

Hybrid Evolutionary-based Sparse Channel Estimation for IRS-assisted mmWave MIMO Systems

Zhen Chen, *Member, IEEE*, Jie Tang, *Senior Member, IEEE*,
Xiu Yin Zhang, *Senior Member, IEEE*, Daniel Ka Chun So, *Senior Member, IEEE*, Shi Jin, *Senior Member, IEEE*, and Kai-Kit Wong, *Fellow, IEEE*

Abstract

The intelligent reflecting surface (IRS)-assisted millimeter wave (mmWave) communication system has emerged as a promising technology for coverage extension and capacity enhancement. Prior works on IRS have mostly assumed perfect channel state information (CSI), which facilitates in deriving the upper-bound performance but is difficult to realize in practice due to passive elements of IRS without signal processing capabilities. In this paper, we propose a compressive channel estimation techniques for IRS-assisted mmWave multi-input and multi-output (MIMO) system. To reduce the training overhead, the inherent sparsity of mmWave channels is exploited. By utilizing the properties of Kronecker products, IRS-assisted mmWave channel is converted into a sparse signal recovery problem, which involves two competing cost function terms (measurement error and sparsity term). Existing sparse recovery algorithms solve the combined contradictory objectives function using a regularization parameter, which leads to a suboptimal solution. To address this concern, a hybrid multiobjective evolutionary paradigm is developed to solve the sparse recovery problem, which can overcome the difficulty in the choice of regularization parameter value. Simulation results show that under a wide range of simulation settings, the proposed method achieves competitive error performance compared to existing channel estimation methods.

Index Terms

Intelligent reflecting surface (IRS), millimeter wave communications, channel estimation, compressed sensing, hybrid evolutionary algorithm.

I. INTRODUCTION

Intelligent reflecting surface (IRS) is a key candidate technology to achieve high spectrum efficiency and the wireless communication coverage, due to it enables the reconfiguration of wireless propagation environment by controlling the signal reflections via the use of massive low-cost passive elements, [1]–[4]. By utilizing the preprogrammed controller, the IRS can change the attenuation and scattering of the incident electromagnetic wave so that the desired wireless channels is reconfigured to reach the intended receiver, which is called as programmable and controllable wireless environment. Different from the conventional relaying and backscatter communications that require power amplifiers to retransmit the signal received, IRS does not process any incident signals but only reflects the incident signals without amplifying or decoding, thus incurring no additional hardware cost and power consumption [5], [6]. Since IRS is not equipped with expensive signal processing equipments such as digital-to-analog converter and modulator/demodulator, IRS is simple and low-cost that can be easily deployed on urban commercial buildings, shopping mall or indoor spaces [6], [7]. These significant advantages make IRS passive beamforming as a green energy-efficient technique for 6G cellular Internet of Things (IoT), and hence it has been investigated in different wireless communication such as multi-cell cellular networks, secure MIMO transmission, simultaneous wireless information and power transfer (SWIPT) etc.

However, accurate channel estimation is challenging in IRS-assisted wireless communication. As opposed to conventional backscatter and relay-assisted wireless communication that channel state information (CSI) can be obtained using transmission/processing pilot symbols, the passive reflection elements of IRS no signal processing capability, and hence it is practically difficult to estimate channels. On the other hand, some of the existing works assume the availability of perfect CSI to design the precoding matrix at the base station (BS) and phase shifts matrix at the IRS. Nevertheless, this assumption is difficult to realize in practice. To address these problem, various channel estimation strategies and algorithms were developed for IRS-assisted mmWave communications. In [8], [9], the least squares (LS)-based channel estimation schemes were proposed for a single-user IRS-assisted MISO system. The idea was extended to the orthogonal frequency division multiplexing (OFDM) system [10] that an IRS with discrete phase shifts was used to improve the achievable spectrum efficiency. The work was also extended in [11], [12] that exploit a pilot-based three phase framework for the IRS-assisted multiuser systems. It

was show that the number of pilot symbols can be significantly reduced in channel estimation processing and effectively used to the more general channel model without assuming any channel property. Besides, a sparse channel estimation scheme with matrix factorization was developed that exploiting the low-rank structure between the BS-IRS and IRS-user channels [13]. However, in IRS-assisted mmWave communication systems, how to reduce the pilot overhead in accurate channel estimation still remains a challenge and an open research issue, because the number of channel parameters to be estimated is large as the passive elements of IRS increase, while pilot frequency sequence adopted by channel estimation is limited to make sure an achievable spectrum efficiency.

An alternative approach to mmWave channel estimation is to adopt the compressed sensing (CS), which can be leveraged to effectively estimate mmWave massive MIMO channels [14]–[16]. More recently, CS techniques combine with deep denoising neural network has been advocated for IRS-assisted mmWave channel estimation that substantiality reduce the pilot overhead [17]. These CS-based approaches exploit the the limited scattering characteristics of mmWave propagation to form a sparse channel recovery problem [18], [19]. Therefore, by exploiting the sparsity of mmWave channels, the corresponding channel estimation would significantly reduce the required pilot overhead. It is particularly beneficial to mmWave MIMO systems, where the number of pilot symbols increases linearly with the number of reflecting elements and transmit antennas. With that approach, the corresponding CS-based channel estimation algorithms including matching pursuit, orthogonal matching pursuit (OMP) [20] and the least absolute shrinkage and selection operator (LASSO) algorithm [21], [22] or its variations were exploited for the pilot-assisted channel estimation rather than using traditional least squares (LS) methods. In particular, a burst LASSO algorithm was developed for robust channel estimation in massive MIMO system, which can greatly improve the channel estimation performance by exploiting the burst-sparse structure of the channel [23]. They also developed a P-CSPI aided burst LASSO algorithm to improve performance gain and effectively alleviate the stringent requirement on the sparsity level for robust CSIT recovery. In [24], a compressed channel estimation for IRS-assisted mmWave system was studied, in which the cascade channel can be cast as a CS-based sparse recovery problem by exploiting the inherent sparse structure of mmWave. However, all of the above CS-based algorithms mentioned are single objective optimization problem that is solved with with the fixed solution path. Moreover, the performance of [21]–[24] is highly dependent on the regularization parameter. Over the recent few years, various strategies and

algorithms for choosing optimal parameter were studied [25]–[27]. To achieve the optimized regularization parameter, the noise standard deviation should be provided in advance. The same strategies also have been studied in [28] that can obtain the more accurate results by sacrificing computational overhead for noise standard deviation estimation. To avoid the noise estimation for the optimized regularization parameter, one possible method is to convert the CS-based optimization problem into multiobjective optimization problem (MOP), and the corresponding multiobjective evolutionary algorithms (MOEAs) is employed to yield better results than the use of conventional single objective optimization solutions. MOP has gained increasing interest [29]–[31] and have been proven to be beneficial for solving multiple contradictory objectives simultaneously. The final solution of MOEAs is to approximate the Pareto optimal front. In [32], a general MOEA based on differential evolution (DE) was proposed. Unlike other MOEAs, the DE can easily solve large-scale optimization problem in massive MIMO system. To achieve the satisfactory solutions of multiobjective schemes, the knee region of weakly Pareto front near the true solution must be preferred for the exploitation of the local search. It is worth noting that the aforementioned works cannot be extended directly to IRS-based mmWave channel estimation, due to the fact that the mmWave channel with IRS will make the problem highly complex and difficult to solve, which is one of the major motivations to devote our endeavour to developing efficient channel estimation for IRS-assisted mmWave MIMO systems.

As mentioned above, most of the existing studies [18], [19], [22], [23], [33]–[36] only focus on the investigation of channel estimation algorithms in the presence of mmWave MIMO channel model without considering the IRS technologies. In fact, IRS may play an important role in providing high speed data throughput for the future mmWave MIMO communication scenarios, such as the autonomous vehicular communications [37], high speed railway and tunnel communications. In particular, as the development of urbanization, the line-of-sight (LoS) link is blocked by more and more high-rise buildings emerge in dense down town. On the other hand, the study in [15], [17], [24] adapts CS-based channel estimation to reliably estimate and feed back the CSI, which achieves accurate CSI acquisition. These prior works are restricted to conventional single objective design, which yield to the fixed solution path in the optimization process and leads to the suboptimal sparse solution by fixed regularization parameter. This motivates us to investigate a simple and yet efficient channel estimation scheme in IRS-assisted mmWave massive MIMO systems.

In this work, we propose a sparse channel estimation scheme that combines CS techniques

and a hybrid multiobjective evolutionary optimization method for IRS-assisted mmWave MIMO systems. The key observation is exploiting CS-based sparse channel estimation to reduce the training overhead and developing the multiobjective evolutionary method to find an optimal tradeoff between competing objective functions. According to the MOP, a hybrid MOEA is proposed to select the final solution of the channel estimation in the knee areas. To accurately estimate knee regions, the maximum slope of the Pareto plot is exploited to determine the final solution, which indicates the best balance between the measurement error and the sparsity constraint whilst providing the best performance in sparse channel estimation. Moreover, formal theoretical understanding of why MOEAs can estimate the sparse channel is largely missing [29]–[31], [38], [39]. To solve this problem, we provide a convergence analysis for the proposed hybrid MOEA by deriving a performance bound on the estimation error. The bound proves that our method can estimate the sparse channel when the measurement matrix satisfies a properly restricted isometry property (RIP). The main contributions are summarized as follows:

- To reduce the training overhead, we put forward a CS-based sparse recovery approach for IRS-assisted mmWave channel estimation. By utilizing the properties of Kronecker products, the original IRS-assisted channel estimation model can be reformulated as a cascade channel, which is a classic CS-based sparse recovery problem.
- To provide adaptability to the sparse solution, the proposed sparse recovery problem is transformed into MOPs. The corresponding hybrid MOEA based on DE is then developed to solve the MOPs. In each iteration, the iterative hard thresholding (IHT) method is integrated into the proposed hybrid MOEA, which is very crucial to significantly reduce the computational cost of the selection process. Moreover, the parallel search scheme is proposed to obtain diversifying the search along weakly Pareto front without optimizing regularizing parameter.
- To support the applicability of the proposed hybrid MOEA, we derive a performance bound on the channel estimation error and prove our method can converge to the neighborhood of the true channel value. Moreover, we conduct numerical simulations to verify the performance of the proposed method in comparison with existing channel estimation methods.

The rest of this paper is organized as follows. In Section II, the detailed description of CS-based channel model in IRS-assisted mmWave system is proposed. Section III proposes a hybrid multiobjective evolutionary approach for mmWave channel estimation. In Section IV,

the convergence analysis of the proposed algorithm is provided. The simulation results on the proposed hybrid MOEA are reported and analyzed in Section V. Section VI concludes the paper and outlines the future research directions.

Notations: The boldface lower-case and upper-case symbols denote vectors and matrices respectively. $\text{vec}(\mathbf{A})$ is a vector obtained through the vectorization of a matrix \mathbf{A} . $\|\mathbf{a}\|_p$ represents the p -norm of \mathbf{a} , where a_i is the i -th component of vector \mathbf{a} and p denotes an integer. The superscripts $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^*$ represent the transpose, conjugate transpose and complex conjugate respectively. $\mathbf{X} \otimes \mathbf{Y}$ denotes the Kronecker product of \mathbf{X} and \mathbf{Y} . \mathbb{R} and \mathbb{C} denote the set of real and complex numbers, respectively. $\mathbb{E}[\cdot]$ denotes the expectation and $\log(\cdot)$ denotes the logarithm.

II. SYSTEM MODEL AND PROBLEM FORMULATION

In this section, the IRS-assisted mmWave MIMO communication system is introduced. Then, we formulate the IRS-assist mmWave channel optimization problem as a CS-based sparse recovery problem in the following.

