MST path.

$$G_\phi = (V, E), \qquad V = \{v_i\}, \qquad E = \{e_{ij}\},$$
$$v_i = U(\mathbf{r}_i), \qquad e_{ij} = \frac{1}{2}[v_i + v_j], \tag{A3}$$

where V is the set of nodes in G_ϕ with v_i the value of the ith node, and E is the set of edges in G_ϕ with e_{ij} the edge connecting v_i and nearest neighbor v_j.

The Dijkstra algorithm results in a path from source to drain through G_ϕ, but this path overestimates the maximum potential energy of an electron along the path. Gate voltages which close the transport channel are then underestimated by our model. As we discuss in the main text, calculating the MST of G_ϕ resolves this issue by providing a unique path connecting source and drain with a minimum sum of edge weights. A comparison of trajectories computed using the full graph and the MST is shown in Fig. 6. Such trajectories allow for the number of quantum dots in the transport channel to be counted.

APPENDIX B: DEEP LEARNING

A deep CNN, denoted $\mathcal{F}_U$, is trained to approximate U^* given gate voltages and a disorder potential. The training dataset contains 85 000 entries with each input being a potential profile $\phi_{\text{in}} = \phi_{\text{g}} + \phi_{\text{d}} + \phi_{\text{s}}$ where ϕ_{g} and ϕ_{d} are randomly generated. The output training data consist of the self-consistent potential ϕ_{tot} and $\phi_{\text{tot}}(\mathbf{r}_*)$ with $U^* = -\phi_{\text{tot}}(\mathbf{r}_*)$. Training is performed for 100 epochs with a learning rate of 1×10^{-3} (dropping to 2.5×10^{-4} in two steps) and a mean squared error loss function. The resolution of each input is reduced from the high resolution required to accurately compute the training data, with test results shown in Table I.

TABLE I. Performance metrics of $\mathcal{F}_U$ for different resolutions of ϕ_{in} computed using a GPU (GTX 1080 Ti). Max-pooling processes are adapted for each resolution, otherwise the networks are identical. The resolution reduction fraction is applied to both input dimensions. Full resolution (269 411) is required only for computing training data, so it is not considered. Time per output is from a batch of 1000 inputs. MAE is chosen so that units remain in mV, and the transport channel error is based on binary classification using the value of $\phi_{\text{tot}}(\mathbf{r}_*)$ produced by $\mathcal{F}_U$.

ϕ_{in} resolution fraction	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{8}$
Time per training epoch (s)	430	130	69	65
Time per output (ms)	1.66	0.46	0.24	0.18
$\phi_{\text{tot}}(\mathbf{r}_*)$: test MAE (mV)	1.41	1.18	1.25	1.27
Transport channel: test error (%)	1.57	1.23	0.96	1.20

Another CNN learns the mapping $\mathcal{F}_D$: $\phi_{\text{in}} \to P(N_{\text{dot}})$ for $N_{\text{dot}} \in \{0, 1, 2, 3\}$, where classification is taken as the maximum value of $P(N_{\text{dot}})$. An intermediate classifier is used to generate a suitable dataset, as discussed in the main text. Training is performed for 100 epochs with a learning rate of 1×10^{-3} dropping to 5×10^{-4} after 70 epochs. Test results for dot classification have an accuracy of 98.0% on random data (dot-sparse) and 75.9% on selected (dot-abundant) data.

Following the notation used in Table I, $\mathcal{F}_U$ and $\mathcal{F}_D$ use $\frac{1}{8}$ and $\frac{1}{6}$ resolution of ϕ_{in}, respectively. The computation time for both networks $\mathcal{F}_U$ and $\mathcal{F}_D$ is approximately 0.2 ms given a 2D potential input and using a GPU. However, processing a vector of gate voltages into a 2D potential increases this time to 0.6 ms. The processing involves determining the total gate potential using Eq. (A1) and summing this with the disorder potential. Schematics of the neural-network architectures are shown in the Supplemental Material [50].

APPENDIX C: DISORDER COVARIANCE

A Gaussian process requiring a covariance function of the random disorder potential is used to generate random disorder potentials in the inference algorithm. The donor plane is divided into cells, with a random variable $I_{ij} \in \mathbb{N}_0$ determining the number of donors in the cell at $\mathbf{r}_{ij} = (x_i, y_j)$. The potential of the 2DEG from the donor ion distribution is given by $\phi_{\text{d}}(\mathbf{r}) = \sum_{ij} I_{ij} \phi_{\text{d}}(\mathbf{r}, \mathbf{r}_{ij})$, summing over each cell in the donor plane.

The covariance between two points in the 2DEG plane can be evaluated numerically, and appropriate kernel parameters are fitted. The covariance of ϕ_{d} between two points in the 2DEG plane can be computed numerically as $\text{cov}(\mathbf{r}, \mathbf{r}') = \text{var}(I) \sum_{ij} \phi_{\text{d}}(\mathbf{r}, \mathbf{r}_{ij}) \phi_{\text{d}}(\mathbf{r}', \mathbf{r}_{ij})$, where I is the distribution from which each I_{ij} is independently drawn.

