Review of "Accelerating 3D Magnetotelluric Forward Modelling with Domain Decomposition and Order-Reduction Methods"

L. Tao, A. Muixi, S. Zlotnik, F. Zyserman, J. Afonso, P. Diez

August 13, 2026

This paper develops a domain decomposition (DD) formulation for solving the 3D magnetotelluric (MT) equations, paired with a POD/Galerkin dimension reduction technique. The coupling is achieved via the non-overlapping Schwarz alternating method with Robin–Robin transmission conditions. The proposed framework allows one to couple subdomain-local finite element models, or subdomain-local projection-based reduced order models (ROMs). The resulting coupled models are termed DD and DD-POD, respectively. It is demonstrated that the DD-based models achieve significant speedups over their non-DD monolithic counterparts on two examples, a canonical toy problem, and a real-world problem.

The approach presented in the paper is timely, as DD-based methods involving ROMs have been taking off in recent years, and this topic is a particular research interest of mine. I believe it is among the first, or maybe the first, to combine DD and ROM for the MT application (caveat: I am not an expert on MT).

The paper is interesting and well-written, but unfortunately there are numerous important details about the methodology and numerical results that are missing. Some important references are missing as well. I also find the results that the DD variant of the method is significantly cheaper than the non-DD extremely surprising, especially since this is a non-overlapping Schwarz coupling framework, and a relatively large number of subdomains is being coupled, I believe. More evidence is needed to justify/explain these results.

My recommendation for the paper is "Major Revision". I ask that the authors revise their paper to address the following comments.

General response: First, we would like to thank the reviewer for taking the time to carefully read the manuscript and for providing constructive comments and suggestions. We believe that addressing these points will improve the clarity of the manuscript and strengthen the presentation and robustness of the proposed methodology.

1 Major comments

1. **There has been a lot of recent work on DD + ROM, including DD + ROM + Schwarz, and it is worth citing some of it. Please refer to**

   - **Ruan, Class, Rozza, *A Structured Review of Reduced Order Modeling for Domain Decomposition Problems: State of the Art and Perspectives*, 2026.**

   for a comprehensive review of what has been done with DD + ROM until now. In particular, I would cite:

   - **Moore, Wentland, Gruber, Tezaur, *Domain Decomposition-based Coupling of Operator Inference Reduced Order Models via the Schwarz Alternating Method*, 2024.**

- **Rodriguez, C., Tezaur, I., Mota, A., Gruber, A., Parish, E., and Wentland, C. (2025).** ***Transmission Conditions for the Non-Overlapping Schwarz Coupling of Full Order and Operator Inference Models.*** **arXiv:2509.12228.**
- **Hoang, C., Choi, Y., and Carlberg, K. (2021).** ***Domain-decomposition least-squares Petrov–Galerkin (DD-LSPG) nonlinear model reduction.*** **Computer Methods in Applied Mechanics and Engineering, 384, 113997.**
- **Diaz, P., Choi, Y., and Heinkenschloss, M. (2024).** ***A fast and accurate domain decomposition nonlinear manifold reduced-order model.*** **Computer Methods in Applied Mechanics and Engineering.**
- **Chung, S. W., Choi, Y., Roy, P., Moore, T., Roy, T., Lin, T. Y., Nguyen, D. T., Hahn, C., Duoss, E. B., and Baker, S. E. (2024).** ***Train small, model big: Scalable physics simulators via reduced order modeling and domain decomposition.*** **Computer Methods in Applied Mechanics and Engineering, 427, 117041.**

A: We thank the reviewer for pointing this out. We agree that the discussion of recent work combining domain decomposition and reduced-order modelling is too limited in the current manuscript. In the revised manuscript, we will expand this part of the introduction to include the suggested literature and better position the present contribution within the broader DD–ROM context.

