ORIGINAL ARTICLE

Reduced order modelling using neural networks for predictive modelling of 3d-magneto-mechanical problems with application to magnetic resonance imaging scanners

S. Miah[1] · Y. Sooriyakanthan[2] · P. D. Ledger[1] · A. J. Gil[2] · M. Mallett[3]

Abstract

The design of magnets for magnetic resonance imaging (MRI) scanners requires the numerical simulation of a coupled magneto-mechanical system where the effects that different material parameters and in-service loading conditions have on both imaging and MRI performance are key to aid with the design and the manufacturing process. To correctly capture the complex physics, and to obtain accurate solutions, finite element simulations with dense meshes and high order elements are needed. Reduced order model approaches, based on the established proper orthogonal decomposition (POD) approach, are attractive as they can rapidly predict the numerical simulations needed under changing parameters or conditions. However, the projected (PODP) approach has an invasive computational implementation, whilst the interpolated (PODI) approach presents challenges when the dimension of the space of parameters to be investigated becomes large. As an alternative, we investigate a POD technique based on using a neural network regression, which is not as invasive as PODP, but has superior approximation properties compared to PODI. We apply this to the coupled magneto-mechanical system to understand three pressing industrial problems: firstly, the accurate and rapid computation of the resonant frequencies associated with this coupled magneto-mechanical system, secondly, the effects of magnet motion on the Ohmic power and kinetic energy curves, and, thirdly, the prediction of the uncertainty in Ohmic power and kinetic energy curves as a function of exciting frequency for uncertain material parameters.

Keywords Magneto-mechanical coupling · Magnetic resonance imaging · Neural networks · Machine learning · Proper orthogonal decomposition · Reduced order model

✉ P. D. Ledger
p.d.ledger@keele.ac.uk

S. Miah
m.s.miah@keele.ac.uk

Y. Sooriyakanthan
y.sooriyakanthan.970070@swansea.ac.uk

A. J. Gil
a.j.gil@swansea.ac.uk

M. Mallett
michael.mallett@siemens-healthineers.com

[1] School of Computing and Mathematics, Keele University, Newcastle under Lyme, UK

[2] Zienkiewicz Centre for Computational Engineering, Faculty of Science and Engineering, Swansea University, Swansea, UK

[3] Siemens Healthineers, MR Magnet Technology, Eynsham, Witney, Oxford, UK

1 Introduction

The use of magnetic resonance imaging (MRI) [1] has been an essential tool in modern medical diagnosis due to the scanners' high in-built resolution when imaging fractures [2], joints [3], and soft tissues, such as damaged cartilage [2] or tumors [4].

Time varying magnetic fields generated as part of the imaging process by an MRI scanner interact with the conducting shields and generate unwanted eddy currents. The purpose of these shields is to isolate or block the strong static magnetic field generated by the main MRI magnet from the region outside the scanner. However, gradient coils, which generate weaker time varying magnetic fields, generate eddy currents and give rise to electromagnetic stresses causing deformations and vibrations within the MRI scanner [5–7]. This, in turn, can lead to patient discomfort and ghosting. Still further, eddy currents and vibrations release energy, which

can cause helium boil-off, an expensive coolant used to keep the magnets superconducting, and this may, ultimately, lead to a magnet quench. Vibrations are also caused by external machinery (e.g. lifts, cleaning equipment and building maintenance), which can lead to eddy currents being generated in conducting components of the scanner and may cause disturbance to the uniformity of the strong magnet field generated by the main magnet and also lead to ghosting effects. Manufacturers of MRI scanners want to predict and understand these vibrations to formulate a better understanding of the coupling phenomena and design new and better scanners, which exploit alternative materials. To do this, they need to carefully optimise the configuration of the model parameters in order to ensure that the new magnets meet their design specification.

To help design scanners, in collaboration with our industrial partner Siemens Healthineers, this paper addresses three pressing industrial problems: firstly, the accurate and rapid computation of the resonant frequencies associated with this coupled magneto-mechanical system, secondly, the effects of magnet motion on the Ohmic power and kinetic energy curves, and, thirdly, the prediction of the uncertainty in Ohmic power and kinetic energy curves as a function of exciting frequency for uncertain material parameters. Accurate full order model simulations require fine finite element discretisations using dense meshes and/or high order elements, and can take significant time to compute [8–12]. To reduce the cost of solutions for new parameters, proper orthogonal decomposition (POD) reduced order model (ROM) schemes [13–16] have been explored that use either Galerkin projection or interpolation [17, 18]. However, the projected POD (PODP) approach has an invasive computational implementation, while the interpolated POD approach (PODI) presents challenges when the dimension of the space of parameters to be investigated becomes large. As an alternative, we investigate a POD technique based on using a neural network regression [19, 20], which is not invasive but has superior approximation properties compared to PODI. Using multilayer perceptrons (MLP) neural networks in combination with POD has been employed by Hesthaven and Ubbiali [21] and this technique is called PODNN (or POD-NN) in the literature. The use of neural networks in the context of POD range from the prediction of wind pressure and wind induced responses for high-rise buildings [22] with a emphasis on back-propagating networks, to the use of recurrent neural networks (RNN) closure of parametric POD-Galerkin ROMs [23]. In the latter, it is shown that a long short-term memory (LSTM) type of RNN can significantly improve accuracy and efficiency for non-linear problems, even beyond the time interval of training data. More recently, the use of other deep learning based reduced order models (DL-ROMs) in POD has been considered in [24, 25] with a more extensive literature review given in [26] covering artificial neural networks (ANN), physics informed neural networks (PINNs) and feed forward neural networks and the differences between them. The application of this technology to the three aforementioned problems will demonstrate the advantages of using the neural network based POD over our previous POD approaches.

The structure of the paper is as follows: Sect. 2 briefly outlines the mathematical model of the problem of the coupled magneto-mechanical problem, in the form of a transmission problem, that we wish to tackle in this project. In this section, we also describe the finite element discretisation and the structure of the algebraic system to be solved. Then, in Sect. 3, we describe the reduced order modelling approach based on POD and describe the offline and online stages. For the online stage, we focus on PODNN and show comparisons with other POD approaches. We also emphasize the importance of how the neural networks should be applied to POD to ensure accurate results efficiently. Sect. 4 presents a series of numerical examples where the PODNN technique is applied to the problems of interest in this work. This includes the prediction of resonant modes in the Ohmic power and kinetic energy curves using different parameter spaces, predicting the response in the same outputs of interest for different external vibrational loading conditions and predicting the response in outputs of interest when model parameters are uncertain. Finally, in Sect. 5 we conclude our findings and recall the work carried out in this paper.
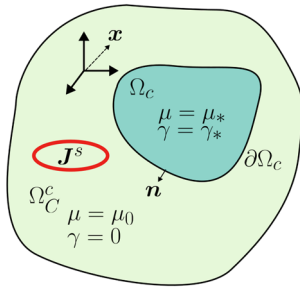
2 Magneto-mechanical model

To computationally simulate the coupled electro-mechanical interactions in a MRI scanner we choose to follow [12] and formulate the equations in a Lagrangian setting. We recall that this mathematical model assumes that both the displacements and the velocities of the conducting components are small. The problem of interest is described by Fig. 1, where Ω_C denotes a conducting region, $\Omega_C^c := \mathbb{R}^3 \backslash \overline{\Omega}_C$ denotes the (unbounded) non conducting region and $\partial \Omega_C$ represents the boundary of the conducting region. The unknown fields for this problem are the magnetic vector potential $\boldsymbol{A}$ and the mechanical displacement $\boldsymbol{u}$, while $\boldsymbol{J}$ is a known external current source.

Continuing to follow [12], by linearising about the static (DC) solution, the transient (AC) problem becomes linear in time (t) dependent fields and the latter can be represented in a time-harmonic fashion so that

$$A(t) = A^{\mathrm{DC}} + \mathrm{Re}\left(\boldsymbol{\mathcal{A}}^{\mathrm{AC}} e^{\mathrm{i}\omega t}\right), \tag{1a}$$

$$u(t) = u^{\mathrm{DC}} + \mathrm{Re}\left(\boldsymbol{\mathcal{U}}^{\mathrm{AC}} e^{\mathrm{i}\omega t}\right), \tag{1b}$$

Fig. 1 General representation of the magneto-mechanical problem, illustrating a conducting region Ω_c with magnetic permeability $\mu = \mu_*$ and electrical conductivity $\gamma = \gamma_*$ that is within an unbounded non-conducting region Ω_C^c with $\mu = \mu_0$ and $\gamma = 0$. An excitation is applied by a current source $\boldsymbol{J}^s$ that is prescribed in the coils

with a similar decomposition for $\boldsymbol{J}$, where $\omega = 2\pi f$ is the angular frequency of the driving excitation, f is the frequency in Hz and $\mathrm{i} := \sqrt{-1}$. A linearised transmission problem has already been derived in [12, 27] for this problem leading to the strong form

$$\operatorname{curl}(\mu^{-1} \operatorname{curl} \boldsymbol{A}^{\mathrm{AC}}) + \mathrm{i}\omega\gamma \boldsymbol{A}^{\mathrm{AC}} = \boldsymbol{J}^{\mathrm{AC}} \qquad \text{in } \mathbb{R}^3, \tag{2a}$$

$$\operatorname{div} \boldsymbol{A}^{\mathrm{AC}} = 0 \qquad \text{in } \Omega_C^c, \tag{2b}$$

$$\operatorname{div}\left(\sigma^m(\boldsymbol{\mathcal{U}}^{\mathrm{AC}}) + \mu^{-1}\boldsymbol{T}(\boldsymbol{A}^{\mathrm{DC}}, \boldsymbol{A}^{\mathrm{AC}})\right) = -\rho\omega^2 \boldsymbol{\mathcal{U}}^{\mathrm{AC}} \quad \text{in } \Omega_C, \tag{2c}$$

$$\boldsymbol{A}^{\mathrm{AC}} = O(|\boldsymbol{x}|^{-1}) \qquad \text{as } |\boldsymbol{x}| \to \infty, \tag{2d}$$

$$\boldsymbol{\mathcal{U}}^{\mathrm{AC}} = \boldsymbol{\mathcal{U}}_D^{\mathrm{AC}} \qquad \text{on } \partial\Omega_C^D, \tag{2e}$$

$$\boldsymbol{n} \times \left[\boldsymbol{A}^{\mathrm{AC}}\right]_{\partial\Omega_C} = \boldsymbol{0} \qquad \text{on } \partial\Omega_C, \tag{2f}$$

$$\boldsymbol{n} \times \left[\mu^{-1} \operatorname{curl} \boldsymbol{A}^{\mathrm{AC}}\right]_{\partial\Omega_C} = \boldsymbol{0} \qquad \text{on } \partial\Omega_C, \tag{2g}$$

$$\left(\sigma^m(\boldsymbol{A}^{\mathrm{AC}}) + \mu^{-1}\boldsymbol{T}(\boldsymbol{A}^{\mathrm{DC}}, \boldsymbol{A}^{\mathrm{AC}})\right)\Big|_{\partial\Omega_C}^{-} \boldsymbol{n}$$
$$= \left(\mu^{-1}\boldsymbol{T}(\boldsymbol{A}^{\mathrm{DC}}, \boldsymbol{A}^{\mathrm{AC}})\right)\Big|_{\partial\Omega_C}^{+} \boldsymbol{n} \qquad \text{on } \partial\Omega_C \setminus \partial\Omega_C^D, \tag{2h}$$

where γ, ρ and μ are the electric conductivity, mass density and magnetic permeability, respectively; $\boldsymbol{n}$ is a unit outward normal (pointing from the conducting to the non-conducting side); $\sigma^e(\boldsymbol{\mathcal{U}}^{\mathrm{AC}})$ is defined as the mechanical contribution to the Cauchy stress tensor; $\left(\mu^{-1}\boldsymbol{T}(\boldsymbol{A}^{\mathrm{DC}}, \boldsymbol{A}^{\mathrm{AC}})\right)$ is the linearised Maxwell stress tensor; $\partial\Omega_C^D$ denotes the (mechanical) Dirichlet part of $\partial\Omega_C$ and $[\cdot]_{\partial\Omega_C} := (\cdot)|_{\partial\Omega_C}^{+} - (\cdot)|_{\partial\Omega_C}^{-}$ denotes the jump, with $(\cdot)|_{\partial\Omega_C}^{+}$ and $(\cdot)|_{\partial\Omega_C}^{-}$ representing the non-conducting and conducting side of $\partial\Omega_C$, respectively. Once (2) has been solved, the complex amplitudes of the electric and magnetic AC fields in the Eulerian setting can be recovered as

$$\boldsymbol{\mathcal{E}}^{\mathrm{AC}} = -\mathrm{i}\omega\boldsymbol{A}^{\mathrm{AC}} + \mathrm{i}\omega\boldsymbol{B}_0^{\mathrm{DC}} \times \boldsymbol{\mathcal{U}}^{\mathrm{AC}} \qquad \text{in } \Omega_C, \tag{3a}$$

$$\boldsymbol{\mathcal{H}}^{\mathrm{AC}} = \mu^{-1}\boldsymbol{\mathcal{B}}_0^{\mathrm{AC}} \qquad \text{in } \mathbb{R}^3, \tag{3b}$$

where $\boldsymbol{\mathcal{B}}_0^{\mathrm{AC}} = \operatorname{curl} \boldsymbol{A}^{\mathrm{AC}}$ and $\boldsymbol{B}_0^{\mathrm{DC}} = \operatorname{curl} \boldsymbol{A}^{DC}$ and the complete physical fields for both stages are

$$\boldsymbol{E} = \operatorname{Re}(\boldsymbol{\mathcal{E}}^{\mathrm{AC}} e^{\mathrm{i}\omega t}) = \operatorname{Re}((-\mathrm{i}\omega\boldsymbol{A}^{\mathrm{AC}}$$
$$+ \mathrm{i}\omega\boldsymbol{B}_0^{\mathrm{DC}} \times \boldsymbol{\mathcal{U}}^{\mathrm{AC}}) e^{\mathrm{i}\omega t}) \quad \text{in } \Omega_C, \tag{4a}$$

$$\boldsymbol{H} = \mu^{-1}\boldsymbol{B}_0 = \mu^{-1}(\boldsymbol{B}_0^{\mathrm{DC}} + \operatorname{Re}(e^{\mathrm{i}\omega t}\boldsymbol{\mathcal{B}}_0^{\mathrm{AC}}))$$
$$= \mu^{-1}(\operatorname{curl} \boldsymbol{A}^{DC} + \operatorname{Re}(e^{\mathrm{i}\omega t} \operatorname{curl} \boldsymbol{A}^{\mathrm{AC}})) \quad \text{in } \mathbb{R}^3. \tag{4b}$$

In this paper, we assume that the DC fields are known. Our goal is to provide a new reduced order model procedure to allow the rapid prediction of approximations to $\boldsymbol{A}^{\mathrm{AC}}$ and $\boldsymbol{\mathcal{U}}^{\mathrm{AC}}$ for new material parameters and loading conditions. In turn, this allows us to obtain associated outputs of interest as well as to understand the uncertainty in outputs of interest for uncertain material parameters. To this end, we truncate Ω_C^c a finite distance from Ω_C, and create the truncated domain Ω. On $\partial\Omega$, we impose $\boldsymbol{n} \times \boldsymbol{A}^{\mathrm{AC}} = \boldsymbol{0}$ as an approximation to the decay condition (2d). To circumvent the Coulomb type gauge in (2b), we add a small perturbation $\varepsilon\boldsymbol{A}^{\mathrm{AC}}$ to the left hand side of (2a) in $\Omega \setminus \Omega_C$ and replace $\boldsymbol{A}^{\mathrm{AC}}$ with the regularised solution $\boldsymbol{A}_\varepsilon^{\mathrm{AC}}$. We then use the finite element method (FEM) to approximate full order model solutions to $\boldsymbol{A}_\varepsilon^{\mathrm{AC}}$ and $\boldsymbol{\mathcal{U}}^{\mathrm{AC}}$ over the Ω. For further details we refer to [12]. In the next section, we briefly recall the finite element (FE) discretisation employed.

2.1 FEM discretisation

FEM discretisation involves partitioning the domain into non-overlapping regions called elements and then approximating the solution over each of these elements by appropriate piecewise polynomials. We will use an unstructured grid of tetrahedral elements due to the automatic procedures that are available for generating these meshes around complex configurations and adopt a higher order (hp-version) FE

framework allowing the combination of refining the mesh spacing (h) and increasing the order (p) of these polynomials in order to improve the accuracy of the solution [8, 10, 11], as was applied in the case in our previous work [12]. The use of such technology allows accurate solutions to be obtained and, in the case of p-refinement for smooth solutions, and combinations of h- and p-refinements, for non-smooth solutions due to sharp edges and corners, is known to lead to exponential convergence of the solution [8]. The mathematics and continuity requirements of the fields dictate that the mechanical displacement and magnetic vector potential should be approximated differently, for details of the weak form of (2) and its derivation we refer to [12]. This then means that $\boldsymbol{H}$(curl) conforming basis functions be used for $\boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}}$ while H^1 conforming basis functions should be used for each component of $\boldsymbol{\mathcal{U}}^{\mathrm{AC}}$. The particular choice of basis functions employed correspond to those proposed by Schöberl and Zaglmayr [28, 29]. This means that fields $\boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}}$ and $\boldsymbol{\mathcal{U}}^{\mathrm{AC}}$ are approximated by

$$\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{\mathrm{AC}} = \sum_{g=1}^{P_{\mathrm{global}}^{\mathrm{AC}}} \mathbf{N}^g \mathcal{A}_\epsilon^{AC,g}, \tag{5a}$$

$$\boldsymbol{u}_{hp}^{\mathrm{AC}} = \sum_{s=1}^{Q_{\mathrm{global}}} \mathbf{L}^s \mathcal{U}^{AC,s}, \tag{5b}$$

where the subscripts hq and hp are used to indicate that these are discrete approximations on a mesh of spacing h with different orders q and p, additionally, $P_{\mathrm{global}}^{\mathrm{AC}}$ and Q_{global} refers to the number of degrees of freedom (DOF) in the AC system in the same respective manner. Still further, $\mathbf{N}^g$ and $\mathbf{L}^s$ are typical $\boldsymbol{H}$(curl) and H^1 basis functions respectively [17].

