

Research Paper

Investigation of wavelength-induced uncertainties in full-wave radar tomography of high contrast domain: An application to small solar system bodies

Yusuf Oluwatoki Yusuf^{a,b,*}, Astrid Dufaure^b, Liisa-Ida Sorsa^a, Christelle Eyraud^b, Sampsa Pursiainen^a

^a Computing Sciences, Tampere University (TAU), PO Box 692, 33101, Tampere, Finland

^b Aix-Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, Marseille, France



ARTICLE INFO

Keywords:

Asteroids
Compositions
Interiors
Radar observations
Image processing

ABSTRACT

This paper aims to reconstruct the internal structure of a two-dimensional test object via numerically simulated full-wave time domain radar tomography with the presence of wavelength-induced (WI) uncertainties, following from a complex domain structure, and domain diameters 21 or 64 times the wavelength of the signal propagating inside the target. In particular, we consider an application in planetary scientific studies of reconstructing the interior structure of an arbitrary high contrast small Solar System Body (SSSB), i.e., an asteroid, with a probing signal wavelength limited by the instrument and mission payload requirements. Our uncertainty reduction model finds the reconstruction via averaging multiple inverse solutions assuming that the WI deviations in the solutions correspond to random deviations, which we assume to be independent and identically distributed (IID). It incorporates error marginalisation via a randomised signal configuration, spatial-averaging of candidate solutions, frequency-based error marginalisation, and the truncated singular value decomposition (TSVD) filtering technique, based on our assumptions of the phase discrepancy of the signal, domain parameters, and the full-wave forward model. The numerical experiments are performed for 20 and 60 MHz centre frequencies proposed for CubeSat-based radars, the latter being the centre frequency of the Juvetas Radar which will be aboard Hera mission to investigate the interior structure of asteroid Dimorphos. A benchmark reconstruction of the target was obtained with the spatial averaging, sparse point density and frequency randomised configuration for both 20 and 60 MHz frequency systems.

1. Introduction

Radar tomography (RT) has emerged as a powerful technique for obtaining high resolution images of complex target domains in recent years. With the application of tomographic radar imaging spanning across several fields of research, such as biomedical imaging, geoscience, and engineering; sophisticated and computationally efficient inversion techniques to characterise the interior properties of such a target domain from the scattered field of the probing radar signal have also been developed extensively (Chew and Wang, 1990; Carlsten et al., 1995; Semenov et al., 2005; Ernst et al., 2007; Pursiainen and Kaasalainen, 2016; Wiskin et al., 2020). The inverse scattering problem of full-wave tomography, however, is an ill-posed inverse problem that can result into a less accurate reconstruction of the target domain as the signal frequency and, thereby, the measurement and modelling uncertainties increase. In particular, the inversion is very sensitive to

the carrier, hence we propose a method to help mitigate this sensitivity. This paper aims at reconstructing the interior structure of a complex domain, e.g., an arbitrary high contrast Small Solar System Body (SSSB), via full-wave time domain RT, and specifically, investigating the wavelength-induced uncertainties, which follow as a consequence of complex multipath signal propagation inside the target due to high contrast details in the permittivity distribution.

The first attempt to use radio waves transmission to infer the deep interior structure of an SSSB was the COMet Nucleus Sounding Experiment by Radio-wave Transmission (CONSERT) (Kofman et al., 2007, 2015), which was part of the European Space Agency's (ESA) Rosetta mission to comet 67P/Churyumov-Gerasimenko. The emergence of small spacecraft technology as a part of deep space missions has improved the future possibilities to perform RT investigations of SSSBs with a sufficient signal coverage (Bambach et al., 2018; CDF, 2018).

* Corresponding author at: Computing Sciences, Tampere University (TAU), PO Box 692, 33101, Tampere, Finland.
E-mail address: yusuf.yusuf@tuni.fi (Y.O. Yusuf).

<https://doi.org/10.1016/j.icarus.2022.115173>

Received 7 October 2021; Received in revised form 20 May 2022; Accepted 3 July 2022

Available online 9 July 2022

0019-1035/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Such a plan is included in ESA's coming Hera mission; the Juventas CubeSat carried by the Hera probe will perform tomographic radar measurements of Dimorphos, the asteroid moon of 65803 Didymos, with its Juventas Radar (JuRa) (Herique et al., 2020a,b).

In this study, we concentrate on full-wave modelling, an important tool to distinguish how the different part of the domain (e.g., the voids or surface) contribute to the observed scattered signals (Eyraud et al., 2020). This requires that the differences between the modelled and measured field are small enough compared to the measurement inaccuracies. Our focus is suppressing the effects of these inaccuracies in the reconstruction process, which is of utmost importance, e.g., based on the recent numerical study by Deng et al. (2021), where the envelope inversion method was used in suppressing local minima points in the observed signal with a misfit function. The asteroid model in this study has an average real relative permittivity of $\epsilon'_r \approx 4$ which is higher than that of some comets at $\epsilon'_r \approx 1.4$, i.e., comet 67P and Wild 2 (Brownlee et al., 2012; Kofman et al., 2015; Hérique et al., 2018). In JuRa, the centre frequency is 60 MHz, hence, from $\lambda = c_0 \epsilon_r^{-1/2} / f$, the wavelength is 5 m outside and approximately 2.5 m inside the target, where c_0 is the speed of light in vacuum. The diameter of Dimorphos is about 160 m, which is approximately 64 times larger than the wavelength of the signal propagating inside it. We define the wavelength-induced uncertainties as modelling errors in form of phase inaccuracies brought about by the nonuniformity of the permittivity distribution, change in wave velocity and irregular propagation path in the target domain.

In full-wave modelling with the finite element time domain (FETD) method (Pursiainen and Kaasalainen, 2016; Takala et al., 2018b), the WI uncertainties in the modulated signal can result in modelling or measurement errors produced by (1) wavelength-induced uncertainties' factors i.e., nonuniform permittivity distribution, change in wave velocity and wave path length in the target domain, (2) discretisation, (3) wave propagation accuracy, and (4) numerical inaccuracy due to noise in demodulation of the carrier signal. This study concentrates on the errors due to (1), wavelength-induced uncertainties, which are caused by the short wavelength of the signal as compared to the target object diameter, and propagation through a high contrast domain. To marginalise these errors, we propose a Gaussian prior model for the uncertainty of the phase in form of a discrepancy function. This discrepancy gives the maximum limit for the applicable baseband frequency considering the real part of the permittivity that can be reconstructed. The numerical experiments performed on a two-dimensional domain show how the discrepancy-based prior model, combined with total variation regularised-inversion approach, allows for the marginalisation of the wavelength-induced uncertainties via independence sampling. The experiments also show how data obtained for a dense spatial point distribution can be effectively utilised in the process of marginalising wavelength-induced uncertainties, taking into account the Nyquist sampling condition which depends on both the modulated and demodulated wave (Khare and George, 2002; Chaparro and Akan, 2018).

This article is organised into Sections 2–4. Section 2 focuses on the model description, and statistical formulations of spatial and frequency-based error marginalisation process. Furthermore, the total variation inversion technique and the truncated singular value decomposition filtering are discussed in Section 2.5. In Section 2.6, the experimental setup and signal configurations are highlighted, and the corresponding parameters related to the numerical implementation are documented. Section 3 includes the numerical results for the two-dimensional analogue model and the discussion of these results is presented in Section 4.

2. Materials and methods

2.1. Inverse wave propagation problem

We consider transverse electric (TE) wave propagation in which the total electric field u propagating in the horizontal plane is oriented

along the vertical direction. The TE-field satisfies the wave equation

$$\epsilon'_r \frac{\partial^2 u}{\partial t^2} + \sigma \frac{\partial u}{\partial t} - \Delta u = \frac{\partial \mathfrak{J}}{\partial t} \quad (1)$$

in a given spatio-temporal domain $[0, T] \times \Omega$, in which the spatial part includes the scattering target Ω_1 and its near surroundings Ω_2 . Here, ϵ'_r is the real part of the relative permittivity ($\epsilon'_r = \epsilon_r - j\epsilon''_r$), ϵ''_r the imaginary part of the relative permittivity, ϵ_r the relative permittivity, σ the conductivity $\sigma = 2\pi f \epsilon''_r$, f denotes the signal frequency, and $\partial \mathfrak{J} / \partial t$ is a point source term for which $\mathfrak{J} = \mathfrak{J}(t, \vec{p}_0)$ represents the current density of a vertical antenna set at point \vec{p}_0 . The spatial scaling is assumed to be such that the velocity of the wave in vacuum is one ($c_0 = 1$). The time t , position \vec{r} , permittivity ϵ'_r , conductivity σ , and velocity $c = \epsilon_r^{-1/2}$ can all be scaled to SI-units through the expressions $s c_0^{-1} t$, $s \vec{r}$, $\epsilon_0 \epsilon'_r$, $s^{-1} \epsilon_0 c_0 \sigma$, and $c_0 c$, respectively. Here, s is a spatial scaling factor (metres), $c_0 = (\epsilon_0 \mu_0)^{-1/2}$ is the speed of the electromagnetic wave in vacuum, $\epsilon_0 = 8.85 \cdot 10^{-12}$ F/m is the electric permittivity of vacuum, and $\mu_0 = 4\pi \cdot 10^{-7}$ H/m is the magnetic permeability which is assumed to be constant in Ω .

The domain is assumed to be decomposed by a triangular mesh $\mathcal{T} = \{T_1, T_2, \dots, T_M\}$, whose j th triangle \mathcal{T}_j corresponds to a set indicator function χ_j , with $\chi_j(\vec{r}) = 1$ if $\vec{r} \in T_j$ and $\chi_j(\vec{r}) = 0$ otherwise. The mesh \mathcal{T} discretises the real relative permittivity as

$$\epsilon'_r = \bar{\epsilon}'_r + \sum_{j=1}^M s_j \chi_j, \quad (2)$$

where $\bar{\epsilon}'_r$ is a constant background permittivity, s_j is the corresponding coefficient of χ_j , and $\bar{\mathbf{x}} = \sum_{j=1}^M s_j \chi_j$ is the perturbation of the discretised real relative permittivity. Replacing the total electric field u in the wave Eq. (1) which is a smooth function, by its partial derivative with respect to s_j results in

$$\epsilon'_r \frac{\partial^2}{\partial t^2} \left(\frac{\partial u}{\partial s_j} \right) + \sigma \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial s_j} \right) - \Delta \frac{\partial u}{\partial s_j} = \frac{\partial}{\partial t} \left(\frac{\partial h}{\partial s_j} \right), \quad (3)$$

where h on the right-hand side is referred to as the scattering source which follows from Ampere's law, and relates directly to \mathfrak{J} , as $\mathfrak{J} = \frac{\partial h}{\partial s_j}$. Here, h is denoted by

$$h = \epsilon'_r \frac{\partial u}{\partial t} \Big|_{s=0} = \sum_{j=1}^M s_j \chi_j \frac{\partial u}{\partial t} \Big|_{s=0} + \bar{\epsilon}'_r \frac{\partial u}{\partial t} \Big|_{s=0}, \quad (4)$$

hence, its partial derivative with respect to s_j gives $\partial h / \partial s_j = \chi_j \frac{\partial u}{\partial t} \Big|_{s=0}$. This scattering source can be interpreted as the source of the differentiated full-wave as, e.g., in Liu et al. (2018). We obtained the measurement data \mathbf{y} as a difference between a full-wave detailed model and a full-wave constant background model ($u(t, \vec{p}_i) - \bar{u}(t, \vec{p}_i)$). Here, the detailed model can be interpreted as the superposition of the background model and the perturbation $\bar{\mathbf{x}}$. Furthermore, a differential operator J_ℓ is used to obtain a linearised approximation of the detailed model taking into account the coupling of the respective scatterers and multi-path effects with the background. However, it omits the coupling between the scatterers (scatterer-to-scatterer interaction). This approximation is similar to the Extended-Born Approximation (EBA) (Abubakar and Habashy, 2005; Gao and Torres-Verdín, 2006), i.e., it is a model which tries to estimate multiple scattering effects in addition to the direct ones. Unlike the classical Born approximation in which the total field u has been replaced with the incident field (Van Hove, 1954; Chew and Wang, 1990; Sorsa et al., 2020), this linearisation incorporates higher-order or multi-path effects between the scattering obstacle and the target body. It is one of the two common ways to approach full-wave inversion, the second method being to solve a local nonlinear least-squares optimisation problem (Virieux and Operto, 2009).

The wavefield is assumed to consist of two complex quadrature amplitude modulated (QAM) components (Sorsa et al., 2021b). The components of QAM consist of the in-phase and quadrature component,

of which, the latter has a $\pi/2$ phase difference compared to the in-phase signal. Consequently, the amplitude of the signal can be obtained accurately anywhere in the spatio-temporal domain. The reconstruction process utilises baseband frequency data \mathbf{y} obtained via QAM demodulation as a full wave recorded for time steps t_1, t_2, \dots, t_N at spatial points $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_K$. The data corresponds to the difference between the wavefields $u(t, \vec{p}_i)$ and $\tilde{u}(t, \vec{p}_i)$ associated with the actual real relative permittivity distribution ϵ'_r and its background estimate $\tilde{\epsilon}'_r$. Using the differentiated signal, this difference can be estimated as follows:

$$y_j(t) = \mathfrak{D}\mathfrak{M}[u(t, \vec{p}_i) - \tilde{u}(t, \vec{p}_i)] = \mathfrak{D}\mathfrak{M} \left[\sum_{j=1}^M s_j \frac{\partial}{\partial s_j} u(t, \vec{p}_i) \right], \quad (5)$$

where $\mathfrak{D}\mathfrak{M}$ denotes a QAM demodulation operator. Defining Jacobian matrices \mathbf{J}_ℓ with

$$(J_\ell)_{i,j} = \mathfrak{D}\mathfrak{M} \left[\frac{\partial}{\partial s_j} u(t_\ell, \vec{p}_i) \right], \quad (6)$$

where $\frac{\partial}{\partial s_j} u(t_\ell, \vec{p}_i)$ denotes a linearised approximation of the full wavefield with respect to the permittivity perturbation of a single point scatterer (triangle), one obtains a linearised forward model

$$\mathbf{y}_\ell = \mathbf{J}_\ell \tilde{\mathbf{x}} + \mathbf{n}_\ell \quad \ell = 1, 2, \dots, N, \quad \text{i.e.,} \quad \mathbf{y} = \mathbf{L}\tilde{\mathbf{x}} + \mathbf{n} \quad (7)$$

for the inverse problem of reconstructing \mathbf{x} given the data \mathbf{y} . Here,

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \\ \vdots \\ \mathbf{J}_N \end{pmatrix}, \quad \text{and} \quad \mathbf{n} = \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_N \end{pmatrix}. \quad (8)$$

The matrix \mathbf{L} determines the linearised effect of the scatterers on the data \mathbf{y} as a function of the perturbations of the exact discretised real relative permittivity $\tilde{\mathbf{x}}$, and the vector \mathbf{n} is the noise term. We assume here that the noise term includes both measurement and modelling inaccuracies. Of these, the latter are associated with the difference $u - u_s$, where u is the actual wave and u_s is its simulated counterpart which can include both discretisation and general modelling errors.

