RESEARCH   **Open Access**

Hybrid modeling: towards the next level of scientific computing in engineering

Stefan Kurz[1]* , Herbert De Gersem[2] , Armin Galetzka[2] , Andreas Klaedtke[1], Melvin Liebsch[3],
Dimitrios Loukrezis[2] , Stephan Russenschuck[3] and Manuel Schmidt[1]

*Correspondence:
stefan.kurz2@de.bosch.com
[1]Robert Bosch GmbH, Renningen,
Germany
Full list of author information is
available at the end of the article

Abstract

The integration of machine learning (Keplerian paradigm) and more general artificial intelligence technologies with physical modeling based on first principles (Newtonian paradigm) will impact scientific computing in engineering in fundamental ways. Such hybrid models combine first principle-based models with data-based models into a joint architecture. This paper will give some background, explain trends and showcase recent achievements from an applied mathematics and industrial perspective. Examples include characterization of superconducting accelerator magnets by blending data with physics, data-driven magnetostatic field simulation without an explicit model of the constitutive law, and Bayesian free-shape optimization of a trace pair with bend on a printed circuit board.

1 Introduction: what is hybrid modeling?

If we take a look at Gartner's 2018 Hype Cycle of Emerging Technologies [43], Deep Learning has been at the *top of inflated expectations* and should now be moving towards the *plateau of productivity*. Indeed, machine learning and more general artificial intelligence technologies recently have spurred a lot of interest in the applied mathematics and industrial communities, see for instance [5, 10, 45], and [21] for an introduction.

The idea of combining physics with data has a long history. Following [10, p. 57], we call modeling based on first principles the *Newtonian paradigm*. Newton's laws of motion provided (within their range of validity) "for the first time a unified quantitative explanation for a wide range of observations" [44]. Conversely, Johannes Kepler[1] started from astronomer Tycho Brahe's and own measurement data, and worked towards a mathematical description to fit the measured data. Following again [10], we call this approach the *Keplerian paradigm*. Both paradigms complement each other. For a simple enough model system, Kepler's laws can be derived from Newton's theory. Conversely, starting from a two-body model system, actual trajectories of celestial bodies can be modeled by Newton's laws plus data-driven terms that correct for perturbations due to effects that are not present in the model.

---

[1]Johannes Kepler was born exactly 450 years ago when this paper was written, namely on 27.12.1571 in Weil der Stadt.

Springer

In modern terms, we call this complementary approach *hybrid modeling.*

Definition *Hybrid models* combine first principle-based models with data-based models into a joint architecture, supporting enhanced model qualities, such as robustness and explainability.

First principles express formalized domain knowledge. For the purpose of this paper the domain knowledge results from physics. But there are other possibilities, such as statistics (e.g., probabilistic graphical models [3, Ch. 8]) or discourse (e.g., ontologies [16]). Data may be obtained from any source, in particular from observation or simulation. We find also the somewhat narrower terms *scientific machine learning* [29], *physics-based machine learning* [39] and *predictive data science* [45], respectively.

Consider a high-dimensional manifold that contains some big data. It might be that a submanifold can be identified, which is dictated to us by the laws of physics, e.g. regarding admissible system dynamics. Learning algorithms can then be used to project the data into this submanifold. In other words, the structure of submanifold embeds physical constraints. A classical example is *Kálmán filtering* [23]. Kálmán filtering was actually an enabling technology for the moon landing in 1969, where the goal was landing within ≈500 m after ≈400,000 km of travel. More preference is given to physics or data, depending on the level of uncertainty. Kálmán filtering can be recognized as recursive Bayesian estimation. A more sophisticated example is presented in Sect. 2. *Ensemble Kálmán filtering* is used in weather forecasting centers worldwide [37]. They have to deal with about 10^6 incoming data points per hour, and mathematical models with about 10^9 states. This use case is similar to digital twinning, since data from the field is acquired and used to update the models. Citing [45, p. 39]: "Learning from data through the lens of models is a way to exploit structure in an otherwise intractable problem."

Looking closer into engineering, we notice that a large class of physics models can be decomposed into conservation laws and constitutive laws [40, Ch. 1.3], [41]. The conservation laws are of topological nature and can therefore be discretized easily, leaving little room for data-driven techniques. The situation is different for the constitutive relations, which are of metric nature, and encode phenomenological material properties. Except for simple media (local, linear) there are many potential complications (non-local, hysteretic, non-linear, multi-scale, multi-physics, etc.). Here, data-driven models can be useful, provided that the models fulfil certain admissibility criteria, which can often be expressed in terms of invariance with respect to symmetry groups (orthogonal group, Lorentz group, etc.). This is demonstrated in Sect. 3.

In physics-based modeling, one often aims at sparse models, by exploiting the principle of locality [18],[2] or at least at low-rank representations. Some popular hybrid approaches model physics by partial differential equations plus boundary conditions, represent the solution space by a deep neural network, and learn the solution in a data-driven fashion (Deep Ritz [11], Physics-Informed Neural Networks, PINNs [32]). In general, the notion of sparsity is then lost, though. Recently, Sparse, Physics-based, and partially Interpretable

---

[2]Citing [18]: *The German term "Nahwirkungsprinzip" is more impressive than the somewhat colourless word "locality". Certainly the idea behind these words, proposed by Faraday around 1830, initiated the most significant conceptual advance in physics after Newton's Principia.*

Neural Networks (SPINNs) were proposed, where sparsity is recovered by exploiting concepts from traditional mesh-free numerical methods [33].
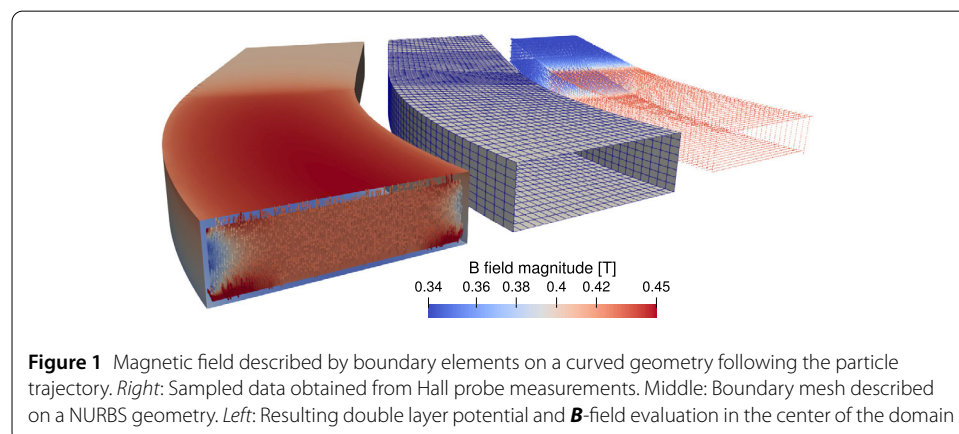
To sum up, hybrid modeling has the potential to improve the Pareto tradeoff between simulation accuracy and simulation cost significantly, and therefore bring scientific computing in engineering to the next level. In the remainder of the paper we will showcase this by some recent achievements.

