Solving Full-Wave Nonlinear Inverse Scattering Problems by Deep Learning Schemes

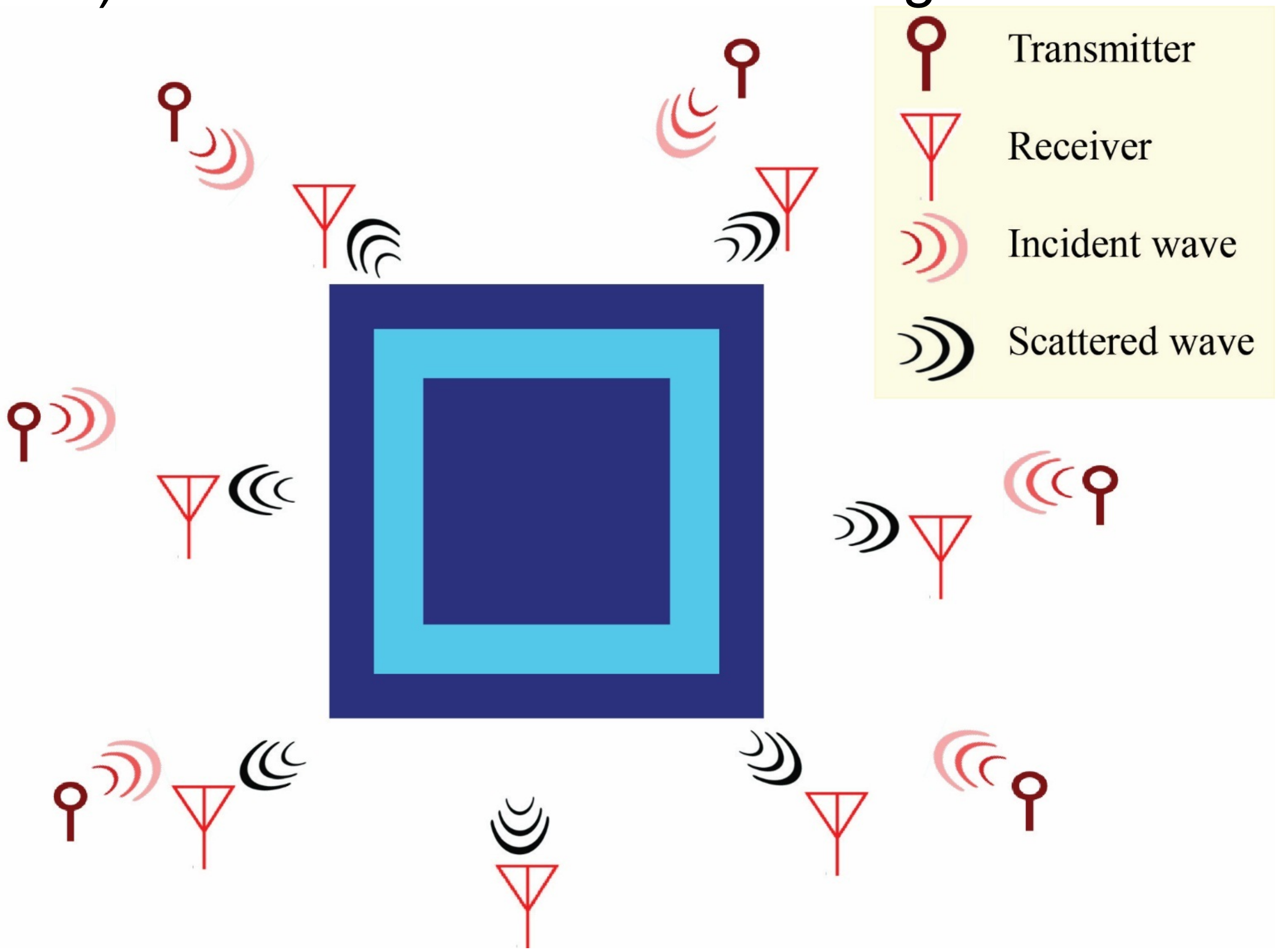
Zhun WEI and Xudong CHEN

Department of Electrical and Computer Engineering, National University of Singapore

