Letters

Cybersecurity in Power Electronics Using Minimal Data – A Physics-Informed Spline Learning Approach

V. S. Bharath Kurukuru, *Member, IEEE*, Mohammed Ali Khan, *Member, IEEE*, and Subham Sahoo, *Member, IEEE*

***Abstract*—Cyberattacks can be strategically counterfeited to replicate grid faults, thereby manipulating the protection system and leading to accidental disconnection of grid-tied converters. To prevent such setbacks, we propose a physics-informed spline learning approach-based anomaly diagnosis mechanism to distinguish between both events using minimal data for the first time in the realm of power electronics. This methodology not only provides compelling accuracy with limited data, but also reduces the training and computational resources significantly. We validate its effectiveness and accuracy under experimental conditions to conclude how data availability problem can be handled.**

***Index Terms*—Anomaly diagnosis, artificial intelligence, cyberattacks, photovoltaic inverters.**

I. Introduction

BASED on the attack disruption resources and model information, cyberattacks on power electronic converters can be deliberately designed to be replicated as grid faults. In this case, the attacker's objective is to maloperate the protection system decision, thereby causing unnecessary converter outage. In [1], a design framework of emulating cyberattacks into faults using the game theory and generative adversarial networks (GANs) was thoroughly discussed. It has further been concluded that a considerably high accuracy of 99.4% can be achieved for emulation of cyberattack in a grid-tied Photovoltaic (PV) system as an asymmetrical fault with limited data using GANs. In addition, the generation of this cyberattack took approximately around 17 min with moderate computational resources. Considering hijacking of the vulnerable attack points in a grid-tied PV system in Fig. 1(a), the mathematical description of the system state might be unclear and is in critical need for observational data. In this condition, there can be many unexplored system dynamics, as the attack can be emulated through any vulnerable points in the system, as shown in Fig. 1(c). This makes it difficult to derive the governing equations as the system transient stability state is found to exhibit discontinuities also during the attack [2]. Furthermore, using historic line–line (LL) fault data in Fig. 1(b), it can be seen in Fig. 1(d) that the generated cyberattack replicates the fault accurately. This problem, usually addressed by fully data-driven discriminators to distill the

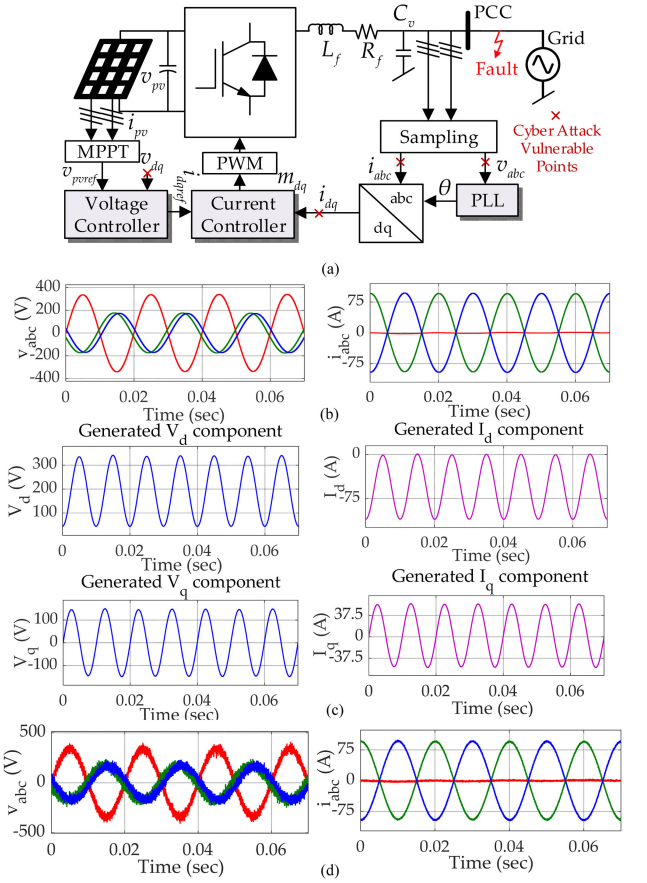

Fig. 1. (a) Single-line diagram of a grid-tied PV system with cyberattack vulnerable points. (b) Response during actual LL fault. (c) Generated cyberattack components using GANs [1]. (d) Response during the cyberattack generated using GANs.