2.2. Wave propagation in high and low contrast medium

We focus on inverting Eq. (7) under the assumption that \mathbf{L} contains inevitable modelling errors following from wavelength-induced uncertainties in the wave simulations. The carrier wave, in which the inputs are imposed, has 4 and 6 times the frequency of the baseband signal for the two systems of this study. In QAM, the transmitter uses the input signal to vary the carrier's amplitude, hence creating a modulated signal.

Specifically, a wavefield propagated through a complex domain undergoes changes as a result of permittivity variations within the spatial domain leading to effects such as refraction and reflection, on the material boundaries. The boundary effects give rise to changes in the wave amplitude, phase, and direction of propagation depending on the contrast of the medium described by the refractive index (Bousquet, 2017). The greater the refractive index is, the stronger the effects are. The refractive index $n = c_0/c = \sqrt{\epsilon'_r}$ is obtained as the ratio of the free space velocity c_0 and the velocity c in the medium, via the real relative permittivity ϵ'_r . Since the frequency is maintained everywhere in the medium, there is a smaller velocity and a shorter wavelength in the high-contrast (high permittivity) target than in free space or low-contrast target. Hence, the complexity of the wave propagation, and thereby the occurrence of wavelength-induced uncertainties increases along with the value of the refractive index causing an increase in the point-wise difference between the actual and simulated wave; See Fig. 1.

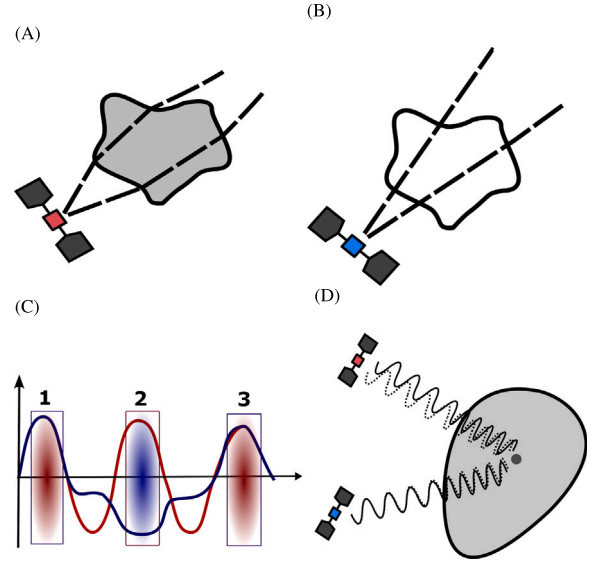


Fig. 1. Forward modelling error sources, their effect and discrepancy. Schematic illustration of signal propagated through a high contrast (A) and low contrast (B) domain. Since the velocity of the wave, and thereby wavelength decreases when the contrast increases, there is a phase shift of π as the signal propagates from a medium of low to a medium of high refractive index (Daniels, 2004). Moreover, the increase in the probing signal energy concentrated at the target surface is likely to increase the amplitude errors, resulting into larger refraction in the high contrast medium due to increased deviations from the signal energy. (C) Wavelength-induced uncertainties as a result of phase misfit can lead to large random fluctuations in the demodulated baseband data; three (1, 2 and 3) of the local maxima of the demodulated red wave coincide with those of the blue wave. The second peak has a flipped sign due to an opposite phase of the carrier (areas with matching peaks shaded in red and flipped peaks shaded in blue). More generally, phase misfit can cause any phase error in the complex plane while the difference between the sign corresponds to the phase angle of π in the complex plane. (D) Two signals (solid and dotted wave) are considered to be in the same phase if the phase angle difference is less than or equal to a quarter of pi, i.e., $\phi \leq \pi/4$. We assume that this discrepancy condition is satisfied for a wavefield propagating the length ℓ inside the target Ω with real relative permittivity ϵ'_r , if the maximum frequency of the wavefield is less than or equal to the baseband frequency (pulse bandwidth) f_B , and if the uncertainty of these parameters satisfies the discrepancy condition in Eq. (17). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.3. Structure detectability in inverse reconstruction

The inverse wave modelling in this study is characterised by a limited spectrum as described in the wave propagation phenomena in Section 2.2. It can be expected that there are intense scattering peaks from different parts of the domain that cannot be predicted by our forward model, meaning that the deviations are replicated or amplified in the inverse solution. Therefore, it is difficult to distinguish the surface and void back scattering data with the full-wave modelling approach in the absence of prior information on the domain structure. While the existence of a scattering obstacle, for example, a void beneath a surface layer might be predicted based on the difference between the simulated and background data, obtaining the magnitude of the scattering perturbation for the full domain is difficult under limited spectral coverage of the data (Dogan et al., 2017). Hence, our focus in this study is on the overlap obtained between the exact structure and reconstruction.

2.4. Prior model for wavelength-induced uncertainties

Wavelength-induced uncertainty in a propagated wave signal can be viewed from the spatial and spectral perspectives. The Nyquist sampling criterion plays an important role in determining the number of points, sampled from a wave signal. We consider a pulse of length

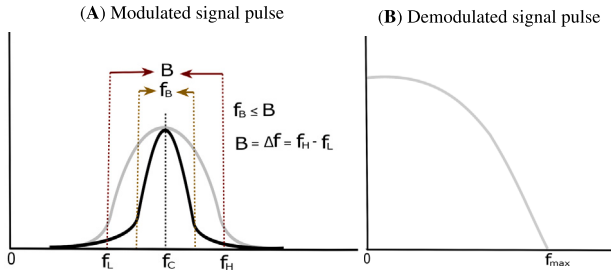


Fig. 2. Schematic diagram of (A) modulated and (B) demodulated signal pulses. The radar bandwidth B bounds the pulse bandwidth f_B from above, hence the propagated pulse is always limited by the instrument capacity. Here, f_H and f_L are the upper and lower cut-off frequencies respectively, f_c is the centre frequency of the modulated signal pulse, while f_{\max} is the maximum frequency of the demodulated signal pulse.

T and pulse bandwidth $1/T = f_B \leq B$ which matches the twice maximal frequency of the demodulated signal and is bounded by the radar bandwidth B from above; See Fig. 2.

The pulse is modulated via the QAM approach which incorporates an in-phase and quadrature components into the signal resulting in the centre frequency f_{centre} , corresponding to the centre of the spectrum with a pulse bandwidth f_B . The Nyquist criterion is defined by the spatial and temporal sampling rate S_s and S_t , stating $S_s \geq 2f_{\max}/c$ and $S_t \geq 2f_{\max}$. This implies that, the QAM demodulated baseband signal can be captured fully if the sampling frequency is greater than double the highest frequency f_{\max} contained in the baseband signal (Orfanidis, 1995). In case of QAM, $f_{\max} = f_B/2$ i.e., the maximum pulse frequency corresponds to half the pulse bandwidth for a smooth pulse. We use the first derivative Blackman–Harris (BH) pulse, which is employed often in geophysical FETD modelling applications as the probing pulse of a radar signal. The modulated signal is transmitted and propagated through the tomographic target Ω_1 , after which it is demodulated to obtain the final simulated measurement data \mathbf{y} containing frequencies on the baseband frequency interval $[0, f_{\max}]$.

2.4.1. Model for grid reduction

The principles of compressed sensing, see, e.g., Egiazarian et al. (2007), provide a viable means to reduce the wavelength-induced uncertainties. That is, a sparse and a stable reconstruction of the unknown permittivity perturbation can be obtained utilising a reduced data coverage in the reconstruction process. To avoid phase shifts in the eventual signal, i.e., the demodulated data wave, a statistically motivated approach is needed to reduce spatial uncertainty based on the assumption that the wavelength-induced uncertainties are random. Among other possible approaches are the classical regularised inversion methods or nonlinear data fitting in the form of least squares optimisation (Liu et al., 2018; Tejero et al., 2018; Dong et al., 2020). We assume that the pulse bandwidth has been set so that any data fluctuation \mathbf{z} related to frequencies higher than that of the demodulated wave, i.e., the pulse bandwidth f_B , are noisy due to modelling inaccuracies, and therefore, constitute a nuisance to be marginalised out of the final reconstruction. Because the higher frequency fluctuations are regarded as noise, we set the spatial sparse grid density to half the wavelength corresponding to the bandwidth frequency. Given a simulated measurement data \mathbf{y} , and the corresponding data fluctuations \mathbf{z} , the permittivity perturbation distribution of a target domain can be approximated by a reconstruction ($\mathbf{x} = \bar{\mathbf{x}} + \eta$) assuming the distribution of the permittivity perturbation is unknown for a target domain with a low-to-moderate complexity in terms of its geometry, where η is the deviation from the exact permittivity perturbation. Associating the unknown permittivity perturbation distribution \mathbf{x} with a conditional posterior probability density of \mathbf{x} given \mathbf{y} and \mathbf{z} , $\pi(\mathbf{x} | \mathbf{y}, \mathbf{z})$, the effect

of \mathbf{z} on \mathbf{x} in which both are dependent, can be marginalised as (Finetti, 1972):

$$\pi(\mathbf{x} | \mathbf{y}) = \int_{\mathbb{R}^m} \pi(\mathbf{z}) \pi(\mathbf{x} | \mathbf{y}, \mathbf{z}) d\mathbf{z}, \quad (9)$$

where $\pi(\mathbf{z})$ is a probability density for the modelling errors, especially wavelength-induced uncertainties; $\pi(\mathbf{x} | \mathbf{y})$ is the marginalised density, and \mathbb{R}^m is the space of \mathbf{z} . Furthermore, the expectation of \mathbf{x} given the data \mathbf{y} can be obtained as:

$$\mathbb{E}(\mathbf{x} | \mathbf{y}) = \mathbf{x}^\dagger = \int_{\mathbb{R}^n} \mathbf{x} \pi(\mathbf{x} | \mathbf{y}) d\mathbf{x}, \quad (10)$$

where \mathbb{R}^n is the space of \mathbf{x} . In our current approach, \mathbf{x} given \mathbf{y} and \mathbf{z} is determined by a classical regularised inversion scheme, that is, the conditional distribution of \mathbf{x} given \mathbf{y} and \mathbf{z} corresponds to Dirac's delta function. Substituting Eq. (9) in Eq. (10) where $\pi(\mathbf{x} | \mathbf{y}, \mathbf{z})$ is replaced with the Dirac's delta function $\delta(F(\mathbf{y}, \mathbf{z}) - \mathbf{x})$. According to the delta function's sifting property, this gives

$$\mathbf{x}^\dagger = \int_{\mathbb{R}^m} \int_{\mathbb{R}^n} \mathbf{x} \delta(F(\mathbf{y}, \mathbf{z}) - \mathbf{x}) d\mathbf{x} \pi(\mathbf{z}) d\mathbf{z}. \quad (11)$$

The simplest model is to assume that the wavelength-induced uncertainties are IID or nearly IID and that the posterior is set by a deterministic function $\mathbf{x} = F(\mathbf{y}, \mathbf{z})$. Thus, the expectation is obtained as

$$\mathbf{x}^\dagger = \int_{\mathbb{R}^m} F(\mathbf{y}, \mathbf{z}) \pi(\mathbf{z}) d\mathbf{z}. \quad (12)$$

Eq. (12) can be resolved by the Monte Carlo sampling to save computing resources (Liu, 2008). However, to ensure convergence of the sampling, the realisations \mathbf{z}_k should be sampled from the distribution $\pi(\mathbf{z})$ or an arbitrary distribution very close to $\pi(\mathbf{z})$ and assumed to be uniform in this case. The law of large numbers and the generalised central limit theorem (Liu, 2008) motivate the convergence if the sample in question is weakly enough correlated and the convergence rate $\mathcal{O}(K^{-1/2})$ of the sample-based mean

$$\frac{1}{K} \sum_{k=1}^K F(\mathbf{y}, \mathbf{z}_k) \rightarrow \mathbf{x}^\dagger, \quad \text{when } K \rightarrow \infty. \quad (13)$$

To formulate this approach for spatial point sets, we define sets \mathcal{P}_k and \mathcal{Q} in which the mutual distance between any two points is less than or equal to $\lambda_{f_{\max}}/2 = c_0/(2f_{\max})$ and $\lambda_{\text{centre}}/2 = c_0/(2f_{\text{centre}})$, respectively. That is, \mathcal{P}_k is dense enough to satisfy the spatio-temporal Nyquist criterion w.r.t. the information content of the scattered field (Bucci and Franceschetti, 1987). It is further assumed that \mathcal{Q} contains the full set of measurement points and \mathcal{P}_k , $k = 0, 1, 2, \dots, k$ with an equal number of points in each is its subset, $\mathcal{P}_k \subset \mathcal{Q}$. As the measurements are inverted after the QAM demodulation, if modelling errors are absent, each dataset to be inverted has a full coverage when the criterion w.r.t. f_{\max} is satisfied. However, due to the existence of the modelling errors, each dataset satisfying this criterion needs to be considered rather as a subset of the full data fulfilling the criterion w.r.t. f_{centre} . As we consider the modelling error for a given set to be a random variable, the sample-based mean in Eq. (13) can be realised by first finding a suitable point set sample \mathcal{P}_k which can be obtained either by a regular or randomised point selection process.