Using the correlation function eliminates the dependence on I,

$$\mathrm{corr}(\mathbf{r},\mathbf{r}') = \frac{\sum_{ij}\phi_{\mathrm{d}}(\mathbf{r},\mathbf{r}_{ij})\phi_{\mathrm{d}}(\mathbf{r}',\mathbf{r}_{ij})}{\sqrt{\sum_{ij}\phi_{\mathrm{d}}(\mathbf{r},\mathbf{r}_{ij})^2}\sqrt{\sum_{ij}\phi_{\mathrm{d}}(\mathbf{r}',\mathbf{r}_{ij})^2}}. \quad \text{(C1)}$$

A rational quadratic kernel function

$$k(\mathbf{r},\mathbf{r}') = \sigma\left(1+\frac{|\boldsymbol{r}-\boldsymbol{r}'|^2}{\rho^2}\right)^{-1} \quad \text{(C2)}$$

is chosen, with fitted values of $\rho = 139.8$ nm and $\sigma = 20.8$ mV. Alternative kernels provide better fits, but the explicit form of the corresponding frequency distribution of random Fourier features is unknown or intractable. Different sources of disorder which may be present in other device realizations can be modeled by choosing an appropriate Gaussian process kernel. Kernel parameters can be fitted using physical models, such as in Ref. [8], which considers charge defects and bond disorder in Si/SiGe devices.

APPENDIX D: REPARAMETRIZATION

Let $X = \{\mathbf{r}_k^X | k = 1,\ldots,n_X\}$ denote the set of dense grid points (34×52 or 45×69 for the experiments in the paper, depending on the CNN model) on the $x-y$ plane, the potential of which is the input of the CNN. Without any measurement, the disorder potential values on X denoted by $\boldsymbol{\phi}^X$ is approximately a random vector following the normal distribution:

$$\boldsymbol{\phi}^X \sim \mathcal{N}(m\mathbf{1}, K_X),$$

where m is the precalculated mean potential level, 1 is a one-filled vector, and K_X is the covariance matrix, whose element (i,j) is $k(\mathbf{r}_i^X,\mathbf{r}_j^X)$. The value of $m = 1184$ mV is determined from the mean values of 1000 random disorder potentials generated using the electrostatic model (with ϕ_{s} absorbed into the disorder potential, $m = 384$ mV). For the sake of simplicity, the derivations below are based on the mean-adjusted potential: $\mathbf{f} = \boldsymbol{\phi}^X - m\mathbf{1}$. In order to generate a random sample from $\mathbf{f}$, we can draw a random sample from $\epsilon_X \sim \mathcal{N}(\mathbf{0}, I_X)$ and then transform it as

$$\mathbf{f} = L_X\epsilon_X, \quad \text{(D1)}$$

where $\mathbf{0}$ is a zero-filled vector, I_X is the $n_X\times n_X$ identity matrix, and L_X is the lower Cholesky decomposition of K_X.

Since n_X is too large for a practical Bayesian inference problem, and we want to make the inference algorithm independent of n_X, the inducing point approach is used. The set of inducing points $Z = \{\mathbf{r}_k^Z | k = 1,\ldots,n_Z\}$ usually has many fewer points than X: $n_Z < n_X$. Let $\mathbf{u}$ denote the vector of the mean-adjusted potential values at Z (i.e., $\mathbf{u} = \boldsymbol{\alpha} - m\mathbf{1}$ using notation from the main text). The two mean-adjusted potential vectors $\mathbf{f}$ and $\mathbf{u}$ are jointly a normal distribution, and the joint distribution can be decomposed into two terms: $p(\mathbf{u},\mathbf{f}) = p(\mathbf{u})p(\mathbf{f}|\mathbf{u})$. The first term is a prior distribution $p(\mathbf{u}) = \mathcal{N}(\mathbf{u};\mathbf{0},K_Z)$, where K_Z is the covariance matrix, whose element (i,j) is $k(\mathbf{r}_i^Z,\mathbf{r}_j^Z)$. The second term is the conditional distribution of $\mathbf{f}$ given $\mathbf{u}$:

$$p(\mathbf{f}|\mathbf{u}) = \mathcal{N}(\mathbf{f}|K_{XZ}K_Z^{-1}\mathbf{u}, K_X - K_{XZ}K_Z^{-1}K_{ZX}). \quad \text{(D2)}$$

The computational complexity of generating samples from $\mathbf{f}|\mathbf{u}$ is $O(n_X^3)$ because of the covariance matrix in Eq. (D2). To reduce the computational complexity, any low-rank approximation can be used. In this paper, we approximate the covariance matrix with spectral features. The idea behind this approach is to let inducing points take account of spatially important locations, and the spectral features control relatively unimportant spatial information. The approximation of many types of covariance kernel functions with spectral features is extensively studied in the context of random Fourier features [30–35].

The spectral feature is

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{q}}[\cos(\omega_1^\top\mathbf{r}),\sin(\omega_1^\top\mathbf{r}),\ldots,\cos(\omega_q^\top\mathbf{r}),\sin(\omega_q^\top\mathbf{r})],$$

where q is an arbitrary chosen integer satisfying $n_Z < 2q < n_X$, and w_i is a random sample whose probability density function depends on the underlying covariance kernel function. We use $q = 300$ in this work. The $n_z < 2q$ inequality ensures $\Psi_Z\Psi_Z^\top$ (defined below) is invertible, and $2q < n_x$ ensures that an advantage is gained in computational complexity when using random Fourier features.