Manuscript received 16 April 2022; revised 12 May 2022 and 1 June 2022; accepted 3 June 2022. Date of publication 8 June 2022; date of current version 26 July 2022. *(Corresponding author: Subham Sahoo.)*

V. S. Bharath Kurukuru is with the Department of Electrical Engineering, Jamia Millia Islamia University, New Delhi 110025, India (e-mail: kvsb272@gmail.com).

Mohammed Ali Khan is with the Department of Electrical Power Engineering, Brno University of Technology, 61600 Brno, Czech Republic (e-mail: khan@vut.cz).

Subham Sahoo is with the Department of Energy, Aalborg University, 9220 Aalborg, Denmark (e-mail: sssa@et.aau.dk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPEL.2022.3180943.

Digital Object Identifier 10.1109/TPEL.2022.3180943

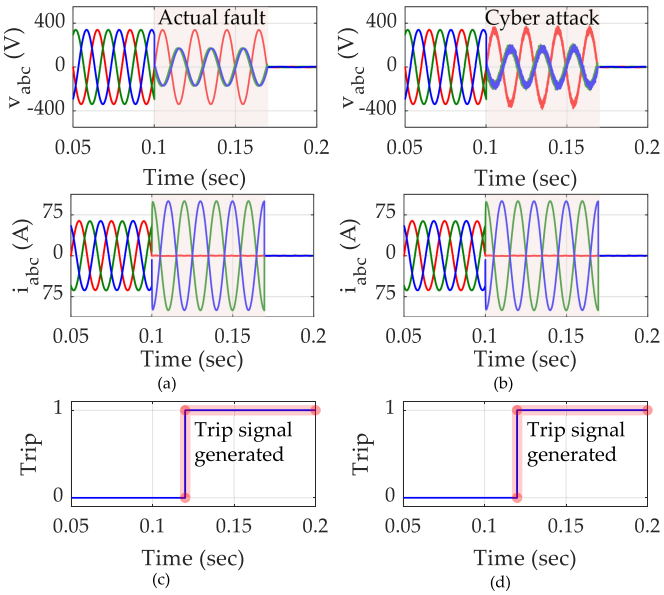

Fig. 2. (a) Voltage and current during actual LL fault. (b) Voltage and current during generated cyberattack using GANs. (c) Trip signal for circuit breaker under actual fault. (d) Trip signal for circuit breaker under cyberattack.

underlying dynamics [3]–[5], still remains a big challenge due to the necessary requirements of high computational resources and observational data. Moreover, considering the data-privacy restraints, distilling the analytical equations from scarce data, commonly seen in practice, adds to this intractable challenge [6]. Classical observers also fail to isolate such anomalies [7], [8], and the protection system settings are unnecessarily triggered. To demonstrate this, a cyberattack fabricated as an LL fault is injected into the vulnerable points in Fig. 1(a). The voltage and current measured under the actual grid fault is shown in Fig. 2(a), and the impact of the disturbance created by cyberattack is shown in Fig. 2(b). From the results, it is identified that, similar to the trip signal generated for an actual fault [see Fig. 2(c)], a trip signal is triggered for the cyberattack modeled as a fabricated fault in Fig. 2(d).

Hence, we propose a physics-informed spline learning (PiSL) approach, which fuses physics and data-derived dynamics to infer the local approximations of a differentiable surrogate model. To do so, we use B-splines [9] to interpolate a discrimination policy in evaluating faults and intelligent cyberattacks in a grid-tied PV system. Hence, for the first time in the realm of power electronics, we realize the collaborative performance of splines and discovered equations to solve the cybersecurity problem with a considerable accuracy using minimal data. As a result, the computational power and dimensionality of data are significantly reduced as compared to the existing solutions.

II. Physics-Informed Spline Learning

This section introduces PiSL with respect to the dynamics of the converter and its control in a grid connected system. By using a generalized model of the considered system [10] in Fig. 1(a), we get

$$\dot{\mathbf{X}}_{\text{sys}} = \mathbf{A}_{\text{sys}}\mathbf{X}_{\text{sys}} + \mathbf{B}_{\text{sys}}\left[V_{\text{dc}_{\text{ref}}} I_{q_{\text{ref}}}\right]^{\text{T}} \tag{1}$$

