Statistical Sparse Channel Modelling for Measured and Simulated Wireless Temporal Channels

Peng-Fei Cui, J. Andrew Zhang, Wen-Jun Lu, Y. Jay Guo, Hong-Bo Zhu

***Abstract*—Time-domain wireless channels are generally modelled by Tapped Delay Line (TDL) model and its variants. These models are not effective for channel representation and estimation when the number of multipath taps is large. Compressive sensing (CS) provides a powerful tool for sparse channel modelling and estimation. Most of research has been focusing on sparse channel estimation, while sparse channel modelling (SCM) is rarely considered for centimeter-wave channels. In this paper, we investigate statistical sparse channel modelling, using both measured and simulated channels over a frequency range of 6 to 8.5 GHz, . We first introduce the triple equilibrium principle to explore the trade-off between sparsity, modelling accuracy, and algorithm complexity in SCM, and provide a methodology for characterizing the sparsity of time-domain channels using single-measurement-vector compressive sensing algorithms. Mainly using mainly the selected wavelet dictionary and various CS reconstruction (aka recovery) algorithms, we then present comprehensive statistical sparse channel models, including channel sparsity, magnitude decaying profile, sparse coefficient distribution and atomic index distribution. Connections between the parameters of sparse and conventional TDL channel models are mathematically established. We also propose three methods for generating simulated channels from the developed sparse channel models, which validates their effectiveness.**

***Index Terms*—Channel model, sparse channel modelling, compressive sensing, wavelet dictionary, ℓ1-norm algorithm, Orthogonal Matching Pursuit (OMP)**

I. Introduction

TIME-DOMAIN wireless channel modelling, which characterizes the propagation of wireless signals, is typically based on the Tapped Delay Line (TDL) model [1]–[3] and its variants such as the cluster-based Saleh-Valenzuela (SV) model [4]. The parameters of these models are described by statistical distributions, such as small-scale Rayleigh and Rician fading, exponential power delay profile (PDP), channel coherent time and bandwidth. Such models are simple and straightforward in structure, and they would provide clear physical interpretation for wireless signal propagation. However, some of the advantages may disappear for large-bandwidth channels. For example, there could be tens and even hundreds of resolvable multipath in latest 5G channel models [5]. In such dense multipath channels, there are many parameters to be characterized for channel modelling, and to be estimated for channel estimation, which renders TDL model and its variants ineffective.

A natural question to ask is then, whether we can develop a channel model to represent such dense multipath channels in a more sparse and compact way. Compressive sensing (CS) [6] provides an effective tool for both sparse channel modelling (SCM) and signal processing based on sparse channel models. However, it is surprising that most work has been focusing on the latter, while SCM, which should have been the basis for sparse signal processing, is somewhat neglected. In the past decade, there have been strong interest in applying CS techniques for sparse channel estimation [7]–[9] and for sparse channel coding [10]–[12]. These works assume that wireless channels are sparse and compressible, and hence less training signals at the transmitter and observations at the receiver can be used to get a complete estimation for the sparse channels. Such sparsity assumption, however, is not very well validated by practical SCM results. The channel sparsity for millimeter-wave channels has been validated by several channel models established from practically measured data [13], [14], particularly in the angular domain. However, for lower-frequency centimeter-wave channels, there are very limited reported results, validating the sparsity assumption – not even to mention detailed statistical analysis for the sparse channel parameters. One of the limited number of examples is [15], where a sparsity pattern expressed by virtual channel representation with the Fourier dictionary is proposed to model the double-selective fading multi-path channels, but the channel statics is not well developed.

The negligence on SCM is probably linked to its unacknowledged importance. SCM, through disclosing the inherent channel sparse structure, can not only provide a potentially simpler method for channel simulation, but also benefit to sparse channel estimation and coding directly. For example, SCM can demonstrate the sparsity and best dictionaries, and hence provide guidance to the training symbol design for channel estimation [9]; the statistics of parameters in sparse channel models can also be exploited for developing better channel estimation algorithms, as has been explored in [16]. SCM can also estimate the sparsity, which is a necessary knowledge required by many sparse reconstruction algorithms such as CoSaMP.

This paper devotes to SCM for centimeter-wave channels, using both practically measured and simulated single-input single-output dense multipath channels in the frequency-band

Peng-Fei Cui is with Nanjing University of Posts and Telecommunications, P. R. China, and he is also with University of Technology Sydney, Australia. Email: pengfei.cui@student.uts.edu.au.

J. Andrew Zhang and Y. Jay Guo are with the Global Big Data Technologies Centre, University of Technology Sydney, Australia. Email: {Andrew.Zhang; Jay.Guo}@uts.edu.au.

Wen-Jun Lu and Hong-Bo Zhu are with the Jiangsu Key Laboratory of Wireless Communications, Nanjing University of Posts and Telecommunications, Nanjing, P. R. China. Email: {wjlu;zhb}@njupt.edu.cn.

This work is supported by National Nature Science Foundation Council (NSFC) Proposal under granted no. 61427801 and 61871233.

The corresponding authors: Hong-bo Zhu and Andrew Zhang.

from 6 to 8.5 GHz. Such channels are known as one type of *ultra wideband* channels. With a 2 GHz bandwidth, the channels consist of a large number of resolvable multipath taps in the time domain, which can be represented directly by the TDL model. The original channel impulse response (CIR) can be downsampled with a low-pass filter to generate channels for systems with lower bandwidth if needed. In this paper, we work on the high-resolution dense multipath directly, and will show that there are great sparsity if we represent the CIRs using different dictionaries (aka, bases).

One main challenge for SCM is that, a sparse model can be affected by quite a few factors, such as the used dictionary, the reconstruction algorithm, and required accuracy of modelling. Actually, the TDL model can also be treated as a special sparse model when insignificant multipath taps are ignored according to a desired accuracy. The dictionary will be a (partial) identity matrix in this case if represented in the time domain, or a partial Discrete Fourier Tranform (DFT) matrix if represented in the frequency domain. But it is not an efficient sparse representation when the channel is rich in multipath. The sparsity is also largely affected by the used reconstruction algorithm with various complexity and the desired accuracy. Hence a good trade-off between sparsity, accuracy and complexity needs to be achieved. This trade-off will be investigated in detail in Section III.

In this paper, we aim to provide a methodology for sparse modelling of time-domain channels, and to establish statistical sparse channel models. Our main contributions are as follows.

- We introduce a *triple equilibrium principle* to characterize the trade-off between sparsity, modelling accuracy, and reconstruction complexity in SCM. Using both measured and simulated channel datasets, we quantitatively demonstrate and practise the triple equilibrium principle in SCM. This principle provides an important guidance to selecting dictionary and reconstruction algorithm in SCM;
- We develop comprehensive statistical sparse channel models, including channel sparsity, dependency of sparsity on channels, magnitude decaying profile, sparse coefficient distribution, and sparse atomic index distribution, which are analogous to those in the conventional TDL model. Novel methods using ordered sparse coefficients and atomic index separation are proposed for statistical modelling;
- We mathematically establish the connection between parameters in the sparse channel model and conventional TDL model;
- We propose three methods for generating simulated channels using the sparse channel models. The effectiveness of the established sparse channel models are validated by comparing these generated channels with those actual ones.

The rest of this paper is organized as follows. In Section II, we describe the channel datasets and our sparse modelling methodology. In Section III, we introduce the triple equilibrium principle and discuss the selection of the dictionary and reconstruction algorithms used for SCM in this paper. In Section IV, we present detailed SCM results, and investigate the connection between parameters for sparse and conventional channel models. In Section V, methods for generating simulated channels from the established sparse models are presented, and the generated channels are compared to the actual ones to verify the effectiveness of the sparse models. Section VI concludes the paper.

II. Channel Datasets and Sparse Modelling Methodology

In this section, we introduce the channel datasets used in this work and present our sparse modelling methodology.

A. Tested Channel Datasets

We conduct the channel sparse modelling using two datasets, i.e., the practically measured off-body CIRs [17], [18] and simulated wireless CIRs following the channel model proposed by IEEE 802.15.3a working group in [3].

1) Measured dataset: The channel data is measured in a typical hospital-type room with a few furnitures. An omni-directional monopole antenna is used to emulate the external access point (AP), and a wearable loop-dipole antenna worn on different positions of the volunteer is used to emulate the receiver (Rx). A vector network analyzer (cf. VNA Agilent 8720) is used to generate a 0 dBm, 801-point sweeping signal with the frequency ranging from 6 to 8.5 GHz. The off-body CIRs are observed in different body-worn parts and large-scale measurement locations. More details on the channel dataset and measurement can be found in [17], [18].

According to the factors affecting the large-scale fading, the measured off-body channels are classified into three classes of datasets: body-parts-dependent (denoted as BAN_Parts), distance-dependent (BAN_Dis) and height-dependent (BAN_Hei) CIRs. Each CIR is denoised and normalized to unity energy. In one class of datasets, there may be measurements under different scenarios, for example, for BAN_Parts, there are measurements obtained from different body parts, and for BAN_Hei, measurements are from the AP placed at different heights.

