Markovian Cascaded Channel Estimation for RIS Aided Massive MIMO Using 1-Bit ADCs and Oversampling

SHAO Zhichao, YAN Wenjing, YUAN Xiaojun

(National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China)

DOI: 10.12142/ZTECOM.202201007

https://kns.cnki.net/kcms/detail/34.1294.TN.20220305.1545.002.html, published online March 8, 2022

Manuscript received: 2021-12-27

Abstract: A reconfigurable intelligent surface (RIS) aided massive multiple-input multiple-output (MIMO) system is considered, where the base station employs a large antenna array with low-cost and low-power 1-bit analog-to-digital converters (ADCs). To compensate for the performance loss caused by the coarse quantization, oversampling is applied at the receiver. The main challenge for the acquisition of cascaded channel state information in such a system is to handle the distortion caused by the 1-bit quantization and the sample correlation caused by oversampling. In this work, Bussgang decomposition is applied to deal with the coarse quantization, and a Markov chain is developed to characterize the banded structure of the oversampling filter. An approximate message-passing based algorithm is proposed for the estimation of the cascaded channels. Simulation results demonstrate that our proposed 1-bit systems with oversampling can approach the 2-bit systems in terms of the mean square error performance while the former consumes much less power at the receiver.

Keywords: massive MIMO; reconfigurable intelligent surface; channel estimation; 1-bit ADCs; oversampling

Citation (IEEE Format): Z. C. Shao, W. J. Yan, and X. J. Yuan, "Markovian cascaded channel estimation for RIS aided massive MIMO using 1-bit ADCs and oversampling," *ZTE Communications*, vol. 20, no. 1, pp. 48 – 56, Mar. 2022. doi: 10.12142/ZTECOM.202201007.

1 Introduction

Massive multiple-input multiple-output (MIMO) has been identified as a key technology for future communication systems[1]. In fact, large spatial degrees of freedoms (DoFs) can increase spectral efficiency without requiring extra spectral resources. Recently, reconfigurable intelligent surfaces (RISs) have been proposed as a cost-effective technology for tuning the wireless propagation channel among transceivers[2]. A RIS consists of a large number of meta-atoms that can be controlled by the software to modify their phase shifts, so that incident electromagnetic waves can be mostly reflected to the desired receiver, which makes the wireless transmission more energy-efficient. The combination of RIS and massive MIMO is treated as one of the promising technologies for the sixth-generation wireless communication systems[3].

The channel estimation method for RIS aided massive MIMO systems is a serious challenge, since there exist two cascaded channels, namely, the channel between the users and the RIS and the channel between the RIS and the base station (BS), to be estimated. The acquisition of channel state information (CSI) has been recently studied in Refs. [4 – 7]. TAHA et al.[4] considered a RIS architecture which is a mixture of active and passive elements. This method facilities the channel estimation but increases the hardware cost and energy consumption. In Ref. [5], a non-iterative two-stage channel estimation framework for passive RIS aided millimeter-wave MIMO systems was proposed, where every stage is formulated as a multi-dimensional direction-of-arrival estimation problem. Similarly, the authors in Ref. [6] have proposed a two-stage channel estimation algorithm, namely, sparse matrix factorization and matrix completion, to exploit the rank-deficient structure of the channel. In Ref. [7], the cascade channel estimation is converted into a sparse signal recovery problem by utilizing the properties of Katri-Rao and Kronecker products.

The receiver design in massive MIMO systems, however, becomes more challenging since the power consumption increases rapidly as the number of antennas grows. Among all the components in the radio frequency (RF) chain, a large portion of the total power consumption lies in the analog-to-digital converters (ADCs), whose power consumption grows exponentially with the number of quantization bits[9]. The deployment of current high-resolution (8 – 12 bits) ADCs is a critical bottleneck for the practical use of large-scale MIMO. To alleviate this issue, the use of low-resolution ADCs (1 – 4 bits) can largely reduce the power consumption and is more suitable for the deployment of large-scale MIMO systems.