2 Blending data with physics [30]

This section deals with the characterization of normal and superconducting magnets of CERN's accelerator complex. Due to the complex interplay of iron saturation, eddy currents and magnetic hysteresis, numerical field simulation alone fails to achieve the required accuracy for particle tracking applications. However, after a well-defined excitation current cycling, the physical state of the magnet is reproducible. The magnets are therefore precycled and the static magnetic field is characterized by measurements. It is beneficial to model the magnetostatic field by the Boundary Element Method (BEM). In this way, the reconstructed field is locally an exact magnetostatic solution, no domain discretization inside the air gap is required and high convergence rates are achieved for field evaluations [30].

In the sequel, the field is represented by a double layer potential. The dipole layer is located on the boundary of the domain of interest, which follows the particle trajectory, such as the curved domain illustrated in Fig. 1. Following an isogeometric approach, the domain itself is described by NURBS (Non-Rational Uniform B-Splines) surfaces and the double layer density is discretized by higher oder basis splines [8]. This gives rise to a state vector $\boldsymbol{v} \in \mathbb{R}^N$, with N degrees of freedom. The measurement data $\boldsymbol{y} \in \mathbb{R}^M$ relates to the state vector $\boldsymbol{v}$, by means of the observation operator $\boldsymbol{H} \in \mathbb{R}^{M \times N} : \boldsymbol{v} \to \boldsymbol{y}$.

Spatial field sampling is achieved by using alignment machines and magnetic field sensors, such as Hall probes or induction coils, which are supported on stages or beams with finite stiffness. To achieve a reasonable total measurement duration the data is almost always acquired on the fly. This means that the machine is performing moves along lines and a measurement is triggered whenever a certain distance is passed. As field gradients are large in the fringe field region of a magnet, most of the measurement uncertainty is due to positioning errors and vibrations. Such errors are perturbing the observation operator $\boldsymbol{H}$. To quantify the impacts, the perturbations $\boldsymbol{d} \in \mathbb{R}^{6M}$ may be introduced, encoding the de-

Figure 1 Magnetic field described by boundary elements on a curved geometry following the particle trajectory. *Right*: Sampled data obtained from Hall probe measurements. Middle: Boundary mesh described on a NURBS geometry. *Left*: Resulting double layer potential and **B**-field evaluation in the center of the domain

grees of freedom of a rigid body motion of the sensor, six for each of the M measurements. As perturbations are small deviations from zero, the linearization $\boldsymbol{y} = \boldsymbol{H}(\boldsymbol{v}) + \partial_{\boldsymbol{d}}\boldsymbol{H}(\boldsymbol{v})\boldsymbol{d}$ is justified. The operator $\partial_{\boldsymbol{d}}\boldsymbol{H}$ encodes the derivatives of $\boldsymbol{H}$ with respect to $\boldsymbol{d}$. A hybrid model is established, by estimating the state vectors $\boldsymbol{v}$ and $\boldsymbol{d}$ from measurement data. Considering both $\boldsymbol{v}$ and $\boldsymbol{d}$ as unknown random variables, Bayesian inference provides the framework to utilize all the available knowledge. To this end, Bayes rule is applied to the joint probability density function $p(\boldsymbol{v}, \boldsymbol{d}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{v}, \boldsymbol{d})p(\boldsymbol{v})p(\boldsymbol{d})$.

The distributions $p(\boldsymbol{v})$ and $p(\boldsymbol{d})$ are called prior distributions as they encode knowledge about $\boldsymbol{v}$ and $\boldsymbol{d}$, which has been available before the data $\boldsymbol{y}$ was collected. Due to the dependency of $\boldsymbol{y}$ on $\boldsymbol{v}$ and $\boldsymbol{d}$ the explicit computation of the statistical moments of $p(\boldsymbol{v}, \boldsymbol{d}|\boldsymbol{y})$ is infeasible. However, under normal distribution assumptions, the conditionals $p(\boldsymbol{v}|\boldsymbol{d}, \boldsymbol{y})$ and $p(\boldsymbol{d}|\boldsymbol{v}, \boldsymbol{y})$ take on the least squares forms,

$$
\left.\begin{array}{r} p(\boldsymbol{v}|\boldsymbol{d}, \boldsymbol{y}) \\ p(\boldsymbol{d}|\boldsymbol{v}, \boldsymbol{y}) \end{array}\right\} \sim \exp\left(-\frac{1}{2}\left\|\boldsymbol{A}_i(\boldsymbol{v}, \boldsymbol{d}) - \boldsymbol{b}_i\right\|^2\right), \quad i \in \{\boldsymbol{v}, \boldsymbol{d}\}, \tag{1}
$$