A. System Model

We focus on the downlink transmission of an IRS-assisted mmWave MIMO system in Fig. 1, where an IRS is deployed to assist the signal transmission from the N_{BS} antennas of the BS to user equipment (UE). It is assumed that the IRS is a planar array with M reflecting elements and the signals are reflected by the IRS once. The IRS is attached to the facade of a building located in the LoS link of the BS. The phase shifts of each reflecting element on the IRS are configurable via a smart controller, which gets this information from the BS over a backhaul link [40]. Let $\mathbf{R} \in \mathbb{C}^{M \times N_{BS}}$ be the channel from the BS to the IRS, $\mathbf{H}_r^H \in \mathbb{C}^{N_{UE} \times M}$ denotes the channel from the IRS to the UE. This is a valid assumption as IRS is practically applied to support UE without direct communication link from the BS due to deep path loss or obstacle blockage. To this end, we neglect the LOS link from the BS to the UE. Furthermore, a time-division duplexing is considered for the CSI acquisition by exploiting the characteristic of the mmWave channel reciprocity.

As shown in Fig. 1, at t -th time frame, the transmitted beamforming at the BS side can be written as

$$\mathbf{x}_t = \mathbf{F}_t \mathbf{s}_t, \quad (1)$$

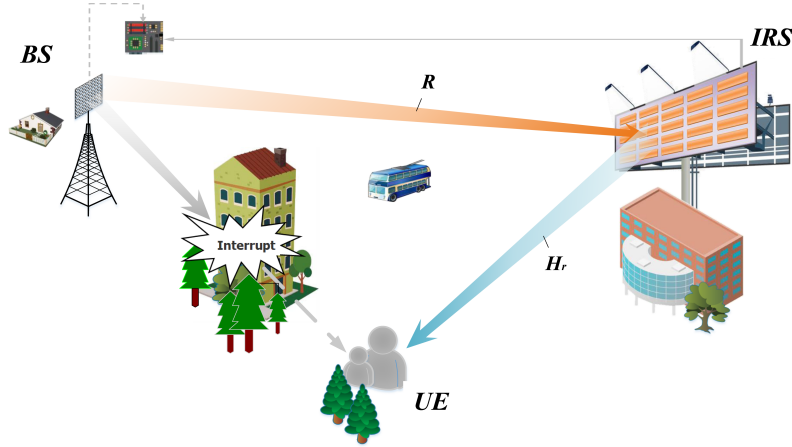


Fig. 1. Illustration of the considered IRS-assisted millimeter-wave MIMO communication systems.

where $s_t \in \mathbb{C}^{N_s}$ denotes the transmitted pilot symbol and $F_t \in \mathbb{C}^{N_{BS} \times N_s}$ is the transmit beamforming at the BS. We suppose the pilot satisfies $\mathbb{E}[s_t s_t^H] = (P_s/N_s)\mathbf{I}_{N_s}$ and P_s is the transmitted power of the BS.

For the channel training scheme in the above mentioned system, the BS sends out training symbols s_t and the observed signal at the BS side in time slot t is

$$\hat{y}_t = \mathbf{H}x_t + z_t, \quad (2)$$

where the z_t is the additive Gaussian noise (AWGN), and its elements are from complex Gaussian distribution $\mathcal{CN}(0, \sigma^2)$ [41]. After being multiplied with the hybrid combining matrix, the received signal at the UE side can be expressed as

$$y_t = \mathbf{W}^H \mathbf{H}x_t + n_t, \quad (3)$$

where $\mathbf{W} \in \mathbb{C}^{N_{UE} \times Q}$ denotes the combining matrix and $n_t = \mathbf{W}z_t$. \mathbf{H} is the cascaded channel matrix of the IRS that is given by

$$\mathbf{H} = \mathbf{H}_r^H \Theta \mathbf{R}, \quad (4)$$

where $\Theta \in \mathbb{C}^{M \times M}$ is diagonal phase-shifting matrix of the IRS, which can be expressed as

$$\Theta = \text{diag}(\beta_1 e^{j\theta_1}, \dots, \beta_M e^{j\theta_M}), \quad (5)$$

and $\theta_m \in [0, 2\pi)$, $\beta_m \in [0, 1]$, $m = 1, \dots, M$ being the m -th phase shift of reflecting element and amplitude reflection, respectively. To simplify hardware design of the IRS, we consider the same β_m for all the reflecting elements of the IRS.

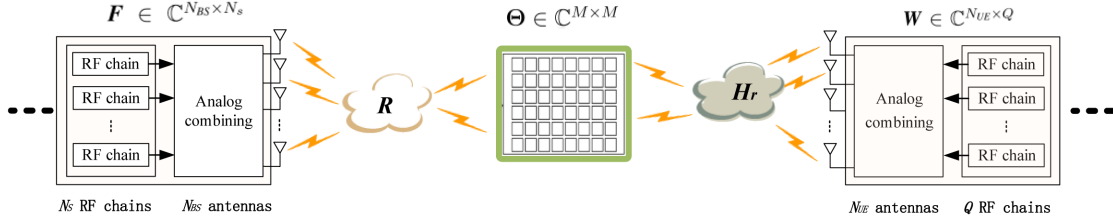


Fig. 2. Channel model of sparse channel estimation in IRS-assisted MIMO system.

With regard to the total T time frames, thus the received matrix at the UE side can be written as

$$\mathbf{Y} = \mathbf{W}^H \mathbf{H} \mathbf{X} + \mathbf{N}, \quad (6)$$

where $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$ is a $Q \times T$ received matrix, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T]$ is a $N_{BS} \times T$ beamforming matrix, $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_T]$ is a $Q \times T$ noise matrix by concatenating the noise vectors. The data path diagram is described in Fig. 2.

B. Problem Formulation

Before formulating the channel estimation problem, the channel model for the IRS-assisted mmWave system should be described firstly. By smartly adjusting the phase shifts of each scattering elements, as illustrated in Fig. 1. Due to the reflecting element of IRS no signal processing, the superposed signal arrived at the IRS will be directly reflected towards the desired UEs without any additional time domain delay. Hence, the time difference between the reflected signal and the directed signal is marginal and thus can be ignored. According to the discussion above, we take into account the cascade channel matrix \mathbf{H} , which is divided into the BS-IRS channel \mathbf{R} and the IRS-UE channel \mathbf{H}_r . The primary challenge is to estimate the BS-IRS and the IRS-UE channel simultaneously. Nevertheless, we discover above channel model can be converted to a CS-based sparse recovery problem. To be specific, the BS-IRS channel can be modeled as [42], [43]

$$\mathbf{R} = \sqrt{\frac{N_{BS} N_{IRS}}{\rho}} \sum_{l=1}^L \alpha_l \mathbf{a}_{IRS,1}(\varphi_l, \gamma_l) \mathbf{a}_{BS}^H(\phi_l), \quad (7)$$

where α_l is the channel complex gain of the l -th path, L denotes the scattering paths of BS-IRS link, and corresponding path loss is denoted as ρ . $\varphi_l(\gamma_l)$ denotes the azimuth (elevation) angle

of arrival (AoA) of the l -th path, ϕ_l is the angle of departure (AoD) of the l -th path; $\mathbf{a}_{IRS,1}$ and \mathbf{a}_{BS} denote the array response associated with the IRS and BS, respectively. We assume that the IRS is an $M_x \times M_y$ uniform planar array (UPA). Then the steering vectors at the IRS can be written as [24]

$$\mathbf{a}_{IRS,1}(\varphi_l, \gamma_l) = \mathbf{a}_x(u) \otimes \mathbf{a}_y(v), \quad (8)$$

where \otimes stands for the Kronecker product. $\mathbf{a}_x(u)$ and $\mathbf{a}_y(v)$ are expressed as

$$\begin{aligned} \mathbf{a}_x(u) &= \frac{1}{\sqrt{M_x}} \left[1, e^{\frac{j2\pi du}{\lambda}}, \dots, e^{\frac{j2\pi(M_x-1)du}{\lambda}} \right]^T, \\ \mathbf{a}_y(v) &= \frac{1}{\sqrt{M_y}} \left[1, e^{\frac{j2\pi dv}{\lambda}}, \dots, e^{\frac{j2\pi(M_y-1)dv}{\lambda}} \right]^T, \end{aligned} \quad (9)$$

where $u \triangleq \cos(\gamma_l)$, $v \triangleq \sin(\gamma_l)\cos(\varphi_l)$, d denotes the antenna spacing between neighboring antenna elements and λ is the signal wavelength.

Due to the sparse scattering nature of mmWave, the number of path L is generally less than the dimensions of \mathbf{R} . The corresponding \mathbf{R} can be expressed as

$$\mathbf{R} = (\mathbf{F}_x \otimes \mathbf{F}_y) \Sigma_\alpha \mathbf{F}_{BS}^H \triangleq \mathbf{F}_{IRS,1} \Sigma_\alpha \mathbf{F}_{BS}^H, \quad (10)$$

where $\mathbf{F}_{IRS,1} = \mathbf{F}_x \otimes \mathbf{F}_y \in \mathbb{C}^{M \times M_L}$, $\mathbf{F}_x \in \mathbb{C}^{M_x \times M_{L,x}}$ and $\mathbf{F}_y \in \mathbb{C}^{M_y \times M_{L,y}}$ are defined with each of its columns having a form of $\mathbf{a}_x(u)$ and $\mathbf{a}_y(v)$, respectively. $u(v)$ chosen from a pre-discretized grid, $\Sigma_\alpha \in \mathbb{C}^{M_L \times M_L}$ is a diagonal matrix with L non-zero entries corresponding to the channel path gains $\left\{ \sqrt{\frac{N_{BS} N_{IRS}}{\rho}} \alpha_l \right\}$, in which $M_L = M_{L,x} \times M_{L,y}$, $\mathbf{F}_{BS} \in \mathbb{C}^{N_{BS} \times M_L}$ is an overcomplete matrix and each of its columns has a form of $\mathbf{a}_{BS}(\phi_l)$, with ϕ_l chosen from a pre-discretized grid.

Similar to the BS-IRS channel \mathbf{R} , the IRS-UE channel can be modeled as

$$\mathbf{H}_r = \sqrt{\frac{N_{UE} N_{IRS}}{\mu}} \sum_{l'=1}^{L'} c_{l'} \mathbf{a}_{UE}(\varphi_{l'}, \gamma_{l'}) \mathbf{a}_{IRS,2}^H(\phi_{l'}), \quad (11)$$

where $c_{l'}$ denotes the complex gain associated with the l' -th path, L' denotes the number of IRS-UE channel path, μ denotes the average path loss of the IRS-UE link and $\varphi_{l'}(\gamma_{l'})$ denotes the azimuth (elevation) angle of departure. Thus, the IRS-UE channel can also be written as

$$\mathbf{H}_r^H \triangleq \mathbf{F}_{UE} \Sigma_c \mathbf{F}_{IRS,2}^H, \quad (12)$$

where $\Sigma_c \in \mathbb{C}^{M_{L'} \times M_{L'}}$ is a sparse matrix with L' non-zero entries corresponding to the channel path gains $\left\{ \sqrt{\frac{N_{UE} N_{IRS}}{\mu}} c_{l'} \right\}$, $\mathbf{F}_{IRS,2} \in \mathbb{C}^{M \times M_{L'}}$ and $\mathbf{F}_{UE} \in \mathbb{C}^{N_{UE} \times M_{L'}}$ are similarly defined with each of its columns having a form of $\mathbf{a}_{IRS,2}(\varphi_{l'}, \gamma_{l'})$ and $\mathbf{a}_{BS}(\phi_{l'})$, respectively.