2.4.2. Model for phase error formation

A complementary approach to reduce uncertainty in the spectral sense is to apply a discrepancy condition. Assuming that the phase angle shift between the modelled and measured demodulated signal is maximally $\pi/4$, the signal is measured for a two-way path whose length inside the target is ℓ . If the signal penetrates to the centre of the target from each direction, providing a full signal coverage for the interior, then ℓ can be associated with the largest diameter of the target. In order to maintain the phase angle difference within the interval $[0, \pi/4]$, the path length per wavelength inside the target, defined as

$$N = \ell / \lambda_f = \frac{\sqrt{\epsilon_r} f \ell}{c_0}, \quad (14)$$

for the modelled and measured wavefield, must not differ more than $\pi/4$ in phase which is equivalent to one-eighth of the pulse cycle. This gives rise for the following definition of the phase discrepancy, d , which is based on the maximum relative difference of N :

$$d = \frac{N + \frac{1}{8}}{N} - 1 = \frac{8\sqrt{\epsilon'_r} f \ell + c_0}{8\sqrt{\epsilon'_r} f \ell} - 1, \quad (15)$$

where ϵ'_r represents the average real relative permittivity of the target. The maximum deviation between N_{sim} and N resulting from the simulated and actual wave propagation is, thereby,

$$1 - d \leq \frac{N_{sim}}{N} \leq 1 + d \quad \text{or} \quad \left| \frac{\Delta N}{N} \right| \leq d, \quad (16)$$

where $\Delta N = N_{sim} - N$. By combining the Eqs. (14) and (16), this can be expressed in terms of absolute total derivative as

$$\frac{|\Delta \epsilon'_r|}{2\sqrt{\epsilon'_r}} + \frac{|\Delta f|}{f} + \frac{|\Delta \ell|}{\ell} \leq d. \quad (17)$$

Firstly, this inequality allows for estimating the discrepancy, given the uncertainty of the *a priori* real relative permittivity estimate ϵ'_r and that of the path length ℓ , i.e., the more complex the geometry and the higher the contrast of the domain the more uncertain ℓ . Secondly, together with the definition of the discrepancy in Eq. (15), it also enables estimating the maximal bandwidth applicable in the measurement; that is, the discrepancy implied by the permittivity, path length and frequency deviation should coincide with the *a priori* uncertainty estimates obtained for these quantities. Thirdly, the uncertainty given by Eq. (17) can be incorporated into the forward modelling process through a surrogate approach, assuming that all uncertainty follows from frequency deviation Δf , which refers to the possible frequency fluctuations of the modulated signal which are reflected to the baseband data through the demodulation process. Since the effect of such deviations can be evaluated via signal demodulation, one can use surrogate approach which allows avoiding repetitive and computationally expensive wave simulations. We define a surrogate frequency f_s to be a Gaussian random variable with mean f_m and standard deviation d matching with the discrepancy, i.e.,

$$f_s = f_m + d\tau, \quad (18)$$

where τ is a zero-mean Gaussian random variable with standard deviation equal to one and f_m is the actual centre frequency f_{centre} . The formulation in Eq. (18) models the uncertainty as a surrogate frequency f_s for other uncertainty types explained in Eq. (17). Eq. (17) shows how frequency, velocity, and path-length uncertainties contribute to the wavelength-induced uncertainties. Thus, we can use frequency perturbations as a surrogate for the velocity and path-length uncertainties. Namely, it is difficult to perturb the permittivity distribution for each case and similarly the domain structure to account for the uncertainties relating to velocity and path-length. This would require performing an extensive set of computationally costly full-wave simulations, increasing the already expensive calculations by a large factor. Hence, perturbing the frequency is used to create similar effect as what would be obtained by perturbing the permittivity and domain structure. This is further described in Section 2.6.2 and implemented in the case of randomised averaging approach experiments in this study.

2.5. Inversion techniques

Full-wave tomography is challenging, partly because there is a need to select an inversion method that appropriately solves the problem of reconstructing scattered waves from the target domain. Expectedly, there is a trade-off between the performance and complexity of several inversion methods available.

2.5.1. Total variation

Total variation (TV) is an inversion method that has evolved from image denoising application to more robust applications of inverse problems (Luisier and Rodríguez, 2013). It has been found to be stable in tomographic reconstruction of distributional information based on sparse data and to even enable finding a robust reconstruction assuming that the distribution to be reconstructed is also sufficiently sparse (Sorsa et al., 2020; Candes et al., 2006; Candès et al., 2006). We obtain a TV regularised solution of the linearised forward model, i.e., $\mathbf{x} = \min_{\mathbf{x}} F(\mathbf{x})$ with

$$F(\mathbf{x}) = \|\mathbf{L}\mathbf{x} - \mathbf{y}\|_2^2 + 2\sqrt{\alpha}\|\mathbf{D}\mathbf{x}\|_1, \quad (19)$$

following the approach of Pursiainen and Kaasalainen (2016) and Takala et al. (2018b), i.e.,

$$\mathbf{x}^\ddagger = \mathbf{L}^* \mathbf{y}, \quad (20)$$

$$\mathbf{x}_{\kappa+1} = (\mathbf{L}^* \mathbf{L} + \mathbf{D}^T \Gamma_\kappa \mathbf{D})^{-1} \mathbf{x}^\ddagger. \quad (21)$$

Here, \mathbf{x}^\ddagger is a backpropagated estimate for \mathbf{x} . $\Gamma_\kappa = \text{diag}(|Dx_\kappa|)^{-1}$ for all $\kappa \geq 0$, where $\Gamma_0 = \mathbf{I}$, is the weighting matrix limiting the magnitude of \mathbf{x} to avoid numerical instability due to division by a very small value. \mathbf{D} is a normalised positive definite and invertible derivative operator computed over the edges of \mathcal{T} , defined as

$$D_{i,j} = \frac{\|\mathbf{L}\|_{\text{fro}}}{\sqrt{N}} \left(\beta I_{i,j} + \alpha \frac{(2I_{i,j} - 1) \int_{\mathcal{T}_i \cap \mathcal{T}_j} ds}{\max_{i,j} \int_{\mathcal{T}_i \cap \mathcal{T}_j} ds} \right), \quad (22)$$

where $I_{i,j} = 1$ if $i = j$, $I_{i,j} = 0$ if $i \neq j$, and α and β are the regularisation and smoothing parameters of the solution, and the average Frobenius or column norm $\|\mathbf{L}\|_{\text{fro}}/\sqrt{N}$ is regarded as a normalising constant. The final reconstruction is found as a vector containing the absolute value of the entries in the final iterate.

According to the general theory of regularisation (Engl et al., 1996; Kaipio and Somersalo, 2006) the noise effects in the reconstruction should be suppressed by ensuring that the regularisation level is great enough to prevent noise corruption of the objective function. Here, the TV regularisation is considered to be balanced when the two terms in Eq. (19) are roughly equal, i.e., if $\sqrt{\alpha}$ approximately matches the expected relative noise fluctuations in the data. For example, if the total effect of known and unknown noise is ≈ 20 dB, then $\alpha \approx 0.01$. The smoothing parameter β defines a diagonal weight for TV, ensuring that \mathbf{D} is invertible. We choose the magnitude of this parameter to be slightly less than α^2 to ensure the appropriate invertibility of \mathbf{D} of the TV regularisation while keeping the contribution of the diagonal weighting below the expected noise effects.

2.5.2. Filtering via truncated singular value decomposition

Singular value decomposition (SVD) (Golub and Van Loan, 2013) is a traditional means of data compression. We combine the statistical approach described in Section 2.4 with the truncated SVD (TSVD) to filter spatio-temporal data (Ludeno et al., 2020). This is motivated by the assumption that scattering consists of separate scattering patterns $n = 1, 2, 3, \dots, m$, organised in descending order by their amplitude. The spatial pattern and time-dependence are described by two orthogonal sets of vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ and $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$. The vector entries in \mathbf{u}_n and \mathbf{v}_n correspond to the spatial points and time steps of the measurements, respectively. Consequently, the matrix $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N)$ contains the full data and the TSVD filter is defined as follows:

$$\mathbf{x}^\ddagger = \mathbf{L}^* \mathbf{y} \approx \sum_{n=1}^p \sigma_n \mathbf{K}_n^* \mathbf{u}_n, \quad \text{where} \quad \mathbf{K}_n = \sum_{\ell=1}^N \bar{v}_{\ell,n} \mathbf{J}_\ell, \quad (23)$$

σ_n denotes the amplitude of the n th most intense pattern, and $p \leq m$ is selected so that the energy corresponding to the remaining sum is approximately that of the noisy signal, i.e., $\sum_{n=p+1}^m \sigma_n^2 = \|\mathbf{n}\|_2^2$. In essence, we choose p such that $\sum_{i=1}^p \sigma_i^2 / \sum_{n=1}^m \sigma_n^2 \geq 1/(1+10^{-\frac{\mu}{20}})^2$, where $i = 1, \dots, m$ and μ is the noise level. The right-hand side provides a lower

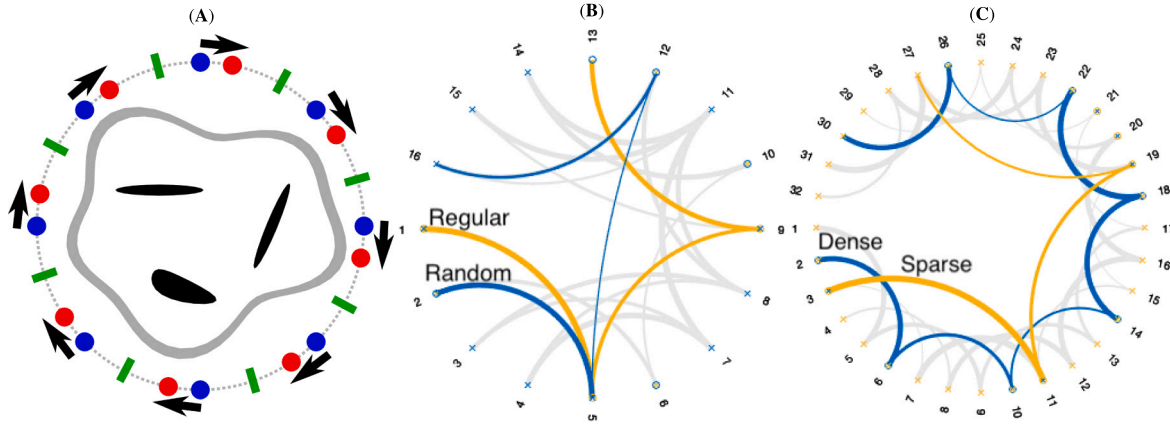


Fig. 3. The target Ω_1 applied in the two-dimensional numerical simulations. (A) The domain consists of an interior part, three voids and a surface layer (mantle) with complex relative permittivity values $\epsilon_r = 4.0 + j0.03$, $\epsilon_r = 1.0$, and $\epsilon_r = 3.0 + j0.02$, respectively. The synthetic orbit used in the numerical experiments is visualised by the circle surrounding the target. The blue and red dots depict two monostatic signal configurations with uniform distance between the points on the synthetic orbit. The sectors are depicted by the intervals between the green lines. The complete measurement point set including all the sectors satisfies the Nyquist criterion w.r.t. the centre frequency of the signal. (B) In the regular turn, such monostatic point sets are obtained by rotating an initial set of points from different sectors as shown by the arrows and, thus, covering each point in the full set of measurements. In the randomised approach, a uniformly distributed random point is picked independently from each sector, whose length satisfies the Nyquist criterion w.r.t. the pulse bandwidth as described in Section 2.4.1. (C) In the sparse density configuration, points are selected from each sector with intervals large enough to cover the total receiver points on the synthetic orbits. The dense point configuration has twice as many points as the sparse case i.e., half the interval for point selection as compared to the sparse configuration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bound by the assumption that the amplitude of the signal is a sum of the noiseless and noisy signals.

In this study, the TSVD approach is applied in the backpropagation part in Eq. (20) of the TV regularisation process. We compare its performance to the case of full data. Here, $\mathbf{K}_n^* \mathbf{u}_n$ can be interpreted as a backpropagated pattern \mathbf{u}_n with intensity σ_n and time dependence \mathbf{v}_n . This formulation of TSVD is used since the scattering patterns are assumed to be well localised in the time domain.

2.6. Numerical experiments

2.6.1. Two-dimensional domain

In the numerical experiments, we investigate inverting full-waveform data in two dimensions using the target Ω_1 of Sorsa et al. (2020) with unitless diameter 0.28 scaled to the estimated diameter 160 m ($s = 571$) of Dimorphos and a synthetic orbit of diameter 0.32; see Fig. 3. The far-field formulation is omitted due to the difference in attenuation between 2- and 3-Dimensional domain (Takala et al., 2018a). We considered a background model with a constant permittivity $\epsilon'_r = 4$ and a detailed (exact) model Ω_1 as shown in Fig. 3. The interior part of the detailed model is given a real relative permittivity $\epsilon'_r = 4$, excluding the three voids having a vacuum permittivity ($\epsilon'_r = 1$) and the surface layer (mantle) with real relative permittivity value of $\epsilon'_r = 3$. The existence of a surface layer for an asteroid target is predicted in the impact studies by Jutzi and Benz (2017) and the overall structure is considered to have a relatively low density (Carry, 2012). The permittivity of asteroid minerals such as kaolinite and dunite match roughly these values regarding solid and powder composition (with grain size distribution of approximately 1-10 μm) for the interior and surface layer, respectively (Herique et al., 2002; Hérique et al., 2018). The loss tangent δ which determines the imaginary part of the complex relative permittivity $\epsilon_r = (1 + i\delta)\epsilon'_r$ is considered as a nuisance parameter, as the internal absorption distribution is known to be challenging to be reconstructed (Barriot et al., 1999). Hence, we are not inverting for this parameter, but it causes an error source in the simulation. We set δ to be in the range from 0.0005 to 0.04 proposed for asteroid minerals in Kofman (2012). Here, we consider the internal absorption simply as a source of error; however, it should be noted that for an average loss tangent $\tan(\delta) = 0.008$ and $\epsilon'_r = 4$, an attenuation of 30 dB/km is fully realistic when using a frequency of 20 MHz and 78 dB/km when using 60 MHz given by -8.68α , where

$$\alpha = \frac{2\pi f}{c_0} \sqrt{\frac{\epsilon'_r}{2} \sqrt{1 + \left(\frac{\epsilon''_r}{\epsilon'_r}\right)^2} - 1} \quad (\text{Ulaby et al., 2014}).$$

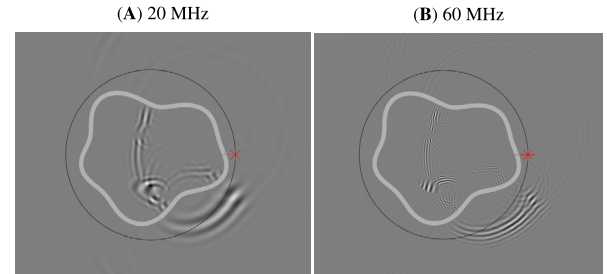


Fig. 4. Monostatic signal propagation path in the two-dimensional domain for the signal pulse (A) 20 and (B) 60 MHz at $0.22 \mu\text{s}$ measurement time ($s c_0^{-1} t$) after propagation from the transmission point marked with red asterisk. The 20 MHz signal has a single strong wavefront propagating in the domain while the 60 MHz signal has multiple weak wavefronts propagating through the domain, which is a potential source of wavelength-induced uncertainty.

