Riccardo Torchio[1], F Conte[1], M Scarpa[1], M Filippini[1], F Pase[1], F Toso[1], P Shahriari Nasab[1], E Marson[1], A Viroli[1], P Posa[1], and M Zangoli[1]

[1]Affiliation not available

April 18, 2024

Digital Twins in Power Electronics: A Comprehensive Approach to Enhance Virtual Thermal Sensing

R. Torchio, F. Conte, M. Scarpa, M. Filippini, F. Pase, F. Toso, P. Shahriari Nasab, E. Marson, A. Viroli, P. Posa, M. Zangoli

Abstract—**The traditional approaches in material research and hardware design are insufficient to address the evolving Operation and Maintenance (O&M) demands in contemporary power electronics. Overengineering and data acquisition practices lead to unsustainable costs and reduced profit margins. Digital Twins (DTs), defined as real-time simulation models of physical systems, emerge as promising solutions to meet stringent O&M requirements. In power electronics, DTs offer significant potential in thermal management, crucial for control performance, safety, and system lifespan. This paper aims to analyze the development of computationally efficient and high-fidelity DTs tailored for power electronics applications, emphasizing their predictive reliability. To achieve this goal, the proposed physics-based approach is enhanced by integrating Data-Driven Artificial Intelligence (AI)-based techniques. The predictive reliability of the DTs produced through this workflow is then experimentally validated against a power electronic converter designed for induction heating applications. Additionally, the feasibility of real-time execution is demonstrated, affirming the practical applicability of the developed DTs.**

Index Terms—**Digital Twins, Power Converters, Real-Time, Physics-Based, Data-Driven, Artificial Intelligence.**

I. Introduction

TRADITIONAL approaches in material research and hardware design are no longer sufficient to meet the evolving Operation and Maintenance (O&M) requirements of contemporary power electronics products and systems. The prevailing industrial practice involves overengineering components and acquiring extensive data, leading to unsustainable costs and diminished profit margins.

The emergence of Digital Twins (DTs), characterized as mathematical models capable of real-time simulation of a system's physical behavior, stands out as the most promising solution to address the increasingly stringent O&M demands in the power electronics market [1]. The maturation of enabling technologies that unify hardware and software has ushered in a new era, focusing on two key performance indicators (KPIs): enhancing maximum functionality through control performance [2] and enabling more precise and effective predictive maintenance through increased information [3].

Specifically within the domain of power electronics, the most promising area for DT technology lies in thermal management [4]. Temperature, particularly junction temperature (T_j), significantly influences control performance, safety, costs, efficiency, and system lifespan [5], [6]. While manufacturers invest in advanced materials (e.g., SiC, GaN) to maximize power density and minimize losses, a critical challenge remains in the thermal management of these increasingly efficient modules [7]–[9].

Current thermal management relies on sensors like Negative-Temperature-Coefficient (NTC) ones placed away from the die junctions for engineering reasons, resulting in significant latency and imprecise observations of junction temperature dynamics [10], [11]. This complicates the implementation of high-performance thermal management control by system manufacturers, who are forced to adopt oversized safety margins, thereby flattening the added value and competitive advantage of utilizing, for example, SiC Mosfet-based modules over Si Mosfet-based ones.

Another critical reason for the importance of junction temperature is aging models, coveted by power electronics system manufacturers for providing predictive maintenance services, enhancing product safety, and optimizing warranty periods [12]. Currently, the most accurate and reliable aging models are in the hands of semiconductor module manufacturers, performing power cycles (PC) and temperature cycles (TC) to extract empirical and physics of failure aging models [13]. These models estimate the remaining useful life (RUL) of modules, classify major fault types (e.g., solder junction wear, wire lift bonding, etc.), and depend on the junction temperature variation in a cycle (ΔT_j) [13]. Various methods exist for measuring junction temperature, but none are implemented in production for obvious reasons [5]. Despite module suppliers providing valuable aging models, system builders struggle to utilize them due to the inability to estimate junction temperature. Hence, real-time estimation of Tj becomes crucial. While literature explores various real-time junction temperature estimation methods for power modules [14], [15], there lacks a structured and unified methodology poised to become an industry standard, a fundamental piece in constructing a DT

R. Torchio and F. Conte are with the Department of Industrial Engineering of the University of Padova, Padova, 35131, Italy (email: riccardo.torchio@unipd.it, filippo.conte@studenti.unipd.it).

R. Torchio and M. Scarpa are with the Department of Information Engineering of the University of Padova, Padova, 35131, Italy (email: mattia.scarpa@studenti.unipd.it).

M. Filippini, F. Pase, and F. Toso are with Newtwen, Padova, 35127, Italy (email: mattia.filippini@newtwen.com, francesco.pase@newtwen.com, francesco.toso@newtwen.com).

P. Shahriari Nasab, E. Marson, and P. Posa, are with Electrolux Italia, Porcia, 33080, Italy (email: pedram.shahriari-nasab@electrolux.com, enrico.marson@electrolux.com, paolo.posa@electrolux.com).

A. Viroli and M. Zangoli are with Electrolux Italia, Forlì, 47122, Italy (email: alex.viroli@electrolux.com, massimo.zangoli@electrolux.com).

for thermal management and predictive maintenance of power electronics systems.

The current state-of-the-art for junction temperature estimation remains thermal network models [12], with limitations, especially for complex systems. Designing a thermal network model without knowledge of the component geometry, i.e., without reverse engineering the power module, is challenging. Some power module manufacturers provide thermal models for T_j estimation, but coupling them to the entire system and the heat exchange model proves difficult. Thermal networks lose physicality since the heat equation is not well represented by an equivalent lumped circuit. High-resolution thermal networks compromise real-time implementation, rendering them computationally burdensome. Designing them is mostly a manual process, requiring expertise that few engineers possess in companies. Certain simulation software companies (e.g., Ansys, Siemens, Newtwen) are investing in tools capable of reducing complex models like finite element models (FEM) into Reduced Order Models (ROM) [16]. These ROMs can be implemented on third-party hardware platforms for real-time execution, striking a balance between accuracy and computational complexity [17].