2. **I am confused by the discussion on p. 7 about enforcing the BC at a single point. The relevant BC is a Robin one, not a Dirichlet one. Thus, by construction, it requires multple points along a subdomain boundary to define. Can you please clarify what is meant by the "single point enforcement"? Are you perhaps integrating over the boundary using a single integration point?**

   A: First it is important to mention that the finite element formulation is adopted from Douglas Jr., Santos, and Sheen, "A Nonconforming Mixed Finite Element Method for Maxwell's Equations" (2000), while the domain decomposition scheme follows Zyserman and Santos, "Parallel Finite Element Algorithm with Domain Decomposition for Three-Dimensional Magnetotelluric Modelling." (2000). We thank the reviewer for this comment and agree that the current wording can be improved. The Robin transmission condition itself is defined over the complete interface between neighbouring subdomains, as introduced in Eqs. (7)–(8). In the weak formulation, the corresponding interface integral is evaluated element-wise using a single integration point located at the midpoint of each interface element. Therefore, the term "single point enforcement" refers to the use of a one-point quadrature rule for each interface element, rather than to imposing the Robin condition at only one point of the entire subdomain interface. We will improve the wording.

3. **I was surprised by the introduction of Lagrange multipliers (LMs) within the Schwarz iteration on p. 7. One of the principal advantages of Robin–Schwarz methods is their minimal intrusiveness: existing subdomain solvers can typically be reused with only modified interface boundary conditions and data exchange. Introducing Lagrange multipliers adds interface unknowns and a saddle-point coupling, which appears to increase implementation complexity. It would be helpful if the authors clarified why this additional machinery is necessary and what advantages it provides over a purely Robin–Schwarz formulation.**

   A: The Lagrange multipliers are introduced mainly to simplify the treatment of the interface coupling in our formulation. They are associated with the tangential magnetic field coefficients

at the subdomain interfaces. In practice, the linear system for each subdomain is solved for the electric field coefficients, while the magnetic field coefficients at the interfaces, which are required for the Robin transmission conditions, are recovered through the Lagrange multipliers. This allows the electric and magnetic field coefficients to be treated separately in the implementation and simplifies the overall interface coupling.

Although this introduces additional interface unknowns, they can be easily and cheaply recovered, since they are computed as a function of the electric field and the Lagrange multiplier from the previous Schwarz iteration.

We agree that the reason for introducing the Lagrange multipliers and their role in the formulation are not sufficiently explained in the current manuscript. We will revise Sect. 2.2.4 to clarify these points.

4. **It is also not clear to me what exactly the LMs are enforcing on p. 7.**

   A: We thank the reviewer for this comment. The Lagrange multipliers are used to weakly enforce the continuity of the tangential magnetic field coefficients across the interfaces between neighbouring subdomains. In the formulation, they represent the tangential magnetic field at the interfaces and are used to construct the Robin transmission conditions exchanged between neighbouring subdomains during the Schwarz iterations. We will clarify this point in Sect. 2.2.4 of the revised manuscript.

5. **Still on the same note, the motivation for introducing the Lagrange multiplier formulation and the single-point treatment of the interface conditions is not entirely clear. In Schwarz methods, the dominant computational cost typically arises from the repeated subdomain solves over multiple interface iterations, whereas the enforcement of the transmission conditions is relatively inexpensive. It would therefore be helpful for the authors to explain what computational or algorithmic benefit these additional interface treatments provide, for example in terms of reduced Schwarz iterations, improved accuracy, or reduced online cost. Did the authors ever try the method without these two "speed-ups"? What was the relative performance? If the LM makes it not that much better, I would say giving up the non-intrusiveness of regular Schwarz by introducing it is not worth it.**

   A: We thank the reviewer for this comment. We would like to clarify that the Lagrange multiplier formulation and the single-point treatment of the interface terms were not introduced as "speed-ups" of the Schwarz method, but rather to simplify the formulation and solution of the problem which can possibly be faster.

   The Lagrange multipliers are part of the hybrid formulation used in our implementation. They represent the tangential magnetic field at the interfaces and allow us to solve the subdomain linear systems for the electric field while recovering the magnetic quantities required by the Robin transmission conditions. Although this introduces additional interface unknowns, their computation is inexpensive, since they are obtained from the electric field and the Lagrange multiplier from the previous Schwarz iteration.

   Similarly, the single-point treatment refers to the numerical integration of the interface terms, the Robin condition is defined over the complete interface, while the corresponding contribution is evaluated using one integration point at the midpoint of each interface element.

   We have not considered these two features as independent acceleration techniques and therefore have not performed a comparison with a formulation in which they are removed. The main computational gains reported in this work come from the domain decomposition and,

particularly, from the reduced-order modelling of the subdomain problems, rather than from these interface treatments. We will revise Sect. 2.2.4 to make this distinction clearer.

