Open access • Journal Article • DOI:10.1109/JSTSP.2011.2169649

Compressive Sensing Based High-Resolution Channel Estimation for OFDM System

— Source link

Jia Meng, Wotao Yin, Yingying Li, Nam Tuan Nguyen ...+1 more authors

Institutions: CGG, Rice University, University of Houston

Published on: 01 Feb 2012 - IEEE Journal of Selected Topics in Signal Processing (IEEE)

Topics: Orthogonal frequency-division multiplexing, Communication channel, Compressed sensing and Equalization (audio)

Compressive Sensing Based High Resolution Channel Estimation for OFDM System

Jia (Jasmine) Meng[1], Wotao Yin[2], Yingying Li[2,3], Nam Tuan Nguyen[3], and Zhu Han[3,4]

[1]CGGVeritas, LLC, Houston, TX
[2]Department of Computational and Applied Mathematics, Rice University
[3]Department of Electrical and Computer Engineering, University of Houston
[4]International Scholar, Department of Electronics and Radio Engineering, Kyung Hee University, South Korea

***Abstract*— Orthogonal frequency division multiplexing (OFDM) is a technique that will prevail in the next generation wireless communication. Channel estimation is one of the key challenges in OFDM, since high-resolution channel estimation can significantly improve the equalization at the receiver and consequently enhance the communication performances. In this paper, we propose a system with an asymmetric DAC/ADC pair and formulate OFDM channel estimation as a compressive sensing problem. By skillfully designing pilots and taking advantages of the sparsity of the channel impulse response, the proposed system realizes high resolution channel estimation at a low cost. The pilot design, the use of a high-speed DAC and a regular-speed ADC, and the estimation algorithm tailored for channel estimation distinguish the proposed approach from the existing estimation approaches. We theoretically show that in the proposed system, a N-resolution channel can be faithfully obtained with an ADC speed at $M = O(S^2 \log(N/S))$, where N is also the DAC speed and S is the channel impulse response sparsity. Since S is small and increasing the DAC speed to $N > M$ is relatively cheap, we obtain a high-resolution channel at a low cost. We also present a novel estimator that is both faster and more accurate than the typical ℓ_1 minimization. In the numerical experiments, we simulated various numbers of multipaths and different SNRs and let the transmitter DAC run at 16 times the speed of the receiver ADC for estimating channels at the 16x resolution. While there is no similar approaches (for asymmetric DAC/ADC pairs) to compare with, we derive the Cramér-Rao lower bound.**

I. Introduction

In a typical wireless scenario, the transmitted signal arrives at the receiver via various paths of different lengths. This leads to inter symbol interference (ISI) and posts a major difficulty to information decoding, for example, in orthogonal frequency division multiplexing (OFDM). OFDM has been widely applied in wireless communication systems because it transmits at a high rate, achieves high bandwidth efficiency, and is relatively robust to multipath fading and delay [1]. OFDM applications can be found in digital audio broadcasting (DAB), HDTV-digital video broadcasting (DVB), wireless LAN network, 3GPP Long Term Evolution (LTE), and IEEE 802.16 broadband wireless access system, etc. Current OFDM based WLAN standards (such as IEEE802.11a/g) require a coherent detection at the OFDM receiver. This requirement needs an accurate multipath channel estimation of channel state information (CSI), which comes with computation and bandwidth overheads. There is rich literature on OFDM channel estimation. Below, we provide a brief overview.

There are two major classes of channel estimation schemes. One does not use pilot symbols and is called decision-directed, and the other uses pilot symbols [13]. The approaches in the former class can be deployed where the sending pilot signals is not applicable (e.g., passive listening in a military context) [14], [15]. On the other hand, they require a large amount of data to converge due to the receiver being "blind". The approaches in the latter class can take advantages of the pilots in the transmitted data, which are the training sequences known by both the transmitter and receiver, and therefore, they achieve more accurate channel estimation and are faster. The approach developed in this paper belongs to this class.

The design of a pilot-assisted approach includes both the pilots and the estimation algorithm. The goal is to achieve an optimal combination of spectrum efficiency and estimation accuracy [16–20]. Among the existing OFDM channel estimation approaches, some are based on the time-multiplexed pilot, frequency-multiplexed pilot, and scattered pilot [21]. They achieve relatively higher estimation accuracy yet use relatively more pilots. There have been attempts to reduce the number of pilots such as J. Byun et al. [22], which sends out a small number of pre-estimation pilots to estimate the number of pilots needed in the main estimation. There is no guaranteed overall reduction of pilots though. Another approach is the adaptive channel estimation proposed in [23], which uses a logic controller to choose among several available training patterns. The controller choice is based on the cross-correlation between the pilot symbols over two consecutive time instants, as well as the deviation from the desired bit error rate (BER). Compared with the traditional least-squares channel estimator, this adaptive channel estimation has the advantages of a low BER and high data rate.

Unlike the aforementioned approaches with pilot symbols on regular lattices, the recent work of P. Fertl and G. Matz [24] proposes irregular pilot arrangements and nonuniform sampling techniques along with a conjugate-gradient based channel estimator. Their proposed system features a low computational complexity while maintaining a similar channel estimation accuracy as the mean-squared-error-minimization (MMSE) channel estimator.

We believe that as a sensing problem, OFDM channel estimation can benefit from compressive sensing (CS), which acquires a sparse signal from fewer samples than what is dictated by the Nyquist-Shannon sampling theorem (cf. a survey of this topic in the setting of wireless communication [27]). CS encodes a sparse signal by taking its "incoherent" linear projections and subsequently decodes the signal using sparse optimization such as ℓ_1 minimization. To maximize the benefits of CS for OFDM channel estimation, one shall carefully design its encoding and decoding steps. They correspond to the two focuses of this paper: the designs of the pilots and the estimator, respectively. We shall note that CS has been applied to channel estimation in [28–31], and some preliminary results with little proof and analysis has been published in [39].

Compared to the existing CS-based work [2–5], our approach is unique in various ways as follows. We use pilots with uniform random phases and offer a novel theoretical guarantee for faithful estimation. Its proof is based on first showing a concentration-of-measure phenomenon for a certain subsampled circulant matrix, subsequently showing its restricted isometry property (RIP), and applying the existing RIP-based results to establish the recovery guarantee. The result shows that one can obtain high-resolution channel by just increasing the transmitter DAC speed while keeping the receiver ADC unchanged. In addition, a novel estimator is tailored for OFDM channel response; instead of using the generic ℓ_1 minimization, we modify it to take advantages of the characteristics of channel response, by using iterative support detection (ISD) [6] and a limited-support least-squares subproblem. The resulting algorithm is very simple and performs noticeably better than generic ℓ_1 minimization. Furthermore, we derive a Cramér-Rao lower bound of the mean square error, which is compared to the actual performance of the estimator. We demonstrate the efficiency and effectiveness of the proposed approach. We hope that the results of this paper convince the reader with the potential of the proposed approach as a low-cost and high-performance channel estimator.

The rest of this paper is organized as follows. Section II reviews the general OFDM system model. Section III relates channel estimation to CS and presents the proposed pilot design with its theoretical properties. In Section IV, introduces our OFDM-tailored estimator, analyzes its complexity, and derives a Cramér-Rao lower bound for performance comparison. Section V presents the simulation results. Finally, Section VI concludes this work and discusses some future work.