A DT transcends a real-time model estimating junction temperature. It must adapt to changes such as material aging, different operating conditions, manufacturing errors, exchange data bidirectionally with the system through sensors and control actions, predict scenarios, optimize control actions, identify anomalies and hazardous conditions, and generate a wealth of information to enhance future design and predictive maintenance models [18]. In this paper, we intend to analyze the entire workflow for constructing a digital twin capable of providing the functionalities necessary to improve control performance and predictive maintenance capabilities—KPIs essential to meeting the current O&M requirements in the new generation of power electronics systems.

II. Digital Twin Life-cycle for Power Electronics Components

In this section, a thorough discussion is conducted on the principal constituents of a DT, encompassing its structure, implementation methodologies, and operational paradigms. Initially, the DT is founded solely upon the physics information about the constituent entity, comprising geometric specifications, material characteristics, and mathematical models delineating its physics dynamics. Subsequent phases necessitate the compression of this informational reservoir to allow real-time execution of the model on a designated microcontroller. To this end, the judicious application of Model Order Reduction (MOR) techniques is imperative, ensuring an optimal equilibrium between model fidelity and computational exigency, particularly concerning memory allocation and computational complexity inherent to the microcontroller environment [19].

Following the reduction process, incorporating a stochastic element rooted in empirical data becomes imperative to reduce unavoidable discrepancies between the physics model and real-world phenomena. These discrepancies typically stem from uncertainties in parameters, dynamic and time-evolving boundary conditions, and approximations introduced by the reduction strategies. Given the stochastic nature of electronic component production processes, the integration of data-driven models alongside physics formulations emerges as a robust strategy, affording a hybrid model architecture that maximizes the generalization and accuracy of power electronic device representations [20], [21]. Upon deployment within a microcontroller framework, the resultant hybrid model furnishes real-time insights of both quantitative and qualitative nature, informing control strategies with critical information such as the identification of temperature hot spots, often situated within inaccessible points like die junctions, thermal exchange dynamics with cooling system, and temporal temperature prognostications. This real-time information is thereby a powerful feedback to be used by thermal management processes, transcending the limitations so far due to lack of adequate sensing. Finally, the expansive repository of real-time insights thus unlocked can be seamlessly disseminated across an entire network of interconnected systems, facilitated by cloud infrastructure integration. This integration engenders enhanced analytical capabilities, particularly concerning degradation and aging metrics, thereby expediting the acquisition of a dependable predictive maintenance model, which is a requisite in contemporary market exigencies aiming to curtail operational costs and minimize downtime.

A. High Fidelity Models

The model of a Power Converter module has an intrinsically multi-physics nature since, in general, electric, thermal, and fluid dynamic effects must be considered to define the overall behavior of the device. With the final objective of generating an embeddable and thus computationally cheap DT of the device for the real-time monitoring of critical quantities (e.g., temperature), dedicated modeling strategies must be used to consider these three physics and the coupling between them.

1) Electric (Loss) Model: The Electric Loss model of the power converter relies on either datasheet specifications or experimental measurements. The methodology, as explicated in [22], delineates a systematic approach to asses both switching and conduction losses.

Conduction power losses are computed by directly multiplying the collector current (I_C) by the corresponding voltage (U_{CE}) from the datasheet, thereby determining $P_{V,COND}$ depending on the current. Furthermore, the method's advantage lies in accurately approximating the loss function with second order polynomial fitting.

Junction temperature dependency becomes paramount in estimating losses dynamically varying with the component's temperature and, so, temperature-dependent coefficients can be incorporated into the polynomials:

$$P_{V,COND}(I_C, T_j) = c \cdot I_C + d \cdot I_C^2, \tag{1}$$

where c and d are, in case of a 2nd order polynomial fitting:

- $c(T_j) = c_0 + c_1 \cdot T_j + c_2 \cdot T_j^2$
- $d(T_j) = d_0 + d_1 \cdot T_j + d_2 \cdot T_j^2$

with c_i and d_i to be defined, see [22].

The accuracy of this approximation increases with a greater number of recorded operating temperatures, allowing for a higher-order approximation.

Switching losses in power electronics are dependent on variables such as current, junction temperature, operating voltage, and switching frequency. An effective evaluation procedure begins by extracting the total switching energy ($E_{tot} = E_{on} + E_{off}$) from referenced datasheets, and subsequently, the power loss expression is derived from this data. Then, these losses are systematically correlated with the respective variables, employing a method similar to that utilized for conduction losses. [22]

2) Thermal Model: The thermal model of a Power Converter must be capable of providing the dynamic evolution of the temperature in critical points of interest, e.g., the junction temperature. The thermal model is described by the following well-known advection–diffusion equation, i.e.,

$$\rho c_p \frac{\partial T}{\partial t} + \rho c_p \mathbf{v} \cdot \nabla T - \nabla \cdot k \nabla T = q, \tag{2}$$

where ρ is the density, c_p is the heat capacity at constant pressure, T is the temperature, k is the thermal conductivity, q is the power density, and $\mathbf{v}$ is the velocity field (which is not zero only in the fluid region). Power losses, i.e., q, are obtained from the Electric (Loss) Model described in Section II-A1 and the velocity field $\mathbf{v}$ of the coolant (if any) is provided by the Fluid Dynamic Model described in Section II-A3.

In (2), the dependence w.r.t. the position has been omitted for simplicity. Equation (2) is then complemented by boundary conditions valid on the border of the model ($\partial\Omega$), e.g., Dirichlet, Neumann, or, more frequently used, convective condition, i.e.,

$$\mathbf{n} \cdot k \nabla T = h(T_{ext} - T), \tag{3}$$

where $\mathbf{n}$ is the unit normal vector of the boundary of the motor, h is the convective coefficient, and T_{ext} is the external/ambient temperature. Depending on the case, radiation boundary conditions can be included too. However, considering them makes the problem non-linear, and this is in general avoided. The interested reader can refer to [23] for more details.