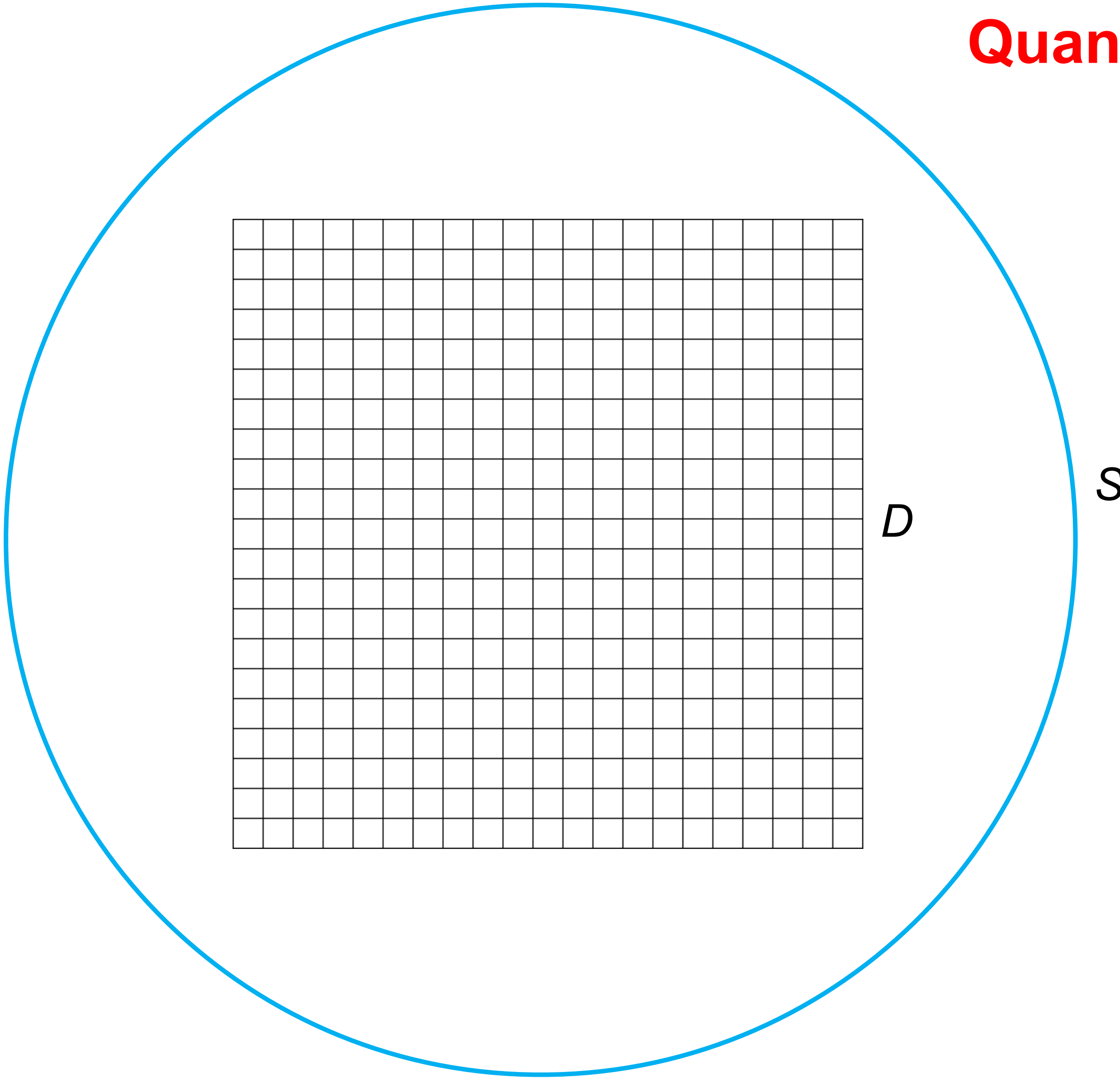
ISP Workshop : Singapore, Sept. 2018

Introduction

- To find unknown scatterers (permittivities, sizes, locations) inside the wall from scattering data

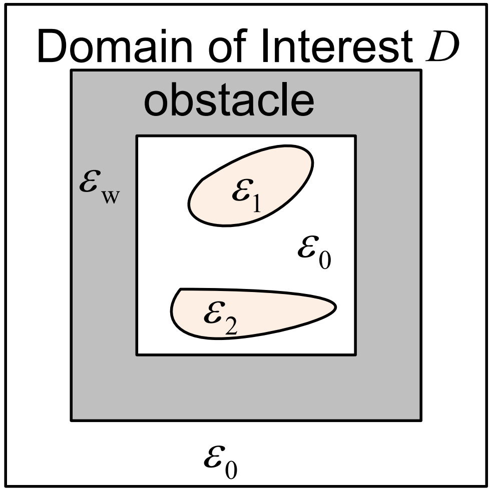


Introduction



Quantitative imaging:

The value of ε_r will be recovered at each pixel of D



Introduction: Forward Scattering (2D)

- Inhomogeneous background with $\varepsilon_b(\mathbf{r})$ μ_0
- Distribution of permittivities $\varepsilon(\mathbf{r}) = \varepsilon_r(\mathbf{r}) \cdot \varepsilon_b(\mathbf{r})$
- Governing equation:

$$\left[\nabla^2 + k_b^2(\mathbf{r})\right] E_z^{\text{inc}}(\mathbf{r}) = 0, \qquad \left[\nabla^2 + k^2(\mathbf{r})\right] E_z^{\text{tot}}(\mathbf{r}) = 0, \quad \mathbf{r} \in D$$

$$\Rightarrow \left[\nabla^2 + k_b^2(\mathbf{r})\right] E_z^{\text{sca}}(\mathbf{r}) = -[k^2(\mathbf{r}) - k_b^2(\mathbf{r})] E_z^{\text{tot}}(\mathbf{r}) = -J(\mathbf{r}) \text{ [defined]}$$

where

$$k_b = \omega\sqrt{\varepsilon_b \mu_0}$$

$$E_z^{\text{tot}} = E_z^{\text{sca}} + E_z^{\text{inc}}$$

- Scattered field

$$E_z^{\text{sca}}(\mathbf{r}) = \int_D g(k_b; \mathbf{r}, \mathbf{r}') J(\mathbf{r}') d\mathbf{r}',$$

$$J(\mathbf{r}) = [k^2(\mathbf{r}) - k_b^2(\mathbf{r})] \cdot \left[E_z^{\text{inc}} + \int_D g(k_b; \mathbf{r}, \mathbf{r}') J(\mathbf{r}') d\mathbf{r}'\right]$$

$$= k_b^2(\mathbf{r})\left[\varepsilon_r(\mathbf{r}) - 1\right] \cdot \left[E_z^{\text{inc}} + \int_D g(k_b; \mathbf{r}, \mathbf{r}') J(\mathbf{r}') d\mathbf{r}'\right]$$

Introduction: Forward Scattering

- Introduce notation

$$g(k_b;\mathbf{r},\mathbf{r}') \text{ is denoted as} \begin{cases} g_S(k_b;\mathbf{r},\mathbf{r}') & \text{for } \mathbf{r} \in S \text{ (at receiver)} \\ g_D(k_b;\mathbf{r},\mathbf{r}') & \text{for } \mathbf{r} \in D \text{ (inside domain)} \end{cases}$$

- Scattering equations:

$$E_z^{\text{sca}}(\mathbf{r}) = \int_D g_S(k_b;\mathbf{r},\mathbf{r}') J(\mathbf{r}') d\mathbf{r}', \qquad \mathbf{r} \in S$$

$$J(\mathbf{r}) = k_b^2(\mathbf{r})\left[\varepsilon_r(\mathbf{r}) - 1\right] \cdot \left[E_z^{\text{inc}} + \int_D g_D(k_b;\mathbf{r},\mathbf{r}') J(\mathbf{r}') d\mathbf{r}'\right], \quad \mathbf{r} \in D$$

A different version of Lippmann-Schwinger equation

Introduction: Forward Problem

- Data equation

$$\overline{E}^{scat} = \overline{\overline{G}}_S \cdot \overline{J}$$

- State equation

$$\overline{J} = \overline{\overline{\chi}} \cdot \left(\overline{E}^{inc} + \overline{\overline{G}}_D \cdot \overline{J} \right)$$

where $\bar{\bar{\chi}}$ is the diagonal matrix consisting of $k_b^2(\varepsilon_r - 1)$, referred to as the contrast (with the background).