Substituting (10) and (12) into (4), the cascade channel model can be expressed as

$$\begin{aligned} \text{vec}(\mathbf{H}) &= \text{vec}(\mathbf{F}_{UE} \Sigma_c \mathbf{F}_{IRS,2}^H \Theta \mathbf{F}_{IRS,1} \Sigma_\alpha \mathbf{F}_{BS}^H) \\ &\stackrel{(a)}{=} (\mathbf{F}_{BS}^* \otimes \mathbf{F}_{UE}) \text{vec}(\Sigma_c \mathbf{F}_{IRS,2}^H \Theta \mathbf{F}_{IRS,1} \Sigma_\alpha) \\ &\stackrel{(b)}{=} (\mathbf{F}_{BS}^* \otimes \mathbf{F}_{UE}) (\Sigma_\alpha^T \otimes \Sigma_c) \text{vec}(\mathbf{F}_{IRS,2}^H \Theta \mathbf{F}_{IRS,1}) \\ &\stackrel{(c)}{=} (\mathbf{F}_{BS}^* \otimes \mathbf{F}_{UE}) (\Sigma_\alpha^T \otimes \Sigma_c) \mathbf{D} \mathbf{u}, \end{aligned} \quad (13)$$

where (a) and (b) follow from the property of Kronecker product, (c) comes from $\mathbf{u} \triangleq \text{vec}(\Theta) \in \mathbb{C}^{M^2}$ and $\mathbf{D} = \mathbf{F}_{IRS,1}^T \otimes \mathbf{F}_{IRS,2}^H \in \mathbb{C}^{M_L M_{L'} \times M^2}$.

With the proposed cascade channel design in (13), the received signal $\mathbf{y} = \text{vec}(\mathbf{Y})$ in (6) can be written as

$$\begin{aligned} \mathbf{y} &= (\mathbf{X}^T \otimes \mathbf{W}^H) \text{vec}(\mathbf{H}) + \text{vec}(\mathbf{N}) \\ &\stackrel{(a)}{=} (\mathbf{X}^T \otimes \mathbf{W}^H) (\mathbf{F}_{BS}^* \otimes \mathbf{F}_{UE}) (\Sigma_\alpha^T \otimes \Sigma_c) \mathbf{D} \mathbf{u} + \text{vec}(\mathbf{N}) \\ &\stackrel{(b)}{=} (\mathbf{D} \mathbf{u})^T \otimes ((\mathbf{X}^T \otimes \mathbf{W}^H) (\mathbf{F}_{BS}^* \otimes \mathbf{F}_{UE})) \mathbf{h} + \text{vec}(\mathbf{N}) \\ &\stackrel{(c)}{=} \Phi \mathbf{h} + \mathbf{e}, \end{aligned} \quad (14)$$

where (a) is based on the property of Kronecker product, (b) is obtained by $\mathbf{h} \triangleq \text{vec}(\Sigma_\alpha^T \otimes \Sigma_c)$ and (c) comes from $\Phi \triangleq (\mathbf{D} \mathbf{u})^T \otimes ((\mathbf{X}^T \otimes \mathbf{W}^H) (\mathbf{F}_{BS}^* \otimes \mathbf{F}_{UE}))$ and $\mathbf{e} = \text{vec}(\mathbf{N})$.

By the sparsity of the mmWave MIMO channel, we can obtain the Σ_α and Σ_c are sparse. If the sparsity of Σ_α and Σ_c are L and L' , then \mathbf{h} is a $M_L^2 M_{L'}^2$ vector with LL' -sparse. Since the $LL' < M_L^2 M_{L'}^2$, the matrix \mathbf{D} contains redundant columns. Based on the above analysis, the channel estimation model in (14) is a classic CS-based sparse recovery problem [44], and thus the CS theory can be employed to estimate \mathbf{h} by solving the following sparse recovery problem:

$$\min_{\mathbf{h}} \left\{ \mathbf{G}(\mathbf{h}) = \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{h}\|_2^2 + \lambda \|\mathbf{h}\|_p \right\}, \quad (15)$$

where λ is regularization parameter to balance the measurement error term $\|\mathbf{y} - \Phi \mathbf{h}\|_2^2$ and sparsity term $\|\mathbf{h}\|_p$ and p is usually set to 0 or 1. Here exists the problem of determining λ , as λ is closely related to the precision of channel estimation. In fact, a relatively high value of λ may lead to a sparse solution, whereas a small value may leads to a noisy result. Therefore, the main

challenge in the optimization problem (15) is how to optimally choose suitable regularization parameter so as to improve the performance of channel estimation.

Fig. 3 illustrates a simple experiment to demonstrate how the parameter λ affects the results of the CS-based sparse recovery algorithm. As illustrated in Fig. 3, the horizontal axis denotes some possible choices of regularization parameter λ , and the vertical axis plots the relative error in terms of the difference between the true channel and the estimated sparse solutions. It can be inferred from the above experiment that the regularization parameter λ significantly affects the recovery quality. To this end, various strategies were studied to determine regularization parameters in [25], [26], when the noise standard deviation is known assumption. There are some existing works that focused on the noise standard deviation estimation [45], [46]. However, in practical applications, system environments are usually changed with time during the measurement, the statistic character of the system noise and the observation noise cannot be estimated or be determined accurately. Therefore, the selection of regularization parameter will directly affects the performance of channel estimation.

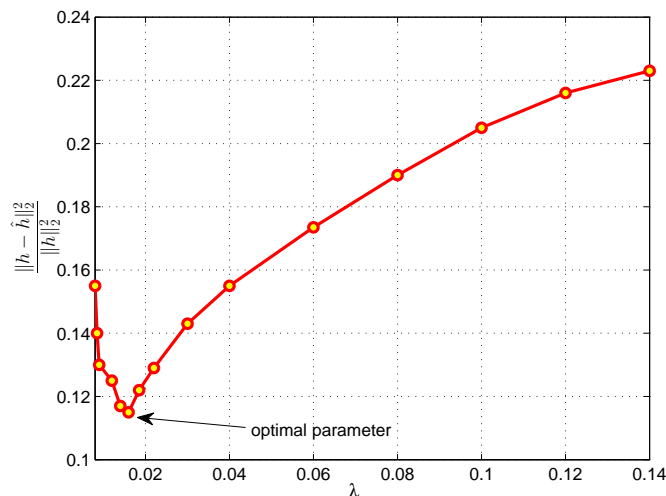


Fig. 3. The performance of the LASSO algorithm with different choices of the regularization parameter λ .

It is interesting that this deficiency can be naturally alleviated by employing MOP. Next, a hybrid MOEA is presented for MOP and show how it can be applied to estimate the IRS-assisted channel estimation problem.

III. MULTIOBJECTIVE EVOLUTIONARY APPROACH TO CHANNEL ESTIMATION

The MOEA has been widely used to solve MOP [32], [47], which can optimize multiple objectives simultaneously and obtain a set of nondominated solutions approximating the Pareto-optimal front. In this section, we first transform the single objective optimization problem (15) to a bi-objective optimization problem; and then a hybrid evolutionary algorithm is proposed by incorporating an IHT-based local search to achieve multiobjective optimization result. Finally, the Pareto knee point is chosen as the solution. However, the selection of a point from the Pareto front depends on the specific problem and the Pareto knee is not always the most appropriate solution. Therefore, to address above difficulties and achieve accurate selection of knee regions, the maximum slope of the Pareto plot is developed to determine the final solution by at a given point.

Considering the measurement error and sparse constraint as two objectives function, the CS-based sparse recovery problem (15) can be formulated as

$$\min_{\mathbf{h}} \mathbf{G}(\mathbf{h}) = \min_{\mathbf{h}} \left(f_1(\mathbf{h}), f_2(\mathbf{h}) \right), \quad (16)$$

where $f_1(\mathbf{h}) = \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{h}\|_2^2$ denotes measurement error term and $f_2(\mathbf{h}) = \|\mathbf{h}\|_0$ denotes sparse constraint term.

A MOEA based on DE [32], [48], [49] is an efficient and reliable multiobjective evolutionary algorithm to solve the MOP (16). However, the DE algorithm exploits the preferences on the knee region of weakly PF, which lead to similar individuals in population. Furthermore, the performance of DE depends on the setting of the original population. To overcome the weakness in DE, we introduce a hybrid MOEA that exploit the the IHT-based local search and maximum slope selection operator to reduce the computational cost and determine the final solution, respectively.

The procedure of the proposed hybrid MOEA scheme is provided in Algorithm 1. Let $P^t = [\mathbf{h}_{1,t}, \mathbf{h}_{2,t}, \dots, \mathbf{h}_{N,t}]$ be the current population with size of N solutions, where $\mathbf{h}_{i,t}$ denotes the i -th solution of the population and t denotes the generation to which the population belongs. P_{mut}^t and P_{Cro}^t are the number of individuals, which are obtained by the corresponding mutation and crossover operators, respectively. Then, an IHT-based local search scheme is used to improve the local search performance of the hybrid MOEA. If the termination condition is satisfied, the optimal PF of algorithm is obtained. For decision process, the maximum slope of the Pareto plot method is used to determine the knee points of the problem.

Algorithm 1 The hybrid evolutionary-based sparse channel estimation

- 1: **Initialization:** $t = 0$, generate P^t randomly and set the stopping criterion.
 - 2: **Input:** \mathbf{y} , Φ
 - 3: **For:** $i = 1$ to N_t .
 - 4: **Implementation of mutation operator:** Select k individual from population P^t . The mutation operator is applied to obtain the $P_{\text{mut}}^t = \mathbf{Mutation}(P^t)$;
 - 5: **Implementation of crossover operator:** Select k individual from population P_{mut}^t . The mutation operator is applied to obtain the $P_{\text{Cro}}^t = \mathbf{Crossover}(P_{\text{mut}}^t)$;
 - 6: **Implementation of local search operator:** Select individual from P_{Cro}^t to generate P_{com}^t according to IHT-based local search operation;
 - 7: **Implementation of selection operator:** The number of solutions dominated by (28);
 - 8: **End For**
 - 9: **Implementation of the final solution:** The Pareto knee point is determined by (29);
 - 10: **Output:** \mathbf{h}
-

A. Mutation and crossover operators

Mutation and crossover aim to generate the new individuals and the corresponding offspring solution controlled by mutation and crossover factor. The mutation factor F and crossover factor C largely affect the searching ability and premature convergence of algorithm. Therefore, adaptive DE algorithm or its variations are exploited, e.g., self-adaptive DE (SaDE) [50]. In DE, the random value is usually used in crossover operator to dynamically generate its offspring. However, this random nature may lead to instability of offspring solution. To address the concern, an improved mutation and crossover operator is exploited to generate promising solutions. We adopt a factor-adaptive mechanism for mutation operation, which is given as follows

$$\mathbf{v}_{i,t} = \mathbf{h}_{i,t} + F (\mathbf{h}_{r2,t} - \mathbf{h}_{r3,t}), \quad (17)$$

where the $\mathbf{h}_{r2,t}$, and $\mathbf{h}_{r3,t}$ are randomly selected within the current population P^t and they are different from $\mathbf{h}_{i,t}$. The corresponding mutation factor F is considered to scale the difference results, that is

$$F = \begin{cases} F_0, & \text{if iteration} = 1 \\ F_0 e^{-2t/G_{\max}} & \text{otherwise} \end{cases} \quad (18)$$

and F_0 is the initial value and G_{\max} is the maximum iterations.

After the mutation operation, a crossover operator is carried out on $\mathbf{v}_{i,t}$ and $\mathbf{h}_{i,t}$ to generate a trial vector, which is expressed as follows

$$\mathbf{z}_{i,t} = \begin{cases} \mathbf{v}_{i,t}, & \text{if } r \leq C \\ \mathbf{h}_{i,t} & \text{otherwise,} \end{cases} \quad (19)$$

where r is a random value within the interval $[0,1]$, C is given by

$$C = \begin{cases} C_0, & \text{if iteration} = 1 \\ C_0 e^{-2t/G_{\max}} & \text{otherwise,} \end{cases} \quad (20)$$

and C_0 is the initial value.

B. IHT-based local search operation

To generate better new individuals, an additional operation is developed, where the generated individuals are further perturbed using a local search strategy improved from IHT [51]. To be specific, we first introduce the IHT algorithm in detail and then describe how the local search strategy is incorporated within the corresponding hybrid MOEA.