To generate a numeric dynamic model of (2) (including boundary conditions), Finite Element Method (FEM) is the most widely used approach. Thus, a computational model of the Power Converter is generated and a mesh is constructed. The discretized model can be finally written as [24]

$$\mathbf{M}\frac{d\mathbf{x}}{dt} + (\mathbf{K} + \mathbf{A} + \mathbf{H})\mathbf{x} = \mathbf{Q_p p} + \mathbf{Q_c} T_{ext}, \tag{4}$$

where $\mathbf{M}$ is the mass matrix, while $\mathbf{K}$, $\mathbf{A}$, and $\mathbf{H}$ are the stiffness matrices related to conductive, advective, and convection terms, respectively. $\mathbf{p}$ is the power loss array of dimension N_p storing the losses (in [W]) for each domain (see Section II-A1), $\mathbf{Q_p}$ is the $N \times N_p$ matrix which maps $\mathbf{p}$ into the rhs of the thermal model, and $\mathbf{Q_c}$ is the array mapping the external temperature T_{ext} into the rhs of the thermal model related to the convective boundary condition.

When the device is liquid-cooled and therefore the advection term is included, it is well known that advection-dominated computational models such as the one in (4) are particularly challenging from the numerical point of view: even fine meshes lead to Peclet number $Pe > 1$, which results in large node to node oscillations. To remove such oscillations, standard stabilization techniques can be adopted (e.g., based on Streamline Upwind Petrov Galerkin (SUPG) [25]). Alternatively, one can eliminate the advective term from the thermal model and replace it with equivalent boundary conditions, specifically convective boundary conditions featuring a substantially high convective coefficient [26]. While this approach streamlines the computational complexity of the model, it concurrently compromises the model's physics accuracy.

Finally, (4) is discretized in time by applying, e.g., a backward Euler scheme, and it is written in state space form , i.e.,

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_{k-1} \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k \end{aligned}, \tag{5}$$

where k indicates the time-step related to the $k\Delta t$ instant, Δt is the time step, and $\mathbf{y}$ is the vector storing the temperature of interest. It is worth noting that more advanced time-stepping techniques may be applied to discretize (4). However, advanced time-stepping techniques may not be compatible with the final on-chip implementation of the DT. The backward Euler scheme is instead simple enough to be implemented in a standard microprocessor and, by choosing a small enough value of Δt, a good level of accuracy can be guaranteed.

3) Fluid Dynamic Model: Often, Power Converters for high-power applications such as automotive ones have active cooling systems to dissipate the heat generated, e.g., based on forced fluid or air flows. Fluids allow to reach higher power densities but generally require a more expensive and complex system (pumps, filters, radiators), thus, when possible, the forced-air solution is preferred. A computational fluid dynamics (CFD) simulation is required to study the velocity and pressure distribution in the fluid domain. Time domain simulations at different flow rates generally can be carried out under the following assumptions:

- in-compressible fluid flow: this simplification is true for fluids and could be adopted also for gases when there are mild pressure changes and temperature variations;
- turbulent flow: $k - \omega$ Reynolds-averaged (RANS) turbulence model;
- wall functions with quadrangular fluid boundary mesh;
- P1 + P1 discretization of velocity and pressure;
- Streamline + crosswind diffusion numerical stabilization;

For instance, the $k - \omega$ formulation based on turbulent kinetic energy k and specific dissipation rate ω can be used:

$$\begin{cases} \rho \frac{\partial k}{\partial t} + \rho(\mathbf{u} \cdot \nabla k) = P_k - \rho\beta^* k\omega + \nabla \cdot (\mu\sigma^* \mu_T \nabla k) \\ \rho \frac{\partial \omega}{\partial t} + \rho(\mathbf{u} \cdot \nabla \omega) = \alpha \frac{\omega}{k} P_k - \rho\beta\omega^2 + \nabla \cdot (\mu\sigma\mu_T \nabla \omega) \end{cases} \tag{6}$$

For the full definition of symbols, the reader is referred to [27]. It is worth mentioning that such simulations can result in high computational effort since the formulation is nonlinear and fine meshes are needed to achieve convergence. Because of these complexities, a common simplified approach is to entirely avoid the CFD simulation by substituting the coolant/wall heat exchange with an equivalent condition as previously mentioned [26].

B. Model Order Reduction

The main property that differentiates a DT from a high-fidelity model is the possibility of implementing the DT in-cloud or on-the-edge hardware for the real-time (or more than real-time) execution of the model, allowing a mutual exchange of information between the physical asset (i.e., the Power Converter) and the corresponding DT. Since, in this paper, DTs are aimed at the real-time monitoring of critical quantities, in-cloud implementations alone may not be a reliable solution, due to unavoidable communication delays. Fortunately, the recent advancements in microprocessor technology pave the way for on-chip DTs, where the digital replicas are directly embedded in the onboard available hardware.

Obviously, due to their large dimension, the high-fidelity models described in the previous section are not directly compatible with the on-chip implementation. To solve this problem, MOR techniques can be used. Indeed, MOR is an enabling technology well known in literature and ripe for the industrial ecosystem with new commercial tools.

While the Electric (Loss) model is already compatible with the on-chip implementation, for the real-time monitoring of critical quantities such as the junction temperature, the discretized state-space thermal model, i.e., (7), must be solved in real-time. A thermal model of a realistic Power Converter module resulting from FEM discretization may have thousands or even millions of unknowns. Thus, its dimensionality must be reduced to allow on-chip implementation. To do that, MOR strategies based, e.g., on Balanced Truncation, Moment Matching, or Proper Orthogonal Decomposition can be used. The interested reader is referred to, e.g., [28] for more details about different MOR strategies, which can be applied to both continuous or discrete models. Regardless of the adopted technique, MOR allows for projecting the original Full Order Model (FOM) (7) into a reduced order space, i.e.,

$$\begin{aligned} \hat{\mathbf{x}}_k &= \hat{\mathbf{A}}\hat{\mathbf{x}}_{k-1} + \hat{\mathbf{B}}\mathbf{u}_{k-1} \\ \mathbf{y}_k &= \hat{\mathbf{C}}\hat{\mathbf{x}}_k \end{aligned}, \tag{7}$$

where $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, and $\hat{\mathbf{C}}$ have been obtained by projecting the corresponding FOM matrices into the reduced order space, while $\hat{\mathbf{x}}$ is the reduced order state, i.e., $\mathbf{x} \approx \mathbf{V}\hat{\mathbf{x}}$, where $\mathbf{V}$ is the projection basis function constructed by the adopted MOR strategy.