2.6.2. Signal specifications

The data \mathbf{y} was obtained as the (noisy) difference between the two simulated full wavefields corresponding to the background and detailed domain described above. To avoid an inverse crime, i.e., an overly accurate data fit, these domains were discretised using two different finite element meshes.

We apply QAM modulated Blackman–Harris window as a signal pulse centred at (A) 20 and (B) 60 MHz frequencies covering the bandwidths of 10 and 20 MHz, respectively; See Table 1 and Fig. 4. The centre frequency of 20 MHz has been proposed as a potentially feasible low end for a CubeSat-based penetrating radar exploration in a recent concept study (Bambach et al., 2018), while 60 MHz will be applied in the JuRa investigation (Herique et al., 2019, 2020a). The general knowledge of asteroid composition and radar signal penetration suggests that the centre frequency should be between 10 and 100 MHz (Binzel and Kofman, 2005; Kofman, 2012) to enable full coverage of the wavefield inside the target.

We consider the following four monostatic signal configurations (I)–(IV) (Fig. 3, Table 2) for obtaining the full-wave data. These configurations differ by the averaging and filtering strategy applied in processing the data. The averaging process utilises the simulated full-wavefield which has been obtained around the tomographic target, with spatial and temporal resolution satisfying the Nyquist criterion with respect to the centre frequency.

Table 1

Blackman–Harris signal pulses and their corresponding frequencies, bandwidths, pulse durations, cycles per pulse, wavelengths, signal wavelength–diameter ratios, and sampling rates (S_s).

| Signal pulse | Centre freq. (MHz) | Bandwidth (MHz) | Pulse duration (μ s) | Cycles per pulse | λ (m) | Diameter/ λ | S_s |
|--------------|--------------------|-----------------|---------------------------|------------------|---------------|---------------------|-------|
| (A) | 20 | 10 | 0.190 | 4 | 7.5 | 21 | 0.067 |
| (B) | 60 | 20 | 0.095 | 5.5 | 2.5 | 64 | 0.134 |

Table 2

Spatial resolution of the full dataset before and after the averaging process with their corresponding averaging techniques. The approach to the spatial averaging in configurations (I)–(IV) of the numerical experiments is described in Fig. 3, and the signal pulses (A) 20 and (B) 60 MHz described in Table 1. The 3rd and 4th columns give the total and averaged number of spatial points in the full dataset and after the averaging process, respectively. The 5th column includes the number of turns in the regular averaging process for each averaging resolution, and the 6th gives the number of spatial point sets and frequencies applied in the randomised averaging. The last column indicates the type of filtering applied to each case in the experiment.

| Signal pulse | Signal configuration | Total points | Spatial averaging | Regular turns | Randomised turns/frequency | Filter |
|--------------|----------------------|--------------|-------------------|---------------|----------------------------|------------|
| (A) | (I) | 128 | dense/64 | 2 | – | Unfiltered |
| | (II) | 128 | dense/64 | 2 | – | TSVD |
| | (III) | 128 | dense/64 | – | 8/5 | Unfiltered |
| | (IV) | 128 | dense/64 | – | 8/5 | TSVD |
| | (I) | 128 | sparse/32 | 4 | – | Unfiltered |
| | (II) | 128 | sparse/32 | 4 | – | TSVD |
| | (III) | 128 | sparse/32 | – | 8/5 | Unfiltered |
| | (IV) | 128 | sparse/32 | – | 8/5 | TSVD |
| (B) | (I) | 384 | dense/128 | 3 | – | Unfiltered |
| | (II) | 384 | dense/128 | 3 | – | TSVD |
| | (III) | 384 | dense/128 | – | 8/5 | Unfiltered |
| | (IV) | 384 | dense/128 | – | 8/5 | TSVD |
| | (I) | 384 | sparse/64 | 6 | – | Unfiltered |
| | (II) | 384 | sparse/64 | 6 | – | TSVD |
| | (III) | 384 | sparse/64 | – | 8/5 | Unfiltered |
| | (IV) | 384 | sparse/64 | – | 8/5 | TSVD |

2.6.3. Averaging approaches

The regular averaging approach refers to a systematic method of selecting measurement points located at uniform angular distances from each other and averaging the (reconstructed) candidate solutions corresponding to such signal configurations. The points are independently selected at intervals corresponding to the spatial sampling rate S_s in Table 2. Two different spatial resolutions, sparse and dense, are considered in the reconstruction averaging process in this study. The sparse cases have an average point density corresponding to the Nyquist criterion for the pulse bandwidth, and the dense cases have twice as much of the point density of the sparse cases. Thus, 32 (sparse) and 64 (dense) points are sampled for system (A), while 64 (sparse) and 128 (dense) points are sampled for system (B), where systems (A) and (B) have a total of 128 and 384 points, respectively. To ensure that the resulting point sets cover the full set of measurement data, multiple constellations of measurement points are sampled. In essence, the regular averaging approach essentially has 4×32 (4 constellations of sparse) and 2×64 (2 constellations of dense) points for system (A), and 6×64 (6 constellations of sparse) and 3×128 (3 constellations of dense) points for system (B). The number of constellations (4, 2, 6, and 3) in the regular averaging approach are referred to as the regular turns in Table 2 and illustrated in Fig. 3.

In the randomised averaging approach, the spatial points are selected in a uniformly distributed random manner (no equal distance between points) from the sectors constituted by the regular turn. In addition, the centre frequency is slightly perturbed by varying the surrogate centre frequency in the demodulation process according to the discrepancy condition in Eq. (18). However, the number of constellations of the measurement points was fixed to 8 for both sparse and dense points in the two systems. In essence, we demodulate the wave-field data for 5 different surrogate centre frequency realisations, using the discrepancy d as the standard deviation and actual centre frequency f_{centre} as the mean, f_m . Hence, the estimated inverse solution for each point is averaged over 5 different frequencies and then all 8 spatial

constellations. The sampling rate of the regular averaging approach was retained, resulting to a system with similar number of measurement points i.e., 32 (sparse) and 64 (dense) points for system (A), and 64 (sparse) and 128 (dense) points for system (B). This guarantees that every point has the chance to be selected and the full coverage of the data is ensured since a uniformly distributed random point is selected from each sector illustrated by the green lines in Fig. 3.

In configurations (I) and (II), regular spatial averaging is applied to unfiltered (full) data in (I) and TSVD filtering (II). Configurations (III) and (IV) utilise a randomised point set and frequency perturbed data with full data (III) and TSVD filtering (IV).

2.6.4. Noise

The measurement noise was simulated by adding a constant noise level relative to the maximum signal amplitude (Peak SNR). The inversion results are obtained for two different magnitudes of the noise; with the lower noise level, the signal-to-noise ratio (SNR) of the simulated data was 20 dB and with the higher one, it was 12 dB. The higher SNR is motivated by the observations of the CONSERT team suggesting that the main signal peak of the measurement was found to have at least 20 dB SNR with respect to noise peaks (Kofman et al., 2015). The lower SNR has been set based on recent numerical and experimental modelling studies (Sorsa et al., 2019; Eyraud et al., 2020; Sorsa et al., 2021b) suggesting that a SNR above 10 dB might be sufficient for reconstructing the internal permittivity of a complex-structured asteroid analogue and that such an accuracy between a numerically modelled field and a laboratory measurement can be achieved experimentally. The experimental setup in Table 2 was implemented for both 20- and 12-dB noise cases.

2.6.5. Similarity and error measures

To analyse the accuracy of the reconstructions in different parts of the target Ω_1 , we define sub-domains \mathcal{S}_1 (voids) and \mathcal{S}_2 (surface layer) restricted by the contours of the permittivity distribution in a

descending order. The structural similarity (SSIM) of the reconstruction to the exact distribution is computed to measure the quality of the reconstructed image relative to the background (Wang et al., 2004). We also evaluate the Root Mean Squared Error (RMSE) between the exact distribution and reconstruction for Ω_1 , \mathcal{S}_1 and \mathcal{S}_2 . The Overlap Error (OE) between the detail \mathcal{S}_i and a reconstructed permittivity perturbation is defined as the ratio

$$\text{area}(\mathcal{S}_i \cup \hat{\mathcal{O}}_i) / \text{area}(\mathcal{S}_i), \quad (24)$$

where $\hat{\mathcal{O}}_i = \bigcup_{i \in I_c} T_i$ is a set composed by triangles T_i in the index set I_c which includes the most intense part of the reconstructed distribution with surface area equals to that of the given detail excluding the sets $\hat{\mathcal{O}}_j$ for $j = 1, 2, \dots, i - 1$, that is, $I_c = \{i \mid |x_i| \geq v, T_i \in \Omega_1 \setminus \bigcup_{j=1}^{i-1} \hat{\mathcal{O}}_j\}$ and $\zeta = \arg \max_v \{\sum_{i \in I_c} \text{area}(T_i) \leq \text{area}(\mathcal{S}_i)\}$.

The Wilcoxon rank sum (WRS) test, a nonparametric alternative to the two-sample t-test for two populations X and Y of assumably independent samples, was used to analyse the significance of the differences observed. The WRS tests the null hypothesis that data in X and Y are samples from continuous distributions of equal medians, against the alternative that they are not, with the assumption that the samples are independent (Gibbons and Chakraborti, 2014).

2.6.6. Numerical implementation

The computations are implemented using the openly available GPU-Torre package,¹ (Sorsa et al., 2020) which utilises the Matlab platform (Mathworks, Inc.). GPU-Torre combines a GPU-accelerated FETD forward routine and a multigrid-based inversion approach, where the triangular mesh \mathcal{T} applied to discretise the permittivity distribution in the inversion stage is coarser than the one applied in the forward wave propagation. Each matrix L was found by a GPU-accelerated deconvolution process, obtained in less than 10 s. Using Dell 5820 Workstation equipped with 256 GB RAM and 8 GB NVIDIA Quadro RTX 8000 GPU RAM, the complete FETD simulation of the signals took approximately 30 and 134 h, while it took 14 and 21 min for a single point of the (A) 20 and (B) 60 MHz frequencies, respectively. The inversion mesh \mathcal{T} consisted of 896 triangles. The total variation regularisation parameter α in Eq. (20) was chosen experimentally to be 0.02 and 0.005 for the 20 and 60 MHz systems respectively, while β in Eq. (22) was chosen to be $0.5\alpha^2$ for both cases. In all, a total of 32 numerical experiments are considered as described in Table 2 given that two noise level cases are implemented in this study.

3. Results

The results from the numerical experiments of Section 2.6 are presented in Figs. 5–11. Of these, Figs. 5 and 6 show the amplitude of the reconstructed real relative permittivity perturbation visualised on a linear scale between 0 and 1 after setting an upper threshold with the 98% quantile of the data as our normalisation. These also show the relative Overlap Error (OE) between the reconstructed details and those of the actual permittivity distribution. In Table 3, Figs. 7 and 8, the robustness of the reconstructions are analysed in a tabular form and via boxplots obtained with 10 different noise vector realisations. The norm of the reconstructed permittivity distribution, \mathbf{x} , for the different sampling steps is shown for the randomised configuration and full data in Fig. 10. The results show that the aim of reconstructing the permittivity perturbation distribution of the complex domain incorporating phase errors formulation, was achieved by a monostatic full-wave simulation and inversion with the spatial averaging, and discrepancy-based frequency averaging methods as stated in Sections 2, 2.4.1 and 2.4.2. The results show that the point selection, spatial density, and noise level used are highly consequential to how robust reconstruction can

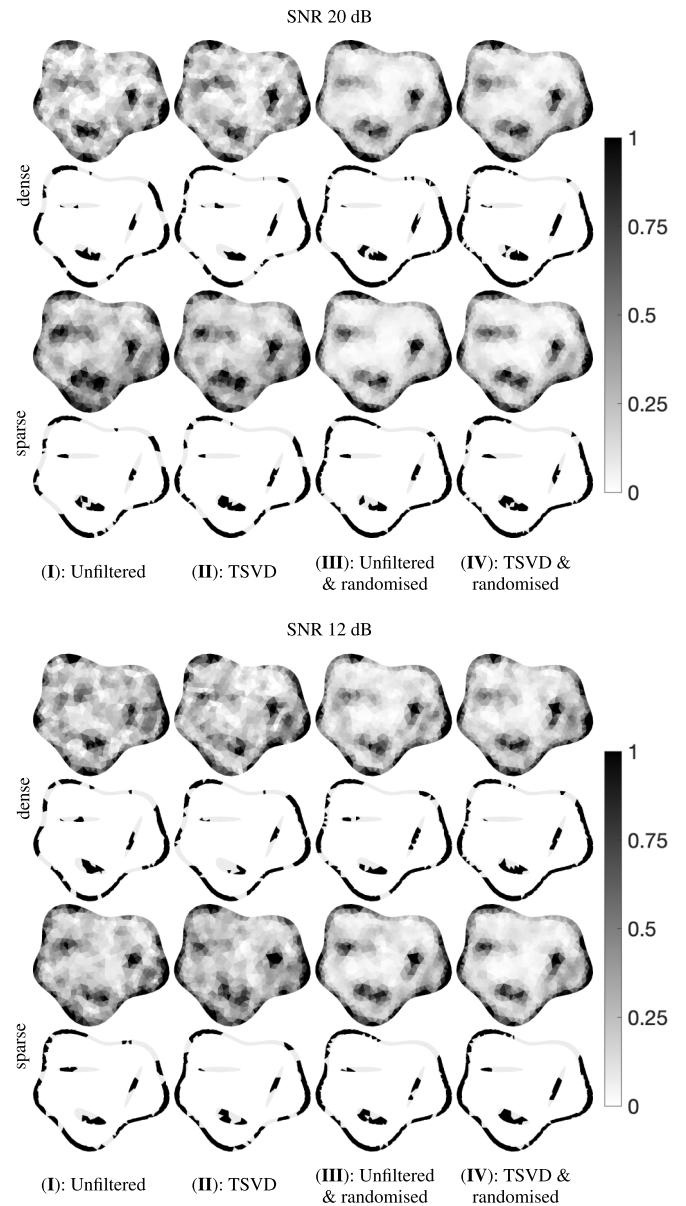


Fig. 5. Domain reconstructions obtained using the 20 MHz signal pulse (A) with SNR of 20 dB (1st–4th row) and 12 dB (7th–10th row). The reconstructions are presented in the odd rows while their corresponding overlap reconstructions (even rows) depict how much of the surface and void details are reconstructed in black (grey parts are those which do not overlap). The top row shows the results for the dense configuration with 64 points and the bottom row for the sparse one with 32 points. The results for configurations (I)–(IV) are shown in columns from left to right, respectively.

be obtained. The ideal reconstructions obtained for the low frequency case is the Low-noise-Sparse-Random-TSVD configuration with surface overlap error of 36.16% and void overlap error of 52.77%, see Fig. 5 column (IV). The ideal reconstruction obtained for the high frequency case is the High-noise-Sparse-Random-Unfiltered configuration with surface overlap error of 27.15% and void overlap error of 78.21%, see Fig. 6 column (III).