II. OFDM System Model

A baseband OFDM system is shown in Figure 1. In this system, the modulated signal in the frequency domain, represented by $\mathbf{X} \in \mathbb{C}^N$, is inserted with pilot signal, and then an N-point IDFT transforms the signal into the time domain, denoted by $x \in \mathbb{C}^N$, where a cyclic extension of time length T_G is added to avoid inter-symbol and inter-subcarrier interferences. The resulting time series data are converted by a digital-to-analog converter (DAC) with a clock speed of $1/T_S$ Hz into an analog signal for transmission. We assume that the channel response comprises P propagation paths, which can

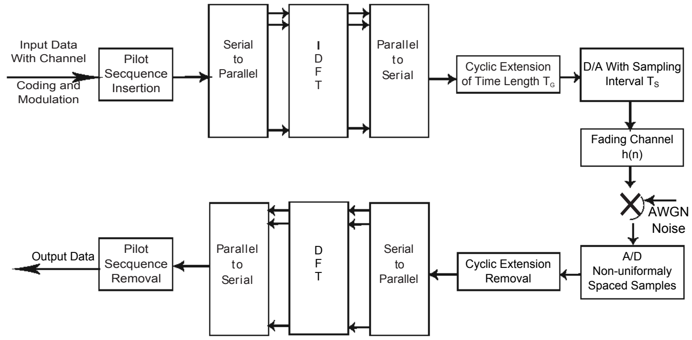

Fig. 1. Baseband OFDM System

be modeled by a time-domain complex-baseband vector with P taps:

$$h_n = \sum_{p=1}^{P} \alpha_p \delta(n - \tau_p T_S),\ n = 1, \dots, N, \tag{1}$$

where α_p is a complex multipath component, δ stands for the Dirac delta, and τ_p is the multipath delay ($0 \leq \tau_p T_S \leq T_G$). Since T_G is shorter than the OFDM symbol duration, the nonzero channel response concentrates at the beginning, which translate to $h = [h_1, h_2, \dots, h_{\tilde{N}}, 0, \dots, 0] \in \mathbb{C}^N$, i.e., only the first $\tilde{N}$ components of h can possibly take nonzero values and $\tilde{N} < N$. Assuming that interferences are eliminated, what arrives at the receiver is the convolution of the transmitted signal and the channel response plus noise, denoted by $z \in \mathbb{C}^N$ given by

$$z = x \otimes h + \xi, \tag{2}$$

where $\otimes$ denotes convolution and ξ denotes the AWGN noise. Passing through the analog-to-digital converter (ADC), $z_n,\ n \in [1, N]$ is sampled as $y_m,\ m \in [1, M]$, and the cyclic prefix (CP) is removed. Traditional OFDM channel estimation schemes assume $M = N$. If $M < N$, then y is a downsample of z. An M-point DFT converts y to $\mathbf{Y} \in \mathbb{C}^M$, where the pilot signal will be removed. The goal is to recover the channel vector h from the measurements $\mathbf{Y}$ (or, equivalently y), given the pilots $\mathbf{X}$ (or, equivalently x). Throughout the paper, we use capital letters for frequency domain signals and lower case letters for time domain signals.

III. Compressive Sensing and Pilot Design

In this section, we present a novel CS based OFDM channel estimation architecture. We first provide the motivation, as well as the CS background. Next, we propose to design pilots with uniform random phases and give the reasons behind. Along with a theoretical guarantee, we present numerical evidence showing that our design achieves an optimal encoding performance. Finally, we compare our proposed approach with the related existing results.

A. Motivation

CS allows sparse signals to be recovered from very few measurements, which often translates to fewer samples and shorter sensing times. Because the channel impulse response

h is very sparse (especially in the outdoor case), we are motivated to apply CS to recover a high-dimensional h from a small number of samples. Since in channel estimation, the sample number is determined by the receiver ADC speed and the dimension of h by the transmitter DAC speed, we propose to obtain a high-dimensional (thus high-resolution) h by employing a pair of high-speed DAC and regular-speed ADC. Here regular-speed means the speed for general data transmission. In today's market, the price for DAC is much lower than that of ADC. Since the ADC runs at a regular speed, we consider our high-speed-DAC approach an inexpensive way to obtain high-resolution channel estimation.

B. CS Background

CS theories [7], [8], [25] state that an S-sparse signal[1] h can be stably recovered from linear measurements $y = \Phi h + \xi$, where Φ is a certain matrix with M rows and N columns, $M < N$, and ξ is noise, by minimizing the ℓ_1-norm of h. Classic CS often assumes that the sensing matrix Φ, after scaling, satisfies the restricted isometry property (RIP)

$$(1 - \delta_S)\|h\|_2^2 \leq \|\Phi h\|_2^2 \leq (1 + \delta_S)\|h\|_2^2 \tag{3}$$

for all S-sparse h, where $0 < \delta_S < 1$ is the RIP parameter. The works in [36], [37], and [43] also study the stable recovery of h from noisy observations based on conditions on δ_S. The RIP is satisfied with high probability by a large class of random matrices such as thoses with entries independently sampled from a subgaussian distribution.

However, the classic random sensing matrices are not admissible in OFDM channel estimation because the channel response h is *not* directly multiplied by a random matrix; instead, as described in Section II, h is *convoluted* with x, followed by noise contamination and uniform downsampling. Because convolution is a circulant linear operator, we can present this process by

$$y = P_\Omega z = P_\Omega(Ch + \xi) = (P_\Omega C)h + \xi_\Omega, \tag{4}$$

where C represents a full circulant (convolution) matrix determined by x, P_Ω denotes the uniform down-sampling from points $[1, N]$ to its subset $\Omega = \{1, 1 + N/M, \ldots, N - N/M + 1\}$, and ξ_Ω is noise. As is widely perceived, CS favors fully random matrices, which enjoy RIPs and thus admit stable recovery from fewest measurements (in terms of order of magnitude), but both P_Ω and C in our case are not as "random". These factors seemingly suggest that $P_\Omega C$ would be unlikely to work well for CS. Nevertheless, carefully designed circulant matrices can deliver the same optimal CS performance.

C. Pilots with Random Phases

To design the sensing matrix C, we propose to generate pilots $\mathbf{X}$ in either one of the following two ways: (i) the real and imaginary parts of $\mathbf{X}(k)$ are sampled independently from the standard Gaussian distribution, $k = 1, \ldots, N$; (ii) (same as [30]) $\mathbf{X}(k)$, $k = 1, \ldots, N$, have independent random phases but a uniform amplitude. Note that since x is the inverse discrete Fourier transform of $\mathbf{X}$, the entries of the resulting x of type (i) are i.i.d. standard Gaussian. Furthermore, $\mathbf{X}$ of type (i) have independent random amplitudes, so type (ii) is more restrictive than type (i). On the other hand, $\mathbf{X}$ of both types have random phases. Let F denote the discrete Fourier transform. From the convolution theorem $x \otimes h = F^{-1}(F(x) \cdot F(h))$ and $x = F^{-1}(\mathbf{X})$, we have $x \otimes h = F^{-1}\mathrm{diag}(\mathbf{X})Fh$, so the measurements y can be written as

$$y = P_\Omega F^{-1}\mathrm{diag}(\mathbf{X})Fh + \xi_\Omega. \tag{5}$$

Note that the proposed sampling is very different from partial Fourier sampling $P_\Omega F$ or $P_\Omega F^{-1}$ widely used in compressive imaging (e.g., MRI). The latter requires a random Ω to avoid aliasing artifacts in the recovered image. In contrast, the proposed scheme permits arbitrary types of Ω including the one corresponding to uniform downsampling, which naturally occurs when the ADC runs at a speed lower than the DAC. Therefore, the proposed scheme is easy to implement in the OFDM system. In the next two subsections, we show the encoding efficiency of this scheme both theoretically and numerically. To keep our exposition general, the discussions in this section do *not* assume that the S nonzero entries of h only occur in its first $\tilde{N} < N$ positions. This property of OFDM channels shall be exploited in the next section to improve both the theoretical and numerical performances.