where $\mathbf{X}_{\text{sys}} = [\mathbf{X}_c\ \mathbf{X}_{\text{pll}}\ \mathbf{X}_l]^{\text{T}}$ and $V_{\text{dc}_{\text{ref}}}$ and $I_{q_{\text{ref}}}$ denote the reference dc voltage and reactive current command, respectively. Furthermore, $\mathbf{X}_c$, $\mathbf{X}_{\text{pll}}$, and $\mathbf{X}_d$ denote the states of the converter, phase locked loop (PLL), and distribution lines, respectively. Furthermore, considering the measured characteristics at vulnerable points in Fig. 1, it can be clearly argued that hijacked v_{dq} and i_{dq} will influence the system dynamics. Hence, whenever there is cyberattack emulated at any of the vulnerable points, the influence can be seen on the measured outputs of the system. These outputs along with system states are a major source of information for building the splines in the PiSL method. Moreover, before proceeding with spline development, it is necessary to investigate the discontinuities in the system state caused by the transients during the attack condition. Besides, the local bifurcation phenomena may occur in such dynamical systems, and they need to be investigated from both the theoretical and physical perspectives. Hence, the local piecewise dynamic points need to be established from the outputs of the influenced system for providing inferences on distinguishing between actual faults and cyberattacks accurately.

A. B-Splines

B-splines are defined as a combination of several piecewise polynomials of degree $k-1$ with at most C^{k-2} continuity at the breakpoints. These breakpoints at which the joints occur are called *knots*, and a set of nondescending breaking points $t_0 \le t_1 \le \cdots \le t_r$ define a knot sequence or a knot vector $\mathbf{T} = \{t_0, t_1, \ldots, t_r\}$, where r indicates the spline sections for a polynomial of degree k. For an odd degree with $2r$ interpolating conditions, the continuity is forced at the knots. Similarly, for an even degree with $r+1$ interpolating conditions, the continuity is forced at the nodes, and for r interpolating conditions, the continuity is forced at knots.

The resultant vector determines the parameterization of the basis function, and has been widely used for curve-fitting and numerical differentiation of experimental data. For a given knot vector $\mathbf{T}$, the associated B-spline basis functions $N_{i,k}(t)$ can be expressed as

$$N_{i,1} = \begin{cases} 1, \text{ if } t_i \le t < t_{i+1} \\ 0, \text{else} \end{cases} \tag{2}$$

for $k = 1$, and

$$N_{i,k} = \frac{t - t_i}{t_{i+k-1} - t_i} N_{i,k-1}(t) + \frac{t_{i+k} - t}{t_{i+k} - t_{i+1}} N_{i+1,k-1}(t) \tag{3}$$

for $k > 1$ and $i = 0,1,\ldots,n$. In (2), t_i denotes the knots and k denotes the polynomial degree. These representations are usually referred to as the Cox–de Boor recursion formula [9]. In this context, three physics-informed models/functions are formed based on the event: 1) normal operation; 2) grid faults; and 3) cyberattacks. If $k = 0$, these basis functions are all step functions, and the basis function $N_{i,0}(t)$ is 1, if t is in the ith knot