$\boldsymbol{A}_{\boldsymbol{v}} \in \mathbb{R}^{(M+N)\times(M+N)}$, $\boldsymbol{b}_{\boldsymbol{v}} \in \mathbb{R}^{M+N}$, $\boldsymbol{A}_{\boldsymbol{d}} \in \mathbb{R}^{(M+6M)\times(M+6M)}$, $\boldsymbol{b}_{\boldsymbol{d}} \in \mathbb{R}^{M+6M}$. With the conditionals at hand, samples from the joint distribution $p(\boldsymbol{v}, \boldsymbol{d}|\boldsymbol{y})$ can be computed by Gibbs sampling [35], according to Algorithm 1. As the mechanical degrees of freedom can be considered as uncorrelated between the moves, samples are drawn move-by-move in a blockwise fashion. In Fig. 2 the sampled mean as well as the maximum and minimum values for the vertical sensor position over a single move are drawn. For verification, an optical measurement was taken. Differences are within measurement precision.
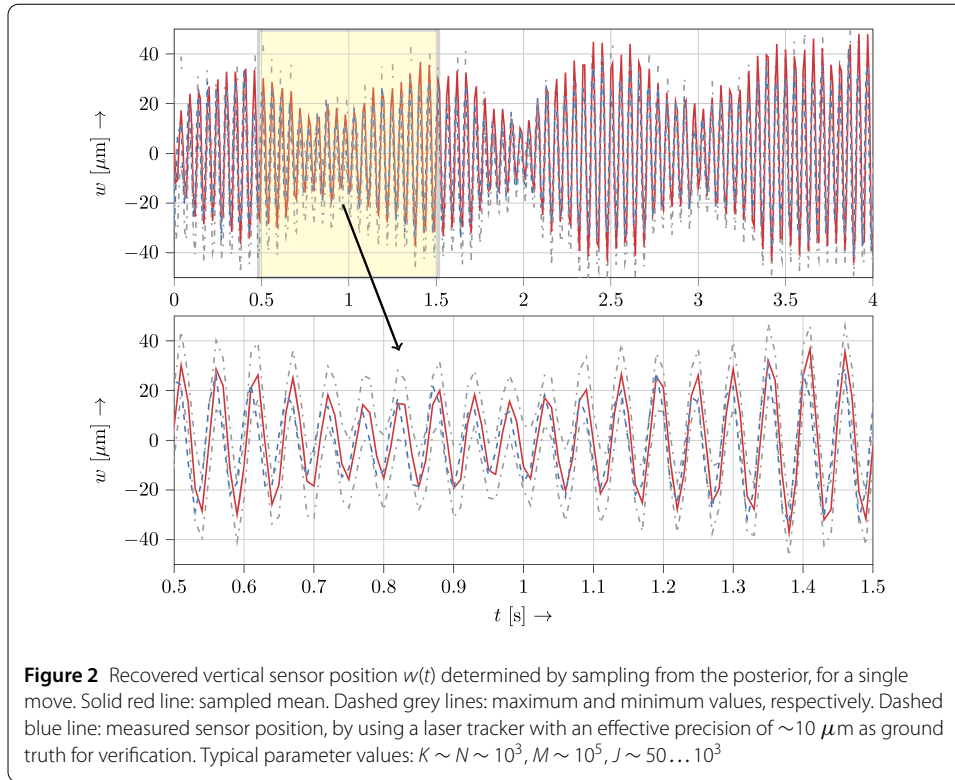
In case of a linear observation operator $\boldsymbol{H}$, sampling $\boldsymbol{v}_k$ from the high dimensional $p(\boldsymbol{v}|\boldsymbol{d}, \boldsymbol{y})$ is efficient, when implemented as the solution of the stochastic linear equation system [31]

$$
\begin{aligned} &\left(\widetilde{\boldsymbol{H}}_k^T \boldsymbol{R}^{-1} \widetilde{\boldsymbol{H}}_k + \boldsymbol{L}_{\boldsymbol{v}}\right)\boldsymbol{v}_k = \widetilde{\boldsymbol{H}}_k^T \boldsymbol{R}^{-1}\boldsymbol{y} + \boldsymbol{L}_{\boldsymbol{v}}\boldsymbol{v}_0 + \left(\widetilde{\boldsymbol{H}}_k^T \boldsymbol{R}^{-1/2}, \boldsymbol{L}_{\boldsymbol{v}}^{1/2}\right)\boldsymbol{\epsilon}, \\ &\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_{M+N}), \end{aligned} \tag{2}
$$

where a Gaussian prior $p(\boldsymbol{v}) \sim \mathcal{N}(\boldsymbol{v}_0, \boldsymbol{L}_{\boldsymbol{v}}^{-1})$ was selected. The vector $\boldsymbol{v}_0 \in \mathbb{R}^N$ is the prior expected value of the BEM state vector and $\boldsymbol{L}_{\boldsymbol{v}} \in \mathbb{R}^{N\times N}$ is the prior precision matrix. The matrix $\boldsymbol{R} \in \mathbb{R}^{M\times M}$ is a sparse, symmetric and positive definite covariance matrix of the sensor noise and $\widetilde{\boldsymbol{H}}_k \in \mathbb{R}^{M\times N}$ is the perturbed observation matrix, which is constructed

---

Algorithm 1 Sampling from the posterior $p(\boldsymbol{v}, \boldsymbol{d}|\boldsymbol{y})$

$\boldsymbol{d}_0 = \boldsymbol{0}$	▷ initialize mechanical state vector
$\boldsymbol{v}_0 \sim p(\boldsymbol{v}|\boldsymbol{d}_0, \boldsymbol{y})$	▷ sample BEM state vector, see (2)
for $k = 1, \ldots,$ number of samples K **do**	
for $j = 1, \ldots,$ number of moves J **do**	
$\boldsymbol{d}_{k,j} \sim p(\boldsymbol{d}_j|\boldsymbol{v}_{k-1}, \boldsymbol{y}_j)$	▷ sample mechanical state vector for move j
end for	
$\boldsymbol{d}_k = (\boldsymbol{d}_{k,1}^T, \ldots, \boldsymbol{d}_{1,J}^T)^T$	▷ collect mechanical state vectors
$\boldsymbol{v}_k \sim p(\boldsymbol{v}|\boldsymbol{d}_k, \boldsymbol{y})$	▷ sample BEM state vector, see (2)
end for	

Figure 2 Recovered vertical sensor position $w(t)$ determined by sampling from the posterior, for a single move. Solid red line: sampled mean. Dashed grey lines: maximum and minimum values, respectively. Dashed blue line: measured sensor position, by using a laser tracker with an effective precision of $\sim 10\,\mu$m as ground truth for verification. Typical parameter values: $K \sim N \sim 10^3$, $M \sim 10^5$, $J \sim 50\ldots10^3$

by changing the sensor position and orientation in $\boldsymbol{H}$ according to the mechanical state vector $\boldsymbol{d}_k$, by leveraging the linearization of the observation operator in $\boldsymbol{d}$.

Equation (2) follows from the conditional $p(\boldsymbol{v}|\boldsymbol{d},\boldsymbol{y})$ and the linearity of the expected value function [2, 1.17]. The structure of $\widetilde{\boldsymbol{H}}_k$ depends on the underlying field representation. Boundary element formulations give rise to dense discrete operators. Fast multipole methods can be applied to compress the products $\widetilde{\boldsymbol{H}}_k^T \boldsymbol{y}$ and $\widetilde{\boldsymbol{H}}_k \boldsymbol{x}$ if required. A solution to (2) can then be obtained by conjugate gradient iterations.