D. CS by Random Convolution

We first review the existing CS results of random convolution. In [28], Toeplitz[2] measurement matrices are constructed with i.i.d. random row 1 (the same as type (i)) but with only ± 1 or $\{-1, 0, 1\}$; their downsampling effectively takes the first M rows; and the number of measurements needed for stable ℓ_1 recovery is shown as $M \geq O(S^3 \cdot \log N/S)$. [29] uses a "partial" Toeplitz matrix, with i.i.d. Bernoulli or Gaussian row 1, for sparse channel estimation where the downsampling effectively also takes the first M rows. Their scheme requires $M \geq O(S^2 \cdot \log N)$ for stable ℓ_1 recovery. In [30], random convolution of type (ii) with either random downsampling or random demodulation is proposed and studied. It is shown that the resulting measurement matrix is incoherent with any given sparse basis with a high probability and ℓ_1 recovery is stable given $M \geq O(S \cdot \log N + \log^3 N)$. Our proposed type (ii) is motivated by [30]. Recent results in [38] show that several random circulant matrices satisfy the RIP in expectation given $M \geq O(\max\{S^{3/2}\log^{3/2} N, S\log^2 S \log^2 N\})$ with arbitrary downsampling. The rest of this subsection focuses on proving the recovery guarantees for the proposed type-(i) sensing scheme. In short, we shall establish stable recovery under the condition $M \geq O(S^2 \log(N/S))$, that is, when the channel is sparse, there can be up to a log difference between the recovered channel resolution and the receiver ADC speed. We note that one might hope to improve S^2 to S, like in the same of i.i.d. Gaussian sensing matrices, but it will require a novel approach.

[1]In our case, S is equal to P, the number of non-zero taps in (1).

[2]which is slightly more general than circulant.

Let the type-(i) CS sensing matrix be denoted by

$$A := (M^{-1/2})P_\Omega C \in \mathbb{C}^{M\times N}, \tag{6}$$

where $M^{-1/2}$ is just a factor for the normalization purpose, P_Ω is a downsampling operator that keeps the entries in an *arbitrary* index set Ω of cardinality M and discards the rest, and

$$C := \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ x_N & x_1 & \cdots & x_{N-1} \\ & & \ddots & \\ x_2 & x_N & \cdots & x_1 \end{bmatrix}$$

is a circulant matrix with complex standard Gaussian random $x = [x_1; x_2; \cdots; x_N]$.

The proof sketch is the following. The main step is a concentration (isometry) result: for an arbitrary S-sparse vector h with $\|h\|_2 = 1$, $\|Ah\|_2^2$ is concentrated around its mean, which equals 1. The unit-norm of h gives the unit mean; they are not essential. The remaining steps follow the arguments in [40], with minor changes to some formulas and numbers: roughly speaking, we fix an arbitrary index set T with $|T| = S$, pick an ϵ-net $Q \subset H_T = \{h \in \mathbb{C}^N : \mathrm{supp}(h) = T,\ \|h\|_2 = 1\}$ — and use the above concentration result for a single h to establish the isometry for $\|Ah\|$ *uniformly* over $h \in Q \subset H_T$; then, based on the ϵ-net trick and a union bound, the isometry is extended from Q to all $h \in H_T$ uniformly; and finally, the union bound is applied again to extend the isometry property from H_T with a fixed T to the set of all S-sparse vectors. This establishes the RIP of A, more accurately, with high probability given $M \geq O(S^2 \log(N/S))$. Quoting existing RIP-based recovery results, we then obtain stable recovery guarantees for all S-sparse vectors h.

The major work to prove the concentration result is based on reducing $\|Ah\|_2^2$ to $Z = \sum_{i=1}^N a_i(Y_i^2 - 1)$ and applying the following result from [42] that relates the concentration of Z to the parameters a_i.

Lemma 1 (Sec. 4.1 of [42]): Assume that $Y_i \sim \mathcal{N}(0,1)$ for $i = 1, 2, \ldots, N$ i.i.d. and $a = [a_1, \ldots, a_N] \geq 0$. Let $Z := \sum_{i=1}^N a_i(Y_i^2 - 1)$. The following inequalities hold for any $t > 0$:

$$\begin{aligned} \mathbb{P}(Z \geq 2\|a\|_2\sqrt{t} + 2\|a\|_\infty t) &\leq e^{-t}, &(7)\\ \mathbb{P}(Z \leq -2\|a\|_2\sqrt{t}) &\leq e^{-t}. &(8)\end{aligned}$$

Therefore, we shall express $\|Ah\|_2^2$ as Z and bound $\|a\|_2$ and $\|a\|_\infty$.

1) A Concentration Result of Random Circulant Matrices: Let h be such that $\|h\|_2 = 1$ and $\|h\|_0 = |\mathrm{supp}(h)| = S$. (We shall remove the unit-norm assumption later.) We break the development into a few steps:

1) Step 1. Based on the symmetry of convolution, we can rewrite

$$Ah = (M^{-1/2})PCh = (M^{-1/2})PB\tilde{c}, \text{ i.e., } Ch = B\tilde{c}$$

where $\tilde{c} = [x_N; x_{N-1}; \ldots; x_1]$ and

$$B = \begin{bmatrix} h_N & h_{N-1} & \cdots & h_1 \\ h_1 & h_N & \cdots & h_2 \\ & & \ddots & \\ h_{N-1} & h_{N-2} & \cdots & h_N \end{bmatrix}.$$

2) Step 2. Let $U\Sigma V^*$ be the *full-size* singular value decomposition (SVD) of matrix PB, and assume $\mathrm{diag}(\Sigma) = [\sigma_1, \sigma_2, \ldots, \sigma_N]$. Introduce $\bar{c} = V^*\tilde{c}$. Since V is unitary, $\bar{c}$ is complex standard Guassian as well. For simplicity, we assume the real-valued $\bar{c} \sim \mathcal{N}(0, I_N)$, which causes a loss of factor of 2 but does not change the results below in any essential way. Hence,

$$\begin{aligned} \|Ah\|_2^2 &= M^{-1}\|PB\tilde{c}\|_2^2 \\ &= M^{-1}\|U\Sigma V^*\tilde{c}\|_2^2 \\ &= M^{-1}\|\Sigma\bar{c}\|_2^2 \\ &= M^{-1}\sum_{i=1}^N \sigma_i^2\bar{c}_i^2. \end{aligned} \tag{9}$$

To apply Lemma 1, we let $Y_i := \bar{c}_i$ and $a_i := M^{-1}\sigma_i^2$. We shall bound $\|a\|_\infty = M^{-1}(\sup_i |\sigma_i|)^2$ and $\|a\|_2$.

3) Step 3. Since $\|h\|_2 = 1$, we have $\|a\|_1 = M^{-1}\sum_{i=1}^N \sigma_i^2 = M^{-1}\|PB\|_F^2 = \|h\|_2^2 = 1$.