2.2 Full order model

Using the FEM recalled above, the application of the linearised strategy in [12] leads to the need to solve parametrised linear systems of the form: Find $\mathcal{A}_\epsilon^{\mathrm{AC}}$ and $\mathcal{U}^{\mathrm{AC}}$ such that

$$\begin{bmatrix} \mathbf{K}_{\mathcal{A}\mathcal{A}}^{\mathrm{AC}} + \mathbf{C}_{\mathcal{A}\mathcal{A}}^{\mathrm{AC}} & \mathbf{0} \\ \mathbf{K}_{\mathcal{U}\mathcal{A}}^{\mathrm{AC}} & \mathbf{K}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}} - \omega^2 \mathbf{M}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}} \end{bmatrix} \begin{pmatrix} \mathcal{A}_\varepsilon^{\mathrm{AC}} \\ \mathcal{U}^{\mathrm{AC}} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{\mathcal{A}}^{\mathrm{AC}} \\ 0 \end{pmatrix}, \tag{6}$$

where $\mathbf{K}_{\mathcal{A}\mathcal{A}}^{\mathrm{AC}}$, $\mathbf{C}_{\mathcal{A}\mathcal{A}}^{\mathrm{AC}}$, $\mathbf{K}_{\mathcal{U}\mathcal{A}}^{\mathrm{AC}}$, $\mathbf{K}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}}$, $\mathbf{K}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}}$ and $\mathbf{M}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}}$ refer to the system matrices that are obtained by discretisation of the weak forms. The explicit definitions for these matrix blocks can be seen in [18] with $\mathbf{K}$, $\mathbf{M}$ and $\mathbf{C}$ referring to stiffness, mass and damping type matrices, respectively. The $\mathbf{R}_{\mathcal{A}}^{\mathrm{AC}}$ term drives the solution to the problem. The structure of (6) means that a computational efficient solution strategy

is first employed to solve

$$\left[\mathbf{A}(\boldsymbol{w}) \right] \left(\mathbf{q}(\boldsymbol{w}) \right) = \left(\mathbf{r}(\boldsymbol{w}) \right), \tag{7}$$

for $\mathbf{q}(\boldsymbol{w}) := \mathcal{A}_\varepsilon^{\mathrm{AC}}(\boldsymbol{w})$, where $\mathbf{A}(\boldsymbol{w}) := \mathbf{K}_{\mathcal{A}\mathcal{A}}^{\mathrm{AC}}(\boldsymbol{w}) + \mathbf{C}_{\mathcal{A}\mathcal{A}}^{\mathrm{AC}}(\boldsymbol{w})$ and $\mathbf{r}(\boldsymbol{w}) := \mathbf{R}_{\mathcal{A}}^{\mathrm{AC}}(\boldsymbol{w})$, we have emphasised how the solution may depend on parameters of interest $\boldsymbol{w}$, and have introduced the simplified notation to help with later sections. Once (7) has been solved, $\mathcal{U}^{\mathrm{AC}}$ can be obtained from solving

$$\left[\mathbf{K}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}} - \omega^2 \mathbf{M}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}} \right] \left(\mathcal{U}^{\mathrm{AC}} \right) = - \left[\mathbf{K}_{\mathcal{U}\mathcal{A}}^{\mathrm{AC}} \right] \left(\mathcal{A}_\varepsilon^{\mathrm{AC}}(\boldsymbol{w}) \right), \tag{8}$$

and, in the case of a non-zero Dirichlet boundary condition in (2c), to the right hand side in (8), an additional contribution corresponding to columns of $\left[\mathbf{K}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}} - \omega^2 \mathbf{M}_{\mathcal{U}\mathcal{U}}^{\mathrm{AC}} \right]$ that have been moved to the right hand side multiplied by the corresponding solution coefficients found from the Dirichlet condition [9], should be subtracted. Of the systems (7) and (8), in practice, the former is larger and more computationally expensive to solve since, typically, $P_{\mathrm{global}}^{\mathrm{AC}} \gg Q_{\mathrm{global}}^{\mathrm{AC}}$ as Ω_{C} is just a small part of Ω and the latter only needs to be solved for those mechanical degrees of freedom in Ω_{C}. Henceforth, following the justification considered in [17], the former will be our focus in the following section.

3 Reduced order modelling

To aid with the development and design of a new MRI scanner, linear systems of the form (6) must be repeatability solved a large number of times for different parameters (such as frequency or conductivity of radiation shields), which can lead to exponential growth in computational cost. To address this, we consider a reduced order model (ROM) to optimise work flow and reduce the overall computational cost. In particular, we consider a form of proper orthogonal decomposition (POD) approach e.g. [14, 30, 31], which seeks to extract modal information from a small number of solution snapshots corresponding to sets of snapshot parameters and use the most important of these to construct a problem of reduced size. These ROMs have been successfully applied to a range of engineering problems and, in the context of the linearised magneto-mechanical problem [17], interpolated (PODI) and projected versions (PODP) of the POD for the system (7) have been considered. In this work, we investigate a POD technique based on using a neural network regression (PODNN). Our implementation is equivalent to that proposed by Hesthaven and Ubbiali [21], and is less invasive than PODP and is amenable to a large number of potential parameters $\boldsymbol{w}$. We briefly outline the offline (shared by all POD approaches) and the differing online stages below. We also provide some practical insights in to how PODNN should be applied to achieve accurate results.

3.1 Offline stage

Computed solutions to (7) for different sets of (snapshot) parameters $\boldsymbol{w} = \mathbf{w}_j \in \mathbb{R}^P$ are arranged in $\mathbf{D}$ as:

$$\mathbf{D} = (\mathbf{q}(\mathbf{w}_1), \mathbf{q}(\mathbf{w}_2), \mathbf{q}(\mathbf{w}_3), \ldots, \mathbf{q}(\mathbf{w}_{N_s})) \in \mathbb{C}^{\eta \times N_s}, \qquad (9)$$

with N_s representing the number of computed snapshots and $\eta = P_{\text{global}}^{\text{AC}}$. To extract the modal information, a singular value decomposition (SVD) is applied so that $\mathbf{D} = \mathbf{H}\Sigma\mathbf{G}^H$, where $\mathbf{H} \in \mathbb{C}^{\eta \times \eta}$ and $\mathbf{G} \in \mathbb{C}^{N_s \times N_s}$ being unitary matrices, $\Sigma \in \mathbb{R}^{\eta \times N_s}$ being a diagonal matrix padded out with zeros and Σ_{ii} being the singular values, arranged in descending order decay towards zero and H denoting the matrix Hermitian. Then, by dropping those singular values corresponding to $\Sigma_{ii}/\Sigma_{11} \leq \text{TOL}$, $\mathbf{D}$ can be well approximated by a truncated singular value decomposition (TSVD)

$$\mathbf{D} \approx \mathbf{D}^m = \mathbf{H}^m \Sigma^m (\mathbf{G}^m)^H, \qquad (10)$$

where m is the number of singular values that are retained. Since $m < N_s << \eta$, the truncated matrices of left and right singular vectors follow the form $\mathbf{H}^m = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_m] \in \mathbb{C}^{\eta \times m}$ and $(\mathbf{G}^m)^H \in \mathbb{C}^{m \times N_s}$ whilst the diagonal matrix follows $\Sigma^m \in \mathbb{R}^{m \times m}$. We would like m to be small whilst also maintaining a good approximation to $\mathbf{D}$.

3.2 Online stage

The online stage involves approximating the solution of (7) for new sets of parameters at reduced computational cost compared to the offline stage. The aim is to calculate an approximation $\mathbf{q}(\boldsymbol{w})$ for new $\boldsymbol{w}$ that are different to the snapshot parameters $\mathbf{w}_j$.

3.2.1 Interpolated POD (PODI)

In the case of interpolated POD (PODI), the approximation is typically chosen as e.g [17, 32]

$$\mathbf{q}^{\text{PODI}}(\boldsymbol{w}) = \sum_{k=1}^{m} \mathbf{h}_k \Sigma_{kk} I_k(\boldsymbol{w}, (\mathbf{G}^m)^H), \qquad (11)$$

where $I_k(\boldsymbol{w}, (\mathbf{G}^m)^H)$ is an interpolation through the kth row of $(\mathbf{G}^m)^H$ or equivalently the kth column of $\overline{\mathbf{G}^m}$, with the overbar denoting the complex conjugate, for the parameter set of interest $\boldsymbol{w}$. For further details of the implementation see [17].

3.2.2 Projected POD (PODP)

In the case of projected POD (PODP), the form of the approximation is chosen as e.g [15, 17, 33]

$$\mathbf{q}^{\text{PODP}}(\boldsymbol{w}) = \mathbf{H}^m \mathbf{p}(\boldsymbol{w}), \qquad (12)$$

where $\mathbf{p}(\boldsymbol{w}) \in \mathbb{C}^m$ is obtained from solving a small parameter dependent linear system obtained by Galerkin projection of (7)

$$\mathbf{A}^m(\boldsymbol{w})\mathbf{p}^m(\boldsymbol{w}) = \mathbf{r}^m(\boldsymbol{w}), \qquad (13)$$

where $\mathbf{A}^m = (\mathbf{H}^m)^H \mathbf{A}(\boldsymbol{w})\mathbf{H}^m$ and $\mathbf{r}^m = (\mathbf{H}^m)^H \mathbf{r}$ for each parameter set of interest. For further details of the implementation see [17].

3.2.3 Neural network POD (PODNN)

In the case of a neural network POD (PODNN), one possible choice is analogous to (11) i.e.

$$\mathbf{q}^{\text{PODNN}}(\boldsymbol{w}) = \sum_{k=1}^{m} \mathbf{h}_k \Sigma_{kk} R_k\left(\boldsymbol{w}, \boldsymbol{p}\left((\mathbf{G}^m)^H\right)\right), \qquad (14)$$

where a Neural Network is constructed to provide a regression (instead of interpolation) through the kth row of $(\mathbf{G}^m)^H$ or equivalently the kth column of $\overline{\mathbf{G}^m}$). An alternative is to use

$$\mathbf{q}^{\text{PODNN}}(\boldsymbol{w}) = \sum_{k=1}^{m} \mathbf{h}_k R_k\left(\boldsymbol{w}, \boldsymbol{p}\left(\Sigma^m(\mathbf{G}^m)^H\right)\right). \qquad (15)$$

The version in (15) can be seen to be equivalent to the strategy proposed by Hesthaven and Ubbiali [21], which, in the above notation, is $R_k\left(\boldsymbol{w}, \boldsymbol{p}\left((\mathbf{H}^m)^H\mathbf{D}\right)\right) \approx R_k\left(\boldsymbol{w}, \boldsymbol{p}\left(\Sigma^m(\mathbf{G}^m)^H\right)\right)$ by (10). In both the case of (14) and (15), the neural network regression involves creating a function defined by parameters $\boldsymbol{p}((\mathbf{G}^m)^H)$ or $\boldsymbol{p}(\Sigma^m(\mathbf{G}^m)^H)$, respectively, which enables the prediction of $\mathbf{q}^{\text{PODNN}}$ at low cost for new parameters $\boldsymbol{w}$. To establish the advantages that (15) offers over (14), we consider the ability of neural networks to capture oscillatory functions in the next section.

3.2.4 Capture of oscillatory functions by neural networks

As a simple example, we wish to consider the ability of a feed forward neural network to capture a vector $\mathbf{y}(x) \in \mathbb{R}^{100}$ of oscillatory functions with components

$$y_k(x) = w_k \sin(\kappa_k x), \qquad (16a)$$

$$\kappa_k = \frac{2\pi}{k}, \qquad (16b)$$

where w_k are amplitudes of the functions and $x \in [0, 1]$. To do this, we set up a dataset with N samples in the form of ordered pairs

$$\mathcal{D} := \left\{ (\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), (\mathbf{x}^{(2)}, \mathbf{y}^{(2)}), \ldots, (\mathbf{x}^{(N)}, \mathbf{y}^{(N)}) \right\}, \quad (17)$$

where $\mathbf{x}^{(j)} = \mathrm{x}^{(j)}$ is a scalar input feature in this case and $\mathbf{y}^{(j)}$ is the corresponding vectorial output labels. Using a mean squared error (MSE) as a loss function, the training of a neural network for this problem corresponds to finding parameters $\boldsymbol{p}$ such that

$$\mathcal{L}(\boldsymbol{p}) := \min_{\boldsymbol{p}} \frac{1}{N} \sum_{j=1}^{N} \left| \mathbf{y}^{(j)} - \boldsymbol{R}(\mathrm{x}^{(j)}, \boldsymbol{p}) \right|^2 \quad (18)$$

where $\boldsymbol{R}(\mathrm{x}^{(j)}, \boldsymbol{p})$ describes the regression function provided by a particular Neural Network architecture. We start with a neural network with $L = 1$ layers and $n = 10$ neurons in each layer and a dataset with $N = 101$ entries. In Fig. 2, the results obtained for the cases where a single network is trained to approximate $\boldsymbol{y}(x)$ with $w_k = 1$, separate networks are trained to approximate $y_k(x)$ with $w_k = 1$ individually and a single network is applied to capture $\boldsymbol{y}(x)$ with $w_k = 1/k^2$ (where the oscillatory functions have reducing amplitude) are compared. The MSE loss for the single network with $w_k = 1$ is $\mathcal{L}(\boldsymbol{p}) = 3.21 \times 10^{-1}$, then using multiple networks with $w_k = 1$ the loss for the first oscillatory function is $\mathcal{L}(\boldsymbol{p}) = 4.76 \times 10^{-13}$ and $\mathcal{L}(\boldsymbol{p}) = 4.99 \times 10^{-1}$ for the last oscillatory function. Finally, the MSE loss for a single network with $w_k = 1/k^2$ is $\mathcal{L}(\boldsymbol{p}) = 5.13 \times 10^{-8}$. In this figure, we observe that a single network is unable to capture the oscillatory functions for $w_k = 1$, and there are clearly errors present even in the prediction of $y_1(x)$. On the other hand, training separate networks capture the behaviour of $\boldsymbol{y}(x)$ well with $w_k = 1$. However, if the oscillatory functions are such that $w_k = 1/k^2$, a single network is able to capture the behaviour of the lowest modes accurately.

As $L = 1$ and $n = 10$ may not represent the best choice of hyper parameters, we now apply a Bayes optimisation parameter search [34] to determine the best choices of L_{opt} and n_{opt} in each case so that $\mathcal{L}(\boldsymbol{p})$ is minimised. Other alternative optimisation strategies could be used, but we expect them to lead to similar results. The resulting loss surfaces are shown in Fig. 3 and, in Fig. 4, we repeat the investigation shown in Fig. 2 for the new hyper-parameters obtained. Again, we can state the MSE loss for the optimised single network with $w_k = 1$ ($L_{\text{opt}} = 3$, $n_{\text{opt}} = 16$) as $\mathcal{L}(\boldsymbol{p}) = 2.19 \times 10^{-1}$, then, using multiple networks with $w_k = 1$, the optimised network trained to the first oscillatory function ($L_{\text{opt}} = 2$, $n_{\text{opt}} = 5$) as $\mathcal{L}(\boldsymbol{p}) = 2.09 \times 10^{-10}$ and the optimised network trained to the last oscillatory function ($L_{\text{opt}} = 2$, $n_{\text{opt}} = 14$) as $\mathcal{L}(\boldsymbol{p}) = 4.22 \times 10^{-1}$. Lastly, using an optimised single

network with $w_k = 1/k^2$ ($L_{\text{opt}} = 1$, $n_{\text{opt}} = 12$) gives $\mathcal{L}(\boldsymbol{p}) = 3.12 \times 10^{-8}$. Similar to Fig. 2, Fig. 4 shows that a single network with optimised hyper-parameters is unable to capture well $\boldsymbol{y}(x)$ when $w_k = 1$, but is able to capture the behaviour well when $w_k = 1/k^2$. When networks are trained to predict $y_k(x)$ individually, the prediction is accurate for the first few oscillatory functions but as they become more oscillatory, the loss is higher, as seen on Fig. 3 for the last oscillatory function, which hence means they are unable to capture $y_k(x)$ for higher k.