1) *Introduction of IHT*: IHT is one of gradient-based iterative algorithm that can solve the optimization problem (15). Here, the IHT-based iterative mechanism is integrated into the proposed MOEA that aims to improve local search results, which can be treat as a special case of the following optimization problem

$$\min_{\mathbf{z}_{i,t}} \{ \mathbf{g}(\mathbf{z}_{i,t}) = f_1(\mathbf{z}_{i,t}) + \eta_i f_2(\mathbf{z}_{i,t}) \}, \quad \mathbf{z}_{i,t} \in P_{\text{Cro}}^t \quad (21)$$

where η_i is a coefficient that balances the tradeoff two competing function.

The optimization problem (21) is solved through a sequence of iterations $\{ \mathbf{z}_{i,t}^{(k)}, k = 1, 2, \dots, K \}$, where $\mathbf{z}_{i,t}^{(k+1)}$ is obtained from the previous solution $\mathbf{z}_{i,t}^{(k)}$ by solving the following optimization problem:

$$\begin{aligned} \mathbf{z}_{i,t}^{(k+1)} &= \arg \min_{\mathbf{z}_{i,t}} f_1 \left(\mathbf{z}_{i,t}^{(k)} \right) + \left(\mathbf{z}_{i,t} - \mathbf{z}_{i,t}^{(k)} \right)^T \nabla f_1 \left(\mathbf{z}_{i,t}^{(k)} \right) \\ &\quad + \left(\mathbf{z}_{i,t} - \mathbf{z}_{i,t}^{(k)} \right)^T \nabla^2 f_1 \left(\mathbf{z}_{i,t}^{(k)} \right) \left(\mathbf{z}_{i,t} - \mathbf{z}_{i,t}^{(k)} \right) + \eta_i f_2(\mathbf{z}_{i,t}) \\ &\approx \arg \min_{\mathbf{z}_{i,t}} \left(\mathbf{z}_{i,t} - \mathbf{z}_{i,t}^{(k)} \right)^T \nabla f_1 \left(\mathbf{z}_{i,t}^{(k)} \right) + \frac{\rho_i^{(k)}}{2} \left\| \mathbf{z}_{i,t} - \mathbf{z}_{i,t}^{(k)} \right\|_2^2 + \eta_i f_2(\mathbf{z}_{i,t}) \\ &= \arg \min_{\mathbf{z}_{i,t}} \frac{1}{2} \left\| \mathbf{z}_{i,t} - \mathbf{d}_{i,t}^{(k)} \right\|_2^2 + \frac{\eta_i}{\rho_i^{(k)}} f_2(\mathbf{z}_{i,t}), \end{aligned} \quad (22)$$

where the matrix $\rho_i^{(k)} \mathbf{I}$ is an estimation of the Hessian matrix $\nabla^2 f_1(\mathbf{z}_{i,t}^{(k)})$ and

$$\mathbf{d}_{i,t}^{(k)} = \mathbf{z}_{i,t}^{(k)} - \frac{1}{\rho_i^{(k)}} \nabla f_1(\mathbf{z}_{i,t}^{(k)}). \quad (23)$$

Now, the following equation can be obtained

$$\mathbf{z}_{i,t}^{(k+1)} = \mathbf{S} \left(\mathbf{d}_{i,t}^{(k)}, \frac{\eta_i}{\rho_i^{(k)}} \right) = \arg \min_{\mathbf{z}_{i,t}} \frac{1}{2} (\mathbf{z}_{i,t} - \mathbf{d}_{i,t}^{(k)})^2 + \frac{\eta_i}{\rho_i^{(k)}} f_2(\mathbf{z}_{i,t}) \quad (24)$$

where the closed-form solution of $\mathbf{S}(\mathbf{d}_{i,t}^{(k)})$ is given by [52]

$$\mathbf{S}(\mathbf{d}_{i,t}^{(k)}) = \begin{cases} (\mathbf{d}_{i,t}^{(k)})_j & \text{if } (\mathbf{d}_{i,t}^{(k)})_j^2 \geq \frac{\eta_i}{\rho_i^{(k)}} \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

and $(\cdot)_j$ denotes the j -th component of a vector.

2) *Determining η_i and ρ_i* : The performance of the algorithm greatly depends on the parameters η_i and ρ_i . It worth noting that a lot of existing strategies have been investigated to choose the optimal coefficient η_i by a continuation procedure. However, traditional IHT algorithms employ the steady-state strategy, in which the same scalar η_i is utilized in iterative processing. The strategy can provide a rather fixed search path but easily fall into suboptimal solutions. To address this concern, some of the existing MOEA execute more effective parallel searching by using a vector $\boldsymbol{\eta}_i = [\eta_{i,1}, \eta_{i,2}, \dots, \eta_{i,J}]$ in each generation, where each element of vector η_i is randomly selected between 0 and 1 respectively. It is worth noting that the $\{\eta_{i,j}\}_{j=1,\dots,J}$ is different from the λ in (15), due to $\boldsymbol{\eta}_i$ is formed by random vector without any optimization processes. The operation is very crucial for diversifying the search along weakly Pareto front. However, it can be noticed in optimization problem (21) that if a large of $\eta_{i,j}$ is chosen ($\eta_{i,j} \geq \|\Phi^T \mathbf{y}_i\|_\infty$), the zero vector would be the unique solution [53]. Therefore, to avoid this phenomenon, a more accurate parallel local search strategy is developed in IHT-based local search operation. Specifically, in each generation, we use a vector $\boldsymbol{\kappa}$ whose elements are independent and identically distributed standard normal random variables, produced by

$$\boldsymbol{\eta}_i = \boldsymbol{\kappa} \|\Phi^T \mathbf{y}_i\|_\infty, \quad (26)$$

which improve the local search and guarantees useless attempts on generating zero solutions.

Next, the parameter $\rho_i^{(k)}$ is to be determined. It is shown that the IHT algorithm has different strategies for choosing the parameter $\rho_i^{(k)}$. Inspired by [54], the Barizilai-Borwein method is suggested to determine the parameter $\rho_i^{(k)}$. That is

$$\begin{aligned} \rho_i^{(k)} &= \min_{\rho_i^{(k)}} \left\| \rho_i^{(k)} \left(\mathbf{z}_{i,t}^{(k)} - \mathbf{z}_{i,t}^{(k+1)} \right) - \left(\nabla f_1 \left(\mathbf{z}_{i,t}^{(k)} \right) - \nabla f_1 \left(\mathbf{z}_{i,t}^{(k+1)} \right) \right) \right\|_2^2 \\ &= \frac{\left(\mathbf{z}_{i,t}^{(k)} - \mathbf{z}_{i,t}^{(k+1)} \right)^T \left(\nabla f_1 \left(\mathbf{z}_{i,t}^{(k)} \right) - \nabla f_1 \left(\mathbf{z}_{i,t}^{(k+1)} \right) \right)}{\left(\mathbf{z}_{i,t}^{(k)} - \mathbf{z}_{i,t}^{(k+1)} \right)^T \left(\mathbf{z}_{i,t}^{(k)} - \mathbf{z}_{i,t}^{(k+1)} \right)}. \end{aligned} \quad (27)$$

The detailed procedure of local search strategy is summarized in Algorithm 2. This procedure is crucial to guarantees improvement in two different ways: one is an improved spread within the set of nondominated solutions and the other is a tendency to move from dominated toward nondominated solutions. This improvement is beneficial to obtain the optimal nondominated solutions on the PF.

Algorithm 2 Implementation of local search operation

- 1: **Initialization:** P_{Cro}^t
 - 2: **Input:** \mathbf{y} , Φ , I_{max}
 - 3: **For:** $i = 1$ to N_t .
 - 4: $\mathbf{z}_{i,t}^{(1)} = P_{\text{Cro}}^t(i)$, $\mathbf{b}^0 = \mathbf{0}^{M \times 1}$
 - 5: **For:** $k = 1$ to I_{max} .
 - 6: $\mathbf{b}^{(k+1)} = \mathbf{b}^{(k)} - \left(\Phi \mathbf{z}_{i,t}^{(k)} - \mathbf{y} \right)$
 - 7: $\mathbf{z}_{i,t}^{(k+1)} = \mathbf{S} \left(\Phi^H \mathbf{b}^{(k+1)}, \frac{\lambda^{(k)}}{\rho^{(k)}} \right)$
 - 8: **End For**
 - 9: $\mathbf{s}_{i,t} = \mathbf{z}_{i,t}^{(I_{\text{max}})}$
 - 10: **End For**
 - 11: **Output:** $\{\mathbf{s}_{1,t}, \mathbf{s}_{2,t}, \dots, \mathbf{s}_{N_t,t}\}$
-

C. Selection operator

The selection scheme also different from that of existing DE algorithms. The solution of current population is selected to generate the next population, and its corresponding temporary

solution is given by using the following rule:

$$\mathbf{h}_{i,t+1} = \begin{cases} \mathbf{s}_{i,t} & \text{if } \mathbf{g}(\mathbf{s}_{i,t}) \leq \mathbf{g}(\mathbf{z}_{i,t}), \\ \mathbf{z}_{i,t} & \text{otherwise,} \end{cases} \quad (28)$$

where $\mathbf{s}_{i,t}$ is the output result of IHT-based local search operation.

By using the (28), each temporary solution is compared with its counterpart in the current population. The tournament selection is carried out to select the lower objective function value that will survive to the population of the next generation. As a result, it is guaranteed that each solution of the next generation are as well as or better than the current generation. It is worth noting that temporary solution is not compared against all the solutions in the current population, but only against its counterpart.

D. Acquisition of the final solution

In MOPs, the Pareto knee region provides promising trade-off between the two conflicting objectives functions, and the near optimal solution is chosen in this region. It proves that the Pareto knee point can be regarded as the final solution [31], [55], [56]. However, the problem being tackled in this paper is more difficult than the problems described in [55], [56]. The mostly problem is that: firstly, the problem is NP-hard, it is not easy to guarantee that the estimated PF produced by the algorithm converges to the true PF. When some of the solutions not lie on the PF, the distort result will be computed and solution strayed away from the knee region, which leading to an incorrect estimation of the global optimal solution. The second difficulty is that two objective functions $\|\mathbf{h}_i\|_0$ and $\|\mathbf{y}_i - \Phi\mathbf{h}_i\|_2^2$ are involved in the PF, but the ranges of these two conflicting functions generally have different magnitudes. This can diminish the significance of the differences in tradeoff angles between neighboring solutions. Since the population size is usually fixed, the third difficulty is hard to obtain the solutions of population that adequately collect the full range of the PF.

To overcome above difficulties and achieve accurate solution of selection, first we normalize the PF by its maximum $\|\mathbf{h}_i\|_0$ and $\|\mathbf{y}_i - \Phi\mathbf{h}_i\|_2^2$ value, which can avoid the second difficulty mentioned above. Next, the maximal slope is computed to choose the Pareto knee point as the final solution. The basic thought is to determine the final solution by the maximum slope of the

Pareto plot at a given point. To be specific, for a solution \mathbf{h}_i and corresponding value of two objective function $(f_1(\mathbf{h}_i), f_2(\mathbf{h}_i))$, its slope variance $\Delta\phi_i$ is

$$\Delta\phi_i = \frac{\phi_i}{\phi_{i-1}}, \quad \phi_i = \arctan\left(\frac{f_1(\mathbf{h}_i) - f_1(\mathbf{h}_{i+1})}{f_2(\mathbf{h}_i) - f_2(\mathbf{h}_{i+1})}\right), \quad (29)$$

where ϕ_i and ϕ_{i-1} represent the slope at point i and $i-1$, respectively. $\Delta\phi_i$ is the slope change at point i and $\Delta\phi_0 = \pi/2$. It follows that the solution with the maximum $\Delta\phi_i$ value is considered as the nearly optimal solution.

