

Joint ML Estimation of Channel and RF Chain Responses with Noisy Side Information

Young Yun Kang, Sang Hyun Mo, and Joon Ho Cho

Department of Electronic and Electrical Engineering

Pohang University of Science and Technology (POSTECH)

Pohang, Kyungbuk 790-784, Korea

Email: {jjanguni,mozziya,jcho}@postech.ac.kr

Abstract—In this paper, a frequency-domain joint maximum-likelihood (ML) estimation of channel and radio frequency (RF) chain responses is considered when noisy side information is available. The channel includes only transmit and receive antennas and radio propagation medium, while the RF chain includes transmitter (Tx) and receiver (Rx) RF chains. A side information vector of the same length as an observation vector is obtained by directly connecting the Tx and the Rx RF chains after removing the antennas. Since the Rx RF chain is a part of the RF chain whose response is unknown, the problem appears ill-posed because the covariance matrices of Gaussian noise components in the observation and the side information vectors are unknown. However, it is shown, by using asymptotically equivalent sequences of matrices, that the simple entry-by-entry division of the discrete Fourier transform (DFT) of the observation vector by the DFT of the side information vector leads to the jointly optimal channel estimate if the length of the vectors is sufficiently large. A typical experimental result that is obtained using wideband channel sounders is also provided.

I. INTRODUCTION

Recently, multiple-input multiple-output (MIMO) systems have attracted considerable attention for their potential dramatic increase in the spectral efficiency of wireless communication systems. Accordingly, various modulation, demodulation, coding, and decoding schemes have been proposed that are suitable for MIMO systems [1]. To provide the performance upper bounds that are theoretically achievable by these schemes, intensive investigation on the capacity of MIMO channels also has been conducted [2].

For a slowly-varying or a quasi-static wideband MIMO channel, the capacity is determined by the frequency responses of the elements of a frequency-selective matrix channel. To measure a channel frequency response, two types of channel sounders have been mostly used. One is a sounder that employs a vector network analyzer (VNA) [3] and the other is a sounder that employs an arbitrary waveform generator and an acquisition board followed by a post processor [4]. In general, a VNA-based channel sounder is easier to implement and simpler to obtain the samples of a channel response in the frequency-domain with high resolution within a broad

frequency range. However, a VNA-based sounder has a limited applicability to MIMO channel sounding due to the small number of input and output ports. To overcome this limitation, a VNA-based sounder with a virtual antenna array realized by an antenna positioner can be used [5].

The other is a sounder that employs a multi-channel arbitrary waveform generator and a multi-channel signal acquisition board. This type of channel sounder can simultaneously measure the frequency responses of the channel matrix of a MIMO channel, though now multiple radio frequency (RF) chains are needed to upconvert the intermediate frequency (IF) signals generated by the arbitrary waveform generator and to downconvert the received RF signals to feed the acquisition board. For this type of channel sounders, the signaling scheme called the correlation method [4] is often used to probe the channel.

A direct measurement using either type of channel sounders produces an estimate of the frequency response between the input and the output ports, which includes radio propagation medium, transmit and receive antennas, and transmitter (Tx) and receiver (Rx) RF chains. To remove the effect of the Tx and the Rx RF chains, in [3] and [6], a calibration procedure using a side information is proposed that performs a simple division of the continuous-time Fourier transform (CTFT) of the observation signal by the CTFT of the side information signal, where the side information signal is obtained by directly connecting the Tx and the Rx RF chains after removing the antennas. This procedure is also adopted in [7] for the estimation of the sampled frequency response of the channel, where the discrete Fourier transform (DFT) of the observation vector is entry-by-entry divided by the DFT of the side information vector.

This division method is optimal if the observation and the side information are both noiseless. However, the thermal noise generated by the low-noise amplifier (LNA) in the receiver is unavoidable and its effect must be taken into account during the estimation procedure. In general, the noise covariance matrix is highly ill-conditioned due to the oversampling performed to obtain the observation and the side information vectors. Moreover, the noise covariance matrices of these vectors are even unknown because the noise is added before the Rx RF chain that is a part of the overall RF chain whose response is unknown. So, the problem appears ill-posed.

This work was supported by the Ministry of Information and Communication, Korea, in part under the grant IITA-2007-C1090-0701-0037 for the BrOMA-ITRC@POSTECH supervised by the IITA and in part under the Korea Research Foundation Grant KRF-2006-331-D00338 funded by the Korean Government (MOEHRD).

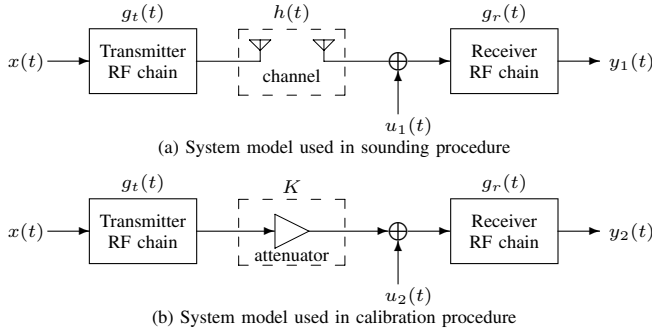


Fig. 1. System models

In [8], the maximum-likelihood (ML) channel estimator with over-sampled noisy data is derived under the assumption of perfect knowledge on the RF chain response. Using an asymptotically equivalent sequence of matrices, it is shown that dividing the DFT of the observation vector by the DFT of the known RF chain response leads to an asymptotic ML estimate of the channel response. However, the joint estimation, based on the over-sampled noisy observation and side information vectors, of the channel and the RF chain responses is not considered. Moreover, only the effect of the Rx RF chain is compensated during the estimation procedure.