3.1. Low vs. high centre frequency

The global structure of the two-dimensional test domain, including the surface layer and voids, was reconstructed with an average surface overlap error of 38.87% and 31.7% and void overlap error of 58.39% and 78.67% for both 20 and 60 MHz signals, respectively, see

¹ <https://github.com/sampsapursainen/GPU-Torre>.

Table 3

Similarity and error estimates of the reconstructions for different numerical configurations. The 1st column presents the two signal pulses (A) 20 and (B) 60 MHz, described in Table 1, while the 2nd column shows several configurations over which these frequencies have been experimented. The corresponding levels of the configurations are also highlighted in the 3rd column. The 4th to 8th columns gives the mean values with respect to the different centre frequencies, configurations, and levels for SSIM (scales between 0 & 1 and 1 shows the best result), RMSE of the void and surface (lower limit of 0, which shows the best result), OE of the void and surface (in %, scales between 0 & 100 and 0 shows the best result)). The WRS test values are also highlighted, with 0 implying no significant differences and 1 implying a significant difference in the median values of the corresponding distributions.

| Signal pulse | Configuration | Levels | SSIM | RMSE | | | | OE % | | | |
|--------------|-------------------|------------|-------|------|-------|---------|-------|------|--------|---------|--------|
| | | | | Void | | Surface | | Void | | Surface | |
| | | | | WRS | WRS | WRS | WRS | WRS | WRS | WRS | WRS |
| (A) 20 MHz | | | 0.530 | | 2.074 | | 1.602 | | 58.387 | | 38.868 |
| | Spatial Averaging | Dense | 0.527 | 1 | 2.286 | 1 | 1.725 | 1 | 60.476 | 1 | 40.741 |
| | | Sparse | 0.533 | | 1.863 | | 1.480 | | 56.298 | | 36.995 |
| | Point Selection | Regular | 0.521 | 1 | 2.224 | 1 | 1.751 | 1 | 60.671 | 1 | 43.176 |
| | | Random | 0.539 | | 1.925 | | 1.453 | | 56.103 | | 34.559 |
| | Filtering | Unfiltered | 0.530 | 0 | 2.060 | 0 | 1.603 | 0 | 58.389 | 0 | 38.435 |
| | | TSVD | 0.530 | | 2.088 | | 1.602 | | 58.384 | | 39.301 |
| | Noise | Low | 0.530 | 0 | 2.085 | 0 | 1.657 | 1 | 55.933 | 1 | 39.795 |
| | | High | 0.530 | | 2.064 | | 1.548 | | 60.840 | | 37.940 |
| (B) 60 MHz | | | 0.673 | | 4.787 | | 3.833 | | 78.668 | | 31.785 |
| | Spatial Averaging | Dense | 0.672 | 0 | 4.797 | 0 | 3.787 | 0 | 79.527 | 1 | 31.612 |
| | | Sparse | 0.674 | | 4.777 | | 3.880 | | 77.810 | | 31.957 |
| | Point Selection | Regular | 0.680 | 0 | 4.840 | 1 | 3.957 | 1 | 80.045 | 1 | 33.058 |
| | | Random | 0.667 | | 4.735 | | 3.708 | | 77.292 | | 30.511 |
| | Filtering | Unfiltered | 0.672 | 0 | 4.780 | 0 | 3.827 | 0 | 78.339 | 0 | 31.560 |
| | | TSVD | 0.673 | | 4.794 | | 3.839 | | 79.000 | | 32.009 |
| | Noise | Low | 0.710 | 1 | 4.855 | 1 | 4.094 | 1 | 76.195 | 1 | 33.483 |
| | | High | 0.636 | | 4.720 | | 3.572 | | 81.142 | | 30.086 |

Table 3. The former gave a more ideal reconstruction outcome, see Fig. 5 columns (I–IV), in terms of RMSE and OE of the void, while the latter with higher SSIM, shows faint details and hot spots on the surface, especially in the concave part of the target and parts close to the receiver, see Fig. 6 columns (I–IV). This is expected in the high frequency case as there is more reflection at shallow depth and fewer wavefronts are received from the deeper part of the target. This anomaly limits the visibility of the voids which is obvious based on their increased RMSE and OE of the void, as there is more contrast at the surface compared to the interior part. However, the surface layer has a lesser OE in comparison to the lower frequency case, see Fig. 5 rows 2 & 4 vs. Fig. 6 rows 2 & 4.

3.2. Signal configuration

For the sparse configurations, the global structure, surface layer, and voids are generally more pronounced and uniform, with an enhanced contrast compared to the dense ones, which are of deteriorated reconstruction quality reflected by, for example, fluctuating artefacts and inconsistent intensity of the voids, see Fig. 5 row 1 vs. 3. Due to these anomalies, the average RMSEs and OEs of the dense configurations are higher than those obtained for the sparse ones, see Table 3. The difference between the results obtained with dense and sparse point density is less pronounced when the randomisation of the points and frequencies is applied. The randomised and frequency-perturbed configurations have a smooth outcome compared to the cases of regular turn configurations, which result in a weaker visibility of the interior details and structure, as well as fractures within the voids and the surface layer, see Fig. 5 columns (III–IV). For the distinguishability of the details, the randomised configuration with sparse density is a preferable set-up, especially in the lower frequency case as shown in Fig. 5 columns (III–IV), which is the benchmark reconstruction in this study. We observe that in the high frequency results, based on the SSIM, RMSE, and OE in Table 3, the global structure of the reconstructed domain is more conspicuous for the regular turn case, but not necessarily with respect to the void and surface layer. The choice

of signal configuration is more significant with higher frequency which is revealed by the comparison between results from dense regular turn and sparse randomised configurations in Fig. 6 columns (I–II) vs. (III–IV) and Fig. 8 columns (I–II) vs. (III–IV). This is also evident in Table 3 where the average OE of the void and surface are 80.04% and 33.06% for the regular turn and 77.29% and 30.51% for the randomised configuration.

3.3. Low vs. high noise

As expected, the reconstruction quality of the void was observed to be maximised with low noise (SNR 20 dB), while the higher noise level (SNR 12 dB) was observed to cause artefacts in the void. However, the latter allowed finding the interior details with a bias towards the surface, see Fig. 6 (SNR 20 dB) vs. (SNR 12 dB). The differences between the low frequency and the high frequency configurations are maintained regardless of the noise level. The structure obtained with the latter remains smooth while the visibility of the void details weakens with higher noise. This is also revealed in the results presented in Table 3, where the low noise has a lower overlap error of the void compared to high noise for both 20 and 60 MHz systems. Similarly, the global structure with respect to SSIM suggests that the difference between the noise levels seems comparable to the difference between the configurations in terms of their significance. The high noise case has lower RMSE and OE of the surface as presented in Table 3, indicating that the surface layer is better reconstructed at higher noise level. This is more pronounced in the high frequency case, see Table 3, and expected since the noise leaves more signature on the surface layer which is coupled with the intense surface reflection thus reducing the OE of the surface. The void RMSE in the 20 MHz system has a small difference between their respective noise levels as compared to the 60 MHz system, see Table 3. The noise levels are insignificant in the 20 MHz system in terms of the SSIM measure, and largely significant in the 60 MHz system, where the low noise level performs better with a similarity of 0.71 compared to the high noise at 0.64. The WRS test shows that all the error measures have significant difference between

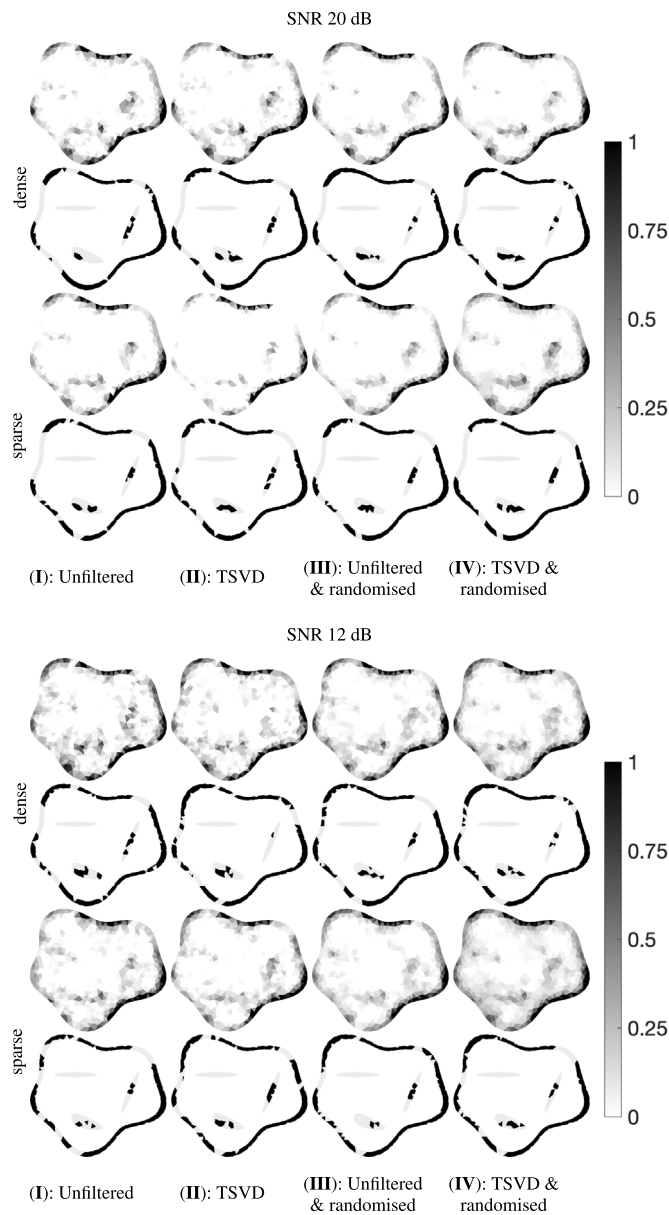


Fig. 6. As in Fig. 5 obtained using the 60 MHz signal pulse (B). The top row shows the results for the dense configuration with 128 points and the bottom row for the sparse one with 64 points. The results for configurations (I)–(IV) are shown in columns from left to right, respectively.

the two noise levels in the high frequency case while only the void RMSE and surface OE have significant differences between the noise levels for the low frequency case.

3.4. Unfiltered vs. TSVD

The TSVD filter can be observed to preserve the main details of the reconstruction, i.e., the global structure, while filtering out minor fluctuations. The effect of the TSVD filter is pronounced in the low frequency case as it smooths out random background noise from the global structure, see Fig. 5. Similarly, the ideal reconstruction in low frequency case is from the set where the TSVD has been applied, hence suggesting this has an effect in reducing background noise from the data. The reconstructions found with TSVD are biased towards the surface layer in the high frequency case, which can be considered as a predominating structure with respect to the scattered energy, and,

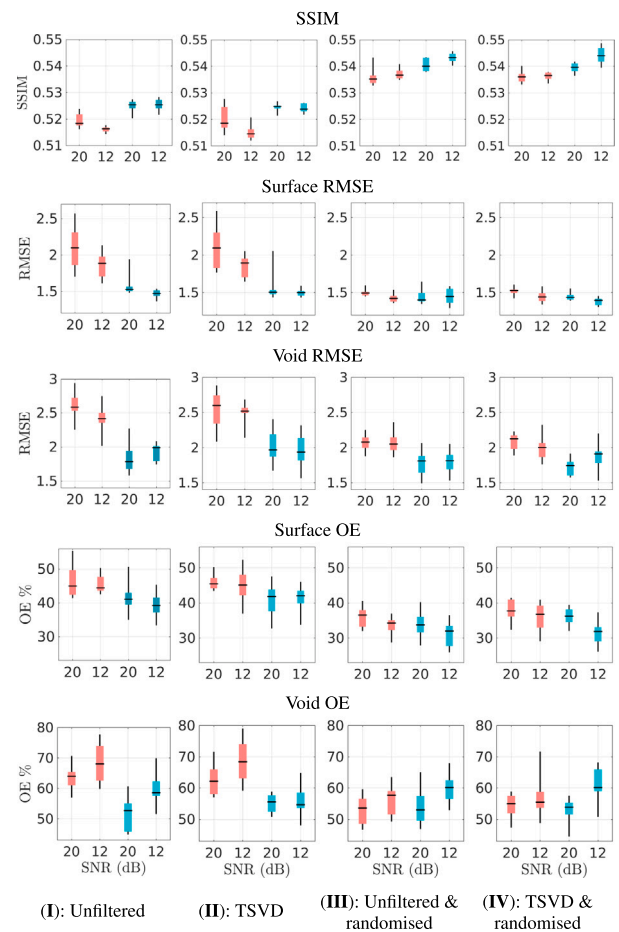


Fig. 7. Structural Similarity (SSIM, scale of 0–1 and 1 shows the best result), Root Mean Squared Error (RMSE, lower limit of 0, which shows the best result) and Overlap Error (OE %, scale of 0–100 and 0 shows the best result) for the 20 MHz signal pulse (A) with SNR of 20 and 12 dB. The red boxes correspond to the dense configuration and the cyan boxes correspond to the sparse configuration, both having their respective noise levels indicated on the horizontal axis. This figure indicates that the RMSE & OE of the surface and void are in most cases (I)–(IV) lower for the sparse as compared to the dense configuration. Similarly, the SSIM is higher for the sparse case across all the levels of configurations (I)–(IV) shown in columns from left to right, respectively. Here the most ideal reconstructions are those obtained from the sparse randomised configurations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

therefore, its contribution to the large singular values and vectors constituting the TSVD is enhanced. The unfiltered set, however, has a superior effect on the average measures in Table 3 especially for the high frequency case, hence limiting the importance of the TSVD filtering. The unfiltered configuration has its RMSE and OE of the surface and void lower than that of the TSVD for the high frequency case, while only the surface RMSE is higher than that of the TSVD configuration in the low frequency case. The WRS test indicates that there is no significant difference in the Unfiltered and TSVD median values with respect to the error measures for both 20 and 60 MHz systems.