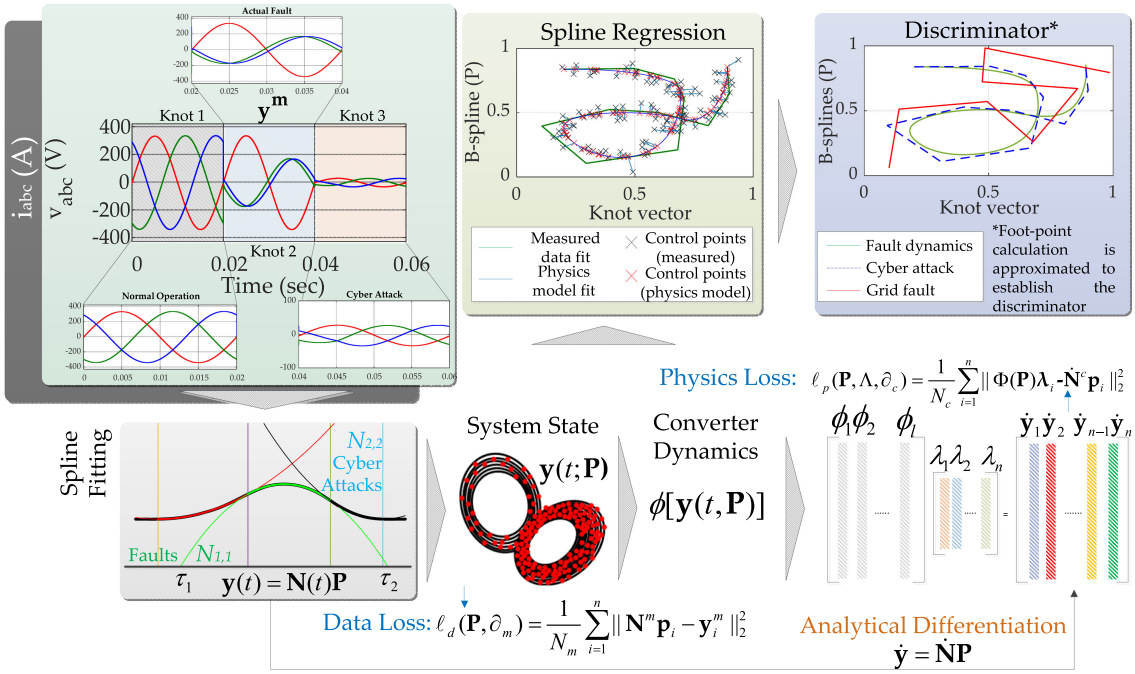

Fig. 3. Schematic architecture of PiSL for discovery of governing equations for the model dynamics to infer between faults and cyberattacks based on scarce data.

span $[t_i, t_{(i+1)})$. Furthermore, the values of the nonzero basis functions are multiplied with an equally spaced control point set $p \in \mathbb{R}^{(r+3)1}$, namely $y(t) = \sum_{i=0}^{r+2} N_{i,3}(t)p_i$, to interpolate the B-splines. Here, the number of control points is empirically chosen according to the frequency of system state such that the computational efficiency can be improved.

B. Architecture Development

Initially, to interpolate n-dimensional system states, n sets of control points are defined for B-splines $\mathbf{P} = \{p_1, p_2, \ldots, p_n\} \in R^{(r+3)n}$ and are multiplied with the spline basis function $N(t)$ to obtain

$$y(t; \mathbf{P}) = \mathbf{N}(t)\mathbf{P}. \tag{4}$$

As shown in Fig. 3, the analytical differentiation can be carried out by differentiating (4). Let $\mathbf{F}(\circ)$ be a function that defines the converter dynamics for different operating states, which are governed by a library of l candidate functions $\Phi(\mathbf{y}) \in \mathbb{R}^{1l}$ [11], given as

$$\Phi = \{1, y, y^2, \ldots, \sin(u), \cos(u)\}. \tag{5}$$

With (4) and its analytical derivatives, the governing equations can thus be given by

$$\dot{\mathbf{y}}(\mathbf{P}) = \Phi(\mathbf{P})\mathbf{\Lambda} \tag{6}$$

where $\Phi(\mathbf{P}) = \Phi(\mathbf{y}(t; \mathbf{P}))$ and $\mathbf{\Lambda} = \{\lambda_1, \lambda_2, \ldots, \lambda_n\} \in \mathbb{R}^{l \times n}$ denotes the coefficient matrix that belongs to constraint subset $\mathbb{S}$. Furthermore, to clarify the discovery problem, the measurement domain m and the measurement data $\partial_m = \{\mathbf{y}_i^m\}_{i=1,\ldots,n} \in \mathbb{R}^{N_m \times n}$ explore the best set of $\mathbf{P}$ and $\mathbf{\Lambda}$, such that (6) holds. Here, the measured response of ith state is $\mathbf{y}_i^m$ and the number of data points in the measurement is denoted by N_m. The loss function to train the PiSL, comprising of the data ℓ_d and physics ℓ_p components, is given by

$$\ell_d(\mathbf{P}, \partial_m) = \sum_{i=1}^{n} \frac{1}{N_m} ||\mathbf{N}^m \mathbf{p}_i - \mathbf{y}_i^m||_2^2 \tag{7}$$

$$\ell_p(\mathbf{P}, \mathbf{\Lambda}, \partial_c) = \sum_{i=1}^{n} \frac{1}{N_c} ||\Phi(\mathbf{P})\boldsymbol{\lambda}_i - \dot{\mathbf{N}}_c \mathbf{p}_i||_2^2 \tag{8}$$

where ∂_c denotes a random set of sampled collocation points (N_c), wherein $N_c \geq 10N_m$ ensures improvement of the physics satisfaction, $\mathbf{N}^m$ defines the basis matrix for splines, and Φ defines the collocation library matrix for the candidate terms. Adhering to all the abovementioned constraints, PiSL training can be analytically formulated as an optimization problem

$$\{\mathbf{P}^*, \mathbf{\Lambda}^*\} = \arg \min_{\{\mathbf{P}, \mathbf{\Lambda}\}} [\ell_d + \alpha \ell_p] \text{ s.t. } \mathbf{\Lambda} \in \mathbb{S} \tag{9}$$

where α is a relative coefficient and the sparsity of Λ is enforced by $\mathbb{S}$. By optimizing (9), we ensure that the splines provide accurate modeling of the system, and its derivatives and candidate function terms to formalize the governing equations.