The matrix $\widetilde{\boldsymbol{H}}_k^T \boldsymbol{R}^{-1} \widetilde{\boldsymbol{H}}_k + \boldsymbol{L}_v$ corresponds to the inverse of the posterior covariance matrix. It therefore needs to be symmetric and positive definite. This can generally be achieved with the selection of a symmetric and positive definite prior precision matrix $\boldsymbol{L}_v$, (see e.g. [2, 5.1.1]). In the results presented in Fig. 2, dense operators were used and the solution of (2) was computed by using the Cholesky solver of the Eigen C++ library [17].

3 Data-driven field simulation [7, 13, 14]

This section is about data-driven field simulation for magnetostatic problems. Here, data-driven simulation is meant in the context of simulations directly on the material data and was first introduced in [26]. Consider Maxwell's equations for magnetostatics,

$$\operatorname{curl} \boldsymbol{H} = \boldsymbol{j}, \qquad \operatorname{div} \boldsymbol{B} = 0 \quad \text{in } \Omega, \tag{3}$$

where $\boldsymbol{H}$ denotes the magnetic field strength, $\boldsymbol{j}$ the imposed source current density, $\boldsymbol{B}$ the magnetic flux density, and Ω the considered domain. The problem formulation (3) is completed with appropriate boundary conditions.

The phase space of the system is denoted by $\mathcal{Z} := \{\boldsymbol{z}(\boldsymbol{x}) := (\boldsymbol{B}(\boldsymbol{x}), \boldsymbol{H}(\boldsymbol{x}))\}$, $\boldsymbol{x} \in \Omega$. The set of all states that fulfill (3) and the boundary conditions is denoted by $\mathcal{M} \subset \mathcal{Z}$, the set

of Maxwell-conforming fields.[3] Yet, to uniquely solve equations (3), a relation between $\boldsymbol{B}$ and $\boldsymbol{H}$ is necessary. Following the conventional approach, first a material law is assumed. Secondly, various fitting techniques are employed to compute the coefficients of the considered material law such that it fits best to the available measurement data. However, this approach suffers from approximation errors and additionally introduces epistemic uncertainties, i.e. the uncertainty is in the actual choice of the material model. Note that in contrast Maxwell's equations are derived from first principles and accepted to be known exactly.

The *hybrid* solver acts directly on material data and circumvents the demand for a fixed material model. The measurement data is collected in a set $\mathcal{D}^* := \{\boldsymbol{z}_i^* := (\boldsymbol{B}_i^*, \boldsymbol{H}_i^*), i = 1, \ldots, N\}$, where N is the number of measurement points. This gives rise to a set of discrete material states $\mathcal{D} := \{\boldsymbol{z} \in \mathcal{Z} | \boldsymbol{z}(\boldsymbol{x}) \in \mathcal{D}^* \forall \boldsymbol{x} \in \Omega\}$. "The material response is not known exactly and, instead, it is imperfectly characterized" [26, p. 95] by the set $\mathcal{D}$.

The solution is given by the states $\mathcal{M} \cap \mathcal{D}$ that fulfill Maxwell's equations, while being compatible with the material states. However, for a finite number of data points, this set is very likely to be empty, $\mathcal{M} \cap \mathcal{D} = \emptyset$. Therefore, we define the solution $\mathcal{S}$ by the relaxed condition

$$\mathcal{S} := \operatorname{argmin}\{d(\boldsymbol{z}, \mathcal{D}), \boldsymbol{z} \in \mathcal{M}\}, \tag{4}$$

where the distance function

$$d(\boldsymbol{z} \in \mathcal{M}, \mathcal{D}) := \min_{\boldsymbol{z}^* \in \mathcal{D}} \left\{ \frac{1}{2} \|\boldsymbol{B} - \boldsymbol{B}^*\|_{\tilde{\nu}}^2 + \frac{1}{2} \|\boldsymbol{H} - \boldsymbol{H}^*\|_{\tilde{\mu}}^2 \right\} \tag{5}$$
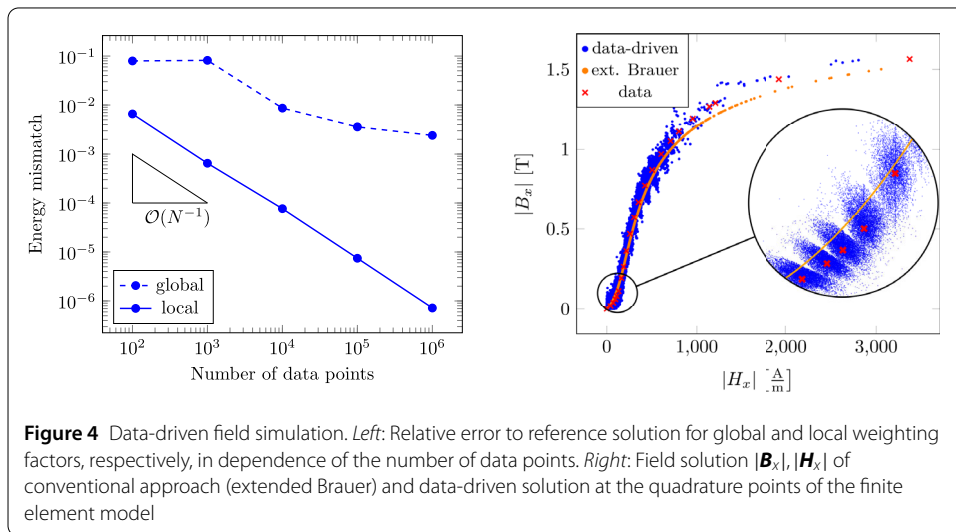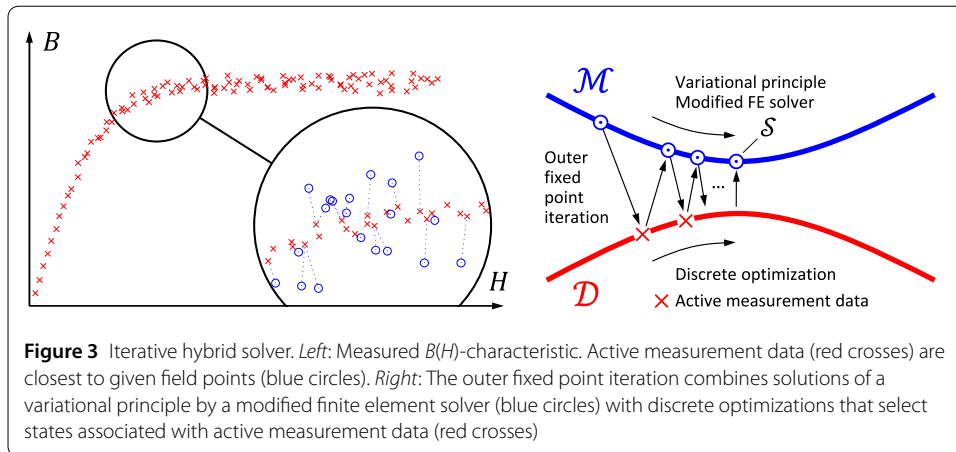
is defined in terms of the auxiliary norms $\|\cdot\|_{\tilde{\nu},\tilde{\mu}}$, which are defined as