3.5. SSIM, RMSE, and OE

The RMSE and OE are two complementary measures with the objective of evaluating the quality of the reconstructed domain. The RMSE shows the exact difference between two different distributions, while OE gives an idea of the localisation and contrast of the reconstructed details. The effect of the signal configurations on the reconstruction can

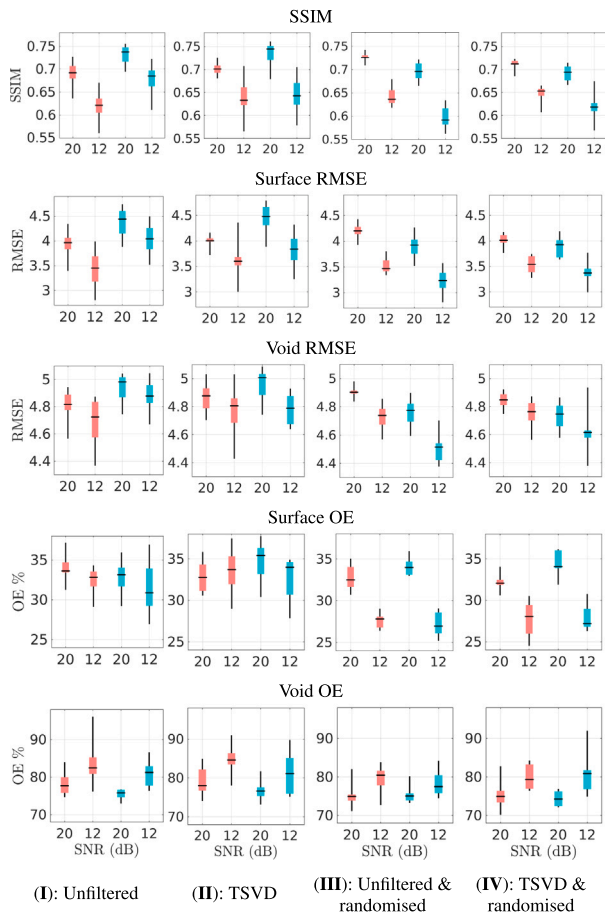


Fig. 8. As in Fig. 7 obtained for the 60 MHz signal pulse (B) with SNR of 20 and 12 dB. The regular configuration has higher SSIM, RMSE & OE of the surface for the sparse (cyan) point density as compared to the dense (red). This is opposite for the random configuration which has lower SSIM, RMSE of the surface and void for the sparse (green) point density as compared to the dense (red). The OE of the surface is lower for both random and regular configuration for the spare density points. Here the most ideal reconstruction are those obtained from the sparse randomised configurations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be observed based on the values of the metrics used in evaluating reconstruction quality, i.e., RMSE and OE. The results show that the RMSE and OE give a similar pattern in evaluating the voids and the surface layer, see Table 3. The SSIM was further used as the evaluating criteria for the global structure of the reconstructed domain, complementing the noise effect in the interior structure, hence a lower SSIM in the low frequency case as compared to the high frequency case in Table 3.

The WRS test shows that the configuration levels, i.e., dense vs. sparse and regular vs. random are significantly different for the 20 MHz case as shown in Table 3. The point selection is significant for all the measures in the low frequency case and significant for 4 out of 5 measures in the high frequency system. Similarly, when evaluating the noise level configurations, the RMSE for the surface and OE for the void also have significantly different medians for the low and high noise configurations. The TSVD vs. unfiltered levels are insignificant for all their corresponding measures. For the 60 MHz case, only the low vs. high noise configurations have significantly different medians for all their corresponding measures.

The correlation heatmap in Fig. 9 shows that the SSIM has a strong negative correlation with other measures ranging from -0.97 for OE of the surface, -0.9 for RMSE of the surface, -0.82 for RMSE of the void, and -0.72 for OE of the void for the 20 MHz case. This is expected since the best SSIM values are the highest and they correspond to the

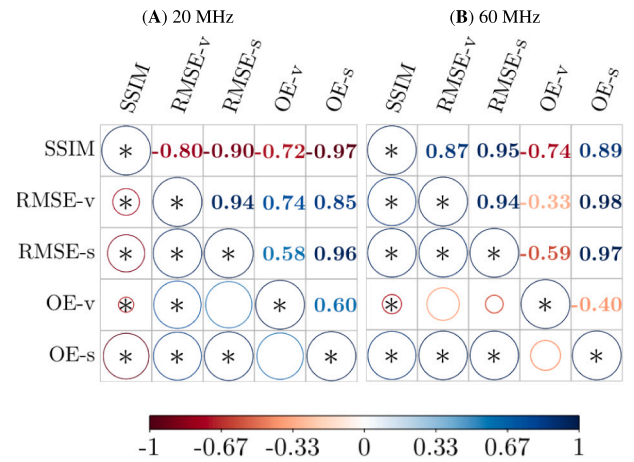


Fig. 9. Correlation heatmap between structural similarity (SSIM), Root Mean Squared Error of the void (RMSE-v), Root Mean Squared Error of the surface (RMSE-s), Overlap Error of the void (OE-v) and Overlap Error of the surface (OE-s) with respect to the mean values in Table 3. The significant correlations are shown with the asterisk sign *, the insignificant correlations are depicted with no asterisk at a 5% significance level. The size of the circle shows the magnitude of significance and insignificance; larger circle implying high significance or insignificance. The colour shows the magnitude of the correlation for both centre frequencies (A) 20 and (B) 60 MHz, configurations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

best values of other measures which are the lowest and vice versa as shown in Table 3. The negative correlation here does not relate to the reconstructions, but the measures used in assessing the reconstructions. Other measures are positively correlated with the highest being 0.96 between RMSE of the surface and OE of the surface, and the least being 0.58 between OE of the void and RMSE of the surface. In the 60 MHz case, OE of the void is also negatively correlated with other measures, the strongest being -0.74 for SSIM and the least being -0.33 for RMSE of the void. The other measures have strong positive inter-measure correlations ranging from 0.98 between RMSE of the void and OE of the surface to 0.87 between RMSE of the void and SSIM. Here, the

3.6. Convergence

An investigation of the convergence and performance of several configurations on the full data is presented in Figs. 10 and 11. The plots in Fig. 10 all start from zero on a norm scale to indicate the variation of the magnitude for each configuration. The upper limits are however retained for clear visibility of the spread in the respective distribution. The uncertainty in the reconstructed estimates x spreads as the noise increases for both 20 and 60 MHz frequencies.

The magnitude of the norm in the high frequency case is greater and its spread extends frequently outside the interquartile range. Fig. 10 also indicates the effectiveness of the sampling process. Since the sample realisations are computationally demanding to obtain, the discrepancy conditions and averaging are selected such that the randomness in the process is present and the IID assumption holds. The plot in Fig. 11 gives a comparison of the perturbation distribution between the dense (magenta histogram) and sparse configurations (cyan histogram). From this plot, it was evident that the sparse configurations have higher spread in the perturbation distribution for both 20 and 60 MHz frequencies. The dense configuration is of lower magnitude and spread compared to the sparse case, which also shows spurious outliers beyond the 10% to 90% quantile range. The effect of the randomised frequency perturbation is less evident in the dense configuration, especially in the low frequency case. The plot also shows that the outliers are smaller in the low noise case as compared to the high noise.

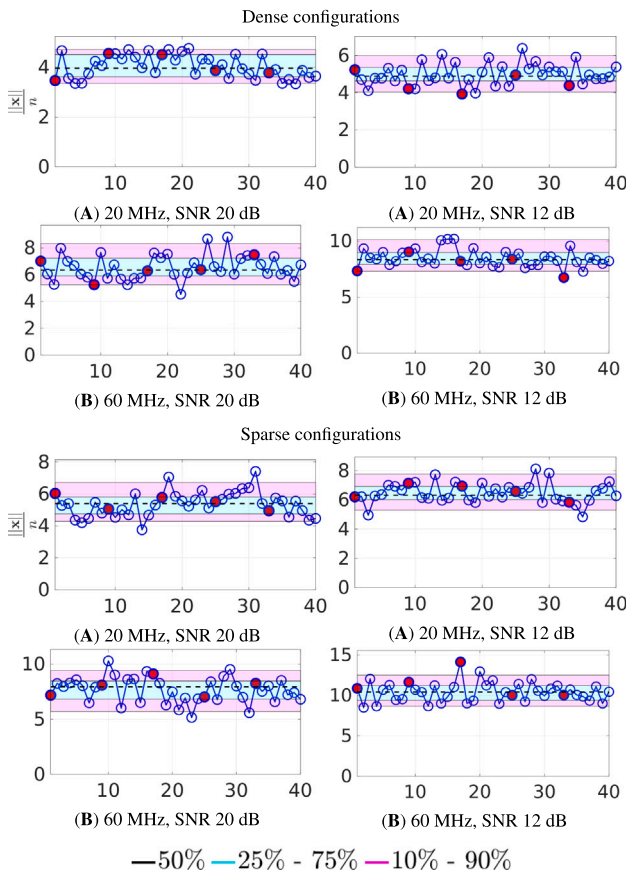


Fig. 10. The average absolute value for the entries of \bar{x} as a function of the different sampling steps in the case of the unfiltered & randomised data. The 1st row shows the results for the dense configuration, while the 2nd row presents the results for the sparse configuration. The red marker indicates an update in the centre frequency applied in the signal demodulation process. The magenta area shows the range from 5 to 95% quantile of the sample variation, the cyan area shows 25 to 75% quantile (interquartile range), and the dashed line shows the median of the samples. The magnitude of \bar{x} can be observed to grow as the noise, signal frequency, and sparsity of the measurement point configuration grows. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.7. Beyond a piece-wise constant domain

We carried out a numerical comparison between the benchmark reconstructions for the 20 and 60 MHz systems and two random field models shown in Fig. 12. The first having background amplitude fluctuation of 1 (i.e., surface, and interior part of the domain excluding the void have a permittivity distribution between 3.5 and 4.5) and the second having a background amplitude fluctuation of 4 (i.e., surface, and interior part of the domain has a permittivity distribution between 1 and 5). The reconstruction obtained for the 20 MHz system for the random field models show clear visibility of the voids and surface layer. However, when compared to the benchmark model as shown in Fig. 12, we see that the benchmark reconstruction has more visibility of the voids and surface layer. This is expected as the background permittivity contrast variation increases the discrepancy formulated in Eq. (17). The benchmark reconstruction of the 60 MHz system outperforms the reconstruction from the two random field models, in the sense that we can see the voids and surface layer of the benchmark model by visual inspection of the reconstruction. The surface layer and a faint strand of the bottom void is visible in the other reconstruction. The OE of the surface and void are also presented for this investigation in Fig. 12. The Figure shows that the OE grows almost linearly from the benchmark to the high background contrast variation case for the 20

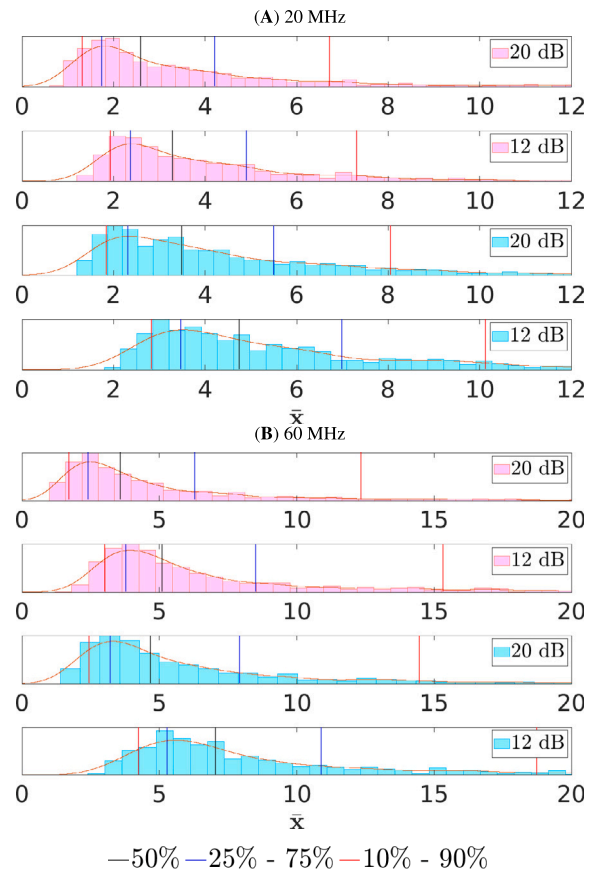


Fig. 11. The average permittivity perturbation distribution for the (A) 20 and (B) 60 MHz systems. The magenta histograms correspond to the dense configuration and the cyan histograms correspond to the sparse configuration, both having their noise levels indicated in their respective legends. The y-axis is the probability density function (PDF) of the estimated permittivity perturbation. This figure indicates that there is a larger median value in the higher frequency estimates and wider spread in the sparse configurations. The lower noise has smaller spread for both density configurations and systems. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

MHz system. In the 60 MHz system, the OE reaches a plateau at low background contrast variation case which either increases or decreases as it progresses to the high background contrast case depending on the quantile considered. The interquartile and interdecile range of the samples are shown in the cyan and magenta respectively. The spread in the void OE is wider for both the 20 and 60 MHz systems as compared to the surface OE, indicating that the uncertainty in reconstructing the void is higher than that of the surface.