The corresponding probability distribution of the samples $\omega_1,\ldots,\omega_q$ given the kernel function (C2) is

$$p(\omega) = \rho e^{-\rho\|\omega\|_1},$$

where ω is an angular frequency, and $\|\cdot\|_1$ is the L1 norm function.

The prior covariance matrices K_X and K_Z are approximated by the spectral features: $K_X \approx \Psi_X\Psi_X^\top$ and $K_Z \approx \Psi_Z\Psi_Z^\top$, where $\Psi_X \in \mathbb{R}^{n_X\times 2q}$ with $\psi(\mathbf{r})^\top$ for $\mathbf{r}\in X$ as rows, and Ψ_Z is defined in a similar fashion. The posterior covariance in Eq. (D2) is approximated as $\mathrm{cov}(\mathbf{f}|\mathbf{u}) \approx \bar{\Psi}_X\bar{\Psi}_X^\top$, where $\bar{\Psi}_X = \Psi_X - \Psi_X\Psi_Z^\top(\Psi_Z\Psi_Z^\top)^{-1}\Psi_Z$. The approximated random field by substituting the approximated covariance matrix into Eq. (D2) is

$$\mathbf{f} \approx K_{XZ}L_Z^{-1\top}\boldsymbol{\epsilon}_Z + \bar{\Psi}_X\boldsymbol{\epsilon}_{2q}, \quad \text{(D3)}$$

where $\boldsymbol{\epsilon}_Z$ and $\boldsymbol{\epsilon}_{2q}$ are standard normal random vectors with length n_Z and $2q$, respectively, and L_Z is the lower Cholesky decomposition of K_Z. The approximated posterior random vector is straightforward,

$$\mathbf{f}|\mathbf{u} \approx K_{XZ}L_Z^{-1\top}\mathbf{u} + \bar{\Psi}_X\boldsymbol{\epsilon}_{2q}. \quad \text{(D4)}$$

The equation defines the reconstruction of $\boldsymbol{\phi}^X$ through the function

$$\phi^X \approx f(Z, \mathbf{u}, \boldsymbol{\epsilon}_{2q}) = K_{XZ} L_Z^{-1\top} \mathbf{u} + \bar{\Psi}_X \boldsymbol{\epsilon}_{2q} + m\mathbf{1}$$

using the mean-adjusted values, and

$$\hat{\phi} = f(Z, \boldsymbol{\alpha}, \boldsymbol{\beta}) = K_{XZ} L_Z^{-1\top} (\boldsymbol{\alpha} - m\mathbf{1}) + \bar{\Psi}_X \boldsymbol{\beta} + m\mathbf{1}$$

using the notation in Eq. (2) where $\boldsymbol{\alpha} = \mathbf{u} + m\mathbf{1}$, and $\boldsymbol{\beta} = \boldsymbol{\epsilon}_{2q}$.

APPENDIX E: DETAILED INFERENCE ALGORITHM

1. Overview

The posterior inference requires two prerequisites with no interdependence: (i) fixing a gate scale factor G_{SF} and (ii) optimizing the inducing points. The gate voltages are multiplied by the gate scale factor. The gate scale factor is optimized by maximizing the likelihood of the observations with assuming the disorder is perfectly flat. Optimized scale factor values range from 3.48 to 3.94 for runs of the inference algorithm on different thermal cycles.

2. Inducing points optimization

Before obtaining any measurements, the inducing point optimization can be conducted with simulated data. The inducing point optimization is expensive to compute, but the computation time is not critical, because the optimization has to be performed only once for a given gate architecture. Algorithm 1 shows the optimization procedure. Line number 10 in Algorithm 1 is important: It generates a posterior random sample with the information $\{Z, \mathbf{u}\}$. The sample retains only information about ϕ_{d} at Z. For the experiments, we use Eq. (D2) for the posterior sampling, and the approximated distribution (D4) can be used if the computation speed matters.

For the input of Algorithm 1, Eq. (D1) is used for generating Φ^X_{sim}. Each element of $\mathcal{V}_{\mathrm{sim}}$ is generated by choosing a disorder randomly from Φ^X_{sim}, then choosing a pair of voltage vectors near the closed-channel boundary with uniform direction sampling in Ref. [9]. The function $\mathrm{KL}(\mathbf{p}, \mathbf{p}')$ computes $p_i \log(p_i/p_i') + (1-p_i)\log((1-p_i)/(1-p_i'))$ elementwise for $\mathbf{p}$ and $\mathbf{p}'$, then it computes the average of them. For the experiments in the paper, n_d and n_v are set to 20, and the ADAM optimizer is used. The current probability prediction $\mathcal{F}_U^{\mathrm{prob}}$ uses the CNN model $\mathcal{F}_U$ and a sigmoid function:

$$\mathcal{F}_U^{\mathrm{prob}}(\phi_{\mathrm{d}}, \mathbf{v}) = \sigma_\xi(\mathcal{F}_U(g(\phi_{\mathrm{d}}, \mathbf{v})); 10),$$

where $g(\phi_{\mathrm{d}}, \mathbf{v})$ computes ϕ_{in} on the dense grid for CNN (see Sec. III C), and σ_ξ is a modified sigmoid function with steepness parameter and margin $\sigma_\xi(\cdot; 10) = \xi + (1-2\xi)\sigma(\cdot; 10)$. The margin $\xi = 0.01$ allows discrepancy between our approximated model and the real-world measurement, and the steepness parameter set to 10 makes a relatively sharp probability of electric current while allowing the function differentiable. The differentiability is required to use the ADAM optimizer.