In this paper, a frequency-domain joint ML estimation of the channel and the RF chain responses is considered when over-sampled noisy observation and side information vectors are available. It is shown, by using asymptotically equivalent sequences of matrices, that the simple entry-by-entry division of the DFT of the observation vector by the DFT of the side information vector again leads to the ML channel estimate, even though the noise in the vectors is added before the unknown Rx RF chain. This result holds regardless of the unknown Rx RF chain response, as far as the power spectral density of the Gaussian noise component at the input to the faster-than-Nyquist sampler has no spectral nulls on the set of frequencies at which the frequency responses are estimated.

The rest of this paper is organized as follows. In Section II, the signal and system model is described. In Section III, the joint ML estimator using over-sampled noisy side information is derived. A typical experimental result that is obtained using wideband channel sounders is also provided in Section IV, and the concluding remarks are offered in Section V.

Throughout this paper, boldface lowercase letters (e.g., \mathbf{x}) represent vectors, and boldface uppercase letters (e.g., \mathbf{X}) represent matrices. Functions with uppercase letters (e.g., $X(f)$) denote Fourier transforms, and $*$ denotes convolution integral. The transpose is denoted by $\{\cdot\}^T$, while the conjugate transpose is denoted by $\{\cdot\}^H$. The diagonal matrix with its diagonal entries given by the entries of a vector \mathbf{d} is denoted by $\text{diag}\{\mathbf{d}\}$.

II. SIGNAL AND SYSTEM MODEL

As depicted in Fig. 1, the channel sounder has two configurations to jointly estimate channel and RF chain responses—

one for sounding procedure and the other for calibration procedure. The parameters to be estimated are the uniform samples of the frequency responses of the channel and the RF chains in the desired frequency band with bandwidth $\frac{1}{T}$ for some $T > 0$. The channel impulse response denoted by $h(t)$ is defined as the combined effect of transmit and receive antennas and radio propagation medium. The RF chain response $g(t)$ including both the Tx and the Rx RF chain responses is defined as $g(t) \triangleq g_t(t) * g_r(t)$, where $g_t(t)$ and $g_r(t)$ are the Tx and the Rx RF chain response, respectively. Then, the parameters to be estimated can be written as $L \times 1$ vectors

$$\mathbf{h} \triangleq [H(f_1), H(f_2), \dots, H(f_L)]^T \quad (1)$$

and

$$\mathbf{g} \triangleq [G(f_1), G(f_2), \dots, G(f_L)]^T \quad (2)$$

where L is the number of samples in the desired frequency band, $f_i \triangleq \frac{1}{T(L-1)}(i-1) - \frac{1}{2T}$ for $i = 1, 2, \dots, L$, and $H(f)$ and $G(f)$ are the CTFTs of the channel $h(t)$ and the RF chain response $g(t)$, respectively.

For the joint estimation of the channel and the RF chain responses, a periodic sounding signal given by

$$x(t) = \sum_{n=-\infty}^{\infty} a[n]p(t - nT) \quad (3)$$

is transmitted in both of the sounder configurations, where $\{a[n]\}_n$ is a periodic training sequence with period $N = L - 1$. The reason for the choice $N = L - 1$ is that only one period of $x(t)$ is to be observed and processed by the estimator, which corresponds to MN time-domain samples with over-sampling factor M and results in N frequency-domain samples in $\frac{1}{T}$ [Hz] bandwidth. The waveform $p(t)$ is a square-root Nyquist pulse that satisfies $\sum_{k=-\infty}^{\infty} |P(f + k/T)|^2 = T$.

In the sounding procedure depicted in Fig. 1-(a), the channel sounding signal $x(t)$ passes through the Tx RF chain with impulse response $g_t(t)$, a frequency-selective channel with impulse response $h(t)$, and the Rx RF chain with impulse response $g_r(t)$. It is assumed that additive white Gaussian noise (AWGN) generated by an LNA corrupts the signal before the Rx RF chain. In the calibration procedure depicted in Fig. 1-(b), the signal $x(t)$ passes through the Tx RF chain, an attenuator having a flat frequency response, and the Rx RF chain. It is also assumed that the signal is corrupted by AWGN before the Rx RF chain. The attenuator is used to prevent the saturation of the LNA at the Rx RF chain. The gain K of the attenuator is assumed known to the estimator. Then, the output signals can be written as

$$y_1(t) = (x(t) * g_t(t) * h(t) + u_1(t)) * g_r(t) \quad (4)$$

and

$$y_2(t) = (x(t) * (Kg_t(t)) + u_2(t)) * g_r(t) \quad (5)$$

where $u_1(t)$ and $u_2(t)$ are zero-mean independent AWGN processes with two-sided power spectral densities σ_1^2 and σ_2^2 , respectively.