4. Discussion

This article describes and evaluates numerically a filtering approach, which allows marginalising wavelength-induced random errors from a large full-waveform dataset. This approach enables obtaining a reconstruction of the surface layer and voids of a 160 m diameter non-convex two-dimensional asteroid model, via a monostatic measurement of a full-wavefield signal, having centre frequencies of 20 and 60 MHz and bandwidths of 10 and 20 MHz, respectively. Finding a reconstruction was feasible with 20- and 12-dB SNR, of which the former has been estimated as the lower bound of the CONCERT measurement of Kofman et al. (2015) and the latter is slightly (2 dB) above the 10 dB limit which has been suggested as the lower bound for finding a reasonable reconstruction for the interior structure of an asteroid model in the recent numerical simulations and laboratory experiments (Takala et al.,

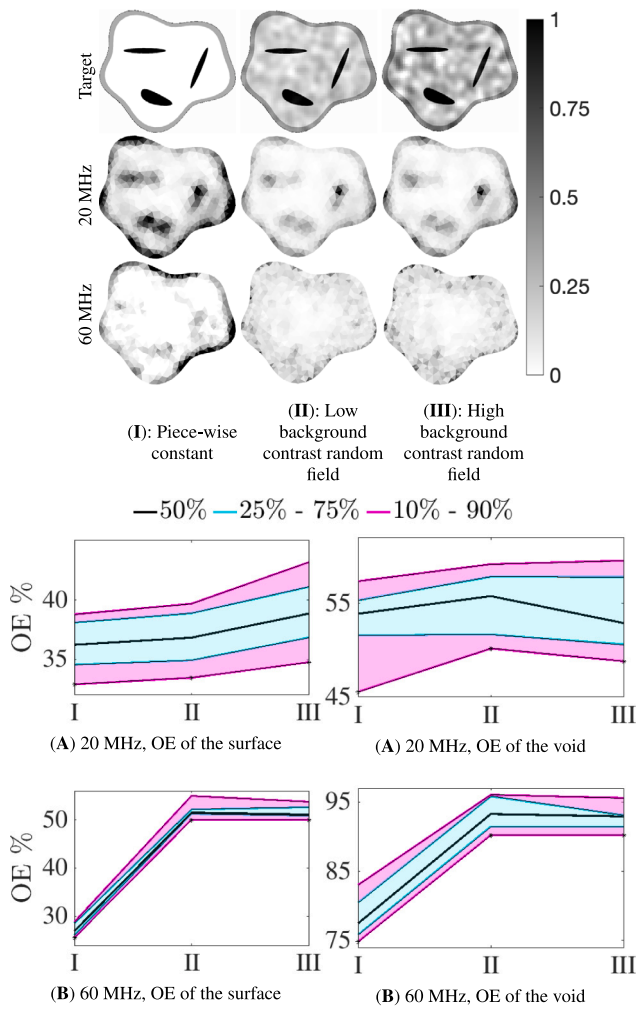


Fig. 12. Benchmark reconstructions compared to the random field model reconstructions for the 20 and 60 MHz systems. The 1st row shows the variation in the background permittivity contrasts for the respective models. The 2nd and 3rd rows present the reconstruction of the models for the 20 and 60 MHz systems, respectively. The surface and void OE for the different background models is shown in the 4th and 5th rows with their interquartile and interdecile range coloured cyan and magenta, respectively. This figure shows that the benchmark reconstruction outperforms the random field models which can be observed by visual inspection and the OE curves. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2018a; Sorsa et al., 2019; Eyraud et al., 2020; Sorsa et al., 2021b; Deng et al., 2021).

We demonstrate through numerical experiments on a two-dimensional domain that, when a centre frequency applicable in a space investigation is used in the numerical simulation of the full-wavefield together with quadrature amplitude modulation and demodulation of the simulated signal (Sorsa et al., 2021b), the reconstruction of the domain is improved i.e., lessening wavelength-induced uncertainty, when it is found via averaging several candidate reconstructions or decomposing the signals into meaningful and noisy parts. The present results show that the expected effect of phase errors is of significantly higher magnitude in comparison to the noise level, and that their effect on the final reconstruction can be reduced via averaging over a sample of candidate reconstructions obtained by a randomising signal configuration in the spatial and spectral domains. Comparing the results obtained with regular and randomised signal configurations, the latter ones demonstrate an overall improved reconstruction quality which we interpret as a consequence of a successfully created larger sample of candidates. Namely, the convergence of the

sample means as stated by the generalised central limit theorem (Liu, 2008) depends only on the sample size K as $\mathcal{O}(K^{-1/2})$, if it is appropriate otherwise. Here the frequency randomisation constituted an important way to extend the sample size, accounting as a surrogate for the uncertainties due to wave propagation and domain structure.

As for the dense and sparse configurations, the wavelength-induced uncertainties can be identified as the cause of the difference in the reconstructions obtained. The sparse configuration yields a higher contrast as compared to a dense one, since phase fluctuations lead to larger fluctuations in the reconstruction the greater the number of data point in the constellation set. Hence, aggregating small collections of phase fluctuations over a large number of constellations preserves the phase information as compared to aggregating large phase fluctuations over a small number of constellations. We can infer from Fig. 11 that obtaining the exact permittivity values from the inversion process is challenging task if the full-wave data is not augmented with some additional information such as the signal travel-time. The permittivity perturbation norm spreads in the sparse configuration case with the low noise level cases having shorter-tailed distribution compared to the high noise cases. The estimated permittivity perturbations are large, since the formalism in Eqs. (2), (9), and (19) show that we can only get an estimate \hat{x} for the permittivity perturbation (\hat{x}), and this estimate is largely dependent on the model and wavelength-induced uncertainties that cannot be totally marginalised. We have shown distribution of the estimated permittivity perturbation in Fig. 11. The challenge of estimating exact permittivity distributions and perturbation distribution i.e., minimising the errors in the estimate \hat{x} obtained in this current study is to be considered as a future research goal. Since we assume that the errors in the perturbation estimates obtained are not flat and cannot be fitted to the exact perturbation distribution as that would be an inverse crime, a more robust marginalisation method or supporting data would be needed to solve this task.

In this study, we assume that the averaged total variation regularised estimates in the marginalisation process are independent (enough) and identically distributed (IID) so that the convergence conditions of the central limit theorem or its generalisation to weakly correlated samples hold (Liu, 2008). This assumption sets a challenge for choosing the measurement point configuration and phase discrepancy appropriately. We covered two possible strategies for spatial point selection, otherwise known as statistical sparsity-based learning with a suitable prior model (Mirbeik et al., 2021). The sample size of the present randomised configuration was found to reduce the phase-dependent fluctuations by 2%–26% relative to the regular configuration in terms of the OE and RMSE measures, see Table 3. Further improvement of the current estimates might be obtained by applying a larger sample, relying on the asymptotic central limit theorem, which states that the accuracy of a sample-based estimate is inversely proportional to the square root of the sample size. Alternatively, as the IID condition is only assumed due to the missing *a priori* information on the wavelength-induced uncertainties the formation of the sample can be supervised by filtering out the possible outliers. In such a procedure, the principal component analysis (Lee et al., 2008), or machine learning and deep neural network (Geng et al., 2021; Li et al., 2019; Ji et al., 2021) might be utilised. In particular, a deeper investigation on the dependence of the wavelength-induced uncertainties on the frequency and receiver positions might be conducted to gain more *a priori* understanding of how the receiver positions could be distributed around the orbit, hence determining if the point selection approach can be improved.

From a potential mission design viewpoint, the results obtained suggest that setting the bandwidth of the modelled signal pulse based on the discrepancy principle and averaging the measurement point density following the Nyquist criterion is crucial w.r.t. the stability of the reconstruction process. If the number of averaging points is limited further from the current setup, the size of the reconstructed details grows, or they tend to get blurred as a natural consequence

of the lower measurement point density. Due to the importance of selecting the measurement point configuration suggested by the WRS test, it is also obvious that a sparse configuration alone will not allow obtaining an optimal reconstruction quality. However, sampling can be applied without having a point density matching with the Nyquist criterion w.r.t. the centre frequency, which can be the case of *in-situ* measurements. This is important if the global internal structure of an asteroid is to be explored during a space mission, to provide answers to the scientific questions on the internal composition asteroids and/or comets (Carr, 2012; Jutzi and Benz, 2017; Hérique et al., 2018).

The TSVD filter is based on excluding the less significant signal components out of the inversion process, which is known to improve the outcome of Ground-penetrating radar (GPR) reconstructions (Ludeno et al., 2020). The TSVD was found to have a regularising effect on the reconstruction, while it did not affect the major fluctuations due to the wavelength-induced uncertainties but rather increased the error values. As our concentration is on the monostatic data, i.e., the points of transmission and measurement coincide, similar to the future JuRa investigation, TSVD was considered as a way to decompose the signal. If more complex signal patterns e.g., bistatic or multistatic measurements will need to be processed as suggested in several previous studies (Sorsa et al., 2019), then more advanced decomposition approaches will be necessary. For example, the decomposition of the time reversal operator DORT (Décomposition de l'opérateur de retournement temporel) (Prada et al., 1996; Fink et al., 2000) finds the most 'reflective' multistatic patterns by finding the eigenvalues and eigenvectors of a symmetric (reciprocal) transfer matrix between the transmission and measurement points.

While our concentration is on reconstructing the real part of the relative permittivity structure, motivating real-valued discretisation of the forward and inverse problem (Sorsa et al., 2020), a complex formulation follows from the complex-valued QAM signal, where the in-phase and quadrature component compose the real and imaginary part, respectively. Inverting a complex-valued system was found to be advantageous to reduce any possible fluctuations due to wavelength-induced uncertainties. Namely, the wavelength-induced uncertainty might affect only one of the QAM signal components, in which case, a real-value Jacobian matrix might contain two opposite fluctuations, potentially cancelling each other, while a complex Jacobian treats both perturbations appropriately as two complex components with a positive absolute value, thereby, preserving the visibility of the scatterer causing the deviation.

The inversion process applied relies on the linearisation of the full wavefield w.r.t. the permittivity perturbation in a given position (Virieux and Operto, 2009). Hence, while the wave propagation is modelled as a nonlinear process, taking into account indirect signal paths including reflection and refraction and higher-order or multi-path scattering effects between the scattering obstacles and the structure of the target, the mutual coupling effects between the obstacles, are omitted in this study. Based on earlier results of the same model, the improvement obtained with a higher-order polynomial or Born approximation via a recursive use of the present linearisation (Sorsa et al., 2020) is assumed to be a minor one. In particular, as the higher-order approximation was shown to be sensitive to noise which cannot be avoided in the present application due to the wavelength-induced uncertainties. Notably, the numerical inaccuracies in the modelled material distribution, which gives rise to the requirement for the discrepancy condition, are always present regardless of the order of Born approximation incorporated in the model. However, another modelling approach for the nonlinear higher-order scattering effects might yield an improvement. Therefore, comparisons between different methods and solvers, e.g., ray-tracing techniques (Gassot et al., 2020; Ciarletti et al., 2015) and frequency-based methods (Eyraud et al., 2019, 2020) will be important.

We have chosen a two-dimensional reference domain for this study, as it allows examining the effects of wavelength-induced uncertainties

on numerical inversion efficiently without the need of high-performance cluster computations. In addition, our domain has a few features that makes it relevant from the mission design viewpoint, in particular, considering asteroids as potential targets. Firstly, our model includes a surface layer which is a potential structure in asteroid interiors as shown by impact simulation studies (Jutzi and Benz, 2017). As a surface layer is known to significantly hinder the detection of buried objects (Dogan et al., 2017), it can be considered as important from the practical applicability aspect. Many studies, however, omit this aspect, e.g., Deng et al. (2021), thus, potentially somewhat exaggerating the detectability of the deep interior structures. In addition to the surface layer, another important modelling feature ensuring the relevance of our model is the realistic value of the real background relative permittivity which, based on the assumed mineral composition of the target (Hérique et al., 2002; Hérique et al., 2018), was set to be 4. A significantly lower value 1.6 has been used in Haynes et al. (2021), hence, reducing the complexity of the wave propagation inside the target as compared to our present study. Furthermore, the extension of the piece-wise constant domain to a random field domain gives an indication of how the complexity of the wavelength-induced uncertainties grows. In essence, the high frequency (60 MHz) is more sensitive to all kinds of uncertainty which in turn, reduces the reconstruction quality of the target domain. The experimental relevance of the present structural model composition can be motivated by the recent successful laboratory experiments with 3D-printed analogue targets (Sorsa et al., 2021a,b; Eyraud et al., 2020) with a comparable structure compared to the present case.

The next step will be to investigate the present methodology with a three-dimensional analogue object and with experimental scattering data similar to Sorsa et al. (2021b,a) and Eyraud et al. (2020). In such a process, the outcome of this study to understand combining point cloud data into a single reconstruction is important. The current two-dimensional implementation can, in principle, be directly extended to the three-dimensional laboratory coordinates as done in the previous analogue studies. This study also allows for further numerical investigation of the uncertainties relating to the wavelength-induced uncertainties, which is essential due to the significance of their magnitude based on the current results.

Acknowledgement

This study was carried out as a part of the Academy of Finland project, ICT 2023 (FETD-Based Tomographic Full-Wave Radar Imaging of Small Solar System Body Interiors; project number 336151) and was supported by the Academy of Finland Centre of Excellence in Inverse Modelling and Imaging, 2018–2025. Y.O. Yusuf was also supported by the Magnus Ehrnrooth Foundation, Finland through the graduate student scholarship.

References

- Abubakar, A., Habashy, T.M., 2005. A green function formulation of the extended born approximation for three-dimensional electromagnetic modelling. *Wave Motion* 41 (3), 211–227.
- Bambach, P., Deller, J., Vilenius, E., Pursiainen, S., Takala, M., Braun, H.M., Lentz, H., Wittig, M., 2018. DISCUS—The deep interior scanning CubeSat mission to a rubble pile near-earth asteroid. *Adv. Space Res.* 62 (12), 3357–3368.
- Barriot, J.-P., Kofman, W., Hérique, A., Leblanc, S., Portal, A., 1999. A two dimensional simulation of the CONSERT experiment (Radio tomography of comet Wirtanen). *Adv. Space Res.* 24 (9), 1127–1138.
- Binzel, R.P., Kofman, W., 2005. Internal structure of near-earth objects. *C. R. Phys.* 6 (3), 321–326.
- Bousquet, M., 2017. Satellite communications and space telecommunication frequencies. In: *Handbook of Satellite Applications*. Springer International Publishing, Cham, pp. 325–357. http://dx.doi.org/10.1007/978-3-319-23386-4_13.
- Brownlee, D., Joswiak, D., Matrajt, G., 2012. Overview of the rocky component of wild 2 comet samples: Insight into the early solar system, relationship with meteoritic materials and the differences between comets and asteroids. *Meteorit. Planet. Sci.* 47 (4), 453–470.

