Near-Optimal Selection of GPR Observations for Linear Inversion

Jianping Wang and Alexander Yarovoy
Microwave Sensing, Signals and Systems
Delft University of Technology, Delft, The Netherlands
Email: J.Wang-4@tudelft.nl; A.Yarovoy@tudelft.nl

***Abstract*—Dense spatial measurement of Ground Penetrating Radar (GPR) guarantees the sampling criterion for the linear inversion. On the other hand, the oversampled data significantly increases the computational load in the inversion operation, or it requires more complex GPR systems. In this paper, the "frame potential" based approach was suggested to near-optimally select GPR measurements for linear inversion. As it is mainly based on the observation matrix, the suggested approach can be used to select the most informative observations before linear inversion or optimize the GPR spatial sampling criterion. As an example, an antenna array topology optimization problem is presented to show the effectiveness of this approach.**

***Index Terms*—Ground penetrating radar (GPR), Sensor selection, sampling design, linear inversion.**

I. Introduction

Ground penetrating radar (GPR) has been widely used for landmine detection, subsurface human utility survey, archaeological investigation, civil engineering studies, etc.. The range of GPR applications is still expanding. One of the recent examples is to integrate GPR system with the tunnel boring machine (TBM) to predict the soil properties and geological structures in front of its cutter-head [1], [2].

In contrary to the regular applications where the GPR systems take measurements on the ground surface over a (quasi-) rectangular grid, the GPR systems used for TBMs are embedded in the cutter-head and GPR measurements are collected during the rotation of the cutter-head. Therefore, in this case, the GPR signals are typically acquired over a polar grid. Although we have derived the polar sampling criteria based on the monochromatic assumption [3], the sampling intervals are too dense to be achieved for practical GPR systems. More specifically, the small sampling intervals in the radial direction require numerous antennas to be used for signal acquisition. However, due to the harsh working environment and mechanical constraints, a very limited space is available for deploying a substantially smaller number of antennas. So dense sampling in the radial direction is infeasible. Fortunately, GPR systems generally operate with wideband/ultrawideband (UWB) signals. According to the UWB technique [4], it is possible to take fewer spatial samples than that derived based on the monochromatic theory. Although UWB array (sampling) design was discussed in [4], the suggested methods are still not straightforwardly applicable to optimize the UWB sampling over a polar grid. Actually, even if the GPR signals can be collected in space according to the sampling intervals derived with the monochromatic theory, the samples close to the center are typically much denser than that closed to the edge of the aperture. So many samples in the central region of the aperture are redundant and they do not provide independent information for estimating the scattering properties of the targets. Thus the problem to effectively eliminate the redundant samples to reduce the computational load in the inversion processing becomes essential.

To address this problem, we formulate the UWB spatial sampling as an observation/sensor selection problem in this paper. By formulating the GPR imaging as a linear inversion problem, the UWB sampling aims to acquire the least number of samples but still provides robust estimation to the scattering properties of the subsurface objects. So from the observation/sensor selection perspective, the UWB sampling problem is to select the most informative samples for a linear inversion problem. In the open literature, the observation/sensor selection problems have been studied for magnetic resonance imaging (MRI)/ magnetic resonance spectroscopic imaging (MRSI) observation selection [5], network sensor selection [6]–[8], radar network topology optimization [9], *etc.*. Two classes of algorithms are proposed: greedy methods and convex relaxed methods. In [10], two sequential selection methods, i.e., sequential backward selection (SBS) and sequential forward selection (SFS), were proposed for observation selection. In [8], a FrameSense greedy method was presented for sensor selection. On the other hand, Joshi *etal.* formulated the sensor selection problem as a convex optimization problem with some relaxation scheme and then convex optimization solvers were used to select the optimal sensors with certain optimality criteria. Both FrameSense approach and convex optimization method have been utilized to optimize the radar network topology. Most of these methods are tackled the problems with the single measurement per sensor. Although selecting sensors with vector measurement was discussed in [6], the convex optimization based method is very computationally expensive, especially for a large number of observation selection which is typical for imaging problem. In this paper, the FrameSense method is extended and applied to GPR spatial sampling.

The rest of the paper is organized as follows. In section II, the linear inversion problem for GPR data processing and the brief description of frame potential are presented. Following that, one example for UWB GPR array optimization is shown to demonstrate the extended method. Finally, some conclusions

are drawn in section IV.