6. **The Schwarz method can technically handle non-conformal meshes in the subdomains, but interpolation must be introduced. Is this use case of interest to you?**

   A: We agree that non-conformal interfaces are an interesting possible extension of the method. The implementation considered in the present work uses a structured conforming mesh and a conforming decomposition of the computational domain. Non-conformal interfaces have not been implemented or tested in the present work. We will mention this explicitly in the revised manuscript as a possible extension of the current implementation.

7. **The reduced systems appear to be solved using an iterative linear solver. Since projection-based ROMs typically yield relatively small dense systems, a direct dense factorization is often more efficient and avoids iterative convergence issues. Could the authors comment on their choice of iterative solver and whether direct dense solvers were considered or compared?**

   A: We thank the reviewer for raising this point. The linear systems in the implementation are solved using the complex symmetric direct solver ZMUMPS from the MUMPS package. The iteration appearing in DD and DD-POD is the final Schwarz iteration, in which interface information is exchanged and updated between subdomain solves; it is not an iterative linear solver used to solve the reduced systems.

   For the reduced systems, the factorization and solution times are very small, as shown in the timing breakdowns. For G-POD and DD-POD, the dominant online cost is generally associated with the construction and projection of the reduced operator rather than with the direct solution of the reduced linear system.

   We will revise the relevant text to distinguish more clearly between the direct algebraic solver and the iterative Schwarz coupling.

8. **The proposed framework appears to couple reduced models on all subdomains. Could the methodology be extended to hybrid FOM–ROM coupling, as in some of the references above, where only a subset of subdomains is reduced? Such a capability would increase flexibility by allowing reduced and full-order subdomains to coexist within the Schwarz iteration.**

   A: We thank reviewer for the suggestion. The structure of the formulation is compatible with a hybrid FOM–ROM strategy in which some subdomains are retained at full order while others are represented by reduced models. In the present formulation, the full-order and reduced subdomain problems are coupled through the same interface exchange within the Schwarz iteration.

   However, the numerical experiments presented in the current manuscript use either full-order models on all subdomains (DD) or reduced-order models on all subdomains (DD-POD). Hybrid FOM–ROM coupling has therefore not been implemented or assessed numerically in this work. We will mention this as a possible extension in the revised manuscript.

9. **The offline training procedure is not described in sufficient detail, especially for the first problem. It would be helpful to specify the sampling strategy for the conductivity models, the number of training realizations, the snapshot collection procedure, the POD truncation criterion, and whether separate reduced**

bases are constructed for different frequencies. These details are important for reproducibility and for assessing the offline cost of the proposed approach.

A: We agree that the offline procedure should be described more systematically. The relevant information is distributed across several sections of the manuscript.

The snapshot generation is described in Sects. 5.1.2 and 5.2.2. For the first problem (DTM1 benchmark), the conductivity model contains three varying parameters that correspond to the resistivity of 3 anomalies. The training set is the complete Cartesian product

$$\rho_i \in \{20000,\, 200,\, 100,\, 2,\, 0.2\}\ \Omega\,\mathrm{m},$$

which gives $5^3 = 125$ full-order training configurations.

For each training configuration, the full-order problem is iterated to convergence. The converged solution is retained as a snapshot, and the corresponding part of each full-order snapshot is extracted for each subdomain to construct the local snapshot matrices.

For the realistic model, five geological units are treated parametrically. Four of them use

$$\{1000,\, 216,\, 47,\, 10\}\ \Omega\,\mathrm{m},$$

while the fifth uses

$$\{100,\, 21.6,\, 4.7,\, 1.0\}\ \Omega\,\mathrm{m}.$$

The complete Cartesian product therefore gives $4^5 = 1024$ full-order training configurations. The parameter ranges were selected to span representative conductivity values without coinciding with the target values used in the corresponding test model, that will be included in the MS. Independent POD bases are constructed for each subdomain, polarization and frequency, and the POD truncation is based on the SVD criterion in Eq. (24). In the revised manuscript, we will include these details into a clearer description of the offline procedure and will state the number of training solutions, the real-world test case and the snapshot-generation procedure, the basis construction and the offline computational cost more explicitly.