2) Simulated CIRs: The simulated CIRs are generated from both the cluster-based model and TDL model. The cluster-based model follows the SV model in [4], and the TDL model is given by

$$h(t) = \sigma_0^2 \sum_{l=0}^{L-1} \alpha_l e^{-lT_s/\tau_{\text{RMS}}} \delta(t - \tau_l), \tag{1}$$

where L is the total number of propagation taps, τ_l is the delay of l-th tap, τ_{RMS} is the Root Mean Square (RMS) delay spread and T_s is the sampling time, α_l is the fading coefficient of the l-th tap, and $\sigma_0^2 e^{-lT_s/\tau_{RMS}}$ represents an exponential power delay profile (PDP) of the taps, with $\sigma_0^2 = \frac{1-e^{-T_s/\tau_{RMS}}}{1-e^{-(L+1)T_s/\tau_{RMS}}}$ ensuring the average CIR energy is unity.

The simulated CIRs include four typical Case Models (CM), i.e., CM1 for 0-4 meters and Line of Sight (LOS), CM2 for 0-4 meters and none Line of Sight (NLOS), CM3 for 4-10 meters and NLOS, and CM4 for extremely bad NLOS with RMS

delay spreads of 5.28, 8.03, 14.28 and 25 nanoseconds [3]. We consider various fading such as Rayleigh, Rician and Nakagami fading for the fading coefficients α.

B. Methodology for Sparse Channel Modelling

For a given sparse channel coefficient vector $\mathbf{x}_0$ of length N, *channel sparsity* here is referred to as either the absolute number of non-zero elements in the vector, given by $K = \|\mathbf{x}\|_0$, or the ratio given by K/N. We will use these two optional definitions interchangeably hereafter. If K or K/N is small (or large), we say the signal is sparser (or less sparser) and the sparsity is small/low (or large/high).

For any given samples of single-channel CIR $\mathbf{y} \in \mathbb{R}^M$, we can formulate the CS representation of the channel as

$$\mathbf{x}_s = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1, \quad \text{subject to } \mathbf{y} = \mathbf{\Phi\Psi x} + \mathbf{r}. \qquad (2)$$

where $\mathbf{x}$ is the sparse coefficient vector, $\|\mathbf{x}\|_1$ is norm-1 of $\mathbf{x}$, i.e., the sum of absolute values of the elements in $\mathbf{x}$, $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ $(M \leq N)$ is the measurement/sensing matrix, $\mathbf{\Psi} \in \mathbb{R}^{N\times N}$ is the dictionary, and $\mathbf{r}$ is the residual vector representing noise and/or residual errors in the sparse approximation. There are a lot of CS algorithms such as Orthogonal Matching Pursuit (OMP), L1-minimization and Bayesian compressive sensing, that can be applied to solve the optimization problem in (2). Different algorithms have different complexity and different reconstruction accuracy.

We have full datasets for the originally simulated and measured channels, sampled at the Nyquist rate. We can design the sensing matrix $\mathbf{\Phi}$ with M much smaller than N, to mimic what is in the practical channel estimation situation. However, this may make the SCM results lose generality. So instead, we use all the channel samples in $\mathbf{y}$ directly for SCM, without using $\mathbf{\Phi}$.

For a selected dictionary, for example the wavelet dictionary, we call each column in the dictionary matrix $\mathbf{\Psi}$ an atom, and its index as atomic index. The n-th sparse coefficient corresponds to the n-th atom. For the TDL model, as mentioned earlier, it can be treated as a special sparse model with the dictionary being an identity matrix and each atom being a delta function.

Our SCM methodology consists of three major stages, which are summarized next and will be detailed in following sections.

- The first stage is *selection of dictionary and reconstruction algorithms*. From the formulation in (2), we can see that the sparsity k is closely related to three factors, the product $\mathbf{\Phi\Psi}$ (or just the dictionary $\mathbf{\Psi}$ in this paper), the reconstruction algorithm, and the constraint of residual vector $\mathbf{r}$ (i.e., the desired the accuracy). Understanding the trade-off between these factors is an important step in SCM. In Section III, we introduce the triple equilibrium principle to elaborate the trade-off and discuss how to select appropriate dictionaries and CS reconstruction algorithms.
- The second stage is *sparse channel modelling* as will be detailed in Section IV. In this stage, analogous to the parameters in the TDL model including significant MultiPath Components (MPC), PDP, small-scale fading, and delay, we analyze the statistics for sparsity, magnitude decaying profile (MDP), distribution of sparse coefficients, and distribution of atomic indexes. Thus a complete statistical sparse channel model is established.
- The third stage is *sparse channel validation*, as will be detailed in Section V. During this stage, we stochastically generate the sparse coefficients and atomic indexes using the developed statistical sparse channel models, and compare the generated channels with the original measured and simulated channels from which the statistical sparse models are developed. This process is used to validate the effectiveness of the developed models.

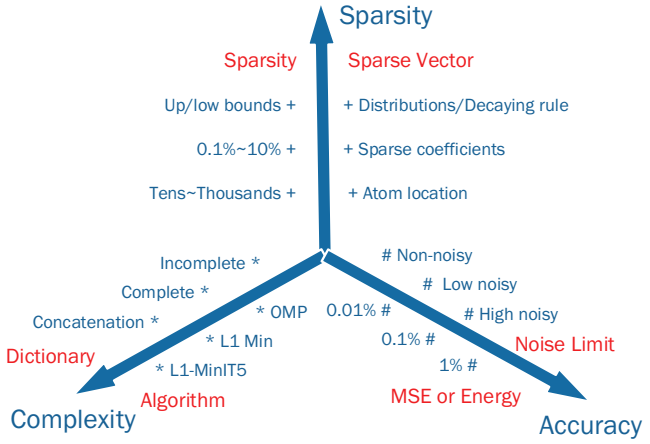

Fig. 1: The triple equilibrium relationship among sparsity, complexity and accuracy.

III. Triple Equilibrium Principle and Selection of Dictionaries and Reconstruction Algorithms

In this section, we first introduce the *triple equilibrium principle* for characterizing the trade-off between sparsity, modelling accuracy and complexity in SCM. We then discuss how we select the dictionary and reconstruction algorithms to achieve a trade-off under the guidance of the principle, by investigating the actual impact of the trade-off on SCM.

A. Triple Equilibrium among Sparsity, Complexity and Accuracy in SCM

In the sparse channel modelling, we will consider the interaction and tradeoff between channel sparsity, reconstruction complexity and accuracy. Reconstruction complexity and accuracy are associated with, e.g., the reconstruction algorithms, times of iterations used in the algorithms, the adopted dictionary and its size, and the number of measurements. According to many research results on compressive sensing including the classical work by Donoho, Candes et al. [6], sparsity, reconstruction complexity and accuracy interact with each other. Generally, the sparsity of practical signals can change with the required reconstruction accuracy and the affordable complexity, and the changes can be significantly different for different signal sources.

For sparse channel modelling, we hence keep in mind the triple equilibrium relationship among sparsity, complexity, and accuracy as illustrated in Fig. 1. Firstly, the achievable channel sparsity is expected to grow from smaller to larger when the required reconstruction accuracy increases and/or the required complexity decreases. Secondly, to achieve a given sparsity level, e.g., 5% sparsity ratio, the sparser the actual signal is, the lower the reconstruction complexity will be to attain certain reconstruction accuracy. Thirdly, there is a complexity tradeoff between the selected recovering algorithm and dictionary. Fig. 1 lists uncomplete, critical complete, concatenation dictionaries and recovering algorithms such as greedy algorithm (like OMP), optimization algorithm (ℓ1-Minimization denoted as ℓ1-Min.) and iterative algorithm (5 step reweighted ℓ1-Minimization denoted as ℓ1-MinIT5) from low to high complexity. It is observed in [19] that with 3 to 5 concatenation dictionaries, even the simplest OMP algorithm can achieve better performance than the optimal algorithm such as iterative ℓ1-MinIT5. Finally, it is noted that the denosing ability of CS algorithms can also have a great impact on the reconstruction accuracy. Measurement noise can be the upper limit of sparse processing ability.

The triple-equilibrium relationship in sparse channel modelling will be explored in detail in Section IV-A and III-C.

B. Impact of Dictionary on Sparse Channel Modelling

Dictionary can have a significant impact on the sparsity and reconstruction accuracy of a general CS problem, as well as our sparse channel modelling. For example, in our experiments, it is widely observed that to achieve the similar modelling accuracy, the sparsity for recovering the same BAN_Part CIR using the ℓ1-Min algorithm are 17, 18, 72, 94 and 210 for the symlet 3-5 wavelet, exponential, truncated cosine wave, Fourier and random Gaussian dictionaries, respectively. The wavelet and exponential dictionaries exhibit much lower sparsity than the others. This is because their fast-fading atoms adapt to the sparse properties of significant clusters, i.e., the properties of compact support and locality have advantages in representing the cluster characteristics. The observation here is consistent with the quantitative dictionary selection formula and the results shown in [20]. So we will mainly present the sparse channel modelling results using wavelet and exponential dictionaries in this paper.