II. Formulation

A. Linear inversion problem

Based on the Born approximation, the GPR signal scattering process can be formulated as a linear system. It is represented as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^N$ is the vector of GPR measurements, $\mathbf{x} \in \mathbb{C}^n$ is the vector of reflectivity functions of scatterers, and $\mathbf{n} \in \mathbb{C}^n$ denotes the noise and measurement errors. $\mathbf{A} \in \mathbb{C}^{N\times n}$ is the observation matrix and each row is an observation vector corresponding to each measurement. The entries of matrix $\mathbf{A}$ represent the propagations of electromagnetic waves from the transmitting antennas to the scatterers and then to the receiving antennas, which can be written as

$$\begin{aligned} &A_{ij}(\mathbf{R}^t, \mathbf{R}^r; \omega) \\ &\quad = \int_V G_T(\mathbf{R}^t, \mathbf{R}^s; \omega) W(\mathbf{R}^s) G_R(\mathbf{R}^s, \mathbf{R}^r; \omega) s(\omega) dV, \\ &\qquad\qquad i = 1, \cdots, N; \quad j = 1, \cdots, n. \end{aligned} \tag{2}$$

where A_{ij} is the entry of matrix $\mathbf{A}$ at i-th row and j-th column. ω is the frequency of the signal and $\mathbf{R}^t$, $\mathbf{R}^r$ and $\mathbf{R}^s$ are the positions of transmitting antenna, receiving antenna and scatterers, respectively. $W(\mathbf{R}^s)$ is a box window centered at the scatterer's position $\mathbf{R}^s$. $G_T(\cdot)$ and $G_R(\cdot)$ are the Green's functions that represent the propagations of electromagnetic waves from transmitting antenna to the scatterer and that from the scatterer to the receiving antenna. $s(\omega)$ is the radiated wavelet.

In (1), N field measurements are needed to fully reconstruct the scattering properties of the targets, which may be determined by the Nyquist sampling criterion based on the resolution requirement. However, sometimes to acquire N measurements of the physical field is impossible or practically infeasible as we indicated in the section I. Assume only L samples can be acquired and the N measurements are the potential samples. Then to acquire L samples for practical system can be converted to select L samples from the N potential ones. Denote the L selected samples as $\mathcal{L} = \{s_1, s_2, \cdots, s_L\}$ and the N potential samples as $\mathcal{N} = \{1, 2, \cdots, N\}$.

Assume the observation vectors associated with the L selected samples $\mathbf{y}_\mathcal{L}$ form an obervation matrix $\mathbf{A}_\mathcal{L}$, then the linear model in (1) can be rewritten as

$$\mathbf{y}_\mathcal{L} = \mathbf{A}_\mathcal{L}\mathbf{x} + \mathbf{n}_L \tag{3}$$

where $\mathbf{y}_\mathcal{L} \in \mathbb{C}^L$ and $\mathbf{A}_\mathcal{L} \in \mathbb{C}^{L\times n}$. $\mathbf{n}_L \in \mathbb{C}^L$ represents the measurement errors and noise but the reduced dimension $L > n$, and it is assumed to be zero-mean Gaussian distribution with the variance σ^2.

Considering the set of measurements in (3), a minimum variance unbiased estimation of $\mathbf{x}$ can be obtained via least squares

$$\hat{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{y}_\mathcal{L} \tag{4}$$

where $\mathbf{A}^\dagger = (\mathbf{A}_\mathcal{L}^H \mathbf{A}_\mathcal{L})^{-1} \mathbf{A}_\mathcal{L}^H$ is the pseudo inverse of $\mathbf{A}_\mathcal{L}$. The mean square error (MSE) of $\hat{\mathbf{x}}$ is

$$\mathrm{MSE}(\hat{\mathbf{x}}) = \|\hat{\mathbf{x}} - \mathbf{x}\|^2 = \sigma^2 \sum_{k=1}^{n} \frac{1}{\lambda_k} \tag{5}$$

where λ_k is the k-th eigenvalue of the matrix $\mathbf{\Psi} = \mathbf{A}_\mathcal{L}^H \mathbf{A}_\mathcal{L}$.

