IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 May 2023, accepted 19 June 2023, date of publication 23 June 2023, date of current version 6 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3288997

RESEARCH ARTICLE

Bayesian Experimental Design for Efficient Sensor Placement in Two-Dimensional Electromagnetic Imaging

ALI IMRAN SANDHU[1], (Member, IEEE), BEN MANSOUR DIA[1], OLIVER DORN[2], (Senior Member, IEEE), AND PANTELIS SOUPIOS[3]

[1]Center for Integrative Petroleum Research, College of Petroleum Engineering and Geosciences, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran 31261, Saudi Arabia
[2]Department of Mathematics, Alan Turing Building, The University of Manchester, M13 9PL Manchester, U.K.
[3]Department of Geosciences, College of Petroleum Engineering and Geosciences, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran 31261, Saudi Arabia

Corresponding author: Ali Imran Sandhu (ali.sandhu@kfupm.edu.sa)

ABSTRACT Careful sensor placement is crucial in electromagnetic imaging experiments as it significantly impacts the quality and accuracy of the measurements. This study examines the placement of a network of sensors to advance the Bayesian learning with the aim of achieving a minimal level of uncertainty in a qualitative imaging regime. The quality of the measured data, associated with a network of sensors, is assessed by computing the expected Kullback-Leibler divergence between the prior and the posterior distributions, wherein the Laplace approximation is invoked to reduce the associated computational cost. The numerical experiment is carried out to evaluate various sensor placement scenarios to identify the network geometry that can enhance the quality of inversion.

INDEX TERMS Electromagnetic imaging, Bayesian experimental design, optimal sensor placement.

I. INTRODUCTION

Electromagnetic (EM) imaging is a non-invasive framework to determine the location, shape, and spatial distribution of the dielectric permittivity and electrical conductivity of a target of interest from the scattered electric and magnetic fields measured away from it [1], [2], [3], [4]. EM imaging finds various applications in biomedical imaging [5], [6], non-destructive testing [7], see-through wall imaging [8], and hydrocarbon exploration [9], [10], [11], where it can be used in conjunction with other geophysical methods [12], [13], [14], [15], such as seismic and non-seismic imaging [16], [17], to provide a more comprehensive understanding of the subsurface environment [18], [19]. EM imaging plays a crucial role in advancing our knowledge of the subsurface at different scales and supports decision-making in the fields of mining, environmental management, and natural resource exploration and exploitation [1].

Optimal sensor placement is fundamental in EM imaging experiments because it directly affects the quality and accuracy of the measured data. Non-optimal sensor placement can result in a loss of valuable information, increase the measurement noise and reduce the signal-to-noise ratio, further compromising the data quality. To ensure a successful EM imaging experiment, careful consideration of the target and the measurement environment is necessary, and the sensors must be positioned in a way that allows for the most complete and accurate measurement of the scattered electric and magnetic fields. Placing sensors at carefully chosen regular or irregular intervals across a predefined grid [20], [21], [22] ensures that the data collected will be of sufficient quality, resolution, and will be evenly distributed. Despite being simple, easy to implement, and providing a uniform and comprehensive coverage of the target area, this approach may not be efficient and may result in over-sampling or under-sampling or collecting unnecessary data in certain areas.

The Bayesian approach [23] to EM imaging characterizes the uncertainty in the learned target of interest via a probability density function (PDF). The degree of uncertainty

The associate editor coordinating the review of this manuscript and approving it for publication was Roberta Palmeri.

depends on various factors such as the characterization of prior knowledge, the noise structure, and, most importantly, the quality of the data used for the knowledge update. Acquiring data for characterizing or monitoring the subsurface using EM, is a challenging and resource-intensive endeavor, with no guarantee of data quality that can be relied upon for accurate parameter estimation. To overcome these budgetary and logistics constraints, the simulations-based optimal experimental design (OED) technique [24] is the preferred method for obtaining reliable data for efficient parameter estimation.

The Bayesian framework for OED minimizes the level of uncertainty associated with the estimation of the parameter of interest by gathering the most informative data available. It has been applied in various contexts, depending on factors such as the analysis objective, the forward model, and the type of uncertainties present in the data. Its primary purpose is to control the experimental conditions for data acquisition, which can be broken down into two main tasks: approximating the objective function and selecting the most informative setup that could be achieved either by exploring the design space in a combinatorial fashion [25] or by incorporating optimization algorithms [26], [27], [28], [29]. In general, the objective function is computationally challenging to approximate due to its nested form of expectations over high-dimensional variables. A double loop Monte Carlo estimator is presented in [30], which involves performing Monte Carlo sampling twice. Various techniques have been proposed to improve the computational cost associated with the latter method. These include using the Laplace approximation [31], [32], [33], [34], employing polynomial chaos expansions with pseudo-spectral projection [28], using a lower-bound estimate of the objective function [35], and leveraging the scalability of averaging the estimates of the information matrices [36]. Bayesian experimental design provides an effective framework for optimizing experimental design and resource allocation, and has shown notable success in various fields including medical science [37], [38], material science [33], [39], and geophysical applications [40], [41], [42], [43]. In [44] a Bayesian framework to statistically optimize reconstruction of 3-D objects using limited number of radio-graphs is presented. Later in [45], the authors applied the theory of optimal experimental design to optimize sensor positions and the operating frequency for sensing buried, conducting and ferrous, targets via a moving EM induction sensor.

This paper aims to investigate the use of Bayesian OED for efficient sensor placement in an EM imaging setup while identifying the most informative data that would result in the minimum uncertainty level in the posterior, given a fixed number of transmitters and receivers. This involves assessing uncertainty levels during the learning process of the actual parameter of interest, which happens to be the center of a circular phantom (cylindrical scatterer). The Bayesian inverse problem involves the development of a framework to update the prior knowledge about the parameter of interest through the use of observational data and a forward predictive mathematical model. By applying Bayes' rule, one can obtain the updated probability distribution, also known as the posterior distribution or the conditional distribution. To evaluate the information content of measured data, a widely employed utility function, called the Kullback-Leibler divergence [46], [47], is applied, which measures the relative information entropy between the prior and the posterior distributions given a particular sensor network setup: wherein higher values indicating more informative data regarding the parameter of interest. Due to the inherent uncertainties present in the data, the expected Kullback-Leibler divergence is utilized as the objective function, also known as the expected information gain (EIG). In reference to sensitivity analysis to advance nonlinear inversion [48], the physical interpretation of EIG is to achieve an efficient estimate of a phantom location. For linear forward predictive model, the approximation of the EIG leads to a standard form A or D-optimality criterion for choosing optimal location for sensors [49]. However, for nonlinear forward model, it needs to mitigate the computational burden associated with approximating the EIG. In this work, the Laplace approximation (LA) [50] is incorporated, which is a second-order Taylor expansion of the negative logarithmic of the posterior distribution around the posterior mode. To circumvent the potential drawbacks of using optimization algorithms on continuous design space, including slower convergence, increased computational time, and the requirement to impose the Lipschitz continuity on the EIG, we pursued an exploration of the design space by utilizing the combinatorial optimization approach [25], which evaluates the EIG over a discretized design space and picks the one having the maximum value. Additional analysis of the quality of the reconstruction will be conducted using the continuous ranked probability score (CRPS) in cases where a direct comparison of posterior distributions may not be obvious. This work showcases numerical experiments that (i) operate in the qualitative imaging regime (methods that only provide information about the shape and location of the scatterers being tested, lacking information about the spatial distribution of the material properties), and (ii) assess the effectiveness of the proposed framework. Furthermore, this study provides a preliminary investigation into three-dimensional geophysical inversion problems, with ongoing efforts underway to extend and build upon the foundation established by this work.

The paper is structured as follows: Section II-A presents the mathematical formulation and discretization of the forward problem. Section II-B introduces the Bayesian inversion framework, including the parameterization of uncertainties in the data, the characterization of the prior, and the posterior probability density function. In Section II-C, a data informativeness criterion for Bayesian experimental design is introduced, along with the Laplace approximation for the expected information gain. The efficacy of the proposed framework is demonstrated with numerical examples in Section III, followed by the conclusions in Section IV.

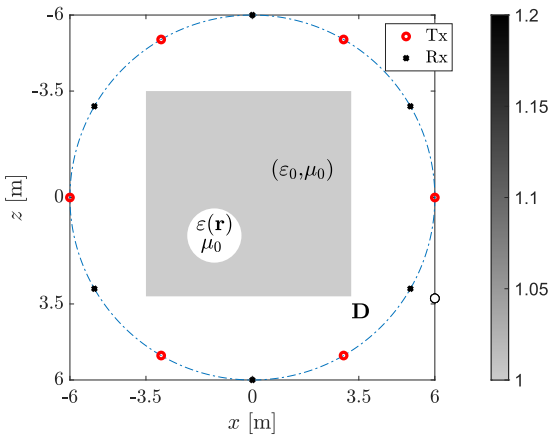


FIGURE 1. Description of the 2D EM imaging problem.

II. FORMULATION

A. ELECTROMAGNETIC SCATTERING EQUATION AND ITS DISCRETIZATION

Let D be an open-bounded subset of $\mathbb{R}^2$, that represents the support of an investigation domain which resides in an unbounded background medium. A point inside D is denoted by $\mathbf{r}$ with coordinates (x, z) in the reference (O, Ox, Oz). The domain D is characterized by having a dielectric permittivity $\varepsilon(\mathbf{r})$ and a constant permeability $\mu(\mathbf{r}) = \mu_0$. In the background medium, the dielectric permittivity and the permeability are $\varepsilon(\mathbf{r}) = \varepsilon_0$ and $\mu(\mathbf{r}) = \mu_0$, as depicted in Figure 1. Without loss of generality, it is assumed that the material properties of the investigation region, D, are invariant along the $z-$coordinate. The investigation domain is surrounded by n_t transmitters and n_r receivers, which are located at $\mathbf{r}_i^{\mathrm{t}}$, $i = 1, \ldots, n_t$ and $\mathbf{r}_m^{\mathrm{r}}$, $m = 1, \ldots, n_r$, respectively. The transmitters generate a $z-$polarized transverse-magnetic incident electric field represented by $E_i^{\mathrm{inc}}(\mathbf{r})$, where the subscript i indicates the ith transmitter. Upon excitation, electric current is induced on D which in turn generates, the total $E_i^{\mathrm{tot}}(\mathbf{r})$ and the scattered electric field $E_i^{\mathrm{sca}}(\mathbf{r})$, that satisfy the volume integral equation [51]:

$$E_i^{\mathrm{sca}}(\mathbf{r}) = -k_0^2 \int_D \tau(\mathbf{r}') E_i^{\mathrm{tot}}(\mathbf{r}') G(\mathbf{r}, \mathbf{r}') d\mathbf{r}'. \tag{1}$$