The computational costs of using a single neural network with $w_k = 1$, multiple networks with $w_k = 1$ and a single network with $w_k = 1/k^2$ are compared in Fig. 5 for both a fixed choice of L and n and also those obtained from the Bayes optimisation L_{opt} and n_{opt}. These, together with the aforementioned plots, illustrate that a single network can perform well and is efficient provided that the oscillatory functions have reducing amplitudes.
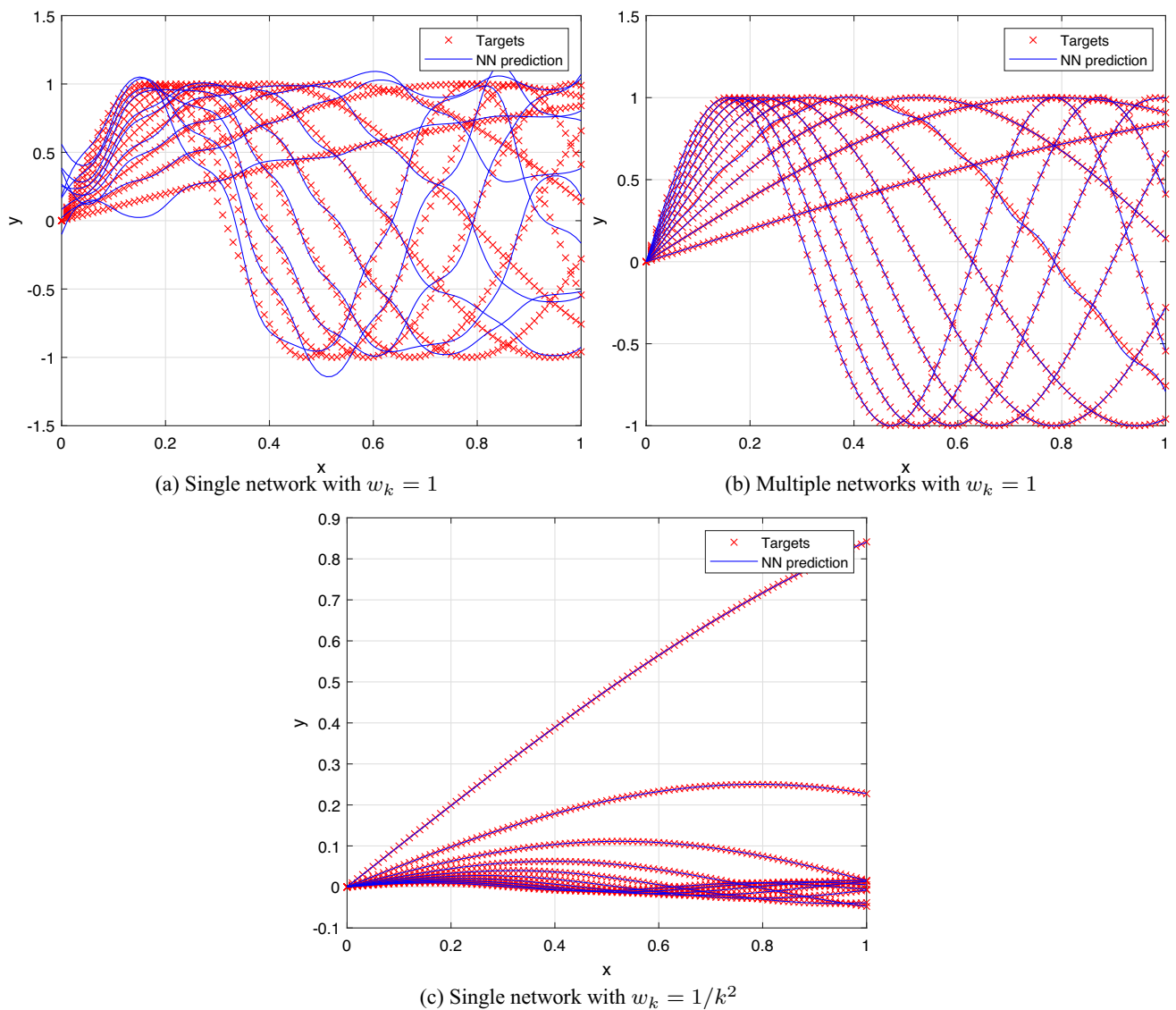
3.2.5 Implications for PODNN

Returning again to (14) and (15), we see that the latter involves $R_k \left(\boldsymbol{w}, \boldsymbol{p} \left(\Sigma^m (\mathbf{G}^m)^H \right) \right)$, where the rows of $\Sigma^m (\mathbf{G}^m)^H$ represent functions that are both oscillatory and reducing in amplitude, whilst the former involves $R_k \left(\boldsymbol{w}, \boldsymbol{p} \left((\mathbf{G}^m)^H \right) \right)$, with the rows of $(\mathbf{G}^m)^H$ being oscillatory but not reducing in amplitude. Thus, by Sect. 3.2.4, the latter case is preferred as it requires less computational effort and requires only the training of a single network. In practice, the datasets we will use for PODNN can be setup in a similar way to $\mathcal{D}$ described above and, for later use, it is useful to introduce

$$\mathbf{x}^{(j)} = \begin{cases} \mathbf{w}_j & \text{for linear spaced snapshots,} \\ \log_{10}(\mathbf{w}_j) & \text{for logarithmically spaced snapshots,} \end{cases} \quad (19)$$

where the $\log_{10}(\mathbf{w}_j)$ should be interpreted as $\left(\log_{10}(\mathbf{w}_{1j}), \log_{10}(\mathbf{w}_{2j}), \ldots, \log_{10}(\mathbf{w}_{Pj}) \right)$ for P parameters and use a similar approach for describing the continuous variable $\boldsymbol{x}(\boldsymbol{w})$. Analogous to (18), we also state the corresponding MSE loss function for this case as

$$\mathcal{L}^{\text{PODNN}}(\boldsymbol{p}) := \min_{\boldsymbol{p}} \frac{1}{N} \sum_{j=1}^{N} \left| \left(\Sigma^m (\mathbf{G}^m)^H \right)_{:,j} \right. \\ \left. - \boldsymbol{R} \left(\mathbf{x}^{(j)}, \boldsymbol{p} \left(\Sigma^m (\mathbf{G}^m)^H \right) \right) \right|^2, \quad (20)$$

where the notation $:, j$ implies all rows and the jth column.

(a) Single network with $w_k = 1$

(b) Multiple networks with $w_k = 1$

(c) Single network with $w_k = 1/k^2$

Fig. 2 Oscillatory functions; feed forward network fit $y_k(x), k = 1, \ldots, 10$ with $L = 1, n = 10$ for **a** single network with $w_k = 1$, **b** multiple networks with $w_k = 1$, **c** single network with $w_k = 1/k^2$

4 Numerical examples

Numerical examples that demonstrate the advantages of the PODP scheme over a PODI scheme in the magneto-mechanical problem described in Sect. 2 has been presented in [17] and so we focus on PODNN and the additional benefits this offers. We begin by describing the test-magnet geometry in Sect. 4.1, which will be the focus of our comparisons presented in this work. Then, in Sect. 4.2, we describe the software used to generate our numerical results followed by a description of the discretisation in Sect. 4.3. We present predictions of resonant modes in the Ohmic power and kinetic energy curves using different parameter spaces for the test magnet geometry in Sect. 4.4 followed by the predictions of the response in the same outputs of interest for different

external vibrational loading conditions in Sect. 4.5. Then, in Sect. 4.6, the PODNN technique is then also used to predict the response in outputs of interest when model parameters are uncertain.

4.1 Test magnet geometry

Our focus will be a simplified MRI configuration called the test magnet problem, which is rotationally symmetric, and was previously considered in [12]. For the simplified model, only the rotational symmetric z-gradient coil is considered and the non-rotationally symmetric x and y gradient coils are removed. If desired, this allows the problem to be reduced to a axisymmetric problem although we will model it as one
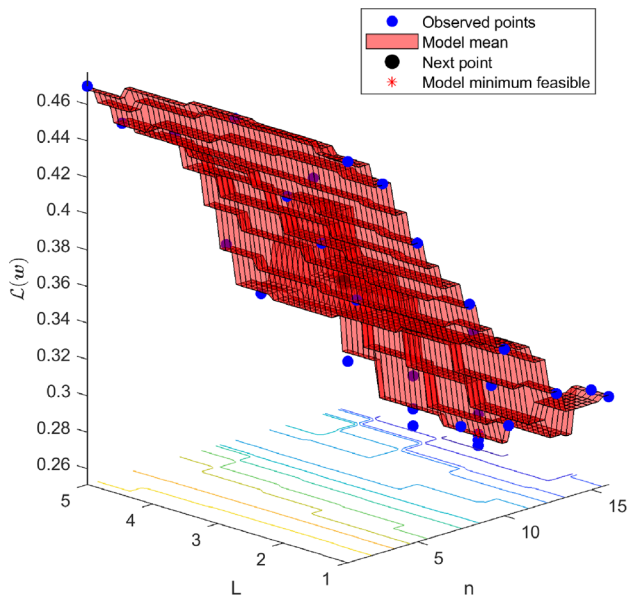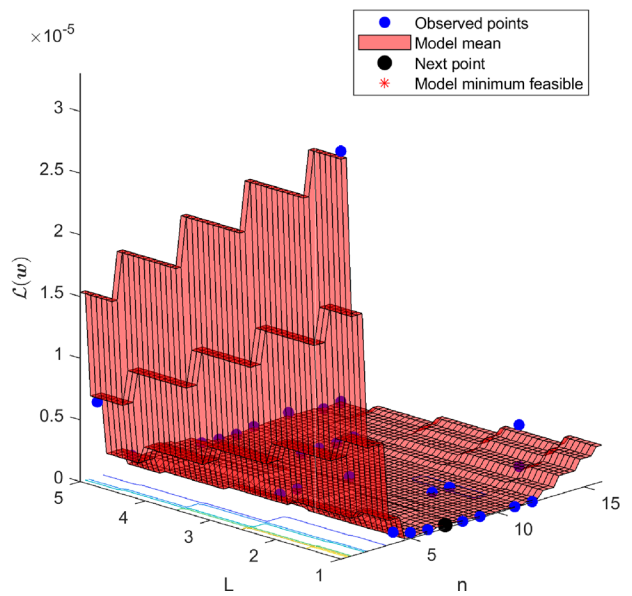
(a) Single network with $w_k = 1$
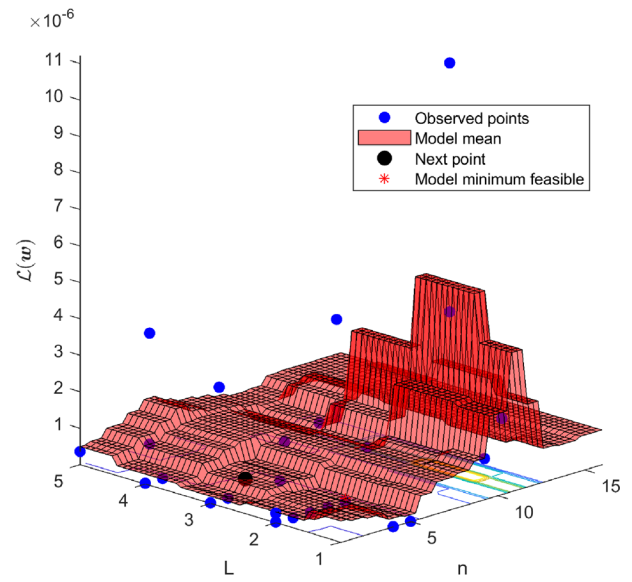
(b) Multiple networks with $w_k = 1$; first oscillatory function
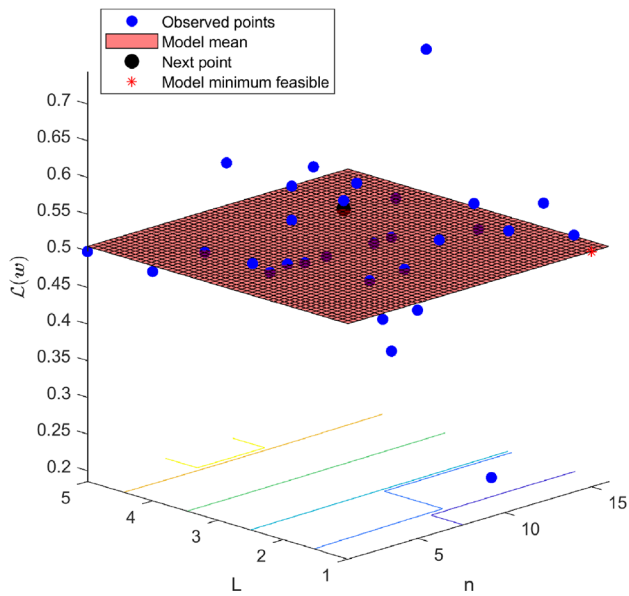
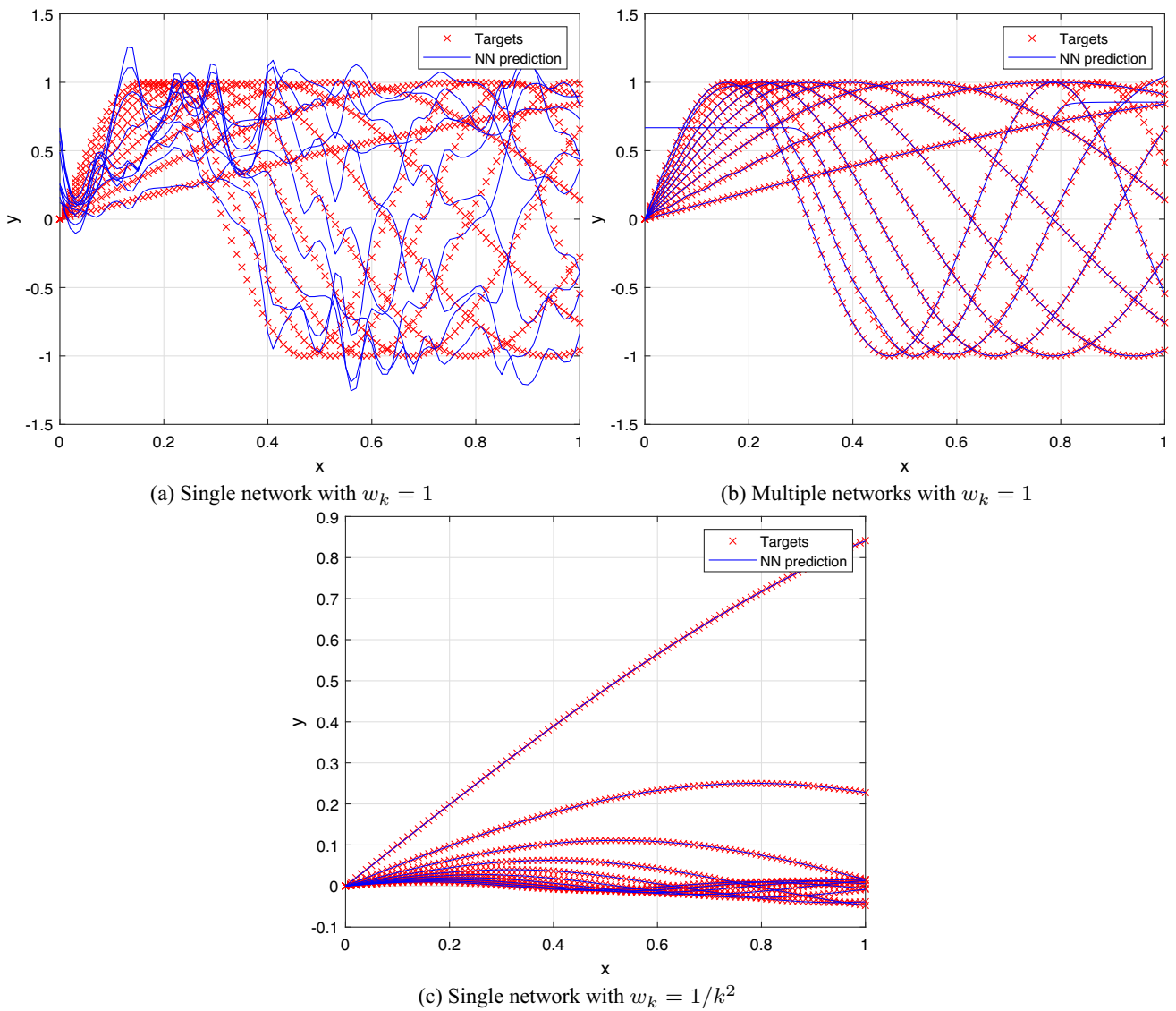(c) Multiple networks with $w_k = 1$; last oscillatory function

(d) Single network with $w_k = 1/k^2$

Fig. 3 Oscillatory functions; the MSE loss function (18) as a function of $L = 1, \ldots, 5$ and $n = 1, \ldots, 16$ used in the Bayes optimisation for **a** single network with $w_k = 1$, **b** multiple networks with $w_k = 1$ for the first oscillatory function, **c** multiple networks with $w_k = 1$ for the last oscillatory function and **d** single neural network $w_k = 1/k^2$

(a) Single network with $w_k = 1$

(b) Multiple networks with $w_k = 1$

(c) Single network with $w_k = 1/k^2$

Fig. 4 Oscillatory functions; feed forward network fit $y_k(x)$, $k = 1, \ldots, 10$ with optimum hyper-parameters for **a** single network with $w_k = 1$, **b** multiple networks with $w_k = 1$ and **c** single network with $w_k = 1/k^2$

(a) Fixed network $L = 1, n = 10$

(b) Bayes optimisation $L = L_{opt}, n = n_{opt}$
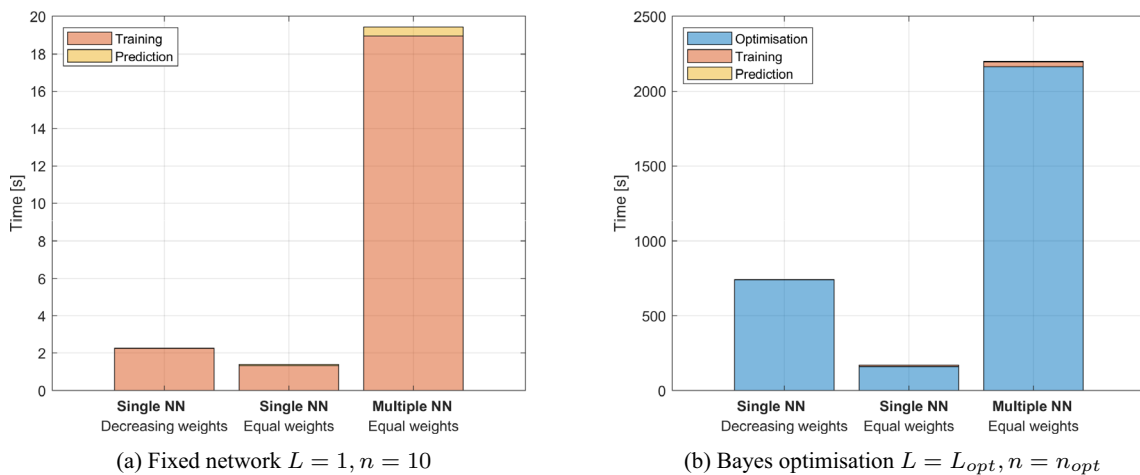
Fig. 5 Oscillatory functions; timing comparison between a single network and multiple networks based on simulations **a** in Fig. 2 and **b** Fig. 4

quarter of the full three-dimensional geometry.[1] An illustration of this MRI configuration can be seen on Fig. 6, which shows a typical 2D cross-section and 3D view of the general setting. In this figure, the conducting shields (4 K and 77 K) and the outer vacuum chamber (OVC) are shown in different shades of blue and these will deform under the presence of electromagnetic stresses generated due to the presence of eddy current currents in the conducting components. The main coils and gradient coils are shown in shades of red. The dimensions, exciting currents and materials of this problem are commercially sensitive, but indicative values are provided in [35]. Unless otherwise stated, we assume that $\boldsymbol{\mathcal{U}}_D^{\mathrm{AC}} = \boldsymbol{0}$ in the boundary condition $\boldsymbol{\mathcal{U}}^{\mathrm{AC}} = \boldsymbol{\mathcal{U}}_D^{\mathrm{AC}}$ on $\partial\Omega_C^D$ in (2e) to reflect the fact that the magnet geometry is fixed in position. Furthermore, the transmission problem described by (2) must be supplemented by the symmetry boundary conditions [17]

$$\boldsymbol{\sigma}^m(\boldsymbol{\mathcal{U}}^{\mathrm{AC}})\boldsymbol{n} = \boldsymbol{0}, \tag{21a}$$

$$\boldsymbol{n} \cdot \boldsymbol{\mathcal{U}}^{\mathrm{AC}} = 0, \tag{21b}$$

$$\boldsymbol{n} \times \boldsymbol{\mathcal{A}}^{\mathrm{AC}} = \boldsymbol{0}, \tag{21c}$$

that are to be imposed on the $x = 0$ and $y = 0$ planes.