10. **How were the basis sizes selected? Was it using an energy criterion?**

A: The basis dimension is selected from the singular values of the snapshot matrix using the truncation criterion given in Eq. (24):

$$\sum_{i=1}^{n_{\mathrm{POD}}} \tau_i \geq (1 - \delta_{\mathrm{SVD}}) \sum_{i=1}^{n_s} \tau_i,$$

where τ_i are the singular values and δ_{SVD} controls the truncation. For the benchmark calculations, we use

$$\delta_{\mathrm{SVD}} = 10^{-3}.$$

This is a cumulative singular-value criterion rather than the conventional POD energy criterion based on the sum of squared singular values. We will clarify this terminology in the revised manuscript.

For some of the realistic-model experiments, n_{POD} is prescribed and varied deliberately in order to study the dependence of the error and computational time on the reduced dimension.

11. **The reported speedup of the domain-decomposed ROM over the global ROM, and the domain-decomposed full order model (FOM) over the global FOM, is extremely surprising. Classical Schwarz methods, particularly non-overlapping variants, generally introduce additional interface iterations and therefore do not**

reduce the total computational work in serial. I agree that the subdomain problems are easier to solve. However, with nine subdomains for the first problem, one would expect the domain decomposition overhead to offset the benefit of the smaller local systems. It would be helpful if the authors could provide additional insight into the source of the observed speedup—for example, by reporting the local and global reduced dimensions, the number of Schwarz iterations, and a timing breakdown separating local solves from interface iteration. Are the comparisons in CPU times apples-to-apples, i.e., is the same amount of computational resources utilized for each set of runs?

A: We thank the reviewer for this comment. The local and global reduced dimensions are currently reported in Tables 3–4 for the first problem and in Table 7 for the realistic case. The Schwarz iteration counts and the timing breakdowns for the XY and YX polarizations are shown in Figs. 5–6 for the first problem and in Fig. 8 for the realistic case. The computational resources are the same and their allocation for the different solver variants are described in Sects. 4.2.2 and 6.1.1, with the aim of making the comparisons between the approaches as fair as possible.

The timing breakdowns also show that, for the global FOM, a large fraction of the execution time is associated with the analysis and factorization of the global left-hand-side matrix. In the DD formulation, the local left-hand-side matrices remain constant throughout the Schwarz iterations, so they are analysed and factorized once and the resulting factorization are reused in the subsequent iterations. Because the local systems are substantially smaller, this can lead to a lower total execution time provided that the number of Schwarz iterations remains sufficiently low. We also chose the hyperparameters that focused on maximizing speed.

We agree that the origin of the measured speed-up should be demonstrated more clearly. In the revised manuscript, we will make the existing information easier to identify and will add a sensitivity analysis comparing the global and DD solvers for different numbers of subdomains and transmission-parameter choices, together with the corresponding iteration counts, execution times and errors.

12. **The reported timing results for the domain-decomposition method are difficult to assess without the corresponding Schwarz iteration counts. Please report the number of iterations required for convergence for each numerical experiment, together with the stopping tolerance, relaxation strategy, and any variation with frequency or subdomain count. This information is essential for interpreting both the computational cost and the scalability of the method.**

A: We agree that these quantities should be easier to identify and assess together. The stopping tolerance, frequency and number of subdomains are currently reported in the descriptions of Tables 3–4 for the first problem and Table 7 for the realistic case. The Schwarz iteration counts for the reported timing cases are shown for each polarization in Figs. 5–6 and Fig. 8.

In the revised manuscript, we will consolidate this information and report the Schwarz iteration counts and the relevant convergence settings more explicitly for each numerical experiment, including any variation with frequency or subdomain count.

13. **The paper appears to use a single reduced basis dimension for each polarization across all subdomains. Since the POD bases are constructed locally, did the authors consider selecting the basis dimension independently for each subdomain using a local truncation criterion? Such an approach could further reduce the online cost and better reflect the varying complexity of the local solution manifolds.**

A: We agree with the point raised by the reviewer and would like to clarify that the POD truncation is already performed locally in the benchmark problem. Each subdomain has its own snapshot matrix and its own POD basis. When the SVD truncation criterion in Eq. (24) is applied, the resulting n_{POD} may therefore differ between subdomains and polarizations. This is why Table 3 reports the largest reduced-system dimension among the subdomains and polarizations rather than a single dimension that is necessarily shared by all of them.

For some of the realistic-model experiments, the same n_{POD} is intentionally prescribed across the subdomains in order to study the computational scaling and error as a controlled function of the reduced dimension. This is not a requirement of the DD-POD formulation.