Here, $\tau(r) = \varepsilon(\mathbf{r})/\varepsilon_0 - 1$ is the contrast, $E_i^{\mathrm{tot}}(\mathbf{r})$ is the total electric field induced inside D, $k_0 = \omega\sqrt{\varepsilon_0\mu_0}$ is the free space wave-number, and ω is the angular frequency. Moreover,

$$G(\mathbf{r}, \mathbf{r}') = \frac{1}{4j} H_0^2(k_0 \left|\mathbf{r} - \mathbf{r}'\right|)$$

is the two-dimensional Green's function, where $j^2 = -1$, and H_0^2 is the Hankel function of second-kind and order 0. The total, incident, and scattered electric fields satisfy:

$$E_i^{\mathrm{tot}}(\mathbf{r}) = E_i^{\mathrm{inc}}(\mathbf{r}) + E_i^{\mathrm{sca}}(\mathbf{r}) \tag{2}$$

which renders uniquely solving for $E_i^{\mathrm{tot}}(\mathbf{r})$ satisfying:

$$E_i^{inc}(r) = E_i^{tot}(r) + k_0^2 \int_D \tau(r') E_i^{tot}(r') G(r, r') d\mathbf{r}'. \tag{3}$$

To solve (1) and (3) numerically, D is discretized using N square cells, having support D_n, and the unknowns $\tau(r)$ and $E_i^{tot}(r)$ are approximated as:

$$\tau(r) = \sum_{n=1}^{N} \tau(\mathbf{r}_n) q_n(r), \quad E_i^{\mathrm{tot}}(r) = \sum_{n=1}^{N} E_i^{\mathrm{tot}}(\mathbf{r}_n) q_n(r), \tag{4}$$

where $\mathbf{r}_n$, $n = 1, \ldots, N$ denote the centers of the cells D_n. Moreover, $q_n(r)$ is the pulse basis function defined as

$$q_n(r) = \begin{cases} 1, & r \in D_n \\ 0, & \text{elsewhere.} \end{cases}$$

Substituting (4) in (3) and evaluating the resulting equation at r_n, $n = 1, \ldots, N$ yield the linear discretized system representing the "forward solver":

$$\bar{E}_i^{\mathrm{tot}} = \left[\bar{I} + \bar{G}^D \mathrm{diag}(\bar{\tau})\right]^{-1} \bar{E}_i^{\mathrm{inc}}. \tag{5}$$

Similarly inserting (4) in (1), and evaluating the resulting equation at the receiver locations, r_m^r, $m = 1, \ldots, n_r$ yield the scattered electric fields away from D:

$$\bar{E}_i^{\mathrm{sca}} = -\bar{G}^{\mathrm{R}} \mathrm{diag}\left(\bar{E}_i^{\mathrm{tot}}\right) \bar{\tau} = \bar{H}_i \bar{\tau}. \tag{6}$$

In (5) and (6), $\bar{E}_i^{\mathrm{tot}} = \left(E_i^{\mathrm{tot}}(\mathbf{r}_1), E_i^{\mathrm{tot}}(\mathbf{r}_2), \cdots, E_i^{\mathrm{tot}}(\mathbf{r}_N)\right)^\top$, $\bar{E}_i^{\mathrm{inc}} = \left(E_i^{\mathrm{inc}}(\mathbf{r}_1), E_i^{\mathrm{inc}}(\mathbf{r}_2), \cdots, E_i^{\mathrm{inc}}(\mathbf{r}_N)\right)^\top$, $\bar{E}_i^{\mathrm{sca}} = \left(E_i^{\mathrm{sca}}(\mathbf{r}_1^r), E_i^{\mathrm{sca}}(\mathbf{r}_2^r), \cdots, E_i^{\mathrm{sca}}(\mathbf{r}_{n_r}^r)\right)^\top$, $\bar{I}$ is an identity matrix, $\bar{\tau} = (\tau(\mathbf{r}_1), \tau(\mathbf{r}_2), \cdots, \tau(\mathbf{r}_N))^\top$, and the entries of the matrices $\bar{G}^D$ and $\bar{G}^{\mathrm{R}}$ are

$$\bar{G}^D_{n,n'} = k_0^2 \int_{D_{n'}} G(\mathbf{r}_n, \mathbf{r}') d\mathbf{r}', \quad \bar{G}^{\mathrm{R}}_{m,n'} = k_0^2 \int_{D_{n'}} G(\mathbf{r}_m^r, \mathbf{r}') d\mathbf{r}'.$$

A few comments about the forward solver (5) and scattering field calculation in (6) are in order: (i) Given $\bar{\tau}$ and $\bar{E}^{inc}$, (5) is solved for $\bar{E}^{\mathrm{tot}}$ induced inside D using the iterative generalized minimal residual (GMRES) method. Since matrix $\bar{G}^D$ is toeplitz, the matrix-vector multiplications required by the iterative solver are accelerated using fast Fourier transform (FFT). (ii) Given $\bar{E}^{\mathrm{tot}}$ and $\bar{\tau}$, $\bar{E}^{sca}$ is computed using (6). (iii) (5) and (6) are derived assuming a single excitation. For multiple transmitters configuration, they can simply be cascaded into larger matrix systems where the scattered fields are computed at receiver locations for one transmitter a time.

B. BAYESIAN FORMULATION

1) DATA REPRESENTATION

The dielectric phantom in Fig.1, hereinafter denoted by $\mathcal{P}_t$, is cylindrical in shape, having a dielectric permittivity different from the rest of D, and is characterized by having a fixed radius ϕ with its center at $\boldsymbol{\theta}_t = (x_t, z_t)$. A typical sensor

network having n_t transmitters and n_r receivers is denoted by $\mathcal{S}_{n_r,n_t}$. The scattered electric field happens to be the output of the forward predictive model $\boldsymbol{g}$, and is a function of $\mathcal{P}_t$. The measurements are then formed by adding Gaussian noise to the scattered electric fields:

$$\boldsymbol{y}_k = \boldsymbol{g}(\mathcal{P}_t, \mathcal{S}_{n_r,n_t}) + \boldsymbol{\epsilon}_k, \quad k = 1, \cdots, N_e, \tag{7}$$

where N_e is the number of repetitive experiments and $\boldsymbol{\epsilon}_k$ are independent and identically distributed (i.i.d.) realizations from an zero-mean Gaussian distribution with covariance $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ i.e., $\boldsymbol{\epsilon}_k \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\epsilon})$. The forward problem is modeled in the frequency domain (i.e., a stationary process) wherein for a given transmitter, the scattered electric field observed at the p^{th} receiver is not directly correlated to the scattered electric field recorded at the q^{th} receiver ($p \neq q$). It is therefore appropriate to consider a diagonal form for the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ ($\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = \sigma_{\varepsilon}^2 \bar{I}_{n_t n_r}$ where the $\bar{I}_{n_t n_r}$ is the unit matrix). Given a phantom $\mathcal{P}$, for multiple transmitters, the output of the forward model is cascaded to form:

$$\boldsymbol{g}(\mathcal{P}, \mathcal{S}_{n_r,n_t}) = \left(\bar{E}_1^{\text{sca}}, \bar{E}_2^{\text{sca}}, \cdots, \bar{E}_{n_t}^{\text{sca}}\right)^{\top}. \tag{8}$$

The overall observational data $\boldsymbol{Y}$ assembles the N_e polluted measurements $\boldsymbol{Y} = \{\boldsymbol{y}_k\}_{k=1}^{N_e}$.

2) PRIOR, LIKELIHOOD AND POSTERIOR

In the proposed Bayesian framework, the parameter of interest $\boldsymbol{\theta}_t$ is treated as a random variable, hereinafter, and is denoted by $\boldsymbol{\theta}$. The knowledge about the possible values, that $\boldsymbol{\theta}$ can take, is characterized by a prior PDF $\pi_{\text{pr}}(\boldsymbol{\theta})$ with $\boldsymbol{\theta} = (x(\xi), z(\xi))^{\top}$, $\xi \in \Omega$ where Ω is the set of elementary events. Bayesian inference addresses the update of the knowledge regarding the position of the phantom by incorporating information of the data $\boldsymbol{Y}$ with density $p(\boldsymbol{Y})$. The resulting density, conditioned with the data, is the posterior PDF, and is given by Bayes' theorem

$$\pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r}) = \frac{p(\boldsymbol{Y}|\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r})\pi_{\text{pr}}(\boldsymbol{\theta})}{p(\boldsymbol{Y}|\mathcal{S}_{n_t,n_r})}, \tag{9}$$

where $p(\boldsymbol{Y}|\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r})$ is the likelihood function measuring the distribution of the data $\boldsymbol{Y}$, collected on the network $\mathcal{S}_{n_t,n_r}$, for each possible value of $\boldsymbol{\theta}$. According to the data representation (7), the likelihood function is Gaussian with weighted L^2-norm given by

$$\begin{aligned} &p(\boldsymbol{Y}|\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r}) \\ &\quad = \det\left(2\pi\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}\right)^{-\frac{N_e}{2}} \\ &\qquad \times \exp\left(-\frac{1}{2}\sum_{k=1}^{N_e} \left\|\boldsymbol{y}_k(\mathcal{S}_{n_t,n_r}) - \boldsymbol{g}(\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r})\right\|^2_{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1}}\right), \end{aligned} \tag{10}$$

where for a given vector $\boldsymbol{x}$ and a covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$, the matrix norm is given by $\|\boldsymbol{x}\|^2_{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1}} = \boldsymbol{x}^T \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1} \boldsymbol{x}$, also known as the Mahalanobis norm. To characterize the prior density $\pi_{\text{pr}}(\boldsymbol{\theta})$, the domain D is assumed rectangular $[-d_0, d_0] \times [-d_0, d_0]$ where d_0 is a positive real number. Moreover an improper uniform prior probability density is utilized for $\boldsymbol{\theta}$:

$$\boldsymbol{\theta} \sim \mathcal{U}\left([-\theta_m, \theta_m]^2\right), \text{ where } \quad \theta_m = d_0 - \phi.$$

Note that, although the uniform form of the prior density seems non-informative, since $\pi_{\text{pr}}(\boldsymbol{\theta})$ is constant, the likelihood setting guarantees a proper posterior distribution.