- Method of solving forward problem: Eliminate $\overline{J}$ and obtain the Incidence-to-Scattering mapping

$$\overline{E}^{scat} = \overline{\overline{P}} \cdot \overline{E}^{inc}$$

Introduction: Inverse Problem

Forward problem

$$H\{x\} = y$$

an image $x \in \mathcal{X}$, a vector of measurements $y \in \mathcal{Y}$

operator $H : \mathcal{X} \rightarrow \mathcal{Y}$

Inverse problem

recover the original image, x, from the measurements, y

$$R : \mathcal{Y} \rightarrow \mathcal{X}$$

Objective function approach

$$R_{\text{obj}}\{y\} = \underset{x \in \mathcal{X}}{\operatorname{argmin}}\, f(H\{x\}, y) + g(x)$$

$f : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}^+$ is an appropriate measure of error

$g : \mathcal{X} \rightarrow \mathbb{R}^+$ is a regularization functional

Introduction: Inverse Problem

Learning approach

Given a training set of ground-truth images and their corresponding measurements $\{(x_n, y_n)\}_{n=1}^{N}$

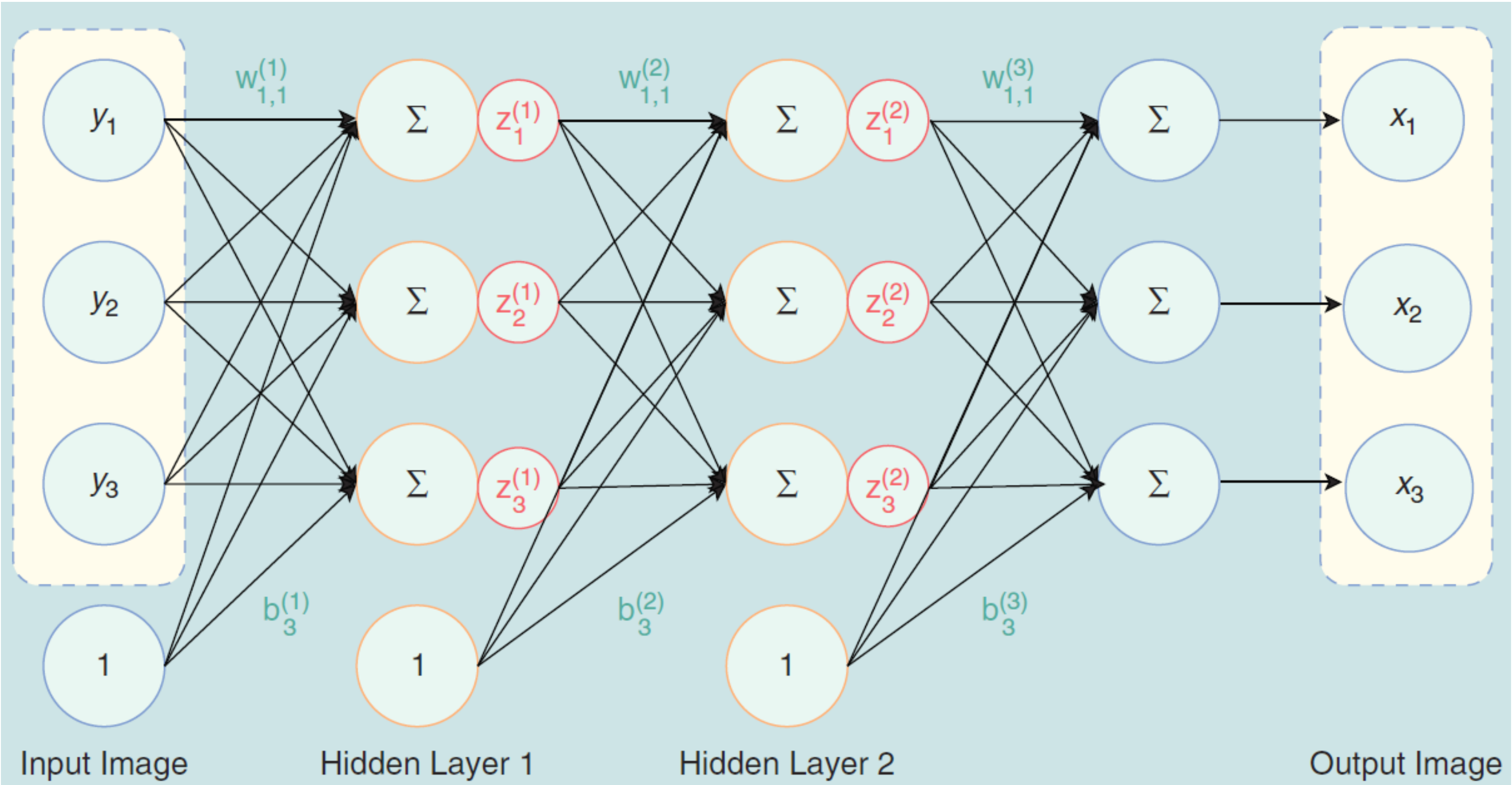
$$R_{\text{learn}} = \underset{R_\theta, \theta \in \Theta}{\operatorname{argmin}} \sum_{n=1}^{N} f(x_n, R_\theta\{y_n\}) + g(\theta)$$

Θ is the set of all possible parameters

$f : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ is a measure of error

$g : \Theta \to \mathbb{R}^+$ is a regularizer

An example of a fully connected neural network with two hidden layers



The activation of the jth output neuron in layer l is defined as

$z_j^l = f\left(\sum_i w_{i,j}^l z_i^{l-1} + b_j^l\right)$, where $f(\cdot)$ is the chosen activation function.

All weights $\mathbf{w}$ and biases $\mathbf{b}$ are learned during the training phase.

Universal approximation theorem

Lucas et.al. IEEE Signal Processing Magazine, 2018

Computational Cost: Objective function approach

- Computational complexity of inverse scattering problem (if a traditional objective function is used):

$$O(N_{\text{opt}} \cdot N_{\text{for}} \cdot N_{\text{inc}} \cdot M \log M)$$

N_{opt} : Number of iterations for optmization

N_{for} : Number of iterations for forward problem

N_{inc} : Number of incidences (Degree of Freedom)

M : Number of pixels (also number of unknowns)

- Bottleneck: Nonlinearity of the objective function
    - Need a large N_{opt} to reach a global minimum;
    - Often pre-converge to a local minimum
- Main objective is to reduce the N_{opt} by rewriting the objective function in a way such that it depends in a much less nonlinear way on unknowns.

