Diagnosis of Intelligent Reflecting Surface in Millimeter-wave Communication Systems

Rui Sun, Weidong Wang, Li Chen, Guo Wei, and Wenyi Zhang, *Senior Member, IEEE*

arXiv:2101.03792v2 [eess.SY] 8 Jul 2021

Abstract

Intelligent reflecting surface (IRS) is a promising technology for enhancing wireless communication systems. It adaptively configures massive passive reflecting elements to control wireless channel in a desirable way. Due to hardware characteristics and deploying environments, an IRS may be subject to reflecting element blockages and failures, and hence developing diagnostic techniques is of great significance to system monitoring and maintenance. In this paper, we develop diagnostic techniques for IRS systems to locate faulty reflecting elements and retrieve failure parameters. Three cases of channel state information (CSI) availability are considered. In the first case where full CSI is available, a compressed sensing based diagnostic technique is proposed, which significantly reduces the required number of measurements. In the second case where only partial CSI is available, we jointly exploit the sparsity of the millimeter-wave channel and the failure, and adopt compressed sparse and low-rank matrix recovery algorithm to decouple channel and failure. In the third case where no CSI is available, a novel atomic norm is introduced as the sparsity-inducing norm of the cascaded channel, and the diagnosis problem is formulated as a joint sparse recovery problem. Finally, the proposed diagnostic techniques are validated through numerical simulations.

Index Terms

Array diagnosis, atomic norm, compressed sensing, intelligent reflecting surface (IRS), millimeter wave communication

The authors are with the CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China, Hefei 230027, China (e-mail: ruisun@mail.ustc.edu.cn; wdwang@ustc.edu.cn; chenli87@ustc.edu.cn; wei@ustc.edu.cn; wenyizha@ustc.edu.cn).

Codes available on: https://github.com/DestinationSR/IRSdiagnosis.

I. Introduction

Intelligent reflecting surface (IRS) is an emerging technology to improve the performance of wireless communication systems. By adaptively adjusting the attenuation and/or phase shift of massive passive reflecting elements, the IRS is able to "program" the wireless propagation channel with low-cost hardware, which shows great potential for improving spectrum efficiency and energy efficiency, especially for millimeter-wave (mmWave) communication system due to its vulnerability to high path-loss and link blockage [1], [2]. Massive research works on the IRS have been conducted recently, including channel modeling [3], [4], beamforming design [5], [6], and hardware experiments [7], [8].

In the meantime, hardware characteristics and deploying environments also bring unique reliability challenges to the IRS. The IRS consists of massive passive "meta-atoms" that are made of two-dimensional (2D) metamaterial [9], [10], which may result in a high failure probability due to large system scale and dense interconnections of tuning circuits [11], [12]. On the other hand, the short wavelength of mmWave band makes the IRS vulnerable to the blockages with comparable sizes like dirt, particles, and flying debris in outdoor deploying environments [13]. Reflecting characteristics of the IRS also make it more sensitive to failures than conventional MIMO systems. It is reported in [14] that the power gain provided by the IRS scales up with the square of the number of reflecting elements. Therefore, the IRS also suffers from "squared power loss" since faulty reflecting elements could cause twice attenuation during incidence and reflection. It is also reported in [12] that faulty reflecting elements may result in various radiation pattern distortion issues like degraded directivity, increased side-lobe level, and beam misalignment. In summary, the existence of faulty reflecting elements could significantly degrade system performance, which necessitates diagnostic techniques for locating faulty reflecting elements and retrieving failure parameters for calibration purposes.

To provide references for the IRS diagnosis, we refer to extensive related works on antenna array diagnosis. Early representative works include [15] and [16]. These techniques exploit the linear relationship between antenna pattern and array excitation, and take samples uniformly on the radiation pattern to recover actual array excitation by solving an inverse problem. Therefore, faulty antenna elements can be identified by comparing the actual array excitation and the ideal one. In order to make the inverse problem solvable, these techniques require the number of measurements no less than the total number of antenna elements, which may be time-consuming

for large-scale systems. About a decade ago, compressed sensing was introduced to array diagnosis [17], which efficiently captures the failure sparsity that the number of faulty antenna elements is usually much smaller than the total number of antenna elements. Hence, one can subtract the ideal antenna pattern from the observed one to obtain the differential pattern. The resulting differential array excitation is a sparse signal that can be recovered from compressed measurements, in which non-zero entries indicate faulty antenna elements. That work has initiated numerous compressed sensing based diagnostic techniques. One type of improvement focuses on the enhancement in sparse recovery algorithms, such as Bayesian compressive sensing [18], [19], non-convex compressed sensing [20], and mixed-norm constrained compressed sensing [21]. Other improvements are mainly the development of new sampling methods, including far-field sampling [22], phase-less sampling [23], phase domain sampling [24], and random beam sampling [13], [25], [26]. In particular, the work [27] extended the above antenna-only diagnoses to the joint diagnosis of antenna, phase shifter, and RF chain of hybrid beamforming (HBF) system.

The above pioneering works provide various approaches to detect faulty antenna elements in conventional MIMO systems. However, they do not directly apply to the IRS diagnosis since they fail to address the following challenges. The first challenge is the coupling between channel and failure. Intuitively speaking, it is difficult to distinguish antenna failure from channel fading since an antenna failure can be treated as a part of channel. To cope with this issue, almost all of currently proposed diagnostic techniques assume a free-space wireless environment in order to calculate the channel using known spatial coordinates, which is then used to generate fault-free reference signal for detecting failures. Such requirement significantly limits their applications in outdoor online diagnosis since wireless channel is subject to inevitable multipath scattering, which makes accurate reference signal unavailable. The second challenge arises from the hardware characteristic of the IRS. The IRS is a passive device that has no signal processing capability except for configurable reflecting coefficients. Hence, one has no direct access to the received signal of each reflecting element. To perform diagnosis, auxiliary transmitter (TX) and receiver (RX) are required to play the role of “signal generator” and “signal analyzer”, respectively. The equivalent channel from the TX through the IRS to the RX is a cascaded channel, which combines the effect of multipath scattering in both TX-IRS and IRS-RX channels. Such channel characteristic makes the channel-failure coupling issue even worse and introduces more channel uncertainties to the diagnosis. Current literature fails to address the above issues.

This motivates us to develop tailored diagnostic techniques that incorporate the hardware and channel characteristics of an IRS system.

In this paper, we develop diagnostic techniques for an IRS system to locate faulty reflecting elements and retrieve failure parameters. We first establish the IRS failure model and discuss the channel state information (CSI) availability. Specifically, three cases of the CSI availability are considered and three different diagnostic techniques are developed accordingly. For the diagnosis with full CSI (all channels are known), a compressed sensing based diagnostic technique is proposed, which significantly shortens the diagnostic time by exploiting the failure sparsity. For the diagnosis with partial CSI (only the channel from the IRS to the RX is known), the channel-failure coupling issue is solved by jointly exploiting the sparsities of failure and mmWave channel, and the diagnosis problem is formulated as a compressed sparse and low-rank matrix recovery (cSLRMR) problem. For the diagnosis without CSI (both channels are unknown), a novel atomic norm is introduced to induce the sparsity of cascaded channel, and an atomic norm minimization (ANM) based diagnostic technique is proposed to decouple the cascaded channel and failure. The proposed ANM-based diagnostic technique is further extended to the multiple measurement vectors (MMV) model to exploit the diversity gain provided by multipath scattering. Two efficient algorithms are also proposed to accelerate the diagnosis. The main contributions of this paper can be summarized as follows:

- **IRS failure modeling.** We establish the failure model for an IRS system, which incorporates the attenuation and phase-shifting effects of reflecting element failure. The proposed failure model is further extended to the multi-antenna setting to exploit the diversity gain provided by multipath scattering.
- **Channel-failure decoupling.** The coupling between channel fading and failure is a major challenge of IRS diagnosis. To overcome this issue, we jointly exploit the sparsity of mmWave channel and failure, and propose a cSLRMR-based diagnostic technique and an ANM-based diagnostic technique for the case of partial CSI and no CSI, respectively.
- **Efficient diagnostic algorithms.** The proposed ANM-based diagnostic technique has strong robustness against multipath scattering, but it also suffers from high computational complexity if conventional solvers are adopted. To accelerate the diagnosis, we develop alternating direction method of multipliers (ADMM) based algorithms to implement the proposed techniques, which significantly shorten the diagnostic time compared with conventional algorithms.

TABLE I
NOTATIONS ADOPTED IN THIS PAPER.

Notation	Description
a, A	a scalar
$\mathbf{a}$	a vector
$\mathbf{A}$	a matrix
$[\mathbf{a}]_i$	i-th element of vector $\mathbf{a}$
$[\mathbf{A}]_{i,j}$	the element in i-th row and j-th column of matrix $\mathbf{A}$
$[\mathbf{A}]_{i,:}$	i-th row of $\mathbf{A}$
$\mathbf{I}_N$	an $N \times N$ identity matrix
$\mathbf{1}_{M\times N}$	an all-one matrix of size $M \times N$
$\|\mathbf{a}\|_p$	l_p norm of $\mathbf{a}$
$\|\mathbf{a}\|_{\mathcal{A}}$	atomic norm of $\mathbf{a}$
$\|\mathbf{A}\|_{\mathrm{F}}$	Frobenius norm of $\mathbf{A}$
$\|\mathbf{A}\|_*$	nuclear norm of $\mathbf{A}$ (sum of the singular values of $\mathbf{A}$)
$(\cdot)^{\mathrm{T}}$	transpose
$(\cdot)^{\mathrm{H}}$	conjugate transpose
$\mathrm{vec}(\cdot)$	vectorization
$\mathrm{diag}(\mathbf{a})$	a diagonal matrix with the elements of $\mathbf{a}$ on its main diagonal
$\mathrm{tr}(\cdot)$	matrix trace
$\circ$	Hadamard product (element-wise product)
$\oslash$	Hadamard division (element-wise division)
$\otimes$	Kronecker product
$\mathcal{CN}(\mu, \sigma^2)$	circularly symmetric complex Gaussian distribution with mean μ and variance σ^2
$U(a, b)$	uniform distribution over the interval (a, b)

The remaining part of this paper is organized as follows. Section II describes system failure model and channel model. Section III, Section IV, and Section V present the diagnostic techniques that require full CSI, partial CSI, and no CSI, respectively. Section VI presents some illustrative instances and numerical results. Finally, Section VII concludes the paper.

The notations adopted in this paper are listed in Table I.

II. SYSTEM MODEL

In this section, we present system failure model and discuss the CSI availability of three cases.

A. Failure Model

The diagnosis setup is shown in Fig. 1, where the IRS is operating at the mmWave band and is the system to be diagnosed. The IRS is in uniform planar array (UPA) structure and consists of $N = H \times W$ reflecting elements, where H and W are the number of elements along vertical and horizontal directions, respectively. Both the TX and the RX are equipped with a single antenna.[1]

[1]The extension to multi-antenna RX will be presented in Sec. V-D.

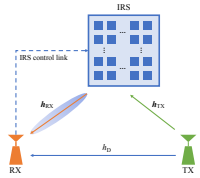


Fig. 1. Diagnosis setup.

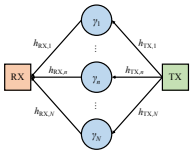


Fig. 2. IRS weighting scheme, where $h_{\mathrm{TX},n}$ and $h_{\mathrm{RX},n}$ are the n-th channel coefficient in $\mathbf{h}_{\mathrm{TX}}$ and $\mathbf{h}_{\mathrm{RX}}$, respectively.

The fault-free received signal at the RX can be expressed as [28]

$$y' = \left(\mathbf{h}_{\mathrm{RX}}^{\mathrm{T}}\mathbf{\Gamma}\mathbf{h}_{\mathrm{TX}} + h_{\mathrm{D}}\right) x + w', \tag{1}$$

where $\mathbf{h}_{\mathrm{RX}} \in \mathbb{C}^N$ and $\mathbf{h}_{\mathrm{TX}} \in \mathbb{C}^N$ are the channel vectors from the IRS to the RX and from the TX to the IRS, respectively; $h_{\mathrm{D}} \in \mathbb{C}$ is the direct channel coefficient from the TX to the RX; x and w' are the transmit symbol and the noise, respectively; $\mathbf{\Gamma} \in \mathbb{C}^{N\times N} = \mathrm{diag}(\mathbf{f}) = \mathrm{diag}([\beta_1 e^{j\gamma_1}, \cdots, \beta_N e^{j\gamma_N}]^{\mathrm{T}})$ is the weighting matrix of the IRS, where β_n and γ_n are the attenuation and phase shift of the n-th reflecting element, respectively. To maximize reflection power, we only adjust phase shift and set $\beta_n = 1$ for $n = 1, \cdots, N$ throughout the diagnosis. An illustrative figure of the IRS weighting scheme is shown in Fig. 2.