C. BAYESIAN EXPERIMENTAL DESIGN

1) EXPECTED INFORMATION GAIN

Bayesian experimental design seeks to pinpoint the optimal network $\mathcal{S}^*_{n_t,n_r}$ from a set of potential networks $\mathcal{S}_{n_t,n_r}$ that provides the most informative data, resulting in the least uncertainty in estimating $\boldsymbol{\theta}$. The utility function to assess the quality of a network is the Kullback-Leibler divergence, denoted by D_{KL} [47], [52]. It evaluates the amount of acquired information about $\boldsymbol{\theta}$ between the prior $\pi_{\text{pr}}(\boldsymbol{\theta})$ and the posterior $\pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})$ and is given by

$$\begin{aligned} &D_{\text{KL}}\left(\pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})||\pi_{\text{pr}}(\boldsymbol{\theta})\right) \\ &\quad = \int_{\Theta} \log\left(\frac{\pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})}{\pi_{\text{pr}}(\boldsymbol{\theta})}\right) \mathrm{d}\pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r}). \end{aligned} \tag{11}$$

The larger the value of D_{KL} is, the more informative the data $\boldsymbol{Y}$ is about the parameter of interest $\boldsymbol{\theta}_t$. The uncertainty in the data is tackled by considering the expectation of D_{KL} over the sample space $\mathcal{Y} \subseteq \mathbb{R}^q$, with $q = n_t n_r$. The EIG $\mathrm{I}(\mathcal{S}_{n_t,n_r})$ is then given by

$$\begin{aligned} \mathrm{I}(\mathcal{S}_{n_t,n_r}) &= \int_{\mathcal{Y}} \int_{\Theta} \log\left(\frac{\pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})}{\pi_{\text{pr}}(\boldsymbol{\theta})}\right) \\ &\qquad \times \pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})\mathrm{d}\boldsymbol{\theta} p(\boldsymbol{Y}|\mathcal{S}_{n_t,n_r})\mathrm{d}\boldsymbol{Y} \\ &= \int_{\Theta} \int_{\mathcal{Y}} \log\left(\frac{p(\boldsymbol{Y}|\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r})}{p(\boldsymbol{Y}|\mathcal{S}_{n_t,n_r})}\right) \\ &\qquad \times p(\boldsymbol{Y}|\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r})\mathrm{d}\boldsymbol{Y}\pi_{\text{pr}}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}. \end{aligned} \tag{12}$$

The latter equality in (12) follows from Bayes' rule. In general, the approximation of I requires a large number of evaluations of the forward model and exhibits several computational issues due to the double integration over high dimensional domains. In the next section, the Laplace approximation is introduced and then explored to reduce the computational cost of estimating I.

2) LAPLACE APPROXIMATION

In the absence of multiple scatterers, the posterior distribution of $\boldsymbol{\theta}$ is concentrated around the MAP (Maximum A Posteriori) estimate $\hat{\boldsymbol{\theta}}$. The Gaussian approximation of the posterior distribution then provides an accurate inversion. The LA [50] is known to be an efficient computational method in the sense that it significantly reduces the computational cost of profiling the posterior distribution. The LA is based on the second-order Taylor expansion of the logarithmic posterior distribution, around its mode. Adopting LA for the estimation

of the EIG reduces the double integration (12) to a single integral over the parameter space. This credits an important asset of computational efficiency to LA although it comes with an extra computational bias that is inversely proportional to the number of repetitive experiments. The posterior distribution of $\boldsymbol{\theta}$, according to Bayes' rule (9) with the likelihood (10), is given by

$$\pi_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r}) = \det(2\pi\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})^{-\frac{N_e}{2}} \times \frac{\exp\left(-\frac{1}{2}\sum_{k=1}^{N_e}\left\|\boldsymbol{y}_k(\mathcal{S}_{n_t,n_r}) - \boldsymbol{g}(\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r})\right\|^2_{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1}}\right)\pi_{\text{pr}}(\boldsymbol{\theta})}{p(\boldsymbol{Y}|\mathcal{S}_{n_t,n_r})},$$

and its Gaussian approximation $\tilde{\pi}_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})$ has the form

$$\tilde{\pi}_{\text{pos}}(\boldsymbol{\theta}|\boldsymbol{Y}, \mathcal{S}_{n_t,n_r}) = (2\pi)^{-\frac{d}{2}}\det\left(\hat{\boldsymbol{\Sigma}}\right)^{-\frac{1}{2}} \times \exp\left(-\frac{1}{2}\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2_{\hat{\boldsymbol{\Sigma}}^{-1}}\right), \tag{13}$$

where $d = 2$, stands for the dimension of the parameter of interest, and $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})$, which by definition is given by

$$\arg\min_{\boldsymbol{\theta}\in\Theta}\left[\frac{1}{2}\sum_{k=1}^{N_e}\left\|\boldsymbol{y}_k(\mathcal{S}_{n_t,n_r}) - \boldsymbol{g}(\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r})\right\|^2_{\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}^{-1}} - h(\boldsymbol{\theta})\right] \tag{14}$$

and

$$\hat{\boldsymbol{\Sigma}}^{-1} = N_e\boldsymbol{J}(\hat{\boldsymbol{\theta}}(\boldsymbol{Y}, \mathcal{S}_{n_t,n_r}))^\top\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}{}^{-1}\boldsymbol{J}(\hat{\boldsymbol{\theta}}(\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})) - \nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}h(\hat{\boldsymbol{\theta}}(\boldsymbol{Y}, \mathcal{S}_{n_t,n_r})) + \mathcal{O}_{\mathbb{P}}\left(\sqrt{N_e}\right)$$

is the inverse Hessian matrix of the negative logarithm of the posterior PDF evaluated at $\hat{\boldsymbol{\theta}}$. Moreover

$$h(\boldsymbol{\theta}) = \log\pi_{\text{pr}}(\boldsymbol{\theta}) \quad \text{and} \quad \boldsymbol{J}(\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r}) = -\nabla_{\boldsymbol{\theta}}\boldsymbol{g}(\boldsymbol{\theta}, \mathcal{S}_{n_t,n_r}).$$

It has been shown [50] that, for a sufficiently large number of repetitive experiments N_e,

$$\hat{\boldsymbol{\theta}}(\mathcal{S}_{n_t,n_r}) = \boldsymbol{\theta}_t + \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{N_e}}\right), \tag{15}$$

where $\mathcal{O}_{\mathbb{P}}$ refers to the big $\mathcal{O}$ in probability and $\boldsymbol{\theta}_t$ is the true value of the parameter of interest. Incorporating (13), the LA formulation of the EIG follows:

$$\mathrm{I} = \int_{\Theta}\left[-\frac{1}{2}\log\left((2\pi)^d\det\left(\hat{\boldsymbol{\Sigma}}(\boldsymbol{\theta})\right)\right) - \frac{d}{2} - h(\boldsymbol{\theta})\right]\pi_{\text{pr}}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} + \mathcal{O}\left(\frac{1}{N_e}\right). \tag{16}$$

In the next section, the Monte Carlo sampling of (16) is detailed.

3) ESTIMATION USING THE MONTE CARLO-LAPLACE APPROXIMATION

The Monte Carlo Laplace Approximation (MCLA) estimator of the EIG is established through the use of Monte Carlo sampling:

$$\mathrm{I}_{la}(\mathcal{S}_{n_t,n_r}) \overset{\text{def}}{=} \frac{1}{N_s}\sum_{l=1}^{N_s}\left(-\frac{1}{2}\log\left((2\pi)^d\det\left(\hat{\boldsymbol{\Sigma}}(\boldsymbol{\theta}_l)\right)\right) - \frac{d}{2} - h(\boldsymbol{\theta}_l)\right), \tag{17}$$

where the samples $\boldsymbol{\theta}_l$, $l = 1, \cdots, N_s$, are realizations from the prior distribution $\pi_{\text{pr}}(\boldsymbol{\theta})$. Note that the MAP $\hat{\boldsymbol{\theta}}$ is replaced by the parameter of interest $\boldsymbol{\theta}_t$, that is viewed as a random variable in the MCLA estimator (17). The Jacobian matrix $\boldsymbol{J}$ is approximated using central finite differences, resulting in an average computational burden $2dN_s\hbar^{n_t,n_r}$, where $\hbar^{n_t,n_r}$ is the computational cost of one evaluation of the forward model for a network of n_t transmitters and n_r receivers. The total error introduced [33], for approximating I by I_{la} can be estimated by

$$|\mathrm{I} - \mathbb{E}[\mathrm{I}_{la}]| \leq C_1\hbar^{n_t,n_r} + \frac{C_2}{N_e} + o(\hbar^{n_t,n_r}), \tag{18}$$

$$\mathbb{V}[\mathrm{I}_{la}] = \frac{C_3}{N_s}, \tag{19}$$

where C_1 is a constant dependent on the total number of elements, N, and the excitation frequency, C_2 is associated with the error introduced with the Laplace approximation, and C_3 is another constant accounting for the statistical error. The EIG approximation using the estimator (17) with a sample size of N_s is described in Algorithm 1.

Algorithm 1 Computation of $\mathrm{I}_{la}(\mathcal{S}_{n_t,n_r})$

Input:

Prior distribution $\pi_{\text{pr}}(\boldsymbol{\theta})$

Forward model $\boldsymbol{g}$

Network of sensors $\mathcal{S}_{n_t,n_r}$

Output: Expected Information gain I_{la}

1) Generate $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_{Ns}$ from $\pi_{\text{pr}}(\boldsymbol{\theta})$
2) **for** l in $\{1, \cdots, N_s\}$ **do**
   a) Evaluate $h(\boldsymbol{\theta}_l)$ and $\nabla_{\boldsymbol{\theta}}\nabla_{\boldsymbol{\theta}}h(\boldsymbol{\theta}_l)$
   b) Approximate the Jacobian $\nabla_{\boldsymbol{\theta}}g(\boldsymbol{\theta}_l)$
   c) Compute the Hessian $\hat{\boldsymbol{\Sigma}}(\boldsymbol{\theta}_l)$
3) Return the arithmetic mean I_{la} given at (17).