Fig. 2 shows the structure of the joint estimator. The input signals $y_1(t)$ and $y_2(t)$ are first fed to identical receive filters

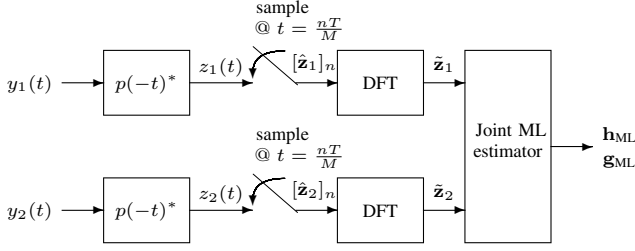


Fig. 2. Structure of joint ML estimator and its front end

whose impulse responses are denoted by $p(-t)^*$. The outputs of the receive filters can be written as

$$z_1(t) = \sum_{n=-\infty}^{\infty} a[n]r_1(t - nT) + v_1(t) \quad (6)$$

and

$$z_2(t) = \sum_{n=-\infty}^{\infty} a[n]r_2(t - nT) + v_2(t) \quad (7)$$

where $r_1(t)$ and $r_2(t)$ are defined as

$$r_1(t) \triangleq p(t) * g_t(t) * h(t) * g_r(t) * p(-t)^* \quad (8)$$

and

$$r_2(t) \triangleq p(t) * (Kg_t(t)) * g_r(t) * p(-t)^*, \quad (9)$$

respectively, i.e., $R_1(f) = H(f)G(f)|P(f)|^2$ and $R_2(f) = KG(f)|P(f)|^2$, where $R_1(f)$, $R_2(f)$, and $P(f)$ are the CTFTs of $r_1(t)$, $r_2(t)$, and $p(t)$, respectively. The noise components $v_1(t)$ and $v_2(t)$ in the receive filter outputs are defined as $v_1(t) \triangleq u_1(t) * g_r(t) * p(-t)^*$ and $v_2(t) \triangleq u_2(t) * g_r(t) * p(-t)^*$, respectively, where $v_1(t)$ and $v_2(t)$ have unknown noise power spectral densities since $g_r(t)$ is unknown to the estimator.

The outputs of the receive filters $z_1(t)$ and $z_2(t)$ are over-sampled at the rate of M/T , which is faster than the Nyquist rate, where the over-sampling factor M is an integer. In both branches, MN consecutive samples are taken from the outputs of the receive filters and form the observation vector $\hat{\mathbf{z}}_1$ and the side information vector $\hat{\mathbf{z}}_2$. It is assumed that the power spectral densities of the Gaussian noise components at the input to the faster-than-Nyquist samplers have no spectral nulls on the set of frequencies at which the frequency responses are estimated. Without loss of generality, we set $[\hat{\mathbf{z}}_1]_n \triangleq z_1((n-1)T/M)$ and $[\hat{\mathbf{z}}_2]_n \triangleq z_2((n-1)T/M)$ for $n = 1, 2, \dots, MN$. The duration of MN samples in the discrete time corresponds to the duration of NT in the continuous time, which is the period of the channel sounding signal $x(t)$. The observation noise vector $[\mathbf{v}_1]_n \triangleq v_1((n-1)T/M)$ and the side information noise vector $[\mathbf{v}_2]_n \triangleq v_2((n-1)T/M)$ are formed by the MN consecutive samples of $v_1(t)$ and $v_2(t)$, respectively. The covariance matrices of the observation noise vector $\mathbb{E}\{\mathbf{v}_1\mathbf{v}_1^H\} \triangleq \mathbf{R}_{\mathbf{v}_1\mathbf{v}_1}$ and the side information noise vector $\mathbb{E}\{\mathbf{v}_2\mathbf{v}_2^H\} \triangleq \mathbf{R}_{\mathbf{v}_2\mathbf{v}_2}$ are $MN \times MN$ Toeplitz matrices whose (i, j) th entries are $[\mathbf{R}_{\mathbf{v}_1\mathbf{v}_1}]_{i,j} = \mathbb{E}\{v_1((T/M)(i-1))v_1((T/M)(j-1))^*\}$

and $[\mathbf{R}_{\mathbf{v}_2\mathbf{v}_2}]_{i,j} = \mathbb{E}\{v_2((T/M)(i-1))v_2((T/M)(j-1))^*\}$, respectively.

In order to rewrite $\hat{\mathbf{z}}_1$ and $\hat{\mathbf{z}}_2$ in terms of \mathbf{h} and \mathbf{g} , we first consider the signal components of the over-sampled receive filter outputs. It can be shown that the discrete-time Fourier transforms (DTFTs) of the periodic extension of the observation vector $\hat{\mathbf{z}}_1$ and the side information vector $\hat{\mathbf{z}}_2$ are given, respectively, by

$$Z_{1,d}(f) = \frac{1}{MN} \sum_{k=-\infty}^{\infty} A_d\left(\frac{k}{MN}\right) \delta\left(f - \frac{k}{MN}\right) \times \frac{M}{T} \sum_{k'=-\infty}^{\infty} R_1\left(\left(f - k'\right)\frac{M}{T}\right) + \text{noise by } \mathbf{v}_1 \quad (10)$$

and

$$Z_{2,d}(f) = \frac{1}{MN} \sum_{k=-\infty}^{\infty} A_d\left(\frac{k}{MN}\right) \delta\left(f - \frac{k}{MN}\right) \times \frac{M}{T} \sum_{k'=-\infty}^{\infty} R_2\left(\left(f - k'\right)\frac{M}{T}\right) + \text{noise by } \mathbf{v}_2 \quad (11)$$

where $Z_{1,d}(f)$, $Z_{2,d}(f)$, and $A_d(f)$ are the DTFTs of the periodic extension of the observation vector $\hat{\mathbf{z}}_1$, that of the side information vector $\hat{\mathbf{z}}_2$, and the periodic training sequence $\{a[n]\}_n$, respectively. The MN -point DFTs of the observation vector $\hat{\mathbf{z}}_1$ and the side information vector $\hat{\mathbf{z}}_2$ are the constant multiples of the coefficients of the impulse trains appearing in a period of the DTFTs $Z_{1,d}(f)$ and $Z_{2,d}(f)$, respectively, i.e.,