Algorithm 1. Inducing point optimization.

Input: Set of randomly generated disorders Φ^X_{sim}, set of randomly generated voltages $\mathcal{V}_{\mathrm{sim}}$, initial inducing points Z_{init}, minibatch size of disorders n_d, minibatch size of voltages n_v, Gaussian process kernel k for disorder, optimizer parameters θ_{opt}, CNN model for current probability prediction $\mathcal{F}_U^{\mathrm{prob}}$
Output: Optimized inducing points Z

```
1:  opt ← Adam optimizer(θ_opt)
2:  Z ← Z_init
3:  while Stopping criterion not satisfied do
4:      Φ^X_mini ← choose random n_d samples from Φ^X_sim
5:      V_mini ← choose random n_v samples from V_sim
6:      loss ← 0
7:      for all φ_d ∈ Φ^X_mini and v ∈ V_mini do
8:          p ← F_U^prob(φ_d, v)
9:          u ← interpolated values of φ_d at Z
10:         φ'_d ← a random sample from the posterior GP with {Z, u}
11:         p' ← F_U^prob(φ'_d, v)
12:         l ← KL(p', p) + KL(p, p')
13:         loss ← loss + l
14:     end for
15:     Z ← opt.update(loss, Z)
16: end while
```

3. MCMC inference

The goal of the MCMC inference is to generate random samples from the posterior distribution of uncertain variables. The uncertain variables in Eq. (D4) are $\mathbf{u}$ and $\boldsymbol{\epsilon}_{2q}$. the posterior probability density function (PDF) of $(\mathbf{u}, \boldsymbol{\epsilon}_{2q})$ given observed current measurement D is

$$p(\mathbf{u}, \boldsymbol{\epsilon}_{2q}|D) \propto \prod_{i=1}^{n} \mathcal{F}_U^{\mathrm{prob}}(f(X, \mathbf{u}, \boldsymbol{\epsilon}_{2q}), \mathbf{v}_i)^{y_i} \times [1 - \mathcal{F}_U^{\mathrm{prob}}(f(X, \mathbf{u}, \boldsymbol{\epsilon}_{2q}), \mathbf{v}_i)]^{1-y_i} p(\mathbf{u}) p(\boldsymbol{\epsilon}_{2q}).$$

The prior distributions $p(\mathbf{u})$ and $p(\boldsymbol{\epsilon}_{2q})$ are defined in Appendix D. A binomial likelihood function is used for the ith measurement where $\mathcal{F}_U^{\mathrm{prob}}(f(X, \mathbf{u}, \boldsymbol{\epsilon}_{2q}), \mathbf{v}_i)$ is the probability of current flowing at voltages $\mathbf{v}_i$. For the experiments in this paper, a Hamiltonian Monte Carlo is used with the posterior PDF. Evaluating the likelihood function with sufficient speed to maintain practical inference times is facilitated by using CNN computation of $\mathcal{F}_U(\cdot)$ on GPU hardware. Each time MCMC inference is performed, a different number of posterior samples are generated. In our work, we find typical values to be $150 < n_{\mathrm{s}} < 320$.

APPENDIX F: INFERENCE RESULTS

The number of inducing points with a low posterior standard deviation increases with the size of the training

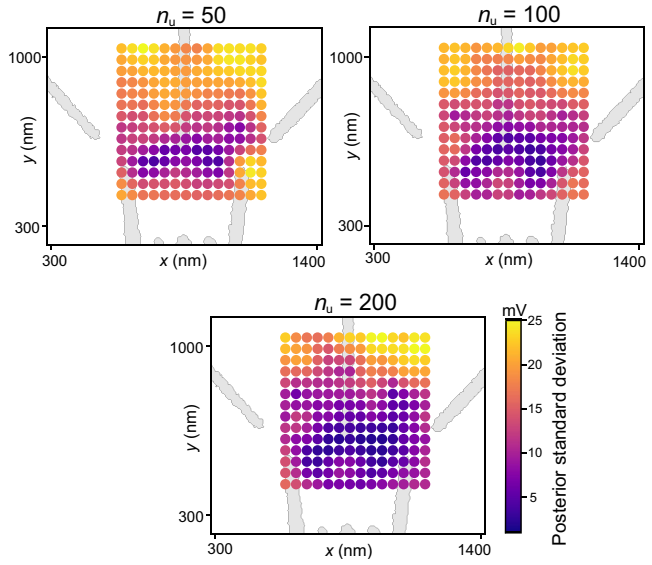

FIG. 7. Disorder inference results using simulated data for three training dataset sizes ($n_u = 50, 100, 200$), with $D = \{(\mathbf{v}_i, y_i)|i = 1, \ldots, 2n_u\}$ as discussed in the main text. The inducing point locations are indicated by circles with the gate structure in the background. The color of each inducing point represents the standard deviation of the posterior disorder potentials at that point.