III. NUMERICAL RESULTS

In particular, we consider (i) studying the EIG profile when only a limited number of transmitters and receivers are available, (ii) fixing the number of transmitters and receivers and varying their location, (iii) what could be the optimum number of transmitter and receivers that would provide enough

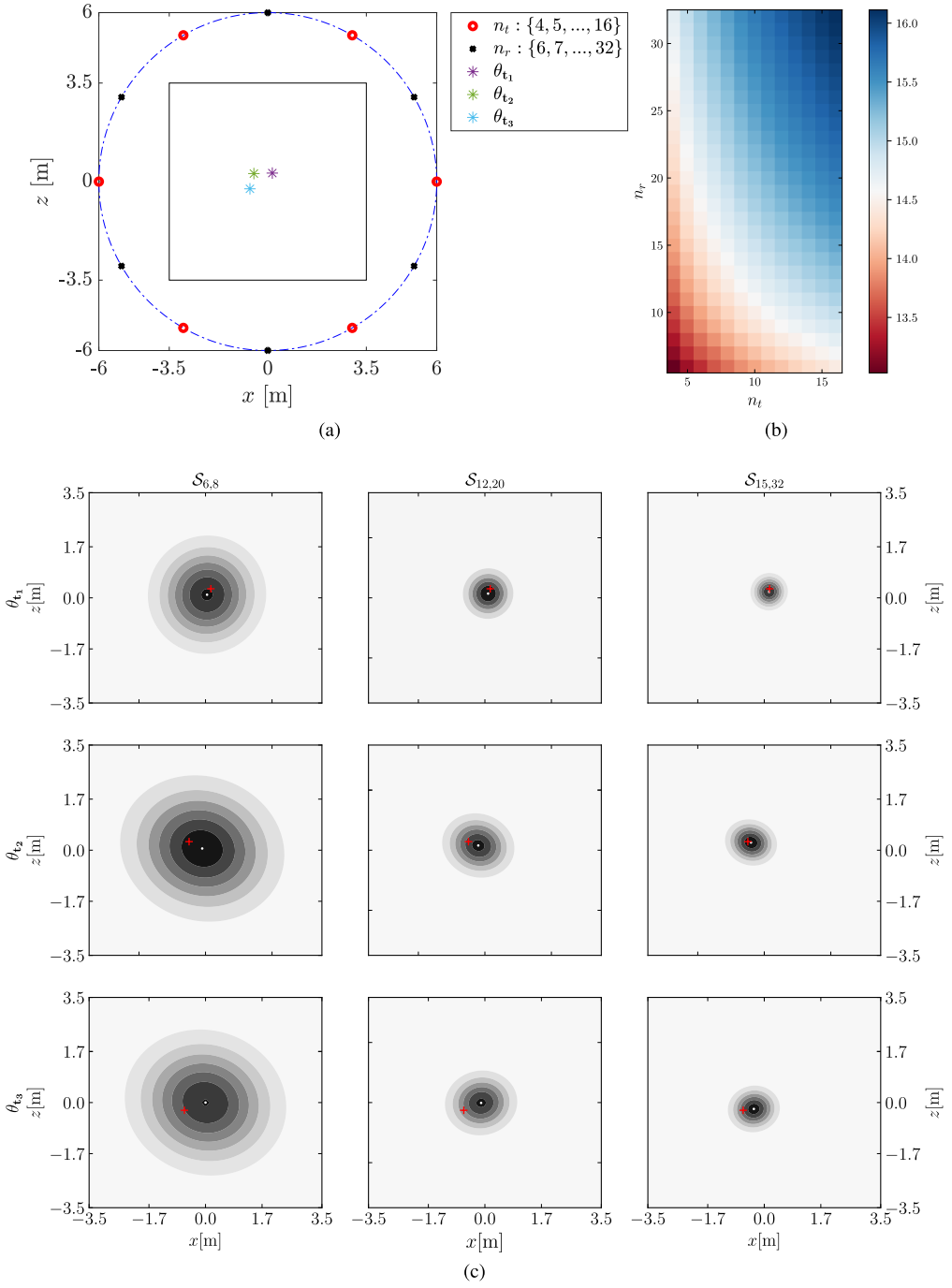


FIGURE 2. (a) problem setup (one phantom is considered at a time), (b) the EIG (approximated by (17), "dimensionless") as a function of varying number of transmitters and receivers, and (c) The uncertainty reflected in the posterior distributions for three different phantom locations (row wise) and for three different sensor setups (column wise). The red + indicates θ_t, the true center location of the cylindrical phantom.

and non-redundant information. By looking at the EIG profile, a few measurement setups are picked, and for each setup, multiple true phantoms (with a random $\boldsymbol{\theta}_t$, and considering only one at a time) are placed in D, and the posterior distribution is plotted to demonstrate the level of uncertainty. The numerical tests presented in this section does not pertain individually to a different physical application, rather they analyze different measurement configurations, even for one common application. The potential applications involve but are not limited to externally monitoring the interior of a chamber filled with dangerous (e.g., radioactive) waste [53], through-wall imaging radar, or choosing few boreholes from many available to collect geophysical measurements [48], and when few sensors are at hand: in such cases, simulations

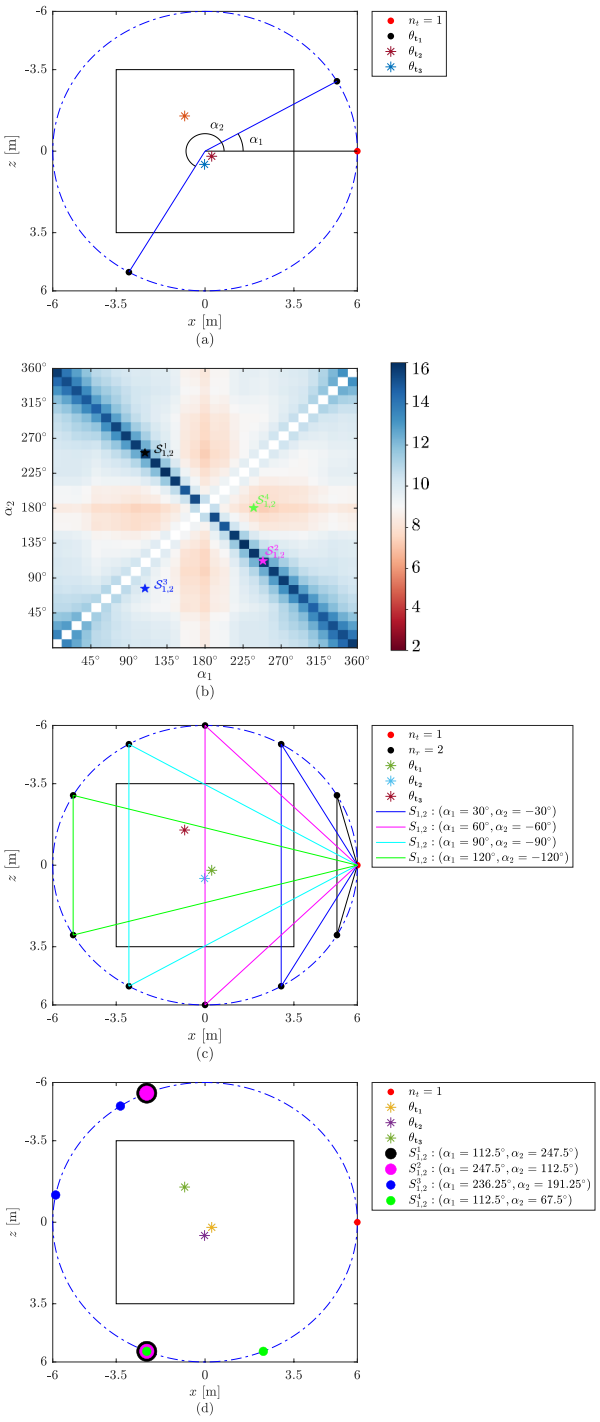


FIGURE 3. (a) problem setup: two receivers are independently rotated w.r.t a fixed transmitter, (b) the EIG (approximated by (17), "dimensionless") as a function of the angular positions (labeled along the horizontal and vertical axes) of both the receivers, (c) conclusive findings: receivers to be positioned at an equal and opposite angular distance w.r.t the transmitter, (d) selected sensor setups to analyze their effect on the accuracy of inversion.

are cheap but deploying sensors into boreholes by trial and error might be expensive.

Moreover, unless stated otherwise, for all the simulations, (i) the frequency of operation is set to 120 MHz, (ii) the scattering object is a dielectric cylinder of radius 0.25λ having permittivity fixed to 1.2, and its position is characterized by the spatial co-ordinates of its center $\boldsymbol{\theta}_t$, (iii) the investigation domain has physical dimension of 7 m $\times$ 7 m which corresponds to electrical dimensions of about $3\lambda \times 3\lambda$ at 120 MHz, and (iv) the transmitters and receivers are setup on a circular path surrounding the investigation domain. Once $\hat{\boldsymbol{\theta}}$ is estimated, $\bar{\tau}$ is readily reconstructed by forming a cylinder of radius 0.25λ at $\hat{\boldsymbol{\theta}}$.

Herein, to simulate realistic electromagnetic (EM) measurements, the synthetic data is contaminated with additive white (wideband) Gaussian noise, such that the signal-to-noise ratio is maintained at 25dB. Additional analysis of the quality of the reconstruction will be conducted using the continuous ranked probability score (CRPS) in cases where a direct comparison of posterior distributions may not be obvious. CRPS is a score function that compares a single ground truth value to a cumulative distribution function F and is given by

$$\mathrm{CRPS}(F, y) = \int \left(F(t) - \mathbb{1}_{\{t \geq y\}}(t)\right)^2 \, dt,$$

where $\mathbb{1}$ stands the indicator function. The CRPS is expressed in the same unit as the observed variable y, that can be the abscissa x or the ordinate z. It generalizes the mean absolute error as it reveals both the bias and the level of uncertainty of each marginal posterior distribution marginal in the way that the small its value is the better is the marginal accuracy of the reconstruction.

A. TEST 1: EXHAUSTING COMBINATIONS OF UNIFORMLY SPACED SENSORS

In this example, the EIG is analyzed as a function of a gradual increase in the number of transmitters and receivers, positioned equally spaced on a circle surrounding the investigation domain, Fig. 2a. The number of transmitters and receivers in a measurement setup exhausted all possible combinations from the sets $\{4, 5, \ldots, 16\}$ and $\{6, 7, \ldots, 32\}$ respectively, Fig. 2a. The EIG in Fig. 2b reflects a gradual increase with the increase in the number of transmitters and receivers. To study the effect of a particular measurement setup (i.e., a given number of transmitters and receivers), on the accuracy of inversion, three different measurement setups $\mathcal{S}_{6,8} : \{n_t = 6, n_r = 8\}$, $\mathcal{S}_{12,20} : \{n_t = 12, n_r = 20\}$, and $\mathcal{S}_{15,32} : \{n_t = 15, n_r = 32\}$ are arbitrarily selected, and the posterior PDFs are evaluated for these setups for three different phantom (i.e., the dielectric cylinders, whose centers are randomly positioned, herein called the "Phantom", Fig. 2a). The posterior distributions are plotted in Fig. 2c, wherein going from top to bottom (i.e., along the rows) the posteriors correspond to three different phantoms $(\mathcal{P}_{t1}, \mathcal{P}_{t2}, \mathcal{P}_{t3})$ for a given measurement setup, and going from left to right (i.e., along the columns) the posteriors correspond to three different measurement setups $(\mathcal{S}_{6,8}, \mathcal{S}_{12,20}, \mathcal{S}_{15,32})$ for a given phantom. The results in Fig. 2c conclude that the uncertainty in the posteriors reduces