$$[\tilde{\mathbf{z}}_1]_k = \frac{1}{T} \sqrt{\frac{M}{N}} A_d\left(\frac{(k-1)}{MN} - \frac{1}{2}\right) R_1\left(\frac{(k-1)}{NT} - \frac{M}{2T}\right) + [\mathbf{W}_{MN}\mathbf{v}_1]_k \quad (12)$$

and

$$[\tilde{\mathbf{z}}_2]_k = \frac{1}{T} \sqrt{\frac{M}{N}} A_d\left(\frac{(k-1)}{MN} - \frac{1}{2}\right) R_2\left(\frac{(k-1)}{NT} - \frac{M}{2T}\right) + [\mathbf{W}_{MN}\mathbf{v}_2]_k \quad (13)$$

for $k = 1, 2, \dots, MN$, where $\tilde{\mathbf{z}}_1 = \mathbf{W}_{MN}\hat{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2 = \mathbf{W}_{MN}\hat{\mathbf{z}}_2$ are the MN -point DFTs of the observation vector $\hat{\mathbf{z}}_1$ and the side information vector $\hat{\mathbf{z}}_2$, respectively, and \mathbf{W}_{MN} is the MN -point DFT matrix whose (l, m) th entry is given by $e^{-j2\pi(l-1-MN/2)(m-1)/(MN)}/\sqrt{MN}$. This effective permutation of the ordinary DFT matrix is necessary because we place the zero frequency at the center of the DFT and the inverse DFT matrices.

The observation vectors $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2$ are $MN \times 1$ vectors that cover the frequency band $[-\frac{M}{2T}, \frac{M}{2T})$. However, the parameters \mathbf{h} and \mathbf{g} to be estimated are $(N+1) \times 1$ vectors that cover the frequency band $[-\frac{1}{2T}, \frac{1}{2T}]$. To estimate \mathbf{h} and \mathbf{g} from the observation vectors $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2$, $MN \times 1$ vectors $\tilde{\mathbf{h}}$ and $\tilde{\mathbf{g}}$ are defined which include the parameters \mathbf{h} and \mathbf{g} , and the nuisance parameters \mathbf{h}_L , \mathbf{h}_U , \mathbf{g}_L and \mathbf{g}_U , respectively, i.e.,

$$\tilde{\mathbf{h}} \triangleq \begin{bmatrix} \mathbf{h}_L \\ \mathbf{h} \\ \mathbf{h}_U \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{g}} \triangleq \begin{bmatrix} \mathbf{g}_L \\ \mathbf{g} \\ \mathbf{g}_U \end{bmatrix} \quad (14)$$

where $\tilde{\mathbf{h}}$ is the $MN \times 1$ vector that covers the frequency band $[-\frac{M}{2T}, \frac{M}{2T})$, i.e., the k th element of $\tilde{\mathbf{h}}$ is defined as $[\tilde{\mathbf{h}}]_k \triangleq H(\frac{(k-1)}{NT} - \frac{M}{2T})$ for $k = 1, 2, \dots, MN$, and $\tilde{\mathbf{g}}$ is the $MN \times 1$ vector that covers the frequency band $[-\frac{M}{2T}, \frac{M}{2T})$, i.e., the k th element of $\tilde{\mathbf{g}}$ is defined as $[\tilde{\mathbf{g}}]_k \triangleq G(\frac{(k-1)}{NT} - \frac{M}{2T})$ for $k = 1, 2, \dots, MN$. The $\frac{MN-(N+1)}{2} \times 1$ vectors \mathbf{h}_L and \mathbf{g}_L cover the lower-side frequency band $[-\frac{M}{2T}, -\frac{1}{2T})$, and the $\frac{MN-(N+1)}{2} \times 1$ vectors \mathbf{h}_U and \mathbf{g}_U cover the upper-side frequency band $(\frac{1}{2T}, \frac{M}{2T})$, respectively.

Let $\tilde{\mathbf{v}}_1 \triangleq \mathbf{W}_{MN}\mathbf{v}_1$ and $\tilde{\mathbf{v}}_2 \triangleq \mathbf{W}_{MN}\mathbf{v}_2$ be the DFTs of \mathbf{v}_1 and \mathbf{v}_2 , respectively. Then, the DFTs $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2$ can be written as

$$\tilde{\mathbf{z}}_1 = \tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} \odot \tilde{\mathbf{h}} + \tilde{\mathbf{v}}_1 \quad \text{and} \quad \tilde{\mathbf{z}}_2 = K\tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} + \tilde{\mathbf{v}}_2 \quad (15)$$

where \odot denotes the entry-by-entry Hadamard product, and an $MN \times 1$ vector \mathbf{a} is defined as

$$\tilde{\mathbf{a}} \triangleq \begin{bmatrix} \mathbf{a}_L \\ \mathbf{a} \\ \mathbf{a}_U \end{bmatrix} \quad (16)$$

with the k th entry given by

$$[\tilde{\mathbf{a}}]_k \triangleq \frac{1}{T} \sqrt{\frac{M}{N}} A_d \left(\frac{(k-1)}{MN} - \frac{1}{2} \right) \left| P \left(\frac{(k-1)}{NT} - \frac{M}{2T} \right) \right|^2 \quad (17)$$

for $k = 1, 2, \dots, MN$. Thus, the vectors \mathbf{a}_L , \mathbf{a} , and \mathbf{a}_U cover the lower-side, the desired, and the upper-side frequency bands, respectively.