As for imaging application, one antenna typically takes more than one, denoted as D, measurements through both spatial and frequency diversity. In the TBM case, the spatial diversity is implemented by the rotation and the frequency diversity by the UWB signal. Assume there are M spatial sampling positions, denoted as $\mathcal{M} = [1, 2, \cdots, M]$, and the associated observation matrices are $\{\Psi_1^\mathsf{T}, \Psi_2^\mathsf{T}, \cdots, \Psi_M^\mathsf{T}\}$, where each $\Psi_i^\mathsf{T} = [\psi_{i1}, \psi_{i2}, \cdots, \psi_{iD}]^\mathsf{T} \in \mathbb{C}^{D\times n}$. Therefore, the spatial sampling optimization problem can be formulated as the following sensor selection problem.

Problem 1: Giving the sensing matrix $\mathbf{A} = [\Psi_1, \Psi_2, \cdots, \Psi_M]^\mathsf{T} \in \mathbb{C}^{(M\cdot D)\times n}$, select K submatrices of $\mathbf{A}$ indexed by $u_1, u_2, \cdots, u_K$ in $\mathcal{M}$ to construct an observation matrix $\mathbf{A}_K = [\Psi_{u_1}, \Psi_{u_2}, \cdots, \Psi_{u_K}]^\mathsf{T} \in \mathbb{C}^{(K\cdot D)\times n}$, such that the estimation error is small enough and the number of selected submatrices is minimized.

B. Frame Potential

Frames are more general concept than bases. In frame theory, frames are the over-complete set of vectors that yield expansions that are not necessarily unique for a vector in the finite dimensional space. So a frame must have at least as many vectors as the number of the dimensions of the space.

Frame potential (FP) is a scalar property of the frame and it describes the orthogonality of the frames. For a matrix $\Psi_\mathcal{L} = [\psi_1, \cdots, \psi_L]^\mathsf{T}$, its frame potential is defined as

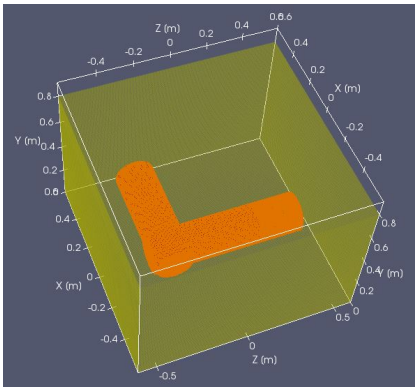
$$\mathrm{FP}(\Psi_\mathcal{L}) = \sum_{i,j\in\mathcal{L}} |\langle \psi_i, \psi_j \rangle|^2 \tag{6}$$

where $\psi_i \in \mathbb{C}^n$ is the i-th row of $\Psi_\mathcal{L}$. $\langle\cdot\rangle$ and $|\cdot|$ represent inner product and 2-norm operators. If the rows of $\Psi_\mathcal{L}$ form a unit-norm tight frame that minimizes the FP, the row vectors of $\Psi_\mathcal{L}$ are as close to orthogonal as possible. According to the Frame theory, the matrix $\Psi_\mathcal{L}$ achieves the minimum MSE in component wise. Actually, as the FP is defined based on the inner production of the observation vectors, minimizing the FP can be regarded as, roughly speaking, to minimize the correlation among the observation vectors.

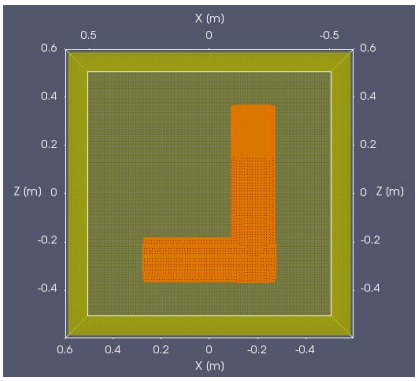
By minimizing the FP of the observation matrix, the FrameSense has been used for sensor selection [8]. Here we apply the FrameSense to optimize spatial sampling positions/sensor with vector measurements (instead of the case single sensor for single measurement). Based on the FrameSense algorithm, each time the spatial sampling position that maximally increases the FP of the observation matrix is removed. The cost function is written as [8]

$$F(S) = \mathrm{FP}(\mathbf{A}) - \mathrm{FP}(\mathbf{A}_{\mathcal{M}\setminus S}) \tag{7}$$

where S is the removed spatial sampling position. Note here eliminating one sampling position indicates to remove D