4) Step 4. Since every row or column of B has a unit 2-norm and at most S nonzero entries, the row or column has a maximal 1-norm of $\sqrt{S}$. Hence, we have $\|B\|_1 = \|B\|_\infty \leq \sqrt{S}$ and $\sup_i \sigma_i = \|PB\|_2 \leq \|P\|_2\|B\|_2 \leq 1\cdot\sqrt{\|B\|_1\|B\|_\infty} \leq \sqrt{S}$. Therefore,

$$\begin{aligned} \|a\|_\infty &\leq S/M &(10)\\ \|a\|_2 &\leq \sqrt{\|a\|_1\|a\|_\infty} \leq \sqrt{S/M} &(11)\end{aligned}$$

and applying Lemma 1 to

$$\begin{aligned} Z &= \sum_{i=1}^N a_i(Y_i^2 - 1) \\ &= M^{-1}\sum_{i=1}^N \sigma_i^2\bar{c} - 1 \\ &= \|Ah\|_2^2 - 1 \end{aligned} \tag{12}$$

gives

$$\begin{aligned} \mathbb{P}\left(\|Ah\|_2^2 - 1 \geq 2\sqrt{\frac{tS}{M}} + \frac{2tS}{M}\right) &\leq e^{-t}, &(13)\\ \mathbb{P}\left(\|Ah\|_2^2 - 1 \leq -2\sqrt{\frac{tS}{M}}\right) &\leq e^{-t}. &(14)\end{aligned}$$

Let $\epsilon := 2\sqrt{tS/M} + 2tS/M$ and obtain $t = \frac{\left(\epsilon+1-\sqrt{2\epsilon+1}\right)M}{2S}$. Combining (13) and (14) and noting

$$\begin{aligned} &\mathbb{P}\left(\left|\|Ah\|_2^2 - 1\right| \geq \epsilon\right)_{\|h\|_2=1, \|h\|_0=S} \\ &= \mathbb{P}\left(\left|\|Ah\|_2^2 - \|h\|_2^2\right| \geq \epsilon\|h\|_2^2\right)_{\|h\|_0=S} \end{aligned} \tag{15}$$

we get concentration inequality

$$\mathbb{P}\left(\left|\|Ah\|_2^2 - \|h\|_2^2\right| \geq \epsilon\|h\|_2^2\right)_{\|h\|_0=S} \leq 2\exp\left(-\frac{M}{S}c_0(\epsilon)\right), \quad (16)$$

where $c_0(\epsilon) = \frac{\epsilon+1-\sqrt{2\epsilon+1}}{2}$.

Theorem 1: A matrix A generated by (6) satisfies the concentration inequality (16) for any S-sparse vector h.

2) From Concentration to RIP: Inequality (16) lets us follow the arguments of [40] and obtain the following two results.

Lemma 2: For any given index set T with $|T| = S < M$ and $0 < \delta < 1$, a matrix A generated by (6) satisfies the inequality

$$(1-\delta)\|h\|_2^2 \leq \|Ah\|_2^2 \leq (1+\delta)\|h\|_2^2 : \forall\, h \in \mathbb{C}^n,\ \mathrm{supp}(h) = T, \quad (17)$$

holds with probability at least

$$1 - 2\left(\frac{12}{\delta}\right)^S e^{-\frac{M}{S}c_0(\delta/2)}.$$

From (17) to the RIP inequality (3), we shall applying the union bound with the multiple $\binom{N}{S} \leq (eN/S)^S$. Hence, (3) fails to hold with probability at most

$$2\left(\frac{eN}{S}\right)^S \left(\frac{12}{\delta}\right)^S e^{-\frac{M}{S}c_0(\delta/2)}$$
$$= \exp\left(-\frac{M}{S}c_0(\frac{\delta}{2}) + S[\log(\frac{eN}{S}) + \log(\frac{12}{\delta})] + \log(2)\right). \quad (18)$$

If we choose $c_1 > 0$ and let $M \geq S^2\log(N/S)/c_1$, then $S\log(N/S) \leq \frac{M}{S}c_1$ and the right-hand side of (18) $\leq \frac{M}{S}\left\{-c_0(\delta/2) + c_1[1 + \log^{-1}(N/S)\cdot(1+\log(12/\delta))]\right\} + \log(2)$. Hence, for each δ we can choose c_1 small enough to ensure $\{\cdots\} < -c_0(\delta/2)/2$. Therefore, we get the following:

Theorem 2: Let matrix A be generated by (6). If $M \geq O\left(S^2\log(N/S)\right)$, then A satisfies the RIP with a prescribed $0 < \delta_S < 1$ with probability at least $1 - e^{-O(M/S)}$, where the constants in $O(\cdot)$ depend only on δ.

From Theorem 2 and the fact [43] that $\delta_{2S} < 0.4931$ is a sufficient condition for ℓ_1-minimization to recover all S-sparse vectors universally and recover all nearly S-sparse vectors stably, we can conclude that universal stably recovery condition for matrix A generated by (6) is $M \geq O\left(S^2\log(N/S)\right)$.

E. Intuitive Explanations

Let us explain intuitively why (5) is an effective encoding scheme for a sparse vector h. The key of successful CS encoding is that no matter where the nonzeros in h are, each measurement must contain a roughly equal amount of information from each nonzero in h; in other words, the information in h must spread out in the measurements, and the spreading must not depend on where the information is localized in h. It is commonly known that as long as h is sparse, Fh is non-sparse (the uncertainty principle) and thus its information is spread over all its components. The challenges are to avoid $F^{-1}\mathrm{diag}(\mathbf{X})Fh$ from de-spreading Fh. The random phases of $\mathbf{X}$ by design are of critical importance. They "scramble" the components of Fh and break the "delicate relations" among these components in a way that, contrary to $F^{-1}Fh = h$ being sparse, $F^{-1}\mathrm{diag}(\mathbf{X})Fh$ is not sparse at all. One can see this by recalling that the phases of Fh encode the location of the information in h. When h is sparse, its information is highly localized. Randomly "scrambling" the phases causes the information to spread over. Due to a phenomenon called concentration of measures, the information in h spreads over the components Ch in a way that, with high probability, the sizes of all S-sparse h are uniformly preserved (scaled by a factor $M/|\Omega|$) by $P_\Omega Ch$ with Ω of a size essentially linear in S^2 and $\log(N/S)$. Preserving size means preserving pair-wise distances, so those downsampled measurements perform stable embedding, which subsequently allows ℓ_1 minimization to obtain a stable recovery of h.

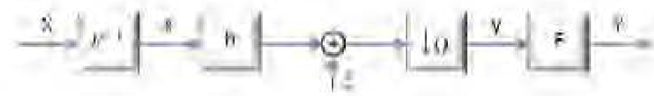

Fig. 2. Logic Block Diagram of the proposed CS-OFDM

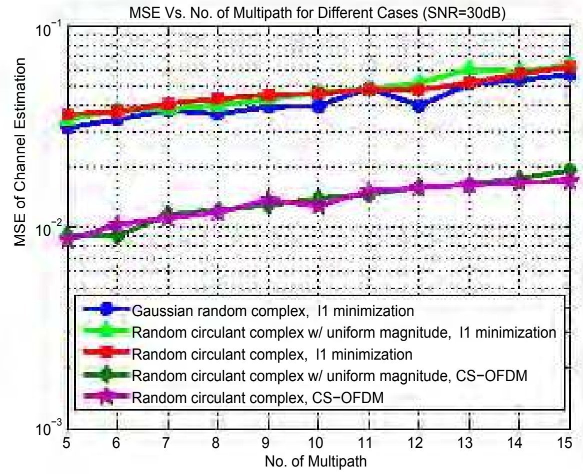

Fig. 3. Mean square error vs. number of multipath (SNR = 30 dB).

F. Numerical Evidence of Effective Random Convolution

CS performance is measured by the number of measurements required for stable recovery. To compare the proposed sensing schemes with the well-established Gaussian random sensing, we conduct numerical simulations and show its results in Figure 3. We compare three types of CS encoding matrices: the i.i.d. Gaussian random complex matrix, and the two circulant random complex matrices corresponding to types (i) and type (ii) above. In addition, the standard ℓ_1 minimization is compared to our proposed algorithm CS-OFDM, which is detailed in the next section. The simulations results show that the random convolutions of both types perform just as well as the Gaussian random sensing matrix, and our algorithm CS-OFDM further improves the performance by half of a magnitude.