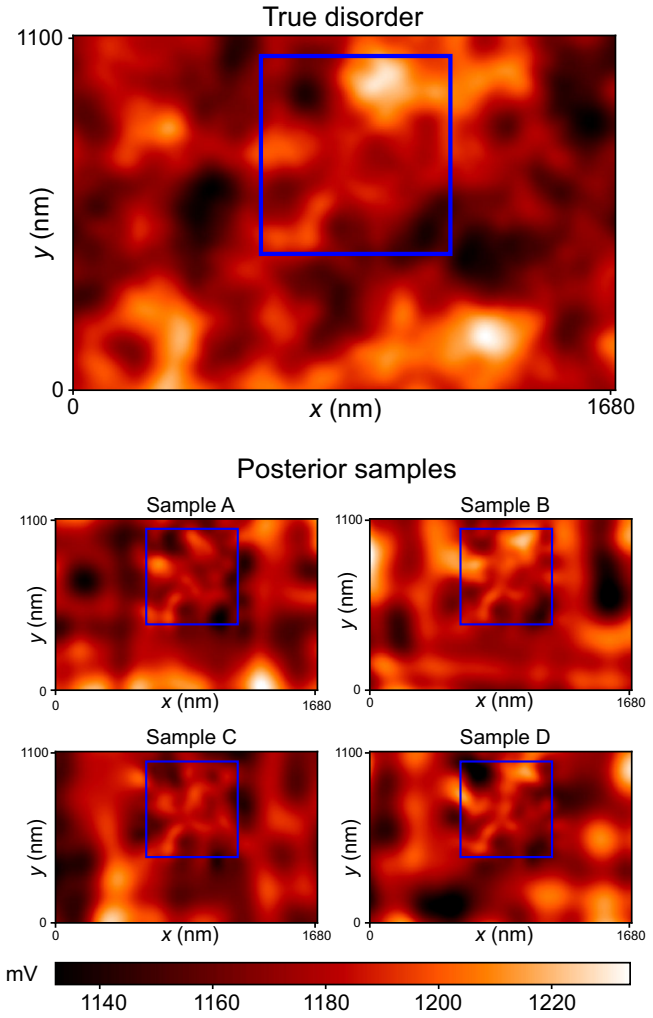

FIG. 8. The true disorder and four randomly selected posterior disorder samples for an iteration of the inference algorithm on a simulated device. The true disorder is generated using the electrostatic model with randomly located donor ions. The blue box on each disorder potential indicates the region spanned by the optimal inducing points. The posterior disorder samples are the same resolution as the true disorder (134×206). The posterior disorders have more detailed features within the box, and qualitative similarities with the true disorder can be observed. All plots share the same scale, as shown at the beneath the posterior samples.

dataset, as shown in Fig. 7. This indicates that the inference algorithm gains information about a larger area of the disorder potential by considering more directions in voltage space. We can also observe that even for a small training dataset, the inference results are most confident about the disorder potential values at the tip of gate G1 which reflects its role in depleting the electron density along the path from source to drain.

Figure 8 shows the true disorder and posterior samples for an iteration of the inference algorithm on a simulated device. The posterior samples exhibit much more detailed features inside the region spanned by the optimized inducing points where qualitative similarities with the true disorder observed. This further demonstrates the information gained at these points (in addition to Fig. 7).

For posterior disorder potentials, features outside the inducing point region are governed by the amplitudes of random Fourier features contained in $\boldsymbol{\beta}$, which are necessary to ensure the posterior samples are continuous and suitable to be used as inputs to $\mathcal{F}_U$ and $\mathcal{F}_D$.

APPENDIX G: DOUBLE-DOT FILTERING

We perform five independent iterations of the inference algorithm on a simulated device, where each iteration uses a different true disorder potential. The output of $\mathcal{F}_D$ with the true disorder is used as the ground truth when determining whether a vector produces a double dot. For each true disorder potential, we generate a dataset of random unit vectors where exactly half of the vectors show double dots (dataset size is either 500 or 1000 vectors depending on the iteration). Using the simulated device, this filtering process allows us to determine the mean false-positive rate (FPR) and mean true-positive rate (TPR) across the five iterations when using different thresholds to determine whether a vector is accepted.

We use the receiver operator characteristic (ROC) curve for posterior, random, and featureless disorders to identify suitable values for the threshold, as shown in Fig. 9(a). A low FPR and high TPR are features of a good classifier. Posterior disorders and random disorders both perform much better than featureless disorders and a random