III. JOINT ML ESTIMATION OF CHANNEL AND RF CHAIN RESPONSES WITH NOISY SIDE INFORMATION

In this section, the frequency-domain joint ML estimator of the channel and the RF chain responses is derived when the DFTs of the over-sampled noisy observation and side information vectors are available. It is shown, by using asymptotically equivalent sequences of matrices, that the joint ML estimate of the sampled channel frequency response is the simple entry-by-entry division of the DFT of the observation vector by the DFT of the side information vector regardless of the unknown Rx RF chain response.

To derive the joint ML estimator, we utilize an asymptotic property of a sequence of Toeplitz matrices. A Toeplitz matrix is often obtained as the covariance matrix of uniform samples from a wide-sense stationary random process. In our case, the covariance matrices of the observation vector and the side information vector are both Toeplitz matrices. We first define the strong and the weak norms of matrices as follows.

Definition 1: The strong norm of a matrix \mathbf{A} is defined as

$$\|\mathbf{A}\| \triangleq \max_{\|\mathbf{x}\|=1} \sqrt{\mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x}}. \quad (18)$$

Definition 2: The weak norm of an $N \times N$ matrix \mathbf{A} is defined as

$$|\mathbf{A}| \triangleq \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} |a_{k,j}|^2}. \quad (19)$$

Using these definitions, the asymptotic equivalence of two sequences of matrices is defined as follows.

Definition 3: Two sequences of square matrices $\{\mathbf{A}_N\}$ and $\{\mathbf{B}_N\}$ are called asymptotically equivalent if the following conditions are satisfied [9]: 1) The strong norm of \mathbf{A}_N and \mathbf{B}_N are uniformly bounded, i.e., there exists a constant C such that

$$\|\mathbf{A}_N\|, \|\mathbf{B}_N\| \leq C < \infty, \quad N = 1, 2, \dots, \quad (20)$$

and 2) the weak norm of $\mathbf{A}_N - \mathbf{B}_N$ vanishes asymptotically as $N \rightarrow \infty$, i.e.,

$$\lim_{N \rightarrow \infty} |\mathbf{A}_N - \mathbf{B}_N| = 0. \quad (21)$$

In what follows, $\mathbf{A}_N \sim \mathbf{B}_N$ denotes the asymptotic equivalence between the sequences $\{\mathbf{A}_N\}$ and $\{\mathbf{B}_N\}$.

A sequence of Toeplitz matrices is asymptotically equivalent to a sequence of circulant matrices [9]. Moreover, a circulant matrix can be written as a product of an inverse DFT matrix, a diagonal matrix, and a DFT matrix. Thus, the Toeplitz matrices $\mathbf{R}_{\mathbf{v}_1\mathbf{v}_1}$ and $\mathbf{R}_{\mathbf{v}_2\mathbf{v}_2}$ are asymptotically equivalent to $\mathbf{W}_{MN}^H \tilde{\mathbf{\Lambda}}_1 \mathbf{W}_{MN}$ and $\mathbf{W}_{MN}^H \tilde{\mathbf{\Lambda}}_2 \mathbf{W}_{MN}$, respectively, where the diagonal matrix $\tilde{\mathbf{\Lambda}}_1$ is defined as $\tilde{\mathbf{\Lambda}}_1 \triangleq \text{diag}\{S_{\mathbf{v}_1}(-1/2), S_{\mathbf{v}_1}(1/MN - 1/2), \dots, S_{\mathbf{v}_1}((MN - 1)/MN - 1/2)\}$ with $S_{\mathbf{v}_1}(f) \triangleq \sum_{k=-\infty}^{\infty} v_1[k] e^{-j2\pi f k}$ and the diagonal matrix $\tilde{\mathbf{\Lambda}}_2$ is similarly defined. Since the strong and the weak norms are unitarily invariant for square matrices and the $MN \times MN$ covariance matrices $\mathbb{E}\{\tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_1^H\} \triangleq \mathbf{R}_{\tilde{\mathbf{v}}_1\tilde{\mathbf{v}}_1}$ and $\mathbb{E}\{\tilde{\mathbf{v}}_2 \tilde{\mathbf{v}}_2^H\} \triangleq \mathbf{R}_{\tilde{\mathbf{v}}_2\tilde{\mathbf{v}}_2}$ are given by $\mathbf{R}_{\tilde{\mathbf{v}}_1\tilde{\mathbf{v}}_1} = \mathbf{W}_{MN} \mathbf{R}_{\mathbf{v}_1\mathbf{v}_1} \mathbf{W}_{MN}^H$ and $\mathbf{R}_{\tilde{\mathbf{v}}_2\tilde{\mathbf{v}}_2} = \mathbf{W}_{MN} \mathbf{R}_{\mathbf{v}_2\mathbf{v}_2} \mathbf{W}_{MN}^H$, respectively, we have the equivalences

$$\mathbf{R}_{\tilde{\mathbf{v}}_1\tilde{\mathbf{v}}_1} \sim \tilde{\mathbf{\Lambda}}_1 \quad \text{and} \quad \mathbf{R}_{\tilde{\mathbf{v}}_2\tilde{\mathbf{v}}_2} \sim \tilde{\mathbf{\Lambda}}_2. \quad (22)$$

Define a $2MN \times 1$ joint observation vector $\tilde{\mathbf{z}}$ as

$$\tilde{\mathbf{z}} \triangleq \begin{bmatrix} \tilde{\mathbf{z}}_1 \\ \tilde{\mathbf{z}}_2 \end{bmatrix}. \quad (23)$$