For the measured voltage v_a during an actual fault in the converter, the piecewise dynamic points are interpolated, as shown in Fig. 4(a). Based on the interpolation data, the possible spline orders are estimated, as shown in Fig. 4(b), to transform the measured variables into basis function, as shown in Fig. 4(c). Furthermore, the continuity of the knots with reference to the measurements in the basis function space is estimated as a spline regression model, as shown in Fig. 4(d). Similarly,

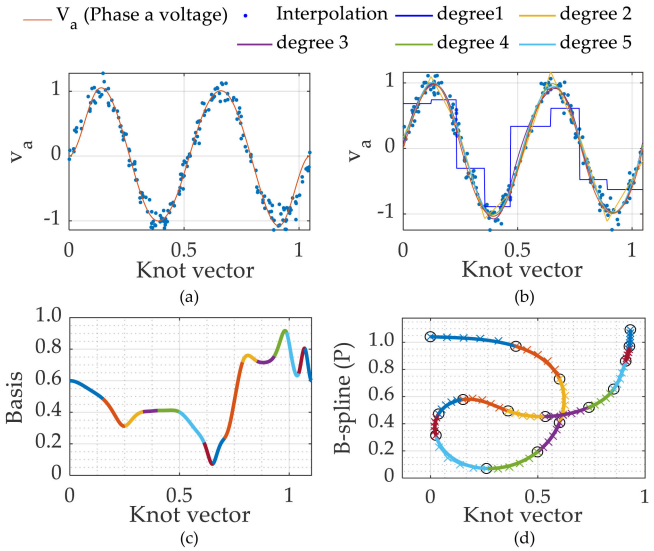

Fig. 4. Continuity of basis for an actual fault. (a) Piecewise dynamic points for interpolation of v_a. (b) Spline orders for the dynamic points. (c) Transformation of the dynamic points in basis function. (d) Continuity with estimated control points.

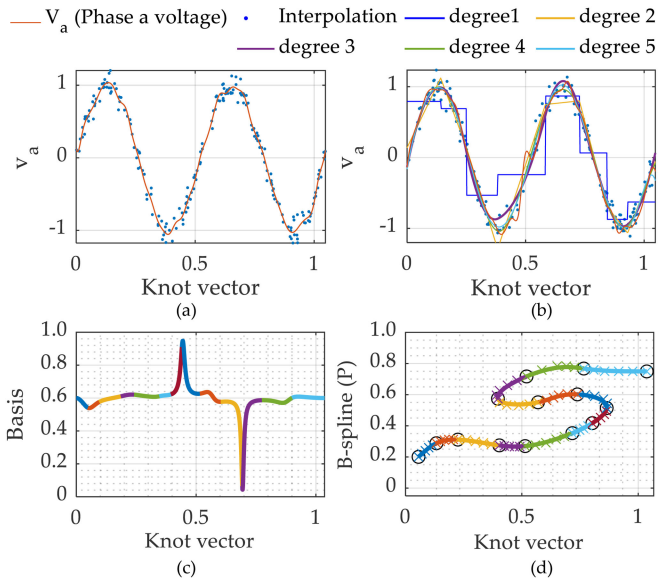

Fig. 5. Continuity of basis for a cyberattack. (a) Piecewise dynamic points for interpolation of v_a. (b) Spline orders for the dynamic points. (c) Transformation of the dynamic points in basis function. (d) Continuity with estimated control points.

for a cyberattack introduced through the vulnerable points in Fig. 1(a), the measured voltage v_a is interpolated, as shown in Fig. 5(a), and the possible spline orders are approximated, as shown in Fig. 5(b). The transformation of interpolated data in basis function is shown in Fig. 5(c), and the continuity between the knots that form a spline regression model is shown in Fig. 5(d). The same approach is followed for all the voltage and current measurements to model the converter dynamics under both actual fault and cyberattack.

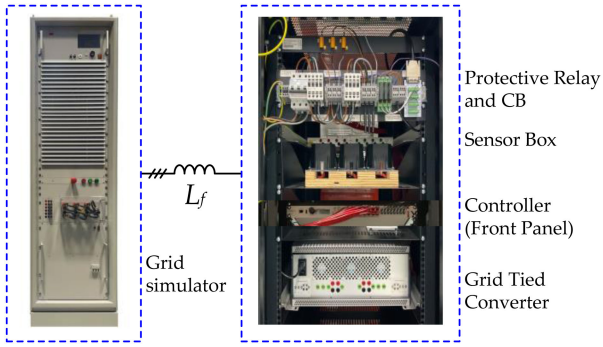

Fig. 6. Experimental setup of the system in Fig. 1(a)—the PiSL network is deployed to identify fault accurately using voltage and current measurements.