Concerning the fluid dynamic model, it is worth noting that the computational effort for this kind of simulation is high, which poses challenges to obtaining a reduced CFD model that can be computed in real-time. Fortunately, in industrial applications, the flow rate is kept constant, or it varies in a prescribed limited range. Thus, the velocity field $\mathbf{v}$ (which is used for the advection term of the thermal model) can be evaluated offline for a set of prescribed conditions and the thermal model can be parameterized to consider different cooling conditions. Of course, this may introduce an unavoidable approximation but allows for avoiding solving in real-time the CFD problem, which may be unfeasible for on-chip implementation. It is worth noting that the literature about MOR for CFD problems is vast and constantly growing [29]. However, due to the complex nature of the CFD problems, incorporating fluid-dynamics

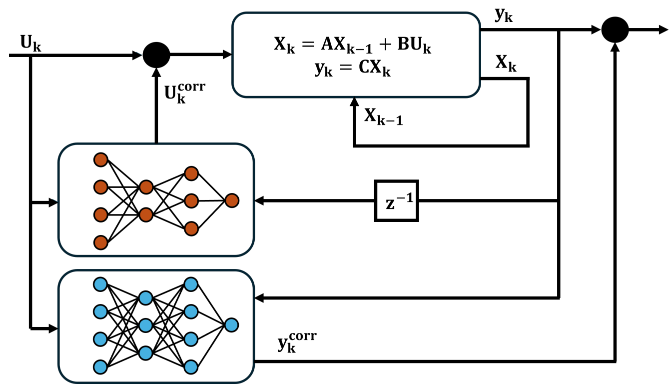

Fig. 1: Hybrid Model Architecture.

ROMs in standard microprocessors for real-time solutions is still a challenge.

C. From Model as Designed to Model as Manufactored

This section elucidates a pivotal aspect of the research, pivotal in clarifying the essence of the digital twin concept within power electronics. Initially, a model, no matter how complex it is, remains an approximation of reality, encountering several challenges in representing the complete dynamics of power electronic systems [30]. These challenges include geometric approximations, limitations of numerical methods in solving partial differential equations governing thermal and electric phenomena, uncertainties in material properties, and the approximations introduced by MOR techniques. Moreover, manufacturing processes introduce unique characteristics into each electronic component, further complicating model fidelity. Additionally, aging, wear, and operational conditions make material parameters time-varying, posing additional modeling challenges. To address these complexities, a comprehensive methodology is proposed, adopting physics-based and data-driven approaches. This hybrid model architecture aims to enhance accuracy and robustness in monitoring and controlling power electronic systems, particularly when integrated within the control and management units.

The proposed hybrid model architecture integrates a reduced physics-based model with two Feed Forward Neural Networks (FFNNs). The first FFNN serves to correct uncertainties in model inputs, while the second FFNN corrects the model's output, effectively mitigating errors in the physics-based model. Notably, the architecture is calibrated based on a substantially reduced experimental dataset, leveraging intrinsic information from the physics-based model to streamline training and reduce computational complexity suitable for real-time integration on a microcontroller or embedded platform.

Training FFNNs involves optimizing model parameters and employing techniques such as gradient-based optimization, regularization methods, and dropout to prevent overfitting. Widely-used frameworks such as PyTorch and TensorFlow provide a robust ecosystem for FFNN development, offering flexibility, extensive support, and efficient computation. Careful selection and application of optimization algorithms, regularization techniques, and appropriate libraries are essential for

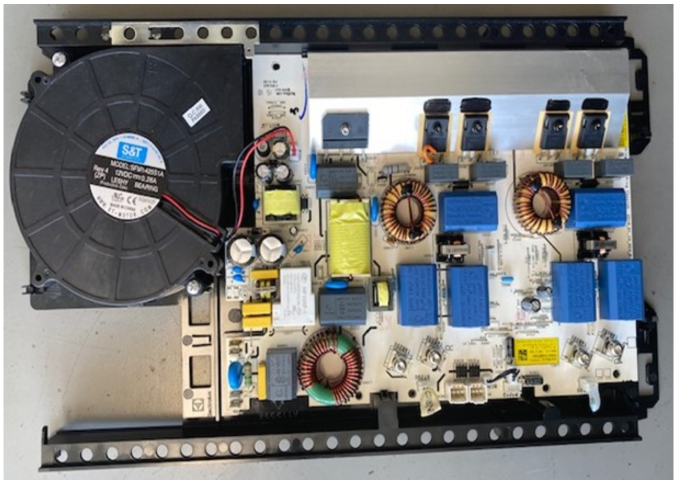

Fig. 2: Test Bench.

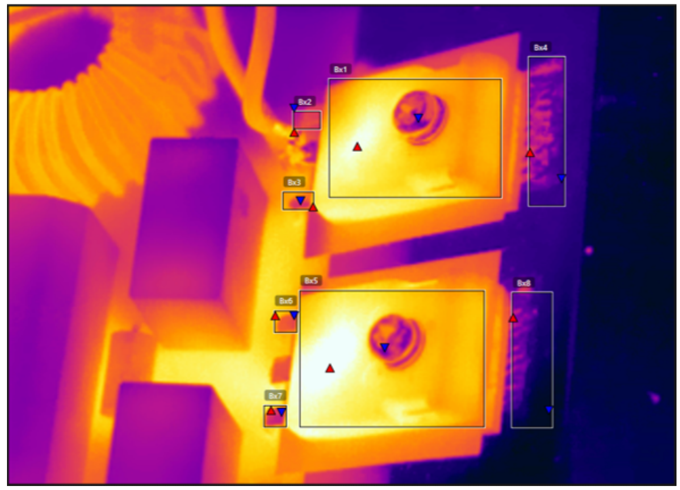

Fig. 3: IR Camera.

effective training and calibration of FFNNs tailored to power electronic applications.

Ensuring the generalization capability of the final model architecture is paramount for robust performance in real-world power electronic applications. Integrating physics-based models with FFNN architectures enhances interpretability and promotes better generalization to unseen data by incorporating domain knowledge and fundamental principles. Maintaining physical constraints within the model architecture prevents overfitting and increases reliability, enhancing confidence in the model's performance across diverse operating scenarios and environmental conditions.