To represent faulty reflecting elements, we multiply each reflecting coefficient by a failure mask $m_n \in \mathbb{C}, n = 1, \cdots N$ [12], [13], which is defined as

$$m_n = \begin{cases} 1, & n\text{-th reflecting element is functioning} \\ \eta_n e^{j\kappa_n}, & n\text{-th reflecting element is faulty} \end{cases}, \tag{2}$$

where $0 \le \eta_n < 1$ and $0 \le \kappa_n < 2\pi$ represent the failure induced attenuation and phase shift of the n-th reflecting element, respectively. Therefore, the transmission model (1) can be rewritten as

$$y' = \left(\mathbf{h}_{\mathrm{RX}}^{\mathrm{T}}\Big(\mathbf{\Gamma}\mathrm{diag}(\mathbf{m})\Big)\mathbf{h}_{\mathrm{TX}} + h_{\mathrm{D}}\right) x + w', \tag{3}$$

where $\mathbf{m} = [m_1, m_2, \cdots, m_N]^{\mathrm{T}}$ is the vector of failure masks. We assume that a *failure sparsity* property holds; that is, the number of faulty reflecting elements is much smaller than the total number of elements [13], [17], [24], and hence most entries in $\mathbf{m}$ are unity.[2]

[2]In practice, the failure masks $\mathbf{m}$ may not be a near all-one vector due to general hardware imperfections (i.e., $\mathbf{m} - \mathbf{1}_{N\times 1}$ may not be strictly sparse). In view of that, we focus on detecting faults and treat hardware imperfections as additive noises, which can be well handled by adjusting the regularization parameters in the proposed diagnostic techniques.

For the purpose of diagnosis, the TX is set to transmit symbol $x = 1$ throughout the diagnosis. Besides, to eliminate the received signal contributed by the direct channel h_{D}, we first set the IRS to the absorbing mode (i.e., $\beta_n = 0$ for $n = 1, \cdots, N$) before the diagnosis and conduct channel estimation to obtain an estimator of h_{D} as $\hat{h}_{\mathrm{D}}$ [1], and then subtract it from y'. Therefore, the transmission model for diagnosis can be rewritten as

$$\begin{aligned} y = y'|_{x=1} - \hat{h}_{\mathrm{D}} &= \mathbf{h}_{\mathrm{RX}}^{\mathrm{T}}\Big(\mathbf{\Gamma}\mathrm{diag}(\mathbf{m})\Big)\mathbf{h}_{\mathrm{TX}} + \underbrace{h_{\mathrm{D}} - \hat{h}_{\mathrm{D}} + w'}_{=w} \\ &= \mathbf{h}_{\mathrm{RX}}^{\mathrm{T}}\Big(\mathrm{diag}(\mathbf{f})\mathrm{diag}(\mathbf{m})\Big)\mathbf{h}_{\mathrm{TX}} + w = \sum_{n=1}^{N}\left[\mathbf{h}_{\mathrm{RX}} \circ \mathbf{f} \circ \mathbf{m} \circ \mathbf{h}_{\mathrm{TX}}\right]_n + w, \end{aligned} \tag{4}$$

where $w \sim \mathcal{CN}(0, 1/\mathrm{SNR})$ is the augmented noise due to the noisy estimation of h_{D} and SNR is the signal-to-noise ratio.

From received signal y, the goal of diagnosis is to recover the failure mask vector $\mathbf{m}$, in which non-unity entries indicate faulty reflecting elements. Besides, recovered failure masks can be utilized for calibration, which has been treated in [29]–[31] and references therein.

In (4), $\mathbf{f}$ is the configurable weighing vector and $\mathbf{m}$ is the failure mask to be estimated, while $\mathbf{h}_{\mathrm{RX}}$ and $\mathbf{h}_{\mathrm{TX}}$ are unknown channel vectors that need to be estimated based on field measurements. Unfortunately, conventional channel estimation techniques like channel training become invalid since channel vectors and failure masks are coupled (as shown in (4)). At first glance, it seems impossible to distinguish channel fading from reflecting element failure, and hence the CSI availability brings a major challenge to IRS diagnosis. In the following, we will discuss how to deal with unknown CSI in three cases.

B. Channel Model

To better understand our channel modeling assumptions, we first introduce mmWave narrow-band clustered channel model [32]

$$\begin{aligned} \mathbf{h}_{\mathrm{TX}} = \mathrm{vec}(\mathbf{H}_{\mathrm{TX}}) &= \mathrm{vec}\left(\sum_{l=1}^{L_{\mathrm{TX}}} \alpha_{\mathrm{TX},l}\mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{TX},l}, \phi_{\mathrm{TX},l})\mathbf{a}_{\mathrm{W}}^{\mathrm{T}}(\theta_{\mathrm{TX},l}, \phi_{\mathrm{TX},l})\right) \\ &= \sum_{l=1}^{L_{\mathrm{TX}}} \alpha_{\mathrm{TX},l}\Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{TX},l}, \phi_{\mathrm{TX},l}) \otimes \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{TX},l}, \phi_{\mathrm{TX},l})\Big), \\ \mathbf{h}_{\mathrm{RX}} = \mathrm{vec}(\mathbf{H}_{\mathrm{RX}}) &= \mathrm{vec}\left(\sum_{l=1}^{L_{\mathrm{RX}}} \alpha_{\mathrm{RX},l}\mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l})\mathbf{a}_{\mathrm{W}}^{\mathrm{T}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l})\right) \end{aligned}$$

$$= \sum_{l=1}^{L_{\mathrm{RX}}} \alpha_{\mathrm{RX},l} \Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l}) \otimes \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l}) \Big), \tag{5}$$

where

$$\begin{aligned} \mathbf{a}_{\mathrm{H}}(\theta, \phi) &= [1, e^{j2\pi d \sin\theta \sin\phi}, \cdots, e^{j2\pi d(H-1)\sin\theta\sin\phi}]^{\mathrm{T}}, \\ \mathbf{a}_{\mathrm{W}}(\theta, \phi) &= [1, e^{j2\pi d \sin\theta \cos\phi}, \cdots, e^{j2\pi d(W-1)\sin\theta\cos\phi}]^{\mathrm{T}}, \end{aligned} \tag{6}$$

and $d = 1/2$ is the element spacing normalized by the wavelength; L_{TX} is the number of sub-paths in the TX-IRS channel; $\alpha_{\mathrm{TX},l}$, $\theta_{\mathrm{TX},l}$, and $\phi_{\mathrm{TX},l}$ are the complex gain, the elevation angle of arrival (AOA), and the azimuth AoA of the l-th sub-path in the TX-IRS channel, respectively. The parameters associated with the IRS-RX channel are defined similarly using the subscript "RX" and the term "angle of departure (AoD)". In addition, we assume a block fading channel that the channel remains constant within a coherence interval [32].

We consider the above channel model in three cases: full CSI (both $\mathbf{h}_{\mathrm{RX}}$ and $\mathbf{h}_{\mathrm{TX}}$ are available), partial CSI (only $\mathbf{h}_{\mathrm{RX}}$ is available), and no CSI (both $\mathbf{h}_{\mathrm{RX}}$ and $\mathbf{h}_{\mathrm{TX}}$ are unavailable).

1) Case I (Full CSI): For factory test, the diagnosis is usually performed in a microwave anechoic chamber where the wireless environment can be precisely controlled. Therefore, the channel vectors $\mathbf{h}_{\mathrm{RX}}$ and $\mathbf{h}_{\mathrm{TX}}$ can be calculated based on known spatial coordinates [13], [21], [22]. In this case, the wireless channel can be simplified as the free-space transmission model, where $L_{\mathrm{TX}} = L_{\mathrm{RX}} = 1$ and all channel parameters in (5) are available. In this way, the channel can be obtained without channel training.

We simply assume that $\mathbf{h}_{\mathrm{TX}}$ and $\mathbf{h}_{\mathrm{RX}}$ are perfectly known for this case and develop the diagnostic technique in Sec. III. The transmission model for this case is

$$\boxed{\text{Case I (full CSI)}: \quad y = \sum_{n=1}^{N} \Big[\underbrace{\mathbf{f} \circ \mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}}}_{\text{known}} \circ \underbrace{\mathbf{m}}_{\text{unknown}} \Big]_n + w.} \tag{7}$$

2) Case II (Partial CSI): For outdoor online diagnosis, the channels are subject to inevitable multipath scattering, which makes the above channel calculation unreliable. Fortunately, for some applications, the IRS is installed in a fixed location (e.g. walls, ceilings, etc.) and one can adopt a base station as the RX which is also fixed. Hence, both the IRS and the RX are located in an open area and the channel $\mathbf{h}_{\mathrm{RX}}$ between them can be assumed as a long-term stationary multipath-free channel [33], which can then be calculated using (5).

In this case, we simply assume that $\mathbf{h}_{\mathrm{RX}}$ is perfectly known, while $L_{\mathrm{TX}} \geq 1$ due to the

multipath scattering environment around the TX and the channel parameters associated with the TX are unknown. Note that since the mmWave channel is sparse in the angular domain [34], there are only a few dominant sub-paths in the channel and we have $L_{\mathrm{TX}} \ll \min(H, W)$. We will develop the diagnostic technique for this case in Sec. IV. The transmission model for this case is

$$\boxed{\text{Case II (partial CSI)}: \quad y = \sum_{n=1}^{N} \Big[\underbrace{\mathbf{f} \circ \mathbf{h}_{\mathrm{RX}}}_{\text{known}} \circ \underbrace{\mathbf{h}_{\mathrm{TX}} \circ \mathbf{m}}_{\text{unknown}} \Big]_n + w.} \tag{8}$$

3) Case III (No CSI): Finally, we consider the most general case where both channels are subject to multipath scattering and thus all channel parameters are assumed unknown, except for the assumption of sparse mmWave channel. We will develop the diagnostic technique for this case in Sec. V. The transmission model for this case is

$$\boxed{\text{Case III (no CSI)}: \quad y = \sum_{n=1}^{N} \Big[\underbrace{\mathbf{f}}_{\text{known}} \circ \underbrace{\mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}} \circ \mathbf{m}}_{\text{unknown}} \Big]_n + w.} \tag{9}$$

III. Case I: Diagnosis with Full CSI

In the case of full CSI, we assume both channels are known, i.e.,

$$y = \sum_{n=1}^{N} \Big[\underbrace{\mathbf{f} \circ \mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}}}_{\text{known}} \circ \underbrace{\mathbf{m}}_{\text{unknown}} \Big]_n + w. \tag{10}$$

Therefore, we can locally generate fault-free reference signal using known channel vectors. By subtracting the reference signal from the received signal, the impact of channel can be eliminated. The diagnosis problem is then formulated as a compressed sensing problem by exploiting the failure sparsity, which requires far fewer measurements than the total number of reflecting elements.

A. Problem Formulation

Absorbing known channel vectors into the weighing vector $\mathbf{f}$, the transmission model (10) can be rewritten as

$$y = \sum_{n=1}^{N} \big[\mathbf{f} \circ \mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}} \circ \mathbf{m} \big]_n + w = \big(\underbrace{\mathbf{f} \circ \mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}}}_{=\mathbf{f}_{\mathrm{h}}} \big)^{\mathrm{T}} \mathbf{m} + w, \tag{11}$$

where $\mathbf{f}_{\mathrm{h}} = \mathbf{f} \circ \mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}}$ is the equivalent sensing vector.

During K time slots within the channel coherence time, the IRS uses different weighting vectors to reflect the signal, yielding the following observation equation:

$$\mathbf{y} = \mathbf{F}_{\mathrm{h}}\mathbf{m} + \mathbf{w}, \tag{12}$$

where $\mathbf{y} = [y_1, \cdots, y_K]^{\mathrm{T}}$, $[\mathbf{F}_{\mathrm{h}}]_{k,:} = \mathbf{f}_{\mathrm{h},k}^{\mathrm{T}}$, and $\mathbf{w} = [w_1, \cdots, w_K]^{\mathrm{T}}$, in which $y_k = \mathbf{f}_{\mathrm{h},k}^{\mathrm{T}}\mathbf{m} + w_k$, $\mathbf{f}_{\mathrm{h},k}$, and w_k are the measured symbol, the sensing vector, and the noise associated with the k-th measurement, respectively.

Recall that the goal of diagnosis is to recover the failure mask vector $\mathbf{m}$ from the received signal $\mathbf{y}$. To recover $\mathbf{m}$ under the linear observation model (12), we can exploit the failure sparsity and turn (12) into a compressed sensing model.