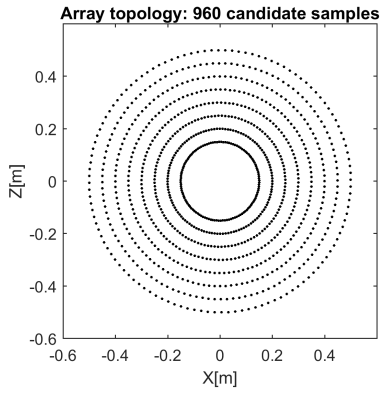


(a)

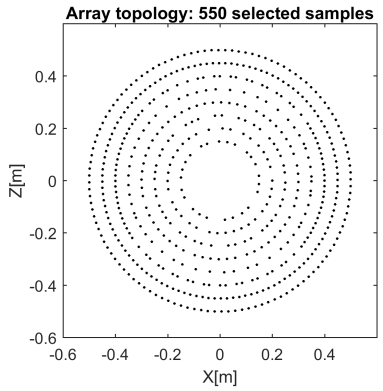


(b)

Fig. 1. Experimental setup for the numerical simulation. (a) 3-D view, (b) top view.



(a)



(b)

Fig. 2. Array topologies of the GPR measurements with: (a) 960 candidate samples, (b) 550 selected samples.

corresponding rows in the observation matrix Ψ. So we name it as clustered FrameSense. Its operation steps are the same as for the traditional FrameSense method in [8].

As the cost function in (7) is evaluated after removing D rows, i.e., a small matrix at a time, the computational cost is reduced compared to the traditional FrameSense method where one observation vector is removed at each time. The computational complexity for the clustered FrameSense can be analyzed as follows. As initially we have an observation matrix $\mathbf{A} \in \mathbb{C}^{(M\times D)\times n}$ associated with the M samples. To select K samples from the M candidates, then $M-K$ samples should be removed through the "worst-out" strategy of the FrameSense. To determine the k-th samples to be removed, the computational cost is $O(2n[(M-k+1)D]^2)$. So the total computational complexity can be estimated as $O\left(2n\sum_{k=1}^{M-K}[(M-k+1)D]^2\right) = O\left(\frac{2}{3}n(MP^2-K^3D^2)\right)$, where $P = M \cdot D$.

III. Example: GPR Spatial Sample Selection

In this section, a numerical simulation was performed to optimize a UWB circular array sampling for GPR imaging.

GPR data were synthesized with gprMax software, which utilizes the Finite Difference Time domain solver (FDTD) [11]. The geometrical configuration is shown in Fig. 1. In the simulation, two dielectric cylinders of radius 0.1 m were used as targets. Their lengths were 0.6m and 0.8m, respectively. The two cylinders were perpendicularly joined together at one end. The joint "L" shape structure was buried in the homogeneous background soil. The relative permittivity and the conductivity of the cylinders were 5.0 and 0.05 S/m, respectively. The permittivity of the background soil was 9.0 and the conductivity was 0.01 S/m. The GPR antennas collected the EM signals over a circular aperture on the ground surface. Within the circular aperture, the GPR signals were measured on a series of concentric circles with the radius ranging from 0.15 to 0.5m. At each spatial sampling position, the antenna axis was parallel to the radius. The Ricker wavelet of 900 MHz was used as the exciting signal.

At first, according to the polar sampling criteria [3], the spatial samples should be measured with the intervals of 5 cm in the radial direction and every 3° in the azimuthal direction, which form the set of candidate spatial samples. The array topology synthesized by all the candidate samples is shown in Fig. 2 (a). In practical systems, this circular aperture can be implemented by rotating a linear array with eight GPR antennas around the center. One can see that the samples close to the center of the aperture are much denser than that on the edge. So great redundancies exist in the central region.