As outputs of interest for the MRI configurations involve the output (dissipated) power $P_\Omega^0(\omega, \boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}}, \boldsymbol{B}_0^{\mathrm{DC}}, \boldsymbol{\mathcal{U}}^{\mathrm{AC}})$ and the kinetic energy $E_\Omega^k(\omega, \boldsymbol{\mathcal{U}}^{\mathrm{AC}})$ which are defined as:

$$
\begin{aligned}
P_{\Omega_C}^0 &:= \frac{1}{2}\int_{\Omega_C}\gamma\left|\boldsymbol{\mathcal{E}}^{\mathrm{AC}}\right|^2\mathrm{d}\Omega \\
&= \frac{1}{2}\int_{\Omega_C}\gamma\left|-\mathrm{i}\omega\boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}} + \mathrm{i}\omega\boldsymbol{B}_0^{\mathrm{DC}} \times \boldsymbol{\mathcal{U}}^{\mathrm{AC}}\right|^2\mathrm{d}\Omega,
\end{aligned} \tag{22a}
$$

$$E_{\Omega_C}^k := \frac{\omega^2}{2}\int_{\Omega_C}\rho\left|\boldsymbol{\mathcal{U}}^{\mathrm{AC}}\right|^2\mathrm{d}\Omega, \tag{22b}$$

as a function of ω. In practice, our approximation to these will be evaluated for a particular conducting shield Ω_C. At an intermediate stage, we will also consider plots of approximations to $\|\boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}}(\omega)\|_{L^2(\Omega_C)}$ as a function ω, which is defined as[2]

$$\|\boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}}(\omega)\|_{L^2(\Omega_C)} := \sqrt{\int_{\Omega_C}\overline{\boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}}(\omega)} \cdot \boldsymbol{\mathcal{A}}_\varepsilon^{\mathrm{AC}}(\omega)\mathrm{d}\Omega}. \tag{23}$$

---

[1] Of course, if desired, a different fraction of the full geometry could be considered provided the correct symmetry boundary conditions are imposed.

[2] Reminder that $\overline{\boldsymbol{a}}$ represents the complex conjugate of a vector field $\boldsymbol{a}$.

4.2 Software

The magneto-mechanical model will be simulated using a simulation tool that has been extended from the previously developed higher order finite element and reduced order model tool developed by our group to simulate the problems described in [12, 17, 18]. The latest additions include the extension to handle the PODI and PODNN reduced order models for large parameter space $\boldsymbol{w}$, uncertainty quantification features and non-zero Dirichlet displacements of the shields. The neural networks were implemented with MATLAB's Deep Learning Toolbox [36], which has similar functionalities to popular python packages such as tensorflow [37], scikit-learn [38] or pytorch [39], since our simulation tool was also written in MATLAB [40]. In particular, the feedforwardnet tool was used to build multilayer perceptron (MLP) neural networks. In the following, we focus on numerical results for the new reduced order models as the accuracy of the linearised strategy in (2) and the hp-finite element discretisation has already been established in [12].

For PODNN, the split between $\mathcal{D}_{\mathrm{train}}$ and $\mathcal{D}_{\mathrm{test}}$ is 85% and 15%, respectively, and is applied throughout. Prior to training using $\mathcal{D}_{\mathrm{train}}$, values were normalised to the standard deviation and mean of $(\Sigma^m(\mathbf{G}^m)^H)_{1,:}$. Additionally, the form of the activation function followed a tan-sigmoid function. To solve the minimisation problems (20) the Bayesian regularisation solver with the option to use the Jacobian for calculations was employed with the maximum number of epochs set to 5000 and min grad$= 10^{-10}$. Bayes optimisation, implemented with MATLAB's Statistics and Machine Learning Toolbox [41], is applied to determine the optimum number of layers L_{opt} and neurons per layer n_{opt}, in a similar way to Sect. 3.2.4 with other hyper-parameters set as per [36]. All computations were performed on a workstation that comprised of a 12-core Intel Xeon W-2265 Processor with 128GB (8×16GB) RAM and a NVIDIA Quadro RTX 4000 8GB graphical processing unit.

4.3 Discretisation

For this problem, Ω_C^c is truncated to form a finite computational domain Ω by excluding the region outside of a suitably sized cylinder. We discretise Ω by an unstructured mesh of 33 805 tetrahedral elements and employ a discretisation with $q = 3$ and $p = 3$ order elements for $\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{\mathrm{AC}}$ and $\boldsymbol{\mathcal{U}}_{hp}^{\mathrm{AC}}$, respectively. To dampen the resonant peaks in the model, we employ a fixed damping ratio of $\xi = 2 \times 10^{-3}$ and to circumvent the Coulomb gauge we use a regularisation parameter of $\varepsilon = 10^{-4}$. This choice of discretisation has been shown to provide accurate full order model solutions to the test magnet geometry over the parameters of interest [12].

Fig. 6 Test magnet problem with z (longitudinal) gradient coil; illustration of the components of the problem **a** in the axisymmetric meridian plane and **b** 3D view

(a) 2D cross-section

(b) 3D view
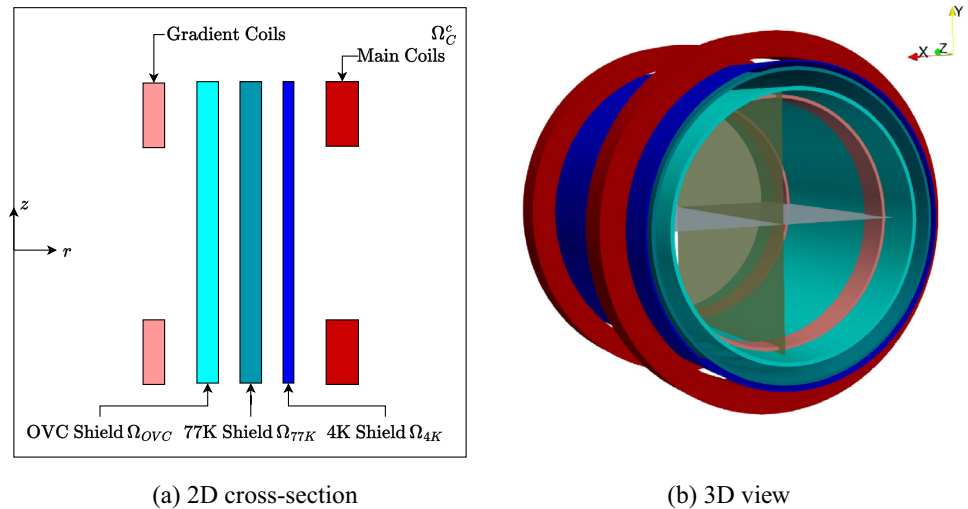
4.4 Prediction of resonant modes

4.4.1 Offline stage

We first consider the effects of different choices of snapshots and the behaviour it has on the decay of Σ_{ii}. Considering the case of $P = 1$ and $\mathbf{w} = (\omega) = 2\pi(f)$ with $5 \leq f \leq 5000$ Hz, we will explore five different selections of snapshot frequencies $\mathbf{w}_i = 2\pi(f_i)$ and decide which will be most appropriate for our problem. From previous analysis in [17], we know that the test-magnet geometry is sensitive to lower frequencies and, hence, may require more frequencies within that lower range.

Secondly, we consider the choice of snapshot parameters for case of $P = 2$ with $\mathbf{w} = (\omega, \gamma_0^{\mathrm{OVC}})$ where $0.1 \leq \gamma_0^{\mathrm{OVC}} \leq 10$ is a non-dimensional factor used to scale the conductivity in the OVC such that

$$\gamma^{\mathrm{OVC}} = \gamma_0^{\mathrm{OVC}}\gamma_{\mathrm{ref}}^{\mathrm{OVC}}, \tag{24}$$

where $\gamma_{\mathrm{ref}}^{\mathrm{OVC}}$ is a reference value for conductivity of the shield.

In [17] a piece-wise linear choice for the snapshot frequencies was made, but it is not clear that this is the best choice. Therefore, we compare the performance of selecting the snapshots f_i in a linear, piece-wise linear and logarithmic fashion in the following.

Linear The linear snapshot frequencies f_i in Hz are chosen such that:

$$f_i = 5 + \frac{5000 - 5}{N_{\mathrm{fs}} - 1}(i - 1), \tag{25}$$

for $i = 1, \ldots, N_{\mathrm{fs}}$, where N_{fs} is the total number of snapshot frequencies and, in the case of $P = 1$, $N_{\mathrm{s}} = N_{\mathrm{fs}}$.

Piece-wise linear The piece-wise linear snapshot frequencies f_i are chosen by first introducing the splitting $N_{\mathrm{fs}} = N_{\mathrm{fs}}^{(1)} + N_{\mathrm{fs}}^{(2)}$, which are then used to calculate the spacings

$$\Delta f^{(1)} = \frac{1000 - 5}{N_{\mathrm{fs}}^{(1)} - 1},$$
$$\Delta f^{(2)} = \frac{5000 - 1000}{N_{\mathrm{fs}}^{(2)} - 1}, \tag{26}$$

where the numerators are determined the range of frequencies that we intend to split by. The frequencies f_i are then constructed as

$$f_i = \begin{cases} 10 + (i - 1)\Delta f^{(1)} & \text{for } 1 \leq i \leq N_{\mathrm{fs}}^{(1)}, \\ 1000 + (i - N_{\mathrm{fs}}^{(1)})\Delta f^{(2)} & \text{for } N_{\mathrm{fs}}^{(1)} < i \leq N_{\mathrm{fs}}^{(2)}, \end{cases} \tag{27}$$

with $N_{\mathrm{fs}}^{(1)}$ and $N_{\mathrm{fs}}^{(2)}$ chosen as in [17].

Logarithmically spaced For logarithmically-spaced frequency snapshots, f_i is chosen such that

$$\log_{10}(f_i) = \log_{10}(5) + \frac{\left(\log_{10}(5000) - \log_{10}(5)\right)}{N_{\mathrm{fs}} - 1}(i - 1), \tag{28}$$

with $i = 1, \ldots, N_{\mathrm{fs}}$.

For $N_{\mathrm{fs}} = 180$, we compare the distribution of f_i according to (25), (27) and (28) in Fig. 7, with the first being equally spaced and second and third showing a clustering of samples for smaller frequencies as expected. Making these different choices of snapshots to generate full order model solutions $\mathbf{q}(\mathbf{w}_j)$, creating $\mathbf{D}$ in (9), and then applying the TSVD in (10) leads to the decay of Σ_{ii} shown in Fig. 8. Comparing the different choices, the log-spaced snapshots in (28) leads to fastest decay of the singular values. Setting $TOL = 10^{-6.5}$ (due to the decay in Fig. 8 indicating that the decay is no
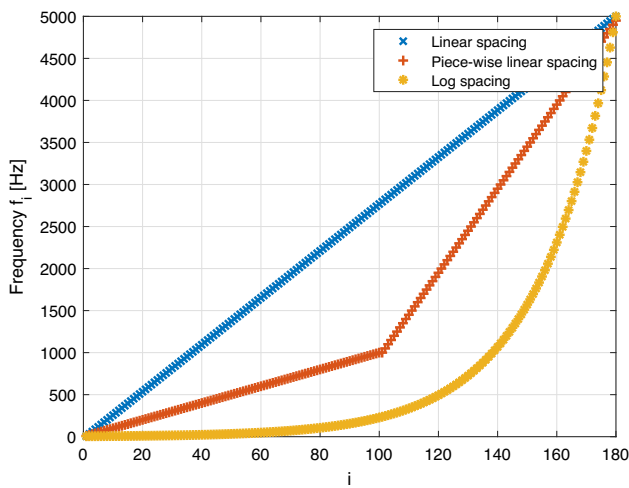
Fig. 7 Test magnet problem with z (longitudinal) gradient coil; Offline stage for $N_{\text{fs}} = 180$. A comparison of linear, piece-wise linear and log-spaced snapshots for $5 \leq f_i \leq 5000$ Hz
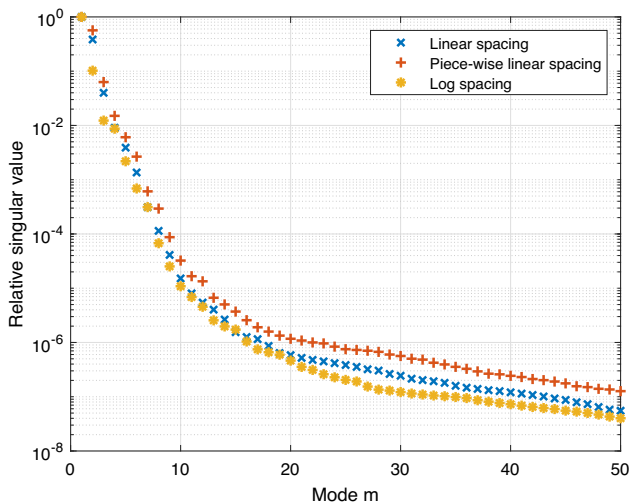
Fig. 8 Test magnet problem with z (longitudinal) gradient coil; Offline stage for $P = 1$, $N_{\text{fs}} = 180$. Relative singular values Σ_{mm}/Σ_{11} obtained using linear, piecewise-linear and log-spacing for f_i

longer exponential after this point), then, by choosing m according to the largest value for which $\Sigma_{mm}/\Sigma_{11} < \text{TOL}$, we obtain $m = 34$ for (25), $m = 28$ for (27) and $m = 20$ for (28), hence, leading to the smallest reduced order model when the log-spaced frequencies are used. An additional benefit of (28) over (27) is that we do not need to define a priori where to split the piece-wise linear spacing of frequencies.

The complex amplitudes $|(\Sigma^m(\mathbf{G}^m)^H)_{ij}|$ shown as lines for a fixed $i = 1, \ldots, 10$ against f_j, $j = 1, \ldots, N_{\text{fs}}$, are included on Fig. 9 for the different choices of snapshots. The general trend for linear, piece-wise linear and logarithmically spaced snapshots is that the lower modes (e.g. $i = 1, 2$) are simple smooth functions that are not very oscillatory, but, as i increases, the modes become increas-

ingly more oscillatory. For higher modes, the oscillatory nature is very pronounced for small f_j and Fig. 9d indicates there are additional benefits by showing $|(\Sigma^m(\mathbf{G}^m)^H)_{ij}|$ against f_j, $j = 1, \ldots, N_{\text{fs}}$ on a log-log graph indicating that $\mathbf{x}^{(j)} = \log_{10} \omega_j = \log_{10}(2\pi f_j)$ may be a good choice in this case. Henceforth, we choose the frequency snapshots according to (28) and use $\mathbf{x}^{(j)} = \log_{10} \omega_j$, but in the following we will not only consider $N_{\text{fs}} = 180$, but also cases of $N_{\text{fs}} = 13, 23, 45, 90$ to reduce computational effort needed for the off-line stage. The snapshots $N_{\text{fs}} = 45, 90$ follow a similar decay and lead to models of size $m = 20$, but, the $N_{\text{fs}} = 13, 23$ cases are seen to decay faster and henceforth a reduced model of size $m = 18$. However, as seen in Sect. 4.4.2 for $N_{\text{fs}} = 13$, it will not be a sufficient amount of snapshots to accurately capture the solution of the problem.