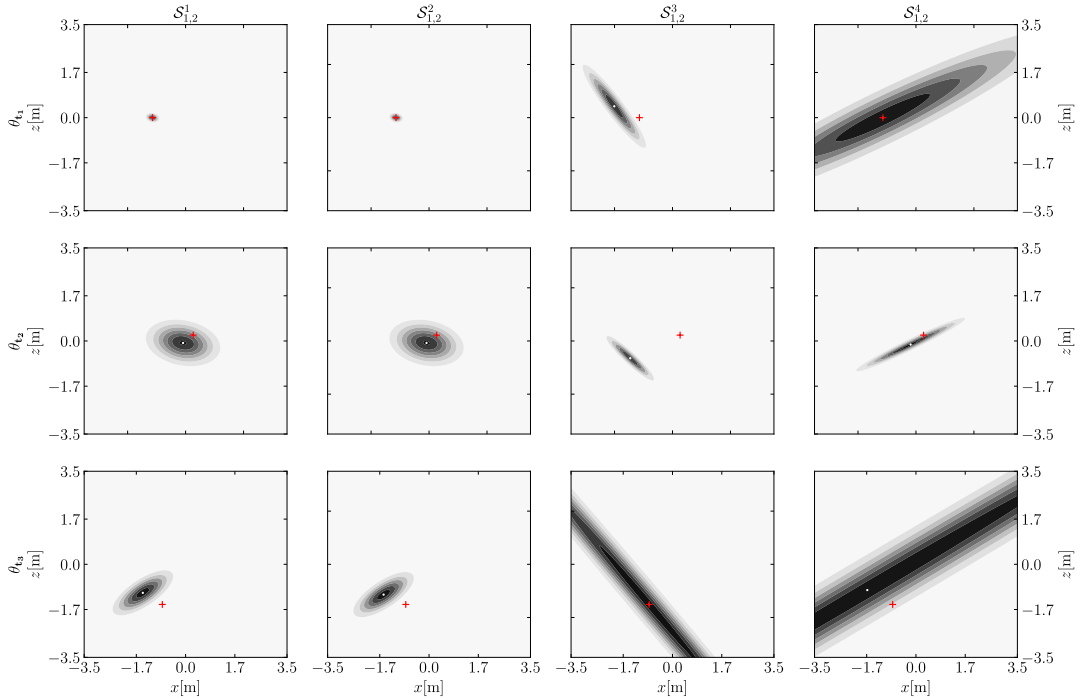


FIGURE 4. Posterior distributions for four different phantom locations (row wise) and for four different sensor setups (column wise).

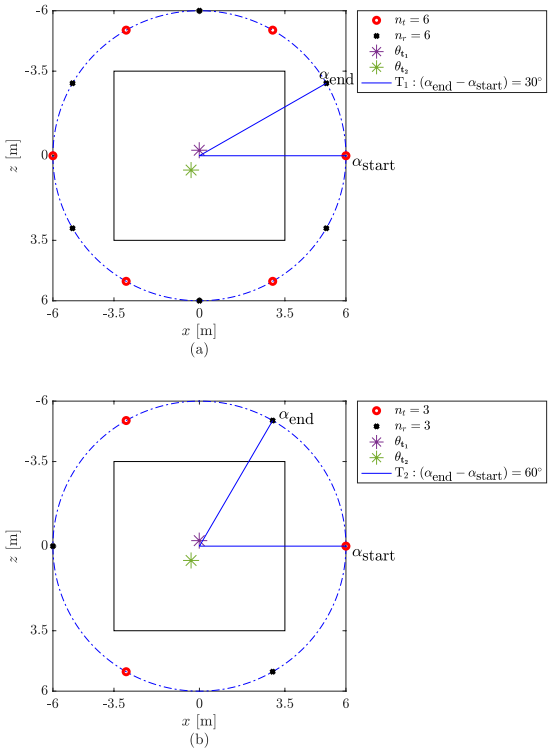


FIGURE 5. Problem setup: the positions of the transmitters are fixed, while the receivers are simultaneously rotated in incremental angular steps (herein called α_{sector}). (a) $n_t = n_r = 6$, ($\alpha_{\mathrm{sector}} = 30°$), (b) $n_t = n_r = 3$, ($\alpha_{\mathrm{sector}} = 60°$).

for an increase in the number of transmitters and receivers in a given measurement setup (e.g., see the posteriors in the third column, showing reduced variance, for all three scenarios). In fact determining the optimal number of sensors to use in an experiment is closely tied to the cost or budget of deploying them. Using $n + 1$ sensors with the optimal configuration always leads to better performance than using only n sensors in their optimal configuration, as more data results in a tighter posterior distribution and less uncertainty in the reconstruction. However, it is also important to note that there is a threshold after which adding more sensors will not significantly improve the reconstruction accuracy. This is because the EIG's surface becomes saturated beyond this point. Such an effect is observed around the top-right corner in Fig. 2b suggesting that the corresponding number of sensors can be deployed given the allocated budget is large enough.

TABLE 1. CRPS scores for the marginal posterior distributions across the three different setups of sensors in case T_1.

	x for $\mathcal{P}_1$	z for $\mathcal{P}_1$	x for $\mathcal{P}_2$	z for $\mathcal{P}_2$
$\mathcal{S}^1_{6,6}$	6.947×10^{-4}	7.917×10^{-4}	1.291×10^{-3}	1.288×10^{-3}
$\mathcal{S}^2_{6,6}$	6.943×10^{-4}	7.915×10^{-4}	1.287×10^{-3}	1.271×10^{-3}
$\mathcal{S}^3_{6,6}$	6.942×10^{-4}	7.913×10^{-4}	1.286×10^{-3}	1.266×10^{-3}

TABLE 2. CRPS scores for the marginal posterior distributions across the three different setups of sensors in case T_2.

	x for $\mathcal{P}_1$	z for $\mathcal{P}_1$	x for $\mathcal{P}_2$	z for $\mathcal{P}_2$
$\mathcal{S}^1_{3,3}$	1.607×10^{-3}	1.588×10^{-3}	2.826×10^{-3}	2.929×10^{-3}
$\mathcal{S}^2_{3,3}$	1.754×10^{-3}	1.858×10^{-3}	2.627×10^{-3}	2.645×10^{-3}
$\mathcal{S}^3_{3,3}$	1.777×10^{-3}	2.007×10^{-3}	2.582×10^{-3}	2.599×10^{-3}

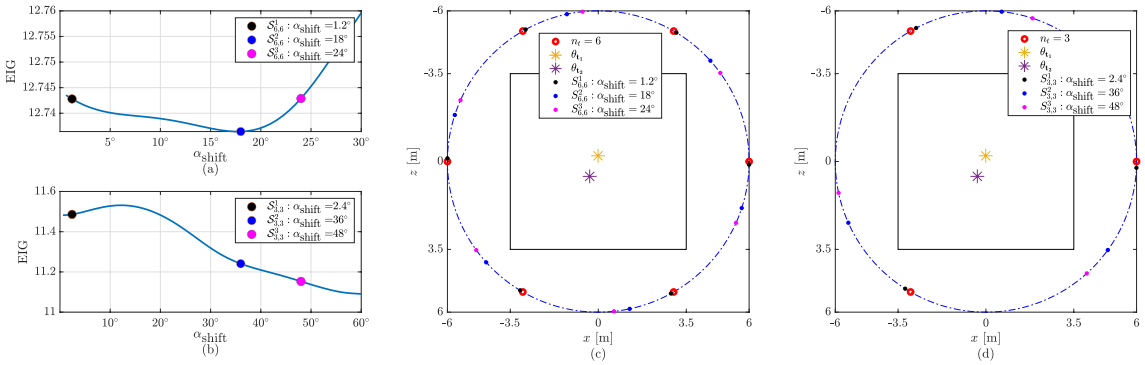


FIGURE 6. Problem setup: (a)&(c) the EIG as a function of α_{shift} for setups in Fig. 5(a) and Fig. 5(b) respectively, (b)&(d) sensor arrangements chosen for setups in Fig. 5(a) and Fig. 5(b) respectively, to analyze their effect on the accuracy of inversion.

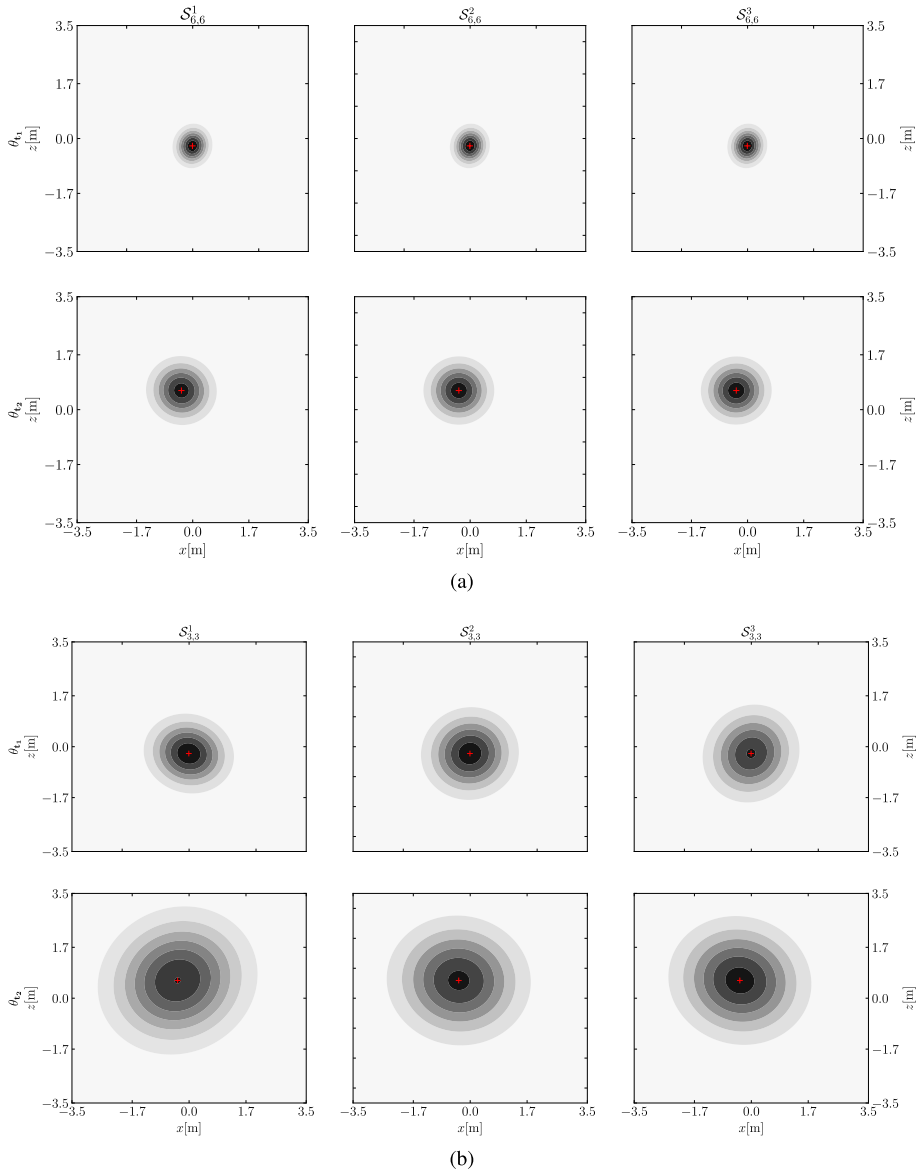


FIGURE 7. Posterior distributions for two phantom locations (row wise) and for three different sensor setups (column wise), (a) for setup in Fig. 6(a), and (b) for setup in Fig. 6(c).