III. Results and Discussions

We validate the discrimination accuracy using the proposed PiSL framework on the experimental prototype shown in Fig. 6. We first obtain the voltage v_{abc} and current i_{abc} dataset $\mathbb{D}_{4001\times 6}$ from this setup by sequentially introducing a fault followed by GANs emulated cyberattack [1]. As the measured data have varying scales and the splines do not make any assumption about their distribution, it is normalized between 0 and 1 to identify the distribution using the minmax approach. Then, we employ a weakly physics-informed gradient-based optimization to pretrain the network using $\mathbb{D}$ and the candidate library Φ in (5). We call it "weakly physics-informed," because we have not included (8) into the optimization yet. Furthermore, the knots in the measured data are identified by picking a random set of data points and interpolating them with the results of the full dataset. To perform the interpolation, a nonuniform rational B-spline (NURBS) of degree 3 and order 4 is employed with the measured voltage and currents. Based on the variability in the data, the NURBS function approximates four control points to identify the knots. Finally, we obtain the PiSL tool upon multiple iterations to interpolate the system states for the given knots, such as normal operation, faults, and cyberattacks. This tool is then deployed into the B-Box RCP 3.0 to provide online inferences. The system and control parameters of the setup in Fig. 6 is provided in Appendix.

It can be seen in Fig. 7 that the proposed PiSL operates accurately to track the system dynamics during a fault. To distill its decision, we first segregate the mapping of system dynamics under faults and cyberattacks into two corresponding models. The cyberattack modeling in this letter is carried out using GANs [1], which can accurately emulate grid faults. Once the data are sampled based on the identified knots, the curve-fitting and analytical differentiation is performed to discriminate the data based on the dynamics of the physical model. As the cyberattack abruptly influences the operation of the system, the corresponding measured characteristics have transients in the initial cycle. This causes the initial errors in the fault model and the estimated set-points. Furthermore, this error increases as the cyberattack tries to maximize its impact at the vulnerable points over a specified range. To minimize this error, the estimated set-points can be clipped at the initial stage, but this

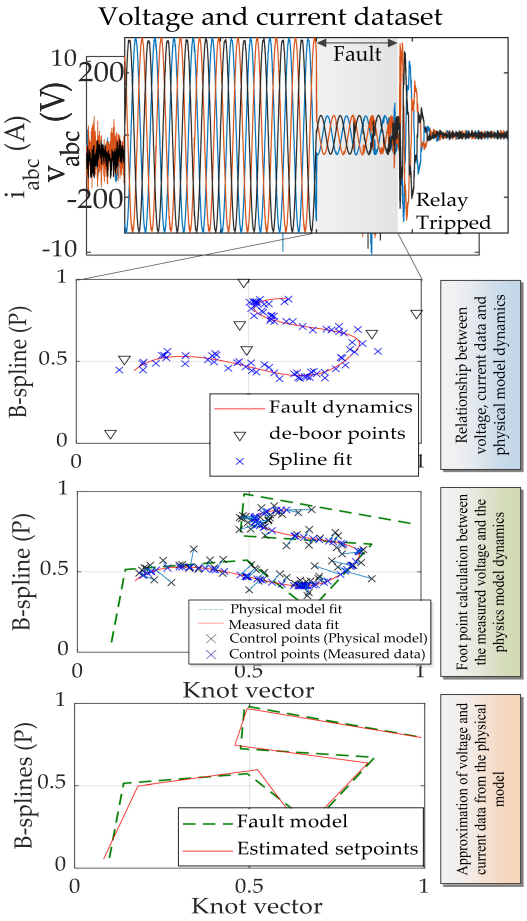

Fig. 7. PiSL operation framework, during an 80% voltage dip (outlined as a fault), it formulates the relationship between $\mathbb{D}$ and the physics-informed model. In the next stage, the footpoint is calculated to evaluate the fitness between the measured data and model dynamics. Finally, an approximation of system states is carried out to infer based on given accuracy levels for fault or cyberattack.