Having found that a logarithmic spacing f_i is preferable, we now consider how best to choose the conductivity factor snapshots $\gamma_{0,i}^{\text{OVC}}$. Given that γ and ω appear as a product in (2a) we argue that the best choice for the spacing of the snapshots $\gamma_{0,i}^{\text{OVC}}$ is also logarithmic. Hence, we propose

$$\log_{10}(\gamma_{0,i}^{\text{OVC}}) = \log_{10}(0.1) + \frac{\left(\log_{10}(10) - \log_{10}(0.1)\right)}{N_{\text{cs}} - 1} i,$$
(29)

with $i = 1, \ldots, N_{\text{cs}}$ and N_{cs} is the chosen number of snapshots for the conductivity factor. This means that for the case of $P = 2$ we have $N_s = N_{\text{cs}} N_{\text{fs}}$ snapshots in total. The decay of singular values, shown in Fig. 8 for $P = 1$, can also be found for $P = 2$ and shows a similar behaviour. In this case, however, setting $TOL = 10^{-6.5}$ corresponds to a reduced model of size $m = 44$ if $N_s = 23 \times 23$ and, hence, still offers a considerable saving.
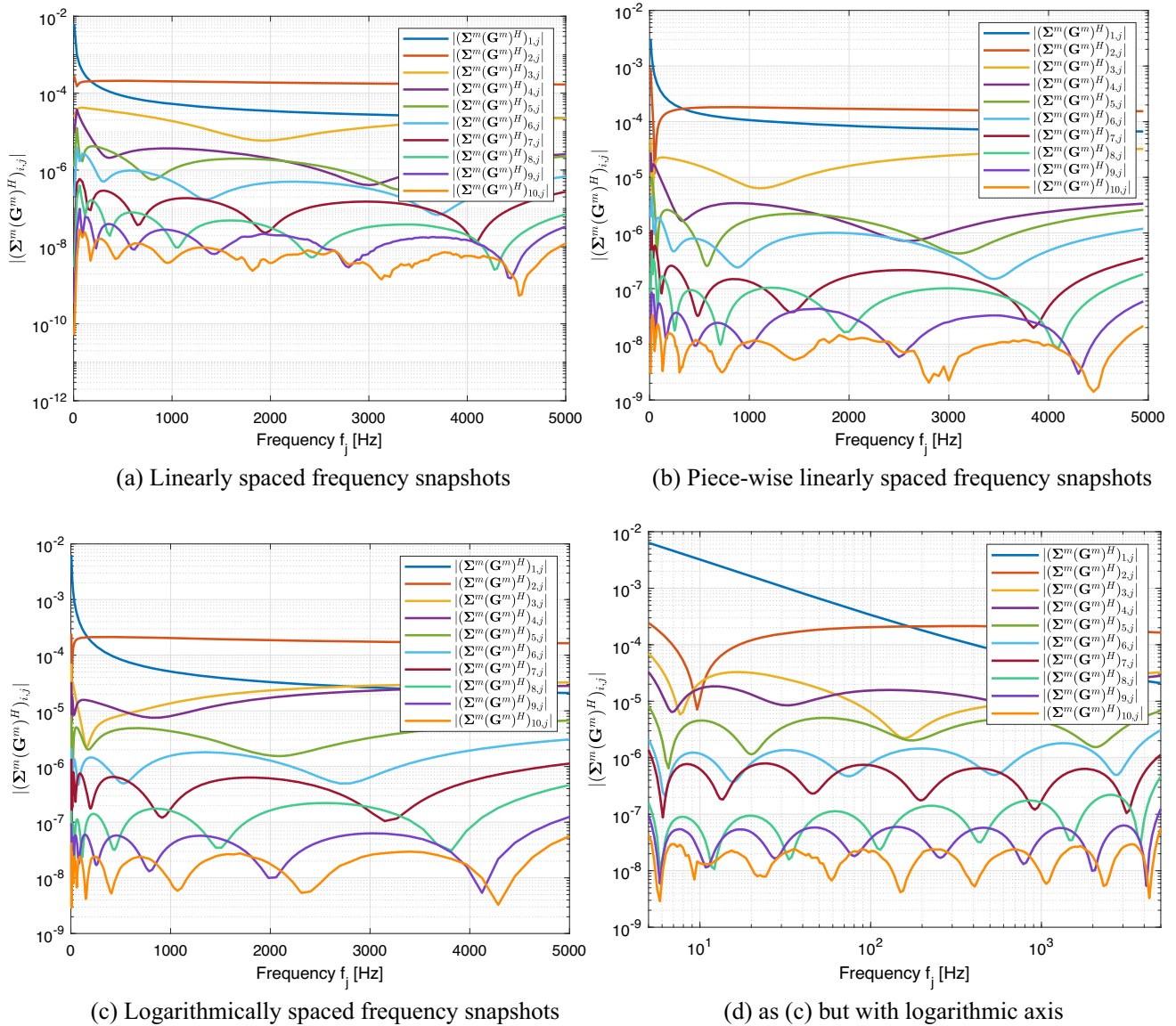
4.4.2 Online stage

Predictions for $\mathbf{w} = (\omega) = 2\pi(f)$ with $P = 1$ In the online stage of either PODI or PODNN we can easily evaluate (23) and (22) for any frequency of interest. However, for the purpose of visualisation, we will show results for the case where the ROMs are evaluated for the following output frequencies in Hz

$$f_i = 15 + \frac{5000 - 15}{N_{\text{fo}} - 1} i,$$
(30)

for $i = 1, \ldots, N_{\text{fo}}$, with $N_{\text{fo}} = 500$, unless otherwise stated.

We use Bayes optimisation as in Sect. 3.2.4 to find L_{opt} and n_{opt} using the search space $L = 1, \ldots, 5$ and $n = 1, \ldots, 16$. The resulting MSE loss $\mathcal{L}^{\text{PODNN}}(\mathbf{p})$ is shown on Fig. 10a for $N_{\text{fs}} = 180$. As previously observed in Fig. 3, a similar $\mathcal{L}^{\text{PODNN}}(\mathbf{p})$ trend is observed with $\mathcal{L}^{\text{PODNN}}(\mathbf{p}) < 10^{-6}$ for $n > 5$ independent of L in this case. This investiga-

(a) Linearly spaced frequency snapshots

(b) Piece-wise linearly spaced frequency snapshots

(c) Logarithmically spaced frequency snapshots

(d) as (c) but with logarithmic axis

Fig. 9 Test magnet problem with z (longitudinal) gradient coil; offline stage for $P = 1$, $N_{fs} = 180$. Complex amplitudes $|(\Sigma^m(\mathbf{G}^m)^H)_{ij}|$ as lines for a fixed $i = 1, \ldots, 10$ against f_j, $j = 1, \ldots, N_{fs}$, for **a** linearly spaced f_j, **b** piece-wise linearly spaced f_j, **c** log-spaced f_j and **d** against log f_j for log-spaced f_j

tion will be repeated for $N_{fs} = 13, 23, 45, 90$ as each will give rise to a slightly different $|(\Sigma^m(\mathbf{G}^m)^H)_{ij}|$. We additionally show in Fig. 10b the ability of the optimised network to fit $|(\Sigma^m(\mathbf{G}^m)^H)_{ij}|$ for a fixed $i = 1, 9, 19$ against f_j, $j = 1, \ldots, N_{fs}$ for $N_{fs} = 180$. As expected, the higher modes are more oscillatory and, hence, the fitting becomes comparatively worse for these modes. For example, $i = 9$ is shown to have reasonable fitting at higher frequencies. However, for $i = 19$, the network is partially able to capture $|(\Sigma^m(\mathbf{G}^m)^H)_{ij}|$ against f_j, $j = 1, \ldots, N_{fs}$. Due to the fact that $\Sigma^m_{19,19}/\Sigma^m_{1,1} \approx 5.9 \times 10^{-7}$, which is close enough to the TOL chosen, the partial fitting of this mode is well enough for a reasonable prediction as seen in Fig. 11.

To illustrate the interdependence between N_{fs} and some of the hyper-parameters of the MLP, we show in Fig. 11 the behaviour of $\|\mathbf{\mathcal{A}}^{AC}_{\varepsilon,hq}(\omega)\|_{L^2(\Omega_C)}$ as a function of $\omega = 2\pi f$ obtained by applying (15) and then (5a) for different N_{fs} and different values of the hyper-parameters $L = L_{opt}$ and $n = n_{opt}$ found by the Bayes optimisation with their respective $\mathcal{L}^{PODNN}(\mathbf{p})$. The different lines in the figure correspond to the cases where Ω_C is considered to be the 4K, 77K shields and the OVC in turn, also shown are the corresponding solution snapshot solutions for $\|\mathbf{\mathcal{A}}^{AC}_{\varepsilon,hq}(\omega)\|_{L^2(\Omega_C)}$. For $N_{fs} = 23, 45, 90, 180$ we get good agreement for all shields, but not in the case of the 4K shield with $N_{fs} = 13$.

Having now established that the performance of PODNN is similar for $N_{fs} > 23$ snapshots, we show, in Fig. 12 a com-

Fig. 10 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 1$. Showing **a** MSE loss as a function of $L = 1, \ldots, 5$ and $n = 1, \ldots, 16$ and **b** comparisons of $|(\Sigma^m (\mathbf{G}^m)^H)_{ij}|$, $j = 1, 9, 19$ for $N_{fs} = 180$

(a) MSE loss

(b) $|(\Sigma^m (\mathbf{G}^m)^H)_{ij}|$
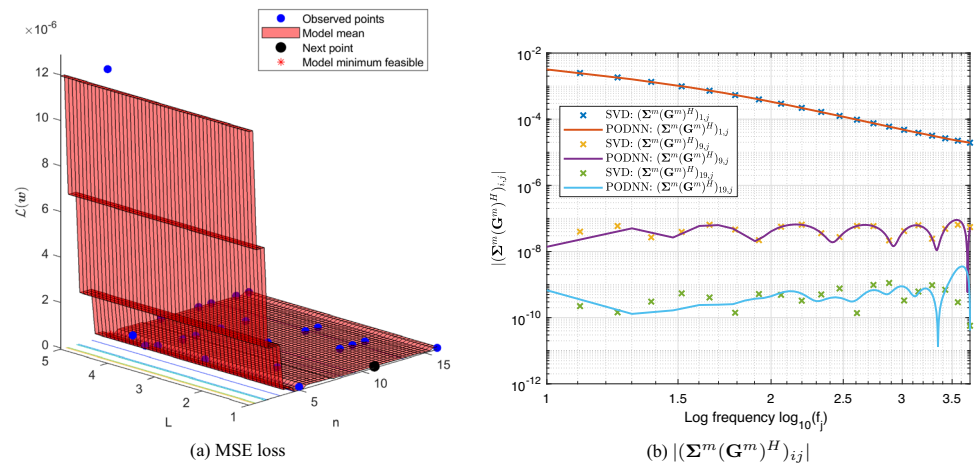
parison between the accuracy and computational efficiencies of PODI, PODP and PODNN. As an approximate measure of accuracy, we introduce

$$
\epsilon(\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC}\|_{L^2(\Omega_C)})
$$

$$
:= \sqrt{\frac{\sum_{j=1}^{N_{fs}} \left(\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, POD}(\omega_j)\|_{L^2(\Omega_C)} - \|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, Full}(\omega_j)\|_{L^2(\Omega_C)} \right)^2}{\sum_{j=1}^{N_{fs}} \|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, Full}(\omega_j)\|_{L^2(\Omega_C)}^2}},
$$

$$
(31)
$$

where $\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, POD}\|_{L^2(\Omega_C)}$ and $\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, Full}\|_{L^2(\Omega_C)}$ are the evaluations of (23) with the reduced techniques and the full order model, respectively. A more precise calculation would involve the

$$
\sqrt{\sum_{j=1}^{N_{fs}} \|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, POD}(\omega_j) - \boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, Full}(\omega_j)\|_{L^2(\Omega_C)}^2 / \sum_{j=1}^{N_{fs}} \|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC, \, Full}(\omega_j)\|_{L^2(\Omega_C)}^2}
$$

which does not average out possible spatial errors. However, we have high confidence in the PODNN solutions given the results in Fig. 11 and those presented later in this section. To compare computational efficiencies, we consider the wall clock time of the online stages of the calculation of $\boldsymbol{q}^{POD}$ for the POD techniques. We see that PODP is most accurate with $\epsilon(\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC}\|_{L^2(\Omega_C)}) \approx 10^{-7}$ independent of the N_{fs} considered. PODI performs worse than PODP, especially for lower N_{fs} by a few orders of magnitude but improves for larger N_{fs}. PODNN does not perform as well as the PODP case but for small N_{fs}, it is shown to perform better than PODI. However, the decrease in accuracy may be due to the tolerances chosen within the network but the accuracy achieved by PODNN is sufficient for this practical application. Additionally, we show the online computational expense for the $\epsilon(\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC}\|_{L^2(\Omega_C)})$, with PODNN showing the best in terms of performance. In Figure 10 and Figure 15 of [17], the total online and offline costs of PODI and a full order solver as well
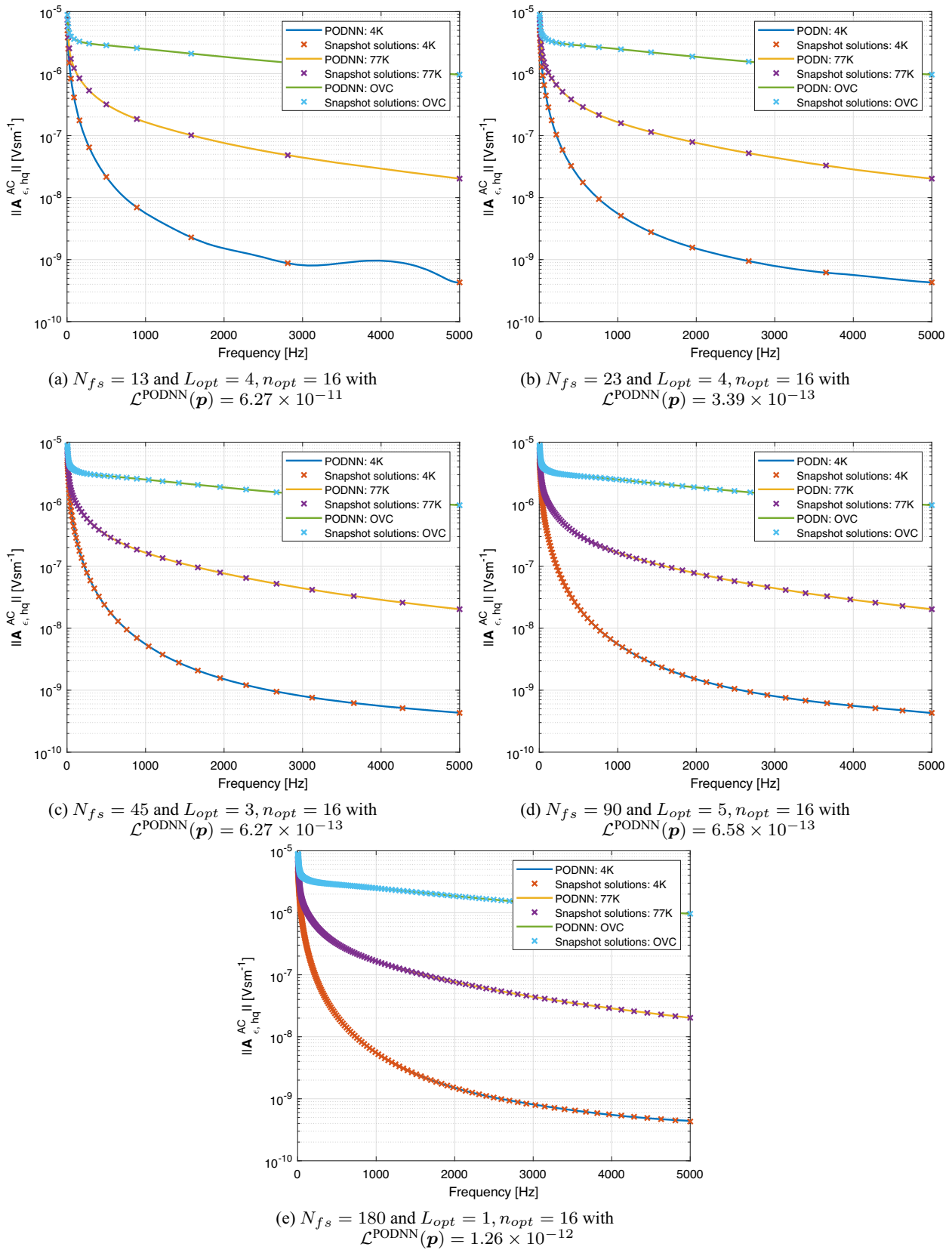
as PODP with a full order solver were compared for the same problem (but using a different machine). In both cases, PODP and PODI offer similar computational costs, but just like in Fig. 12b, the accuracy of PODP was substantially better than PODI. Discounting training and optimisation time, the total costs for PODNN would be similar to PODP and PODI shown in [17] if the same machine was used. However, the cost of optimisation and training can be significant as Fig. 12c shows. Whilst the training and optimisation forms part of the offline stage in PODNN, this is significantly more expensive than the offline stages of PODP and PODI if included in the costs.

Using the most efficient and accurate PODNN strategy with $N_{fs} = 23$, we show, in Fig. 13, the dissipated power $P_{\Omega_C}^0(\omega)$ and kinetic energy $E_{\Omega_C}^k(\omega)$ obtained by additionally solving (8) for each output frequency f_i of interest using the reduced order model (15), field representations (5) and applying (22). The results are in good-agreement with the full order model and, importantly, accurately predict the resonant modes.
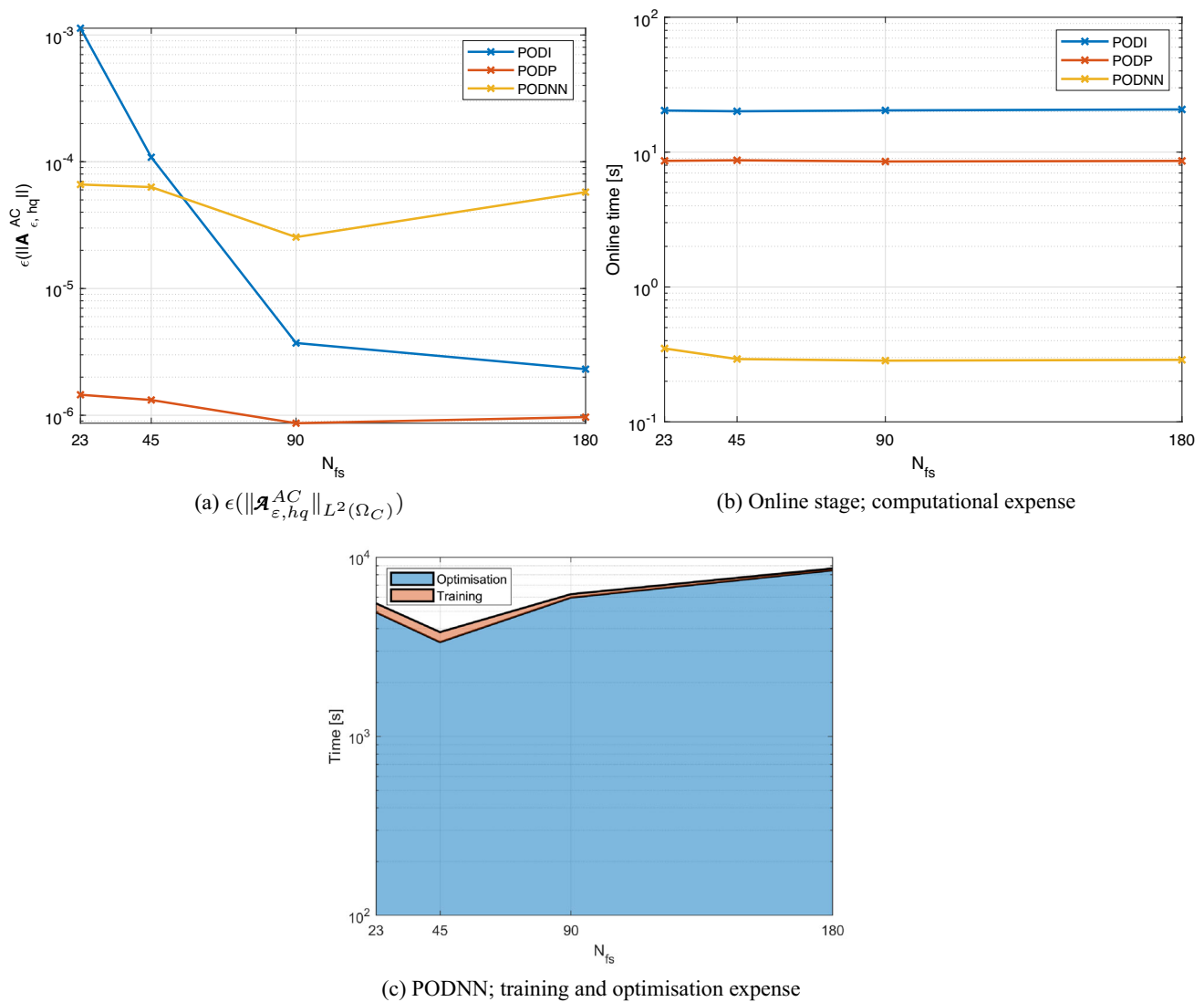
To emphasise that this is a 3D magneto-mechanical problem, we visualise the displacement and magnetic flux density fields obtained from solving the PODNN using the aforementioned network within a conducting domain. Figure 14 showcases both the variation in the magnitudes of the real displacement field and the real magnetic flux density field in the outer 4 K shield at output frequencies of $f = 1000$ Hz and $f = 5000$ Hz. Bands of higher and lower magnitudes can be seen in the displacement field at 5000 Hz, but at 1000 Hz, there is a higher displacement near the outer edges of the domain. A similar effect can be seen with the real magnetic flux density for 1000 Hz but at 5000 Hz, the domain shows almost uniform field of low displacement in the conducting domain.