D. On Chip Implementation

The last step, which is crucial to defining the digital twin as such, involves real-time implementation on a microcontroller, specifically on the hardware platform controlling the actual converter. Once this step is achieved, the digital twin and its corresponding real counterpart have the ability to exchange data and information in a bidirectional flow through sensor readings and control actions. Therefore, it is important to ensure the following functionalities: synchronization of feedback and control actions with the integration time step of the digital twin, which partly consists of a state-space system to be integrated over time; stability properties of the final digital twin architecture, e.g., checking the eigenvalues of the state matrix $\hat{\mathbf{A}}$; and finally, numerical conditioning of the model matrices to avoid truncation and rounding errors when implementing the model in a fixed-point 32-bit architecture, for example.

III. Case Study: Power Converter for Induction Heating Applications

A. Test Case Description

In this section, we describe the experiments conducted to test and validate the effectiveness of the DT in representing the physical behavior of a power electronic converter for induction heating applications in the home appliance sector. The converter consists of a diode rectifier and two single-phase half-bridge inverters connected to different coils. To verify accuracy, the adopted measurement system utilized an infrared (IR) thermal camera on the open device, as depicted in Fig.2. Specifically, the IR camera was used to observe temperature hotspots on the IGBT cases, output pins, and heatsink near the soldering points, providing a granular temperature map around the point of interest, i.e., the junction temperature which cannot be directly measured Fig.3.

Today, particularly given the shortage issues faced by industries, it is crucial to have multiple semiconductor component suppliers. However, this increases the variability of device performance, making monitoring and control increasingly challenging. In this study, three different suppliers of discrete power modules were adopted. However, validating and certifying firmware with three different DT systems, each tailored to a different semiconductor module supplier, proves to be cumbersome. Therefore, developing a single integrated model that accurately and reliably represents a power electronic system potentially employing components from various suppliers in serial production poses a technological challenge. In the following, it is described how this is accounted for.

B. Applied Approach

For the specific application, a nominal conduction and switching loss model, which averages the behaviors of different datasheets of the components, has been designed as described in the previous section. The outputs of the loss model constitute part of the inputs of the thermal model, specifically the heat sources. The thermal model comprises a 3D finite element thermal model that solves the heat equation neglecting the radiation component, making MOR techniques more effective.

To reduce the finite element model, commercial software produced by Newtwen® has been adopted, implementing MOR techniques, as mentioned in the previous section, optimized for finite element matrices, which are typically large, sparse, and numerically ill-conditioned.

The initial high-fidelity model comprises approximately 10^5 Degrees of Freedom (DoF), see Fig.4, while the final reduced order model (ROM) (obtained by using the Moment Matching

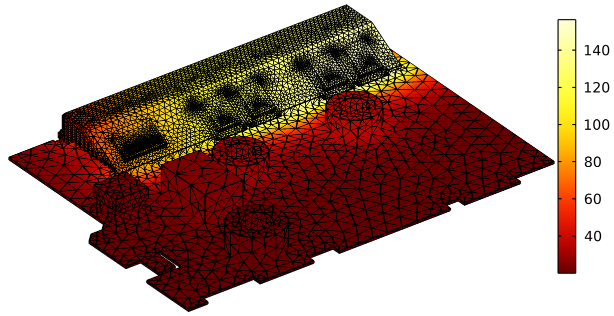

Fig. 4: Full Order Model. Temperature distribution in °C.

technique [31], [32] with a convergence tolerance of 10^{-3} on the relative residual of the rhs of the problem) comprises only 20 DoF and is capable of describing the temperature at each node of the full order model (FOM) mesh with a maximum error of 2.5 Kelvin.

However, a single thermal model is not sufficient to define a DT capable of accurately estimating the behavior of the real device, which can incorporate different components and operate under various load and boundary conditions, including varying cooling and environmental conditions over time. To address this issue, the physical model has been augmented with a data-driven model consisting of two feedforward neural networks (FFNNs) with distinct functions. As depicted in Fig.1, the first neural network takes as input the outputs of the loss model and the thermal dynamics of temperature estimated by the thermal model at over 10 nodes of the mesh. Its output is the correction of the input vector for the thermal model, aiming to mitigate the errors of the loss model and uncertainties of boundary conditions, such as cooling fan speed and external temperature for convective heat exchange. In addition, the second neural network serves to correct the final estimates of the thermal model-FFNN1 architecture, mitigating model errors stemming from material parameter uncertainties such as thermal capacity and conductivity, as well as variance due to multiple suppliers of power modules. Therefore, FFNN2 acts as the data-driven discrepancy model that enhances the generalization of the DT, maximizing its capability to represent the system under study and analysis. This hybrid model architecture allows for training the neural networks on a reduced dataset and, importantly, designing them with a limited number of layers and neurons, making them suitable for real-time implementation on microcontrollers.

C. Model Accuracy and Computational Effort

The calibration dataset was generated from six different tests at various current levels and load conditions for each type of discrete power module supplier. In Figs. 5 - 10, one can observe the results of the calibration. The loss function used is the Euclidean norm of the error between measurement and estimation at each moment of acquisition during the heating transient at three different geometric points corresponding to the temperature hot spots on the IGBT cases and the positioning of NTC sensors in the system. This allows for appropriately

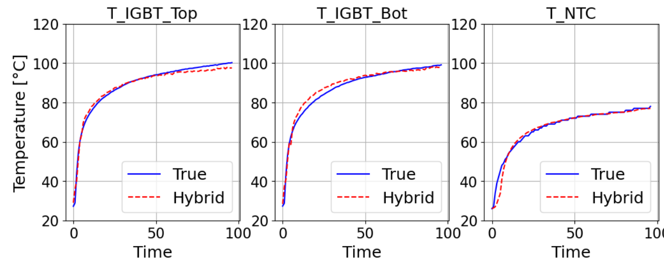

Fig. 5: Calibration Test: IGBT supplier 1, current-frequncy operational load 1.

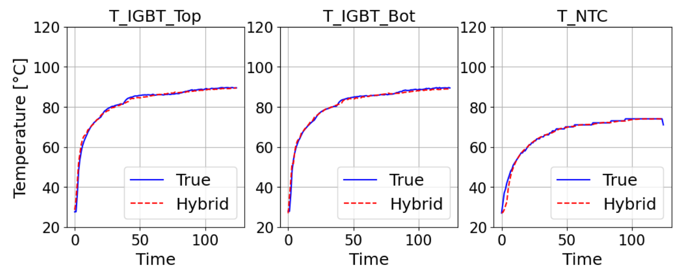

Fig. 6: Calibration Test: IGBT supplier 1, current-frequency operational load 2.