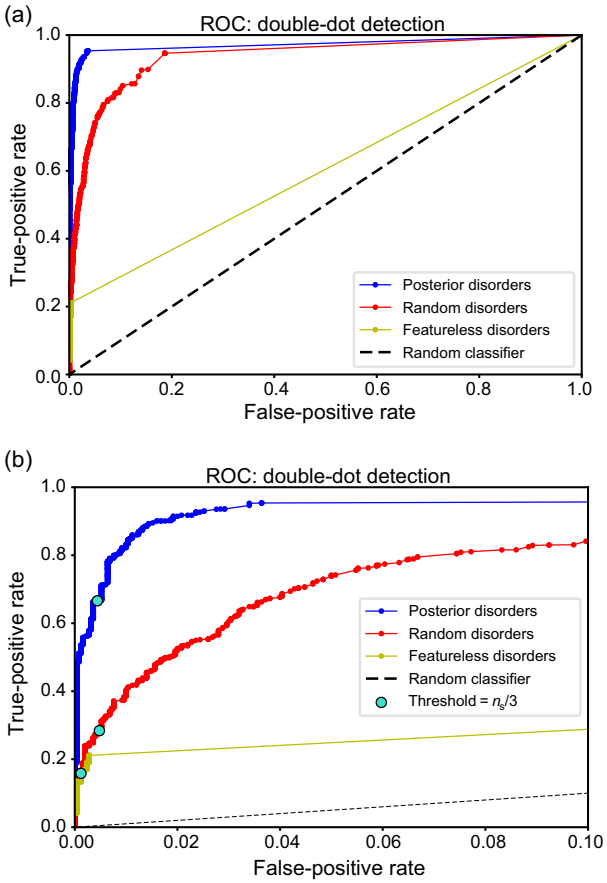

FIG. 9. The ROC curve comparing the FPR and TPR across the domain of possible thresholds $[0, n_s]$ for n_s disorder samples. The mean FPR and TPR from five independent iterations of the inference algorithm and filtering process is shown. (a) The ROC curves for the double-dot filter process using posterior (blue line), random (red line), and flat (yellow line) disorders, along with the ROC curve associated with a random classifier (dashed line). (b) The same ROC curves as shown in (a), but with a truncated false-positive axis. The location corresponding to a threshold of $n_s/3$ as used in the experiment is indicated by a turquoise circle on each curve.

classifier. Posteriors perform better than random disorders. This is further evidence of the success of our inference algorithm. A more detailed comparison in Fig. 9(b) shows that the posterior curve has a much sharper increase in the true-positive rate. We choose a threshold value of $n_s/3$ for the experiment as it allows both posteriors and random disorders to have a moderate true-positive rate, while maintaining a low false-positive rate. A low false-positive rate is desirable because experimental 2D current scans required for human labeling are slow. Other threshold values can be chosen depending on the desired acceptance rate.

For the real device, as discussed in the main text, we generate 5000 random unit vectors which are filtered using posterior and random disorders. The mean number of accepted unit vectors across two iterations of filtering is 38 using posterior disorders, and 26 using random disorders. Featureless disorder potentials are also used to filter vectors for testing on the real device. We use the optimized scale factor from each iteration of the inference algorithm along with 100 featureless disorders, where the constant value is determined by the mean of a random disorder potential. The mean number of accepted unit vectors across the two iterations is 8.5 using featureless disorders, indicating that double dots are not often formed in the model using only the gate potentials. Of the 11 unique vectors which are accepted by featureless disorders across both iterations of filtering, two are labeled as producing quantum dots by human experts. Filtering using featureless disorders fails to find the top scoring vectors for double quantum dots, as identified by the posterior samples. One of the accepted vectors is not tested due to experimental difficulties.

[1] J. Tremblay, A. Prakash, D. Acuna, M. Brophy, V. Jampani, C. Anil, T. To, E. Cameracci, S. Boochoon, and S. Birchfield, *Training deep networks with synthetic data: Bridging the reality gap by domain randomization*, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA,* (2018), pp. 1082–10828, 10.1109/CVPRW.2018.00143.

[2] X. B. Peng, M. Andrychowicz, W. Zaremba, and P. Abbeel, *Sim-to-real transfer of robotic control with dynamics randomization*, in *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA)* (IEEE, New York, 2018), pp. 3803–3810.

[3] B. L. Webster and P. Murdin, *Cygnus X-1—a spectroscopic binary with a heavy companion?*, Nature (London) **235**, 37 (1972).

[4] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. A. Khalek, A. A. Abdelalim, R. Aben, B. Abi, M. Abolins, O. AbouZeid *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716**, 1 (2012).

[5] M. Zhao, T. Li, M. Abu Alsheikh, Y. Tian, H. Zhao, A. Torralba, and D. Katabi, *Through-wall human pose estimation using radio signals*, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA,* (2018), pp. 7356–7365, 10.1109/CVPR.2018.00768.

[6] J. R. Petta, A. C. Johnson, J. M. Taylor, E. A. Laird, A. Yacoby, M. D. Lukin, C. M. Marcus, M. P. Hanson, and A. C. Gossard, *Coherent manipulation of coupled electron spins in semiconductor quantum dots*, Science **309**, 2180 (2005).

[7] M. G. Borselli, K. Eng, E. T. Croke, B. M. Maune, B. Huang, R. S. Ross, A. A. Kiselev, P. W. Deelman, I. Alvarado-Rodriguez, A. E. Schmitz *et al.*, *Pauli spin blockade in undoped* Si/SiGe *two-electron double quantum dots*, Appl. Phys. Lett. **99**, 063109 (2011).

[8] J. Klos, F. Hassler, P. Cerfontaine, H. Bluhm, and L. R. Schreiber, *Calculation of tunnel couplings in open gate-defined disordered quantum dot systems*, Phys. Rev. B **98**, 155320 (2018).