We will clarify this distinction in the revised manuscript.

14. **I am surprised the local ROMs are so much faster given that the number of modes in the local ROMs seems to be quite large (over 100 possibly). I assume the global ROM had the same number of modes? Please comment. It is not entirely clear whether the reported ROM assessments are reproductive or predictive. Please clarify whether the test conductivity models are independent of the training set and, if so, how the training and testing parameter sets were partitioned. If the assessments are primarily reproductive, this should be stated explicitly. Moreover, I would like to see some predictive results, as the predictive regime is the true regime in which the ROM should be assessed.**

A: Regarding computational performance, the speed-up of the local ROMs does not primarily result from having fewer POD modes. As noted above, in the benchmark the local reduced dimensions can be comparable to, or even slightly larger than, the global reduced dimension. The main benefit comes from applying the POD projection to much smaller subdomain matrices and carrying out these operations in parallel. In the realistic example, the global matrix contains approximately 18^2 times as many entries as an individual local matrix. Consequently, projection of the global operator becomes considerably more expensive as n_{POD} increases, whereas the corresponding local operations remain substantially cheaper as seen Table 7 and Figs. 7 and 8. For the realistic case the POD-DD and POD-G were compared for the same number of modes.

We agree that the distinction between reproductive and predictive ROM tests should be stated explicitly.

For the DTM1 benchmark, the test is predictive in parameter space. The true anomalous resistivities used for the target model are

$$10, \quad 1, \quad 10{,}000\ \Omega\,\mathrm{m},$$

whereas the training grid contains only

$$20{,}000, \quad 200, \quad 100, \quad 2, \quad 0.2\ \Omega\,\mathrm{m}.$$

Therefore, the target conductivity configuration is not one of the 125 training realizations.

For the realistic example, the five resistivities varied in the target model are

$$500, \quad 200, \quad 100, \quad 20, \quad 10\ \Omega\,\mathrm{m}.$$

The corresponding training grid is defined by

$$\rho_1, \rho_2, \rho_3, \rho_4 \in \{1000,\ 216,\ 47,\ 10\}\ \Omega\,\mathrm{m},$$

and

$$\rho_5 \in \{100,\ 21.6,\ 4.7,\ 1.0\}\ \Omega\,\mathrm{m},$$

resulting in $4^5 = 1024$ training realizations. Therefore, the target parameter configuration is not one of the training realizations.

In the revised manuscript, we will state the training and testing parameter sets explicitly and identify these tests as predictive.

15. **I don't believe the authors say how many subdomains are used in the realistic example, or how the DD is done. Please add this information.**

A: We agree that this information is currently distributed across different parts of the manuscript and could be presented more clearly. Sect. 4.1 and Fig. 1 describe how the computational domain is partitioned, while the number of subdomains used in each numerical experiment is reported in the corresponding table descriptions through N_s. The first problem uses $N_s = 9$, whereas the realistic case uses $N_s = 20$.

In the revised manuscript, we will state the number and arrangement of the subdomains directly in the description of each numerical example so that the DD configuration is immediately clear.

16. **An interesting observation is that DD-POD is more accurate than G-POD for the first example but less accurate for the second. This reversal is not discussed in the manuscript. Some insight into the factors governing when local POD bases outperform a global basis (or vice versa) would strengthen the paper and help readers understand the advantages and limitations of the proposed approach.**

A: We agree that this observation should be discussed in the analysis of the numerical results. The relative accuracy of G-POD and DD-POD may be influenced by several aspects of the numerical setup, including the complexity of the model, the snapshots used to construct the reduced bases and the parameters used in the domain-decomposition scheme.

In the present experiments, the DD and DD-POD parameters were selected primarily to favour computational efficiency rather than to minimize the approximation error. In addition, the snapshot parameter values for each problem were selected heuristically to span a representative range. These differences may contribute to the behaviour observed between the two examples, although the present experiments do not isolate the effect of each factor.

We will add a discussion of these points in the revised manuscript and will avoid attributing the observed reversal to a single cause that has not been independently demonstrated.