Then, due to (22), the complex Gaussian distribution of $\tilde{\mathbf{z}}$ given by

$$\mathcal{CN} \left(\begin{bmatrix} \tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} \odot \tilde{\mathbf{h}} \\ K\tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} \end{bmatrix}, \begin{bmatrix} \mathbf{R}_{\tilde{\mathbf{v}}_1\tilde{\mathbf{v}}_1} & \mathbf{0}_{MN \times MN} \\ \mathbf{0}_{MN \times MN} & \mathbf{R}_{\tilde{\mathbf{v}}_2\tilde{\mathbf{v}}_2} \end{bmatrix} \right) \quad (24)$$

is well approximated as

$$\mathcal{CN} \left(\begin{bmatrix} \tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} \odot \tilde{\mathbf{h}} \\ K\tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} \end{bmatrix}, \begin{bmatrix} \tilde{\mathbf{\Lambda}}_1 & \mathbf{0}_{MN \times MN} \\ \mathbf{0}_{MN \times MN} & \tilde{\mathbf{\Lambda}}_2 \end{bmatrix} \right) \quad (25)$$

for sufficiently large MN , where $\mathbf{0}_{MN \times MN}$ denotes the $MN \times MN$ zero matrix. Since the mean vectors of $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2$ can be written as

$$\tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} \odot \tilde{\mathbf{h}} = \begin{bmatrix} \mathbf{a}_L \odot \mathbf{g}_L \odot \mathbf{h}_L \\ \mathbf{a} \odot \mathbf{g} \odot \mathbf{h} \\ \mathbf{a}_U \odot \mathbf{g}_U \odot \mathbf{h}_U \end{bmatrix} \quad \text{and} \quad K\tilde{\mathbf{a}} \odot \tilde{\mathbf{g}} = \begin{bmatrix} K\mathbf{a}_L \odot \mathbf{g}_L \\ K\mathbf{a} \odot \mathbf{g} \\ K\mathbf{a}_U \odot \mathbf{g}_U \end{bmatrix}, \quad (26)$$

respectively, the upper and lower parts of $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2$ that cover the frequency bands $[-\frac{M}{2T}, -\frac{1}{2T})$ and $(\frac{1}{2T}, \frac{M}{2T})$, respectively, are functions only of the nuisance parameters \mathbf{h}_L , \mathbf{h}_U , \mathbf{g}_L and \mathbf{g}_U . Moreover, the covariance matrices $\tilde{\mathbf{\Lambda}}_1$ and $\tilde{\mathbf{\Lambda}}_2$ are diagonal matrices, i.e., the noise entries are all independent to each other. Therefore, the upper and lower parts of $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2$ are irrelevant to the estimation of \mathbf{h} and \mathbf{g} .

To eliminate the nuisance parameters, an $(N + 1) \times (MN)$ matrix \mathbf{B} is defined as

$$\mathbf{B} \triangleq \begin{bmatrix} \mathbf{0}_{(N+1) \times \frac{MN-(N+1)}{2}} & \mathbf{I}_{N+1} & \mathbf{0}_{(N+1) \times \frac{MN-(N+1)}{2}} \end{bmatrix}, \quad (27)$$

where $\mathbf{I}_{(N+1)}$ denotes the $(N + 1) \times (N + 1)$ identity matrix. Let \mathbf{z}_1 and \mathbf{z}_2 be defined as

$$\mathbf{z}_1 \triangleq \mathbf{B}\tilde{\mathbf{z}}_1 \quad \text{and} \quad \mathbf{z}_2 \triangleq \mathbf{B}\tilde{\mathbf{z}}_2, \quad (28)$$

respectively. Then, for sufficiently large MN , the $2(N+1) \times 1$ joint observation vector \mathbf{z} defined as

$$\mathbf{z} \triangleq \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} \quad (29)$$

follows the complex Gaussian distribution

$$\mathcal{CN} \left(\begin{bmatrix} \mathbf{a} \odot \mathbf{g} \odot \mathbf{h} \\ K\mathbf{a} \odot \mathbf{g} \end{bmatrix}, \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0}_{MN \times MN} \\ \mathbf{0}_{MN \times MN} & \mathbf{\Lambda}_2 \end{bmatrix} \right) \quad (30)$$

where the covariance matrices $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ are defined as

$$\mathbf{\Lambda}_1 \triangleq \mathbf{B}\tilde{\mathbf{\Lambda}}_1\mathbf{B}^H \quad \text{and} \quad \mathbf{\Lambda}_2 \triangleq \mathbf{B}\tilde{\mathbf{\Lambda}}_2\mathbf{B}^H. \quad (31)$$

Thus, for sufficiently large MN , the likelihood function of the parameters \mathbf{h} and \mathbf{g} given \mathbf{z} is well approximated by

$$f_{\mathbf{z}}(\mathbf{z}|\mathbf{h}, \mathbf{g}) = \frac{1}{(\pi)^{2(N+1)} \det(\mathbf{\Lambda}_1) \det(\mathbf{\Lambda}_2)} \times \exp \left\{ -(\mathbf{z}_1 - \mathbf{a} \odot \mathbf{g} \odot \mathbf{h})^H \mathbf{\Lambda}_1^{-1} (\mathbf{z}_1 - \mathbf{a} \odot \mathbf{g} \odot \mathbf{h}) - (\mathbf{z}_2 - K\mathbf{a} \odot \mathbf{g})^H \mathbf{\Lambda}_2^{-1} (\mathbf{z}_2 - K\mathbf{a} \odot \mathbf{g}) \right\}, \quad (32)$$

where the covariance matrices are invertible because the power spectral density of the Gaussian noise component at the input to the faster-than-Nyquist sampler has no spectral nulls on the set of frequencies at which the frequency responses are estimated. So, the log-likelihood function, after dropping constants, becomes

$$\mathcal{L}_{\mathbf{z}}(\mathbf{z}|\mathbf{h}, \mathbf{g}) = -(\mathbf{z}_1 - \mathbf{a} \odot \mathbf{g} \odot \mathbf{h})^H \mathbf{\Lambda}_1^{-1} (\mathbf{z}_1 - \mathbf{a} \odot \mathbf{g} \odot \mathbf{h}) - (\mathbf{z}_2 - K\mathbf{a} \odot \mathbf{g})^H \mathbf{\Lambda}_2^{-1} (\mathbf{z}_2 - K\mathbf{a} \odot \mathbf{g}). \quad (33)$$