Since $\mathbf{h}_{\mathrm{RX}}$ and $\mathbf{h}_{\mathrm{TX}}$ are known, the RX can locally generate the fault-free reference signal as

$$\mathbf{y}^{\mathrm{r}} = \mathbf{F}_{\mathrm{h}}\mathbf{1}_{N\times 1}. \tag{13}$$

Subtracting the fault-free reference signal $\mathbf{y}^{\mathrm{r}}$ from the received signal $\mathbf{y}$, we can obtain the differential signal

$$\mathbf{y}^{\mathrm{d}} = \mathbf{y} - \mathbf{y}^{\mathrm{r}} = \mathbf{F}_{\mathrm{h}}\mathbf{m} - \mathbf{F}_{\mathrm{h}}\mathbf{1}_{N\times 1} + \mathbf{w} = \mathbf{F}_{\mathrm{h}}(\underbrace{\mathbf{m} - \mathbf{1}_{N\times 1}}_{=\mathbf{x}}) + \mathbf{w}. \tag{14}$$

Due to the failure sparsity that most entries in $\mathbf{m}$ are unity, $\mathbf{x} = \mathbf{m} - \mathbf{1}_{N\times 1}$ is a sparse vector. Therefore, we can recover $\mathbf{x}$ from compressed measurements (i.e., $K \ll N$) and then retrieve $\mathbf{m}$, in which non-unity entries indicate faulty reflecting elements.[3]

B. Failure Mask Recovery

We propose to solve the following LASSO problem to recover $\mathbf{x}$ [35]:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y}^{\mathrm{d}} - \mathbf{F}_{\mathrm{h}}\mathbf{x}\|_2^2 + \lambda_1\|\mathbf{x}\|_1, \tag{15}$$

where l_2 and l_1 norms represent the approximation error and the penalty on sparsity, respectively, and λ_1 is a regularization parameter controlling the balance between them. Once $\mathbf{x}$ is recovered,

[3]Due to noise, hardware imperfections, and limited measurements, the recovered $\mathbf{x} = \mathbf{m} - \mathbf{1}_{N\times 1}$ may not be strictly sparse. One can use thresholding to identify faulty reflecting elements. For example, elements satisfying $|\hat{m}_n - 1| \geq \mathrm{th}$ can be identified as faulty, where th is a threshold and can be set to the maximum tolerable failure intensity. The diagnostic techniques proposed in Sec. IV and V can follow the same procedure to identify faulty reflecting elements.

we can retrieve $\mathbf{m}$ by

$$\hat{\mathbf{m}} = \hat{\mathbf{x}} + \mathbf{1}_{N\times 1}. \tag{16}$$

If each element of the weighting vector $\mathbf{f}$ is drawn uniformly randomly from a set containing all possible weighing coefficients, low mutual correlation of $\mathbf{F}_{\mathrm{h}}$ can be achieved and thus $\mathbf{m}$ can be accurately recovered from about $s \log N$ random incoherent measurements with high probability [36], where s is the number of faulty reflecting elements (i.e., the sparsity of $\mathbf{x}$). Therefore, the diagnostic time can be significantly shortened by exploiting the failure sparsity.

IV. Case II: Diagnosis with Partial CSI

In the case of partial CSI, we assume that only $\mathbf{h}_{\mathrm{RX}}$ is known and $\mathbf{h}_{\mathrm{TX}}$ is unavailable, i.e.,

$$y = \sum_{n=1}^{N} \Big[\underbrace{\mathbf{f} \circ \mathbf{h}_{\mathrm{RX}}}_{\text{known}} \circ \underbrace{\mathbf{h}_{\mathrm{TX}} \circ \mathbf{m}}_{\text{unknown}} \Big]_n + w. \tag{17}$$

Therefore, both channel vector $\mathbf{h}_{\mathrm{TX}}$ and failure mask $\mathbf{m}$ need to be estimated, which can be challenging since they are coupled. In this section, we jointly exploit the sparsities of mmWave channel and failure, and formulate the diagnosis problem as a cSLRMR problem.

A. Problem Formulation

The transmission model (17) can be reformulated as

$$y = \sum_{n=1}^{N} \big[\mathbf{f} \circ \mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}} \circ \mathbf{m}\big]_n + w = \sum_{i,j} \big[\mathbf{H}_{\mathrm{RX}} \circ \mathbf{F}_{\mathrm{mat}} \circ \mathbf{M}_{\mathrm{mat}} \circ \mathbf{H}_{\mathrm{TX}}\big]_{i,j} + w, \tag{18}$$

where $\mathbf{F}_{\mathrm{mat}} \in \mathbb{C}^{H\times W}$ and $\mathbf{M}_{\mathrm{mat}} \in \mathbb{C}^{H\times W}$ are the matrices reshaped from $\mathbf{f}$ and $\mathbf{m}$, respectively, i.e., $\mathrm{vec}(\mathbf{F}_{\mathrm{mat}}) = \mathbf{f}$ and $\mathrm{vec}(\mathbf{M}_{\mathrm{mat}}) = \mathbf{m}$; $\sum_{i,j}$ denotes the summation over all entries in the matrix.

The known channel matrix $\mathbf{H}_{\mathrm{RX}}$ can be absorbed into the weighting matrix $\mathbf{F}_{\mathrm{mat}}$ as

$$y = \sum_{i,j} \big[\mathbf{F}_{\mathrm{RX}} \circ \mathbf{M}_{\mathrm{mat}} \circ \mathbf{H}_{\mathrm{TX}}\big]_{i,j} + w, \tag{19}$$

where $\mathbf{F}_{\mathrm{RX}} = \mathbf{H}_{\mathrm{RX}} \circ \mathbf{F}_{\mathrm{mat}}$.

In addition, the multiplicative failure mask $\mathbf{M}_{\mathrm{mat}}$ can be equivalently replaced by an additive failure deviation $\mathbf{D}_{\mathrm{mat}}$ as

$$y = \sum_{i,j} \left[\mathbf{F}_{\mathrm{RX}} \circ (\underbrace{\mathbf{D}_{\mathrm{mat}} + \mathbf{H}_{\mathrm{TX}}}_{=\mathbf{X}}) \right]_{i,j} + w, \tag{20}$$

where $\mathbf{D}_{\mathrm{mat}} = \mathbf{H}_{\mathrm{TX}} \circ (\mathbf{M}_{\mathrm{mat}} - \mathbf{1}_{H\times W})$.

Equation (20) is a linear observation model where $\mathbf{F}_{\mathrm{RX}}$ and $\mathbf{X}$ are the configurable weighting matrix and the signal to be estimated, respectively, and we denote it as $y = \mathcal{F}_{\mathbf{F}_{\mathrm{RX}}}(\mathbf{X})$. We use K different weighting matrices to measure the signal $\mathbf{X}$ and obtain the observation equation

$$\mathbf{y} = \mathcal{F}(\mathbf{X}) + \mathbf{w}, \tag{21}$$

where $\mathbf{y} = [y_1, \cdots, y_K]^{\mathrm{T}}$ and $\mathbf{w} = [w_1, \cdots, w_K]^{\mathrm{T}}$; $y_k = \mathcal{F}_{\mathbf{F}_{\mathrm{RX},k}}(\mathbf{X}) + w_k$, $\mathbf{F}_{\mathrm{RX},k}$, and w_k are the measured symbol, the weighting matrix, and the noise associated with the k-th measurement, respectively.

The goal of diagnosis is to recover $\mathbf{D}_{\mathrm{mat}}$ and $\mathbf{H}_{\mathrm{TX}}$ simultaneously, and thus the failure mask $\mathbf{M}_{\mathrm{mat}}$ can be retrieved from $\mathbf{M}_{\mathrm{mat}} = \mathbf{D}_{\mathrm{mat}} \oslash \mathbf{H}_{\mathrm{TX}} + \mathbf{1}_{H\times W}$, in which non-unity entries indicate faulty reflecting elements. However, the channel and the failure deviation are coupled since both of them are unknown. To recover $\mathbf{H}_{\mathrm{TX}}$ and $\mathbf{D}_{\mathrm{mat}}$ simultaneously, we need to introduce *a-prior* information on them. Note that $\mathbf{D}_{\mathrm{mat}} = \mathbf{H}_{\mathrm{TX}} \circ (\mathbf{M}_{\mathrm{mat}} - \mathbf{1}_{H\times W})$ is a sparse matrix due to the failure sparsity that most entries in $\mathbf{M}_{\mathrm{mat}}$ are unity. According to the channel assumption in Sec. II-B2 (sparse mmWave channel), $\mathbf{H}_{\mathrm{TX}}$ is a low-rank matrix since $\mathrm{rank}(\mathbf{H}_{\mathrm{TX}}) \leq L_{\mathrm{TX}} \ll \min(H, W)$. Therefore, $\mathbf{X} = \mathbf{D}_{\mathrm{mat}} + \mathbf{H}_{\mathrm{TX}}$ is the sum of a sparse and a low-rank matrix, which allows $\mathbf{H}_{\mathrm{TX}}$ and $\mathbf{D}_{\mathrm{mat}}$ to be recovered simultaneously from compressed measurements thanks to the rank-sparsity incoherence [37], [38].

B. Failure Mask Recovery

One can solve the following optimization problem to recover $\mathbf{H}_{\mathrm{TX}}$ and $\mathbf{D}_{\mathrm{mat}}$ [39]:

$$\{\hat{\mathbf{H}}_{\mathrm{TX}}, \hat{\mathbf{D}}_{\mathrm{mat}}\} = \arg \min_{\mathbf{H}_{\mathrm{TX}}, \mathbf{D}_{\mathrm{mat}}} \|\mathbf{H}_{\mathrm{TX}}\|_* + \lambda_2 \|\mathbf{D}_{\mathrm{mat}}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathcal{F}(\mathbf{D}_{\mathrm{mat}} + \mathbf{H}_{\mathrm{TX}})\|_2 \leq \delta, \tag{22}$$

where $\|\cdot\|_*$ denotes the nuclear norm (sum of singular values) and represents the penalty on the rank; λ_2 is a regularization parameter controlling the balance between the low-rank component and the sparse component; δ is the noise level and is related to the SNR. The optimization

problem (22) is convex and can be solved efficiently using the algorithm proposed in [39]. Once $\mathbf{H}_{\mathrm{TX}}$ and $\mathbf{D}_{\mathrm{mat}}$ are recovered, the failure mask $\mathbf{m}$ can be retrieved by

$$\hat{\mathbf{m}} = \mathrm{vec}\left(\hat{\mathbf{M}}_{\mathrm{mat}}\right) = \mathrm{vec}\left(\hat{\mathbf{D}}_{\mathrm{mat}} \oslash \hat{\mathbf{H}}_{\mathrm{TX}} + \mathbf{1}_{H\times W}\right), \tag{23}$$

where $\oslash$ denotes Hadamard division (element-wise division).

V. Case III: Diagnosis without CSI

In the case of no CSI, we assume both $\mathbf{h}_{\mathrm{TX}}$ and $\mathbf{h}_{\mathrm{RX}}$ are unavailable, i.e.,

$$y = \sum_{n=1}^{N} \Big[\underbrace{\mathbf{f}}_{\text{known}} \circ \underbrace{\mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}} \circ \mathbf{m}}_{\text{unknown}} \Big]_n + w. \tag{24}$$

This is the most challenging case since both channels are coupled with the failure mask. In this section, we introduce a novel atomic norm to induce the sparsity of the cascaded channel $\mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}}$ and propose an ANM-based diagnostic technique to jointly recover the cascaded channel and the failure mask. In addition, we extend the proposed diagnostic technique to the multi-antenna setting. Two efficient algorithms are also developed to accelerate the diagnosis.

A. Problem Formulation

The transmission model (24) can be rewritten as

$$y = \sum_{n=1}^{N} \Big[\mathbf{f} \circ \underbrace{\mathbf{h}_{\mathrm{RX}} \circ \mathbf{h}_{\mathrm{TX}}}_{\text{cascaded channel } \mathbf{h}} \circ \mathbf{m} \Big]_n + w = \mathbf{f}^{\mathrm{T}}(\mathbf{h} \circ \mathbf{m}) + w = \mathbf{f}^{\mathrm{T}}(\mathbf{h} + \mathbf{d}) + w, \tag{25}$$

where $\mathbf{h} = \mathbf{h}_{\mathrm{TX}} \circ \mathbf{h}_{\mathrm{RX}}$ is the cascaded channel and $\mathbf{d} = \mathbf{h} \circ (\mathbf{m} - \mathbf{1}_{N\times 1})$ is the additive failure deviation.