17. **The performance of Robin–Schwarz methods can depend strongly on the choice of the Robin transmission parameter. The manuscript would benefit from additional discussion of how this parameter was selected. Was a single value used throughout, or was it tuned for each frequency and test case? Some discussion of the sensitivity of the convergence and runtime to the Robin parameter would help readers assess the robustness and reproducibility of the proposed approach.**

A: We agree with this observation. For each model, the Robin transmission parameter was selected to reduce the number of Schwarz iterations and, consequently, the execution time. The same Robin transmission parameter was used across all frequencies within each model.

We agree that the sensitivity to this parameter should be documented more clearly. In the revised manuscript, we will add a sensitivity study showing how the choice of the Robin transmission parameter affects the number of Schwarz iterations, the execution time and the resulting error.

2 Minor comments

- **In the introduction, the authors say that "matrix-based approaches generally lead to ill-posed inverse problems". This statement is misleading and conflates the issue. The problem is not that the approaches are matrix-based, it is that the inverse problem is ill-posed when posed in the deterministic setting. Often regularization is used to mitigate the issue. Please reword this phrase to make it more accurate.**

  A: We agree with the reviewer. The ill-posedness is a property of the inverse problem and is not a consequence of using a matrix-based numerical method. We will rephrase this sentence in the revised manuscript to remove this misleading implication and to describe the role of regularization more accurately.

- **The error metric in equation (12) is a bit weird to me. Normally, you would create a relative error, so you would divide simply by**

$$\sum_j \left\| E_j^{(n+1)} \right\|.$$

  I am curious, why do the authors adopt the metric in (12) rather than the usual one?

  A: We thank the reviewer for pointing this out. The metric in Eq. (12) is used as a convergence criterion between two consecutive Schwarz iterations, rather than as an error with respect to a reference solution. It is the convergence criterion adopted in the DD formulation and is written in the symmetric form

$$2\sqrt{\frac{\sum_j \left| E_j^{(n+1)} - E_j^{(n)} \right|^2}{\sum_j \left| E_j^{(n+1)} + E_j^{(n)} \right|^2}}.$$

  With this definition, neither $E^{(n)}$ nor $E^{(n+1)}$ is taken as the reference solution. When two consecutive solutions are close, $E^{(n+1)} \approx E^{(n)}$, the criterion behaves similarly to a conventional relative change between successive iterates.

  We will add a short explanation after Eq. (12) to clarify the purpose and motivation of this convergence criterion.

- **In Figure 1, why is the DD only done into horizontal subdomains?**

  A: The horizontal decomposition shown in Fig. 1 is the decomposition adopted in the present implementation and in the numerical examples considered in this work. The solver uses a structured mesh, and the selected partition preserves the horizontal discretization between neighbouring subdomains.

  We will clarify in the revised manuscript that the decomposition shown in Fig. 1 corresponds to the implementation used in this study.

- **In Figure 5, please clarify what the integers next to "XY=" and "YX=" mean.**

  A: The integers next to "XY=" and "YX=" indicate the number of Schwarz iterations required to reach the prescribed convergence criterion for each polarization. We agree that this is not sufficiently clear in the current figure. We will state this explicitly in the revised caption of Fig. 5 and in the corresponding figures.

- **Tables 5–6 report absolute errors, which are generally not very informative. Please switch to relative errors.**

  A: We agree with the reviewer. In the revised manuscript, we will replace the maximum absolute errors currently reported in Tables 5–6 with relative errors, which will provide a more informative comparison across the different quantities and mesh resolutions. For consistency with the accuracy analysis used later in the manuscript, we will report the relative L_2 error between the reduced-order and corresponding full-order solutions.

- **Please change "validation" to "verification". The former requires comparing to real-world data from the field or experiments. What you are doing is effectively verification.**

  A: We agree with the reviewer. The numerical comparisons performed in this work constitute verification rather than validation. In the revised manuscript, we will replace "validation" with "verification" wherever this terminology is used in this context.

- **At first, I was wondering about extensions to nonlinear problems, but when I looked it up, it appears the linaer MT equations are the state-of-the-art ones to use; there is not a relevant nonlinear version. You could highlight this when you present the equations or in the introduction, in case people criticize the approach for being linear.**

  A: We thank the reviewer for this suggestion. For a prescribed conductivity model, the frequency-domain MT forward problem considered in this work is linear with respect to the electromagnetic fields. The nonlinearity relevant to MT inversion arises from the dependence of the predicted responses on the conductivity model rather than from a nonlinear electromagnetic state equation.

  We will add a short clarification when introducing the governing equations in the revised manuscript.