IV. CONVERGENCE ANALYSIS

In this section, the convergence analysis of the proposed MOEA algorithm is provided to show that how the proposed approach can be applied to recover a sparse channel when the measurement matrix satisfies a properly RIP condition. For a given matrix Φ , the $2s$ -th order Restricted Isometry Constants (RIC) are the largest α and smallest β , such that [57]

$$\alpha\|\mathbf{x}_1 + \mathbf{x}_2\|_2^2 \leq \|\Phi(\mathbf{x}_1 + \mathbf{x}_2)\|_2^2 \leq \beta\|\mathbf{x}_1 + \mathbf{x}_2\|_2^2 \quad (30)$$

holds for all s -sparse vector \mathbf{x}_1 and \mathbf{x}_2 . In the (30), the vectors \mathbf{x}_1 and \mathbf{x}_2 are said to be s -sparse if the number of nonzero components of \mathbf{x}_1 and \mathbf{x}_2 are smaller than or equal to s , i.e., $\|\mathbf{x}_1\|_0 \leq s$, $\|\mathbf{x}_2\|_0 \leq s$. Next, the convergence of the proposed hybrid MOEA is as follows. To simplify the expression, the generation number t of the population is omitted in proof.

Theorem 1 For an arbitrary s -sparse channel \mathbf{h}_i , given $\mathbf{y}_i = \Phi\mathbf{h}_i + \mathbf{e}_i$, where Φ satisfies RIC with $J_\theta < \alpha$ and $\frac{2-\sqrt{2}}{4} < \theta < \frac{1}{2}$, at least

$$k^* = \max \left\{ \left\lceil \frac{2 \ln \left((\sqrt{\alpha}\|\mathbf{n}_i\|_2) / \sqrt{J_\theta}\|\mathbf{y}_i\|_2 \right)}{\ln\left(\frac{\theta^2+2\theta-1}{\theta^2}\right)} \right\rceil, 1 \right\} \quad (31)$$

iterations, the proposed hybrid evolutionary algorithm calculates a solution satisfying

$$\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \leq \left(1 + \sqrt{\frac{J_\theta}{\alpha}}\right) \|\mathbf{n}_i\|_2 + \delta, \quad (32)$$

where $\delta = \sqrt{\frac{J_\theta \lambda N \theta^2}{\alpha(1-2\theta)}}$, $J_\theta = \frac{(2+8\theta)^2(2\theta-\theta^2)}{(8\theta-8\theta^2-1)^2}$ and $\|\mathbf{n}_i\|_2 = \sqrt{\frac{2-2\theta}{1-2\theta}}\|\mathbf{e}_{v_i}\|_2 + \sqrt{\frac{2-2\theta}{\theta^2} + \frac{(\theta-\theta^2)^2}{(1-2\theta)(2\theta-\theta^2)}}\|\mathbf{e}_i\|_2$ with $\|\Phi(\mathbf{h}_i - \mathbf{v}_i)\|_2 \leq \|\mathbf{e}_{v_i}\|_2$.

Proof: The convergence result is based on the RIP in (30), under the constraint that $\|\mathbf{h}_i\|_0 \leq s$, where s is the maximum sparsity of channel. By the sparsity of the RIS-assisted mmWave channel, it is observed in (15) that the channel \mathbf{h}_i is sparse and the sparsity is less than its

dimensionality. However, \mathbf{h}_i^{k+1} is the intermediate result of the algorithm at the $k+1$ -th iteration. With this case, \mathbf{h}_i^{k+1} is no longer guaranteed to be s -sparse. This issue has brought challenges to convergence analysis. To deal with this issue, we let $\mathbf{h}_{i,s}^{k+1}$ be the best approximation to \mathbf{h}_i^{k+1} with no more than s non-zero elements. This implies that

$$\|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 \leq \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 \quad (33a)$$

$$\Leftrightarrow \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2 \leq 2\|\mathbf{h}_i^{k+1} - \mathbf{h}_i\|_2. \quad (33b)$$

Thus we first derive an upper bound for the error $\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2$. This will be done using a chain of inequalities, the first of which is a simple triangle inequality:

$$\begin{aligned} & \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 \\ &= \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 + \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 - 2\langle \mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}, \mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i \rangle \\ &= \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 - \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 - 2\langle \mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}, \mathbf{h}_i^{k+1} - \mathbf{h}_i \rangle \\ &\stackrel{(a)}{=} \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 - \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 + 2 \times \frac{1}{4} (\|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1} - (\mathbf{h}_i^{k+1} - \mathbf{h}_i)\|_2^2 - \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2) \\ &\stackrel{(b)}{\leq} \frac{1}{2} \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 - \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 + \frac{1}{2\theta} (\|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 - \|\mathbf{h}_i^{k+1} - \mathbf{h}_i\|_2^2) + \frac{1}{2\theta(1-\theta)} \|\mathbf{h}_i^{k+1} - \mathbf{h}_i\|_2^2 \\ &\stackrel{(c)}{\leq} \frac{1}{2} \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 - \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 + \frac{1}{\theta} \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2 + \frac{1}{2\theta(1-\theta)} \|\mathbf{h}_i^{k+1} - \mathbf{h}_i\|_2^2 \\ &\stackrel{(d)}{\leq} \frac{3}{2} \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 - \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 + \frac{1}{\theta} \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2 + \frac{1}{2\theta(1-\theta)} \|\mathbf{h}_i^{k+1} - \mathbf{h}_i\|_2^2 \\ &\stackrel{(e)}{\leq} \frac{4\theta + 1}{\theta} \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2 \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 + \left(\frac{6\theta^2 - 6\theta + 1}{2\theta(1-\theta)} \right) \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2, \end{aligned} \quad (34)$$

where (a) is from the fact that $\langle \mathbf{a}, \mathbf{b} \rangle = \frac{1}{4} (\|\mathbf{a} + \mathbf{b}\|_2^2 - \|\mathbf{a} - \mathbf{b}\|_2^2)$ and (b) is from

$$\|(1-\theta)\mathbf{a} + \theta\mathbf{b}\|_2^2 \geq 0 \quad (35a)$$

$$\Leftrightarrow (1-\theta)\|\mathbf{a}\|_2^2 + \theta\|\mathbf{b}\|_2^2 - \theta(1-\theta)\|\mathbf{a} - \mathbf{b}\|_2^2 \geq 0 \quad (35b)$$

$$\Leftrightarrow \|\mathbf{a} - \mathbf{b}\|_2^2 \leq \frac{1}{\theta} \|\mathbf{a}\|_2^2 + \frac{1}{1-\theta} \|\mathbf{b}\|_2^2 \quad (35c)$$

$$\Leftrightarrow \|\mathbf{a} - \mathbf{b}\|_2^2 \leq \frac{1}{\theta} (\|\mathbf{a}\|_2^2 - \|\mathbf{b}\|_2^2) + \frac{1}{\theta(1-\theta)} \|\mathbf{b}\|_2^2, \quad (35d)$$

for any $0 < \theta < 1$, $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$, (c) is from the fact that

$$\begin{aligned} \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 - \|\mathbf{h}_i^{k+1} - \mathbf{h}_i\|_2^2 &\stackrel{(33a)}{\leq} 2\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 (\|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2 - \|\mathbf{h}_i^{k+1} - \mathbf{h}_i\|_2) \\ &\leq 2\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2 \end{aligned} \quad (36)$$

and (d) is from $\|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1}\|_2^2 \geq \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 - \|\mathbf{h}_i - \mathbf{h}_{i,s}^{k+1}\|_2^2$, and (e) is from (33b) and

$$\|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 - \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 \stackrel{(33b)}{\leq} 3\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2(\|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2 - \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2). \quad (37)$$

After rearranging the terms (34), considering the condition $\frac{2-\sqrt{2}}{4} < \theta < \frac{1}{2}$ in Theorem 1, we obtain

$$\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \leq \frac{(2+8\theta)(1-\theta)}{8\theta-8\theta^2-1} \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2. \quad (38)$$

On the other hand, for the vector $\hat{\mathbf{h}}_i = (1-\theta)\mathbf{h}_{i,s}^{k+1} + \theta\mathbf{h}_i^{k+1}$, we have

$$\begin{aligned} \left\| \Phi \left((1-\theta)\mathbf{h}_{i,s}^{k+1} + \theta\mathbf{h}_i^{k+1} \right) - \mathbf{y}_i \right\|_2^2 &\geq \|\Phi\mathbf{h}_{i,s}^{k+1} - \mathbf{y}_i\|_2^2 - \theta^2 \|\Phi(\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i^{k+1})\|_2^2 \\ &\stackrel{(35c)}{\geq} \|\Phi\mathbf{h}_{i,s}^{k+1} - \mathbf{y}_i\|_2^2 + \theta^2 \left(\frac{1}{\theta-1} \|\Phi\mathbf{h}_i^{k+1} - \mathbf{y}_i\|_2^2 - \frac{1}{\theta} \|\Phi\mathbf{h}_{i,s}^{k+1} - \mathbf{y}_i\|_2^2 \right) \\ &= \frac{\theta^2}{\theta-1} \|\Phi\mathbf{h}_i^{k+1} - \mathbf{y}_i\|_2^2 + (1-\theta) \|\Phi\mathbf{h}_{i,s}^{k+1} - \mathbf{y}_i\|_2^2, \end{aligned} \quad (39)$$

and

$$\left\| \Phi \left((1-\theta)\mathbf{h}_{i,s}^{k+1} + \theta\mathbf{h}_i^{k+1} \right) - \mathbf{y}_i \right\|_2^2 \leq (1-\theta)^2 \|\Phi\mathbf{h}_{i,s}^{k+1} - \mathbf{y}_i\|_2^2 + \theta^2 \|\Phi\mathbf{h}_i^{k+1} - \mathbf{y}_i\|_2^2. \quad (40)$$

Combining (39) and (40), considering the condition $\frac{2-\sqrt{2}}{4} < \theta < \frac{1}{2}$ in Theorem 1, we obtain

$$\|\Phi\mathbf{h}_{i,s}^{k+1} - \mathbf{y}_i\|_2^2 \leq \tau \|\Phi\mathbf{h}_i^{k+1} - \mathbf{y}_i\|_2^2. \quad (41)$$

where $\tau = \frac{2\theta-\theta^2}{(1-\theta)^2}$.

Therefore, the (38) can be further expressed as

$$\begin{aligned} \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 &\leq \left(\frac{(2+8\theta)(1-\theta)}{8\theta-8\theta^2-1} \right)^2 \|\mathbf{h}_{i,s}^{k+1} - \mathbf{h}_i\|_2^2 \\ &\stackrel{(a)}{\leq} \frac{1}{\alpha} \left(\frac{(2+8\theta)(1-\theta)}{8\theta-8\theta^2-1} \right)^2 (\|\mathbf{y}_i - \Phi\mathbf{h}_{i,s}^{k+1}\|_2^2 + \|\mathbf{e}_i\|_2^2) \\ &\stackrel{(b)}{\leq} \frac{J_\theta}{\alpha} \left(\|\mathbf{y}_i - \Phi\mathbf{h}_i^{k+1}\|_2^2 + \frac{1}{\tau} \|\mathbf{e}_i\|_2^2 \right). \end{aligned} \quad (42)$$

where (a) is from the RIP condition in (30) and (b) follows from the fact that (41) and $J_\theta = \frac{(2+8\theta)^2(2\theta-\theta^2)}{(8\theta-8\theta^2-1)^2}$.

According to (28) in the proposed algorithm, we consider the value \mathbf{h}_i^{k+1} of (42) in the following two cases.

1) *Case 1*: if $\mathbf{g}(\mathbf{h}_i^{k+1}) \leq \mathbf{g}(\mathbf{z}_i^k)$ in (28), we have

$$\|\mathbf{y}_i - \Phi \mathbf{h}_i^{k+1}\|_2^2 - \|\mathbf{y}_i - \Phi \mathbf{z}_i^k\|_2^2 \leq \lambda(\|\mathbf{z}_i^k\|_0 - \|\mathbf{h}_i^{k+1}\|_0) \leq \lambda N, \quad (43)$$

where N is dimension of the vectors \mathbf{h}_i^{k+1} , \mathbf{z}_i^k , $k = 0, \dots, K - 1$.