Similarly, after K measurements, we have the observation equation

$$\mathbf{y} = \mathbf{F}(\mathbf{h} + \mathbf{d}) + \mathbf{w}, \tag{26}$$

where $\mathbf{y} = [y_1, \cdots, y_K]^{\mathrm{T}}$, $[\mathbf{F}]_{k,:} = \mathbf{f}_k^{\mathrm{T}}$, and $\mathbf{w} = [w_1, \cdots, w_K]^{\mathrm{T}}$, in which $y_k = \mathbf{f}_k^{\mathrm{T}}(\mathbf{h} + \mathbf{d}) + w_k$, $\mathbf{f}_k$, and w_k are the measured symbol, the sensing vector, and the noise associated with the k-th measurement, respectively.

Following a similar idea to that in Sec. IV, we need to introduce *a-prior* information on the unknown cascaded channel $\mathbf{h}$ to decouple it from the failure deviation $\mathbf{d}$. One possible

approach is to use the nuclear norm as adopted in the previously proposed diagnostic technique for Case II. However, the cascaded channel may no longer be low-rank since $\mathrm{rank}(\mathbf{H}_{\mathrm{RX}} \circ \mathbf{H}_{\mathrm{TX}}) \leq \mathrm{rank}(\mathbf{H}_{\mathrm{RX}})\mathrm{rank}(\mathbf{H}_{\mathrm{TX}}) \leq L_{\mathrm{RX}} L_{\mathrm{TX}}$, which means that even moderate multipath scattering may lead to a full-rank cascaded channel. To cope with this issue, in the following, we introduce a novel atomic norm as the sparsity-inducing norm of the cascaded channel.

B. The Atomic Norm of Cascaded Channel

Using the Kronecker product form of channel model (5), the cascaded channel $\mathbf{h}$ can be expressed as

$$\begin{aligned}
\mathbf{h} =& \mathbf{h}_{\mathrm{TX}} \circ \mathbf{h}_{\mathrm{RX}} \\
=& \left(\sum_{l=1}^{L_{\mathrm{TX}}} \alpha_{\mathrm{TX},l} \Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{TX},l}, \phi_{\mathrm{TX},l}) \otimes \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{TX},l}, \phi_{\mathrm{TX},l}) \Big) \right) \circ \left(\sum_{l=1}^{L_{\mathrm{RX}}} \alpha_{\mathrm{RX},l} \Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l}) \otimes \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l}) \Big) \right) \\
=& \sum_{i=1}^{L_{\mathrm{TX}}} \sum_{j=1}^{L_{\mathrm{RX}}} \alpha_{\mathrm{TX},i} \alpha_{\mathrm{RX},j} \Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{TX},i}, \phi_{\mathrm{TX},i}) \otimes \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{TX},i}, \phi_{\mathrm{TX},i}) \Big) \circ \Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{RX},j}, \phi_{\mathrm{RX},j}) \otimes \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{RX},j}, \phi_{\mathrm{RX},j}) \Big) \\
=& \sum_{i=1}^{L_{\mathrm{TX}}} \sum_{j=1}^{L_{\mathrm{RX}}} \alpha_{\mathrm{TX},i} \alpha_{\mathrm{RX},j} \Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{TX},i}, \phi_{\mathrm{TX},i}) \circ \mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{RX},j}, \phi_{\mathrm{RX},j}) \Big) \otimes \Big(\mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{TX},i}, \phi_{\mathrm{TX},i}) \circ \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{RX},j}, \phi_{\mathrm{RX},j}) \Big) \\
=& \sum_{i=1}^{L_{\mathrm{TX}}} \sum_{j=1}^{L_{\mathrm{RX}}} \alpha_{\mathrm{TX},i} \alpha_{\mathrm{RX},j} \Big(\mathbf{a}'_{\mathrm{W}}(\theta_{\mathrm{TX},i}, \phi_{\mathrm{TX},i}, \theta_{\mathrm{RX},j}, \phi_{\mathrm{RX},j}) \otimes \mathbf{a}'_{\mathrm{H}}(\theta_{\mathrm{TX},i}, \phi_{\mathrm{TX},i}, \theta_{\mathrm{RX},j}, \phi_{\mathrm{RX},j}) \Big),
\end{aligned} \tag{27}$$

where

$$\begin{aligned}
\mathbf{a}'_{\mathrm{W}}(\theta_i, \phi_i, \theta_j, \phi_j) &= \mathbf{a}_{\mathrm{W}}(\theta_i, \phi_i) \circ \mathbf{a}_{\mathrm{W}}(\theta_j, \phi_j) \\
&= [1, e^{j2\pi d(\sin\theta_i \cos\phi_i + \sin\theta_j \cos\phi_j)}, \cdots, e^{j2\pi d(W-1)(\sin\theta_i \cos\phi_i + \sin\theta_j \cos\phi_j)}]^{\mathrm{T}}, \\
\mathbf{a}'_{\mathrm{H}}(\theta_i, \phi_i, \theta_j, \phi_j) &= \mathbf{a}_{\mathrm{H}}(\theta_i, \phi_i) \circ \mathbf{a}_{\mathrm{H}}(\theta_j, \phi_j) \\
&= [1, e^{j2\pi d(\sin\theta_i \sin\phi_i + \sin\theta_j \sin\phi_j)}, \cdots, e^{j2\pi d(H-1)(\sin\theta_i \sin\phi_i + \sin\theta_j \sin\phi_j)}]^{\mathrm{T}}.
\end{aligned} \tag{28}$$

Therefore, the cascaded channel can be treated as a channel that has $L = L_{\mathrm{TX}} L_{\mathrm{RX}}$ equivalent sub-paths. To lighten the notation, we rewrite the cascaded channel $\mathbf{h}$ as

$$\mathbf{h} = \sum_{l=1}^{L} \alpha_l \mathbf{a}_{\mathrm{W}}(g_{\mathrm{W},l}) \otimes \mathbf{a}_{\mathrm{H}}(g_{\mathrm{H},l}) = \sum_{l=1}^{L} \alpha_l \mathbf{a}(\mathbf{g}_l), \tag{29}$$

where

$$\begin{aligned}\mathbf{a}(\mathbf{g}_l) &= \mathbf{a}_{\mathrm{W}}(g_{\mathrm{W},l}) \otimes \mathbf{a}_{\mathrm{H}}(g_{\mathrm{H},l}),\\ \mathbf{a}_{\mathrm{W}}(g_{\mathrm{W},l}) &= [1, e^{j2\pi d g_{\mathrm{W},l}}, \cdots, e^{j2\pi d(W-1)g_{\mathrm{W},l}}]^{\mathrm{T}},\\ \mathbf{a}_{\mathrm{H}}(g_{\mathrm{H},l}) &= [1, e^{j2\pi d g_{\mathrm{H},l}}, \cdots, e^{j2\pi d(H-1)g_{\mathrm{H},l}}]^{\mathrm{T}},\end{aligned} \tag{30}$$

and $\alpha_l \in \{\alpha_{\mathrm{TX},i}\alpha_{\mathrm{RX},j}, (i = 1, \cdots, L_{\mathrm{TX}}, j = 1 \cdots, L_{\mathrm{RX}})\}$, $\mathbf{g}_l = [g_{\mathrm{W},l}, g_{\mathrm{H},l}]^{\mathrm{T}}$, $g_{\mathrm{W},l} \in \{\sin\theta_i \cos\phi_i + \sin\theta_j \cos\phi_j, (i = 1, \cdots, L_{\mathrm{TX}}, j = 1 \cdots, L_{\mathrm{RX}})\}$, $g_{\mathrm{H},l} \in \{\sin\theta_i \sin\phi_i + \sin\theta_j \sin\phi_j, (i = 1, \cdots, L_{\mathrm{TX}}, j = 1 \cdots, L_{\mathrm{RX}})\}$.

From the above derivation, we can observe that the cascaded channel $\mathbf{h}$ is the sum of L 2D frequency components, which has a sparse representation in the 2D Fourier dictionary since $L \ll N$. To represent $\mathbf{h}$ in a sparse form, a conventional approach is to discretize the angular domain into a grid and form a 2D discrete Fourier transform (DFT) dictionary [40]. However, this approach incurs off-grid error since frequency components (AOAs/AoDs) do not lie exactly on the grid, and hence the performance of sparse recovery algorithms may degrade considerably [41]. Although increasing the grid density can alleviate the off-grid problem to a certain extent, the computational complexity also increases dramatically, especially under the 2D setting. To bypass this issue, we adopt the atomic norm as the sparsity promoter of the cascaded channel, which does not discretize the angular domain and thus no off-grid error will be induced. A comprehensive survey on the atomic norm can be found in [42].

Recall that the cascaded channel $\mathbf{h}$ has a sparse representation in the 2D Fourier dictionary. To be more specific, $\mathbf{h}$ is sparse in the continuous 2D Fourier dictionary[4]

$$\mathcal{A} = \{\mathbf{a}(\mathbf{g}) | \mathbf{g} \in [-1, 1) \times [-1, 1)\}, \tag{31}$$

which contains the building blocks (atoms) of the signal $\mathbf{h}$. Analogous to the definition of l_1 norm, the atomic norm of $\mathbf{h}$ over the atom set $\mathcal{A}$ is defined as

$$\|\mathbf{h}\|_{\mathcal{A}} = \inf\left\{\sum_l |\alpha_l| \middle| \mathbf{h} = \sum_l \alpha_l \mathbf{a}_l, \ \ \mathbf{a}_l \in \mathcal{A}\right\}, \tag{32}$$

which is the minimum cost (in terms of l_1 norm measure) of representing a sparse signal.

Although the definition of atomic norm is simple and clear, it can not be directly minimized since it involves the optimization of infinite-dimensional variables. To minimize the atomic norm

[4]Strictly speaking, $\mathbf{g}$ lies in the range $[-2, 2) \times [-2, 2)$. However, due to the periodic property of the exponent, only one period (i.e., $[-1, 1) \times [-1, 1)$) is actually required.

(32) in a tractable manner, one can exploit the Vandermonde structure of signal $\mathbf{h}$ and solve the following positive semidefinite program [43]:

$$\arg\min_{\substack{\mathbf{u}\in\mathbb{C}^{N_u}\\ v\in\mathbb{R}}} \frac{1}{2}\left(\frac{1}{N}\mathrm{tr}\big(T(\mathbf{u})\big)+v\right) \quad \text{s.t.} \quad \begin{bmatrix} T(\mathbf{u}) & \mathbf{h} \\ \mathbf{h}^{\mathrm{H}} & v \end{bmatrix} \succeq 0, \tag{33}$$

where $T(\mathbf{u})$ is a twofold Toeplitz operator, which maps a given vector

$$\mathbf{u} = [\underbrace{u_{0,0}, u_{0,1}, \cdots, u_{0,W-1}}_{\text{block } 0}, \underbrace{u_{1,-(W-1)}, \cdots, u_{1,W-1}}_{\text{block } 1}, \cdots\cdots, \underbrace{u_{H-1,-(W-1)}, \cdots, u_{H-1,W-1}}_{\text{block } H-1}] \tag{34}$$

of length $N_u = (H-1)(2W-1)+W$ into a $H \times H$ block Toeplitz matrix

$$T(\mathbf{u}) = \begin{bmatrix} \mathbf{T}_0 & \mathbf{T}_1^{\mathrm{H}} & \cdots & \mathbf{T}_{H-1}^{\mathrm{H}} \\ \mathbf{T}_1 & \mathbf{T}_0 & \cdots & \mathbf{T}_{H-2}^{\mathrm{H}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{T}_{H-1} & \mathbf{T}_{H-2} & \cdots & \mathbf{T}_0 \end{bmatrix}, \tag{35}$$

with each block $\mathbf{T}_i$ $(0 \le i < H)$ being a $W \times W$ Toeplitz matrix

$$\mathbf{T}_i = \begin{bmatrix} u_{i,0} & u_{i,-1} & \cdots & u_{i,-(W-1)} \\ u_{i,1} & u_{i,0} & \cdots & u_{i,-(W-2)} \\ \vdots & \vdots & \ddots & \vdots \\ u_{i,W-1} & u_{i,W-2} & \cdots & u_{i,0} \end{bmatrix}. \tag{36}$$

Note that (33) is just an approximation for minimizing the atomic norm since there is a fundamental difficulty in generalizing the classical Caratheodory's theorem to higher dimensions [43]. Nonetheless, (33) exhibits excellent empirical performance as illustrated later in numerical simulations (Sec. VI).