Fig. 2 and 3 illustrate the original, recovered and residuals of one BAN_Part CIR using different dictionaries and CS algorithms. The combination of wavelet-OMP and exponent-ℓ1-Min both perform well, achieving similar sparsity levels around 20 under the same MSE of 10^{-4}. It is noticed that the exponential dictionary achieves less accurate results in the segment of the primary cluster. This is mainly because all atoms of the exponential dictionary have the same waveform, i.e., $d_k(t) = \sqrt{1-\beta^2}e^{-\beta t}, t \geq k$, while the wavelet dictionaries have multiple different waveforms.

Fig. 4 depicts the compact support length and waveforms for six atoms of the symlet 3-5 wavelet dictionary. Waveform 1 is the scale (father) wavelet of Symlet 3, and waveforms 2 to 6 are the different scales of level 5 to 1 of the mother

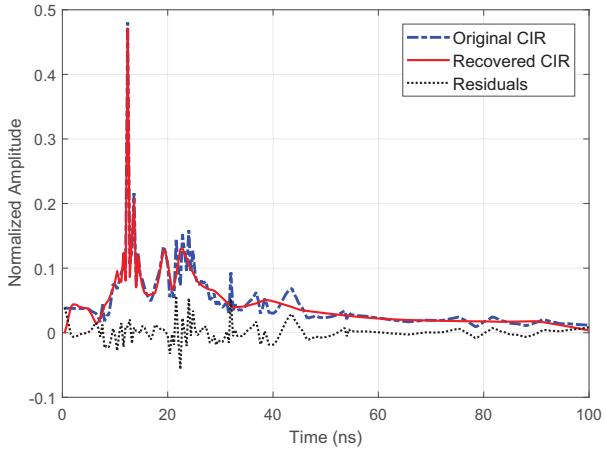

Fig. 2: Comparison of original, recovered and residual signals using wavelet dictionary and OMP algorithm. Sparsity $k = 20$.

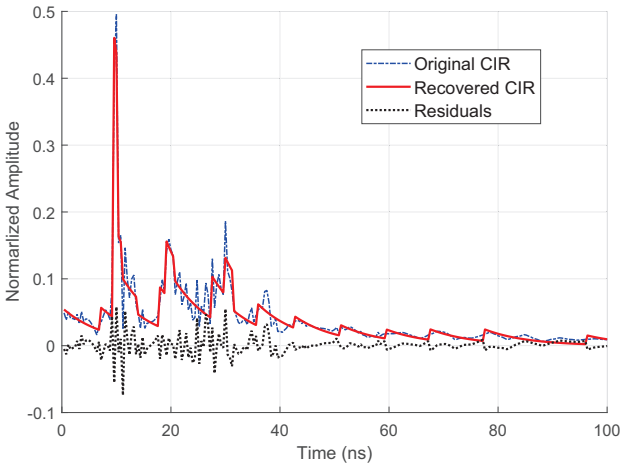

Fig. 3: Comparison of original, recovered and residual signals using exponential dictionary and L1-Minimization algorithm. Sparsity $K = 18$.

wavelet. These waveforms also have good quasi-orthogonal characteristics. Such versatile waveform combinations make wavelet dictionary adapt to different signal types including the clustered channels.

Of course, single-shape exponential dictionary also has its unique advantages. For example, if using three different β-shaped concatenation exponential dictionaries to form multi-shape over-complete dictionary, the sparsity can even be reduced to 10 taps [21]. However, a combination of different wavelet dictionaries does not seem to work well. Thus, it is important to select/train adaptive projection atoms to form dictionaries for specific signals.

C. Sparsity-Complexity Relationship Under Fixed Accuracy

The achievable channel sparsity can be affected by different reconstruction algorithms for the same residual error, according to the sparsity-accuracy analysis in Section IV-A. Different algorithms, or the same algorithm with different iterations, can

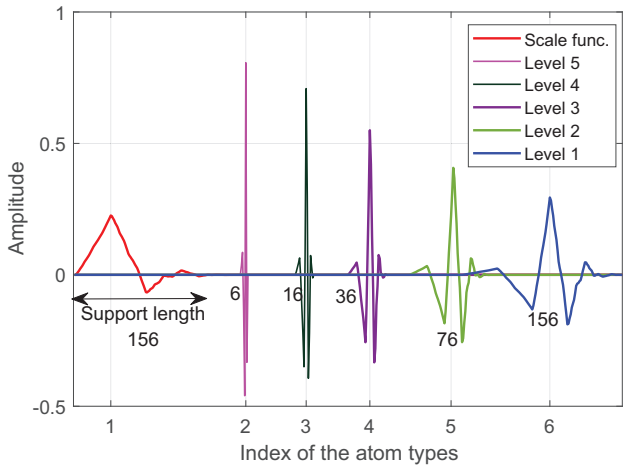

Fig. 4: Waveforms for six types of atoms in the symlet 3 level 5 wavelet dictionary. For the index of atom types, the first (with index 1) is the father wavelet, the last (with index 6) is the mother wavelet and the child wavelets ranging from level 2 to 5 are indexed as type 5 to 2. The number beside each waveform denotes the support length of the wavelet.

have very different complexity. Here, we quantitatively analyze the complexity associated with sparse channel modelling and reconstruction, using the ℓ1-Min and its iterative algorithms as examples. The required multiplications for ℓ1-Min and ℓ1-Min ITq are $O(qMN^2), O((T+1)qMN^2)$, where T is the number of iteration or reweight times, and $T = 0$ for ℓ1-Min [22].

Fig. 5 depicts sparsity variations for many BAN_Parts CIRs, using ℓ1-Min reconstruction algorithms with iterations ranging from 0 to 10. For all datasets the mean sparsity decreases as the iteration number increases, and the first iteration always achieves the largest sparsity reduction. The average sparsity value for a particular dataset also quickly converges, and further increasing the number of iterations only leads to negligible reduction on sparsity. For all datasets, the number of iterations up to 5 is shown to be sufficient. We can also see that all CIRs under good LOS conditions have lower sparsity than those under NLOS cases. Similarly, experimental results for simulated CM1-4 channels lead to an average sparsity of 7, 9, 14 and 25 taps, respectively, which is in line with the trend of the increasing RMS delay spread (5, 8, 14, 25 nanoseconds). The connection between the average sparsity and RMS delay will be mathematically shown in IV-D.

D. Summary of the Triple Equilibrium Principle

It is found that the residual error in sparse modelling exponentially decreases with the desired sparsity increasing, and the desired sparsity can be reduced up to certain bounds at the cost of linearly increasing complexity. The triple equilibrium among sparsity, complexity, and accuracy is thus qualitatively verified. It is nearly impossible to improve any one without impacting the other two.

In next section, we use the selected wavelet and exponential dictionaries and the algorithm sets (OMP, ℓ1-Min and ℓ1-Min ITq) to conduct SCM for both measured and simulated CIRs.

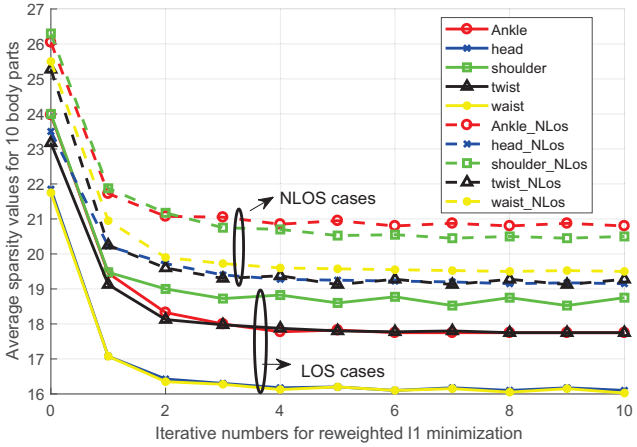

Fig. 5: Sparsity variation for ten body parts CIRs reconstructed using the symlet 3-5 dictionary and reweighted ℓ1 minimization algorithms with 0 to 10 iterations.

IV. Sparse Channel Modelling

In this section, we conduct detailed sparse channel modelling, focusing on sparsity, "magnitude decaying profile" (MDP), "fading" and atomic index distribution, which have analogies in conventional TDL modelling. We will also investigate the dependency of sparse coefficients on channel parameters.

Ideally, we would like to investigate the MDP and fading for each sparse coefficient corresponding to each atom (column) in the dictionary. This will give us the results directly analogous to the TDL modelling. However, this is hard to do for the following reason: the size of the dictionary could be very large and there are frequently insufficient samples for a number of coefficients due to limited measured datasets. The fluctuation of atomic indexes can largely reduce the reliability of obtained sparse statistics. Some dictionaries do not have clear physical meanings, which also makes direct processing not necessary. On the other hand, organizing the non-zero taps in a proper order is found to be able to better reveal the statistics of sparse signals [23]. Therefore in this paper, we organize the obtained sparse channel coefficients in a descending order according to their magnitudes, and conduct MDP and fading analysis based on the ordered coefficients. At the same time, we collect the atomic indexes for these coefficients, and propose an *atom-index-division method*, which will be detailed in Section IV-C, to get the statistical distribution for the indexes of these ordered coefficients.