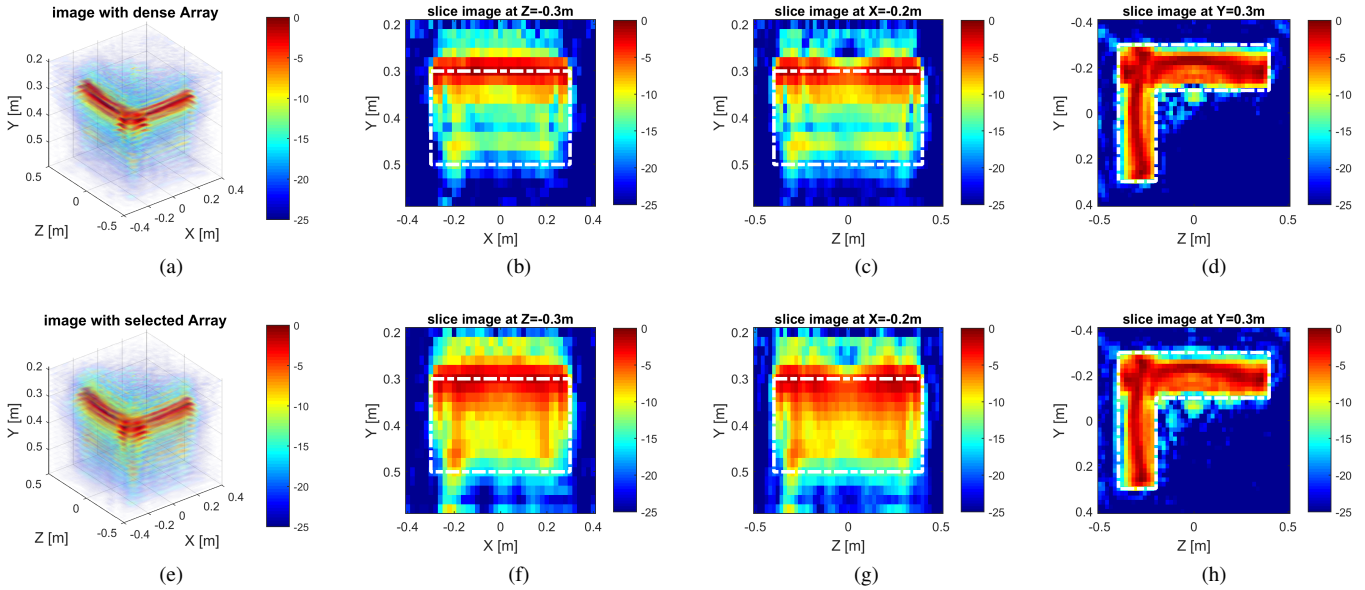


Fig. 3. Reconstructed images with two arrays. (a) is the 3-D image obtained with the dense array; (b), (c) and (d) are its three intersection slices at the objects. (e) is the 3-D image obtained with the selected array; (f), (g) and (h) are its three intersection slices at the objects.

To reduce the redundant samples in the central region by taking advantage of the clustered FrameSense, the imaging volume should be discretized and then a linear system of equations is built for image formation. Here a spatial volume of dimensions 0.8m$\times$0.4m$\times$1m in the x, y and z directions was considered as the region of interest. It was discretized as voxels with sizes of 2cm $\times$2cm$\times$2cm. So the region of interest contains 43911 voxels in total. As linear inversion based image formation is typically performed in the frequency domain, the signals were also discretized in the frequency domain. Considering the effective bandwidth of the Ricker wavelet, 79 discrete frequency samples were taken. So 79 measurements were acquired at each spatial sampling position. Therefore, a candidate observation matrix $\mathbf{A} \in \mathbb{C}^{(960\times 79)\times 43911}$ was built. The entries of $\mathbf{A}$ were computed via accurate Green's functions for rotated antennas [12].

Taking advantage of the clustered FrameSense, 550 spatial samples were selected and their distribution within the aperture is presented in Fig. 2 (b). It is clear that tremendously redundant samples at the central area were removed, which significantly reduces the data acquisition(about 43% in this example) as well as the computation cost for the inversion processing. Here it has to mention that 550 spatial samples were selected with slight redundancy compared to the number of voxels (i.e., the number of unknowns to be estimated). In principle, the number of spatial samples can be further reduced if some prior information is available.