C. Failure Mask Recovery

Using the atomic norm and the l_1 norm as the sparsity promoters for the cascaded channel $\mathbf{h}$ and the failure deviation $\mathbf{d}$, respectively, we can simultaneously recover $\mathbf{h}$ and $\mathbf{d}$ by solving the following joint sparse recovery problem:

$$\{\hat{\mathbf{h}}, \hat{\mathbf{d}}\} = \arg\min_{\mathbf{h},\mathbf{d}} \frac{1}{2}\|\mathbf{y} - \mathbf{F}(\mathbf{h}+\mathbf{d})\|_2^2 + \tau_1\|\mathbf{h}\|_{\mathcal{A}} + \lambda_3\|\mathbf{d}\|_1, \tag{37}$$

where τ_1 and λ_3 are regularization parameters controlling the penalty on the sparsity of $\mathbf{h}$ and $\mathbf{d}$, respectively.

According to (33), the optimization problem (37) can be rewritten as

$$\{\hat{\mathbf{h}}, \hat{\mathbf{d}}\} = \arg \min_{\mathbf{h},\mathbf{d},\mathbf{u},v} \frac{1}{2}\|\mathbf{y} - \mathbf{F}(\mathbf{h}+\mathbf{d})\|_2^2 + \frac{\tau_1}{2}\left(\frac{1}{N}\mathrm{tr}\big(T(\mathbf{u})\big) + v\right) + \lambda_3\|\mathbf{d}\|_1, \ \text{ s.t. } \begin{bmatrix} T(\mathbf{u}) & \mathbf{h} \\ \mathbf{h}^{\mathrm{H}} & v \end{bmatrix} \succeq 0. \tag{38}$$

Once $\mathbf{h}$ and $\mathbf{d}$ are recovered, the failure mask $\mathbf{m}$ can be retrieved by

$$\hat{\mathbf{m}} = \hat{\mathbf{d}} \oslash \hat{\mathbf{h}} + \mathbf{1}_{N\times 1}. \tag{39}$$

The optimization problem (38) is a positive semidefinite program and can be solved by off-the-shelf solvers like SDPT3 in the CVX toolbox, as suggested in numerous related works on atomic norm. SDPT3 is based on the interior point method, which tends to be extremely time-consuming even for medium-scale systems. In Appendix A, we develop an efficient algorithm for solving (38), which is based on the ADMM algorithm. One can refer to [35] for a thorough survey of the ADMM algorithm.

D. Extension to Multiple Antennas at RX

Recall that in this case we assume that there may be multipath scattering in the IRS-RX channel, which can be exploited by a multi-antenna RX to provide diversity gain. We consider a single antenna at the TX and N_{RX} antennas at the RX. For simplicity, we assume that antennas at the RX form a uniform linear array (ULA). Note that the following proposed diagnostic technique can be easily extended to other array structures like UPA.

With multiple antennas at the RX, the transmission model (4) can be extended to

$$\mathbf{y} = \mathbf{H}'_{\mathrm{RX}}(\mathrm{diag}(\mathbf{f})\mathrm{diag}(\mathbf{m}))\mathbf{h}_{\mathrm{TX}} + \mathbf{w}, \tag{40}$$

where $\mathbf{y} \in \mathbb{C}^{N_{\mathrm{RX}}}$ is the received signal, $\mathbf{w} \in \mathbb{C}^{N_{\mathrm{RX}}}$ is the noise. $\mathbf{H}'_{\mathrm{RX}} \in \mathbb{C}^{N_{\mathrm{RX}}\times N}$ is the channel matrix between the RX and the IRS, which can be expressed as

$$\mathbf{H}'_{\mathrm{RX}} = \sum_{l=1}^{L_{\mathrm{RX}}} \alpha_{\mathrm{RX},l}\mathbf{a}_{\mathrm{RX}}(\theta'_{\mathrm{RX},l})\Big(\mathbf{a}_{\mathrm{W}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l}) \otimes \mathbf{a}_{\mathrm{H}}(\theta_{\mathrm{RX},l}, \phi_{\mathrm{RX},l})\Big)^{\mathrm{T}}, \tag{41}$$

where $\mathbf{a}_{\mathrm{RX}}(\theta) = [1, e^{j2\pi d\sin\theta}, \cdots, e^{j2\pi d(N_{\mathrm{RX}}-1)\sin\theta}]^{\mathrm{T}}$ and $\theta'_{\mathrm{RX},l}$ is the AoA of the l-th sub-path in the IRS-RX channel. Note that each entry in $\mathbf{y}$ can be expressed in the form of (25) (transmission

model with single-antenna RX) with different channel vectors $\mathbf{h}_{\mathrm{RX}}$. Denoting the received symbol and the channel vector associated with the i-th antenna at the RX as y_i and $\mathbf{h}_{\mathrm{RX},i}$, respectively, we can represent $\mathbf{y}$ and $\mathbf{H}'_{\mathrm{RX}}$ as

$$\mathbf{y} = [y_1, \cdots, y_{N_{\mathrm{RX}}}]^{\mathrm{T}}, \quad \mathbf{H}'_{\mathrm{RX}} = [\mathbf{h}_{\mathrm{RX},1}, \cdots, \mathbf{h}_{\mathrm{RX},N_{\mathrm{RX}}}]^{\mathrm{T}}. \tag{42}$$

According to the observation equation (26), after K measurements, the i-th antenna at the RX observes a vector of measurements

$$\mathbf{y}_i = \mathbf{F}(\mathbf{h}_i + \mathbf{d}_i) + \mathbf{w}_i, \tag{43}$$

where $\mathbf{h}_i = \mathbf{h}_{\mathrm{TX}} \circ \mathbf{h}_{\mathrm{RX},i}$, $\mathbf{d}_i = \mathbf{h}_i \circ (\mathbf{m} - \mathbf{1}_{N\times 1})$, and $\mathbf{w}_i$ are the cascaded channel, the failure deviation, and the measurement noise associated with the i-th antenna at the RX, respectively.

Stacking the measurement vectors from all RX antennas into a matrix, we obtain the following MMV model:

$$\mathbf{Y} = \mathbf{F}(\mathbf{H} + \mathbf{D}) + \mathbf{W}, \tag{44}$$

where $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_{N_{\mathrm{RX}}}]$, $\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_{N_{\mathrm{RX}}}]$, $\mathbf{D} = [\mathbf{d}_1, \cdots, \mathbf{d}_{N_{\mathrm{RX}}}]$, and $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_{N_{\mathrm{RX}}}]$.

Note that all column vectors in $\mathbf{H}$ share the same sparse support (i.e., location of non-zero components) since AoAs/AoDs are the same with respect to each receiving channel vector $\mathbf{h}_{\mathrm{RX},i}$. Besides, due to multipath scattering in the IRS-RX channel, $\mathbf{H}$ is of rank L_{RX}, which provides a diversity gain of order L_{RX}. The failure deviation $\mathbf{D}$ also reserves such diversity gain since it is associated with $\mathbf{H}$ as $\mathbf{d}_i = \mathbf{h}_i \circ (\mathbf{m} - \mathbf{1}_{N\times 1})$. Therefore, we can exploit the diversity gain provided by multipath scattering in the IRS-RX channel to improve the recovery performance.

The MMV-form sparsity-inducing norms for the channel $\mathbf{H}$ and the failure deviation $\mathbf{D}$ are introduced as follows. The MMV-form atomic norm is defined as [44]

$$\|\mathbf{H}\|_{\mathcal{A}_{\mathrm{MMV}}} = \inf\left\{\sum_l |\alpha_l| \middle| \mathbf{H} = \sum_l \alpha_l \mathbf{A}_l, \ \ \mathbf{A}_l \in \mathcal{A}_{\mathrm{MMV}}\right\}, \tag{45}$$

where

$$\mathcal{A}_{\mathrm{MMV}} = \{\mathbf{a}(\mathbf{g})\mathbf{b}^{\mathrm{T}} | \mathbf{g} \in [-1, 1) \times [-1, 1), \|\mathbf{b}\|_2 = 1\} \tag{46}$$

and $\mathbf{b} \in \mathbb{C}^{N_{\mathrm{RX}}}$. Similar to the single measurement vector (SMV) model (33), the MMV-form

atomic norm of $\mathbf{H}$ can be minimized by solving the positive semidefinite program [44]

$$\arg \min_{\substack{\mathbf{u}\in\mathbb{C}^{N_u} \\ \mathbf{V}\in\mathbb{C}^{N_{\mathrm{RX}}\times N_{\mathrm{RX}}}}} \frac{1}{2}\left(\frac{1}{N}\mathrm{tr}\big(T(\mathbf{u})\big)+\mathrm{tr}(\mathbf{V})\right) \quad \text{s.t.} \quad \begin{bmatrix} T(\mathbf{u}) & \mathbf{H} \\ \mathbf{H}^{\mathrm{H}} & \mathbf{V} \end{bmatrix} \succeq 0. \tag{47}$$

For the MMV-form sparsity-inducing norm of $\mathbf{D}$, we adopt the notation [45]

$$\|\mathbf{D}\|_{2,1} = \sum_{n=1}^{N} \|[\mathbf{D}]_{n,:}\|_2, \tag{48}$$

where $[\mathbf{D}]_{n,:}$ denotes the n-th row of $\mathbf{D}$. It can be observed that the MMV-form atomic norm $\|\cdot\|_{\mathcal{A}_{\mathrm{MMV}}}$ and sparsity-inducing norm $\|\cdot\|_{2,1}$ reduce exactly to the SMV-form atomic norm $\|\cdot\|_{\mathcal{A}}$ and l_1 norm $\|\cdot\|_1$ when the number of RX antennas is $N_{\mathrm{RX}}=1$, respectively.

Using the above sparsity-inducing norms, the optimization problem for jointly recovering $\mathbf{H}$ and $\mathbf{D}$ can be expressed as

$$\{\hat{\mathbf{H}},\hat{\mathbf{D}}\} = \arg\min_{\mathbf{H},\mathbf{D}} \frac{1}{2}\|\mathbf{Y}-\mathbf{F}(\mathbf{H}+\mathbf{D})\|_{\mathrm{F}}^2 + \tau_2\|\mathbf{H}\|_{\mathcal{A}_{\mathrm{MMV}}} + \lambda_4\|\mathbf{D}\|_{2,1}, \tag{49}$$

which admits positive semidefinite program

$$\begin{aligned} \{\hat{\mathbf{H}},\hat{\mathbf{D}}\} = \arg \min_{\mathbf{H},\mathbf{D},\mathbf{u},\mathbf{V}} & \frac{1}{2}\|\mathbf{Y}-\mathbf{F}(\mathbf{H}+\mathbf{D})\|_{\mathrm{F}}^2 + \frac{\tau_2}{2}\left(\frac{1}{N}\mathrm{tr}\big(T(\mathbf{u})\big)+\mathrm{tr}(\mathbf{V})\right) + \lambda_4\|\mathbf{D}\|_{2,1}, \\ \text{s.t.} \quad & \begin{bmatrix} T(\mathbf{u}) & \mathbf{H} \\ \mathbf{H}^{\mathrm{H}} & \mathbf{V} \end{bmatrix} \succeq 0. \end{aligned} \tag{50}$$

Once $\mathbf{H}$ and $\mathbf{D}$ are recovered, the failure mask $\mathbf{m}$ can be retrieved by

$$\hat{\mathbf{m}} = \mathrm{mean}\left(\hat{\mathbf{D}} \oslash \hat{\mathbf{H}} + \mathbf{1}_{N\times N_{\mathrm{RX}}}\right), \tag{51}$$

where $\mathrm{mean}(\cdot)$ denotes the mean of each row.

The optimization problem (50) also enjoys an efficient ADMM-based algorithm, which is summarized in Appendix B.

It is worth noting that no diversity gain can be obtained if the TX equips multiple antennas, since the signal impinging on the IRS is fixed, no matter how many antennas are at the TX. We can also consider using different precoding matrices to transmit the test symbol. However, it will lead to a cascaded channel (signal to be estimated) that varies with each measurement, which even makes the diagnosis completely failed.

TABLE II
DEFAULT SIMULATION PARAMETERS.