A. Sparsity

We characterize the channel sparsity and investigate its dependency on accuracy and types of channels. We will demonstrate in Section IV-A1 that the sparsity and reconstruction accuracy are bonded, and their relationship can be well characterized by an exponential function. We will show in Section IV-A2 that the sparsity also varies from channel to channel as expected, and is dependent of statistical fading distributions.

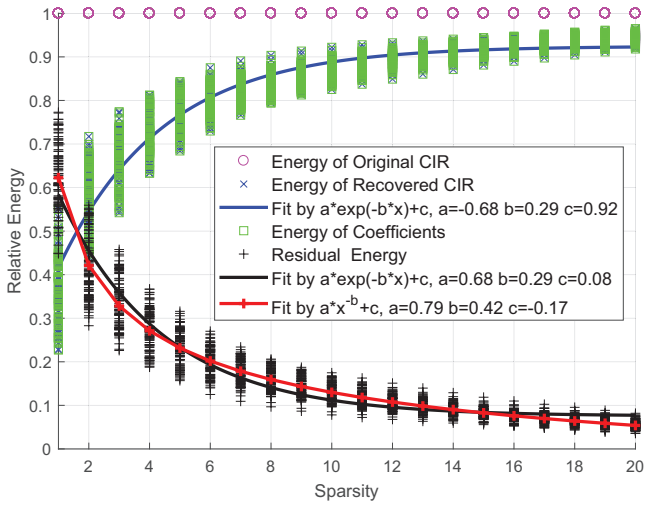

Fig. 6: Relative energy of reconstructed and residual signals for measured BAN_Dis CIR subset using wavelet dictionary and OMP algorithm.

TABLE I: Parameters of exponential fitting function for residual energy for different dictionaries and algorithms.

Cases	Dictionary	Algorithm	a	b	c
BAN_Parts	Wavelet	ℓ1-Min	0.51	0.11	0.12
BAN_Hei	Wavelet	ℓ1-Min	0.58	0.14	0.15
BAN_Dis	Wavelet	ℓ1-Min	0.49	0.12	0.12
	Wavelet	OMP	0.68	0.29	0.08
	Exponent	OMP	0.43	0.25	0.08
Sim_Rayl	Wavelet	ℓ1-Min	0.39	0.20	0.05
	Wavelet	OMP	0.76	0.45	0.07
	Exponent	OMP	0.35	0.50	0.06
Sim_Rice	Exponent	OMP	0.27	0.65	0.08
Sim_Naka	Exponent	OMP	0.28	0.61	0.06

1) Dependency of Sparsity on Reconstruction Accuracy: The required sparsity in channel modelling has a direct impact on the accuracy, and vice verse. In [21], it is theoretically proved that the residual energy for any sparse approximation is linearly bounded by the residual of the best sparse approximation. Such a boundary constraint, however, is often very loose for practical sparse reconstruction algorithms, and is also dependent on the actual dataset. In our work, to understand this relationship and its impact on sparse channel modelling, we conducted various signal recovery experiments using different sparsity constraints, different datasets, dictionaries and recovery algorithms. Next, we present some exemplified results.

Fig. 6 shows the relative energy of reconstructed signal and the residual (error) between the reconstructed and original signals, normalized to the original signal power. The dataset includes 400 measured BAN_Dis CIRs. The sparse signal is reconstructed using the symlet 3 level 5 wavelet dictionary and the OMP algorithm. Due to the unit energy constraint and the Restrict Isometric Property (RIP) property of wavelet dictionary, the energy of reconstructed signal, represented by the sparse coefficients, increases with the sparsity increasing steadily, while the residual error steadily declines.

Plotted together in the figure are the exponential and polynomial curve fitting functions. Using the Akaike Information Criteria (AIC) [24], it is found that the relative signal and residual energy can both be well fitted by the exponential function given by

$$\delta_{Res} = ae^{-bk} + c, \tag{3}$$

where δ_{Res} is the residual energy and k is the desired sparsity. The polynomial (or power) function $\delta_{Res} = ak^{-b}+c$, which is widely used to bound residuals [21], is also found to provide good fitness here.

To summarize, it is found that *the exponential function can fit well the residual energy for almost all measured and simulated datasets*, for different channel fading distributions, dictionaries or algorithms. Table I presents the parameters of the exponential fitting function for the residual energy for different dictionaries and algorithms. In our experiments, we observed that for simulated channels, the fitting accuracy is low when the sparsity is large, when larger decaying exponents (parameters b in (3)) are obtained, especially for Rician or Nakagami fading channels. This indicates that simulated channels with simpler multipath structures such as a dominating LOS component usually lead to smaller residual-sparsity product and hence better recovery performance.

According to the MSE metric $E_{\text{MSE}} = E_{Res}/M$, $E_{\text{MSE}} = 10^{-4}$ is equivalent to a residual energy of 5% with the number of measurements $M = 500$ in most cases. Fig. 6 indicates that the 5% residual error corresponds to an average sparsity between 18 and 20, or 4% sparsity ratio. We use 4% sparsity ratio ($E_{\text{MSE}} = 10^{-4}$) as a baseline for distribution analysis in this paper. The CIR modelling for all pairs of residual error and sparsity can be implemented as long as they are above the asymptotic line indicated by (3).

2) Dependency of Sparsity on Channels: Fig. 7 presents the Cumulative Probability Distributions (CDFs) of reaching a desired sparsity for different algorithms. Results for using all the measured data sets and for using specific data sets (waist only and ankle only) are presented for comparison. Each curve is obtained from over 4000 sparse reconstruction experiments. When all the measured datasets are used, the curves for algorithms ℓ1-Min, ℓ1-Min IT1 and ℓ1-Min IT5 using the same wavelet 3-5 dictionary can be well fitted by Gaussian functions with decreasing mean and variance values. We can also see that using recovery algorithms with more iterations can effectively reduce the mean sparsity and reduce the fluctuation of sparsity. Generally the smaller the sparsity fluctuation is, the smaller the modeling error will be. Thus for a selected dataset, there is a tradeoff between the complexity of the recovery algorithm and the accuracy of the modeling. For waist only and ankle only dataset, the CDF follows distinct distributions due to reduced variation of the sparsity. For example, the CDF for waist and ankle datasets follows Weibull and Log-Normal distributions respectively. Compared to the full datasets results, the sparsity is reduced as can be seen from the CDF curves.

Fig. 8 compares the CDFs for simulated Rayleigh, Rice and Nakagami fading channels using the same wavelet dictionary and OMP algorithm. The sparsity of Rayleigh fading channels has a maximum mean value of 17 taps and the largest sparsity variation. The mean sparsity for Rician fading CIRs gradually

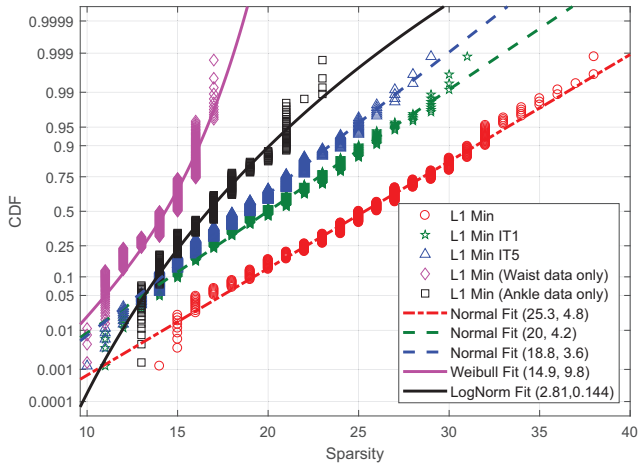

Fig. 7: CDFs of the recovered sparsity for different reconstruction algorithms and measured channel data using Symlet 3-5 dictionary. For the first three curves (ordered according to the legends), all measured datasets are used; while for the rest two, only specific datasets for waist and ankle are used respectively.

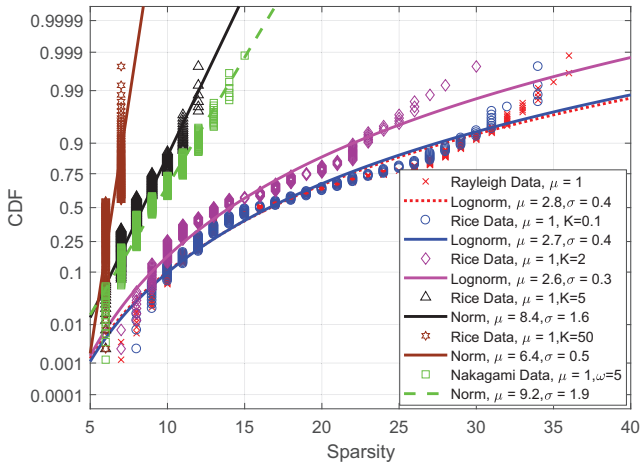

Fig. 8: CDF of the sparsity for simulated fading channels with different parameters.

decreases from 17 taps (nearly the same with Rayleigh fading channel) to only 7 taps, with the Rician K-factor changing from 0.1 to 50. Under the same LOS condition, Nakagami fading channel is less sparser than similar Rician channels (curves marked in black triangles). It seems that the channel condition has a notable effect on sparsity distributions, and channels with larger LOS components demonstrates lower sparsity.