G. Relationship to Existing CS-based Channel Estimation

Our work is closely related to [29] and [31]. In [29], i.i.d. Bernoulli or Gaussian vector is used as training sequence, and downsample is carried out by taking only the first M rows, while channel estimation is obtained as a solution to the Dantzig selector. In [31], MIMO channels are estimated by activating all sources simultaneously. The receivers measure the cumulative response, which consists of random convolutions between multiple pairs of source signals and channel responses. Their goal is to reduce the channel estimation time. ℓ_1 minimization is used to recover the channel response.

Our current work is limited to estimating a single h-vector. Although our work is based on similar random convolution techniques, we have proposed to use a pair of high-speed DAC transmitter and regular-speed ADC receiver for the novel goal of high-resolution channel estimation. Furthermore, we derive theoretical guarantees and apply a novel algorithm tailored for the OFDM channel, which is described in details in Section IV below.

IV. OFDM Channel Estimator

In this section, we first formulate the problem for the OFDM channel estimator. Then, we present the numerical algorithm, as well as its complexity analysis. Finally, an estimated performance lower bound is given to evaluate the proposed algorithm.

A. Problem Formulation

As a result of rapid decaying of wireless channels, P — the number of significant multipath components — is small, so the channel response h is highly sparse. Recall that the nonzero components of h only appear in the first $\tilde{N}$ components[3]. We shall recover a sparse high-resolution signal h with a constraint from the measurements y at a lower resolution of M. We define $|\cdot|$ as the amplitude of a complex number, $\|h\|_0$ as the total number of nonzeros of $|h|$, and $\|h\|_1 = \sum_i |h_i|$. The corresponding model is

$$\min_{h\in\mathbb{C}^N} \|h\|_0, \quad \text{s.t.} \begin{cases} y = \phi h, \\ h_i = 0, \ \forall i > \tilde{N}, \end{cases} \tag{19}$$

where ϕ denotes $P_\Omega C = P_\Omega F^{-1}\mathrm{diag}(\mathbf{X})F$ in (5). Generally speaking, problem (19) is NP-hard and is impossible to solve even for moderate N. A common alternative is its ℓ_1 relaxation model with the same constraints.

$$\min_{h\in\mathbb{C}^N} \|h\|_1, \quad \text{s.t.} \begin{cases} y = \phi h, \\ h_i = 0, \ \forall i > \tilde{N}, \end{cases} \tag{20}$$

which is convex and has polynomial time algorithms. If y has no noise, both (19) and (20) can recover h exactly given enough measurements, but (20) requires more measurements than (19).

[3] $\tilde{N}$ is know. Compared with N, the ratio is $1/5$ in the WiFi system (3.2μs for data and 0.8μs for cyclic prefix). Even for 0.8μs, the number of multipath is still relatively small especially in the outdoor environment. Therefore, the channel taps are still sparse.

B. Algorithm

Instead of using a generic algorithm for (20), we design an algorithm specially to exploit the OFDM system features, including the special structure of h and noisy measurements y. At the same time, we maintain algorithm simplicity to achieve low complexity and match with easy hardware implementation.

First of all, we can simply combine two constraints into one by letting the variables be $\tilde{h} = [h_1, h_2, \ldots, h_{\tilde{N}}]$ and dropping the rest components of h. Let $\tilde{\phi}$ be the matrix formed by the first $\tilde{N}$ columns of ϕ. Hence, the only constraints are $\tilde{\phi}\tilde{h} = y$. Since the solution sparsity P remains to be much smaller than $\tilde{N}$, the sparse optimization is still needed. The RIP result in the last section tells us the number of required measurements is $O(S^2 \log(\tilde{N}/S))$, where $S = P$ for OFDM, instead of $O(S^2 \log(N/S))$. Since $N > \tilde{N}$, with the same number of measurements (receiver ADC speed) one can estimate the channel with a large $\tilde{N}$ and thus an even larger N. Moveover, from the computational point of view, it reduces the size and complexity of our problem and thus makes the algorithm faster.

We also develop our algorithm CS-OFDM for the purpose of handling noisy measurements. The iterative support detection (ISD) scheme proposed in [6] has a very good performance for solving (20) even with noisy measurements. Our algorithm uses the ISD, as well as a final denoising step. In the main loop, it estimates a support set I from the current reconstruction and reconstructs a new candidate solution by solving the minimization problem $\min\{\sum_{i\in I^c} |\tilde{h}_i| : \tilde{\phi}\tilde{h} = y\}$, and it iterates these two steps for a small number of iterations. The idea of iteratively updating the index set I helps catch missing spikes and erase fake spikes. This is an ℓ_1-based method but outperforms the standard ℓ_1 minimization. Because the measurements have noise, the reconstruction is never exact. Our algorithm uses a final denoising step, which solves least-squares over the final support T, to eliminate tiny spikes likely due to noise. The pseudocode of the proposed algorithm is listed in Algorithm 1.

In Algorithm 1, at each iteration j, (21) solves a weighted ℓ_1 problem, and the solution h^j is used for support detection to generate a new I^{j+1}. After the main loop is done, a support T is estimated above a threshold, which is selected based on empirical experiences. If the support detection is executed successfully, T would be the set of all channel multipath delay. Finally, $\tilde{h}$ is constructed by solving a small least-squares problem, and $\tilde{h}_i$, $\forall i \notin T$, fall to zero.

C. Complexity Analysis

This algorithm is efficient since every step is simple and the total number of iterations needed is small. The subproblem is a standard weighted ℓ_1 minimization problem, which can be solved by various ℓ_1 solvers. As ϕ is a convolution operator, we choose YALL1 [11] since (i) it allows us to customize the operators involving $\tilde{\phi}$ and its adjoint to take advantages of the

Algorithm 1 CS-OFDM

Input: ϕ, y;

Initalize:

$\tilde{\phi} \leftarrow$ the first $\tilde{N}$ columns of ϕ.

$I^0 \leftarrow \emptyset$

$j \leftarrow 0$

and $w_i^0 \leftarrow 1, \forall i \in \{1, 2, \ldots, \tilde{N}\}$

while the stopping condition is not met, **do**

Subproblem:

$$\tilde{h} \leftarrow \arg\min \sum_{i \notin I^j} |\tilde{h}_i|, \tag{21}$$

$$\text{s.t. } \tilde{\phi}\tilde{h} = y.$$

Support detection:

$I^{j+1} \leftarrow \{i : |\tilde{h}_i| \geq 2^{-j}\|\tilde{h}\|_\infty\}$, where $\|\tilde{h}\|_\infty = \max_i\{|\tilde{h}_i|\}$.

Weights update:

$w_i^{j+1} \leftarrow 0, \forall i \in I^{j+1}$; $w_i^{j+1} \leftarrow 1$, otherwise.

$j \leftarrow j + 1$

end while

Support-restricted least-squares:

$T \leftarrow \{i : |\tilde{h}_i| >$ threshold$\}$; solve

$$\tilde{h}_T \leftarrow \arg\min_{\tilde{h}} \|\tilde{\phi}_T\tilde{h} - y\|_2^2, \tag{22}$$

$$\text{and } \tilde{h}_{T^c} \leftarrow 0.$$

Return $\tilde{h}$ and $h = (\tilde{h}, 0, \ldots, 0)$.

FFT, making it easier to implement the algorithm on hardware, (ii) YALL1 is asymptotically geometrically convergent and efficient even when the measurements are noisy. With our customization, all YALL1 operations are either FFTs or one dimensional vector operations, so the overall complexity is $O(N \log N)$. Moreover, for support detection, we run YALL1 with a more forgiving stopping tolerance and always restart it from the last step solution. Furthermore, YALL1 converges faster as the index I^j gets closer to the true support. The total number of YALL1 calls is also small since the detect support threshold decays exponentially and bounded below by a positive number. Numerical experience shows that the total number of YALL1 calls never exceeds P, which is the number of taps.