Computational Cost: Objective function approach

- The traditional inversion algorithm: The objective function involves an inversion of a matrix that contains the unknowns (permittivity or contrast)
- Inversion methods that rewriting the objective function
  - Dependence: Unknowns' 4^{th} order polynomial
  - Contrast Source Inversion (CSI)
  - Contrast Source Extended Born (CSEB)
  - Subspace-based Optimization Method (SOM)
  - Compression by alternative bases:
  - Fourier
  - Wavelet

Inverse problem: Learning approach

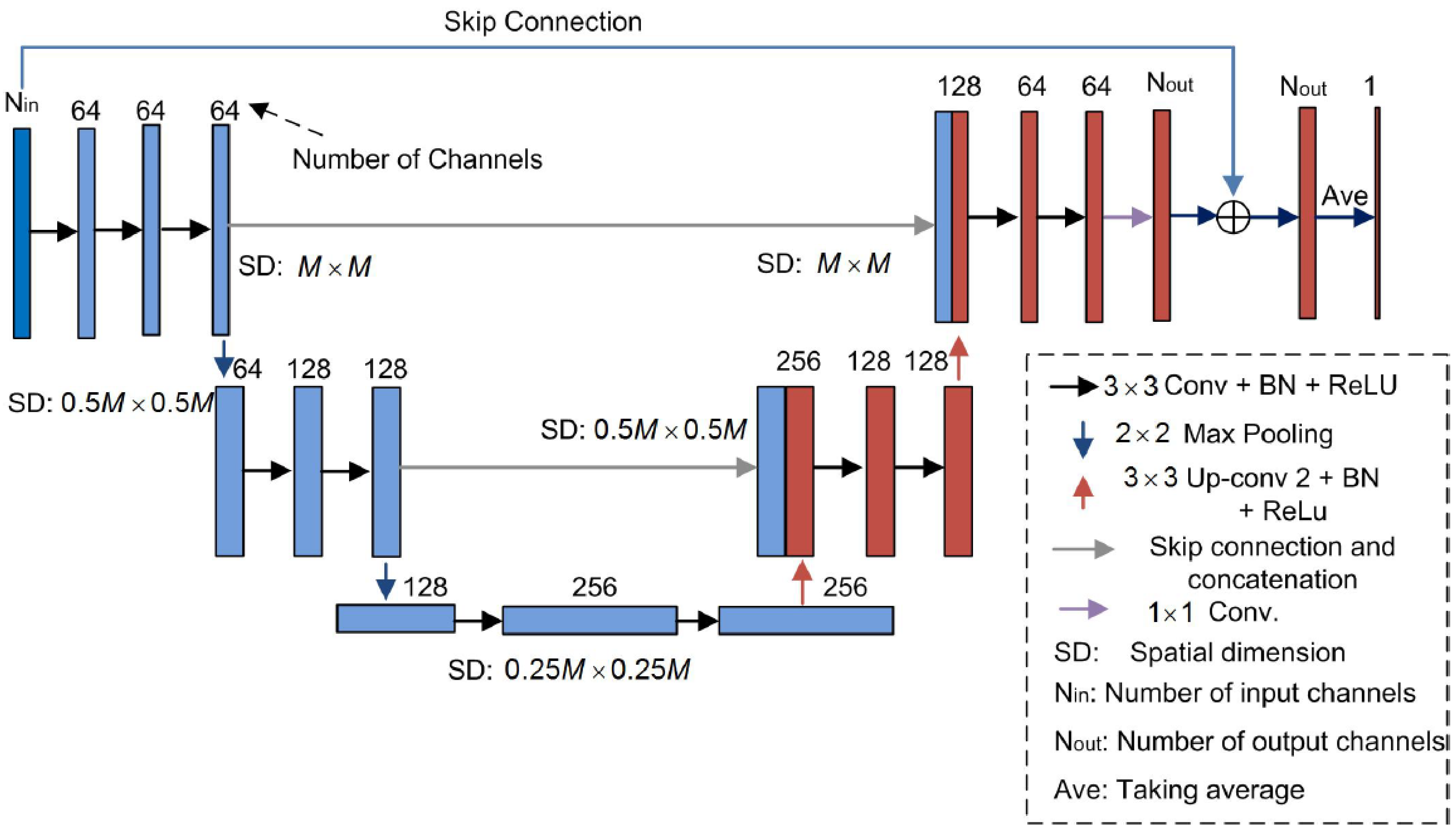
Deep learning approach has **<u>not</u>** had the profound impact on inverse problems that they have had for object classification

Three categories:

- Direct learning: (x, y)
  Comment: Black-box: no insight
- Hybrid approach: still use the <u>objective function approach</u> but learn some operators in each iteration of optimization. [such as: gradient, in Adler, Inverse Problems, 2017]
  Comment: Overall difficulty may not be reduced
- New-representation: $(\hat{x},\ \hat{y})$
  Comment: needs mathematical and physical insights;
  most promising

ISP: non-pixel representation:

Bermani, TGRS, 2003; Rekanos, TM, 2002; Caorsi, TGRS, 1999



The U-net architecture for the proposed three CNN schemes: DIS, BPS, and DCS

Three CNN schemes

1. Direct Inversion Scheme (DIS)

Inputs: γ Scattered field; Output: the contrasts $\bar{\bar{\chi}}$

2. Back-Propagation Scheme (BPS)

$$\overline{I}^{b} = \gamma \cdot \overline{\overline{G}}_{S}^{H} \cdot \overline{E}^{s}$$

$$\gamma = \frac{(\overline{E}^{s})^{T} \cdot (\overline{\overline{G}}_{S} \cdot (\overline{\overline{G}}_{S}^{H} \cdot \overline{E}^{s}))^{*}}{||\overline{\overline{G}}_{S} \cdot (\overline{\overline{G}}_{S}^{H} \cdot \overline{E}^{s})||}$$

$$\overline{\chi}^{b}(n) = \frac{\sum_{p=1}^{N_i} \overline{I}_{p}^{b}(n) \cdot [\overline{E}_{p}^{t,b}(n)]^{*}}{\sum_{p=1}^{N_i} ||\overline{E}_{p}^{t,b}(n)||^{2}}$$