$$\left.\begin{array}{l} \|\boldsymbol{B} - \boldsymbol{B}^*\|_{\tilde{\nu}}^2 = \int_{\Omega} (\boldsymbol{B} - \boldsymbol{B}^*) \cdot \tilde{\boldsymbol{\nu}} (\boldsymbol{B} - \boldsymbol{B}^*) \, d\Omega, \\[2mm] \|\boldsymbol{H} - \boldsymbol{H}^*\|_{\tilde{\mu}}^2 = \int_{\Omega} (\boldsymbol{H} - \boldsymbol{H}^*) \cdot \tilde{\boldsymbol{\mu}} (\boldsymbol{H} - \boldsymbol{H}^*) \, d\Omega, \end{array}\right\} \tag{6}$$

where $\tilde{\boldsymbol{\nu}}, \tilde{\boldsymbol{\mu}}$ are diagonal tensors carrying the so-called weighting factors $\tilde{\nu}, \tilde{\mu}$. Note that $\tilde{\nu}, \tilde{\mu}$ are of computational nature and do not represent physical material properties. Yet, their choice affects the convergence rate of the numerical scheme [13]. The optimization problem (4) is essentially a double minimization problem, where we accept a solution $\boldsymbol{z}$ that conforms to Maxwell's equations, while at the same time being closest to available measurement data. The solution of (4) is organized as a fixed point iteration, see Fig. 3 right. Let $\tilde{d}(\boldsymbol{z}_1, \boldsymbol{z}_2)$ denote the distance function between the two states $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ defined over the auxiliary norms (6). For any given state $\boldsymbol{z} \in \mathcal{Z}$ a modified finite element solver is used to compute the state $\boldsymbol{z}^{\odot} \in \mathcal{M}$ such that $\tilde{d}(\boldsymbol{z}, \boldsymbol{z}^{\odot}) = \min$. The solver is based on a variational principle discussed in [34]. The idea is to solve Ampère's and Faraday's laws exactly and shift the discretization error entirely into the constitutive relation.[4] Given $\boldsymbol{z}^{\odot} \in \mathcal{M}$, a discrete optimization selects the closest measurement data points, see Fig. 3

---

[3]We do not delve into regularity considerations or functional analytical frameworks here.

[4]This formulation was proposed on the Compumag Conference 1983 in Genoa. The related variational principle was called "Ligurian", in honor of the Genoa region, and in similarity to "Lagrangian" [42, p. 49].

Figure 3 Iterative hybrid solver. *Left*: Measured $B(H)$-characteristic. Active measurement data (red crosses) are closest to given field points (blue circles). *Right*: The outer fixed point iteration combines solutions of a variational principle by a modified finite element solver (blue circles) with discrete optimizations that select states associated with active measurement data (red crosses)

Figure 4 Data-driven field simulation. *Left*: Relative error to reference solution for global and local weighting factors, respectively, in dependence of the number of data points. *Right*: Field solution $|\boldsymbol{B}_x|$, $|\boldsymbol{H}_x|$ of conventional approach (extended Brauer) and data-driven solution at the quadrature points of the finite element model

left. These so-called active measurement data are associated with a state $\boldsymbol{z}^{\times} \in \mathcal{D}$ such that $\tilde{d}(\boldsymbol{z}^{\times}, \boldsymbol{z}^{\odot}) = \min$. State $\boldsymbol{z} = \boldsymbol{z}^{\times}$ is the starting point for the next iteration. Under convexity assumptions this algorithm converges to the solution of (4). Furthermore, it has been shown in [4] that for linear elasticity, the conventional solution is recovered with measurement data sets of increasing size.

The hybrid solver was applied to the model of a two-dimensional quadrupole magnet in [7] and [13]. The first work [7] extended the hybrid solver to cope with problems that feature domains where only measurement data is at hand and domains where the material law is known, e.g. in the case of vacuum. Such known material laws can be naturally included into the data-driven iterations. To analyze the performance of the hybrid solver, a strongly nonlinear soft-magnetic material was considered in the iron part of the quadrupole magnet. As a baseline, a standard magnetostatic finite element solution was considered, by discretizing 1/8 of the geometry with 6k piecewise linear elements. For the novel method, data was created by an equidistant sampling of the given non-linear $B(H)$ characteristic, without adding noise. Global weighting factors $\tilde{\nu}, \tilde{\mu}$ were employed throughout. The results show that the relative error to the reference solution, the so-called energy-mismatch, decreases as more measurement data is employed, see the dashed line in Fig. 4. However,

depending on $N = 10^2 \ldots 10^4$ the proposed method required $50 \ldots 1300$ outer iterations for convergence. In [13] local weighting factors $\tilde{\nu}(\boldsymbol{x}), \tilde{\mu}(\boldsymbol{x})$ have been introduced and defined as to be the local tangent on the current operation point $\boldsymbol{z}^\times$. Particularly in the case of unbalanced and/or data-starved sets $\mathcal{D}$, local weighting factors yield significant improvement in accuracy, see the solid line in Fig. 4. Moreover, approximately 55 iterations are needed irrespective of the cardinality of $\mathcal{D}$.

A computationally demanding 3D inductor model was considered in [14], where real-world measurement data [1] was employed for the iron part. For comparison, an extended Brauer model [22] was constructed on the measurement data and utilized in a standard Newton solver. Figure 4, right shows the absolute value of the x-component of the magnetic fields for the data-driven solution as well as the solution computed with the extended Brauer model. Furthermore, the available measurement data is depicted in red. The solution of the hybrid solver clusters around the measurement data, which is expected due to the sparse data set. Here, the data-driven solution is next to the available measurement data in all regimes of the *BH*-curve. That does not hold for the extended Brauer model which provides a good approximation in the linear part, but fails to properly model the Rayleigh part as well as the saturation part of the *BH*-curve. Both approaches lead to an acceptable solution, yet the employed material is well understood and significant effort has already been spent in the material model. If a more complex material is utilized, the hybrid solver should outperform the conventional approach.