Therefore, the joint ML estimates of \mathbf{h} and \mathbf{g} that maximize the log-likelihood function

$$(\mathbf{h}_{\text{ML}}, \mathbf{g}_{\text{ML}}) = \arg \max_{\mathbf{h}, \mathbf{g}} \mathcal{L}_{\mathbf{z}}(\mathbf{z}|\mathbf{h}, \mathbf{g}) \quad (34)$$

are obtained as

$$\mathbf{g}_{\text{ML}} = (K \text{diag}\{\mathbf{a}\})^{-1} \mathbf{z}_2 \quad (35)$$

and

$$\mathbf{h}_{\text{ML}} = (\text{diag}\{\mathbf{a} \odot \mathbf{g}_{\text{ML}}\})^{-1} \mathbf{z}_1 = K(\text{diag}\{\mathbf{z}_2\})^{-1} \mathbf{z}_1. \quad (36)$$

Note that, although the noise covariance matrix is unknown to the estimator, it is just shown that the simple entry-by-entry division of the DFT of the observation vector by the DFT of the side information vector leads to the jointly optimal estimate of the channel \mathbf{h}_{ML} if the length of the observation and side information vectors is sufficiently large.

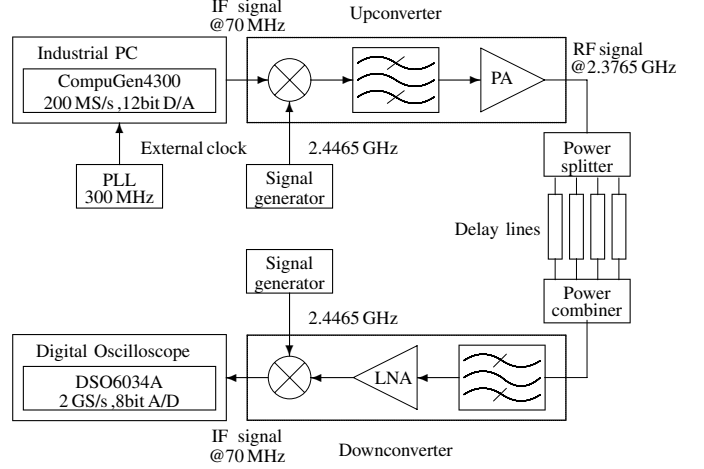


Fig. 3. Block diagram of the correlation-based channel sounder

IV. EXPERIMENTAL RESULTS

In this section, channel measurement results using a VNA-based channel sounder and a correlation-based channel sounder are compared to verify the accuracy of the derived joint ML estimator of the channel and the RF chain responses. Although our result finds the joint ML estimates of the channel and the RF chain responses, only the channel measurement results are compared, since measuring channel frequency responses is the major task performed by a channel sounder.

A. Measurement Setup

Fig. 3 shows the block diagram of the correlation-based channel sounder employed in the measurements. The channel sounder operates at a center frequency of 2.38 GHz with a bandwidth of 20 MHz. The output power of the transmit RF chain is 20 dBm. At the transmitter, a GaGe CompuGen 4300 installed in and controlled by an industrial PC directly generates an IF signal at 70 MHz, and an RF chain upconverts this IF signal. At the receiver, after an RF chain downconverts the received signal to an IF signal, a GaGe CompuScope 12400 digitizes the IF signal and the postprocessing in baseband is performed offline. A pseudo-noise (PN) sequence correlation method is used for channel estimation, where the m -sequence with $N = 1023$ is used as the PN sequence that results in 30 dB processing gain. Phase locked loops (PLLs) are also used to enhance the accuracy of the clock frequencies of the GaGe boards. A calibration procedure is conducted for each channel measurement by taking the IF samples with the transmitter and the receiver RF chains connected by a coaxial cable and compensating the effect of the RF chains in the postprocessing. A VNA-based channel sounder is also employed in the measurements to find the S-parameters of the channel.

For a fair comparison of the two methods, a deterministic channel is needed. However, a deterministic channel is hard to obtain in a radio propagation environment. For this purpose, we construct a deterministic frequency-selective channel by

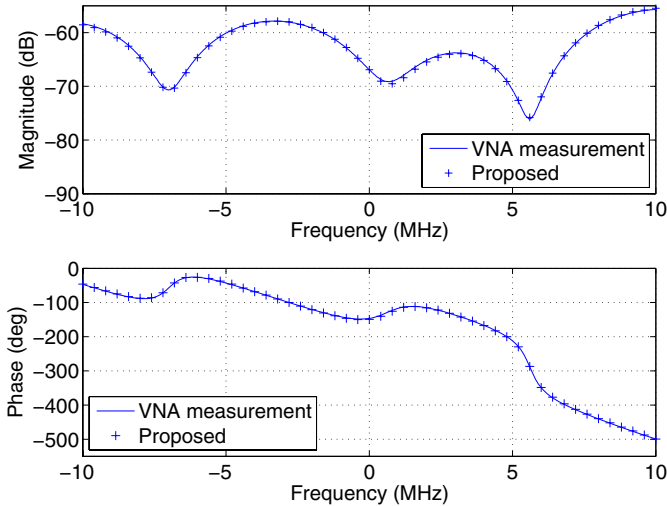


Fig. 4. A typical result obtained using wideband channel sounders when noise level is low.