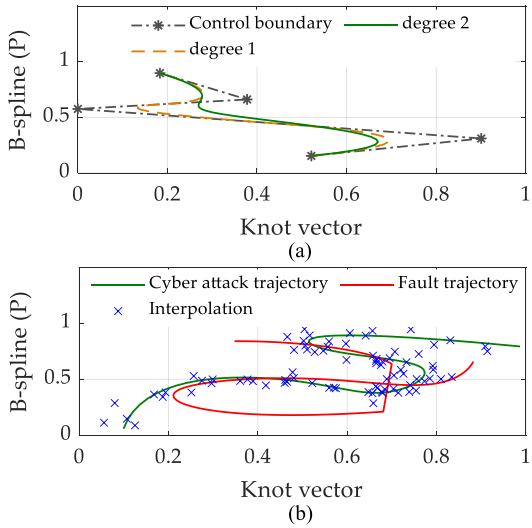

Fig. 8. (a) Periodic boundary and spline orders of degree 2 for a cyberattack. (b) Screenshot of the PiSL operation to discriminate between faults and intelligent cyberattacks.

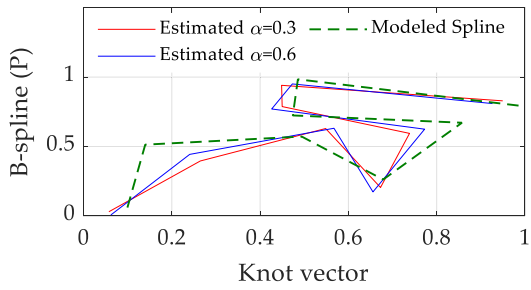

Fig. 9. Varying accuracy of PiSL for different values of α.

overfits the estimated data and results in high inaccuracy during the discrimination process. Hence, the clipping of the data is avoided in this letter. As the fault is confirmed, the decision is routed to trip the relays for ensuring safety. Furthermore, in a practical environment, PiSL will allow real-time monitoring of such events with highest accuracy using minimal data. When a cyberattack is introduced into the vulnerable points in Fig. 6, PiSL is provided with the measured voltage and current, along with the converter dynamics. Here, the piecewise polynomial is used to interpolate the provided inputs through a possible set of spline orders. Generally, for a normal operation of the converter or for a system fault, the interpolation provides a closed spline, which is a combination of several linear spline regions, as shown in Fig. 4. Here, each of these spline regions may be smooth and forms a local bifurcation point, where the curve converges with its previous trajectory. Whereas in the case of a cyberattack, the interpolation provides a spline with open curve and arbitrary smoothness, as shown in Fig. 5. The spline regions in this curve represent a discrete-time dynamical system with discontinuities at the dynamic points. This differentiates the provided input information between an actual fault and a cyberattack.

Besides, during the interpolation process, the control points also provide a way for defining the underlying dynamics to form a spline. To achieve this, the data corresponding to the cyberattack is interpolated using piecewise polynomials. From here, the spline regression selects a series of points to create a fusion of smooth curves that pass through the interpolated data. These curves try to fit the data through different periodic boundaries, which are iteratively derived from various degrees of splines, as shown in Fig. 8(a). Based on the periodic conditions, the endpoints of the spline orders are hinged, and breaks are introduced to return evenly spaced samples over a specified interval in the basis function. These break points will define a control frame from which a spline is derived. To smooth the resultant spline, the break points that form the control frame are adjusted. It can be seen in Fig. 8(b) that the interpolated data form an open curve to correlate with the spline of a cyberattack. Here, the accuracy for a small dataset with irregular patterns around the cyberattack trajectory is around 97.81%, guaranteeing the presence of false data. In this case, PiSL will act as a precursor to activate the equipped cybersecurity tool. Although the dataset is normalized, the relative weighting factor α plays a big role in allocating the search space, which may limit the accuracy. This is evident in Fig. 9, where the accuracy is improved from 89.56% to 91.76%, when α is shifted from 0.3 to 0.6. Hence, tuning α becomes a design tradeoff to improve the accuracy.

TABLE I
ACCURACY LEVELS OF PISL WITH REDUCED DATA

Downsampling factor	Phase offset	Accuracy	Detection time
3	2	96.14%	0.0259 s
7	2	97.44%	0.0254 s
11	2	98.23%	0.0251 s

TABLE II
COMPARATIVE EVALUATION OF THE PROPOSED STRATEGY

Features	[12]	[3]	[4]	This letter
Data requirements	No	Large	Medium	Low
Accuracy	–	98%	91.7%	98.23%
Design time	Low[1]	High	Medium	Low
Comp. burden	Medium[2]	High	High	Medium

[1] As the data requirement is low, the design time intuitively becomes low.
[2] Based on the bounded uncertainty associated with the design process the computation burden increases.