Inputs: the BP contrasts $\bar{\bar{\chi}}^{b}$; Output: the contrasts $\bar{\bar{\chi}}$

3. Dominant Current Scheme (DCS)

- Recall the two equations

$$\overline{E}^{scat} = \overline{\overline{G}}_S \cdot \overline{J}$$

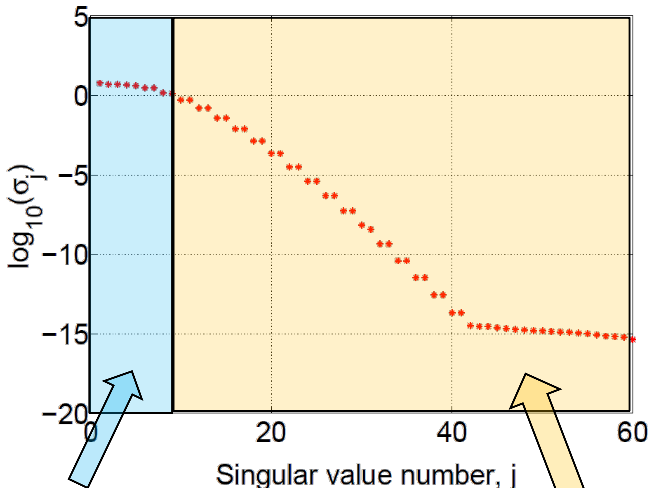
$$\overline{J} = \overline{\overline{\chi}} \cdot \left(\overline{E}^{inc} + \overline{\overline{G}}_D \cdot \overline{J} \right)$$

- Important: both the $\overline{\overline{G}}_S$ and $\overline{\overline{G}}_D$ operators are independent of unknown scatterers
  - Motivate us to analyze the property of these two operators before reconstructing the contrast $\overline{\overline{\chi}}$
  - The computational overhead of such analysis should not be large

Dominant Current Scheme (DCS)

- Singular Value Decomposition (SVD)

$$\overline{\overline{G}}_S = \overline{\overline{U}}_S \cdot \overline{\overline{\Sigma}}_S \cdot \overline{\overline{V}}_S^{\,*}$$



First L leading singular values

Other smaller singular values

$$\overline{J} = \overline{\overline{V}}_S \cdot \overline{\alpha} = \boxed{\overline{\overline{V}}_S^{+} \cdot \overline{\alpha}^{d}} + \boxed{\overline{\overline{V}}_S^{-} \cdot \overline{\alpha}^{a}}$$

Deterministic part

Ambiguous part

Dominant Current Scheme (DCS)

- Obtain the deterministic part using the linear relation

$$\bar{J}^{det} = \overline{V}_S^{+} \cdot \bar{\alpha}^{d} = \sum_{j=1}^{L} \bar{v}_j^{+} \frac{\bar{u}_j^{+*} \cdot \overline{E}^{scat}}{\sigma_j}$$

- Dominant part current is defined as:

$$\bar{J}^{d} = \bar{J}^{det} + \bar{J}^{l}$$

$$\bar{J}^{l} = \bar{\bar{F}} \cdot \bar{\alpha}$$ *Low-frequency Fourier components*

$$\overline{\chi}_p^{d}(n) = \frac{\overline{J}_p^{d}(n) \cdot [\overline{E}_p^{t,d}(n)]^*}{||\overline{E}_p^{t,d}(n)||^2}$$

p: index of incidence

Wei, TGRS, Accepted, 2018

DCS: Inputs: the dominant contrasts $\bar{\bar{\chi}}_p^{d}$; Output: the contrasts $\bar{\bar{\chi}}$

Computational cost

- Assume *M* pixels in the domain of interest & N_r receivers: $M >> N_r$
- Computational cost of SVD

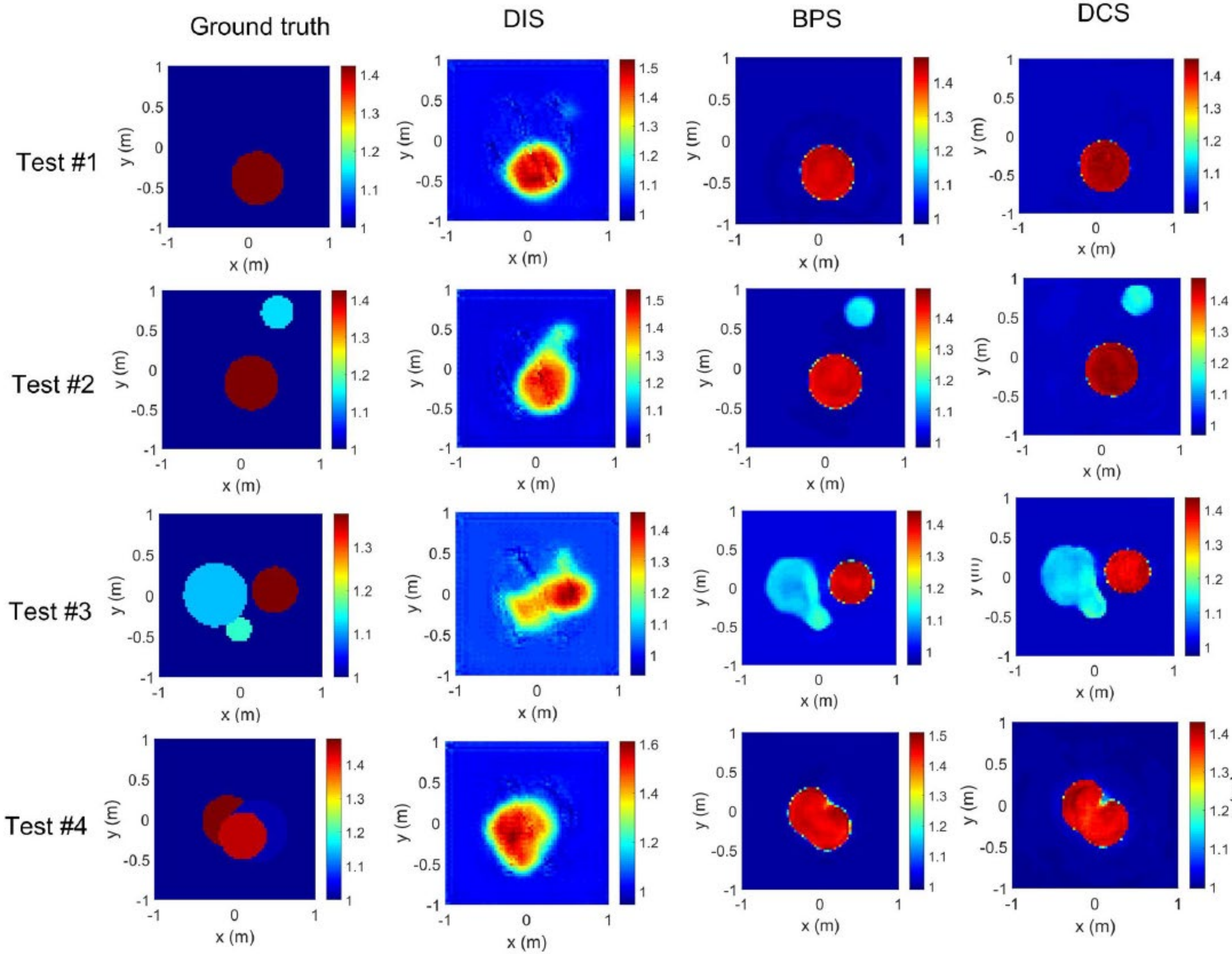
$$\overline{\overline{G}}_S : \quad O(M^2 N_r)$$