Parameter	Symbol	Setting
Number of reflecting elements	$N = H \times W$	16×16
Resolution of phase weighing		2-bit, $[1+1j, 1-1j, -1+1j, -1-1j]/\sqrt{2}$
Distribution of sub-path gain	α_{TX}, α_{RX}	$\alpha_{\mathrm{TX}} \sim \mathcal{CN}(0, 1/L_{\mathrm{TX}})$, $\alpha_{\mathrm{RX}} \sim \mathcal{CN}(0, 1/L_{\mathrm{RX}})$
Distribution of sub-path AoA/AoD	θ, ϕ	$\theta \sim U(0, 2\pi)$, $\phi \sim U(0, 2\pi)$
Distribution of failure mask	$m = \eta e^{j\kappa}$	$\eta \sim U(0, 1)$, $\kappa \sim U(0, 2\pi)$
Regularization Parameters	λ_1, λ_2, λ_3, λ_4	$\lambda_1 = 0.65K/\mathrm{SNR}$, $\lambda_2 = 0.35$, $\lambda_3 = 0.006K$, $\lambda_4 = 0.006K/N_{\mathrm{RX}}$
	τ_1, τ_2	$\tau_1 = 0.004K$, $\tau_2 = 0.004K$

VI. NUMERICAL SIMULATIONS

This section presents several illustrative instances and numerical simulations to demonstrate the performance of proposed diagnostic techniques. The MATLAB codes for implementing all the proposed algorithms are available online.[5]

A. Parameter Setting and Performance Metric

Default simulation parameters are listed in Table II. We adopt the NMSE (normalized mean squared error) of recovered failure masks as the performance metric, which is defined as

$$\mathrm{NMSE} = \frac{\|\mathbf{m} - \hat{\mathbf{m}}\|_2^2}{\|\mathbf{m}\|_2^2}. \tag{52}$$

Besides, the measurement number shown in the figures is normalized by the number of reflecting elements N (nominal signal dimension).

B. Illustrative Instances

We first provide three instances in Fig. 3 to illustrate the main ideas of our proposed diagnostic techniques, where SNR$=$ 20dB and $K = 0.8N$. For better visualization, all vectors are reshaped into matrices with size $H \times W$ in order to keep them consistent with the array structure. The diagnostic results are presented in the form of failure intensity, which is defined as $|m_n - 1|$. Besides, only the real part of the channel is displayed.

[5]https://github.com/DestinationSR/IRSdiagnosis.

The instance of the diagnosis with full CSI is shown in Fig. 3(a). The main idea of the diagnosis with full CSI is to compare the received signal with the reference signal generated from known channel vectors. Therefore, non-zero entries in the sparse differential signal indicate faulty reflecting elements, which can be recovered from compressed measurements.

The instance of the diagnosis with partial CSI is shown in Fig. 3(b), where the number of sub-paths in the TX-IRS channel is set to $L_{\mathrm{TX}} = 4$. In the case of partial CSI, the channel fading and the failure are coupled and hence difficult to distinguish. Note that the unknown channel matrix is low-rank since the number of sub-paths is limited. Therefore, by jointly exploiting the channel low-rank property and the failure sparsity, $\mathbf{H}_{\mathrm{TX}}$ and $\mathbf{D}$ can be recovered simultaneously using the cSLRMR algorithm.

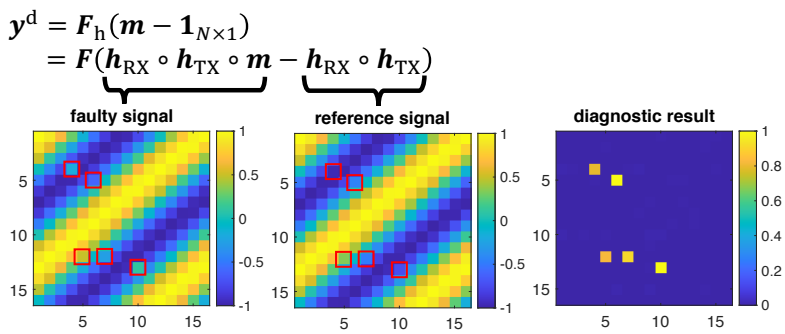
The instance of the diagnosis without CSI is shown in Fig. 3(c), where the number of sub-paths is $L_{\mathrm{TX}} = L_{\mathrm{RX}} = 4$. We can see that the cascaded channel is not low-rank anymore since $\mathrm{rank}(\mathbf{H}_{\mathrm{RX}} \circ \mathbf{H}_{\mathrm{TX}}) = L_{\mathrm{RX}} L_{\mathrm{TX}} = 16$, while the 2D spectrum indicates that the cascaded channel can be sparsely represented in the 2D Fourier dictionary. Therefore, using the atomic norm as the sparsity promoter, the cascaded channel and the failure mask can be recovered simultaneously.

C. Performance Evaluation

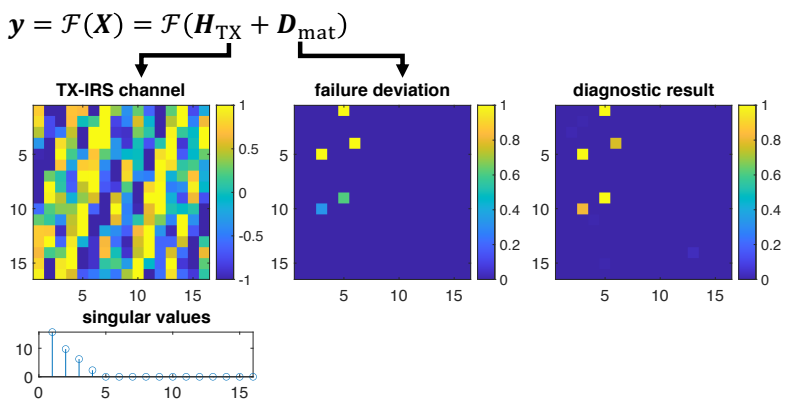
In the following, we evaluate the performance of proposed diagnostic techniques. The results shown are derived from 500 independent Monte-Carlo simulations.

We first evaluate the performance of diagnosis with full CSI. The NMSE versus the normalized number of measurements under different numbers of faulty reflecting elements and SNRs is shown in Fig. 4. It can be observed that about $K = 0.3N$ measurements are sufficient for the diagnosis with full CSI. Besides, a higher SNR and/or a lower number of faulty reflecting elements will result in better performance. We can also observe that the number of faulty reflecting elements has a major impact on the performance when the number of measurements is extremely low ($K = 0.1N$), while the SNR dominates the lower bound of the NMSE when the measurement number becomes sufficiently large.

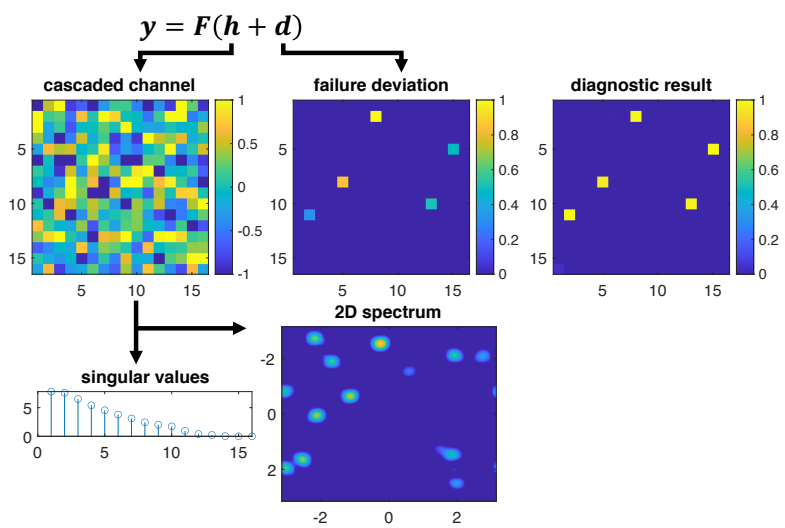
Next, we evaluate the performance of diagnosis with partial CSI. Setting the number of sub-paths in the TX-IRS channel to $L_{\mathrm{TX}} = 4$, the NMSE versus the normalized number of measurements under different numbers of faulty reflecting elements and SNRs is shown in Fig. 5. Comparing with the diagnosis with full CSI, the diagnosis with partial CSI requires more measurements to deliver satisfactory performance. In addition, the impact of SNR appears when



(a) Diagnosis with full CSI, where red squares indicate the location of faulty reflecting elements. The impact of channel can be eliminated by comparing the received signal with the known reference signal, and thus the differential signal is a sparse signal that can be recovered from compressed measurements.



(b) Diagnosis with partial CSI. Although no reference signal is available, we can jointly exploit the low-rank property of the channel $\mathbf{H}_{\mathrm{TX}}$ and the sparsity of the failure deviation $\mathbf{D}$ to recover them simultaneously.



(c) Diagnosis without CSI. The cascaded channel $\mathbf{h}$ is not low-rank anymore, but it can be sparsely represented in the 2D Fourier dictionary. Using the atomic norm as the sparsity promoter, both cascaded channel and failure mask can be recovered.

Fig. 3. Instances of proposed diagnostic techniques.

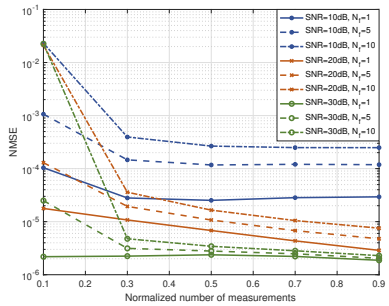


Fig. 4. NMSE performance of the diagnosis with full CSI, where N_{f} is the number of faulty reflecting elements.

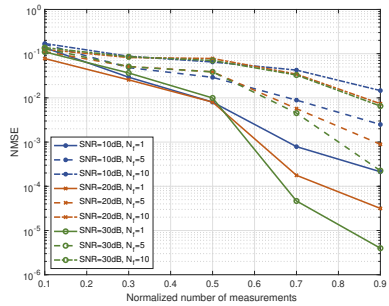


Fig. 5. NMSE performance of the diagnosis with partial CSI, where N_{f} is the number of faulty reflecting elements.

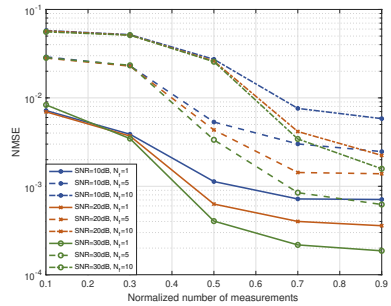


Fig. 6. NMSE performance of the diagnosis without CSI, where N_{f} is the number of faulty reflecting elements.

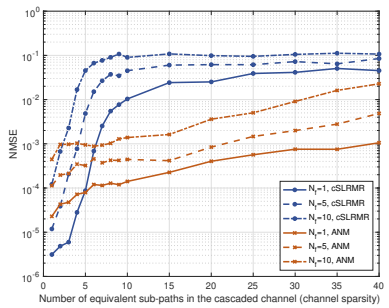


Fig. 7. Comparison of the NMSE performance between cSLRMR-based diagnosis and ANM-based diagnosis, where N_{f} is the number of faulty reflecting elements.

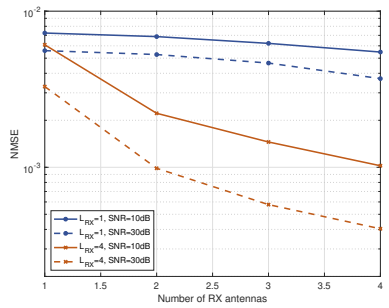


Fig. 8. NMSE performance of the ANM-based diagnostic technique using multiple antennas at the RX.

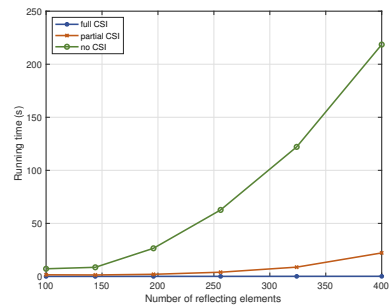


Fig. 9. Comparison of the running time between different diagnostic techniques.

the number of measurements is sufficiently large, while the performance is sensitive to the number of faulty reflecting elements under all numbers of measurements.

In the following, we evaluate the performance of diagnosis without CSI. Setting the number of sub-paths to $L_{\mathrm{TX}} = L_{\mathrm{RX}} = 4$, we plot the NMSE versus the normalized number of measurements under different numbers of faulty reflecting elements and SNRs in Fig. 6. Similar to the previous results, the impact of SNR becomes significant when the measurements are sufficient, while the number of faulty reflecting elements has a major impact across all numbers of measurements. Besides, the diagnosis without CSI has better performance in the regime of low measurement number compared with the diagnosis with partial CSI. The reason may be that the ANM-based diagnosis can tolerate a higher channel sparsity than the cSLRMR-based diagnosis, which reduces the requirement on the measurement number.