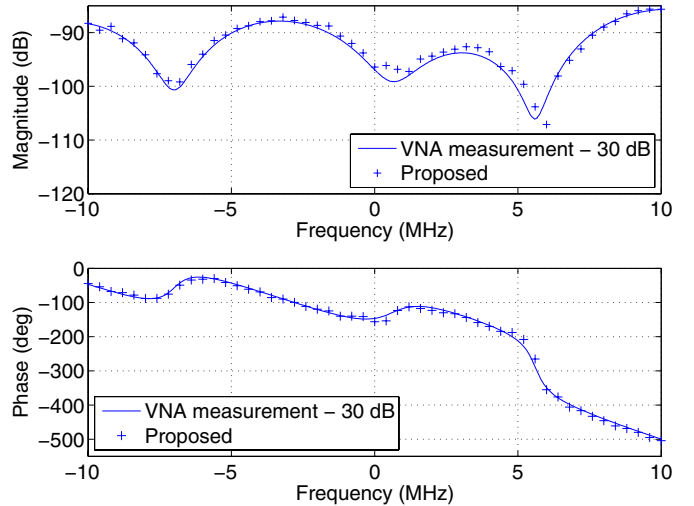


Fig. 5. A typical result obtained using wideband channel sounders when noise level is high.

using a power splitter, four delay lines, attenuators, and a power combiner. So, the impulse response of this channel is given by

$$h(t) = \sum_{i=1}^4 \alpha_i \delta(t - t_i) \quad (37)$$

where α_i and t_i are the gains and the delays of individual paths.

B. Postprocessing

A MATLAB code is written to process the collected data offline. The over-sampled IF signal with $M = 10$ is first downconverted to the baseband, and filtered by a receive filter. Then, $MN = 10230$ samples are fast Fourier transformed to obtain the DFT of the observation vector. The DFT of the side information vector is obtained using the data collected during the calibration procedure. The estimate of the sampled channel frequency response is finally obtained by the entry-by-entry division the DFT of the observation vector by the DFT of the side information vector, which is performed on the samples within the frequency band of interest.

C. A Typical Experimental Result

Fig. 4 shows a typical channel measurement result where the low-noise estimate is obtained by using the proposed method. Fig. 5 shows the result with the same channel attenuated by 30 dB, where the high-noise estimate is in good agreement with the 30 dB attenuated version of the VNA measurement result in Fig. 4, which demonstrates the accuracy of the proposed estimation method. Since the correlation-based channel sounder used in this experiment is not equipped with a time-reference, the ambiguity in linear phase shift is resolved by using a linear regression-like fitting that minimizes the phase error relative to the VNA measurement result [7].

V. CONCLUSIONS

In this paper, a frequency-domain joint ML estimation of channel and RF chain responses is considered when over-sampled noisy observation and side information vectors are available. A side information vector of the same length as an observation vector is obtained by directly connecting the Tx and the Rx RF chains after removing the antennas. It is shown that the simple entry-by-entry division of the DFT of the observation vector by the DFT of the side information vector leads to the joint ML channel estimate. This result holds regardless of the unknown Rx RF chain response. A typical experimental result shows that the proposed method produces very accurate channel estimates both in magnitude and phase.

REFERENCES

- [1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. NY: Cambridge University Press, 2003.
- [2] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. NY: Cambridge University Press, 2005.
- [3] M. S. Varela and M. G. Sanchez, "RMS delay and coherence bandwidth measurements in indoor radio channels in the UHF band," *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 515–525, Mar. 2001.
- [4] Y. Y. Kang, S. H. Mo, and J. H. Cho, "RMS delay spread vs. mean-square path gain for characterization of channel capacity," in *Proc. IEEE Radio and Wireless Symp.*, Orlando, FL, Jan. 22–24, 2008.
- [5] J. Medbo, M. Riback, H. Asplund, and J. Berg, "MIMO channel characteristics in a small macrocell measured at 5.25 GHz and 200 MHz bandwidth," in *Proc. IEEE Veh. Technol. Conf.*, Dallas, TX, Sep. 25–28, 2005, pp. 372–376.
- [6] T. Zwick, T. J. Beukema, and H. Nam, "Wideband channel sounder with measurements and model for the 60 GHz indoor radio channel," *IEEE Trans. Veh. Technol.*, vol. 54, no. 4, pp. 1266–1277, July 2005.
- [7] S. H. Mo, Y. Y. Kang, and J. H. Cho, "Calibrated frequency-domain channel estimation for indoor MIMO channel sounding," in *Proc. 2007 Autumn Microwave and Radio Wave Conference*, Jeju, Korea, Sep. 13–14, 2007, pp. 785–788 (in Korean).
- [8] Y. Jiang, F. W. Sun, and J. S. Baras, "Maximum likelihood slow frequency-selective fading channel estimation using the frequency domain approach," in *Proc. IEEE GLOBECOM'00*, San Francisco, CA, Nov. 27–Dec. 1, 2000, pp. 955–959.
- [9] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Trans. Inf. Theory*, vol. IT-18, pp. 725–730, Nov. 1972.