Predictions for $\boldsymbol{w} = (\omega, \gamma_0^{OVC}) = (2\pi f, \gamma_0^{OVC})$ with $P = 2$

In the online stage of either PODI or PODNN we can easily evaluate (23) and (22) for any frequencies and conductivity

(a) $N_{fs} = 13$ and $L_{opt} = 4, n_{opt} = 16$ with
$\mathcal{L}^{\text{PODNN}}(\boldsymbol{p}) = 6.27 \times 10^{-11}$

(b) $N_{fs} = 23$ and $L_{opt} = 4, n_{opt} = 16$ with
$\mathcal{L}^{\text{PODNN}}(\boldsymbol{p}) = 3.39 \times 10^{-13}$

(c) $N_{fs} = 45$ and $L_{opt} = 3, n_{opt} = 16$ with
$\mathcal{L}^{\text{PODNN}}(\boldsymbol{p}) = 6.27 \times 10^{-13}$

(d) $N_{fs} = 90$ and $L_{opt} = 5, n_{opt} = 16$ with
$\mathcal{L}^{\text{PODNN}}(\boldsymbol{p}) = 6.58 \times 10^{-13}$

(e) $N_{fs} = 180$ and $L_{opt} = 1, n_{opt} = 16$ with
$\mathcal{L}^{\text{PODNN}}(\boldsymbol{p}) = 1.26 \times 10^{-12}$

Fig. 11 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 1$, comparison of $\|\boldsymbol{\mathcal{A}}^{\text{AC}}_{\varepsilon,hq}(\omega)\|_{L^2(\Omega)}$. Results for **a** $N_{\text{fs}} = 13$, **b** $N_{\text{fs}} = 23$, **c** $N_{\text{fs}} = 45$, **d** $N_{\text{fs}} = 90$ and **e** $N_{\text{fs}} = 180$

(a) $\epsilon(\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC}\|_{L^2(\Omega_C)})$

(b) Online stage; computational expense

(c) PODNN; training and optimisation expense

Fig. 12 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 1$, comparisons of **a** $\epsilon(\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC}\|_{L^2(\Omega_C)})$, **b** online stage computational expense and **c** PODNN training and optimisation expense for $N_{fs} = [23, 45, 90, 180]$

factors of interest. But, for the purpose of visualisation, we evaluate the ROMs at the frequencies stated in (30) and at the following non-dimensional conductivity factors

$$\gamma_{0,i}^{OVC} = [0.35, 0.75, 3.5, 7.5], \tag{32}$$

for $i = 1, \ldots, N_{co}$, with $N_{co} = 4$, unless otherwise stated. This means that we evaluate the ROM for a total of $N_{fo}N_{co}$ sets of output parameters.

A similar investigation to that shown in Fig. 11 was repeated for the two-parameter case, again using the Bayes optimisation but with a search space $L = 1, \ldots, 10$ and $n = 1, \ldots, 32$. It should be noted that the optimisation time for this case takes considerably longer than for $P = 1$ and is more costly than the PODP method, with a time of 5.29×10^5s, which is several days. It was found that,

in this case, using a $L_{opt} = 3$, $n_{opt} = 32$ network for $N_{fs} = 23$ and $N_{cs} = 23$ produces accurate results provided that `min grad`$=10^{-10}$ and the number of singular values that is retained is $m = 44$ using the same TOL $= 10^{-6.5}$ as carried out for $P = 1$. As an example, we show the behaviour of $\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{AC}(\omega)\|_{L^2(\Omega_C)}$ as a function of $\omega = 2\pi f$ with lines corresponding to different conductivity factors $\gamma_0^{OVC} = 0.35, 0.75, 3.5, 7.5$, which is obtained by applying (15) and then (5a) in Fig. 15, with an achieved MSE loss of $\mathcal{L}^{PODNN}(p) = 3.56 \times 10^{-13}$. Results are shown for the 4K, 77K shields and the OVC with a comparison against the snapshot solutions. The graphs shown in the left hand column correspond to cases where the conductivity factors are lower than unity, while those in the right hand column correspond to showing the predictive the ROM where the conductivity factors are greater than unity. The main reason

(a) Dissipated Power $P_{\Omega_C}^0$

(b) Kinetic Energy $E_{\Omega_C}^k$
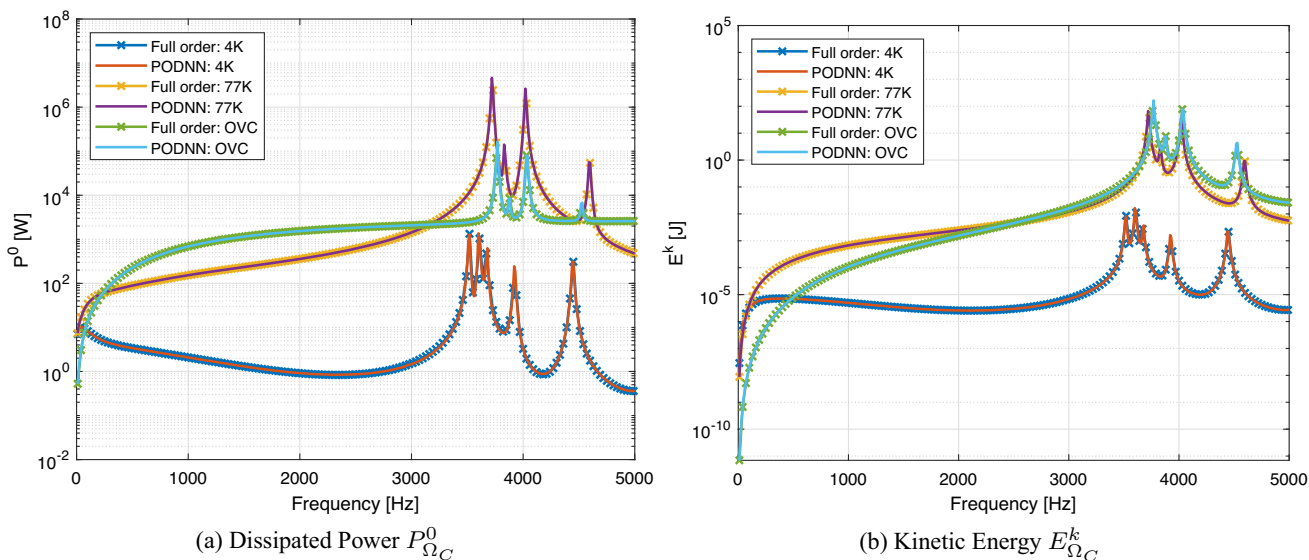
Fig. 13 Test magnet problem with z (longitudinal) gradient coil; Online stage for $P = 1$, comparisons of Full order and PODNN using $N_{\mathrm{fs}} = 23$. Results for **a** dissipated power $P_{\Omega_C}^0$ and **b** kinetic energy $E_{\Omega_C}^k$

Fig. 14 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 1$ and PODNN using $N_{\mathrm{fs}} = 23$. Results for **a** $|\operatorname{Re}\left(\boldsymbol{U}^{\mathrm{AC}}(\omega)\right)|$ at 1000 Hz, **b** $|\operatorname{Re}\left(\boldsymbol{\mathcal{B}}_{hq}^{\mathrm{AC}}(\omega)\right)|$ at 1000 Hz, **c** $|\operatorname{Re}\left(\boldsymbol{U}^{\mathrm{AC}}(\omega)\right)|$ at 5000 Hz and **d** $|\operatorname{Re}\left(\boldsymbol{\mathcal{B}}_{hq}^{\mathrm{AC}}(\omega)\right)|$ at 5000 Hz

(a) $|\operatorname{Re}\left(\boldsymbol{U}_{hp}^{AC}(\omega)\right)|$ at 1000Hz

(b) $|\operatorname{Re}\left(\boldsymbol{U}_{hp}^{AC}(\omega)\right)|$ at 5000Hz

(c) $|\operatorname{Re}\left(\boldsymbol{\mathcal{B}}_{hq}^{AC}(\omega)\right)|$ at 1000Hz

(d) $|\operatorname{Re}\left(\boldsymbol{\mathcal{B}}_{hq}^{AC}(\omega)\right)|$ at 5000Hz

(a) 4K shield at $\gamma_0^{OVC} = 0.35, 0.75$

(b) 4K shield at $\gamma_0^{OVC} = 3.5, 7.5$

(c) 77K shield at $\gamma_0^{OVC} = 0.35, 0.75$

(d) 77K shield at $\gamma_0^{OVC} = 3.5, 7.5$

(e) OVC shield at $\gamma_0^{OVC} = 0.35, 0.75$

(f) OVC shield at $\gamma_0^{OVC} = 3.5, 7.5$
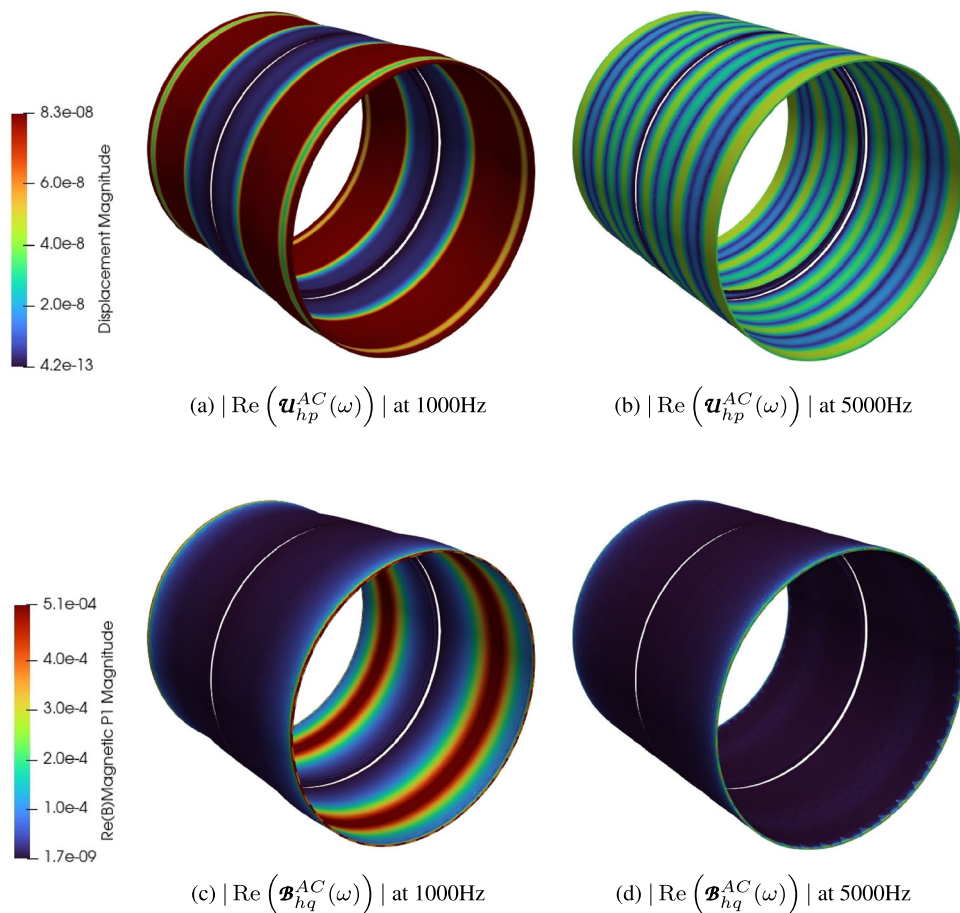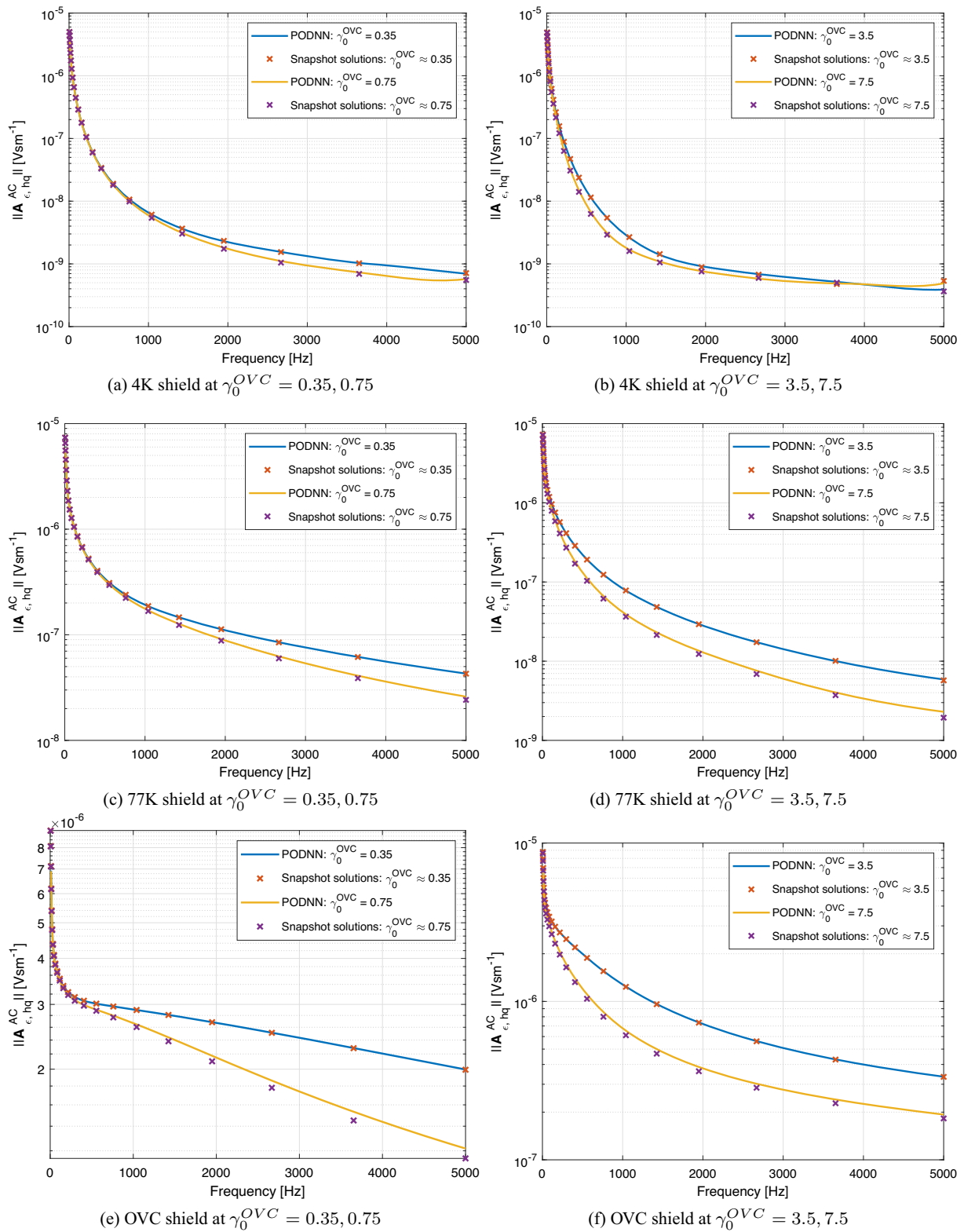
Fig. 15 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 2$ and $N_{\mathrm{fs}} = 23$, $N_{\mathrm{cs}} = 23$ comparing $\|\boldsymbol{\mathcal{A}}_{\varepsilon,hq}^{\mathrm{AC}}(\boldsymbol{w})\|_{L^2(\Omega_\mathrm{C})}$ for each of the three conducting shields. Results for **a** 4 K shield at

$\gamma_0^{\mathrm{OVC}} = 0.35, 0.75$, **b** 4 K shield at $\gamma_0^{\mathrm{OVC}} = 3.5, 7.5$, **c** 77 K shield at $\gamma_0^{\mathrm{OVC}} = 0.35, 0.75$, **d** 77 K shield at $\gamma_0^{\mathrm{OVC}} = 3.5, 7.5$, **e** OVC shield at $\gamma_0^{\mathrm{OVC}} = 0.35, 0.75$ and **f** OVC shield at $\gamma_0^{\mathrm{OVC}} = 3.5, 7.5$

for this separation is due to the log spaced nature of the conductivity snapshots, with a higher abundance of snapshots occurring less than unity than above, the idea here is to see how well the PODNN model with both those cases. In both conductivity factor regions, the PODNN model is seen to match reasonably well with the snapshot solutions at these specified conductivities.

Henceforth, we fix a PODNN using a $L = 3$, $n = 32$ network and snapshot data corresponding to $N_{\mathrm{fs}} = 23$ and $N_{\mathrm{cs}} = 23$ using this choice we show, in Figs. 16 and 17, the dissipated power $P^0_{\Omega_C}(\omega)$ and kinetic energy $E^k_{\Omega_C}(\omega)$ obtained by additionally solving (8) for each output frequency f_i and conductivity factor $\gamma^{\mathrm{OVC}}_{0,i}$ of interest using the reduced order model (15), field representations (5) and applying (22), respectively. Alongside, a comparison with the snapshot solutions is evaluated and can be seen to match well with the PODNN solutions. We observe that the higher conducting factors generally lead lower values of $P^0_{\Omega_C}(\omega)$ and $E^k_{\Omega_C}(\omega)$ for each shield and frequency of interest, although the position of the resonant peaks is not substantially affected.