B. TEST 2: LIMITING THE NUMBER OF SENSORS

In this test, a transmitter is fixed, and two receivers are independently rotated with respect to the fixed transmitter, in a circular path surrounding D at a range of angular positions, Fig. 3(a). The EIG in Fig. 3(b) is plotted as a function of the angular positions (labeled along the horizontal and

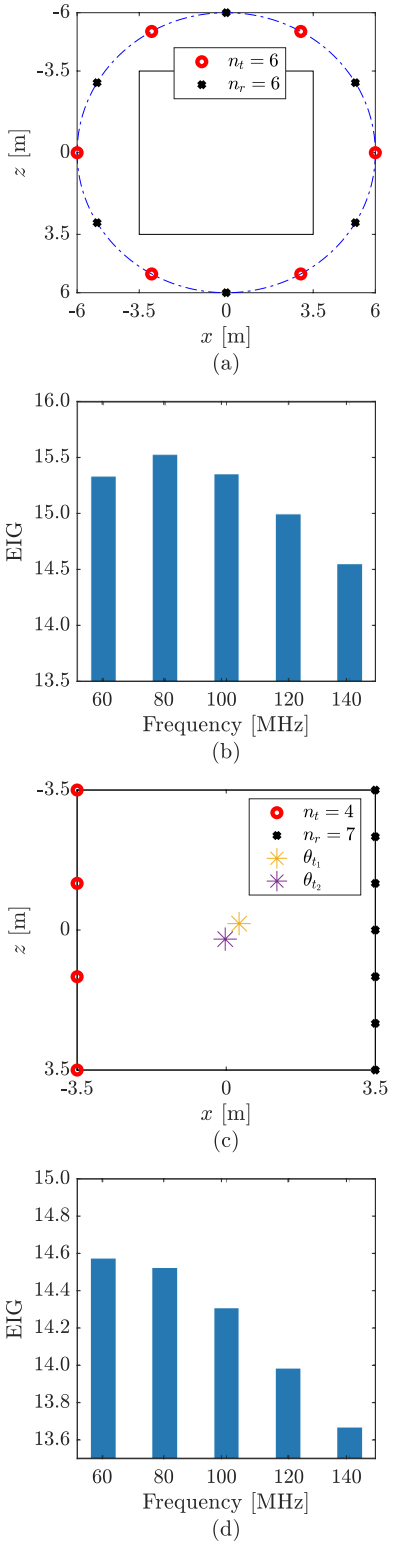


FIGURE 8. problem setup (a) sensors places in a circle surrounding *D*, (b) EIG for a circle setting, (c) sensors in a bore-hole configuration. Measurements are cascaded by switching the bore-holes, (d) EIG for the borehole configuration.

vertical axes) of both the receivers. The EIG is observed maximum whenever both the receivers are positioned at an equal and opposite angular distance with respect to the transmitter (shown by the dark blue pixels along the diagonal), whereas when the two receivers are located at the same position, it does not add new/independent information, instead the predicted scattered electric field will have repeated values, that is, the posterior covariance matrix will become rank deficient which means that its determinant becomes zero. As a result the EIG will be negative, as shown by the bright white pixels along the diagonal in the EIG plot. The blue pixelated corner positions on the EIG plot represent receivers placed at an angular offset relative to the transmitter (placed in between them). These positions could be important for collecting back-scattered information. The angular offset is not so small as to cause ill-conditioning or redundant measurements. These three observations turned out to be the conclusive findings from this experiment, and are summarized in Fig. 3(c).

To study the effect of the position of the two receivers, on the accuracy of inversion, four different setups, $\mathcal{S}^1_{1,2}$: $(\alpha_1 = 112.5°, \alpha_2 = 247.5°)$, $\mathcal{S}^2_{1,2}$: $(\alpha_1 = 247.5°, \alpha_2 = 112.5°)$, $\mathcal{S}^3_{1,2}$: $(\alpha_1 = 236.25°, \alpha_2 = 191.25°)$, and, $\mathcal{S}^4_{1,2}$: $(\alpha_1 = 112.5°, \alpha_2 = 67.5°)$, are arbitrarily selected, Fig. 3(d), and the posterior probability density functions are evaluated for these setups for three different phantoms

The posterior distributions are plotted in Fig. 4, wherein going from top to bottom (i.e., along the rows) the posteriors correspond to three different phantoms for a given measurement setup, and going from left to right (i.e., along the columns) the posteriors correspond to four different measurement setups $(\mathcal{S}^1_{1,2}, \mathcal{S}^2_{1,2}, \mathcal{S}^3_{1,2}, \mathcal{S}^4_{1,2})$ for a given phantom.

It can be noted that $\mathcal{S}^1_{1,2}$ and $\mathcal{S}^2_{1,2}$ correspond to the scenario where the receiver positions are interchanged, therefore the posterior distributions in the first two columns behaves identically. An interesting outcome of the analysis in Fig. 4 is that the inverse solution has less uncertainty when the receivers are placed on opposite sides (not necessarily at equal distances) relative to the transmitter. This is consistent with the analysis of the EIG in Fig. 3(b). The posterior distributions corresponding to $\mathcal{S}^3_{1,2}$ and $\mathcal{S}^4_{1,2}$ shows relatively higher levels of uncertainty, since both the receivers are positioned along the same arc, on either side, with respect to the transmitter. It is observed in Fig. 4 that the predictions for the phantom $\boldsymbol{\theta}_{t_3}$ are far off. The CRPS scores, which measures how each marginal of the posterior distribution approximates the abscissa and the ordinate of the phantom center using the data collected across the setups, are evaluated, for the corresponding posteriors, to be (0.663, 0.363), (0.659, 0.363), (0.798, 0.377) and (0.763, 0.388), for all four measurement setups respectively, which indicated that setups $\mathcal{S}^1_{1,2}$ and $\mathcal{S}^2_{1,2}$ have similar scores and moreover better than the other two setups.

C. TEST 3: ANGULARLY OFFSETTING RECEIVERS RELATIVE TO TRANSMITTERS

In this test, equal number of transmitters and receivers, (i.e., $n_t = n_r$), that are equally spaced on a circle surrounding D,

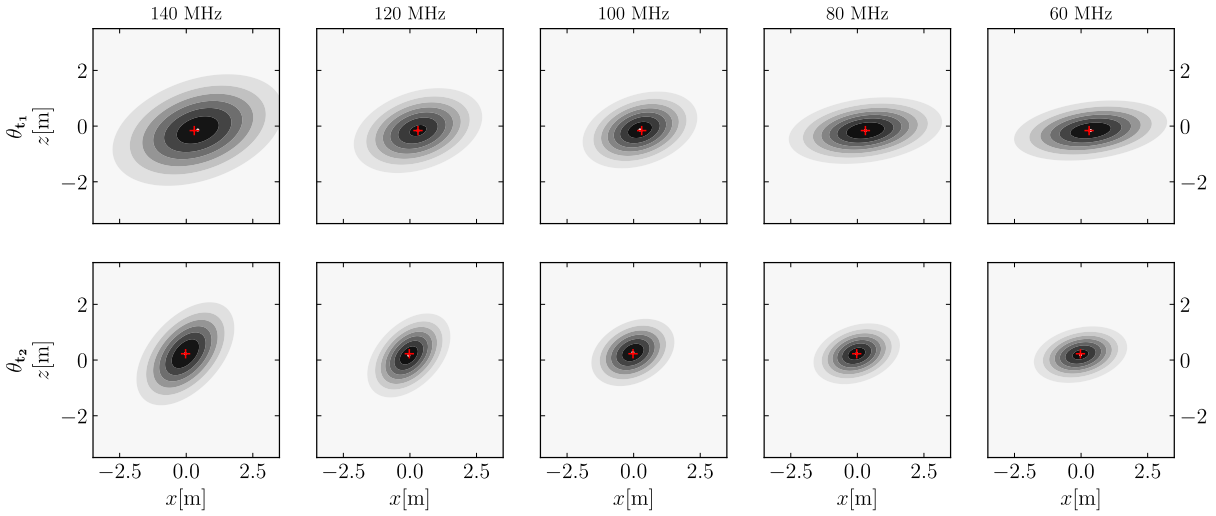


FIGURE 9. Posterior distributions for two phantom locations (row wise) and for five different excitation frequencies (column wise) for the borehole setup.

are considered. The EIG is analyzed as a function of simultaneous rotation of all the receivers, in incremental angular steps (herein called α_{shift}), in a predefined sector around the transmitters (having fixed positions). The transmitter receiver arrangement corresponding to the maximum angular displacement for the two different test cases, ($T_1 : \alpha_{\text{sector}} = 30°$) and ($T_2 : \alpha_{\text{sector}} = 60°$) is shown in Fig. 5(a) and Fig. 5(b). Fig. 6(a) and Fig. 6(c) represent the variation of the EIG, respectively, in case T_1 and case T_2. In both cases, small variation in the magnitude is noticed, although the magnitude of the EIG in case T_1 ($n_t = n_t = 6$) is higher. This is aligned with the findings in Section III-A: more observational data, reduces the uncertainty in the inversion. To study the effect of the position of the receivers, on the accuracy of inversion, three different setups, for each test case. For the setting $n_t = 6$ ($T_1 \rightarrow \mathcal{S}^1_{6,6} : \alpha_{\text{shift}} = 1.2°$, $\mathcal{S}^2_{6,6} : \alpha_{\text{shift}} = 18°$, and, $\mathcal{S}^3_{6,6} : \alpha_{\text{shift}} = 24°$) and for the setting $n_t = 3$ ($T_2 \rightarrow \mathcal{S}^1_{3,3} : \alpha_{\text{shift}} = 2.4°$, $\mathcal{S}^2_{3,3} : \alpha_{\text{shift}} = 36°$, and, $\mathcal{S}^3_{3,3} : \alpha_{\text{shift}} = 48°$), are arbitrarily selected. The said arrangement is shown in Fig. 6(b) and Fig. 6(d). The posterior distribution for both setups is plotted in Fig. 7. A comparison of the posteriors also shows that adding more observational data, reduces the uncertainty in the inversion. However for a better understanding of what is the difference in the quality of posterior for the three chosen setups, the CRPS score is calculated. No significant change is observed. The results of this numerical experiment indicate that the quality of inversion is not significantly impacted when sensors are regularly placed between a set of transmitters.