The computational cost of the final least-squares step is negligible because the associated matrix $\tilde{\phi}_T$ has its number of columns approximately equal to P, namely, the associated matrix for least-squares has size $M \times P$. Generally speaking, the complexity for this least-squares is $O(MP + P^3)$. Since P and M are much smaller than N, the complexity of the entire algorithm is dominated by that of YALL1, which is $O(PN \log N)$.

D. Cramér-Rao Lower Bound

The Cramér-Rao Lower Bound (CRLB) is an indicator of the performance of any unbiased estimator, which has been used in many applications [12]. In this subsection, we derive a CRLB under the assumption that the tap locations (the support of h) are *known*. We are not aware of ways to drop the support assumption. Since our estimator does not know the support, the support aware CRLB derived is pessimistic. It has a value lower than the CRLB with an unknown support. Nevertheless, the pessimistic CRLB does serve the comparison purpose.

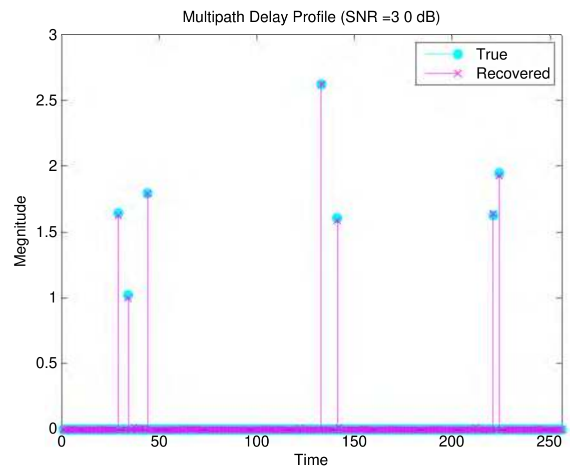

Fig. 4. Example of Reconstructed Multipath Delay Profile.

The CRLB for each entry of h is $\text{CRLB}(h_i) = [\mathbf{I}^{-1}(h)]_{ii}$, where $\mathbf{I}(h)$ is the Fisher information matrix, written as $\mathbf{I}(h) = -\mathbb{E}\left\{\frac{\partial}{\partial h}\log f(y|h)\left[\frac{\partial}{\partial h}\log f(y|h)\right]^*\right\}$, where $\mathbb{E}$ denotes expectation and $f(y|h)$ is the conditional PDF of y given h.

With known $T = \text{supp}(h)$, the channel estimation model can be written as

$$y = \phi_T h_T + \xi, \tag{23}$$

where $\phi = P_\Omega C$, ϕ_T denotes is the sub-matrix of ϕ with columns corresponding to the indices in T, and ξ is the AWGN noise with distribution $N(0, \sigma^2 I_{M\times M})$. Following equation (23), we can derive the conditional PDF of y given h_T:

$$f(y|h_T) = \frac{1}{(2\pi\sigma^2)^{M/2}}\exp\left\{-\frac{1}{2\sigma^2}\|y - \phi_T h_T\|^2\right\}. \tag{24}$$

It is a standard exercise to obtain the overall CRLB:

$$\text{CRLB}(h_T) = \sum_{i=1}^{P}\text{CRLB}[(h_T)_i] = \sigma^2\text{trace}[(\phi_T^*\phi_T)^{-1}]. \tag{25}$$

The above CRLB is compared to the actual performance in the numerical study in the next section.

V. Numerical Simulations

In this section, we present numerical simulations to illustrate the performance of the proposed CS-OFDM algorithm for high-resolution OFDM channel estimation. Our evaluations are based on the mean square error (MSE) of channel estimation and the rate of successful multipath delay detections with respect to different channel profiles and signal-to-noise ratios (SNRs).

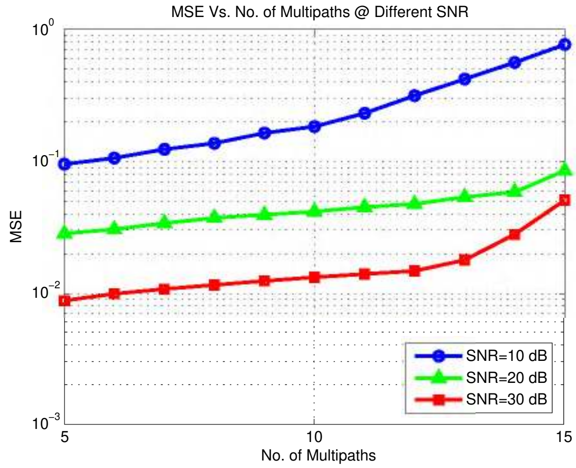

Fig. 5. MSE Performance vs. No. of Multipath

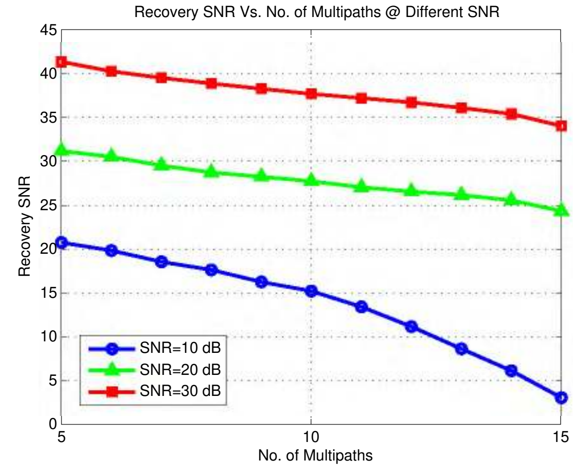

Fig. 6. Reconstructed SNR vs. No. of Multipath

A. Simulation Settings

We consider an OFDM system with $1k$-point IDFT ($N = 1024$) at the transmitter and 64-point DFT ($M = 64$) at the receiver. This gives a compression ratio of 16. The number of silent sub-carrier that acts as guard band is 256 among 1024 sub-carriers. The channel is estimated based on 768 pilot tones with uniformly random phases and a unit amplitude (recall that the unit amplitude does not change estimation results but makes our algorithm faster), with measurement SNRs ranging from 10dB to 30dB. We assume the usage of cyclic prefix and that the impulse response of the channel is shorter than cyclic prefix, i.e., there is no inter-symbol interference. For all simulations, we vary the total number of multipath from 5 to 15. We do not consider the compensation of in phase/quadrature phase (I/Q) imbalance and carrier frequency offset (CFO), and leave them for future work.

B. MSE Performance

Figure 4 is a snapshot of one channel estimation simulation. It shows that the proposed pilot arrangement and CS-OFDM successfully detect an OFDM channel with 7 multipath and SNR=30dB. Our method not only exactly estimates the multipath delays but also correctly estimates the values of the corresponding multipath components.

Figure 5 depicts the MSE performance on OFDM channels with the numbers of multipath varying from 5 to 15 and SNR levels from 10dB to 30dB. As the number of multipath grows, the MSE increases. When there are only a moderate number of multipath on the OFDM channel, the MSE is very low. In addition, the increase of SNR also reduces the MSE for about 10 times per 20dB.

Figure 6 shows the reconstructed SNRs versus the number of multipath at different input SNRs. We can see that CS-OFDM achieves a gain in SNR. For example, when the input SNR is 10dB, we obtain a reconstructed SNR higher than 20dB for 5 multipath. As the number of multipath increases, the SNR gain decreases. However, even when the number of multipath is 10, we still have a 5dB gain, e.g., the reconstructed SNR is 15dB when the input signal SNR is 10dB. The similar SNR gain appears for input SNR$=$ 20dB and SNR$=$ 30dB cases. Over the set of SNRs and multipath numbers in our tested, there is an average gain of 6dB from the input SNR to the recovered SNR.