- Only a thin SVD of $\overline{\overline{G}}_S$ is needed (first *L* singular vectors)
- Computational cost of thin SVD:

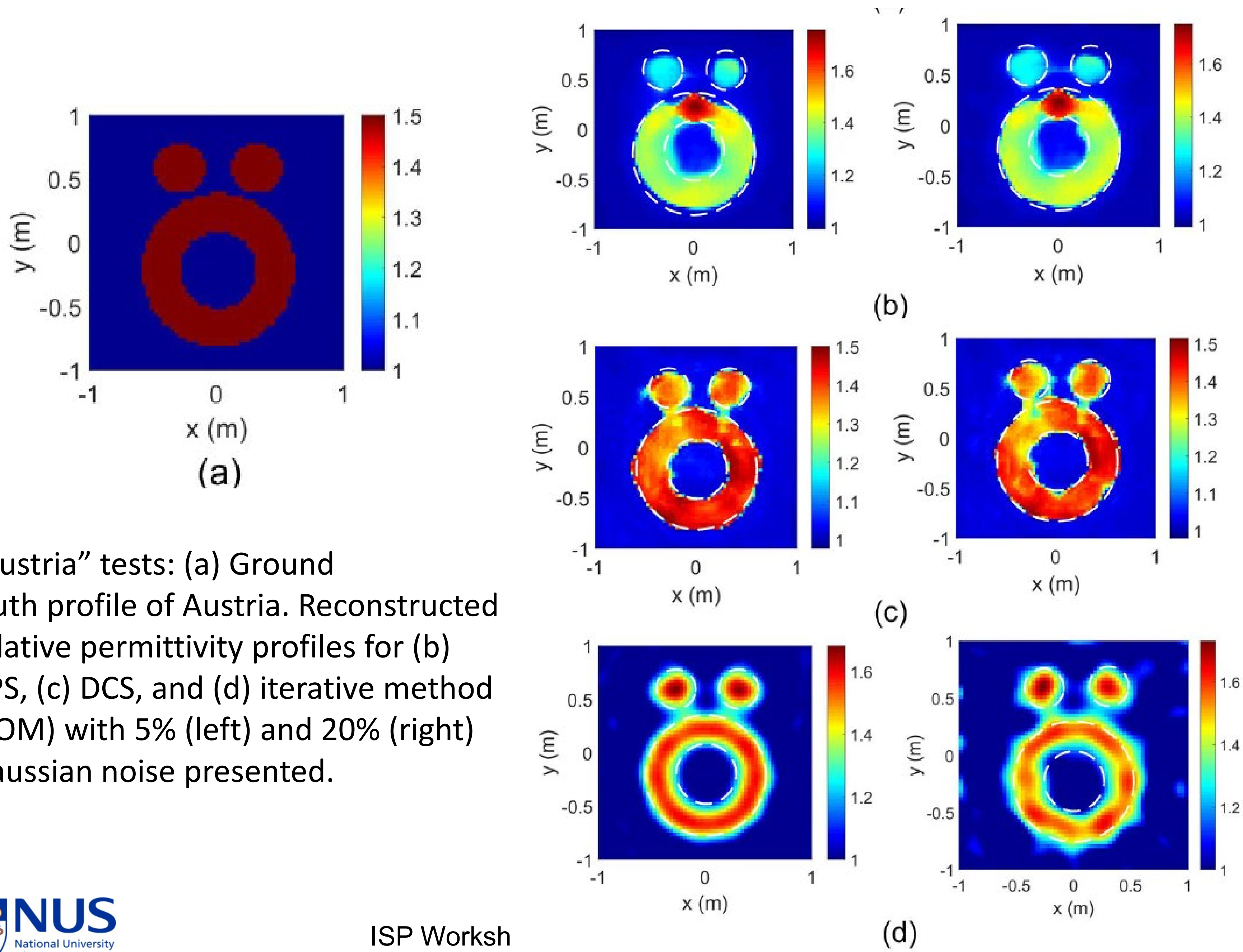
$$\overline{\overline{G}}_S : \quad O(LMN_r^2) \propto O(M)$$

- $\overline{\overline{F}} \cdot \overline{\alpha}$ can be directly calculated by fast Fourier transform