D. TEST 4: SENSOR'S OPERATING FREQUENCY

In this test, the operating frequency of the illumination is varied from 60 MHz to 140 MHz in steps of 20 MHz, while the physical dimensions of the investigation domain as well as the discretization size are kept fixed. Moreover a bore-hole like measurement setup, Fig. 8(c) is also considered in addition to the circular one, Fig. 8(a). The positions of the transmitters and receivers are interchanged for the bore-hole setup, to complete the measurement set.

The EIG profiles for both setups in Fig. 8(b) and 8(d), have shown a decrease with the increase in frequency and vice versa. Note that the EIG for both the setups are individual entities and does not present any correlated information and the variation of the magnitude of the EIG is more noticeable and is monotone for the borehole configuration. The EIG can be related to the ability of the imaging system to distinguish between different features or structures within the object being imaged and this ability is often related to the spatial resolution of the imaging system, which is generally higher at higher frequencies. However, other factors, such as the penetration depth and scattering properties of the object, can also affect the information gain. A higher EIG for low frequencies, suggests that these other factors may be more important for the specific object being imaged than the spatial resolution.

To study the effect of choosing a particular operating frequency, given a bore hole measurement setup, on the accuracy of inversion, the posterior PDFs are evaluated for two different phantoms, Fig.9. The posterior distributions are horizontally elongated representing relatively a higher level of uncertainty in resolving target of interest in the top and the bottom regions. This is expected, because, for a borehole setup, there were no measurements from the top and the bottom directions. This experiment concluded with that there is no significant impact on the quality of measurements, given a particular operating frequency as long as it satisfies the requirements for a certain penetration depth and spatial resolution.

IV. CONCLUSION

In conclusion, this study investigated the sensor placement in electromagnetic imaging experiments using Bayesian

experimental design. The results showed that an increase in the number of transmitters and receivers reduces uncertainty in the posteriors, while given the sensors are arranged in a circle around the investigation domain, the EIG is maximized when the receivers are positioned at equal and opposite angular distances with respect to the transmitter. The quality of inversion is not significantly impacted when sensors are regularly placed between a set of transmitters, and the operating frequency does not significantly impact the quality of measurements as long as it satisfies the requirements for a certain penetration depth and spatial resolution. This study provides a preliminary investigation into three-dimensional geophysical inversion problems, with ongoing efforts underway to extend and build upon the foundation established by this work.

We have investigated some standard situations but the methodology used is very general and flexible and can be applied to a large variety of practical sensor configurations in the characterization (2D and 3D) or the monitoring (time-lapse) phase of any subsurface geophysical exploration phase. Moreover, this methodology can be applied at any scale, from near-surface (environmental, engineering, or agriculture geophysics) to deep (geothermal, mining, and Oil and Gas) geophysical applications. The decision for proper sensor placement and density (number of sensors used) is critical for acquiring accurate data collection with the highest resolution, affecting the final interpreted subsurface model.

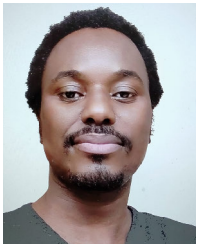
REFERENCES

[1] M. Pastorino, *Microwave Imaging*. Hoboken, NJ, USA: Wiley, 2010.
[2] A. Randazzo and M. Pastorino, *Microwave Imaging Methods and Applications*. Norwood, MA, USA: Artech House, 2018.
[3] A. I. Sandhu, S. A. Shaukat, A. Desmal, and H. Bagci, "ANN-assisted CoSaMP algorithm for linear electromagnetic imaging of spatially sparse domains," *IEEE Trans. Antennas Propag.*, vol. 69, no. 9, pp. 6093–6098, Sep. 2021.
[4] A. I. Sandhu, A. Desmal, and H. Bagci, "An accelerated nonlinear contrast source inversion scheme for sparse electromagnetic imaging," *IEEE Access*, vol. 9, pp. 54811–54819, 2021.
[5] A. K. Pattem, A. Illa, A. Afshan, and P. K. Ghosh, "Optimal sensor placement in electromagnetic articulography recording for speech production study," *Comput. Speech Lang.*, vol. 47, pp. 157–174, Jan. 2018.
[6] N. Irishina, M. Moscoso, and O. Dorn, "Microwave imaging for early breast cancer detection using a shape-based strategy," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1143–1153, Apr. 2009.
[7] O. Dorn and A. Hiles, "A level set method for magnetic induction tomography of 3D boxes and containers," in *Electromagnetic Non-Destructive Evaluation (XXI)*. Amsterdam, The Netherlands: IOS Press, 2018, pp. 33–40.
[8] G. Incorvaia and O. Dorn, "Stochastic optimization methods for parametric level set reconstructions in 2D through-the-wall radar imaging," *Electronics*, vol. 9, no. 12, p. 2055, Dec. 2020.
[9] R. Villegas, O. Dorn, M. Moscoso, and M. Kindelan, "Reservoir characterization using stochastic initializations and the level set method," *Comput. Math. Appl.*, vol. 56, no. 3, pp. 697–708, Aug. 2008.
[10] R. R. Hayes, P. A. Newill, F. J. W. Podd, T. A. York, B. D. Grieve, and O. Dorn, "An investigation into the use of a mixture model for simulating the electrical properties of soil with varying effective saturation levels for sub-soil imaging using ECT," *J. Phys., Conf.*, vol. 255, Nov. 2010, Art. no. 012002.
[11] G. Panagopoulos, P. Soupios, A. Vafidis, and E. Manoutsoglou, "Integrated use of well and geophysical data for constructing 3D geological models in shallow aquifers: A case study at the Tymbakion basin, crete, Greece," *Environ. Earth Sci.*, vol. 80, no. 4, pp. 1–17, Feb. 2021.
[12] W. Hu, A. Abubakar, and T. M. Habashy, "Joint electromagnetic and seismic inversion using structural constraints," *Geophysics*, vol. 74, no. 6, pp. R99–R109, Nov. 2009.
[13] A. Abubakar, G. Gao, T. M. Habashy, and J. Liu, "Joint inversion approaches for geophysical electromagnetic and elastic full-waveform data," *Inverse Problems*, vol. 28, no. 5, May 2012, Art. no. 055016.
[14] Y. Zhang, Z. Zhao, Z. Nie, and Q. H. Liu, "Approach on joint inversion of electromagnetic and acoustic data based on structural constraints," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 7672–7681, Nov. 2020.
[15] P. Edigbue, I. Akca, I. Demirci, A. Al-Shuhail, H. A. Hamdan, P. Kirmizakis, E. Candansayar, S. Hanafy, and P. Soupios, "Testing different stochastic methods to model direct current resistivity and seismic refraction geophysical data using a combined local and global optimization algorithm," *Arabian J. Sci. Eng.*, vol. 48, pp. 7925–7938, Feb. 2023.
[16] P. Edigbue, I. Demirci, H. Hamdan, A. Al-Shuhail, I. Akca, P. Soupios, and E. Candansayar, "Joint inversion of seismic and DC geophysical data using local and global optimization algorithms," in *Proc. Int. Petroleum Technol. Conf.*, Feb. 2022.
[17] İ. Demirci, M. E. Candansayar, A. Vafidis, and P. Soupios, "Two dimensional joint inversion of direct current resistivity, radio-magnetotelluric and seismic refraction data: An application from Bafra Plain, Turkey," *J. Appl. Geophysics*, vol. 139, pp. 316–330, Apr. 2017.
[18] W. Liao, R. Peng, X. Hu, W. Zhou, and G. Huang, "3-D joint inversion of MT and CSEM data for imaging a high-temperature geothermal system in Yanggao region, Shanxi province, China," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5925813.
[19] G. Panagopoulos, A. Vafidis, P. Soupios, and E. Manoutsoglou, "A study on the gas-bearing miocene sediments of MESSARA basin in Crete (Greece) by using seismic reflection, geochemical and petrophysical data," *Arabian J. Sci. Eng.*, vol. 47, no. 6, pp. 7449–7465, Jun. 2022.
[20] Y. Martinez, A. Ashadi, H. Hinojosa, P. Soupios, and K. Strack, "New high-power controlled source electromagnetic system for geothermal applications," in *Proc. Geothermal Rise Conf.*, Reno, NV, USA, 2022, pp. 28–31.
[21] K. M. Strack, C. Barajas-Olalde, S. Davydycheva, Y. Martinez, and P. Soupios, "Surface-to-borehole electromagnetics using an array system: A case study for Co_2 monitoring and the energy transition," in *Proc. SPE Annu. Tech. Conf. Exhib.*, Sep. 2022.
[22] U. Autio, M. Y. Smirnov, A. Savvaidis, P. Soupios, and M. Bastani, "Combining electromagnetic measurements in the mygdonian sedimentary basin, Greece," *J. Appl. Geophys.*, vol. 135, pp. 261–269, Dec. 2016.
[23] A. M. Stuart, "Inverse problems: A Bayesian perspective," *Acta Numerica*, vol. 19, pp. 451–559, May 2010.
[24] K. Chaloner and I. Verdinelli, "Bayesian experimental design: A review," *Stat. Sci.*, vol. 10, no. 3, pp. 273–304, Aug. 1995.
[25] K. Bernhard and J. Vygen, *Combinatorial Optimization: Theory and Algorithms*, 3rd ed. Cham, Switzerland: Springer, 2008.
[26] J. P. Kaipio, A. Seppänen, E. Somersalo, and H. Haario, "Posterior covariance related optimal current patterns in electrical impedance tomography," *Inverse Problems*, vol. 20, no. 3, pp. 919–936, Apr. 2004.
[27] J. P. Kaipio, A. Seppänen, A. Voutilainen, and H. Haario, "Optimal current patterns in dynamical electrical impedance tomography imaging," *Inverse Problems*, vol. 23, no. 3, pp. 1201–1214, Jun. 2007.
[28] X. Huan and Y. M. Marzouk, "Simulation-based optimal Bayesian experimental design for nonlinear systems," *J. Comput. Phys.*, vol. 232, no. 1, pp. 288–317, Jan. 2013.
[29] A. G. Carlon, B. M. Dia, L. Espath, R. H. Lopez, and R. Tempone, "Nesterov-aided stochastic gradient methods using Laplace approximation for Bayesian design optimization," *Comput. Methods Appl. Mech. Eng.*, vol. 363, May 2020, Art. no. 112909.