[9] H. Moon, D. T. Lennon, J. Kirkpatrick, N. M. van Esbroeck, L. C. Camenzind, L. Yu, F. Vigneau, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne *et al.*, *Machine learning enables completely automatic tuning of a quantum device faster than human experts*, Nat. Commun. **11**, 4161 (2020).

[10] M. Stopa, *Quantum dot self-consistent electronic structure and the Coulomb blockade*, Phys. Rev. B **54**, 13767 (1996).

[11] J. A. Nixon, J. H. Davies, and H. U. Baranger, *Breakdown of quantized conductance in point contacts calculated using realistic potentials*, Phys. Rev. B **43**, 12638 (1991).

[12] G. J. Percebois and D. Weinmann, *Deep neural networks for inverse problems in mesoscopic physics: Characterization of the disorder configuration from quantum transport properties*, Phys. Rev. B **104**, 075422 (2021).

[13] L. C. Camenzind, L. Yu, P. Stano, J. D. Zimmerman, A. C. Gossard, D. Loss, and D. M. Zumbühl, *Spectroscopy of quantum dot orbitals with in-plane magnetic fields*, Phys. Rev. Lett. **122**, 207701 (2019).

[14] X. G. Croot, S. J. Pauka, M. C. Jarratt, H. Lu, A. C. Gossard, J. D. Watson, G. C. Gardner, S. Fallahi, M. J. Manfra, and D. J. Reilly, *Gate-sensing charge pockets in the semiconductor-qubit environment*, Phys. Rev. Appl. **11**, 064027 (2019).

[15] T. Fujita, T. A. Baart, C. Reichl, W. Wegscheider, and L. M. K. Vandersypen, *Coherent shuttle of electron-spin states*, npj Quantum Inf. **3**, 22 (2017).

[16] A. Mills, D. Zajac, M. Gullans, F. Schupp, T. Hazard, and J. Petta, *Shuttling a single charge across a one-dimensional array of silicon quantum dots*, Nat. Commun. **10**, 1063 (2019).

[17] V. Langrock, J. A. Krzywda, N. Focke, I. Seidler, L. R. Schreiber, and Ł. Cywiński, *Blueprint of a scalable spin qubit shuttle device for coherent mid-range qubit transfer in disordered* $Si/SiGe/SiO_2$, PRX Quantum **4**, 020305 (2023).

[18] K. Takeda, J. Yoneda, T. Otsuka, T. Nakajima, M. Delbecq, G. Allison, Y. Hoshi, N. Usami, K. Itoh, S. Oda *et al.*, *Optimized electrical control of a* Si/SiGe *spin qubit in the presence of an induced frequency shift*, npj Quantum Inf. **4**, 1 (2018).

[19] S. G. Philips, M. T. Mądzik, S. V. Amitonov, S. L. de Snoo, M. Russ, N. Kalhor, C. Volk, W. I. Lawrie, D. Brousse, L. Tryputen *et al.*, *Universal control of a six-qubit quantum processor in silicon*, Nature (London) **609**, 919 (2022).

[20] D. Jirovec, P. M. Mutter, A. Hofmann, A. Crippa, M. Rychetsky, D. L. Craig, J. Kukucka, F. Martins, A. Ballabio, N. Ares *et al.*, *Dynamics of hole singlet-triplet qubits with large g-factor differences*, Phys. Rev. Lett. **128**, 126803 (2022).

[21] M. Sunnåker, A. G. Busetto, E. Numminen, J. Corander, M. Foll, and C. Dessimoz, *Approximate Bayesian computation*, PLoS Comput. Biol. **9**, e1002803 (2013).

[22] M. A. Beaumont, *Approximate Bayesian computation*, Annu. Rev. Stat. Appl. **6**, 379 (2019).

[23] S. Barber, J. Voss, and M. Webster, *The rate of convergence for approximate Bayesian computation*, Electron. J. Stat. **9**, 80 (2015).

[24] M. Järvenpää, M. U. Gutmann, A. Pleska, A. Vehtari, and P. Marttinen, *Efficient acquisition rules for model-based approximate Bayesian computation*, Bayesian Anal. **14**, 595 (2019).

[25] J. Jiang and J. A. Fan, *Global optimization of dielectric metasurfaces using a physics-driven neural network*, Nano Lett. **19**, 5366 (2019).

[26] Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, *Machine-learning-assisted metasurface design for high-efficiency thermal emitter optimization*, Appl. Phys. Rev. **7**, 021407 (2020).

[27] M. Titsias, *Variational learning of inducing variables in sparse Gaussian processes*, in *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, AISTATS 2009, Clearwater Beach, Florida, USA,* (2009), pp. 567–574.

[28] T. D. Bui, J. Yan, and R. E. Turner, *Streaming sparse Gaussian process approximations*, in *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA* (Curran Associates Inc., Red Hook, NY, USA, 2017), pp. 1533–1541.

[29] M. Bauer, M. van der Wilk, and C. E. Rasmussen, *Understanding probabilistic sparse Gaussian process approximations*, in *Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS'16)* (Curran Associates, Inc., Red Hook, NY, USA, 2016), pp. 1533–1541.