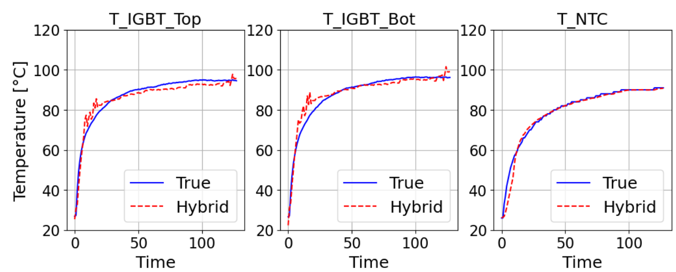

Fig. 7: Calibration Test: IGBT supplier 1, current-frequency operational load 3.

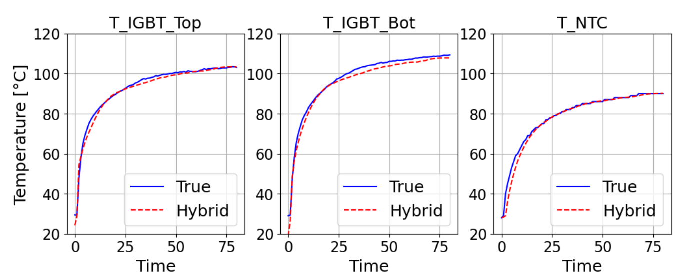

Fig. 8: Calibration Test: IGBT supplier 2, current-frequency operational load 1.

modeling the thermal gradient in the area of interest to make the estimation of junction temperature as reliable as possible, which is engineering-wise impossible to measure in the case of discrete power modules.

Maintaining a certain degree of generalization in the model architecture is essential to prevent overfitting and increase reliability. Overfitting occurs when the model learns to memorize training data rather than capturing underlying patterns,

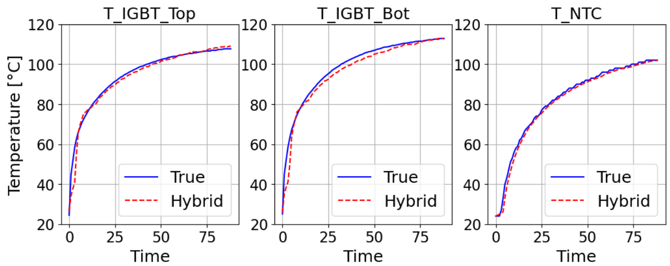

Fig. 9: Calibration Test: IGBT supplier 3, current-frequency operational load 1.

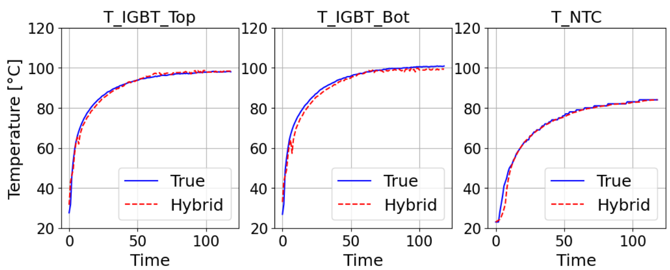

Fig. 10: Calibration Test: IGBT supplier 3, current-frequency operational load 2.

leading to poor generalization. By incorporating physics-based constraints, the model is less likely to extrapolate erroneously and more capable of making accurate predictions in diverse operating scenarios. To evaluate the generalization capability of the final model architecture, it is crucial to test its performance in operating scenarios never seen during the training phase. This ensures that the model can effectively extrapolate beyond the training data and provides confidence in its reliability for real-world applications. For this purpose, in Fig. 11 and Fig. 11 one can observe the two validation tests at different operating conditions that have been used to check the generality of the approach, which is crucial for the reliability of the final model architecture to be implemented in the production firmware. The total of the FFNNs parameters is 237 adopting Leaky Rectified Linear Unit (ReLU) activation functions, suitable for embedded implementation, while the reduced state space model is described by the following matrices: $\mathbf{A} = [20X20]$, $\mathbf{B} = [20X5]$, $\mathbf{C} = [10X20]$. Further, the entire model architecture is executed on an STM32-based evaluation board with an execution time of about 100 μs (65% of the overall time is required for the computation of the ROM and the remaining 35% for the data-driven part, i.e., FFNNs) and 5 kB memory footprint in total.

Figs. 5-12 show the high accuracy of the developed physics-based data-driven augmented DT w.r.t. the measurements collected from three power converters, each one equipped with one of the three discrete power modules. Moreover, it is worth noting that, because using three different IGBTs, the temperature measured in the three power converters is very different, discrepancies of about 20 K can be spotted by comparing results of, e.g., Fig. 5, Fig. 8, and Fig. 9. However, for all of these conditions, the physics-based data-driven augmented DT is in very good agreement with measurements.

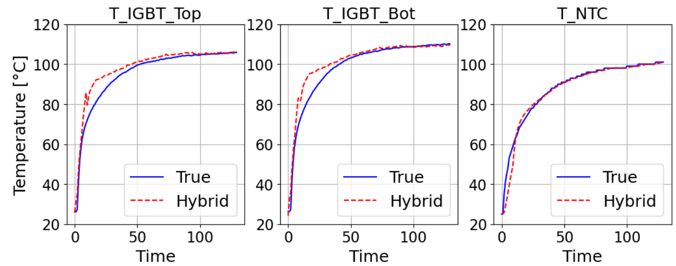

Fig. 11: Validation Test: IGBT supplier 1, current-frequency operational load 4.

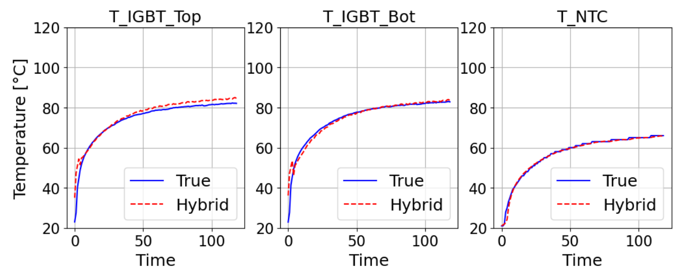

Fig. 12: Validation Test: IGBT supplier 3, current-frequency operational load 3.