[30] K. J. Ryan, "Estimating expected information gains for experimental designs with application to the random fatigue-limit model," *J. Comput. Graph. Statist.*, vol. 12, no. 3, pp. 585–603, Sep. 2003.
[31] E. Ryan, C. Drovandi, and A. Pettitt, "Fully Bayesian experimental design for pharmacokinetic studies," *Entropy*, vol. 17, no. 3, pp. 1063–1089, Mar. 2015.
[32] E. G. Ryan, C. C. Drovandi, J. M. McGree, and A. N. Pettitt, "A review of modern computational algorithms for Bayesian optimal design," *Int. Stat. Rev.*, vol. 84, no. 1, pp. 128–154, Apr. 2016.
[33] J. Beck, B. M. Dia, L. F. R. Espath, Q. Long, and R. Tempone, "Fast Bayesian experimental design: Laplace-based importance sampling for the expected information gain," *Comput. Methods Appl. Mech. Eng.*, vol. 334, pp. 523–553, Jun. 2018.
[34] J. Beck, B. Mansour Dia, L. Espath, and R. Tempone, "Multilevel double loop Monte Carlo and stochastic collocation methods with importance sampling for Bayesian optimal experimental design," *Int. J. Numer. Methods Eng.*, vol. 121, no. 15, pp. 3482–3503, Aug. 2020.
[35] P. Tsilifis, R. G. Ghanem, and P. Hajali, "Efficient Bayesian experimentation using an expected information gain lower bound," *SIAM/ASA J. Uncertainty Quantification*, vol. 5, no. 1, pp. 30–62, Jan. 2017.
[36] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, "A fast and scalable method for a-optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems," *SIAM J. Sci. Comput.*, vol. 38, no. 1, pp. A243–A272, Jan. 2016.
[37] M. Seeger, H. Nickisch, R. Pohmann, and B. Schölkopf, "Optimization of K-space trajectories for compressed sensing by Bayesian experimental design," *Magn. Reson. Med.*, vol. 63, no. 1, pp. 116–126, Jan. 2010.
[38] J. Liepe, S. Filippi, M. Komorowski, and M. P. H. Stumpf, "Maximizing the information content of experiments in systems biology," *PLOS Comput. Biol.*, vol. 9, no. 1, pp. 1–13, Jan. 2013, doi: 10.1371/journal.pcbi.1002888.
[39] A. Bartuska, L. Espath, and R. Tempone, "Small-noise approximation for Bayesian optimal experimental design with nuisance uncertainty," *Comput. Methods Appl. Mech. Eng.*, vol. 399, Sep. 2022, Art. no. 115320. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782522004194
[40] E. Cherkaeva and A. C. Tripp, "Source optimization in the inverse geoelectrical problem," in *Inverse Problems in Geophysical Applications*, H. W. Engl, A. Louis, and W. Rundell, Eds. Philadelphia, PA, USA: SIAM, 1997, pp. 240–256.
[41] Q. Long, M. Motamed, and R. Tempone, "Fast Bayesian optimal experimental design for seismic source inversion," *Comput. Methods Appl. Mech. Eng.*, vol. 291, pp. 123–145, Jul. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045782515001310
[42] B. M. Dia, "Bayesian experimental design for CO_2 sequestration with leakage risk assessment," *Eng. Computations*, vol. 38, no. 3, pp. 1385–1401, Jun. 2021.
[43] R. Thibaut, N. Compaire, N. Lesparre, M. Ramgraber, E. Laloy, and T. Hermans, "Comparing well and geophysical data for temperature monitoring within a Bayesian experimental design framework," *Water Resour. Res.*, vol. 58, no. 11, Nov. 2022, Art. no. e2022WR033045.
[44] K. Sauer, J. Sachs, and C. Klifa, "Bayesian estimation of 3-D objects from few radiographs," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 5, pp. 1780–1790, Oct. 1994.
[45] X. Liao and L. Carin, "Application of the theory of optimal experiments to adaptive electromagnetic-induction sensing of buried targets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 961–972, Aug. 2004.
[46] S. Kullback, *Information Theory and Statistics*. Hoboken, NJ, USA: Wiley, 1959.
[47] J. Ginebra, "On the measure of the information in a statistical experiment," *Bayesian Anal.*, vol. 2, no. 1, pp. 167–211, Mar. 2007.
[48] O. Dorn, H. Bertete-Aguirre, J. G. Berryman, and G. C. Papanicolaou, "Sensitivity analysis of a nonlinear inversion method for 3D electromagnetic imaging in anisotropic media," *Inverse Problems*, vol. 18, no. 2, pp. 285–317, Apr. 2002.
[49] T. Helin, N. Hyvönen, and J.-P. Puska, "Edge-promoting adaptive Bayesian experimental design for X-ray imaging," *SIAM J. Scientific Comput.*, vol. 44, no. 3, pp. B506–B530, Jun. 2022.
[50] Q. Long, M. Scavino, R. Tempone, and S. Wang, "Fast estimation of expected information gains for Bayesian experimental designs based on Laplace approximations," *Comput. Methods Appl. Mech. Eng.*, vol. 259, pp. 24–39, Jun. 2013.
[51] J. Richmond, "Scattering by a dielectric cylinder of arbitrary cross section shape," *IEEE Trans. Antennas Propag.*, vol. AP-13, no. 3, pp. 334–341, May 1965.
[52] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.
[53] E. Miller, "Electromagnetics activities at Lawrence Livermore National Laboratory," *IEEE Antennas Propag. Soc. Newslett.*, vol. APSN-23, no. 6, pp. 4–7, Dec. 1981.

ALI IMRAN SANDHU (Member, IEEE) received the B.S. degree in electronics engineering from COMSATS University Islamabad (CUI), Lahore, Pakistan, in 2007, the M.S. degree in communication engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 2010, and the Ph.D. degree in electrical engineering from the Division of Computer, Electrical, and Mathematical Sciences and Engineering, King Abdullah University of Science and Technology, (KAUST), Thuwal, Saudi Arabia, in 2020.

From 2011 to 2013, he was a Research Engineer with the Microwave Laboratory, University of Calabria, Rende, Italy. From August 2007 to September 2021, he was a Lecturer with the Electrical and Computer Engineering Department, CUI, where he taught at undergraduate and graduate levels. He is currently a Postdoctoral Fellow with the Center of Integrative Petroleum Research, King Fahd University of Petroleum and Minerals (KFUPM). His research interests include applied computational electromagnetics with an emphasis on the characterization of electromagnetic fields and wave interactions on complex geometries, and solutions for 2-D and 3-D joint EM and seismic inverse problems incorporating signal processing and machine learning techniques. He was the Finalist in the Student Poster Competition at the IEEE IST Conference, in 2016, and secured the Best Student Paper Nomination at the IEEE Applied Computational Electromagnetics Society Conference, in 2017. He received two Bronze medals for securing distinction in his undergraduate discipline at the campus as well as at institute levels at CUI.



BEN MANSOUR DIA received the M.Sc. and master's degrees in applied mathematics from Gaston Berger University, and the Ph.D. degree from Gaston Berger University and the Royal Institute of Technology, Stockholm, under the International Science Program (ISP) Mathematical Program.

In October 2017, he joined the King Fahd University of Petroleum and Minerals (KFUPM) after completing his postdoctoral fellowship with Prof. Raul Tempone at the King Abdullah University of Science and Technology (KAUST) on uncertainty quantification. He is currently a Research Assistant Professor of applied mathematics with the Center for Integrative Petroleum Research, College of Petroleum Engineering and Geosciences, KFUPM. His focus is on PDE-based science with high public value challenged by computer model uncertainty. That is formalizing imprecise mathematical models as high-dimensional probabilistic models, and developing numerical methods for solving them with reliable uncertainty quantification via inverse problems. Typically, he investigates computational statistics methods, which are accurate, reliable, and cost-efficient, as well as suitable for addressing complex and large-scale physical problems governed by PDEs and prone to high uncertainty from multiple sources. His field of applications covers imaging in geophysical problems, environmental management, and multiphase flow in porous media with a specific focus on CO_2 sequestration.

OLIVER DORN (Senior Member, IEEE) is an internationally recognized authority on the application of level set and shape-based methods to inverse problems. During his career, he has been working at various prestigious research institutions in the USA, Canada, Spain, France, Germany, and the U.K. on many important inverse problems, including geophysical tomography, reservoir characterization, biomedical imaging, remote sensing, and non-destructive evaluation of materials. He has been collaborating with several international and national companies, such as Shell, Repsol/YPF, Schlumberger, Rohde und Schwarz, ITS, AWE, DSTL, and LLNL, amongst others. He spent several extended invited research stays at international research centers, such as MGSS (Stanford University), MSRI (University of California at Berkeley), IPAM (University of California at Los Angeles), IMA (University of Minneapolis), BIRS (Banff), and CenSSIS (Boston University). His current research projects include large-scale electromagnetic inverse problems, through-the-wall radar imaging, history matching and data assimilation, seismic full waveform inversion, and AI in imaging. He has been participating in a large variety of nationally and internationally funded research projects and consortia in the USA, Canada, Spain, France, and the U.K. He has presented over 150 mostly invited conference presentations worldwide and has published over 80 peer-reviewed journals and conference papers, book chapters, and review papers. He has edited various special issues dedicated to electromagnetic inverse problems in the journal *Inverse Problems*, including a recent one, in 2023. Three of his highly cited publications have been selected by the journal *Inverse Problems* for its "Highlights Collections." He has supervised seven Ph.D. projects and over 20 master's research projects on a variety of relevant practical topics. He is a Platinum Member of the EAGE. He is also a long-standing Editorial Board Member of the journal *Inverse Problems*.

PANTELIS SOUPIOS received the B.S. degree in geology, the M.S. degree in applied geophysics, and the Ph.D. degree in applied geophysics from the Aristotle University of Thessaloniki, Greece, and Uppsala University, Sweden, in 2000. He started his carrier with the Laboratory of Geophysics and Seismology, Aristotle University of Thessaloniki. His expertise is in the application of several geophysical methods for environmental, hydrogeological, and engineering applications, and geohazard studies. From 2002 to 2018, he was a Faculty Member with the Department of Environmental and Natural Resources Engineering, Technological Educational Institute of Crete, Greece. He has established a strong international research network for collaboration. He is currently a Professor of geophysics with the Geosciences Department, King Fahd University of Petroleum and Minerals (KFUPM), and a Program Leader in geophysics with the Center of Integrative Petroleum Research (CIPR), KFUPM. Since joining KFUPM, in 2018, he has been a PI or a Co-PI in several externally funded projects carried out by the Department of Geosciences and the Center of Integrated Petroleum Research, CPG, KFUPM. He has taught undergraduate and graduate courses with the Department of Geosciences. He has published more than 100 technical articles and many reports on various areas of engineering and environmental geophysics. He has attended several conferences related to his area. His papers are published in prestigious international journals (115 documents in SCOPUS (March 2023); 1782 citations and H-index: 23 in SCOPUS; 3138 citations and H-index: 27 in Google Scholar).