As a summary for the sparsity of studied channels, it is found that *an average sparsity of 25 for the measured CIRs and 20 for simulated cases are the least values for most cases* with different dictionaries, algorithms and channel conditions. The maximum sparsity is 38, which correspond to the sparsity ratio of 7.6%, for a total of $M = 500$ samples. This reflects the great advantage of compressed channel sampling over traditional Nyquist sampling method.

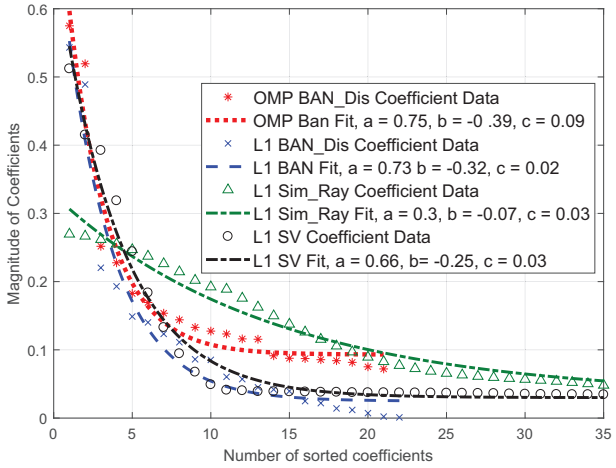

Fig. 9: Exponential fitting for sorted sparse coefficients in different datasets and algorithms using symlet 3-5 wavelet dictionary.

B. Statistics of Sorted Sparse Coefficients: Magnitude Decaying Profile and Coefficient Distribution

We now characterize the statistics of sorted sparse channel coefficients, including the *magnitude decaying profile (MDP)* and *coefficient distribution ("fading")*. There are good analogies between them and the statistics of conventional TDL model. The MDP and coefficient distribution correspond to the power delay profile (PDP) and small-scale fading in TDL models, respectively. The main difference is that the statistics of sparse coefficients could be dependent on the dictionaries, and is also slightly related to the reconstruction algorithms.

The MDP of sorted sparse channel coefficients can be well described by the following exponential function of the index of these coefficients:

$$C_i = a\, e^{b\, i} + c \tag{4}$$

where $i \in [1, K]$ is the index of the i-th largest sparse coefficient, and a, b and c are parameters that can be obtained via curve fitting. Note that (4) is different to (3) in the physical meaning of these parameters. As have been shown in Fig. 6 in Section IV-A, the energy of the coefficients for different atoms (i.e., columns in a dictionary) are very unbalanced. The maximal none-zero coefficient (generally the first output in the OMP algorithm) contains around 40% of the total energy, but the 10-th only occupies 0.2%. This is mainly due to the exponential decaying laws for sorted sparse coefficients as discussed in Section IV-A. In Fig. 9, we show another example for sorted sparse coefficients for four typical channel datasets, using different reconstruction algorithms. The figure shows that, for all measured and simulated channels, their MDPs can be well characterized by exponential functions. For the two measured channels, different reconstruction algorithms have insignificant impact on the decaying speed.

Since the coefficients in the exponential function are found to be very similar for different datasets in the same class,

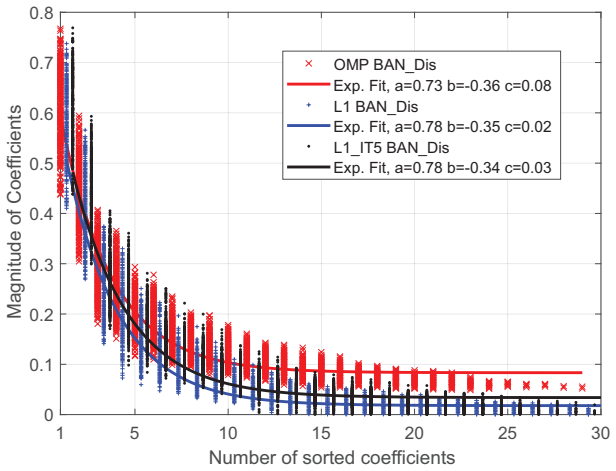

Fig. 10: MDP of aggregated datasets and the fitting by exponential functions for three reconstruction algorithms with the symlet 3-5 wavelet dictionary.

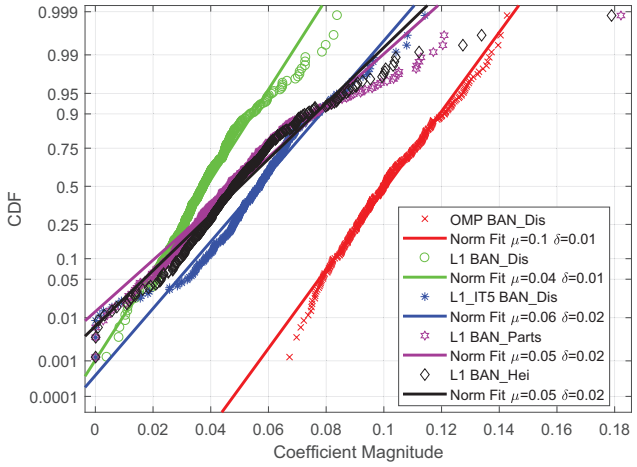

Fig. 11: CDF of the 10-th ordered sparse channel coefficient for different classes of datasets and algorithms using symlet 3-5 wavelet dictionary. "Norm Fit" is short for fitting with Normal (Gaussian) distribution.

we aggregate the sparse coefficients in the same class for analysis, to make the model more general. Fig. 10 shows the MDP of the sorted sparse coefficients for the aggregated BAN_Dis datasets. They are shown to be well fitted by three exponential functions with similar parameters. More advanced algorithms, such as L1_IT5 generally achieves better fitness with smaller tail.

The sparse coefficients with the same index for the same class of datasets are also found to have similar CDFs. Hence we also aggregate all the datasets in each class and analyze their CDFs. Fig. 11 illustrates the CDF for the 10-th largest sparse coefficient in different classes of datasets. The figure shows that for all datasets and all reconstruction algorithms, the CDF can be well approximated by that of Normal distribution. Similar matching distribution has been observed for other sparse coefficients. This indicates that *the magnitude of sparse channel coefficients experience "Gaussian fading" - the variation of magnitude can be characterized by a Gaussian function*.

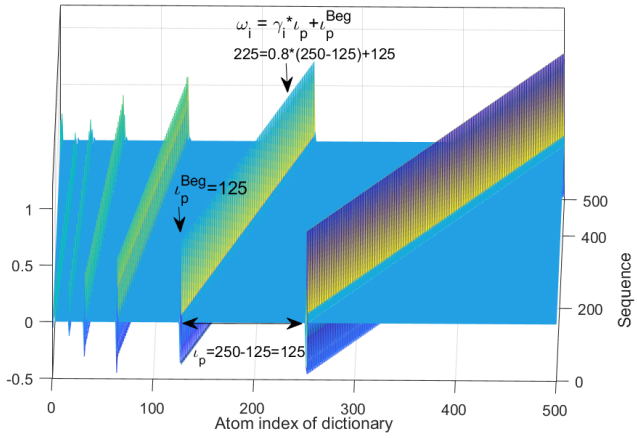

Fig. 12: An example for the atom-index-division method using the symlet level 5 wavelet dictionary. The selected atomic index is 225 in the whole dictionary with the shape factor $p = 3$ and the location factor $\gamma_i = 0.8$.

C. Statistics of Atomic Indexes for Sparse Channel Coefficients

Corresponding to the ordered sparse channel coefficients, we study the statistical properties of their atomic indexes. Due to the large size of the sparse vector, there could only be a small number of samples for each index associated with the ordered channel coefficients. For example, for a channel of length 500, the index set for the 3rd sorted sparse coefficients is found to be concentrated on atom 1 and 6, i.e., father wavelet and level-1 mother wavelet in (4). Therefore directly looking into the statistics for each index will lead to inaccurate results.

Instead, inspired by the wavelet structure, we propose an atom-index-division method which splits any atomic index into a *shape factor* and a *location ratio factor*. This is represented by

$$\omega_i = \gamma_i \, \iota_p + \iota_p^{\text{Beg}} \tag{5}$$

where ω_i is the original atomic index corresponding to the i-th sorted non-zero sparse coefficient, p is the shape factor, γ_i is the corresponding location ratio factor, and ι_p and ι_p^{Beg} are the length and the beginning number of Shape p atoms. It is found that different dictionaries in the same wavelet class only causes slight changes to the specific values of ι_p and ι_p^{Beg}, without significantly affecting the statistics of the two factors. Similarly, concatenating another sub-dictionary to the existing dictionary (such as wavelet 3-5) only adds 1 to the shape factor, without dramatically changing the ranging and mean value of the two factors.