In summary it can be said that so far hybrid solvers fail to be as computationally efficient as standard methods, albeit lots of effort is spent to improve the performance [24, 27]. Yet, hybrid solvers are advantageous if complex or novel materials are considered, since the effort for matching a material model to the measurement data is abolished. Furthermore, by circumventing the demand for a fixed material model, no epistemic uncertainties arise.

4 Bayesian free-shape optimization [36]

Bayesian optimization (BO) is an optimization method to optimize a given function which is expensive to evaluate [12]. It is built upon a hybrid architecture that blends intricate physical models with a Bayesian machine learning technique, such as Gaussian Process (GP) regression. The resulting surrogate models are cheap to evaluate, including derivatives, and keep track of their interpolation uncertainty. The core idea of BO is to successively refine those surrogates in regions of design space that are close to optimal, which are however not known beforehand. Regions with high surrogate uncertainty *might* be optimal even though the mean interpolation says otherwise. Thus, surrogate refinement requires balancing exploration against exploitation during sampling, the so-called bandit problem.[5] There are different strategies for achieving a good balance. One strategy considers the best value achieved so far and computes the *Expected Improvement* (EI). The next sample is taken at the point with the largest EI; this yields yet another optimization problem. The BO algorithm stops if the EI drops below some threshold. The BO approach can be generalized in various ways, such as BO with noise, BO in several dimensions, and BO for several objectives.

---

[5]In this model problem, an array of slot machines is considered. The gambler must balance the goal to find the slot machine with the highest gain (exploitation) with the goal to achieve good results on every play (exploration).

Figure 5 Schematic for system simulation: extracted equivalent electrical circuit (EEC) of the trace pair, surrounded by mode converters

Figure 6 Trace pair with bend. Left: The outer trace is fixed, the inner trace has a free interior surface. Right: Shape gradient vectors at the free surface for the two objectives (i) and (ii)
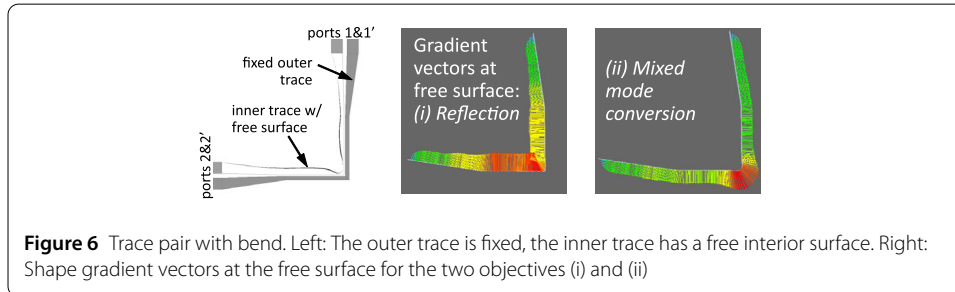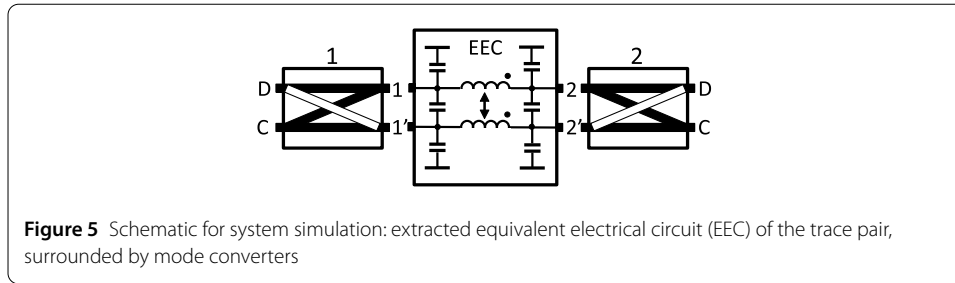
As an industrial example we consider BO of a differential trace pair on a printed circuit board. Differential signalling benefits from high immunity against electromagnetic interference and low crosstalk. However, bend discontinuities in transmission lines introduce (i) reflection and (ii) differential-to-common-mode conversion. An optimal design hence requires multi-objective optimization of the geometry. A parametric case was studied in [15], while we aim at *free-shape optimization*. Figure 5 shows a schematic for system simulation. The trace pair with ports 1,1' and 2,2', respectively, is described by an equivalent electrical circuit (EEC). Mode converters admit a separation of differential mode (D) and common mode (C) signal components. The optimization objectives can be stated in terms of *S*-parameters:[6]

(i) reflection $|S_{\mathrm{DD11}}| \stackrel{!}{=} \min$;　　(ii) mixed mode conversion $|S_{\mathrm{CD21}}| \stackrel{!}{=} \min$.

The geometric setting is depicted in Fig. 6 left. The outer trace is fixed, while the inner trace has a free interior surface. The geometry is described by a finite element mesh, and the free surface can be re-shaped by mesh morphing. This corresponds to a high-dimensional design space with ≈200 dimensions. This should be put in contrast to the six-dimensional design space that was considered in [15]. The optimization problem is: *Find the Pareto front for the shape of the free surface that minimizes the objectives.*

The ingredients for solving the optimization problem are: finite element electromagnetic field solver, EEC extraction, and adjoint sensitivity analysis. Figure 6 right shows the shape gradient vectors at the free surface for the two objectives (i) and (ii). The two gradient vector fields point in opposite directions, so the objectives are conflicting. However, the gradient fields are not exactly negatives of each other, so there is still subtle room for improvement.

The BO is extended to the multi-objective case as follows. The Pareto front is approached via a sequence of auxiliary optimization problems, each with respect to a certain

---

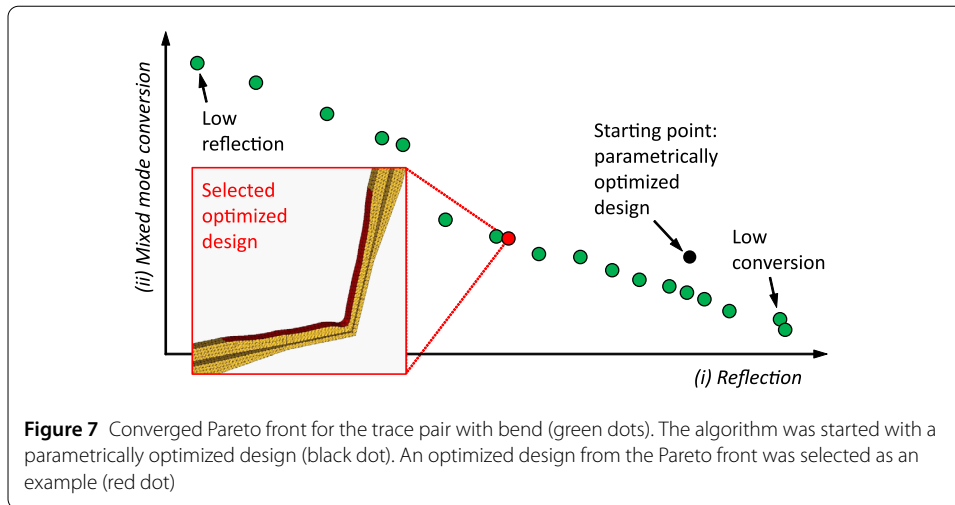[6]For simplicity evaluated at a fixed frequency of $f = 500$ MHz.