To illustrate the tolerance against multipath scattering of the ANM-based diagnostic technique

(diagnosis without CSI), Fig. 7 compares the NMSE performance of the ANM-based diagnostic technique and the diagnosis technique using the cSLRMR under different numbers of channel sub-paths and faulty reflecting elements, where $K = 0.8N$, SNR $= 30$dB, and all channels are assumed unknown. It can be observed that the diagnostic technique using the cSLRMR applies to the case without CSI provided that the cascaded channel is sufficiently sparse. However, as the number of sub-paths increases, the performance of cSLRMR-based diagnostic technique degrades rapidly. In contrast, the ANM-based diagnostic technique shows strong robustness against the multipath scattering and performs well even when the channel matrix is full-rank ($L \geq \min(H, W)$). Since the failure sparsity increases as the number of faulty reflecting elements grows, all diagnostic techniques achieve better NMSE performance when the fault proportion is low.

In Fig. 8, we evaluate the effect of multiple antennas at the RX under different numbers of sub-paths in the IRS-RX channel, where $K = 0.5N$, $N_{\mathrm{f}} = 5$, and the number of equivalent sub-paths in the cascaded channel is set to $L = L_{\mathrm{TX}} L_{\mathrm{RX}} = 16$ to maintain a constant channel sparsity. It can be observed that multiple antennas at the RX significantly improve the performance when there is sufficient multipath scattering in the IRS-RX channel ($L_{\mathrm{RX}} > 1$). The comparison under different SNRs also verifies that the effect of multiple antennas is mainly the diversity gain rather than the SNR gain.

Finally, we evaluate the time complexity of different diagnostic techniques. Fig. 9 plots the running time versus the number of reflecting elements N, where all MATLAB codes are executed on a desktop PC with an Intel Core i7-4790K CPU running at 4.0GHz. The results show that there is a considerable gap in running time between different diagnostic techniques, just as they have different requirements of the CSI. Although the ANM-based diagnostic technique (diagnosis without CSI) also applies to the case of full/partial CSI, we recommend selecting a proper diagnostic technique based on the CSI availability to avoid unnecessary computation.

VII. Conclusions

In this paper, we have developed several diagnostic techniques for IRS system. We have shown that the CSI availability is the most challenging factor for diagnosis since the channel and the failure are intimately coupled. To perform diagnosis in different scenarios, three diagnostic techniques have been developed, which have different requirements for CSI as well as different computational complexities. In the first case where full CSI is available, the impact of the channel

can be eliminated by subtracting the reference signal from the received signal, and the diagnosis problem has been formulated as a compressed sensing problem by exploiting the failure sparsity. In the second case where only partial CSI is available, the channel-failure decoupling has been achieved by jointly exploiting the low-rank property of the mmWave channel and the sparsity of the failure. The third case where no CSI is available is the most challenging since the cascaded channel may not be sparse in conventional dictionaries, which is a unique challenge introduced by IRS hardware characteristics. To cope with this issue, we have shown that the cascaded channel can be sparsely represented in the continuous 2D Fourier dictionary, and introduced the atomic norm to induce the sparsity of cascaded channel. Then, an ANM-based diagnostic technique has been proposed to decouple the cascaded channel and the failure. Two ADMM-based algorithms have also been proposed to accelerate the computation. Numerical simulations have demonstrated that the proposed diagnostic techniques can deliver satisfactory performance.

APPENDIX A
AN EFFICIENT ALGORITHM FOR SOLVING (38)

We first rewrite (38) into an ADMM-compliant form:

$$\{\hat{\mathbf{h}}, \hat{\mathbf{d}}\} = \arg\min_{\mathbf{h},\mathbf{d},\mathbf{u},v} \frac{1}{2}\|\mathbf{y} - \mathbf{F}(\mathbf{h}+\mathbf{d})\|_2^2 + \frac{\tau_1}{2}\left(\frac{1}{N}\mathrm{tr}\big(T(\mathbf{u})\big) + v\right) + \lambda_3\|\mathbf{d}\|_1, \ \text{ s.t. } \mathbf{Z} = \begin{bmatrix} T(\mathbf{u}) & \mathbf{h} \\ \mathbf{h}^{\mathrm{H}} & v \end{bmatrix} \succeq 0, \tag{53}$$

where $\mathbf{Z} \in \mathbb{C}^{(N+1)\times(N+1)}$ is an auxiliary matrix.

The augmented Lagrangian function for (53) is of the form

$$\begin{aligned}\mathcal{L}_\rho(v, \mathbf{u}, \mathbf{h}, \mathbf{d}, \mathbf{Z}, \mathbf{\Lambda}) =& \frac{1}{2}\|\mathbf{y} - \mathbf{F}(\mathbf{h}+\mathbf{d})\|_2^2 + \frac{\tau_1}{2}\left(\frac{1}{N}\mathrm{tr}\big(T(\mathbf{u})\big) + v\right) + \lambda_3\|\mathbf{d}\|_1 \\ &+ \left\langle \mathbf{\Lambda}, \mathbf{Z} - \begin{bmatrix} T(\mathbf{u}) & \mathbf{h} \\ \mathbf{h}^{\mathrm{H}} & v \end{bmatrix} \right\rangle + \frac{\rho}{2}\left\| \mathbf{Z} - \begin{bmatrix} T(\mathbf{u}) & \mathbf{h} \\ \mathbf{h}^{\mathrm{H}} & v \end{bmatrix} \right\|_{\mathrm{F}}^2,\end{aligned} \tag{54}$$

where $\mathbf{\Lambda} \in \mathbb{C}^{(N+1)\times(N+1)}$ is the Lagrangian multiplier, ρ is a penalty parameter, and $\langle \cdot, \cdot \rangle$ denotes the real inner product.

The ADMM algorithm minimizes the augmented Lagrangian function by iteratively performing

the following updates:

$$
\begin{aligned}
&v^{(l+1)} = \arg\min_{v} \mathcal{L}_\rho(v, \mathbf{u}^{(l)}, \mathbf{h}^{(l)}, \mathbf{d}^{(l)}, \mathbf{Z}^{(l)}, \mathbf{\Lambda}^{(l)}), \quad \mathbf{u}^{(l+1)} = \arg\min_{\mathbf{u}} \mathcal{L}_\rho(v^{(l+1)}, \mathbf{u}, \mathbf{h}^{(l)}, \mathbf{d}^{(l)}, \mathbf{Z}^{(l)}, \mathbf{\Lambda}^{(l)}),\\
&\mathbf{h}^{(l+1)} = \arg\min_{\mathbf{h}} \mathcal{L}_\rho(v^{(l+1)}, \mathbf{u}^{(l+1)}, \mathbf{h}, \mathbf{d}^{(l)}, \mathbf{Z}^{(l)}, \mathbf{\Lambda}^{(l)}), \quad \mathbf{d}^{(l+1)} = \arg\min_{\mathbf{d}} \mathcal{L}_\rho(v^{(l+1)}, \mathbf{u}^{(l+1)}, \mathbf{h}^{(l+1)}, \mathbf{d}, \mathbf{Z}^{(l)}, \mathbf{\Lambda}^{(l)}),\\
&\mathbf{Z}^{(l+1)} = \arg\min_{\mathbf{Z}} \mathcal{L}_\rho(v^{(l+1)}, \mathbf{u}^{(l+1)}, \mathbf{h}^{(l+1)}, \mathbf{d}^{(l+1)}, \mathbf{Z}, \mathbf{\Lambda}^{(l)}), \quad \mathbf{\Lambda}^{(l+1)} = \mathbf{\Lambda}^{(l)} + \rho\left(\mathbf{Z}^{(l+1)} - \begin{bmatrix} T\left(\mathbf{u}^{(l+1)}\right) & \mathbf{h}^{(l+1)} \\ \mathbf{h}^{(l+1)\mathrm{H}} & v^{(l+1)} \end{bmatrix}\right)
\end{aligned}
\tag{55}
$$

where the superscript (l) denotes the l-th iteration.

To perform the above updates in an efficient way, we first introduce the partitions

$$
\mathbf{Z}^{(l)} = \begin{bmatrix} \mathbf{Z}_0^{(l)} & \mathbf{z}_1^{(l)} \\ \mathbf{z}_1^{(l)\mathrm{H}} & Z_{N+1,N+1}^{(l)} \end{bmatrix}, \quad \mathbf{\Lambda}^{(l)} = \begin{bmatrix} \mathbf{\Lambda}_0^{(l)} & \boldsymbol{\lambda}_1^{(l)} \\ \boldsymbol{\lambda}_1^{(l)\mathrm{H}} & \Lambda_{N+1,N+1}^{(l)} \end{bmatrix}. \tag{56}
$$

The augmented Lagrangian function is convex and differentiable with respect to the variables v, $\mathbf{u}$, and $\mathbf{h}$. Setting their gradients to zero, these updates admit closed-form expressions as

$$
\begin{aligned}
&v^{(l+1)} = Z_{N+1,N+1}^{(l)} + \frac{1}{\rho}\left(\Lambda_{N+1,N+1}^{(l)} - \frac{\tau_1}{2}\right), \quad \mathbf{u}^{(l+1)} = \mathbf{\Psi}^{-1}\left(T^*\left(\mathbf{Z}_0^{(l)} + \frac{1}{\rho}\mathbf{\Lambda}_0^{(l)}\right)\right) - \frac{\tau_1}{2\rho}\mathbf{e}_1,\\
&\mathbf{h}^{(l+1)} = \left(\mathbf{F}^{\mathrm{H}}\mathbf{F} + 2\rho\mathbf{I}_N\right)^{-1}\left(\mathbf{F}^{\mathrm{H}}(\mathbf{y} - \mathbf{F}\mathbf{d}) + 2\boldsymbol{\lambda}_1^{(l)} + 2\rho\mathbf{z}_1^{(l)}\right),
\end{aligned}
\tag{57}
$$

where

$$
\begin{aligned}
\mathbf{\Psi} = \mathrm{diag}([&HW, H(W-1), \cdots, H,\\
&[H-1, H-2, \cdots, 1] \otimes [1, 2, \cdots, W, W-1, W-2, \cdots, 1]]) \in \mathbb{R}^{N_u \times N_u},
\end{aligned}
\tag{58}
$$

$\mathbf{e}_1$ is a zero vector with the first entry being one, and $T^*(\cdot)$ is the adjoint of $T(\cdot)$. For a given matrix $\mathbf{Q} \in \mathbb{C}^{N\times N}$,

$$
T^*(\mathbf{Q}) = [\mathrm{tr}((\mathbf{\Theta}_{k_1} \otimes \mathbf{\Theta}_{k_2})\mathbf{Q}) | (k_1, k_2) \in \mathsf{H}]^{\mathrm{T}}, \tag{59}
$$

where $\mathbf{\Theta}_k$ is an elementary Toeplitz matrix with ones on the k-th diagonal and zeros elsewhere, H is a halfspace of $(H-1, W-1)$, and $(k_1, k_2) \in \mathsf{H}$ means $k_2 \in \{0, 1, \cdots, W-1\}$ when $k_1 = 0$ and $k_2 \in \{-(W-1), -(W-2), \cdots, W-1\}$ when $k_1 \in \{1, 2, \cdots, H-1\}$.

The update of $\mathbf{d}$ can be written as

$$\begin{aligned}\mathbf{d}^{(l+1)} &= \arg\min_{\mathbf{d}} \mathcal{L}_\rho(v^{(l+1)}, \mathbf{u}^{(l+1)}, \mathbf{h}^{(l+1)}, \mathbf{d}, \mathbf{Z}^{(l)}, \mathbf{\Lambda}^{(l)}) = \arg\min_{\mathbf{d}} \frac{1}{2}\|\mathbf{y} - \mathbf{F}(\mathbf{h}^{(l+1)} + \mathbf{d})\|_2^2 + \lambda_3\|\mathbf{d}\|_1 \\ &= \arg\min_{\mathbf{d}} \frac{1}{2}\|(\mathbf{y} - \mathbf{F}\mathbf{h}^{(l+1)}) - \mathbf{F}\mathbf{d}\|_2^2 + \lambda_3\|\mathbf{d}\|_1,\end{aligned} \tag{60}$$

which is the well-known LASSO problem and can be solved efficiently by the algorithm proposed in [35].

Finally, the update of $\mathbf{Z}$

$$\mathbf{Z}^{(l+1)} = \arg\min_{\mathbf{Z}} \mathcal{L}_\rho(v^{(l+1)}, \mathbf{u}^{(l+1)}, \mathbf{h}^{(l+1)}, \mathbf{d}^{(l+1)}, \mathbf{Z}, \mathbf{\Lambda}^{(l)}) = \arg\min_{\mathbf{Z}\succeq 0} \left\| \mathbf{Z} - \left(\underbrace{\begin{bmatrix} T\left(\mathbf{u}^{(l+1)}\right) & \mathbf{h}^{(l+1)} \\ \mathbf{h}^{(l+1)\mathrm{H}} & v^{(l+1)} \end{bmatrix} - \frac{1}{\rho}\mathbf{\Lambda}^{(l)}}_{=\mathbf{G}} \right) \right\|_{\mathrm{F}}^2 \tag{61}$$

amounts to projecting the Hermitian matrix $\mathbf{G}$ onto the positive semidefinite cone, which can be performed by computing the eigenvalue decomposition of $\mathbf{G}$ and setting all negative eigenvalues to zero.