4.5 Prediction of the response to external vibrational loading

The consideration of external vibrational loading for the mathematical model described in Sect. 2 introduces a non-zero boundary condition $\boldsymbol{u}^{\mathrm{AC}} = \boldsymbol{u}^{\mathrm{AC}}_D$ on $\partial\Omega^D_C$ in (2e), which reflects the amplitude of the external vibrational loading. Given the nature of the rotational symmetry of the problem considered in Sect. 4.1, we consider only the application of non-zero amplitudes in the z direction since specifying other non-zero components would break the rotational symmetry of the problem and necessitate a full three-dimensional computation rather than simulating quarter geometry and using the symmetry boundary conditions in (21) on the $x = 0$ and $y = 0$ planes.

4.5.1 Offline stage

The offline stage follows the approach described in Sect. 4.4.1 where a PODNN network with $L = 3$, $n = 32$ and $N_{\mathrm{fs}} = 23$ snapshots for $\boldsymbol{w} = (\omega) = 2\pi(f)$ with $P = 1$ is employed as described previously.

4.5.2 Online stage

Once (15) has been evaluated for a set of parameters of interest, the smaller dimensional mechanical problem (8) is solved. The magnitudes of the external vibrations are usually very small and so present results obtained by imposing $\mathcal{U}^{\mathrm{AC}}_{D_z} = 0, 10^{-6}, 10^{-5}, 10^{-4}$ m, in turn, simultaneously on all the conducting shields. To obtain, we solve (8) for these frequencies and each $\mathcal{U}^{\mathrm{AC}}_{D_z}$ of interest. Following this, we obtain the dissipated power and kinetic energy using (5) and (22), respectively.

In Fig. 18, we show the resulting the dissipated power $P^0_{\Omega_C}(\omega)$ and kinetic energy $E^k_{\Omega_C}(\omega)$ curves, where the different lines correspond to different amplitudes of displacement. In these figures, we observe that the dissipated power in the 77K and OVC shields are not substantially affected by introducing non-zero $\mathcal{U}^{\mathrm{AC}}_{D_z}$, however, there is a noticeable increase in the magnitude of the kinetic energy curves for these shields when using $\mathcal{U}^{\mathrm{AC}}_{D_z} = 10^{-4}$ m, although, the frequencies of the resonant peaks remain unchanged. In the case of the 4K shield, we see the most noticeable changes, with the resonant peaks all but disappearing for the case of $\mathcal{U}^{\mathrm{AC}}_{D_z} = 10^{-4}$ m and the kinetic energy curve.

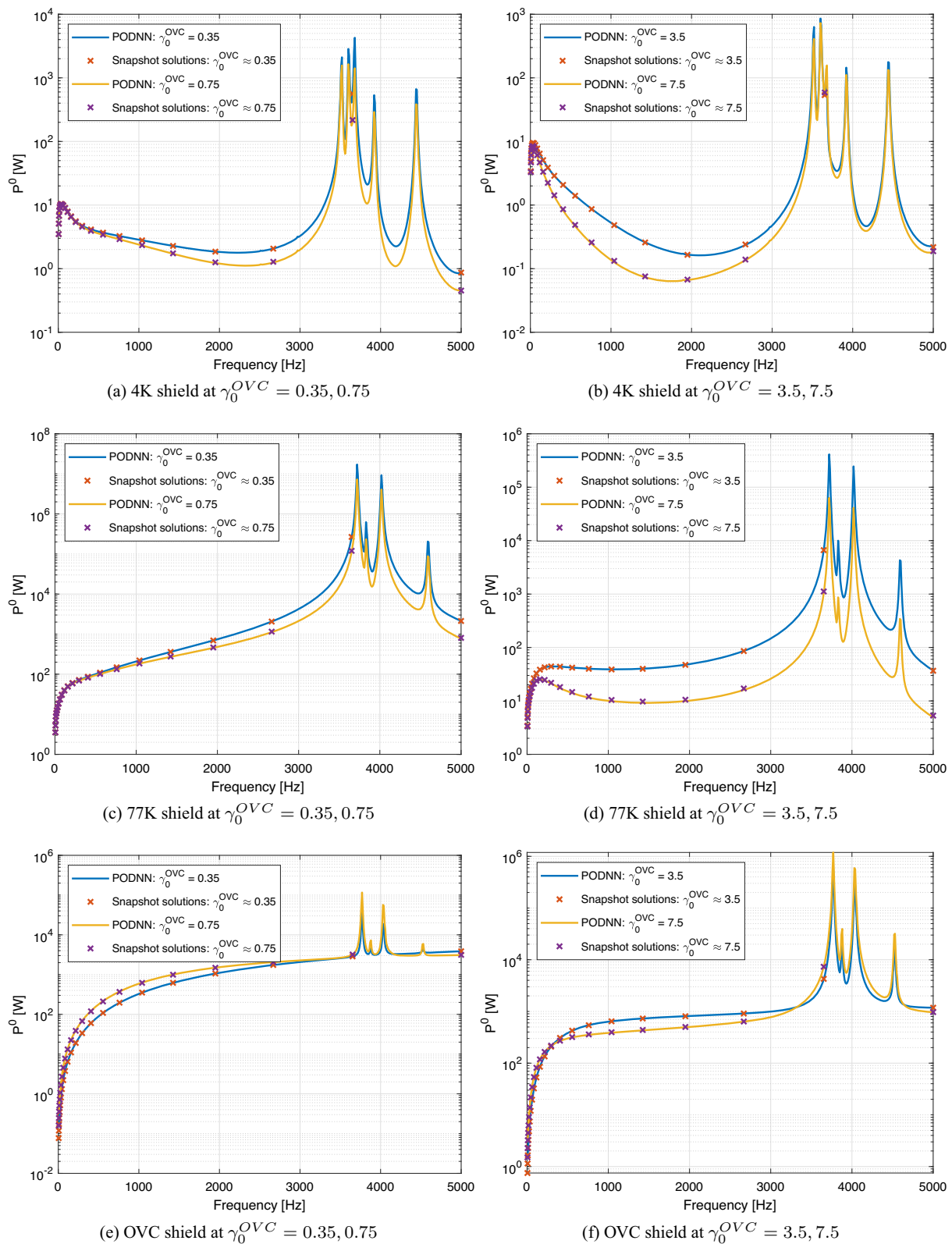## 4.6 Response when model parameters are uncertain

To understand the influence of an uncertain material parameter on the Ohmic power and kinetic energy curves, we may use the previously developed ROM. We focus on the situation where the conductivity $\gamma^{\mathrm{OVC}} = \gamma^{\mathrm{OVC}}_0 \gamma^{\mathrm{OVC}}_{\mathrm{ref}}$ in the OVC shield is uncertain. We consider a non-dimensional $\gamma^{\mathrm{OVC}}_0 \sim \mathcal{N}(m, s)$ with population mean $m = 1$, population standard deviation $s = \log_e(2)$ and $\gamma^{\mathrm{OVC}}_{\mathrm{ref}}$ is a reference value with units of S/m and of order 10^6.

4.6.1 Offline stage

The offline stage follows the approach described in Sect. 4.4.1 where a PODNN network with $L = 3$, $n = 32$ network for $N_{\mathrm{fs}} = 23$ and $N_{\mathrm{cs}} = 23$ snapshots is employed for $\boldsymbol{w} = (\omega, \gamma^{\mathrm{OVC}}_0) = (2\pi f, \gamma^{\mathrm{OVC}}_0)$ with $P = 2$, as described previously.
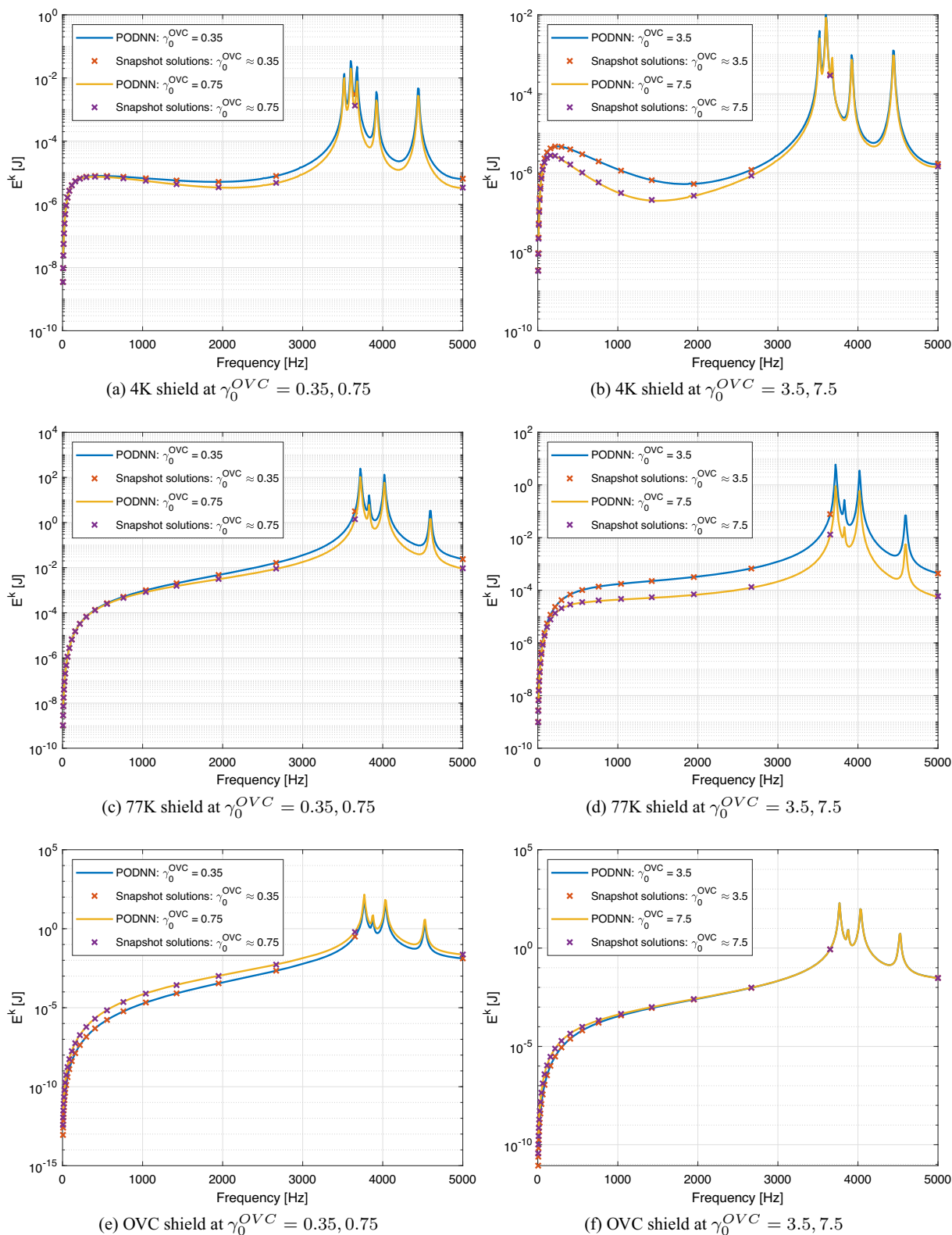
4.6.2 Online stage

In this case, $N_{\mathrm{co}} = 50$ samples of $\gamma^{\mathrm{OVC}}_0 \sim \mathcal{N}(m, s)$ are drawn and for each sample, $N_{\mathrm{fo}} = 500$ output frequencies considered. For each combination of f_i and conductivity factor $\gamma^{\mathrm{OVC}}_{0,i}$, the dissipated power $P^0_{\Omega_C}(\omega)$ and kinetic energy $E^k_{\Omega_C}(\omega)$ are obtained by solving (8) using the reduced order model (15), field representations (5) and applying (22), respectively. Using this data, the mean values and 95% confidence intervals of $P^0_{\Omega_C}(\omega)$ and $E^k_{\Omega_C}(\omega)$, at each output frequency of interest, are evaluated and the results shown in Fig. 19. We see, in general, the position of the peaks are not significantly changed, but there is an noticeable difference in the amplitudes, particularly between the frequencies of $1000 \leq f \leq 3000$ Hz for the 4 K and 77 K shields. The OVC shield does not seem to show as much variation, with the mean and 95% confidence intervals for $P^0_{\Omega_C}$ and $E^k_{\Omega_C}$ being almost indistinguishable on this scale.
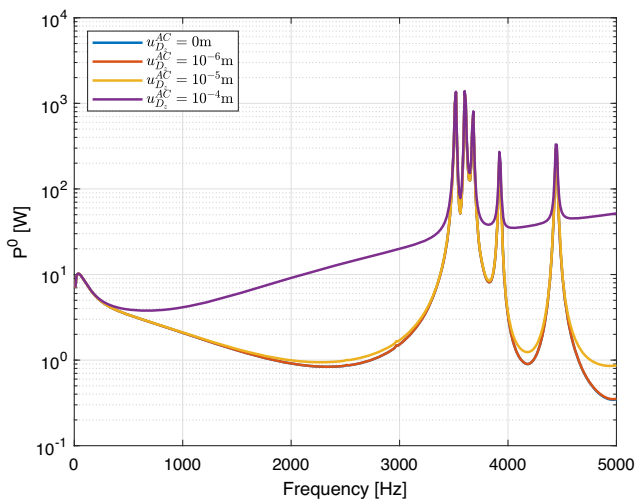
(a) 4K shield at $\gamma_0^{OVC} = 0.35, 0.75$

(b) 4K shield at $\gamma_0^{OVC} = 3.5, 7.5$

(c) 77K shield at $\gamma_0^{OVC} = 0.35, 0.75$

(d) 77K shield at $\gamma_0^{OVC} = 3.5, 7.5$

(e) OVC shield at $\gamma_0^{OVC} = 0.35, 0.75$

(f) OVC shield at $\gamma_0^{OVC} = 3.5, 7.5$

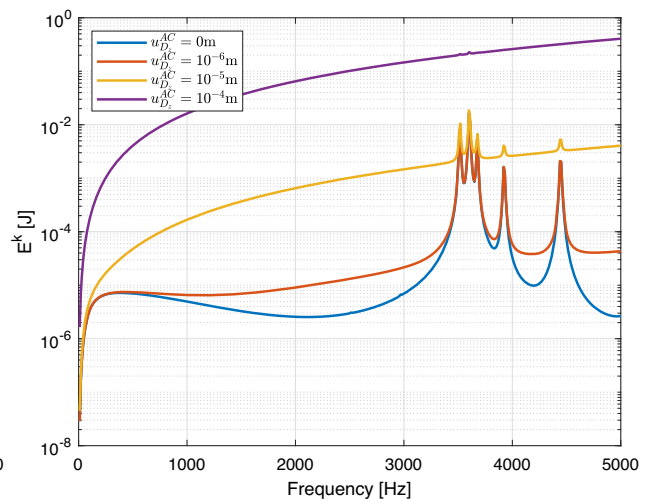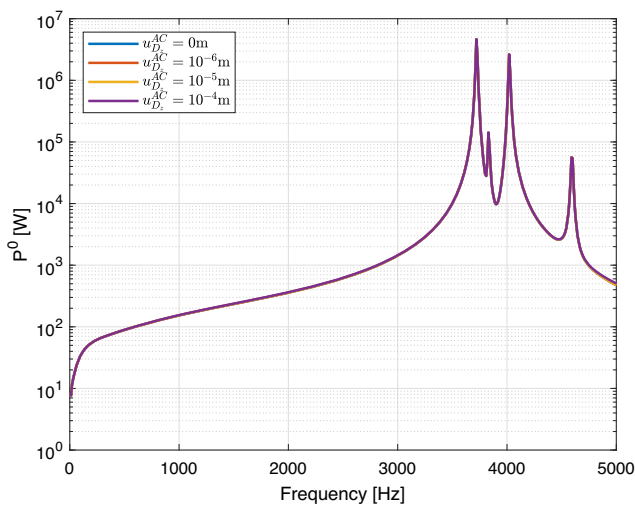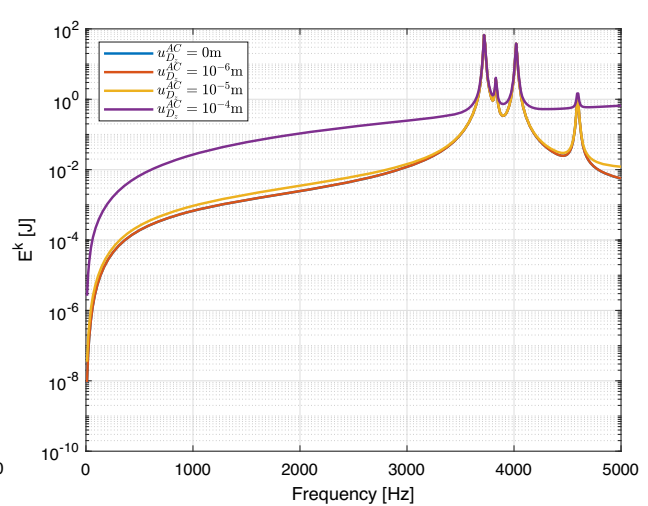**Fig. 16** Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 2$ and $N_{\text{fs}} = 23$, $N_{\text{cs}} = 23$ comparing the dissipated power $P_{\Omega_C}^0$ for each of the three conducting shields. Results for **a** 4 K shield at $\gamma_0^{\text{OVC}} = 0.35, 0.75$, **b** 4 K shield at $\gamma_0^{\text{OVC}} = 3.5, 7.5$, **c** 77 K shield at $\gamma_0^{\text{OVC}} = 0.35, 0.75$, **d** 77 K shield at $\gamma_0^{\text{OVC}} = 3.5, 7.5$, **e** OVC shield at $\gamma_0^{\text{OVC}} = 0.35, 0.75$ and **f** OVC shield at $\gamma_0^{\text{OVC}} = 3.5, 7.5$