Substituting (43) into (42) gives

$$\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 \leq \frac{J_\theta}{\alpha} \left(\|\mathbf{y}_i - \Phi \mathbf{z}_i^k\|_2^2 + \lambda N + \frac{1}{\tau} \|\mathbf{e}_i\|_2^2 \right). \quad (44)$$

According to (19) in the proposed algorithm, each iteration of the value \mathbf{z}_i^k depends on \mathbf{v}_i^k and \mathbf{h}_i^k . Thus, if $r \leq C$, the value of \mathbf{v}_i^k satisfies

$$\begin{aligned} \|\Phi F(\mathbf{h}_{r2}^k - \mathbf{h}_{r3}^k)\|_2^2 &\stackrel{(a)}{=} \|\Phi(\mathbf{v}_i^k - \mathbf{h}_i^k)\|_2^2 \\ &= \frac{1}{\theta^2} (\|\theta(\Phi \mathbf{v}_i^k - \mathbf{y}_i) - \theta(\Phi \mathbf{h}_i^k - \mathbf{y}_i)\|_2^2) \\ &\stackrel{(b)}{<} \frac{1}{\theta^2} ((1 - \theta)\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2 + \theta\|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2)^2 \\ &= \frac{1 - \theta}{\theta^2} (\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2 - \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2) \\ &\quad \times ((1 - \theta)\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2 + (1 + \theta)\|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2) + \frac{1}{\theta^2} \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2, \end{aligned} \quad (45)$$

where (a) is from the (17) and (b) is from the condition $\theta < \frac{1}{2}$.

Moreover, it is easy to see that

$$\begin{aligned} \|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2^2 - \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2 &= 2\langle \Phi(\mathbf{h}_i^k - \mathbf{v}_i^k), \mathbf{y}_i - \Phi \mathbf{h}_i^k \rangle + \|\Phi(\mathbf{h}_i^k - \mathbf{v}_i^k)\|_2^2 \\ &\stackrel{(a)}{=} \frac{1}{2} (\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2^2 - \|\Phi(\mathbf{h}_i^k - \mathbf{v}_i^k) - (\mathbf{y}_i - \Phi \mathbf{h}_i^k)\|_2^2) + \|\Phi(\mathbf{h}_i^k - \mathbf{v}_i^k)\|_2^2 \\ &\geq \frac{1}{2} (\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2^2 - \|\Phi(\mathbf{h}_i^k - \mathbf{v}_i^k)\|_2^2 - \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2) + \|\Phi(\mathbf{h}_i^k - \mathbf{v}_i^k)\|_2^2 \\ &= \frac{1}{2} (\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2^2 - \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2) + \frac{1}{2} \|\Phi(\mathbf{h}_i^k - \mathbf{v}_i^k)\|_2^2, \end{aligned} \quad (46)$$

where (a) is from the fact that $\langle \mathbf{a}, \mathbf{b} \rangle = \frac{1}{4}(\|\mathbf{a} + \mathbf{b}\|_2^2 - \|\mathbf{a} - \mathbf{b}\|_2^2)$.

From the inequality (46), we notice that $\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2 \geq \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2$, which along with inequality (45) yields the desired result given by

$$\begin{aligned} \|\Phi F(\mathbf{h}_{r2}^k - \mathbf{h}_{r3}^k)\|_2^2 &\leq \frac{2(1 - \theta)}{\theta^2} (\|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2^2 - \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2 \|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2) + \frac{1}{\theta^2} \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2 \\ &\leq \frac{2\theta - 1}{\theta^2} \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2 + \frac{2(1 - \theta)}{\theta^2} \|\mathbf{y}_i - \Phi \mathbf{v}_i^k\|_2^2 \\ &\leq \frac{2\theta - 1}{\theta^2} \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2 + \frac{2(1 - \theta)}{\theta^2} \|\Phi(\mathbf{h}_i^k - \mathbf{v}_i^k)\|_2^2 + \frac{2(1 - \theta)}{\theta^2} \|\mathbf{e}_i\|_2^2. \end{aligned} \quad (47)$$

Defining $\|\Phi(\mathbf{h}_i - \mathbf{v}_i)\|_2 \leq \|\mathbf{e}_{v_i}\|_2$, substituting (47) into (44) gives

$$\begin{aligned} \|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 &\leq \frac{J_\theta}{\alpha} \left(\left\| \mathbf{y}_i - \Phi \left(\mathbf{h}_i^k + F(\mathbf{h}_{r2}^k - \mathbf{h}_{r3}^k) \right) \right\|_2^2 + \lambda N + \frac{1}{\tau} \|\mathbf{e}_i\|_2^2 \right) \\ &\leq \frac{J_\theta}{\alpha} \left(\left(\frac{\theta^2 + 2\theta - 1}{\theta^2} \right) \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2 + \lambda N + \frac{2 - 2\theta}{\theta^2} \|\mathbf{e}_{v_i}\|_2^2 + \left(\frac{2 - 2\theta}{\theta^2} + \frac{1}{\tau} \right) \|\mathbf{e}_i\|_2^2 \right), \end{aligned} \quad (48)$$

With the condition in Theorem 1, $J_\theta < \alpha$, $\theta < \frac{1}{2}$, we can iterate the above expression as

$$\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \leq \sqrt{\frac{J_\theta}{\alpha}} \left(\frac{\theta^2 + 2\theta - 1}{\theta^2} \right)^{\frac{k}{2}} \|\mathbf{y}_i\|_2 + \sqrt{\frac{J_\theta}{\alpha}} \|\mathbf{n}_i\|_2 + \delta. \quad (49)$$

where $\|\mathbf{n}_i\|_2 = \sqrt{\frac{2-2\theta}{1-2\theta}} \|\mathbf{e}_{v_i}\|_2 + \sqrt{\frac{2-2\theta}{1-2\theta} + \frac{\theta^2}{\tau(1-2\theta)}} \|\mathbf{e}_i\|_2$ and $\delta = \sqrt{\frac{J_\theta \lambda N \theta^2}{\alpha(1-2\theta)}}$.

From (31), by recursive iteration, $\sqrt{\frac{J_\theta}{\alpha}} \left(\frac{\theta^2 + 2\theta - 1}{\theta^2} \right)^{\frac{k}{2}} \|\mathbf{y}_i\|_2 \leq \|\mathbf{n}_i\|_2$ should be satisfied. Clearly, for each $k > k^*$, there exists

$$\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \leq \left(1 + \sqrt{\frac{J_\theta}{\alpha}} \right) \|\mathbf{n}_i\|_2 + \delta. \quad (50)$$

2) *Case 2*: From (28), it is easy to show that \mathbf{z}_i^k is equivalent to \mathbf{h}_i^{k+1} , and thus (42) is rewritten as

$$\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2^2 \leq \frac{J_\theta}{\alpha} \left(\left(\frac{\theta^2 + 2\theta - 1}{\theta^2} \right) \|\mathbf{y}_i - \Phi \mathbf{h}_i^k\|_2^2 + \frac{2 - 2\theta}{\theta^2} \|\mathbf{e}_{v_i}\|_2^2 + \left(\frac{2 - 2\theta}{\theta^2} + \frac{1}{\tau} \right) \|\mathbf{e}_i\|_2^2 \right). \quad (51)$$

Similarly, by recursive iteration, $\sqrt{\frac{J_\theta}{\alpha}} \left(\frac{\theta^2 + 2\theta - 1}{\theta^2} \right)^{\frac{k}{2}} \|\mathbf{y}_i\|_2 \leq \|\mathbf{n}_i\|_2$ will be satisfied. Thus, we can still obtain the following inequality:

$$\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2 \leq \left(1 + \sqrt{\frac{J_\theta}{\alpha}} \right) \|\mathbf{n}_i\|_2 < \left(1 + \sqrt{\frac{J_\theta}{\alpha}} \right) \|\mathbf{n}_i\|_2 + \delta. \quad (52)$$

By comparing the upper bound of the $\|\mathbf{h}_i - \mathbf{h}_i^{k+1}\|_2$ obtained in the above two cases, the one with higher bound value is chosen to satisfy the inequality (32). On the other hand, it is found that if the condition $r > C$ in (17), we have $\mathbf{z}_i^k = \mathbf{h}_i^k$. In this case, above convergence analysis is still valid by setting $\|\Phi F(\mathbf{h}_{r2}^k - \mathbf{h}_{r3}^k)\|_2^2 = 0$. Therefore, the Theorem 1 is guaranteed to converge according to the proposed hybrid evolutionary algorithm. ■

Remark: (i) In general, the convergence of the algorithm relies on the number of iterations. However, in the realistic situation incorporating sparsity defect and measurement error, the

number of iterations is difficult to control. To this end, we show that the algorithm can converge when the number of iterations satisfies (31).

(ii) The error constant δ is considered in convergence result, which is more reasonable for the proposed hybrid evolutionary algorithm. Since the value of $\delta = \sqrt{\frac{J_\theta \theta^2}{\alpha(1-2\theta)} \lambda N}$ relies primarily on (43) with respect to $\mathbf{g}(\mathbf{h}_i^{k+1}) \leq \mathbf{g}(\mathbf{z}_i^k)$, this procedure is implemented in mutation and crossover processing of the algorithm. Our aim is to take the advantage of mutation and crossover operator to generate better new individuals. It is thus theoretically feasible to employ the error constant δ in (32), which would be benefit some of individuals to approximate the Pareto optimal front.

V. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed channel estimation scheme for IRS-assist MIMO system with IRS. We consider the IRS equipped with UPA at both BS and UE sides and operating over a mmWave channel. The performance of proposed hybrid MOEA is experimentally compared to existing channel estimation algorithms, i.e., OMP [20], two-stage estimation exploiting both sparsity and low rankness (TSSR) [58]. We also consider the oracle LS scheme as our benchmark, where the support of the IRS-assisted mmWave channel is assumed to be perfectly known. To validate the generalization of the proposed channel estimation, some randomly generated 1-D sparse $h_{i,t} \in P^0$ are used which are obtained in the following way [59]. At first, the nonzero entries of the channel are randomly selected which constitute the active set. Then, the magnitudes of these entries are obtained from the standard normal distribution.

TABLE I
THE PROPOSED CHANNEL FRAMEWORK PARAMETERS FOR SIMULATION

Parameters	Parameters
Mutation factor: $F_0 = 0.9$	Path loss for BS-IRS link: $\alpha_\rho = 1.4$
Crossover factor: $C_0 = 0.1, 0.2, 0.3, 0.4$	Path loss for IRS-UE link: $\alpha_\mu = 2$
Passive reflecting elements: $M = 8 \times 8$	The BS-IRS horizontal distance: from 20 to 60 m
Carrier frequency: $f_c = 28$	The training length T : 10, 15, 20, 25, 30, 35, 40
The reference distance: $D_0 = 1m$	The number of data path: $L = 6$

To evaluate the performance of channel estimation schemes, we compare the perfect CSI estimation in terms of both the normalized mean square error (NMSE) and the achievable spectral

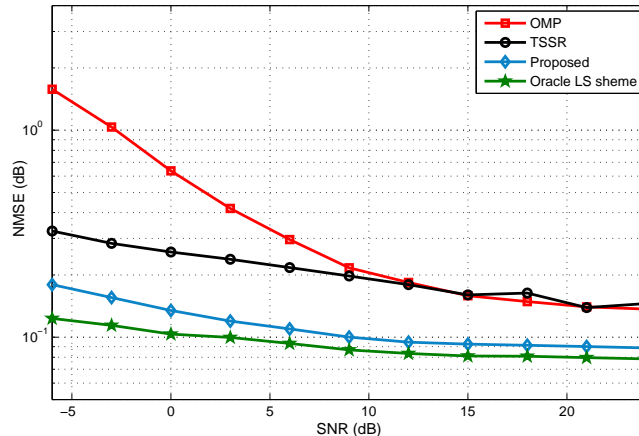


Fig. 4. NMSEs of respective algorithms versus the SNR ($N_t = 36$).

efficiency (SE). The NMSE is considered to quantify the accuracy of channel estimation, which is mathematically defined as $\text{NMSE} \triangleq 10 \log_{10} \frac{\|\mathbf{h} - \hat{\mathbf{h}}\|_2^2}{\|\mathbf{h}\|_2^2}$, where $\hat{\mathbf{h}}$ denotes the estimation for the true channel \mathbf{h} with the considered techniques. Apparently, when the value of NMSE is smaller, the estimated channel $\hat{\mathbf{h}}$ is closer to the true channel \mathbf{h} . In other words, the NMSE value is zero, which indicates that the estimated $\hat{\mathbf{h}}$ is the perfect estimation of \mathbf{h} .