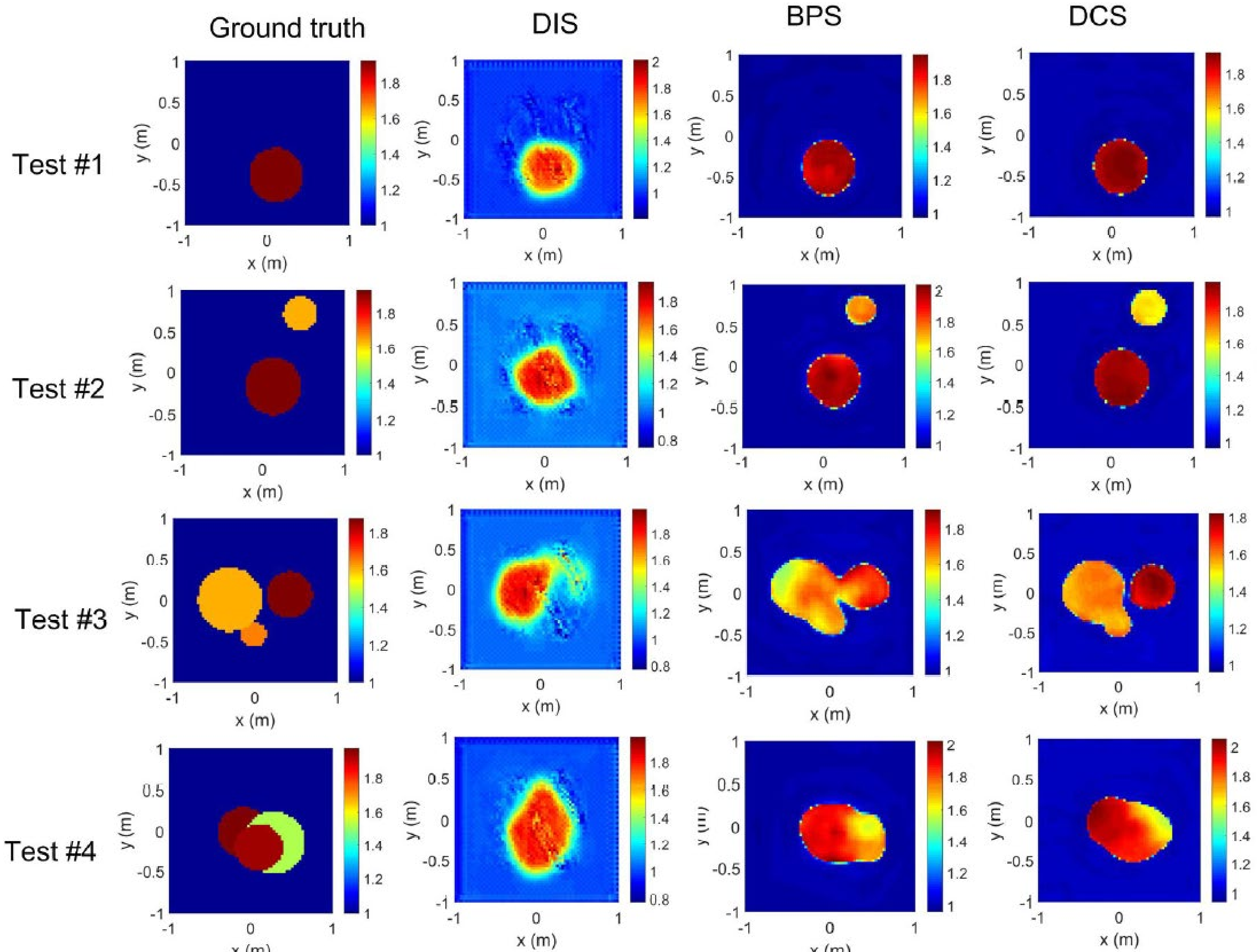
Numerical results



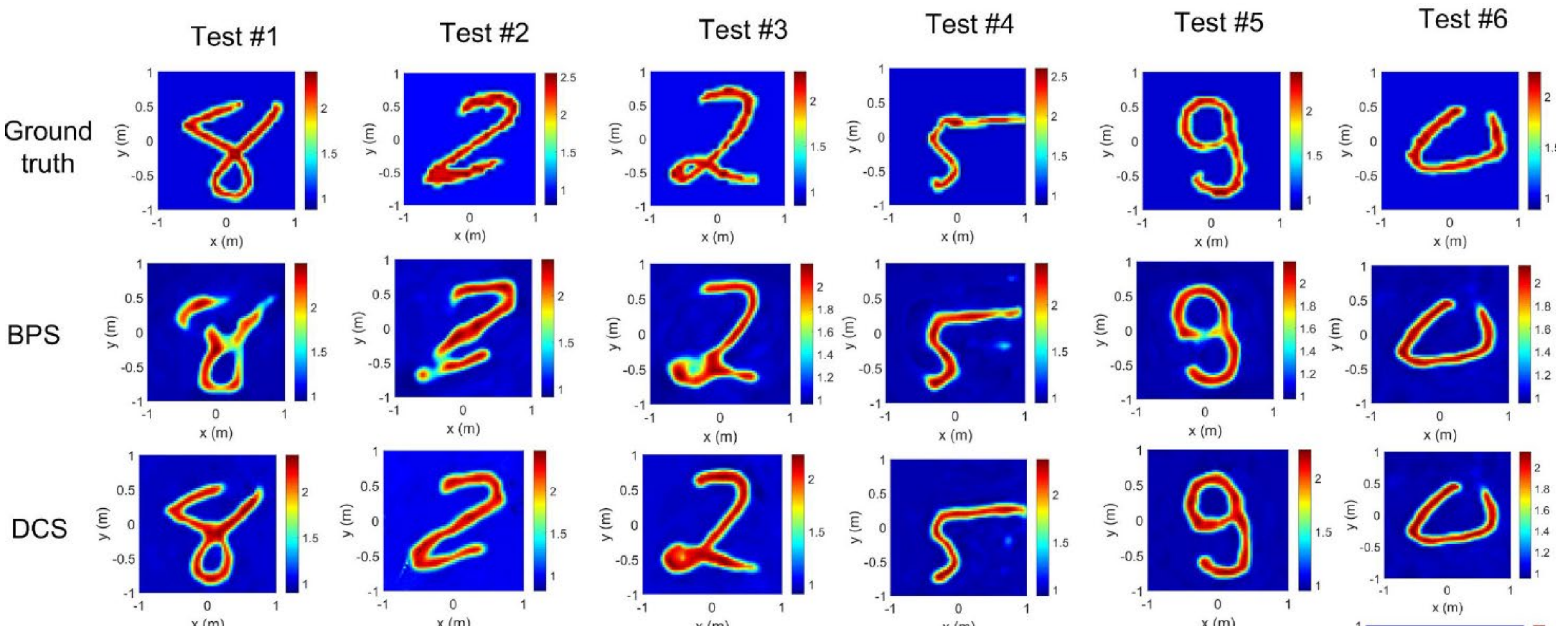
Example One: relative permittivity is between 1 and 1.5