To evaluate the imaging performance of the array selected with clustered FrameSense, least squares estimation was performed for imaging formation in which the biconjugate gradient stabilized method (BiCGSTAB) was used for computing matrix inverse. The 3-D images as well as their slices at objects reconstructed with the arrays of all candidate spatial samples and the selected samples are presented in Fig. 3, where the profile boundaries of objects are indicated with white dash-doted lines. In both cases, the shapes of the dielectric cylinders are well reconstructed (see 3-D images in Fig. 3 (a) and (e), and their slices in Fig. 3 (d) and (h)). So it verifies that the removed spatial samples in the selected arrays are redundant which bring no extra information for image formation. On the other hand, one may notice that the artifacts/sidelobes in Fig. 3 (e) are slightly larger than that in Fig. 3(a), which is more clearly observed by comparing the slice images Fig. 3 (b), (c) with (f) and (g). This is due to the reduced redundancy in the measured data. Furthermore, one can see that the reduced redundancy in the data has larger influence in the artifacts/sidelobe levels in the down-range direction than in the cross-range. For example, the artifacts/side lobe levels in the area below the dash-doted box in Fig. 3 (c) decrease rapidly to be lower than 20 dB while the side lobe levels in the same area in Fig. 3 (g) are around or less than 15 dB. To suppress the large artifacts in the image of the selected array, more advanced inversion approaches, for example, regularization methods can be utilized for image formation. Another natural solution is to keep proper redundancy in the measured data by selecting slightly extra spatial samples. So a trade-off should be considered between the redundancy and the computational load of the inversion operation.

IV. Conclusion

In this paper, the problem of spatial samples selection for electromagnetic linear inversion (e.g., GPR imaging) is studied. To minimize the number of required measurements in the spatial domain and optimize their positions, we extended the FrameSense method to optimize the clustered observation vectors for linear inversion problems. Similar to the traditional

FrameSense method, the clustered FrameSense method gradually removes the observation vectors that maximally increase the frame potential of the observation matrix. However, the observation vectors associated with each spatial samples are eliminated as a group at each time in clustered FrameSense, which results in lower computational complexity compared to the traditional FrameSense. To demonstrate the performance of the clustered FrameSense for spatial sample selection, a GPR imaging experiment with a rotated antenna array was carried out. The imaging results show that the dense sampling according to traditional Nyquist-based approach and the sparse sampling based on the clustered FrameSense method result in comparable imaging performance. So the clustered FrameSense effectively eliminate redundant/less informative spatial samples, thus significantly reducing the computational load for linear inversion.

ACKNOWLEDGMENT

This work was partly supported by the NeTTUN project funded by the European Commission within the FP-7 framework programme under the grant 280712.

REFERENCES

[1] T. Sato, K. Takeda, T. Nagamatsu, T. Wakayama, I. Kimura, and T. Shinbo, "Automatic signal processing of front monitor radar for tunneling machines," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 35, no. 2, pp. 354–359, 1997.

[2] J. Wang, H. Cetinkaya, A. Yarovoy, I. I. Vermesan, and S. Reynaud, "Investigation of forward-looking synthetic circular array for subsurface imaging in tunnel boring machine application," in *Advanced Ground Penetrating Radar (IWAGPR), 2015 8th International Workshop on*, pp. 1–4.

[3] J. Wang, H. Cetinkaya, and A. Yarovoy, "On polar sampling of gpr for tunneling boring machine," in *Ground Penetrating Radar (GPR), 2014 15th International Conference on*, pp. 330–333.

[4] J. L. Schwartz and B. D. Steinberg, "Ultrasparse, ultrawideband arrays," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 45, no. 2, pp. 376–393, 1998.

[5] Y. Gao and S. J. Reeves, "Optimal k-space sampling in mrsi for images with a limited region of support," *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1168–1178, 2000.

[6] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[7] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 684–698, 2015.

[8] J. Ranieri, A. Chebira, and M. Vetterli, "Near-optimal sensor placement for linear inverse problems," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1135–1146, 2014.

[9] I. Ivashko, G. Leus, and A. Yarovoy, "Radar network topology optimization for joint target position and velocity estimation," *Signal Processing*, vol. 130, pp. 279–288, 2017.

[10] S. J. Reeves and Z. Zhao, "Sequential algorithms for observation selection," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 123–132, 1999.

[11] C. Warren, A. Giannopoulos, and I. Giannakis, "gprmax: Open source software to simulate electromagnetic wave propagation for ground penetrating radar," *Computer Physics Communications*, vol. 209, pp. 163–170, 2016.

[12] J. Wang, P. Aubry, and A. Yarovoy, "Linear inversion of polarization-varied GPR data based on exact radiation patterns," *Geoscience and Remote Sensing, IEEE Transactions on*, 2017, under review.