(a) 4K shield at $\gamma_0^{OVC} = 0.35, 0.75$

(b) 4K shield at $\gamma_0^{OVC} = 3.5, 7.5$

(c) 77K shield at $\gamma_0^{OVC} = 0.35, 0.75$

(d) 77K shield at $\gamma_0^{OVC} = 3.5, 7.5$

(e) OVC shield at $\gamma_0^{OVC} = 0.35, 0.75$

(f) OVC shield at $\gamma_0^{OVC} = 3.5, 7.5$

Fig. 17 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 2$ and $N_{\mathrm{fs}} = 23$, $N_{\mathrm{cs}} = 23$ comparing the kinetic energy $E_{\Omega_C}^k$ for each of the three conducting shields. Results for **a** 4 K shield at $\gamma_0^{\mathrm{OVC}} = 0.35, 0.75$, **b** 4 K shield at $\gamma_0^{\mathrm{OVC}} = 3.5, 7.5$, **c** 77 K shield at $\gamma_0^{\mathrm{OVC}} = 0.35, 0.75$, **d** 77 K shield at $\gamma_0^{\mathrm{OVC}} = 3.5, 7.5$, **e** OVC shield at $\gamma_0^{\mathrm{OVC}} = 0.35, 0.75$ and **f** OVC shield at $\gamma_0^{\mathrm{OVC}} = 3.5, 7.5$
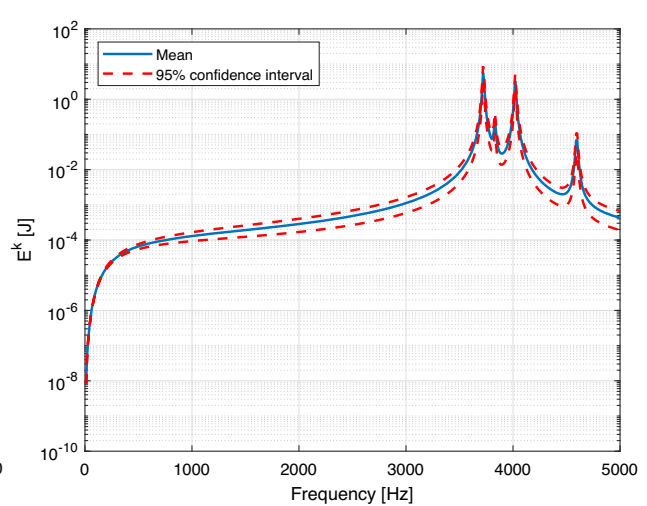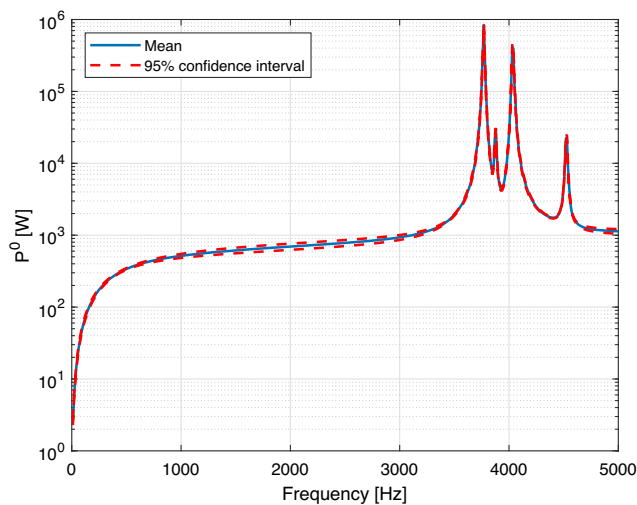
(a) Dissipated Power $P^0_{\Omega_C}$ in 4K shield

(b) Kinetic Energy $E^k_{\Omega_C}$ in 4K shield

(c) Dissipated Power $P^0_{\Omega_C}$ in 77K shield

(d) Kinetic Energy $E^k_{\Omega_C}$ in 77K shield

(e) Dissipated Power $P^0_{\Omega_C}$ in OVC shield

(f) Kinetic Energy $E^k_{\Omega_C}$ in OVC shield

Fig. 18 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 1$ and $N_{\mathrm{fs}} = 23$; comparing the distribution of kinetic energy $E^k_{\Omega_C}$ and dissipated power $P^0_{\Omega_C}$ for Dirichlet displacements in z such that $\mathcal{U}^{\mathrm{AC}}_{D_z} = 0$, 10^{-6}, 10^{-5}, 10^{-4} m, in turn Result for **a** $P^0_{\Omega_C}$ 4 K shield, **b** $E^k_{\Omega_C}$ 4 K shield, **c** $P^0_{\Omega_C}$ 77 K shield, **d** $E^k_{\Omega_C}$ 77 K shield, **e** $P^0_{\Omega_C}$ OVC shield and **f** $E^k_{\Omega_C}$ OVC shield

(a) Dissipated Power $P_{\Omega_C}^0$ in 4K shield
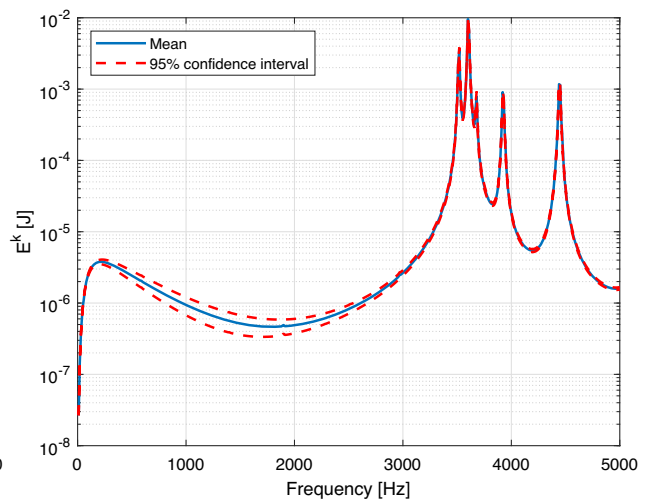
(b) Kinetic Energy $E_{\Omega_C}^k$ in 4K shield
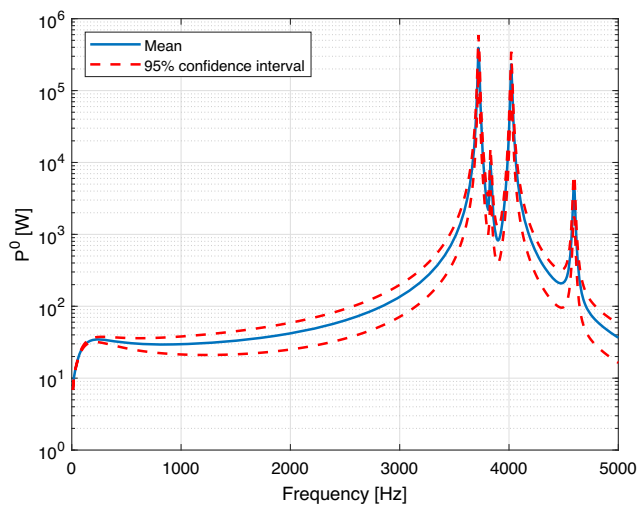
(c) Dissipated Power $P_{\Omega_C}^0$ in 77K shield

(d) Kinetic Energy $E_{\Omega_C}^k$ in 77K shield

(e) Dissipated Power $P_{\Omega_C}^0$ in OVC shield
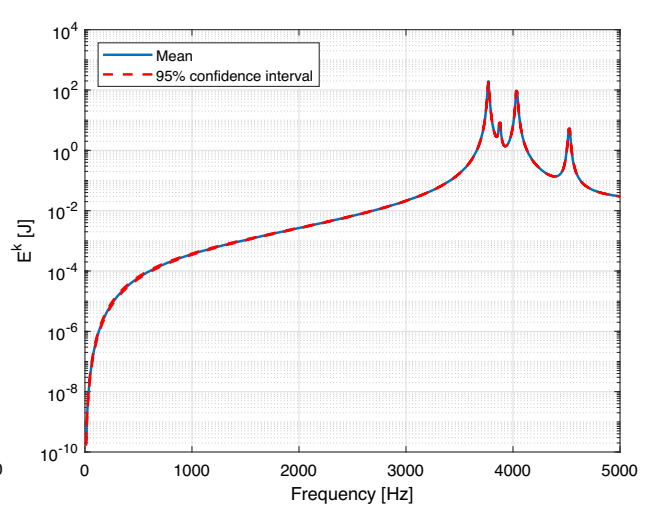
(f) Kinetic Energy $E_{\Omega_C}^k$ in OVC shield

Fig. 19 Test magnet problem with z (longitudinal) gradient coil; online stage for $P = 2$ and $N_{\mathrm{fs}} = 23$, $N_{\mathrm{cs}} = 23$, $N_{\mathrm{co}} = 50$; comparing the distribution of kinetic energy $E_{\Omega_C}^k$ and dissipated power $P_{\Omega_C}^0$ for each of the three conducting shields Results for **a** $P_{\Omega_C}^0$ 4 K shield, **b** $E_{\Omega_C}^k$ 4 K shield, **c** $P_{\Omega_C}^0$ 77 K shield, **d** $E_{\Omega_C}^k$ 77 K shield, **e** $P_{\Omega_C}^0$ OVC shield and **f** $E_{\Omega_C}^k$ OVC shield

5 Conclusion

This paper has presented a practical use of a POD technique based on using a neural network (PODNN), to aid in the design phase of new MRI scanner configurations. We compare the performance of PODNN to different POD techniques including projection (PODP) and interpolating (PODI) for this challenging example. The PODNN methodology employed is equivalent to that proposed by Hesthaven and Ubbiali [21]. The PODP and PODI techniques build on our groups earlier work for the coupled magneto-mechanical problem of interest [12, 17]. The PODP approach produced relatively accurate results with a small numbers of snapshots (N_{fs} and N_{cs}) but was heavily intrusive on the software and required recalculation of FEM matrices. The PODI method allowed the use of simple interpolants (such as a Lagrangian) and was non-intrusive, but had limitations when considering small snapshots. Instead, the PODNN method, does not have an invasive implementation and is well suited to this and produces accurate results but the optimisation time (using Bayes optimisation to obtain hyper-parameters) in the offline stage increases significantly when the parameter space is $P = 2$, making the PODP technique more computationally efficient if overall time is considered.

Results have been presented for the prediction of resonant modes in the Ohmic power and kinetic energy curves using different parameter spaces and the prediction of the response in the same outputs of interest for different external vibrational loading conditions. The PODNN technique has also been applied to predict the response in outputs of interest when model parameters are uncertain. These results have shown that an appropriate network architecture, can provide high-fidelity solutions to the magneto-mechanical problem for both the $P = 1$ (frequency) and $P = 2$ (frequency and conductivity) parameter cases and provide rapid online prediction of the outputs of interest for the design of MRI scanners. In further work, we also plan to make experimental comparisons with the fields induced in moving magnets and make improvements to the mathematical model presented in Sect. 2 that no longer assumes that the velocities of the conducting components are small.

Declarations

Conflict of interest The authors are grateful for the financial support received from EPSRC CASE Award Studentships in partnership with Siemens Healthineers, which are supporting S. Miah and Y. Sooriyakanthan. The authors have no relevant financial or non-financial interests to disclose.

References

1. Clarke L, Velthuizen R, Camacho M, Heine J, Vaidyanathan M, Hall L, Thatcher R, Silbiger M (1995) MRI segmentation: methods and applications. Magn Reson Imaging 13(3):343–368
2. Matzat SJ, van Tiel J, Gold GE, Oei EH (2013) Quantitative MRI techniques of cartilage composition. Quant Imaging Med Surg 3(3):162
3. Savnik A, Malmskov H, Thomsen HS, Graff LB, Nielsen H, Danneskiold-Samsøe B, Boesen J, Bliddal H (2002) MRI of the wrist and finger joints in inflammatory joint diseases at 1-year interval: MRI features to predict bone erosions. Eur Radiol 12(5):1203–1210
4. Kozlowski P, Chang SD, Jones EC, Berean KW, Chen H, Goldenberg SL (2006) Combined diffusion-weighted and dynamic contrast-enhanced MRI for prostate cancer diagnosis: correlation with biopsy and histopathology. J Magn Reson Imaging Off J Int Soc Magn Reson Med 24(1):108–113
5. Ledger PD, Gil AJ, Poya R, Kruip M, Wilkinson I, Bagwell S (2016) Solution of an industrially relevant coupled magneto-mechanical problem set on an axisymmetric domain. Appl Math Model 40:1959–1971
6. Bagwell S, Ledger PD, Gil AJ, Mallett M, Kruip M (2017) A linearised *hp*-finite element framework for acousto-magneto-mechanical coupling in axisymmetric MRI scanners. Int J Numer Methods Eng 112:1323–1352
7. Bagwell S, Ledger PD, Gil AJ, Mallett M (2018) Transient solutions to nonlinear acousto-magneto-mechanical coupling for axisymmetric MRI scanner design. Int J Numer Methods Eng 115:209–237
8. Szabó B, Babuška I (1991) Finite Element Analysis. Wiley, New York
9. Karniadakis GE, Karniadakis G, Sherwin S (2005) Spectral/hp Element Methods for Computational Fluid Dynamics. Oxford University Press on Demand, Oxford
10. Demkowicz L (2006) Computing with hp-Adaptive Finite Elements: Volume 1 One and Two Dimensional Elliptic and Maxwell Problems. Chapman and Hall/CRC, Boca Raton
11. Demkowicz L, Kurtz J, Pardo D, Paszynski M, Rachowicz W, Zdunek A (2007) Computing with hp-Adaptive Finite Elements, vol. 2: Frontiers Three Dimensional elliptic and Maxwell problems with applications, 1st edn. Chapman & Hall/CRC, Boca Raton
12. Seoane M, Ledger PD, Gil AJ, Mallett M (2019) An accurate and efficient three-dimensional high-order finite element methodology for the simulation of magneto-mechanical coupling in MRI scanners. Int J Numer Methods Eng 119:1185–1215
13. Liang Y, Lee H, Lim S, Lin W, Lee K, Wu C (2002) Proper orthogonal decomposition and its applications-part i: theory. J Sound Vib 252(3):527–544

14. Chatterjee A (2000) An introduction to the proper orthogonal decomposition. Curr Sci 78:808–817
15. Hesthaven JS, Rozza G, Stamm B et al (2016) Certified Reduced Basis Methods for Parametrized Partial Differential Equations, vol 590. Springer, Berlin
16. Amsallem D, Farhat C (2008) Interpolation method for adapting reduced-order models and application to aeroelasticity. AIAA J 46(7):1803–1813
17. Seoane M, Ledger PD, Gil AJ, Zlotnik S, Mallett M (2020) A combined reduced order-full order methodology for the solution of 3D magneto-mechanical problems with application to magnetic resonance imaging scanners. Int J Numer Methods Eng 121:3529–3559
18. Seoane M (2020) 3D simulation of magneto-mechanical coupling in MRI scanners using high order FEM and POD. Ph.D. thesis, Swansea University, UK
19. Bishop C (2006) Pattern Recognition and Machine Learning. Springer, Berlin
20. Higham CF, Higham DJ (2019) Deep learning: an introduction for applied mathematicians. SIAM Rev 61:860–891
21. Hesthaven JS, Ubbiali S (2018) Non-intrusive reduced order modeling of nonlinear problems using neural networks. J Comput Phys 363:55–78
22. Dongmei H, Shiqing H, Xuhui H, Xue Z (2017) Prediction of wind loads on high-rise building using a BP neural network combined with POD. J Wind Eng Ind Aerodyn 170:1–17
23. Wang Q, Ripamonti N, Hesthaven JS (2020) Recurrent neural network closure of parametric POD-Galerkin reduced-order models based on the Mori–Zwanzig formalism. J Comput Phys 410:109402
24. Fresca S, Manzoni A, Dedè L, Quarteroni A (2020) Deep learning-based reduced order models in cardiac electrophysiology. PLoS One 15(10 October):1–32
25. Fresca S, Dede' L, Manzoni A (2021) A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs. J Sci Comput 87(2):1–36
26. Fresca S, Manzoni A (2022) POD-DL-ROM: enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition. Comput Methods Appl Mech Eng 388:114181
27. Bagwell S (2018) A Numerical Multi-Physics Approach to Understanding MRI Scanners and their Complex Behaviour. Ph.D. thesis, Swansea University, UK
28. Schöberl J, Zaglmayr S (2005) High order Nédélec elements with local complete sequence properties. Int J Comput Math Electr Electron Eng (COMPEL) 24:374–384
29. Zaglmayr S (2006) High Order Finite Element Methods for Electromagnetic Field Computation. Ph.D. thesis, Institüt für Numerische Mathematik, Johannes Kepler Univerität Linz, Austria
30. Buljak V (2011) Inverse Analyses with Model Reduction: Proper Orthogonal Decomposition in Structural Mechanics. Springer Science & Business Media, Berlin
31. Wang Y, Yu B, Cao Z, Zou W, Yu G (2012) A comparative study of POD interpolation and POD projection methods for fast and accurate prediction of heat transfer problems. Int J Heat Mass Transf 55(17–18):4827–4836
32. Rama R, Skatulla S, Sansour C (2015) Real-time modelling of the heart using the proper orthogonal decomposition with interpolation. In: VI international conference on computational bio-engineering, ICCB, Barcelona, pp 1–12
33. Stabile G, Hijazi S, Mola A, Lorenzi S, Rozza G (2017) POD-galerkin reduced order methods for CFD using finite volume discretisation: vortex shedding around a circular cylinder. Commun Appl Ind Math 8(1):210–236
34. Mockus J (1975) On the Bayes methods for seeking the extremal point. IFAC Proc Vol 8(1):428–431
35. Barroso G, Seoane M, Gil AJ, Ledger PD, Mallett M, Huerta A (2020) A staggered high-dimensional proper generalised decomposition for coupled magneto-mechanical problems with application to MRI scanners. Comput Methods Appl Mech Eng 370:113271
36. Mathwork's deep learning toolbox. https://uk.mathworks.com/products/deep-learning.html. Accessed 10 Apr 2022
37. Tensorflow. https://www.tensorflow.org/. Accessed 10 Apr 2022
38. scikit-learn. https://scikit-learn.org/stable/. Accessed 10 Apr 2022
39. pytorch. https://pytorch.org. Accessed 10 Apr 2022
40. Matlab. https://uk.mathworks.com/products/matlab.html. Accessed 10 Apr 2022
41. Mathwork's statistics and machine learning toolbox. https://uk.mathworks.com/products/statistics.html. Accessed 15 June 2023