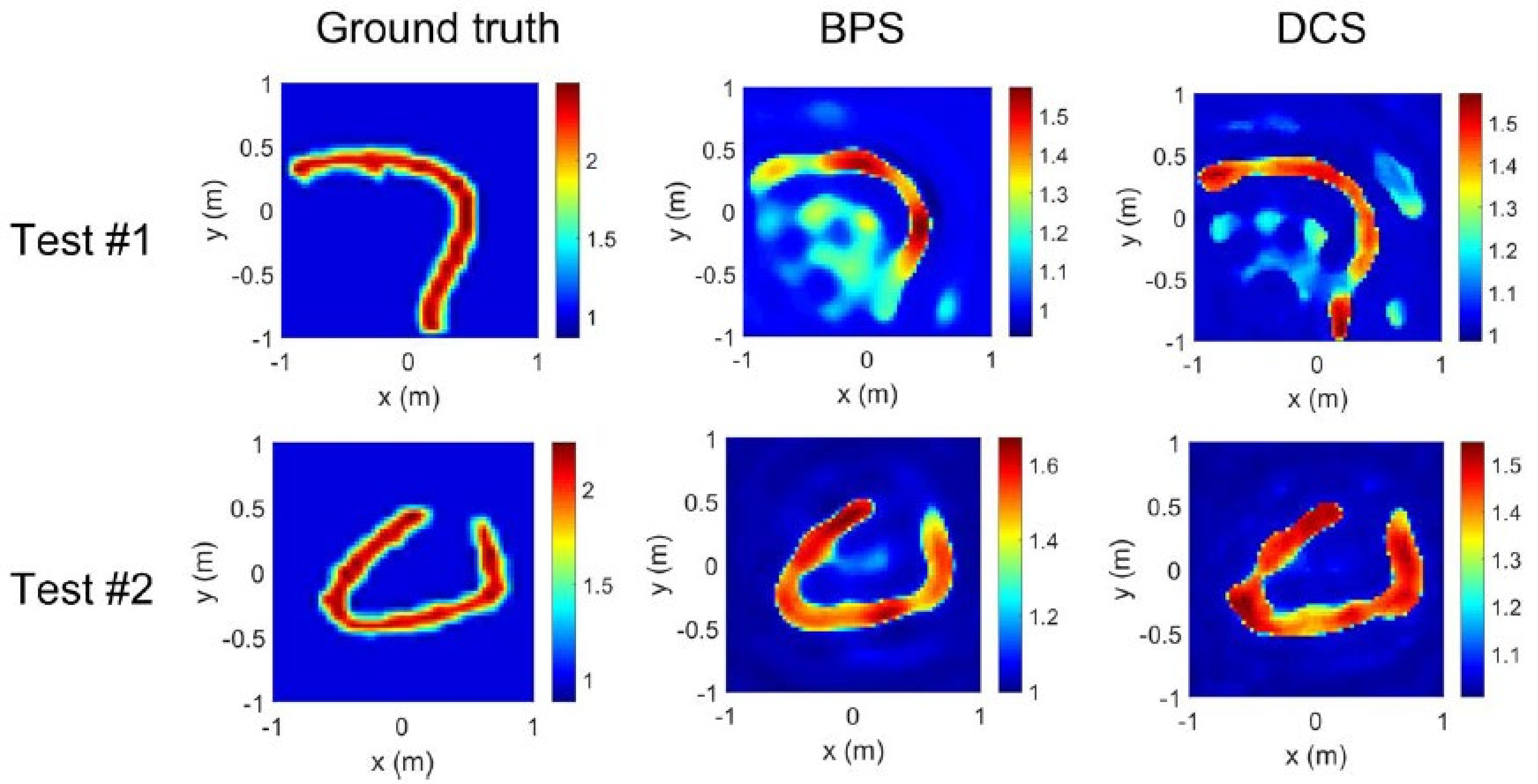
"Austria" tests: (a) Ground truth profile of Austria. Reconstructed relative permittivity profiles for (b) BPS, (c) DCS, and (d) iterative method (SOM) with 5% (left) and 20% (right) Gaussian noise presented.



Example Two: relative permittivity is between 1.5 and 2



Example Three: Tests with MNIST database, the relative permittivity is between 2 and 2.5



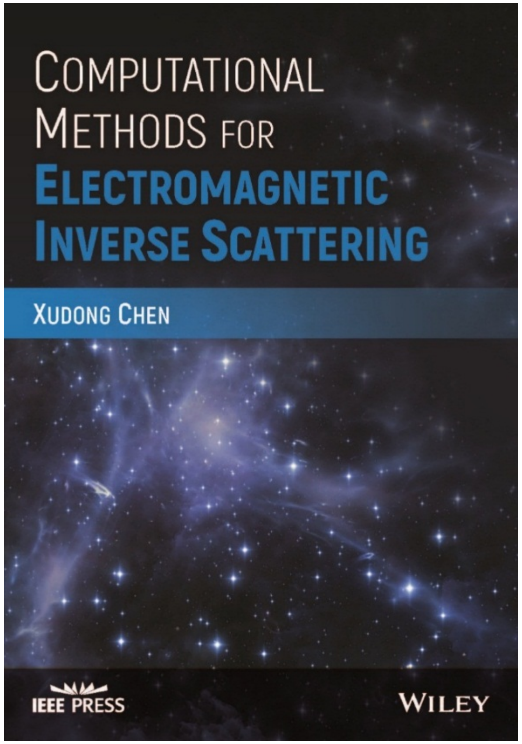
Use the network trained with circular-cylinders in Example 1 to test the MNIST database in Example 3

Take-home message

- No matter using objective function approach or learning approach, the key is to construct corresponding target function in a way such that it depends in a much less nonlinear way on unknowns.
- Avoid directly dealing with measurement data, where CNN has to spend unnecessary cost to train and learn underlying wave physics. Extract out as much as possible what people can do and leave the remaining to machine.
- The above two need a fairly good understanding of the forward problem (physical and mathematical insights)

Reference:

Z. Wei and **X. Chen**, "Deep Learning Schemes for Full-Wave Nonlinear Inverse Scattering Problems, " *IEEE Trans. on Geoscience and Remote Sensing*, Accepted, 2018



X. Chen,

Computational Methods for Electromagnetic Inverse Scattering.

Wiley, 2018

Chapter 6:

Reconstructing Dielectric Scatterers

Thank you!