We now analyze the time complexity of the above algorithm. The computational complexity is mainly contributed by the updates of $\mathbf{u}, \mathbf{h}, \mathbf{d}$, and $\mathbf{Z}$. Note that the update of $\mathbf{u}$ can be simplified by avoiding unnecessary matrix multiplications. In (57), the calculation of $\mathbf{\Psi}^{-1}$ is independent of the input and thus can be cached before the iteration. The adjoint operator $T^*(\cdot)$ is essentially an "extract and sum" operator, where each output element of $T^*(\cdot)$ is the sum of multiple specific elements of the input, whose indices can also be cached before the iteration. Therefore, the time complexity of the update of $\mathbf{u}$ can be reduced to $O(N_u)$ (the multiplication of an $N_u \times N_u$ diagonal matrix $\mathbf{\Psi}^{-1}$ and an N_u-dimensional vector $T^*(\cdot)$). The update of $\mathbf{h}$ has time complexity of $O(N^3)$ due to the matrix inversion $(\mathbf{F}^{\mathrm{H}}\mathbf{F} + 2\rho\mathbf{I}_N)^{-1}$. The update of $\mathbf{d}$ (the LASSO problem) also has time complexity of $O(N^3)$ introduced by N-dimensional matrix inversion [35]. The update of $\mathbf{Z}$ amounts to eigenvalue decomposition, whose time complexity is upper bounded by $O(N^3)$. Combining the above results, the overall complexity of the proposed algorithm is $O(N^3)$.

Algorithm 1: An efficient ADMM-based algorithm for solving (50)

Initialization: $\mathbf{Z}^{(0)} = \mathbf{0}$, $\mathbf{\Lambda}^{(0)} = \mathbf{0}$, $\mathbf{D}^{(0)} = \mathbf{0}$;

while *not converged* **do**

$\mathbf{V}^{(l+1)} = \mathbf{Z}_{N_{\mathrm{RX}}}^{(l)} + \frac{1}{\rho}\left(\mathbf{\Lambda}_{N_{\mathrm{RX}}}^{(l)} - \frac{\tau_2}{2}\mathbf{I}_{N_{\mathrm{RX}}}\right)$;

$\mathbf{u}^{(l+1)} = \mathbf{\Psi}^{-1}\left(T^*\left(\mathbf{Z}_0^{(l)} + \frac{1}{\rho}\mathbf{\Lambda}_0^{(l)}\right)\right) - \frac{\tau_2}{2\rho}\mathbf{e}_1$;

$\mathbf{H}^{(l+1)} = (\mathbf{F}^{\mathrm{H}}\mathbf{F} + 2\rho\mathbf{I}_N)^{-1}\left(\mathbf{F}^{\mathrm{H}}(\mathbf{Y} - \mathbf{F}\mathbf{D}^{(l)}) + 2\mathbf{\Lambda}_1^{(l)} + 2\rho\mathbf{Z}_1^{(l)}\right)$;

$\mathbf{D}^{(l+1)} = \arg\min_{\mathbf{D}} \frac{1}{2}\|\left(\mathbf{Y} - \mathbf{F}\mathbf{H}^{(l)}\right) - \mathbf{F}\mathbf{D}\|_2^2 + \lambda_4\|\mathbf{D}\|_{2,1}$, solved by the M-FOCUSS algorithm [45];

Perform eigenvalue decomposition on $\begin{bmatrix} T\left(\mathbf{u}^{(l+1)}\right) & \mathbf{H}^{(l+1)} \\ \mathbf{H}^{(l+1)^{\mathrm{H}}} & \mathbf{V}^{(l+1)} \end{bmatrix} - \frac{1}{\rho}\mathbf{\Lambda}^{(l)}$ and set all negative eigenvalues to zero, yielding $\mathbf{Z}^{(l+1)}$;

$\mathbf{\Lambda}^{(l+1)} = \mathbf{\Lambda}^{(l)} + \rho\left(\mathbf{Z}^{(l+1)} - \begin{bmatrix} T\left(\mathbf{u}^{(l+1)}\right) & \mathbf{H}^{(l+1)} \\ \mathbf{H}^{(l+1)^{\mathrm{H}}} & \mathbf{V}^{(l+1)} \end{bmatrix}\right)$;

end

where

$\mathbf{Z}^{(l)} = \begin{bmatrix} \mathbf{Z}_0^{(l)} \in \mathbb{C}^{N\times N} & \mathbf{Z}_1^{(l)} \in \mathbb{C}^{N\times N_{\mathrm{RX}}} \\ \mathbf{Z}_1^{(l)^{\mathrm{H}}} \in \mathbb{C}^{N_{\mathrm{RX}}\times N} & \mathbf{Z}_{N_{\mathrm{RX}}}^{(l)} \in \mathbb{C}^{N_{\mathrm{RX}}\times N_{\mathrm{RX}}} \end{bmatrix}$, $\mathbf{\Lambda}^{(l)} = \begin{bmatrix} \mathbf{\Lambda}_0^{(l)} \in \mathbb{C}^{N\times N} & \mathbf{\Lambda}_1^{(l)} \in \mathbb{C}^{N\times N_{\mathrm{RX}}} \\ \mathbf{\Lambda}_1^{(l)^{\mathrm{H}}} \in \mathbb{C}^{N_{\mathrm{RX}}\times N} & \mathbf{\Lambda}_{N_{\mathrm{RX}}}^{(l)} \in \mathbb{C}^{N_{\mathrm{RX}}\times N_{\mathrm{RX}}} \end{bmatrix}$.

Appendix B
An Efficient Algorithm for Solving (50)

An efficient ADMM-based algorithm for solving (50) is presented in Algorithm 1. One can refer to Appendix A for a similar derivation.

Under the MMV model, the time complexity of the proposed algorithm for solving (50) increases to $O\left((N + N_{\mathrm{RX}})^3\right)$ due to the increased size of the eigenvalue decomposition problem (the update of $\mathbf{Z}$).

References

[1] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2019.

[2] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.

[3] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, 2020.

[4] M. Di Renzo, F. H. Danufane, X. Xi, J. De Rosny, and S. Tretyakov, "Analytical modeling of the path-loss for reconfigurable intelligent surfaces–anomalous mirror or scatterer?" in *IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5.

[5] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, 2019.

[6] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, "Intelligent reflecting surface enhanced wireless networks: Two-timescale beamforming optimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 2–17, 2020.

[7] W. Tang, J. Y. Dai, M. Z. Chen, K.-K. Wong, X. Li, X. Zhao, S. Jin, Q. Cheng, and T. J. Cui, "MIMO transmission through reconfigurable intelligent surface: System design, analysis, and implementation," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2683–2699, 2020.

[8] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. Di Renzo *et al.*, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45 913–45 923, 2020.

[9] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light Sci. Appl.*, vol. 3, no. 10, pp. e218–e218, 2014.

[10] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, 2018.

[11] D. Kouzapas, C. Skitsas, T. Saeed, V. Soteriou, M. Lestas, A. Philippou, S. Abadal, C. Liaskos, L. Petrou, J. Georgiou *et al.*, "Towards fault adaptive routing in metasurface controller networks," *J. Syst. Archit.*, vol. 106, p. 101703, 2020.

[12] H. Taghvaee, A. Cabellos-Aparicio, J. Georgiou, and S. Abadal, "Error analysis of programmable metasurfaces for beam steering," *IEEE J. Emerging Sel. Top. Circuits Syst.*, vol. 10, no. 1, pp. 62–74, 2020.

[13] M. E. Eltayeb, T. Y. Al-Naffouri, and R. W. Heath, "Compressive sensing for millimeter wave antenna array diagnosis," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2708–2721, 2018.

[14] Y.-C. Liang, R. Long, Q. Zhang, J. Chen, H. V. Cheng, and H. Guo, "Large intelligent surface/antennas (LISA): Making reflective radios smart," *J. Commun. Inf. Networks*, vol. 4, no. 2, pp. 40–50, 2019.

[15] J. Lee, E. M. Ferren, D. P. Woollen, and K. M. Lee, "Near-field probe used as a diagnostic tool to locate defective elements in an array antenna," *IEEE Trans. Antennas Propag.*, vol. 36, no. 6, pp. 884–889, 1988.

[16] O. M. Bucci, M. D. Migliore, G. Panariello, and P. Sgambato, "Accurate diagnosis of conformal arrays from near-field data using the matrix method," *IEEE Trans. Antennas Propag.*, vol. 53, no. 3, pp. 1114–1120, 2005.

[17] M. D. Migliore, "A compressed sensing approach for array diagnosis from a small set of near-field measurements," *IEEE Trans. Antennas Propag.*, vol. 59, no. 6, pp. 2127–2133, 2011.

[18] G. Oliveri, P. Rocca, and A. Massa, "Reliable diagnosis of large linear arrays—A Bayesian compressive sensing approach," *IEEE Trans. Antennas Propag.*, vol. 60, no. 10, pp. 4627–4636, 2012.

[19] M. Salucci, A. Gelmini, G. Oliveri, and A. Massa, "Planar array diagnosis by means of an advanced Bayesian compressive processing," *IEEE Trans. Antennas Propag.*, vol. 66, no. 11, pp. 5892–5906, 2018.

[20] T. Ince and G. Ögücü, "Array failure diagnosis using nonconvex compressed sensing," *IEEE Antennas Wirel. Propag. Lett.*, vol. 15, pp. 992–995, 2015.

[21] B. Fuchs, L. Le Coq, and M. D. Migliore, "Fast antenna array diagnosis from a small number of far-field measurements," *IEEE Trans. Antennas Propag.*, vol. 64, no. 6, pp. 2227–2235, 2016.

[22] M. D. Migliore, "Array diagnosis from far-field data using the theory of random partial Fourier matrices," *IEEE Antennas Wirel. Propag. Lett.*, vol. 12, pp. 745–748, 2013.

[23] R. Palmeri, T. Isernia, and A. F. Morabito, "Diagnosis of planar arrays through phaseless measurements and sparsity promotion," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 6, pp. 1273–1277, 2019.

[24] C. Xiong, G. Xiao, Y. Hou, and M. Hameed, "A compressed sensing-based element failure diagnosis method for phased array antenna during beam steering," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 9, pp. 1756–1760, 2019.

[25] O. Aldayel, "Constrained optimization for millimeter array diagnosis," in *IEEE Radar Conference (RadarConf)*. IEEE, 2019, pp. 1–6.

[26] S. Ma, W. Shen, J. An, and L. Hanzo, "Antenna array diagnosis for millimeter-wave MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4585–4589, 2020.

[27] R. Sun, W. Wang, L. Chen, G. Wei, and W. Zhang, "Hybrid beamforming system diagnosis: Failure modeling and identification," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3415–3429, 2021.

[28] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, 2019.

[29] R. J. Mailloux, "Array failure correction with a digitally beamformed array," *IEEE Trans. Antennas Propag.*, vol. 44, no. 12, pp. 1543–1550, 1996.

[30] B.-K. Yeo and Y. Lu, "Array failure correction with a genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 47, no. 5, pp. 823–828, 1999.

[31] M. Li and Y. Lu, "Source bearing and steering-vector estimation using partially calibrated arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 4, pp. 1361–1372, 2009.

[32] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, 2014.

[33] G. Zhou, C. Pan, H. Ren, K. Wang, M. Di Renzo, and A. Nallanathan, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1658–1662, 2020.

[34] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[35] S. Boyd, N. Parikh, and E. Chu, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. Now Publishers Inc, 2011.

[36] E. J. Candès and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7235–7254, 2011.

[37] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.

[38] J. Wright, A. Ganesh, K. Min, and Y. Ma, "Compressive principal component pursuit," *Information and Inference: A Journal of the IMA*, vol. 2, no. 1, pp. 32–68, 2013.

[39] J. Tanner and S. Vary, "Compressed sensing of low-rank plus sparse matrices," *arXiv preprint arXiv:2007.09457*, 2020.

[40] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, 2012.

[41] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182–2195, 2011.

[42] Y. Chi and M. F. Da Costa, "Harnessing Sparsity Over the Continuum: Atomic norm minimization for superresolution," *IEEE Signal Process Mag.*, vol. 37, no. 2, pp. 39–57, 2020.

[43] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1030–1042, 2014.

[44] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1257–1269, 2015.

[45] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, 2005.