To better illustrate this concept, in Fig. 12 we provide an example for the atom-index-division method using the symlet level 5 wavelet dictionary. The selected atomic index is 225 in the whole dictionary, and it belongs to the Shape 2 atom types and locates at the 0.8 ratio in all Shape 2 atoms. Thus,

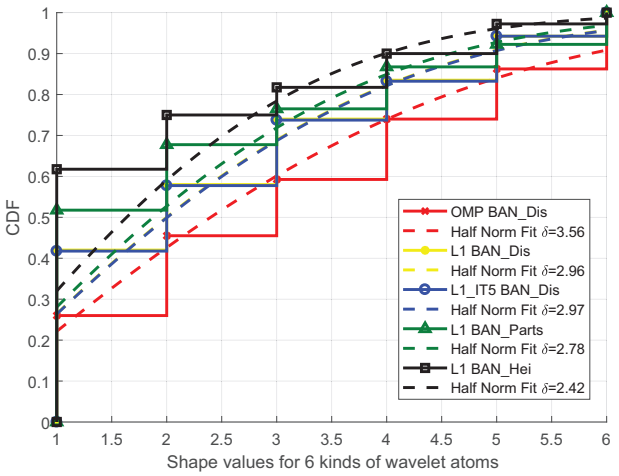

Fig. 13: Distributions of the shape-factor values of the 10-th sorted sparse coefficient for different datasets and algorithms using the symlet 3-5 wavelet dictionary.

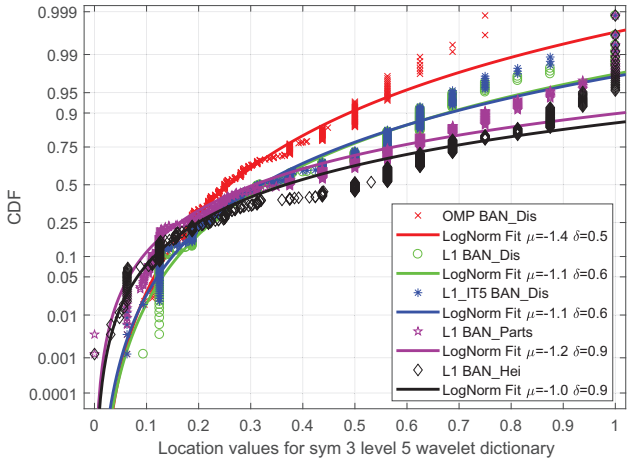

Fig. 14: Distributions of the location-ratio-factor values of the 10-th sorted sparse coefficient for different datasets and algorithms using the symlet 3-5 wavelet dictionary.

the atomic index of 225 is divided to the shape factor $p = 3$ and the location factor $\gamma_i = 0.8$.

Compared to directly working on the statistics of non-zero atomic indexes, the proposed atom-index-division method has three advantages: 1) It needs fewer samples but provides better distribution fitness; 2) It can provide more stable statistical results for scalable dictionaries; and 3) It provides deeper insights into propagation characteristics (e.g., the location ratio factor can be transformed to tap delay of a specific atom).

Figs. 13 and 14 depict the statistics of shape factor and location ratio factor for wavelet 3-5 dictionary under different scenarios, respectively. The shape values for different scenarios and recovery algorithms all follow half-Normal distributions. Compared to the OMP algorithm, the ℓ1-Min and its iterative methods tend to have a higher probability of selecting shape-1 wavelet atom (i.e., father wavelet) and lower probability of selecting high-level wavelets. No significant distinctions are observed among different scenarios. Similarly, the location ratio factors for different algorithms and different datasets all follow the log-Normal distribution very well. Compared to the OMP algorithm; the iterative algorithms have a higher probability of generating high ratio values. No significant statistical difference is observed from the location-ratio-factor values for different non-zero coefficients.

D. Relationship between Sparse Coefficients and Propagation Parameters

There are some close correlation between the significant sparse coefficients and some major propagation parameters, such as the significant MPCs, mean delay and RMS delay in the TDL model. The significant MPCs are defined as the number of multipath in the TDL model with power larger than a set threshold. The correlation between the channel sparsity and the mean delay and RMS delay can be mathematically established below.

According to the TDL and sparse channel models, we can relate the mean delay to the sparse parameters via

$$\begin{aligned}\overline{\tau} = \frac{\sum \tau_i P(\tau_i)}{P(\tau_i)} &= \frac{r_0 \mathbf{Tr}(\mathbf{D_M}(\mathbf{Ax})(\mathbf{Ax})')}{1-\delta_{Res}} \\ &= \frac{r_0 \mathbf{Tr}(\mathbf{xx'A'D_M A})}{1-\delta_{Res}} \\ &\approx \frac{r_0 (\mathbf{x'D_M x}) \mathbf{Tr}(\mathbf{A'A})}{1-\delta_{Res}} \\ &\approx \frac{r_0 \mathbf{Tr}(\mathbf{A'A})}{1-\delta_{Res}} \sum_{j=1}^{K} \omega_j x_j^2, \end{aligned} \tag{6}$$

where $P(\tau_i)$ is the power of the i-th tap, r_0 is the time resolution for sampled CIR signals, $\boldsymbol{D}_M$ is a diagonal matrix with the diagonal elements $\{1, 2, ...M\}$, x_j^2 is the non-zero sparse coefficient and its corresponding atomic index is ω_j.

The RMS delay spread can be represented by

$$\begin{aligned}\tau_{\text{RMS}} &= \sqrt{\frac{\sum \tau_i^2 P(\tau_i)}{P(\tau_i)} - \overline{\tau}} \\ &= \sqrt{\frac{r_0^2 \mathbf{Tr}(\mathbf{D_M^2}(\mathbf{Ax})(\mathbf{Ax})')}{1-\delta_{Res}} - \overline{\tau}} \\ &= \sqrt{\frac{r_0^2 \mathbf{Tr}(\mathbf{xx'A'D_M^2 A}) - r_0 \mathbf{Tr}(\mathbf{xx'A'D_M A})}{1-\delta_{Res}}} \\ &\approx \sqrt{\frac{(r_0^2 \mathbf{x'D_M^2 x} - r_0 \mathbf{x'D_M x}) \mathbf{Tr}(\mathbf{A'A})}{1-\delta_{Res}}} \\ &\approx \sqrt{\frac{\mathbf{Tr}(\mathbf{A'A})}{1-\delta_{Res}}} \sqrt{\sum_{j=1}^{K} r_0 \omega_j (r_0 \omega_j - 1) x_j^2}. \end{aligned} \tag{7}$$

Referring to (6) and (7), the correlation between channel sparsity and mean delay can be explained intuitively as follows. According to the RIP in CS [6] and the desired reconstruction accuracy, the square sum of sparse coefficients is nearly a constant $\sum_{j=1}^{K} x_j^2 \approx 1 - \delta_{Res}$ for a given channel,

TABLE II: Correlation between Sparsity and Propagation Parameters.

Cases	Types	MPC #	RMS τ_{RMS}	Mean Delay
Simulated	Rayleigh	0.90	0.95	0.93
	Rice (K=3)	0.80	0.89	0.91
	SV	0.49	0.53	0.71
Measured	Distance	0.43	0.62	0.56
	Body Parts	0.54	0.61	0.50
	Heights	0.43	0.57	0.50

no matter what the specified sparsity K is. The trace of a properly selected dictionary is also nearly a constant, i.e., trace$(\mathbf{A}'\mathbf{A}) \approx M$, due to the low coherence requirement between different atoms. Thus, the mean delay in (6) is mainly determined by the weights ω_j. Generally, a dispersive channel generates more well-spaced multipath taps which result in more non-zero sparse coefficients. This is why complex NLOS or fast fading channels tend to show larger mean delay and larger sparsity compared to LOS or flat fading channels. The connection between RMS delay and sparsity can be explained in a similar way.

Table II presents the correlation between sparsity and the three TDL parameters, which is obtained by averaging over different datasets and algorithms. The correlation is defined as

$$r_{s,p} = \frac{N\sum s_i^{d,a} p_i - \sum s_i^{d,a} \sum p_i}{\sqrt{N\sum (s_i^{d,a})^2 - (\sum s_i^{d,a})^2}\sqrt{N\sum p_i - (\sum p_i)^2}}, \tag{8}$$

where i represents the i-th selected CIR profile, d and a are indexes of the dictionary and algorithms used for sparse analysis, respectively, $s_i^{d,a}$ represents the sparsity of the i-th CIR reconstructed by using the d-th dictionary and the a-th algorithm, and p_i is one of the studied channel statistics for the i-th CIR.