- Bucci, O., Franceschetti, G., 1987. On the spatial bandwidth of scattered fields. *IEEE Trans. Antennas Propag.* 35.
- Candès, E.J., Romberg, J., Tao, T., 2006. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory* 52 (2), 489–509.
- Candès, E.J., Romberg, J.K., Tao, T., 2006. Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.: J. Issued Courant Inst. Math. Sci.* 59 (8), 1207–1223.
- Carlsten, S., Johansson, S., Wörman, A., 1995. Radar techniques for indicating internal erosion in embankment dams. *J. Appl. Geophys.* 33 (1–3), 143–156.
- Carry, B., 2012. Density of asteroids. *Planet. Space Sci.* 73 (1), 98–118.
- CDF, 2018. Study report. In: *SPP NEO Inactive Body*, Vol. CDF-A(178). pp. 15–19, January.
- Chaparro, L., Akan, A., 2018. *Signals and Systems using MATLAB*. Elsevier Science, URL: <https://books.google.fi/books?id=Unh1DwAAQBAJ>.
- Chew, W.C., Wang, Y.-M., 1990. Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method. *IEEE Trans. Med. Imaging* 9 (2), 218–225.
- Ciarletti, V., Levasseur-Regourd, A.C., Lasue, J., Statz, C., Plettemeier, D., Hérique, A., Rogez, Y., Kofman, W., 2015. CONSERT Suggests a change in local properties of 67P/Churyumov-Gerasimenko's nucleus at depth. *Astron. Astrophys.* 583, A40. <http://dx.doi.org/10.1051/0004-6361/201526337>, URL: <https://doi.org/10.1051/0004-6361/201526337>.
- Daniels, D.J., 2004. *Ground Penetrating Radar*, Vol. 1. Iet.
- Deng, J., Kofman, W., Zhu, P., Hérique, A., Liu, R., Zheng, S., 2021. EI+ FWI Method for reconstructing interior structure of asteroid using lander-to-orbiter bistatic radar system. *IEEE Trans. Geosci. Remote Sens.*
- Dogan, M., Gumus, S., Turhan-Sayan, G., 2017. Detection of conducting and dielectric objects buried under a layer of asphalt or concrete using simulated ground penetrating radar signals. In: *2017 International Conference on Electromagnetics in Advanced Applications (ICEAA)*. IEEE, pp. 1535–1538.
- Dong, S., Han, L., Zhang, P., Feng, Q., Yin, Y., 2020. Convolution coding and amplitude attenuation-based full waveform inversion. *IEEE Access* 8, 182996–183013.
- Egiazarian, K., Foi, A., Katkovnik, V., 2007. Compressed sensing image reconstruction via recursive spatially adaptive filtering. In: *2007 IEEE International Conference on Image Processing*, Vol. 1. IEEE, pp. 1–549.
- Engl, H.W., Hanke, M., Neubauer, A., 1996. *Regularization of Inverse Problems*, Vol. 375. Springer Science & Business Media.
- Ernst, J.R., Green, A.G., Maurer, H., Holliger, K., 2007. Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data. *Geophysics* 72 (5), J53–J64.
- Eyraud, C., Saleh, H., Geffrin, J.-M., 2019. Influence of the description of the scattering matrix on permittivity reconstruction with a quantitative imaging procedure: polarization effects. *J. Opt. Soc. Amer. A* 36, 234–244.
- Eyraud, C., Sorsa, L.-I., Geffrin, J.-M., Takala, M., Henry, G., Pursiainen, S., 2020. Full wavefield simulation versus measurement of microwave scattering by a complex 3D-printed asteroid analogue. *Astron. Astrophys.* 643, A68.
- Finetti, B., 1972. *Probability, Induction and Statistics*. John Wiley and Sons.
- Fink, M., Cassereau, D., Derode, A., Prada, C., Roux, P., Tanter, M., Thomas, J.-L., Wu, F., 2000. Time-reversed acoustics. *Rep. Progr. Phys.* 63 (12), 1933.
- Gao, G., Torres-Verdín, C., 2006. High-order generalized extended Born approximation for electromagnetic scattering. *IEEE Trans. Antennas and Propagation* 54 (4), 1243–1256.
- Gassot, O., Hérique, A., Rogez, Y., Kofman, W., Zine, S., Ludimbulu, P.-P., 2020. SPRATS: A versatile simulation and processing RADAR tools for planetary missions. In: *2020 IEEE Radar Conference (RadarConf20)*. IEEE, pp. 1–5.
- Geng, Z., Yan, H., Zhang, J., Zhu, D., 2021. Deep-learning for radar: A survey. *IEEE Access* 9, 141800–141818.
- Gibbons, J.D., Chakraborti, S., 2014. *Nonparametric Statistical Inference: Revised and Expanded*. CRC Press.
- Golub, G.H., Van Loan, C.F., 2013. *Matrix Computations*, Vol. 3. JHU Press.
- Haynes, M., Fenni, I., Gim, Y., Davidsson, B., Raymond, C., Hernandez, S., Bandyopadhyay, S., Stuart, J., Karimi, R., 2021. Small body radar inverse scattering in monostatic and bistatic geometries. In: *Lunar and Planetary Science Conference*, Vol. 2548. p. 1281.
- Hérique, A., Agnus, B., Asphaug, E., Barucci, A., Beck, P., Bellerose, J., Biele, J., Bonal, L., Bousquet, P., Bruzzone, L., Buck, C., Carnelli, I., Cheng, A., Ciarletti, V., Delbo, M., Du, J., Du, X., Eyraud, C., Fa, W., Fernandez, J.G., Gassot, O., Granados-Alfaro, R., Green, S., Grieger, B., Grundmann, J., Grygorczuk, J., Hahnel, R., Heggy, E., Ho, T.-M., Karatekin, O., Kasaba, Y., Kobayashi, T., Kofman, W., Krause, C., Kumamoto, A., Küppers, M., Laabs, M., Lange, C., Lasue, J., Levasseur-Regourd, A., Mallet, P., Michel, P., Mottola, S., Murdoch, N., Mütze, M., Oberst, J., Orosi, R., Plettemeier, D., Rochat, S., RodriguezSuquet, R., Rogez, Y., Schaffer, P., Snodgrass, C., Souyris, J.-C., Tokarz, M., Ulamec, S., Wahlund, J.-E., Zine, S., 2018. Direct observations of asteroid interior and regolith structure: Science measurement requirements. *Adv. Space Res.* 62, 2141–2162.
- Hérique, A., Gilchrist, J., Kofman, W., Klinger, J., 2002. Dielectric properties of comet analog refractory materials. *Planet. Space Sci.* 50 (9), 857–863.
- Hérique, A., Plettemeier, D., Goldberg, H., Kofman, W., 2020a. JuRa: The juvenas radar on hera to fathom didymoon. In: *European Planetary Science Congress*. pp. EPSC2020–595.
- Hérique, A., Plettemeier, D., Kofman, W., Rogez, Y., Buck, C., Goldberg, H., 2019. A low frequency radar to fathom asteroids from juvenas cubesat on HERA. In: *Proc. EPSC-DPS Joint Meeting*. In: *EPSC-DPS Joint Meeting 2019*, vol. 13, Geneva, Switzerland, p. 807.
- Hérique, A., Plettemeier, D., Kofman, W., Rogez, Y., Goldberg, H., 2020b. Low frequency radar (LFR) on the JUVENTAS CubeSat for HERA/ESA mission. In: *EGU General Assembly Conference Abstracts*. p. 19957.
- Ji, Y., Zhang, F., Wang, J., Wang, Z., Jiang, P., Liu, H., Sui, Q., 2021. Deep neural network-based permittivity inversions for ground penetrating radar data. *IEEE Sens. J.* 21 (6), 8172–8183.
- Jutzi, M., Benz, W., 2017. Formation of bi-lobed shapes by sub-catastrophic collisions-A late origin of comet 67P's structure. *Astron. Astrophys.* 597, A62.
- Kaipio, J., Somersalo, E., 2006. *Statistical and Computational Inverse Problems*, Vol. 160. Springer Science & Business Media.
- Khare, K., George, N., 2002. Direct sampling and demodulation of carrier-frequency signals. *Opt. Commun.* 211 (1–6), 85–94.
- Kofman, W., 2012. Radar techniques to study subsurfaces and interiors of the solar system objects. In: *2012 19th International Conference on Microwaves, Radar Wireless Communications*, Vol. 2. pp. 409–412. <http://dx.doi.org/10.1109/MIKON.2012.6233605>.
- Kofman, W., Hérique, A., Barbin, Y., Barriot, J.-P., Ciarletti, V., Clifford, S., Edenhofer, P., Elachi, C., Eyraud, C., Goutail, J.-P., Heggy, E., Jorda, L., Lasue, J., Levasseur-Regourd, A.-C., Nielsen, E., Pasquero, P., Preusker, F., Puget, P., Plettemeier, D., Rogez, Y., Sierks, H., Statz, C., Svedhem, H., Williams, I., Zine, S., Van Zyl, J., 2015. Properties of the 67p/Churyumov-Gerasimenko interior revealed by CONSERT radar. *Science* 349 (6247), aab0639–1 – aab0639–6.
- Kofman, W., Hérique, A., Goutail, J.-P., Hagfors, T., Williams, I., Nielsen, E., Barriot, J.-P., Barbin, Y., Elachi, C., Edenhofer, P., Levasseur-Regourd, A.-C., Plettemeier, D., Picardi, G., Seu, R., Svedhem, V., 2007. The comet nucleus sounding experiment by radiowave transmission (CONSERT): A short description of the instrument and of the commissioning stages. *Space Sci. Rev.* 128 (413–432).
- Lee, K.-C., Ou, J.-S., Fang, M.-C., 2008. Application of SVD noise-reduction technique to PCA based radar target recognition. *Prog. Electromagn. Res.* 81, 447–459.
- Li, X., He, Y., Jing, X., 2019. A survey of deep learning-based human activity recognition in radar. *Remote Sens.* 11 (9), 1068.
- Liu, J.S., 2008. *Monte Carlo Strategies in Scientific Computing*. Springer Science & Business Media.
- Liu, X., Liu, S., Fu, L., 2018. Full waveform inversion of cross-hole radar data using envelope objective function. In: *2018 17th International Conference on Ground Penetrating Radar (GPR)*. IEEE, pp. 1–5.
- Ludeno, G., Gennarelli, G., Lambot, S., Soldovieri, F., Catapano, I., 2020. A comparison of linear inverse scattering models for contactless GPR imaging. *IEEE Trans. Geosci. Remote Sens.* 58 (10), 7305–7316.
- Luisier, F., Rodríguez, P., 2013. Total variation regularization algorithms for images corrupted with different noise models: A review. *J. Electr. Comput. Eng.* 2013, 217021. <http://dx.doi.org/10.1155/2013/217021>.
- Mirbeik, A., Lawrence, V., LaPeruta, T.A., Popescu, C.M., Hohli, M.E., Tavassolian, N., 2021. Statistical sparsity-based learning for ultra-wideband radar signal reconstruction. In: *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III*, Vol. 11746. International Society for Optics and Photonics, 1174619.
- Orfanidis, S.J., 1995. *Introduction to Signal Processing*. Prentice-Hall, Inc..
- Prada, C., Manneville, S., Spoliensky, D., Fink, M., 1996. Decomposition of the time reversal operator: Detection and selective focusing on two scatterers. *J. Acoust. Soc. Am.* 99 (4), 2067–2076.
- Pursiainen, S., Kaasalainen, M., 2016. Orbiter-to-orbiter tomography: a potential approach for small solar system bodies. *IEEE Trans. Aerosp. Electron. Syst.* 52 (6), 2747–2759.
- Semenov, S.Y., Bulyshev, A.E., Abubakar, A., Posukh, V.G., Sizov, Y.E., Souvorov, A.E., Van Den Berg, P.M., Williams, T.C., 2005. Microwave-tomographic imaging of the high dielectric-contrast objects using different image-reconstruction approaches. *IEEE Trans. Microw. Theory Tech.* 53 (7), 2284–2294.
- Sorsa, L.-I., Eyraud, C., Hérique, A., Takala, M., Pursiainen, S., Geffrin, J.-M., 2021a. Complex-structured 3D-printed wireframes as asteroid analogues for tomographic microwave radar measurements. *Mater. Des.* 198, 109364. <http://dx.doi.org/10.1016/j.matdes.2020.109364>.
- Sorsa, L.-I., Pursiainen, S., Eyraud, C., 2021b. Analysis of full microwave propagation and backpropagation for a complex asteroid analogue via single-point quasi-monostatic data. *Astron. Astrophys.* 645, A73. <http://dx.doi.org/10.1051/0004-6361/202039380>.
- Sorsa, L.-I., Takala, M., Bambach, P., Deller, J., Vilenius, E., Pursiainen, S., 2019. Bistatic full-wave radar tomography detects deep interior voids, cracks and boulders in a rubble-pile asteroid model. *Astrophys. J.* 872 (1), 44.
- Sorsa, L.-I., Takala, M., Eyraud, C., Pursiainen, S., 2020. A time-domain multigrid solver with higher-order Born approximation for full-wave radar tomography of a complex-shaped target. *IEEE Trans. Comput. Imaging* 6, 579–590.
- Takala, M., Bambach, P., Deller, J., Vilenius, E., Wittig, M., Lentz, H., Braun, H., Kaasalainen, M., Pursiainen, S., 2018a. Far-field inversion for the deep interior scanning cubesat. *IEEE Trans. Aerosp. Electron. Syst.*

- Takala, M., Us, D., Pursiainen, S., 2018b. Multigrid-based inversion for volumetric radar imaging with asteroid interior reconstruction as a potential application. *IEEE Trans. Comput. Imaging* 4 (2), 228–240.
- Tejero, C.E.J., Sallares, V., Ranero, C.R., 2018. Appraisal of instantaneous phase-based functions in adjoint waveform inversion. *IEEE Trans. Geosci. Remote Sens.* 56 (9), 5185–5197.
- Ulaby, F.T., Long, D.G., Blackwell, W.J., Elachi, C., Fung, A.K., Ruf, C., Sarabandi, K., Zebker, H.A., Van Zyl, J., 2014. *Microwave Radar and Radiometric Remote Sensing*, Vol. 4. University of Michigan Press, Ann Arbor, MI, USA.
- Van Hove, L., 1954. Correlations in space and time and Born approximation scattering in systems of interacting particles. *Phys. Rev.* 95 (1), 249.
- Virieux, J., Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74 (6), WCC1–WCC26.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Wiskin, J., Malik, B., Borup, D., Pirshafiey, N., Klock, J., 2020. Full wave 3D inverse scattering transmission ultrasound tomography in the presence of high contrast. *Sci. Rep.* 10 (1), 1–14.