[30] Z. Li, J.-F. Ton, D. Oglic, and D. Sejdinovic, *Towards a unified analysis of random Fourier features*, in *Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, 2019*, edited by K. Chaudhuri and R. Salakhutdinov (2019), Vol. 97, pp. 3905–3914.

[31] H. Avron, M. Kapralov, C. Musco, C. Musco, A. Velingker, and A. Zandieh, *Random Fourier features for kernel ridge regression: Approximation bounds and statistical guarantees*, in *Proceedings of the 34th International Conference on Machine Learning, International Convention Centre, Sydney, Australia, 2017*, edited by D. Precup and Y. W. Teh (2017), Vol. 70, pp. 253–262.

[32] J. Hensman, N. Durrande, and A. Solin, *Variational Fourier features for Gaussian processes*, J. Mach. Learn. Res. **18**, 1 (2018).

[33] B. Sriperumbudur and Z. Szabo, *Optimal rates for random Fourier features*, in *Advances in Neural Information Processing Systems 28*, edited by C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett (MIT Press,, Cambridge MA, USA, 2015), pp. 1144–1152.

[34] K. Choromanski, M. Rowland, T. Sarlos, V. Sindhwani, R. Turner, and A. Weller, *The geometry of random features*, in *Proceedings of the International Conference on Artificial Intelligence and Statistics, Playa Blanca, Lanzarote, Canary Islands, 2018*, edited by A. Storkey and F. Perez-Cruz (2018), Vol. 84, pp. 1–9.

[35] A. Rahimi and B. Recht, *Random features for large-scale kernel machines*, in *Neural Information Processing Systems* (Curran Associates Inc., Red Hook, NY, 2007).

[36] T. Palm, *Effects of remote impurity scattering including donor correlations in a branching electron waveguide*, Phys. Rev. B **52**, 11284 (1995).

[37] L. C. Camenzind, L. Yu, P. Stano, J. D. Zimmerman, A. C. Gossard, D. Loss, and D. M. Zumbühl, *Hyperfine-phonon*

spin relaxation in a single-electron GaAs *quantum dot*, Nat. Commun. **9**, 3454 (2018).

[38] J. H. Davies, I. A. Larkin, and E. Sukhorukov, *Modeling the patterned two-dimensional electron gas: Electrostatics*, J. Appl. Phys. **77**, 4504 (1995).

[39] J. A. Nixon and J. H. Davies, *Potential fluctuations in heterostructure devices*, Phys. Rev. B **41**, 7929 (1990).

[40] J. Davies, *Electronic states in narrow semiconducting wires near threshold*, Semicond. Sci. Technol. **3**, 995 (1988).

[41] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, MA, 2009).

[42] K. Ryczko, D. A. Strubbe, and I. Tamblyn, *Deep learning and density-functional theory*, Phys. Rev. A **100**, 022512 (2019).

[43] K. Mills, M. Spanner, and I. Tamblyn, *Deep learning and the Schrödinger equation*, Phys. Rev. A **96**, 042113 (2017).

[44] K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778, https://doi.ieeecomputersociety.org/10.1109/CVPR.2016.90.

[45] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, *Convolutional LSTM network: A machine learning approach for precipitation nowcasting*, Adv. Neural Inf. Process. Syst. **28**, 802 (2015).

[46] H.-C. R. Euler, M. N. Boon, J. T. Wildeboer, B. van de Ven, T. Chen, H. Broersma, P. A. Bobbert, and W. G. van der Wiel, *A deep-learning approach to realizing functionality in nanoelectronic devices*, Nat. Nanotechnol. **15**, 992 (2020).

[47] E. Flurin, L. S. Martin, S. Hacohen-Gourgy, and I. Siddiqi, *Using a recurrent neural network to reconstruct quantum dynamics of a superconducting qubit from physical observations*, Phys. Rev. X **10**, 011006 (2020).

[48] R. M. Neal *et al.*, *Handbook of Markov chain Monte Carlo* (2011), Vol. 2, p. 2, 10.1201/b10905.

[49] J. V. Dillon, I. Langmore, D. Tran, E. Brevdo, S. Vasudevan, D. Moore, B. Patton, A. Alemi, M. Hoffman, and R. A. Saurous, *TensorFlow distributions*, arXiv:1711.10604.

[50] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevX.14.011001 for additional data regarding machine learning models, data labels, and disorder samples.

[51] F. Fedele, A. Chatterjee, S. Fallahi, G. C. Gardner, M. J. Manfra, and F. Kuemmeth, *Simultaneous operations in a two-dimensional array of singlet-triplet qubits*, PRX Quantum **2**, 040306 (2021).

[52] F. Borsoi, N. W. Hendrickx, V. John, M. Meyer, S. Motz, F. van Riggelen, A. Sammak, S. L. de Snoo, G. Scappucci, and M. Veldhorst, *Shared control of a 16 semiconductor quantum dot crossbar array*, Nat. Nanotechnol., 1 (2023). 10.1038/s41565-023-01491-3

[53] H. Bohuslavskyi, A. Ronzani, J. Hätinen, A. Rantala, A. Shchepetov, P. Koppinen, M. Prunnila, and J. S. Lehtinen, *Scalable on-chip multiplexing of low-noise silicon electron and hole quantum dots*, arXiv:2208.12131.