Figure 7 Converged Pareto front for the trace pair with bend (green dots). The algorithm was started with a parametrically optimized design (black dot). An optimized design from the Pareto front was selected as an example (red dot)

Table 1 Three examples of hybrid modeling

Sect.	Physics	Data	Approach
2	3D magnetostatics: BEM discretization	Fields measured by Hall probe or moving coil	Synthesis of physics and data by Bayesian update
3	2D/3D magnetostatics: Ampère's & Gauss's law	Measured material data points ($\boldsymbol{B}_i^*, \boldsymbol{H}_i^*$)	Projection of data into admissible physics manifold
4	Electromagnetic Darwin model	Data sampling from physics model	Bayesian Optimization: GP machine learning

2D affine subspace of the high-dimensional design space. This particular affine subspace is spanned by the adjoint-based gradients; it is the subspace of maximum objective variance. For each optimization problem of the sequence, BO learns and optimizes GP surrogate models for the objective functions, restricted to this subspace. Once the intermediate Pareto front is converged in this subspace, new subspaces may be chosen on the intermediate Pareto front. Figure 7 shows the result of this algorithm, after only ≈ 100 design evaluations. Note that even subtle improvement potentials will be exploited by the hybrid free-shape optimizer.

5 Summary and outlook

We highlighted three examples of hybrid modeling, cf. Table 1. Further examples stem from other applied mathematics and industrial domains, and involve artificial neural networks, such as Physics Informed Neural Networks (PINNs) [25], Deep Learning (DL) with embedded physics simulation [6], or Field Inversion and Machine Learning (FIML) [9]. They will soon find their way into scientific computing in electrical engineering, too.

Abbreviations
BEM, Boundary Element Method; BO, Bayesian Optimization; DL, Deep Learning; FIML, Field Inversion and Machine Learning; GP, Gaussian Process; EEC, Equivalent Electical Circuit; EI, Expected Improvement; NURBS, Non-Rational Uniform B-Splines; PINN, Physics-Informed Neural Network; SPINN, Sparse, Physics-based, and partially Interpretable Neural Network.

Availability of data and materials
The datasets analysed during the study in Sect. 2 are available in the repository https://github.com/MelvinLie/BEMMM.git. For Sect. 3, the measured material data can be found in [1]. Further information on the simulated geometries is given in [7, 13, 14]. The CAD model and network simulation setup explained in Sect. 4 are property of the Robert Bosch GmbH. Details are available on reasonable request, subject to permission by Robert Bosch GmbH.

Declarations

Competing interests
The authors declare that they have no competing interests. AK is one of the inventors of a pending patent titled "Verfahren und Vorrichtung zur Analyse eines Erzeugnisses" which touches the gradient computation used in Sect. 4. The invention was made within the scope of his employment contract with Robert Bosch GmbH and was duly reported to his employer as a service invention under the German law on employees' inventions.

Authors' contributions
The concept of the paper, its historical framing and look in the future, as well as the synthesis of the three hybrid modeling techniques are attributed to SK as main author. Section 2 mainly results from the work of ML, SR and SK. ML performed the measurements, analyzed the data, implemented the software for post-processing and was a major contributor in writing Sect. 2. SR made substential contributions to the concepts and revised the work in this section. Section 3 mainly results from the work of AG, DL and HDG. Section 4 mainly results from the work of AK and MS. AK created the CAD design of the model, parametrized the parasitic extraction and electric network simulation. MS parameterized and performed the Bayesian optimization. Both, AK and MS, were major contributors in writing Sect. 4 of the manuscript. The original draft has been prepared by SK, ML, AG, MS and AK. All authors have been involved in the review and revision of the paper. All authors have read and agreed to the final version of the manuscript.

Author details
[1]Robert Bosch GmbH, Renningen, Germany. [2]TEMF, Technical University of Darmstadt, Darmstadt, Germany. [3]CERN, Geneva, Switzerland.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Anglada JR, Arpaia P, Buzio M, Pentella M, Petrone C. Characterization of magnetic steels for the FCC-ee magnet prototypes. In: 2020 IEEE international instrumentation and measurement technology conference (I2MTC). 2020. p. 1–6.
2. Bardsley JM. Computational uncertainty quantification for inverse problems. vol. 19. Philadelphia: SIAM; 2018.
3. Bishop CM. Pattern recognition and machine learning. Berlin: Springer; 2006.
4. Conti S, Müller S, Ortiz M. Data-driven problems in elasticity. Arch Ration Mech Anal. 2018;229(1):79–123.
5. Coveney PV, Dougherty ER, Highfield RR. Big data need big theory too. Philos Trans R Soc A, Math Phys Eng Sci. 2016;374(2080):20160153.
6. de Avila Belbute-Peres F, Smith K, Allen K, Tenenbaum J, Kolter JZ. End-to-end differentiable physics for learning and control. In: Advances in neural information processing systems. 2018. p. 7178–89.
7. De Gersem H, Galetzka A, Ion IG, Loukrezis D, Römer U. Magnetic field simulation with data-driven material modeling. IEEE Trans Magn. 2020;56(8):1–6.
8. Dölz J, Harbrecht H, Kurz S, Schöps S, Wolf F. A fast isogeometric BEM for the three dimensional Laplace- and Helmholtz problems. Comput Methods Appl Mech Eng. 2018;330:83–101.
9. Duraisamy K. Perspectives on machine learning-augmented Reynolds-averaged and large eddy simulation models of turbulence. Phys Rev Fluids. 2021;6:050504.
10. E W. Machine learning: mathematical theory and scientific applications. ICIAM—international congress on industrial and applied mathematics. 2019. https://web.math.princeton.edu/~weinan/ICIAM.pdf. Accessed 18 Aug 2021.
11. E W, Yu B. The Deep Ritz method: a deep learning-based numerical algorithm for solving variational problems. Commun Math Stat. 2018;6(1):1–12.
12. Frazier PI. Bayesian optimization. In: Recent advances in optimization and modeling of contemporary problems. INFORMS; 2018. p. 255–78.
13. Galetzka A, Loukrezis D, De Gersem H. Data-driven solvers for strongly nonlinear material response. Int J Numer Methods Eng. 2021;122(6):1538–62.