D. Discussion and Added Value

The real-time feasibility, accuracy, and level of generalization achieved thus allow for the implementation of a virtual temperature sensor system in production, which can be utilized for the following strategies:

- Enhancing power derating performance by finely modulating switching frequency and current to maximize the product's state of function, i.e., increasing the functional burden of components by reducing safety margins through continuous monitoring of junction temperature;
- Increasing the quantity and quality of relevant information through virtual sensors, e.g., estimates in inaccessible points, spatial temperature gradients, cross-play of data between virtual and real sensors, to develop aging and degradation models and identify reliable patterns with reduced costs and time.

Moreover, exploiting in real-time a computationally efficient and high-fidelity DT allows for unlocking advanced control strategies and predictive maintenance as shortly discussed in the following.

1) Thermal Management Control: Real-time monitoring of junction temperatures in semiconductor power modules offers a paradigm shift in power electronics control strategies by minimizing reliance on traditional safety margins. Conventionally, control strategies incorporate wide safety margins to accommodate uncertainties, relying on measurements from Negative Temperature Coefficient (NTC) sensors, which often lack precision and dynamic responsiveness regarding junction temperatures. By contrast, real-time junction temperature monitoring provides accurate and dynamic insights into the thermal behavior of critical components. This enables precise control of derating mechanisms, allowing for proactive adjustments in operating parameters such as switching frequency and maximum currents. With reliable real-time junction temperature estimation, control loops can be implemented to dynamically optimize system performance while ensuring safe thermal operation. By reducing reliance on conservative safety margins and leveraging real-time temperature data, power electronic systems can operate more efficiently and reliably. This approach not only maximizes performance but also minimizes downtime and maintenance costs associated with overheating issues. Thus, real-time junction temperature monitoring represents a pivotal advancement in power electronics control strategies, enabling high efficiency, reliability, and performance optimization [33].

2) Predictive Maintenance: The physical modeling of failures involves analyzing and modeling actual failure mechanisms, such as semiconductor aging under thermal cycling. Typically, calibrating physics-based lifetime models requires only a few power cycling tests. Consequently, the Physics-of-Failure (PoF) approach is emerging as a novel methodology that enhances lifetime estimation and allows for the integration of reliability engineering into the development and research cycles of the overall design process [34]. This approach has the potential to significantly enhance the understanding of failure processes and contribute to the overall robustness of power electronics converters. It is crucial to note that both empirical and PoF aging and degradation models are intricately linked to the junction temperature swing—the critical state variable that profoundly affects aging phenomena such as solder degradation and wire bond lifting [13]. The integration of DTs directly into the products themselves, facilitating bidirectional data exchange through sensors and control actions, represents a significant advancement in predictive maintenance strategies [34]. Real-time monitoring of the junction temperature swing introduces a pivotal element by enabling the utilization of specific aging models, thereby augmenting the quality and quantity of data access. This enhanced data granularity and accuracy empower ML/AI algorithms to discern intricate degradation patterns and anticipate impending failures more precisely. Consequently, this technological synergy fosters the development of more robust and effective predictive maintenance strategies, elevating the reliability and longevity of power electronics converters.

IV. Conclusions

In this paper, a comprehensive approach to constructing highly accurate and computationally efficient Digital Twins (DTs) of power electronics applications for the real-time monitoring of critical temperature has been proposed. Physics-based models are the starting point of the proposed work-flow, that are then reduced by using Model Order Reduction techniques to make the DT compatible with real-time execution on microprocessors. Finally, the real-time DT model is augmented by using Data-Driven Artificial Intelligence (AI)-based technique to improve its predictive reliability.

The effectiveness of the approach is verified using real-world power electronic converters intended for induction heating home appliance applications. These converters incorporate IGBTs sourced from various suppliers. Due to industry-wide supply shortages, it is common for manufacturers to utilize components from different suppliers, resulting in potential variability between otherwise equivalent products. Consequently, model-based monitoring becomes more complex. Nevertheless, the physics-based AI-augmented DT developed through the proposed approach exhibits excellent predictive reliability even in such realistic scenarios. This underscores the maturity and practical applicability of the proposed methodology in addressing challenges encountered in industrial settings.

References

[1] H. Chen, Z. Zhang, P. Karamanakos, and J. Rodriguez, “Digital twin techniques for power electronics-based energy conversion systems: A survey of concepts, application scenarios, future challenges, and trends,” *IEEE Industrial Electronics Magazine*, vol. 17, no. 2, pp. 20–36, 2023.

[2] P. Karamanakos, E. Liegmann, T. Geyer, and R. Kennel, “Model predictive control of power electronic systems: Methods, results, and challenges,” *IEEE Open Journal of Industry Applications*, vol. 1, pp. 95–114, 2020.

[3] S. Zhao and H. Wang, “Enabling data-driven condition monitoring of power electronic systems with artificial intelligence: Concepts, tools, and developments,” *IEEE Power Electronics Magazine*, vol. 8, no. 1, pp. 18–27, 2021.

[4] B. Rodriguez, E. Sanjurjo, M. Tranchero, C. Romano, and F. Gonzalez, “Thermal parameter and state estimation for digital twins of e-powertrain components,” *IEEE Access*, vol. 9, pp. 97 384–97 400, 2021.

[5] N. Baker, M. Liserre, L. Dupont, and Y. Avenas, “Improved reliability of power modules: A review of online junction temperature measurement methods,” *IEEE Industrial Electronics Magazine*, vol. 8, no. 3, pp. 17–27, 2014.

[6] M. Andresen, K. Ma, G. Buticchi, J. Falck, F. Blaabjerg, and M. Liserre, “Junction temperature control for more reliable power electronics,” *IEEE Transactions on Power Electronics*, vol. 33, no. 1, pp. 765–776, 2018.

[7] S. Lohrasbi, R. Hammer, W. Essl, G. Reiss, S. Defregger, and W. Sanz, “A comprehensive review on the core thermal management improvement concepts in power electronics,” *IEEE Access*, vol. 8, pp. 166 880–166 906, 2020.