To account for small-scale fading, the Rician fading model is adopted for channel \mathbf{R} between the IRS and the UE, which can be modeled as [4]

$$\mathbf{R}_r = \sqrt{\frac{K}{K+1}} \mathbf{R}_{LoS} + \sqrt{\frac{1}{K+1}} \mathbf{R}_{NLoS} \quad (53)$$

where \mathbf{R}_{LoS} and \mathbf{R}_{NLoS} denote the LoS component and NLoS component components, respectively. The counterparts for the channel \mathbf{H}_r can also be modeled similarly.

In the first simulation, the NMSE performance of our proposed strategy for channel estimation is studied, and some benchmark schemes are introduced for performance comparison and analysis. In Fig. 4 and Fig. 5, we compare the NMSEs and spectral efficiency of different channel estimation schemes versus the SNR levels. It can be observed that the oracle LS estimator achieves the best NMSE performance for all channel estimation method. On the other hand, one can note that, using the hybrid MOEA to solve the proposed IRS-based channel estimation model, which can obtain significant NMSE improvement compared with conventional OMP and TSSR. The reason is that the classical OMP adopts fixed search path and easily fall into suboptimal solutions. On the contrary, the proposed hybrid MOEA exploit the parallel local search strategy for diversifying the search, which can achieve effective performance improvement. In addition,

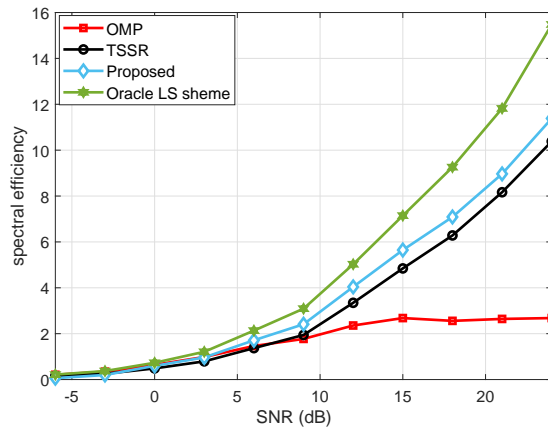


Fig. 5. Spectral efficiency of respective algorithms versus the SNR.

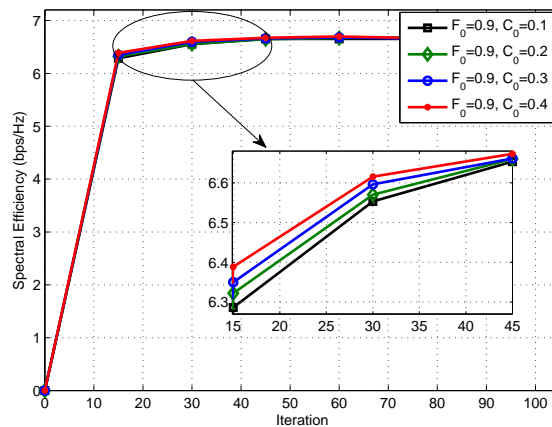
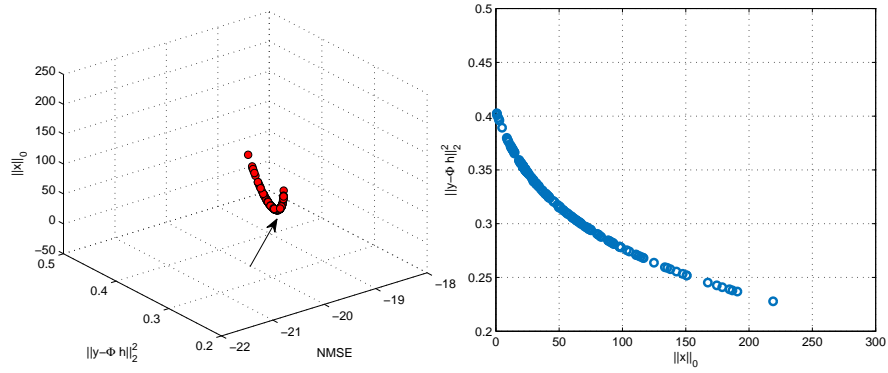


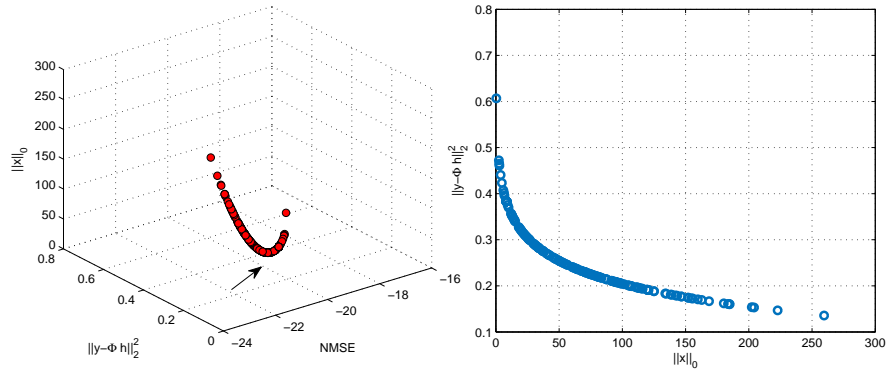
Fig. 6. Convergence behavior of the hybrid MOEA in terms of spectral efficiency.

one can observe that the proposed algorithm performs remarkably with a small gap compared to the benchmark of the oracle LS estimator. The performance gap is attributed to the presence of path loss and noise. These results demonstrate that the proposed hybrid MOEA is more suitable for IRS-based channel estimation.

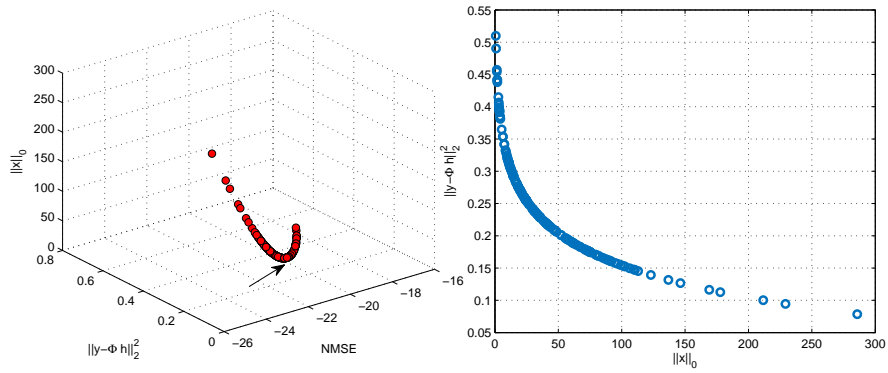
Next, the convergence behavior of the hybrid MOEA is investigated to illustrate how the spectral efficiency behaves with the number of iterations. From Fig. 6, it is observed that the hybrid MOEA can converge to a stable value, and the initiation of crossover operator slightly affects the convergence speed of the hybrid MOEA. Furthermore, the fixed value of spectral efficiency is obtained when the number of iteration over 60 iterations under the crossover factor $C_0 = 0.2, 0.3, 0.4$. It worth noting that if $C_0 = 0.1$ is chosen, the iterative number for stable



(a) SNR = 15db



(b) SNR=25db



(c) SNR=35db

Fig. 7. Relationship between NMSE, measurement error and the sparsity of the solutions on PF and the position of knee regions.

point is reduced to 45. This result demonstrates that the proposed hybrid MOEA can converge to a stable point nearly within 60 iterations.

In the following, we study the existence of a knee area for the proposed hybrid MOEA algorithm. Fig. 7 shows the relationship between the NMSE, the measurement error term and the sparse constraint term in the corresponding solutions. The corresponding left-hand column

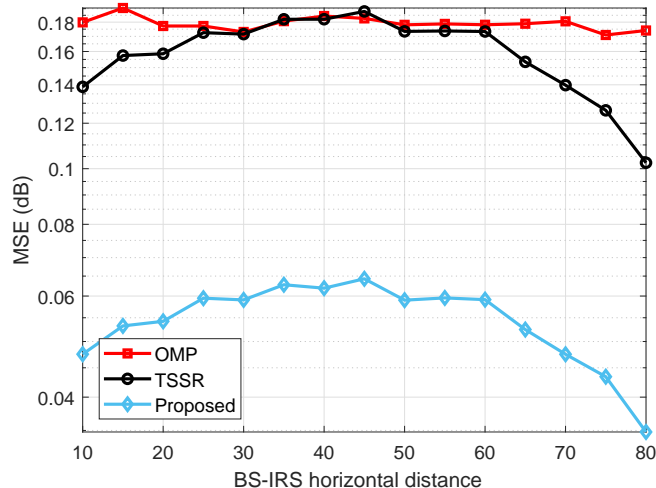


Fig. 8. Achievable rate versus BS-IRS horizontal distance (SNR=15 dB).

in Fig. 7 shows 3-D views of the three variables together. The right-hand column of Fig. 7 depicts 2-D plots of the measurement error $\|\mathbf{y} - \Phi\mathbf{h}\|_2^2$ with change in sparsity. From the of arrow direction in left-hand column of Fig. 7, it is not difficult to find the knee region of Pareto front. On the other hand, (a), (b) and (c) in Fig. 7 show the horizontal position of knee points varies with different SNR levels. We can notice that under the high SNR levels, the sparsity estimation rapidly deteriorates for points to the right of the knee region, while providing a little improvement in measurement error. The results indicate that the Pareto front in the knee areas are consistent with the solution of the true channel. However, it can be found that under the low SNR case, the the knee points are not clear in SNR=15dB of Fig. (7), which may lead to difficult to chose the final solution in knee areas. Therefore, the proposed hybrid MOEA exploiting the maximum slope of the Pareto plot can effectively resolve the issue. It noted that the maximal slope is chosen as an optimal solution of the knee areas to the CS-based sparse recovery problem, because any further attempt to improve sparsity will result in rapid deterioration in $\|\mathbf{y} - \Phi\mathbf{h}\|_2^2$ for only very slightly increases in accuracy.

In Fig.8, we plot the achievable MSE of different channel estimation schemes versus the BS-IRS distance d . From the results, we can see that compared to the OMP, TSSR can achieve nearly the performance of OMP when the BS-IRS horizontal distance from 25 to 60m. On the other hand, it is observed that the proposed scheme using the hybrid MOEA yields a significant gain over the other counterparts, no matter the BS-IRS horizontal distance.

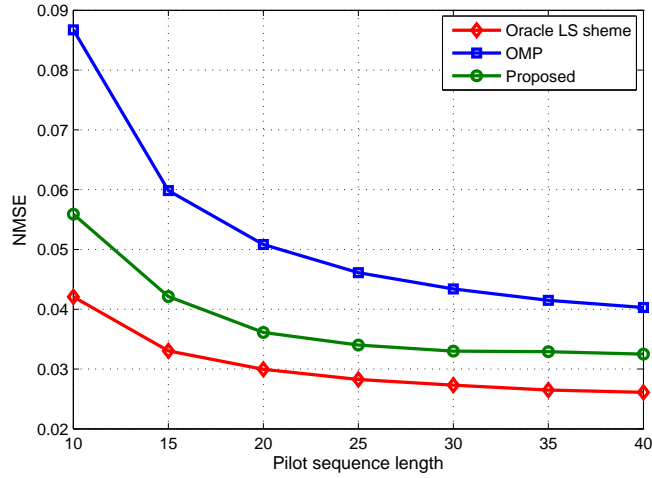


Fig. 9. The NMSE performance versus the pilot sequence length T .