The average correlation for simulated channels is about 0.8, while it is about 0.5 for the measured channels. The correlation coefficients between sparsity and the mean delay and RMS delay spread are greater than the one between sparsity and significant MPC numbers. This is because the MPC number is directly related to the sparsity, while the delay and RMS delay spread are related to the sparsity indirectly via the matrix $\boldsymbol{D}_M$.

It is worth noting that the first and second-order channel parameter characteristics in the temporal, frequency and spatial domains have similar mathematical forms. Thus, this relationship can be easily extended to other domains. Such connections between sparse parameters and traditional TDL-model-based first- and second-moment parameters can have valuable applications in sparse processing and performance evaluation areas. For example, using such correlations, RMS delay spread can be used to estimate channel sparsity roughly. On the other hand, the known sparsity can also be used to evaluate the singular value spread or the ergodic capacity of MIMO applications.

V. Validation and Simulation

In this section, we use the obtained sparse channel models to generate simulated channels and compare them with the originally measured/simulated ones, in order to verify the effectiveness of the sparse channel models.

A. Generating Simulation Channels Using SCM Results

Using the statistics of sparsity, MDP and the atomic indexes in the develped sparse models, we can simulate channels with high accuracy particularly in the cluster structure with a limited number of parameters. Mathematically, the major steps for the proposed SCM model can be represented in the set of equations below,

$$\begin{cases} \widehat{\mathbf{y}_N} = \sqrt{P_0/P_L}\ \boldsymbol{\Psi}_{\widehat{\boldsymbol{\omega}}}\widehat{\mathbf{x}}, \\ \widehat{\boldsymbol{\omega}} = \{\widehat{\omega}_i\} = \{\widehat{\gamma}_i * \iota_{\widehat{p}} + \iota_{\widehat{p}}^{\text{Beg}}\}, \\ \widehat{\gamma} \sim \textit{LogNorm}(\mu_\gamma, \sigma_\gamma), \\ \widehat{\mathbf{p}} \sim \textit{Round}(\textit{HalfNorm}(0, \sigma_p), \\ \widehat{\mathbf{x}} \sim \textit{Norm}(\mu_{\widehat{x}}, \sigma_{\widehat{x}}), \end{cases} \tag{9}$$

In (9), P_0 is the transmission power and P_L is the predicted path loss (e.g., (1) in [17]); $\widehat{\mathbf{x}}$ is the sorted non-zero coefficient vector in descending order; $\widehat{\omega}_\mathbf{i}$, $\widehat{\mathbf{p}}$ and $\widehat{\gamma}$ are the randomly generated atomic index, shape and location ratio factors corresponding to (5); $\boldsymbol{\Psi}_{\widehat{\boldsymbol{\omega}}}$ means taking the atoms/columns with indexes $\widehat{\boldsymbol{\omega}}$ from $\boldsymbol{\Psi}$; *Norm*$(\cdot)$, *LogNorm*$(\cdot)$, and *HalfNorm*$(\cdot)$ represent Normal, Log-Normal, and half-Normal distributions, respectively; and *Round*$(\cdot)$ denotes the rounding operation.

So the process of channel generation mainly includes two steps: Firstly generating the sorted sparse coefficients $\widehat{\mathbf{x}}$ using Normal distributions and then generating corresponding sparse coefficient index set $\widehat{\omega}$. The atom shape factor set $\widehat{\gamma}$ and location ratio factor set $\widehat{\mathbf{p}}$ are individually generated by Log-Normal and half-Normal variables. For a given shape factor, the beginning index $\iota_{\widehat{p}}^{\text{Beg}}$ and the total number $\iota_{\widehat{p}}$ of shape atoms $\widehat{\mathbf{p}}$ can be determined. Thus, the index set is generated and the CIR set is finally synthesized.

The basic method described in (9) can be simplified by using the prior information on, e.g., the exponential decaying rules of the sorted sparse coefficients. We present two examples next.

Firstly, assuming that $\widehat{\mathbf{x}}$ can be approximated by a negative exponential decaying function, the sorted coefficients $\widehat{\mathbf{x}}$ can be generated at a reduced complexity using the following function instead of the Gaussian function in (9)

$$\widehat{\mathbf{x}} = (a\, e^{-b\mathbf{k}} + c)(1 + \sigma_x), \tag{10}$$

where σ_x is the shadowing variable of the sorted non-zero sparse coefficients, and $\mathbf{k}$ is the sparse sequence set $\{1, 2, ...K\}$. By replacing the last equation in (9) with (10), an improved SCM predicted sparse coefficients (*SCM-PSC*) model is obtained. It is called "predicted" because these coefficients can be determined once the exponential function is selected.

Secondly, when modelling any new scenario with exponential decaying but unknown parameters, we can try estimating the parameters using any two sorted sparse coefficients. Without loss of generality, let them be the first and k-th sparse coefficients. The parameters for the exponential decaying can be obtained by solving the following ternary equation set

$$\begin{cases} x_1 = a\, e^{-b} + c, \\ x_k = a\, e^{-b\,k} + c \approx c, \\ \sum_{i=1}^{k} (a\, e^{-b\,k} + c) = 1 - \delta_{Res}. \end{cases} \tag{11}$$

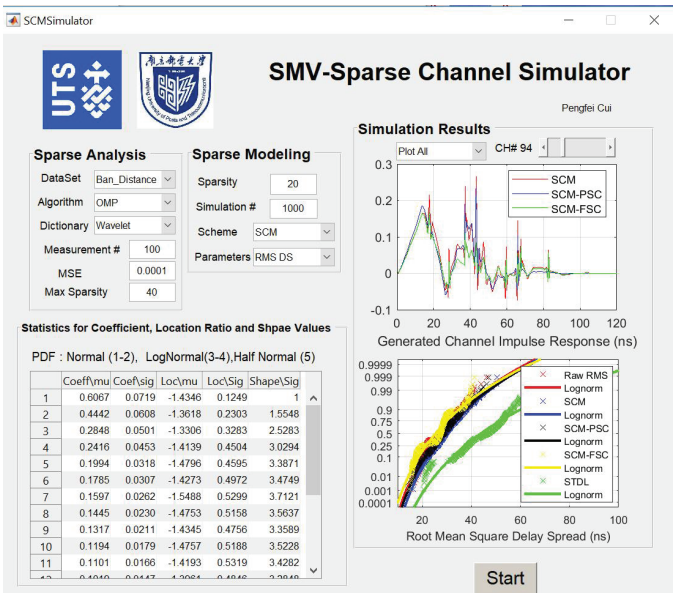

Fig. 15: Graphic user interface for our developed single-measurement-vector sparse channel simulator that is available from Github [25].

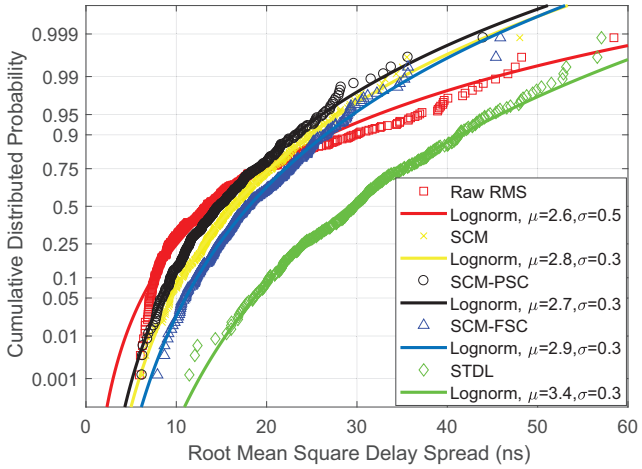

Fig. 16: Model validation for the distributions of the generated root mean square delay spread of BAN_BONBO dataset.

The solution to (11) is shown in (12). Once the parameters are obtained, we can then generate all sorted coefficients for $\widehat{\mathbf{x}}$. We call it SCM Fitted Sparse Coefficients (*SCM-FSC*) scheme for generating channels when detailed exponential decaying parameters are unknown.

By exploiting the magnitude decaying rules, the *SCM-PSC* and *SCM-FSC* schemes reduce the coefficients to be statistically generated from K taps to only 1 or 2 taps. This can largely simplify channel simulation, particularly for high dimensional channels such as massive MIMO channels.

B. Sparse Channel Validation

With the three proposed SCM channel generation schemes in Section V-A, we now generate 1000 Monte Carlo simulations for each dataset to verify the effectiveness of the developed sparse channel models. The widely used statistical TDL (STDL) model is also implemented for comparison [1], [2].

We developed and shared a channel simulator on Github for this work [25]. Fig. 15 depicts the Matlab based Graphic User Interface (GUI) with the corresponding modelling program can be accessed on Github. By setting up the parameters of sparse analysis and modelling modules, the statistics for sorted non-zero coefficients, and shape and location ratio factors are extracted. An example is shown in the left bottom table in Fig. 15. The generated CIRs corresponding to the selected scenario and extracted statistical parameters are shown in the module of modelling results. By changing the plotting drop-down menu and channel number slider, one can configure the desired illustration results. The right upper figure in Fig. 15 shows an example. Three significant clusters which are widely observed in selected BAN_Distance dataset can be seen from generated CIR in 10, 40 and 70 nanoseconds. Such cluster structures are hard to be simulated by traditional modelling methods with limited parameters.