[8] S. Jones-Jackson, R. Rodriguez, Y. Yang, L. Lopera, and A. Emadi, “Overview of current thermal management of automotive power electronics for traction purposes and future directions,” *IEEE Transactions on Transportation Electrification*, vol. 8, no. 2, pp. 2412–2428, 2022.

[9] M. Andresen and M. Liserre, “Impact of active thermal management on power electronics design,” *Microelectronics Reliability*, vol. 54, no. 9, pp. 1935–1939, 2014, sI: ESREF 2014.

[10] Y. Zhou, W. Shi, J. Tang, X. Wang, W. Li, X. He, C. Zhang, and Z. Li, “Dynamic junction temperature estimation via built-in negative thermal coefficient (ntc) thermistor in high power igbt modules,” in *2017 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2017, pp. 772–775.

[11] S. Pyo and K. Sheng, “Junction temperature dynamics of power mosfet and sic diode,” in *2009 IEEE 6th International Power Electronics and Motion Control Conference*, 2009, pp. 269–273.

[12] H. Chen, B. Ji, V. Pickert, and W. Cao, “Real-time temperature estimation for power mosfets considering thermal aging effects,” *IEEE Transactions on Device and Materials Reliability*, vol. 14, no. 1, pp. 220–228, 2014.

[13] I. F. Kovačević, U. Drofenik, and J. W. Kolar, “New physical model for lifetime estimation of power modules,” in *The 2010 International Power Electronics Conference - ECCE ASIA -*, 2010, pp. 2106–2114.

[14] F. Stella, G. Pellegrino, E. Armando, and D. Daprà, “Online junction temperature estimation of sic power mosfets through on-state voltage mapping,” *IEEE Transactions on Industry Applications*, vol. 54, no. 4, pp. 3453–3462, 2018.

[15] M. H. M. Sathik, J. Pou, S. Prasanth, V. Muthu, R. Simanjorang, and A. K. Gupta, “Comparison of igbt junction temperature measurement and estimation methods-a review,” in *2017 Asian Conference on Energy, Power and Transportation Electrification (ACEPT)*, 2017, pp. 1–8.

[16] D. Hartmann, M. Herz, and U. Wever, “Model order reduction a key technology for digital twins,” *Reduced-Order Modeling (ROM) for Simulation and Optimization: Powerful Algorithms as Key Enablers for Scientific Computing*, pp. 167–179, 2018.

[17] F. Di Napoli, A. Magnani, M. Coppola, P. Guerriero, V. D’Alessandro, L. Codecasa, P. Tricoli, and S. Daliento, “On-line junction temperature monitoring of switching devices with dynamic compact thermal models extracted with model order reduction,” *Energies*, vol. 10, no. 2, 2017.

[18] M. G. Kapteyn, D. J. Knezevic, and K. Willcox, *Toward predictive digital twins via component-based reduced-order models and interpretable machine learning*. AIAA Scitech 2020 Forum, 2020.
[19] D. Hartmann, M. Herz, and U. Wever, *Model Order Reduction a Key Technology for Digital Twins*. Cham: Springer International Publishing, 2018, pp. 167–179.
[20] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019.
[21] M. Tezzele, N. Demo, A. Mola, and G. Rozza, *An integrated data-driven computational pipeline with model order reduction for industrial and applied mathematics*. Cham: Springer International Publishing, 2022, pp. 179–200.
[22] U. Drofenik and J. Kolar, "A general scheme for calculating switching- and conduction-losses of power semiconductors in numerical circuit simulations of power electronic systems," *Power Electronic Systems Laboratory (PES), ETH Zurich ETH-Zentrum, Zurich*, 2005.
[23] M. Zorzetto, R. Torchio, F. Lucchini, S. Maseei, L. Robol, and F. Dughiero, "Reduced order modeling for thermal simulations of electric components with surface-to-surface radiation," *TechRxiv*, 2024.
[24] L. Codecasa, "A novel approach for generating boundary condition independent compact dynamic thermal networks of packages," *IEEE Transactions on Components and Packaging Technologies*, vol. 28, no. 4, pp. 593–604, 2005.
[25] R. Codina, "On stabilized finite element methods for linear systems of convection diffusion reaction equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 188, no. 1, pp. 61–82, 2000.
[26] A. P. Catalano, C. Scognamillo, V. d'Alessandro, and A. Castellazzi, "Numerical simulation and analytical modeling of the thermal behavior of single- and double-sided cooled power modules," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 10, no. 9, pp. 1446–1453, 2020.
[27] D. C. Wilcox, "Formulation of the k-w turbulence model revisited," *AIAA Journal*, vol. 46, no. 11, pp. 2823–2838, 2008.
[28] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. Schilders, and L. M. Silveira, Eds., *Model Order Reduction: Volume 1: System-and Data-Driven Methods and Algorithms*. Berlin, Germany: De Gruyter, 2021.
[29] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza, *Model Order Reduction in Fluid Dynamics: Challenges and Perspectives*. Cham: Springer International Publishing, 2014, pp. 235–273.
[30] M. G. Kapteyn, J. V. R. Pretorius, and K. E. Willcox, "A probabilistic graphical model foundation for enabling predictive digital twins at scale," *Nature Computational Science*, vol. 1, no. 5, pp. 337–347, May 2021.
[31] X. Peng, Q. Bai, X. Xia, Z. Huang, K. Saenko, and B. Wang, "Moment matching for multi-source domain adaptation," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1406–1415.
[32] L. Codecasa, D. D'Amore, and P. Maffezzoni, "Multipoint moment matching reduction from port responses of dynamic thermal networks," *IEEE Transactions on Components and Packaging Technologies*, vol. 28, no. 4, pp. 605–614, 2005.
[33] J. Falck, G. Buticchi, and M. Liserre, "Thermal stress based model predictive control of electric drives," *IEEE Transactions on Industry Applications*, vol. 54, no. 2, pp. 1513–1522, 2018.
[34] J. Huber, L. Imperiali, D. Menzi, F. Musil, and J. W. Kolar, "Energy efficiency is not enough!" *IEEE Power Electronics Magazine*, vol. 11, no. 1, pp. 18–31, 2024.