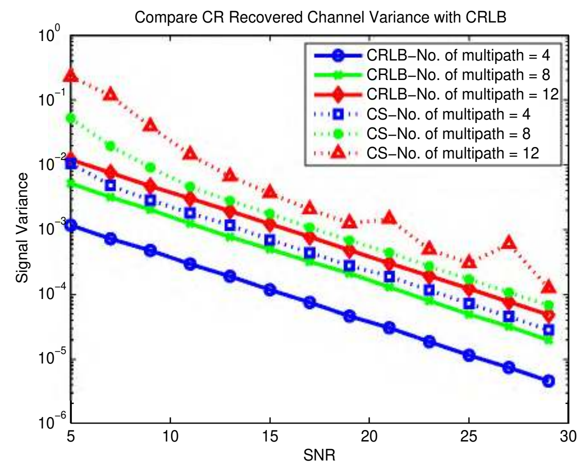

Fig. 7. CS Recovered Channel Variance vs. CRLB

C. CRLB Performance

Figure 7 depicts the estimated channel variance versus the support-known CRLB, corresponding to different SNRs and multipath numbers. Since the algorithm does not know the support while the CRLB does, we believe that the small gaps indicate a strong performance of the algorithm.

D. Multipath Delay Detection Performance

Figures 8 and 9 depict the probability of correct detection (POD) and the false alarm rate (FAR) of multipath delays corresponding to different SNRs and multipath numbers. When the SNR is above 10dB, simulation shows 100% POD for no more than 12 multipath. For the large number of multipath 15, the probability of correct multipath delay detection is higher than 95% for SNR$\geq$ 10dB. Even when SNR is as low as 10dB, as long as the number of multipath does not exceed 10, we

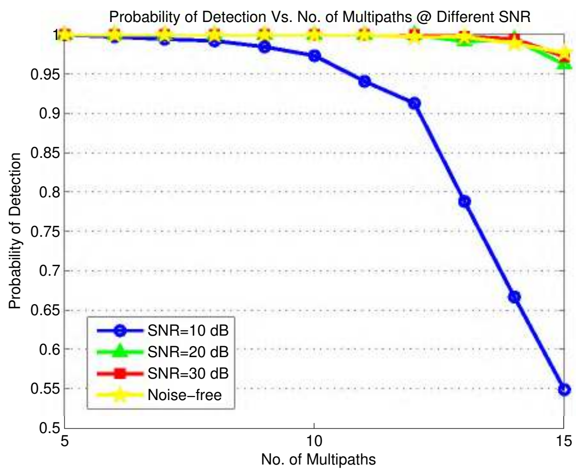

Fig. 8. Probability of Detection vs. No. of Multipath

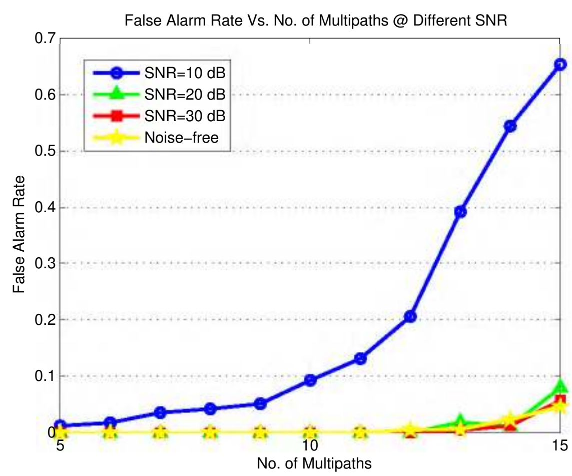

Fig. 9. Probability of False Alarm vs. No. of Multipath

still have a POD of greater than 95%. The FAR performance shows the consistant results: as the SNR decreases and the number of multipath increases, the performance decreases. For SNR$\geq$ 10dB and the number of multipath ≤ 10, we obtain nearly zero FAR.

VI. Conclusions

Efficient OFDM channel estimation will drive OFDM to carry the future of wireless networking. A great opportunity for high-efficiency OFDM channel estimation is lent by the sparse nature of channel response. Riding on the recent development of CS, we propose a design of probing pilots with random phases, which preserves the information of channel response during the convolution and down-sampling processes, and a sparse recovery algorithm, which returns the channel response in high SNR. These benefits translate to the high resolution of channel estimation, as well as shorter probing times. In this paper, the presentation is limited to an idealized OFDM model and simulated experiments. In the future, we will fuse them into more realistic OFDM frameworks. The results presented here hint a high efficiency improvement for OFDM in practice.

Acknowledgements

The work of Zhu Han was partially supported in part by NSF CNS-0910461, NSF CNS-0901425, NSF ECCS-1028782, and NSF CAREER Award CNS-0953377. The work of Wotao Yin was partially supported in part by NSF ECCS-1028790, NSF CAREER Award DMS-07-48839, and ONR Grant N00014-08-1-1101.

References

[1] O. Edfors, M. Sandell, J.-J. Van de Beek, D. Landström, and F. Sjöberg, "An introduction to orthogonal frequency division multiplexing," *Luleȧ Sweden: Luleȧ Tekniska Universitet*, pp. 1–58, September 1996.

[2] C. R. Berger, S. Zhou, P. Willett, B. Demissie, and J. Heckenbach, "Compressed sensing for OFDM/MIMO radar," in *proceedings of the 42nd Annual Asilomar Conference on Signals, Systems and Computers*, pp.213–217, Asilomar, CA, October 2008.

[3] C. R. Berger, S. Zhou, and P. Willett, "Signal extraction using compressed sensing for passive radar with OFDM signals," in *proceedings of the 11th Int. Conf. on Information Fusion*, Cologne, Germany, July 2008.

[4] G. Tauböck and F. Hlawatsch, "A compressed sensing technique for OFDM channel estimation in mobile environments: exploiting channel sparsity for reducing pilots," in *proceedings of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp.2885–2888, Las Vegas, Nevada, April 2008.

[5] C. R. Berger, S. Zhou, W. Chen, and P. Willett, "Sparse channel estimation for OFDM: over-complete dictionaries and super-resolution methods," in *proceedings of IEEE Intl. Workshop on Signal Process. Advances in Wireless Comm.*, pp. 196–200, Perugia, Italy, June 2009.

[6] Y. Wang and W. Yin, "Sparse signal reconstruction via iterative support detection," *Siam Journal on Imaging Sciences*, issue 3, no. 3, p.p. 462–491, August 2010.

[7] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

[8] E. Candes and T. Tao, "Near optimal signal recovery from random projections: universal encoding strategies," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, December 2006.

[9] M. Ledoux, *The Concentration of Measure Phenomenon*, American Mathematical Society. ISBN 0821828649, 2001.

[10] W. Guo and W. Yin. "EdgeCS: an edge guided compressive sensing reconstruction," *in proceedings of Visual Communications and Image Processing (VCIP)*, Huang Shan, An Hui, China, July 2010.

[11] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressives sensing," Rice CAAM Report TR09-37.

[12] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "Bayesian Cramer-Rao bound for noisy. non-blind and blind compressed sensing," *Computing Research Repository(CoRR)*, abs/1005.4316, 2010.

[13] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Transactions on Vehicular Technology* , vol. 49, no. 4, p.p. 1207–1215, July 2000.

[14] P. Ciblat and L. Vandendorpe, "Non-data aided carrier frequency offset estimation for OFDM and downlink DS-CDMA systems," in *proceedings of IEEE 54th Vehicular Technology Conference*, vol. 4, pp. 2618–2622, Atlantic City, USA, October 2001.