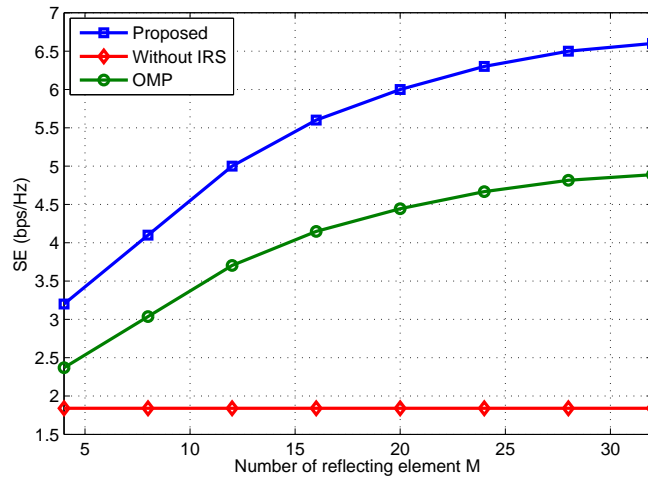


Fig. 10. The SE performance versus the number of reflecting element M .

Next, to see the influence of the training length T on the performance of the algorithms, the NMSE performance versus the pilot sequence length T is provided in Fig. 9. It is obvious that the proposed hybrid MOEA achieves a suboptimal solution and significantly better than the OMP and closely approaches the oracle LS, under the same training length T . This result indicates that to achieve the same NMSE performance, the proposed scheme required training length are far less than the benchmark OMP scheme. Furthermore, we also investigate that the improvement of the NMSE performance of the proposed hybrid MOEA can be negligible when the value of L large than 25, while the OMP need more training to improve the NMSE performance.

Finally, we further analyze and discuss the impact of the proposed channel estimation on the number of IRS elements. Fig. 10 illustrates the SE performance versus the number of reflecting element M under the proposed channel estimation. It can be seen that the SE performance gains achieved by the IRS scheme over the ‘without IRS’ scheme, especially for the large number of phase shifters M . This is mainly related to the fact that an additional strong channel is reflected by the IRS, which can be harvested by the UE. These results also demonstrate that the proposed IRS-based channel estimation is effective to expand the communication range. In addition, as expected, the performance of the method ‘without IRS’ is not affected by the number of the IRS elements in the system.

On the other hand, it is observed that the proposed hybrid MOEA scheme generally increases as the number of reflecting element M . The reasons for the signals reflected by IRS can be added to increase the desired signal strength at the UE, and hence the different frequencies have more degree-of-freedom for the reflection element of IRS design. In next research work, we will consider the proposed method applying to the wideband mmWave systems that the wideband channel is partitioned into multiple subchannels, the original channel will be split in the frequency domain to achieve channel estimation.

VI. CONCLUSION

In this paper, we studied the mmWave channel estimation problem for IRS-assisted communication systems. Under the proposed framework, the cascade channel estimation approach is developed using the properties of Kronecker products. By exploiting the inherent sparse structure of the mmWave channel, we formulate the cascaded channel estimation problem into a CS-based sparse recovery problem that can achieve a substantial training overhead reduction. Under the proposed channel estimation problem, a hybrid multiobjective evolutionary paradigm is developed to achieve high resolution channel estimation. Furthermore, we provide the convergence analysis for the hybrid evolutionary-based channel estimation algorithm. Simulation results show that the proposed method can be significantly improved in terms of NMSE as well as achievable SE.

REFERENCES

- [1] Y. Han, W. Tang, S. Jin, C. Wen, and X. Ma, “Large intelligent surface-assisted wireless communication exploiting statistical CSI,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8238–8242, 2019.

- [2] 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 Transactions on Wireless Communications*, vol. 20, no. 1, pp. 421–439, 2021.
- [3] L. Zhang, Y. Wang, W. Tao, Z. Jia, T. Song, and C. Pan, "Intelligent reflecting surface aided MIMO cognitive radio systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11445–11457, 2020.
- [4] W. Yan, X. Yuan, Z. He, and X. Kuai, "Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1793–1808, 2020.
- [5] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.
- [6] K. Feng, Q. Wang, X. Li, and C. Wen, "Deep reinforcement learning based intelligent reflecting surface optimization for MISO communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 745–749, 2020.
- [7] L. Dong and H. Wang, "Enhancing secure MIMO transmission via intelligent reflecting surface," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2020.
- [8] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4659–4663, 2019.
- [9] T. L. Jensen and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5000–5004, 2020.
- [10] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Communications Letters*, vol. 9, no. 4, pp. 518–522, 2020.
- [11] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6607–6620, 2020.
- [12] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2020.
- [13] Z. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 210–214, 2020.
- [14] X. Rao and V. K. N. Lau, "Compressive sensing with prior support quality information and application to massive MIMO channel estimation with temporal correlation," *IEEE Transactions on Signal Processing*, vol. 63, no. 18, pp. 4914–4924, 2015.
- [15] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning," *IEEE Access*, vol. 9, pp. 44304–44321, 2021.
- [16] H. V. C. J. Chen, Y. C. Liang and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," in *Available: arXiv 1912.03619.*, pp. 1–6, 2019.
- [17] S. Liu, Z. Gao, J. Zhang, M. Di Renzo, and M. Alouini, "Deep denoising neural network assisted compressive channel estimation for mmwave intelligent reflecting surfaces," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9223–9228, 2020.
- [18] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3959–3973, 2018.

- [19] M. Wang, F. Gao, N. Shlezinger, M. F. Flanagan, and Y. C. Eldar, "A block sparsity based estimator for mmwave massive MIMO channels with beam squint," *IEEE Transactions on Signal Processing*, vol. 68, pp. 49–64, 2020.
- [20] K. Venugopal, A. Alkhateeb, N. Gonzalez Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1996–2009, 2017.
- [21] Z. Chen, Y. Fu, Y. Xiang, and R. Rong, "A novel iterative shrinkage algorithm for CS-MRI via adaptive regularization," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1443–1447, 2017.
- [22] A. Liu, V. K. N. Lau, and M. L. Honig, "Compressive RF training for massive MIMO with channel support side information," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3628–3641, 2019.
- [23] A. Liu, V. K. N. Lau, and W. Dai, "Exploiting burst-sparsity in massive MIMO with partial channel support information," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7820–7830, 2016.
- [24] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Processing Letters*, vol. 27, pp. 905–909, 2020.
- [25] P. J. Bickel, Y. Ritov, and A. B. Tsybakov, "Simultaneous analysis of lasso and dantzig selector," *Annals of Statistics*, vol. 37, no. 4, pp. 1705–1732, 2009.
- [26] E. Candes and T. Tao, "The dantzig selector: Statistical estimation when p is much larger than n ," *Annals of Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.
- [27] J. Wright and Y. Ma, *High-Dimensional Data Analysis with Low-Dimensional Models: Principles, Computation, and Applications*. Cambridge University Press, 2021.
- [28] I. Selesnick, "A derivation of the soft-thresholding function." [Online] Available:https://eeweb.engineering.nyu.edu/iselesni/lecture_notes/SoftThresholding.pdf.
- [29] S. Zeng, R. Jiao, C. Li, X. Li, and J. S. Alkassabeh, "A general framework of dynamic constrained multiobjective evolutionary algorithms for constrained optimization," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2678–2688, 2017.
- [30] Z. Liu and Y. Wang, "Handling constrained multiobjective optimization problems with constraints in both the decision and objective spaces," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 5, pp. 870–884, 2019.
- [31] L. Li, X. Yao, R. Stolkin, M. Gong, and S. He, "An evolutionary multiobjective approach to sparse reconstruction," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 6, pp. 827–845, 2014.
- [32] S. Das and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, 2011.
- [33] J. Lee, G. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370–2386, 2016.
- [34] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.
- [35] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6169–6183, 2015.
- [36] J. Dai, A. Liu, and H. C. So, "Non-uniform burst-sparsity learning for massive MIMO channel estimation," *IEEE Transactions on Signal Processing*, vol. 67, no. 4, pp. 1075–1087, 2019.
- [37] X. Ge, "Ultra-reliable low-latency communications in autonomous vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 5005–5016, 2019.
- [38] Y. Wang and Z. Cai, "Combining multiobjective optimization with differential evolution to solve constrained optimization problems," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pp. 117–134, 2012.

- [39] Z. Cai and Y. Wang, "A multiobjective optimization-based evolutionary algorithm for constrained optimization," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 6, pp. 658–675, 2006.
- [40] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.
- [41] Y. Wang, Y. Zhang, Z. Tian, G. Leus, and G. Zhang, "Super-resolution channel estimation for arbitrary arrays in hybrid millimeter-wave massive MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 947–960, 2019.
- [42] X. Li, S. Jin, H. A. Suraweera, J. Hou, and X. Gao, "Statistical 3-D beamforming for large-scale MIMO downlink systems over rician fading channels," *IEEE Transactions on Communications*, vol. 64, no. 4, pp. 1529–1543, 2016.
- [43] S. Jin, W. Tan, M. Matthaiou, J. Wang, and K. Wong, "Statistical eigenmode transmission for the MU-MIMO downlink in rician fading," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6650–6663, 2015.
- [44] Z. Tan, P. Yang, and A. Nehorai, "Joint sparse recovery method for compressed sensing with structured dictionary mismatches," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4997–5008, 2014.
- [45] A. M. Leroy and P. J. Rousseeuw, *Robust regression and outlier detection*. Wiley-Interscience., 2003.
- [46] J. Coon, M. Sandell, M. Beach, and J. McGeehan, "Channel and noise variance estimation and tracking algorithms for unique-word based single-carrier systems," *IEEE Transactions on Wireless Communications*, vol. 5, no. 6, pp. 1488–1496, 2006.
- [47] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.
- [48] V. Santucci, M. Bairoletti, and A. Milani, "Algebraic differential evolution algorithm for the permutation flowshop scheduling problem with total flowtime criterion," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 5, pp. 682–694, 2016.
- [49] W. Gao, G. G. Yen, and S. Liu, "A dual-population differential evolution with coevolution for constrained optimization," *IEEE Transactions on Cybernetics*, vol. 45, no. 5, pp. 1108–1121, 2015.
- [50] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 398–417, 2009.
- [51] Z. Lu, "Iterative hard thresholding methods for l_0 regularized convex cone programming," *Mathematical Programming*, vol. 147, no. 1-2, pp. 125–154, 2012.
- [52] Z. Dong and W. Zhu, "Homotopy methods based on l_0 -norm for compressed sensing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1132–1146, 2018.
- [53] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1 -regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.
- [54] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.
- [55] B. Yan, Q. Zhao, Z. Wang, and X. Zhao, "A hybrid evolutionary algorithm for multiobjective sparse reconstruction," *Signal Image and Video Processing*, pp. 993–1000, 2017.
- [56] J. Branke, K. Deb, H. Dierolf, and M. Osswald, "Finding knees in multi-objective optimization," in *Parallel Problem Solving from Nature - PPSN VIII*, (Berlin, Heidelberg), pp. 722–731, Springer Berlin Heidelberg, 2004.
- [57] T. Blumensath, "Accelerated iterative hard thresholding," *Signal Processing*, vol. 92, no. 3, pp. 752–756, 2012.
- [58] X. Li, J. Fang, H. Li, and P. Wang, "Millimeter wave channel estimation via exploiting joint sparse and low-rank structures," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1123–1133, 2018.
- [59] Y. Zhou, S. Kwong, H. Guo, X. Zhang, and Q. Zhang, "A two-phase evolutionary approach for compressive sensing reconstruction," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2651–2663, 2017.