14. Galetzka A, Loukrezis D, De Gersem H. Three-dimensional data-driven electromagnetic field simulation using real world measurement data. Compel. 2021. https://doi.org/10.1108/COMPEL-06-2021-0219

15. Gazda C, Couckuyt I, Rogier H, Ginste DV, Dhaene T. Constrained multiobjective optimization of a common-mode suppression filter. IEEE Trans Electromagn Compat. 2012;54(3):704–7.

16. Guarino N, Oberle D, Staab S. What is an ontology? In: Handbook on ontologies. Berlin: Springer; 2009. p. 1–17.

17. Guennebaud G, Jacob B, et al. Eigen v3. 2010. http://eigen.tuxfamily.org.

18. Haag R. The principle of locality in classical physics and the relativity theories. Local quantum physics, texts and monographs in physics. Berlin: Springer; 1992. Chap. I.2.

19. Haußmann M, Gerwinn S, Look A, Rakitsch B, Kandemir M. Learning partially known stochastic dynamics with empirical PAC Bayes. In: Banerjee A, Fukumizu K, editors. Proceedings of the 24th international conference on artificial intelligence and statistics. Proceedings of machine learning research. vol. 130. 2021. p. 478–86.

20. He R, Xu Y, Walunj S, Yong S, Khilkevich V, Pommerenke D, Aichele HL, Boettcher M, Hillenbrand P, Klaedtke A. Modeling strategy for EMI filters. IEEE Trans Electromagn Compat. 2020;62(4):1572–81.

21. Higham CF, Higham DJ. Deep learning: an introduction for applied mathematicians. SIAM Rev. 2019;61(4):860–91.

22. Hülsmann T, Bartel A, Schöps S, De Gersem H. Extended Brauer model for ferromagnetic materials: analysis and computation. Compel. 2014;33(4):1251–63.

23. Kálmán RE. A new approach to linear filtering and prediction problems. J Basic Eng. 1960;82:35–45.

24. Kanno Y. Mixed-integer programming formulation of a data-driven solver in computational elasticity. Optim Lett. 2019;13:1505–14.

25. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. Nat Rev Phys. 2021;3(6):422–40.

26. Kirchdoerfer T, Ortiz M. Data-driven computational mechanics. Comput Methods Appl Mech Eng. 2016;304:81–101.

27. Korzeniowski TF, Weinberg K. A multi-level method for data-driven finite element computations. Comput Methods Appl Mech Eng. 2021;379:113740.

28. Kurz S. Hybrid modeling: towards the next level of scientific computing in engineering. In: van Beurden M, Budko N, Schilders W, editors. Scientific computing in electrical engineering SCEE 2020. Eindhoven, the Netherlands, February 2020. Mathematics in industry. vol. 36. ECMI—The European Consortium for Mathematics in Industry, Springer; 2022.

29. Lee S, Baker N. Basic research needs for scientific machine learning: core technologies for artificial intelligence. Tech. rep., USDOE Office of Science. 2018. https://www.osti.gov/servlets/purl/1484362. Accessed 18 Aug 2021.

30. Liebsch M, Russenschuck S, Kurz S. Boundary-element methods for field reconstruction in accelerator magnets. IEEE Trans Magn. 2020;56(3):1–4.

31. Nishimura A, Suchard MA. Prior-preconditioned conjugate gradient method for accelerated Gibbs sampling in "large n & large p" Bayesian sparse regression. 2021. arXiv:1810.12437.

32. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys. 2019;378:686–707.

33. Ramabathiran AA, Ramachandran P. SPINN: sparse, physics-based, and partially interpretable neural networks for PDEs. J Comput Phys. 2021;445:110600.

34. Rikabi JAH, Bryant CF, Freeman EM. An error-based approach to complementary formulations of static field solutions. Int J Numer Methods Eng. 1988;26(9):1963–87.

35. Robert C, Casella G. Monte Carlo statistical methods. Berlin: Springer; 2013.

36. Schuhmacher S, Klaedtke A, Keller C, Ackermann W, De Gersem H. Adjoint technique for sensitivity analysis of coupling factors according to geometric variations. IEEE Trans Magn. 2018;54(3):1–4.

37. Stuart AM. The legacy of Rudolph Kálmán—blending data and mathematical models. In: Boeing distinguished colloquia. Univ. Washington. 2019. https://www.sfb1294.de/fileadmin/user_upload/Kalman_Lectures/1st_Kalman_Lecture_2018_Andrew_Stuart.pdf. Accessed 18 Aug 2021.

38. Stysch J, Klaedtke A, De Gersem H. Low-frequency stabilization for FEM impedance computation. IEEE Trans Electromagn Compat. 2021. 1–10. https://doi.org/10.1109/TEMC.2021.3134323.

39. Swischuk R, Mainini L, Peherstorfer B, Willcox K. Projection-based model reduction: formulations for physics-based machine learning. Comput Fluids. 2019;179:704–17.

40. Tonti E. The mathematical structure of classical and relativistic physics. Berlin: Springer; 2013.

41. Tonti E. Discrete physics—algebraic formulation of physical fields. 2014. http://www.discretephysics.org/en/. Accessed 18 Aug 2021.

42. Trowbridge B. Compumag conference—the first 25 years. 2001. https://www.compumag.org/wp/wp-content/uploads/2018/07/TwentyFiveYearsOfCompumag.pdf. Accessed 18 Aug 2021.

43. Walker M. Hype cycle for emerging technologies. Tech. rep. G00340159, Gartner Research. 2018. https://www.gartner.com/en/documents/3885468/hype-cycle-for-emerging-technologies-2018. Accessed 18 Aug 2021.

44. Wikipedia contributors: Newton's laws of motion—Wikipedia, the free encyclopedia. 2019. https://en.wikipedia.org/w/index.php?title=Newton%27s_laws_of_motion. Accessed 18 Aug 2021.

45. Willcox K. Predictive data science for physical systems—from model reduction to scientific machine learning. ICIAM—international congress on industrial and applied mathematics. 2019. https://kiwi.oden.utexas.edu/papers/Willcox-Predictive-Data-Science-ICIAM-2019.pdf. Accessed 18 Aug 2021.