[15] P. Ciblat and L. Vandendorpe, "Blind, carrier, frequency offset estimation for noncircular constellation-based transmissions," *IEEE Transactions on Signal Processing*, vol. 51, no. 5, pp. 1378–1389, May 2003.

[16] P. Ciblat and L. Vandendorpe, "On the maximum-likelihood based data-aided frequency offset and channel estimates," in*proceedings of European Signal Processing Conference*, Toulouse, France, vol. 1, pp. 627–630, September 2002.

[17] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. 223–229, September 2002.

[18] H. Wu and X. Huang, "Joint phase/amplitude estimation and symbol detection for wireless ICI self-cancellation coded OFDM systems," *IEEE Transactions on Broadcasting*, vol. 50, issue 1, pp. 49–55, March 2004.

[19] A. Punchihewa, Q. Zhang, O. Dobre, C. Spooner, S. Rajan, and R. Inkol, "On the cyclcostationarity of OFDM and single carrier linearly digitally modulated signals in time dispersive channels: theoretical developments and application," *IEEE Transactions on Wireless Communications*, vol. 9, no. 8, pp. 2588–2599, August 2010.

[20] R. Carrasco-Alvarez, R. Parra-Michel, A. Orozco-Lugo, and J. Tugnait, "Enhanced channel estimation using superimposed training based on universal basis expansion," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 1217–1222, March 2009.

[21] S. Takaoka and F. Adachi, "Pilot-assisted adaptive interpolation channel estimation for OFDM signal reception," in *proceedings of Vehicular Technology Conference*, vol.3, pp. 1777–1781, Milan, Italy, May 2004.

[22] J. Byun and N. P. Natarajan, "Adaptive pilot utilization for OFDM channel estimation in a time varying channel," in *proceedings of Wireless and Microwave Technology Conference*, Clearwater, FL, April 2009.

[23] W. M. Afifi and H. M. Elkamchouchi, "A new adaptive channel estimation for frequency selective time varying fading OFDM channels," in *proceedings of International Conference on Computer Engineering and Systems*, Cairo, Egypt, December 2009.

[24] P. Fertl and G. Matz, "Channel estimation in wireless OFDM systems with irregular pilot distribution," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3180–3194, June 2010.

[25] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[26] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications On Pure and Applied Mathematics*, vol. 59, no.8, pp. 1207–1223, August 2006.

[27] Y. Li, Z. Han, H. Li, and W. Yin, *Compressive Sensing for Wireless Networks*, contract with Cambridge University Press, UK, 2012.

[28] W. U. Bajwa, J. D. Haupt, G. M. Raz, S. J. Wright, and R. D. Nowak, "Toeplitz-structures compressed sensing matrices," in *proceedings of IEEE/SP 14th Workshop on Statistical Signal Processing*, pp. 294–298, Madison, WI, August 2007.

[29] J. D. Haupt, W. U. Bajwa, G. M. Raz, and R. D. Nowak, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," Submitted: August 29, 2008, Revised: March 17, 2010.

[30] J. Romberg, "Compressive sensing by random convolution," *SIAM J. Imaging Sci.*, vol. 2, no. 4, pp. 1098-1128, November 2009.

[31] J. K. Romberg and R. Neelamani, "Sparse channel separation using random probes," to be submitted. `http://www.ece.rice.edu/~jdh6/publications/sub08_toep_rev1.pdf`

[32] Z. Tian and G. Giannakis, "Compressed sensing for wideband cognitive radios", in proceedings of *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1357-1360, April 2007.

[33] J. Meng, W. Yin, H. Li, E. Houssain and Z. Han, "Collaborative spectrum sensing from sparse observations in cognitive radio networks", *IEEE JSAC Special Issue on Cognitive Radio Networking and Communications*, special issue on Advances in Cognitive Radio Networking and Communications, vol.29, no.2, p.p.327-337, February 2011.

[34] J. Paredes, G.Arce, and Z. Wang, "Ultra-wideband compressed sensing: channel estimation", *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 3, p.p. 383–395, October 2007.

[35] J. Meng, J. Ahmadi-Shokouh, H. Li, Z. Han, S. Noghanian, and E. Hossain, "Sampling rate reduction for 60 GHz UWB communication using compressive sensing", in proceedings of *Asilomar Conference on Signals, Systems & Computers*, Asilomar, CA, November 2009.

[36] D. L. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vo. 52, no. 1, p.p. 6–18, January 2006.

[37] J. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," *IEEE Transactions on Information Theory*, in press.

[38] H. Rauhut, J. Romberg, and J. A. Tropp, "Restricted isometries for partial random circulant matrices," arXiv:1010.1847, 2010.

[39] J. Meng, Y. Li, Nam Nguyen, W. Yin, and Z. Han, "High resolution OFDM channel estimation with low speed ADC using compressive sensing," in *proceedings of IEEE International Conference on Communications*, Kyoto, Japan, June 2011.

[40] R. Baraniuk, M. Davenport, R. DeVore, and M.Wakin. "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vo. 28, no. 3, p.p. 253–263, 2008.

[41] E. Candes and T. Tao. "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol.51, no.12, p.p. 4203–4215, December 2005.

[42] B. Laurent and P. Massart. "Adaptive estimation of a quadratic functional by model selection," *Annals of Statistics*, p.p. 1302–1338, 2000.

[43] Q. Mo and S. Li. "New bounds on the restricted isometry constant δ_{2k}," *Applied and Computational Harmonic Analysis*, 2011.

Jia (Jasmine) Meng (S'10) received the B.S. and M. S. degrees in Electrical Engineering from the Southwest Petroleum University (China) in 2004 and 2007, and Ph.D in Electrical Engineering from the University of Houston in 2010. Her research interests are the framework of compressive sensing and implementation for communication and signal processing.

Wotao Yin received the B.S. in mathematics from Nanjing University in 2001, and M.S. and Ph.D. in operations research from Columbia University in 2003 and 2006, respectively. Since 2006, he has been with the faculty of Rice University, the Department of Computational and Applied Mathematics, in Houston, Texas, the United States. Dr. Yin's research interests include optimization, as well as its applications in inverse problems, compressed sensing, signal processing, and variational image processing. Dr. Yin won the NSF CAREER Award in 2008 and the Alfred P. Sloan Research Fellowship in 2009.

Yingying Li received the B.S. degree in mathematics from Peking University in 2006 and the Ph.D. in applied mathematics from the University of California Los Angeles. Dr. Li is currently a postdoc jointly at the University of Houston and Rice University. Her research interests include sparse optimization, machine learning, and algorithm design.

Nam Tuan Nguyen (S'11) received his B.E. degree in Electrical and Computer Engineering from Hanoi University of Technology, Hanoi, Vietnam in 2002 and his M.E. in Electrical and Computer Engineering from the Southern Illinois University at Edwardsville in 2009. In August 2009, he started his PhD studies at the University of Houston. His research interests include applications of machine learning and information theory in cognitive radio, wireless security, wireless networks, and wireless communication systems.

Zhu Han (S'01-M'04-SM'09) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R and D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an assistant professor in Boise State University, Idaho. Currently, he is an Assistant Professor in Electrical and Computer Engineering Department at University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, wireless multimedia, security, and smart grid communication.

Dr. Han is an Associate Editor of IEEE Transactions on Wireless Communications since 2010. Dr. Han is the winner of Fred W. Ellersick Prize 2011. Dr. Han is an NSF CAREER award recipient 2010. Dr. Han is the coauthor for the papers that won the best paper awards in IEEE International Conference on Communications 2009 and 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt09).