Since the dataset also contains a lot of noise, we inspect PiSL accuracy with respect to downsampled data in $\mathbb{D}$. By downsampling the original data by multiple factors in Table I, it can be seen that the accuracy rather improves with almost no significant change in the anomaly detection time. As a result, PiSL allows higher accuracy with scarce yet qualitative data. In addition, we evaluate its performance in comparison to the existing tools in Table II, which suffices that PiSL is data and computational light without incuring any overheads on its design time and accuracy.

IV. CONCLUSION

This letter proposed a cybersecurity diagnosis approach for grid-connected systems using minimal data by combining physics- and data-based knowledge in reducing the computational and data requirements simultaneously. To the best of our knowledge, this was the first contribution in the realm of power electronics, which used physics-informed machine learning to handle scarce and noisy data. The experimental results not only illustrated its effectiveness in comparison to the existing methods, but also provided apparent insights on handling the data unavailability problem. As a future scope of work, we aim to propose an index to quantify the qualitative features in a given dataset, such that any adversarial data can be eliminated before the training process required for explainability of data-driven cybersecurity tools in power electronics [13].

APPENDIX

An experimental prototype of two-level three-phase grid-tied converter of 7.5 kVA is connected to the grid simulator via an interfacing filter L_f.

1) System: L_f= 1.5 mH, V_n = 230 V/50 Hz, voltage loop gains are $K_{pv} = 0.04$ and $K_{iv} = 168$, and current loop gains are $K_{pi} = 10.5$ and $K_{iv} = 16\,000$.

2) PiSL: $\alpha = 0.9$ and the training dataset $\mathbb{D}_{4001\times 6}$ comprises of v_{abc} and i_{abc} setpoints sampled at a rate of 10 kHz.

REFERENCES

[1] M. A. Khan, V. S. B. Kurukuru, S. Sahoo, and F. Blaabjerg, "From physics to data oriented cyber attack profile emulation in grid connected PV systems," in *Proc. IEEE 22nd Workshop Control Modelling Power Electron.*, 2021, pp. 1–8.

[2] M. Leng, S. Sahoo, F. Blaabjerg, and M. Molinas, "Projections of cyber attacks on stability of DC microgrids–Modeling principles and solution," *IEEE Trans. Power Electron.*, early access, May 16, 2022, doi: 10.1109/TPEL.2022.3175237.

[3] M. Ganjkhani, M. Gilanifar, J. Giraldo, and M. Parvania, "Integrated cyber and physical anomaly location and classification in power distribution systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 7040–7049, Oct. 2021.

[4] Y. Zhang, J. Wang, and B. Chen, "Detecting false data injection attacks in smart grids: A semi-supervised deep learning approach," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 623–634, Jan. 2021.

[5] K. Bhatnagar, S. Sahoo, F. Iov, and F. Blaabjerg, "Physics guided data-driven characterization of anomalies in power electronic systems," in *Proc. 6th IEEE Workshop Electron. Grid*, 2021, pp. 1–6.

[6] M. Taddeo, T. McCutcheon, and L. Floridi, "Trusting artificial intelligence in cybersecurity is a double-edged sword," *Nat. Mach. Intell.*, vol. 1, no. 12, pp. 557–560, 2019.

[7] M. Jamei *et al.*, "Anomaly detection using optimally placed μPMU sensors in distribution grids," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3611–3623, Jul. 2018.

[8] G. Anagnostou, F. Boem, S. Kuenzel, B. C. Pal, and T. Parisini, "Observer-based anomaly detection of synchronous generators for power systems monitoring," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4228–4237, Jul. 2018.

[9] M. G. Cox, "The numerical evaluation of B-splines," *IMA J. Appl. Math.*, vol. 10, no. 2, pp. 134–149, 1972.

[10] V. S. B. Kurukuru *et al.*, "A review on artificial intelligence applications for grid-connected solar photovoltaic systems," *Energies*, vol. 14, no. 15, 2021, Art. no. 4690.

[11] S. Brunton, J. Proctor, and J. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Nat. Acad. Sci.*, vol. 113, no. 15, pp. 3932–3937, 2016.

[12] K. Gupta, S. Sahoo, R. Mohanty, B. K. Panigrahi, and F. Blaabjerg, "Decentralized anomaly characterization certificates in cyber-physical power electronics based power systems," in *Proc. IEEE 22nd Workshop Control Modelling Power Electron.*, 2021, pp. 1–6.

[13] S. Sahoo, H. Wang, and F. Blaabjerg, "On the explainability of black box data-driven controllers for power electronic converters," in *Proc. IEEE Energy Convers. Congr. Expo.*, 2021, pp. 1366–1372.