Three Key Performance Indicators (KPIs), including RMS delay spread, mean delay and signal levels, can be chosen on the sparse modelling panel. The right bottom figure in Fig. 15, which is also enlarged in Fig. 16, shows the CDF of the RMS delay spread for the generated CIRs using the standard SCM, SCM-PSC, SCM-FSC and STDL models, for the measured CIRs in BAN_Dis and BAN_Parts. The sparsity is about 20. All RMS delay spreads are best fitted by log-Normal distributions with parameters μ and σ. Referring to the measured RMS delay spread, three SCM schemes all achieve better accuracy than the STDL model. No significant difference is observed among the three SCM schemes.

Table III summarizes the parameters of fitting Log-Normal distribution for extensive Monte Carlo experiments for three measured and two simulated channel datasets, using the SMV-Sparse channel simulator tool. All the RMS delay spread of generated CIRs are found to be better fitted with Log-Normal distributions. It is interesting to see the SCM-PSC model performs slightly better than the other two SCM schemes, with Log-Normal distribution parameters closer to originally measured/simulated data. This reflects the high accuracy of the exponential decaying function in modelling the MDP. The ℓ1-Min recovery algorithm achieves better accuracy than OMP. These results indicate that the proposed SCM schemes work well for different channels, dictionaries, and recovery algorithms.

VI. Conclusion

In this paper, we highlighted the trade-off between sparsity, modelling accuracy and reconstruction complexity in SCM, and introduced a universal three-stage methodology for SCM. We also developed comprehensive statistical sparse channel models, using both measured and simulated channel datasets representing ultra-wideband channels over the frequency band from 6 to 8.5 GHz. For both datasets, channels generated from the developed statistical models match original ones in many

$$\begin{cases} a = \frac{y_1 - y_k}{e^{-b}}, \\ b = -\ln \frac{y_k(y_1-y_k)+\sqrt{y_k^2(y_1-y_k)^2+(1-\sigma_{Res}-ky_k)^2(1-\sigma_{Res}-y_1^2-(k-1)y_k^2)}}{(1-\sigma_{Res}-ky_k^2)^2}, \\ c = y_k, \end{cases} \tag{12}$$

TABLE III: Extracted parameters for the fitting Log-Normal distribution for the CDF of root mean squared delay spread.

Cases	dataset	Algorithm	Measured Data		SCM		SCM-PSC		SCM-FSC		STDL [2]	
			μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Measured	BAN_Parts	OMP	3.17	0.38	3.22	0.26	3.13	0.30	3.14	0.27	3.59	0.30
		ℓ1-Min	3.17	0.38	3.21	0.26	3.18	0.27	3.14	0.28	3.72	0.23
	BAN_Dis	OMP	3.29	0.22	3.39	0.19	3.35	0.21	3.36	0.19	3.79	0.24
		ℓ1-Min	3.29	0.22	3.38	0.20	3.39	0.20	3.26	0.24	3.79	0.17
	BAN_Hei	OMP	3.32	0.51	3.42	0.35	3.34	0.44	3.38	0.39	3.87	0.20
		ℓ1-Min	3.32	0.51	3.42	0.37	3.37	0.40	3.18	0.59	3.84	0.20
Simulated	SV model [4]	OMP	2.23	0.16	2.66	0.35	2.64	0.38	2.38	0.38	2.97	0.28
		ℓ1-Min	2.99	0.20	2.51	0.41	2.57	0.42	2.72	0.42	3.17	0.29
	Rayleigh [3]	OMP	2.17	0.47	2.45	0.45	2.42	0.51	2.06	0.39	2.79	0.34
		ℓ1-Min	3.24	0.36	2.35	0.52	2.46	0.51	2.23	0.53	2.98	0.37

aspects, which demonstrates the robustness of the methodology and the developed models. A summary of important SCM results for the channel datasets in this paper are as follows:

- The sparsity generally conforms to the Normal distribution and can be significantly affected by the channel fading types and LOS conditions;
- Average sparsity is found to be approximately 20 under a modelling accuracy of MSE=10^{-4};
- The sorted non-zero sparse coefficients, the corresponding location ratio factors and shape values follow Normal, Log-Normal and half-Normal distributions, respectively;
- The magnitudes of sorted non-zero sparse coefficients (i.e., the MDP) follow exponentially decaying rule and the decaying speed is mainly affected by channel conditions.

We also established the connection mathematically between parameters for the sparse and conventional TDL models. The averaged correlation coefficients are found to be 0.5 and 0.8 for measured and simulated CIRs, respectively, which demonstrates a high correlation between these parameters.

Based on the developed models, we also proposed three methods for generating simulated channels. It is found that these methods are particularly effective in generating clustered channels, at much lower complexity compared to conventional channel models. The proposed models and channel simulation methods have been implemented in Matlab and shared on Github, which allows open access for promising applications in channel modelling, sparse channel estimation and system design.

ACKNOWLEDGEMENT

The authors would like to thank Yu Yu, Jun She, Yang Liu and Bai Xue etc. for their great contribution in collecting the measured data.

REFERENCES

[1] X. Zhao, J. Kivinen, and P. Vainnikainen, "Tapped delay line channel models at 5.3 ghz in indoor environments," in *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, vol. 1. IEEE, 2000, pp. 1–5.

[2] D. Cassioli, M. Z. Win, and A. F. Molisch, "The ultra-wide bandwidth indoor channel: from statistical model to simulations," *IEEE Journal on selected areas in Communications*, vol. 20, no. 6, pp. 1247–1257, 2002.

[3] G. R. Hiertz, Y. Zang, J. Habetha, and H. Sirin, "Ieee 802.15.3a wireless personal area networks - the mboa approach," in *11th European Wireless Conference 2005 - Next Generation wireless and Mobile Communications and Services*, April 2006, pp. 1–7.

[4] A. A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on selected areas in communications*, vol. 5, no. 2, pp. 128–137, 1987.

[5] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "Quadriga: A 3-d multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 6, pp. 3242–3256, 2014.

[6] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[7] 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.

[8] N. T. Son, S. Guo, and H. Chen, "Impact of channel models on compressed sensing recovery algorithms-based ultra-wideband channel estimation," *IET Communications*, vol. 7, no. 13, pp. 1322–1330, 2013.

[9] C. R. Berger, Z. Wang, J. Huang, and S. Zhou, "Application of compressive sensing to sparse channel estimation," *IEEE Communications Magazine*, vol. 48, no. 11, 2010.

[10] L. Bo, X. Ren, and D. Fox, "Multipath sparse coding using hierarchical matching pursuit," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 660–667.

[11] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by v1?" *Vision research*, vol. 37, no. 23, pp. 3311–3325, 1997.

[12] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 1794–1801.

[13] A. I. Sulyman, A. Alwarafy, G. R. MacCartney, T. S. Rappaport, and A. Alsanie, "Directional radio propagation path loss models for millimeter-wave wireless networks in the 28-, 60-, and 73-ghz bands," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6939–6947, 2016.

[14] D. Zhu, J. Choi, Q. Cheng, W. Xiao, and R. W. Heath Jr, "High-resolution angle tracking for mobile wideband millimeter-wave systems with antenna array calibration," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7173–7189, 2018.

[15] W. U. Bajwa, A. Sayeed, and R. Nowak, "Sparse multipath channels: Modeling and estimation," in *Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009. IEEE 13th*. IEEE, 2009, pp. 320–325.

[16] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 255–271, 2010.

[17] P.-F. Cui, Y. Yu, W.-J. Lu, Y. Liu, and H.-B. Zhu, "Measurement and modeling of wireless off-body propagation characteristics under hospital environment at 6–8.5 ghz," *IEEE Access*, vol. 5, pp. 10 915–10 923, 2017.

[18] P.-F. Cui, Y. Yu, Y. Liu, W.-J. Lu, and H.-B. Zhu, "Body obstruction characteristics for off-body channel under hospital environment at 6 8.5 ghz," in *Ubiquitous Wireless Broadband (ICUWB), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–4.

[19] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Transactions on Information theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[20] P.-F. Cui, J. A. Zhang, W.-J. Lu, Y. J. Guo, and W.-J. L. Zhu, Hong-Bo, "Sparse channel modelling using multi-measurement vector compressive sensing," in *Global Communications (Globecom), 2018 IEEE International Conference on*. IEEE, accepted, 2018.

[21] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

[22] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ? 1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5-6, pp. 877–905, 2008.

[23] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM review*, vol. 51, no. 1, pp. 34–81, 2009.

[24] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Selected papers of hirotugu akaike*. Springer, 1998, pp. 199–213.

[25] P. Cui, https://github.com/PerfeyCui/SMV-SparseChannelSimulator, single Measurement Vector Sparse Channel Simulator program including GUI, accessed 6-October-2018.