In this paper, we consider the extreme case of 1-bit resolution, where the in-phase and quadrature components of the received samples are separately quantized to 1 bit. This solution is particularly attractive to massive MIMO systems, since each of the RF chains only contains simple limiting amplifiers (LAs) without the automatic gain control (AGC). This hardware change can largely reduce both the power consumption and the hardware cost at the BS. Prior works on 1-bit massive MIMO have analyzed the sum rate[10], channel estimation[11], and signal detection[12]. Moreover, oversampling is applied to further compensate for the performance loss caused by the coarse quantization[13–14]. Furthermore, the distortion caused by 1-bit quantization and the sample correlation caused by oversampling make the cascaded channel estimation problem even more challenging.

In this paper, we develop an approximate message passing (AMP) based algorithm to solve the considered cascaded channel estimation problem, where the received signal is sampled at a rate beyond Nyquist sampling and quantized to 1-bit. Bussgang decomposition is applied to deal with the coarse quantization and a Markov chain is developed to characterize the correlation of adjacent oversampled samples. The corresponding factor graph is presented and the AMP algorithm is derived. Unlike prior works on AMP-based cascaded channel estimation[7–8], this work considers the statistical characteristic of 1-bit quantization and uses the oversampling technique to increase the estimation accuracy. Simulations show that our proposed algorithm outperforms the method in Refs. [7 – 8] and can even approach the 2-bit Nyquist-sampled systems in terms of the normalized mean square error (NMSE) while the former consumes less power at the receiver.

2 System Model and Problem Statement

In this work, a single-cell uplink RIS aided multi-user 1-bit massive MIMO system with N_t single-antenna users, a RIS with L passive reflecting elements, and a BS with N_r receive antennas are considered, where $N_r \gg N_t$. The system model is depicted in Fig. 1[1], where $p(t)$ is the pulse shaping filter for transmission and $m(t)$ is the matched filter for detection. The received data signal at the n_r-th receive antenna $y^d_{n_r}(t)$ is

$$y^d_{n_r}(t) = m(t) * \sum\nolimits_{l=1}^{L} h_{n_r,l}(t) * \Big(s_l(t)\big(g_{l,n_t}(t) * p(t) * x_{n_t}(t)\big)\Big), \quad (1)$$

where $x_{n_t}(t)$ is the transmitted signal from the n_t-th user; $s_l(t)$ represents the l-th reflecting element at the RIS; $g_{l,n_t}(t)$ and $h_{n_r,l}(t)$ are the channel impulse responses from the user n_t to the l-th reflecting element and from the l-th reflecting element to the n_r-th receive antenna, respectively; * denotes the operation of convolution.

Flat fading channels are considered in this work, i. e., $g_{n_t,l}(t)$ and $h_{n_r,l}(t)$ can be written as

$$g_{l,n_t}(t) = g_{l,n_t}\delta(t) \quad \text{and} \quad h_{n_r,l}(t) = h_{n_r,l}\delta(t), \quad (2)$$

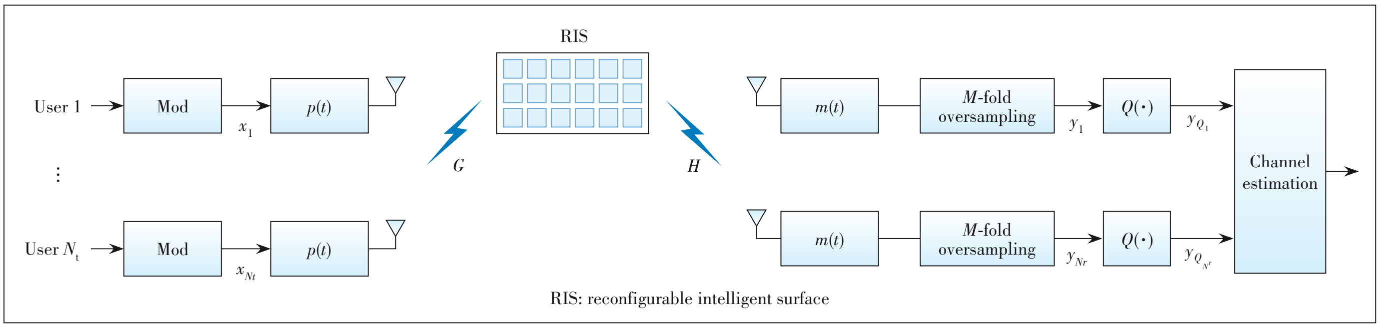
where $g_{n_t,l}$ and $h_{n_r,l}$ are the corresponding channel gains and $\delta(t)$ is the Dirac delta function. Consider a transmission block with length N:

$$x_{n_t}(t) = \sum_{i=0}^{N-1} x_{n_t,i}\delta(t - iT_s) \quad \text{and} \quad s_l(t) = \sum_{i=0}^{N-1} s_{l,i}e^{j\theta_{l,i}}\delta_{t,iT_s}, \quad (3)$$

where T_s is the symbol duration; $x_{n_t,i}$ is the transmitted symbol at the time instant iT_s; $s_{l,i} \in \{0,1\}$ is the on/off state[2] and $\theta_{l,i} \in (0,2\pi]$ is the phase shift of the l-th reflecting element of RIS at the time instant iT_s; $\delta_{t,t'}$ is the Kronecker delta function, where $\delta_{t,t'} = 1$ for $t = t'$ and $\delta_{t,t'} = 0$ otherwise. Eq. (1) can be simplified as:

$$y^d_{n_r}(t) = \sum\nolimits_{l=1}^{L} h_{n_r,l}g_{l,n_t}z(t) * \big(s_l(t)x_{n_t}(t)\big), \quad (4)$$

where $z(t) = p(t) * m(t)$. In oversampled systems, Eq. (4) can be discretized as:



▲**Figure 1. System model of RIS aided multi-user 1-bit massive MIMO system with oversampling at the receiver**

1. Note that the channel matrix in the direct link can be estimated by turning off the RIS[13]. Therefore, the direct-link channel estimation is omitted throughout the paper.

2. In this paper, we assume that the change of the reflection coefficients is synchronized with the transmitted signal. Moreover, the response time of the PIN diode at each reflecting element is assumed to be small enough, so that the duration of each reflection coefficient is the same as the symbol duration. With the above ideal assumptions, no extra harmonics are generated from the surface. The non-ideal case that extra harmonics are generated is beyond the scope of this paper.

$$y_{n_r}^d\left[\frac{i}{M}\right]=\sum_{l=1}^{L}h_{n_r,l}g_{l,n_t}\sum_{k=-M}^{M}z\left[\frac{k}{M}\right]s_l\left[\frac{i-k}{M}\right]x_{n_t}\left[\frac{i-k}{M}\right],$$
$$0\leqslant i\leqslant MN-1, \tag{5}$$

where M is the oversampling factor. Since there are no data symbols at time instants $\frac{i}{M}T_s(i\neq 0,M,2M,...)$, each pair of transmitted data symbols at adjacent time instants are interpolated with $M-1$ zeros. Let $\boldsymbol{x}_{n_t}\triangleq\left[x_{n_t,0},...,x_{n_t,N-1}\right]^T\in\mathbb{C}^{N\times 1}$, $\boldsymbol{s}_l\triangleq\left[s_{l,0}e^{j\theta_{l,0}},...,s_{l,N-1}e^{j\theta_{l,N-1}}\right]^T\in\mathbb{C}^{N\times 1}$ and $y^d_{n_r,\frac{i}{M}}\triangleq y^d_{n_r}\left[\frac{i}{M}\right]$, and we rewrite Eq. (5) in a matrix form as:

$$\boldsymbol{y}_{n_r}^d=\left[y^d_{n_r,0},y^d_{n_r,\frac{1}{M}},...,y^d_{n_r,\frac{MN-1}{M}}\right]^T=$$
$$\sum\nolimits_{l=1}^{L}h_{n_r,l}g_{l,n_t}\boldsymbol{Z}\left(\boldsymbol{I}_N\otimes\boldsymbol{u}\right)\mathrm{diag}\left(\boldsymbol{s}_l\right)\boldsymbol{x}_{n_t}, \tag{6}$$

where $\boldsymbol{Z}\in\mathbb{R}^{MN\times MN}$ is the Toeplitz matrix with the form as

$$\boldsymbol{Z}=\begin{bmatrix} z[0] & z\left[\frac{1}{M}\right] & \cdots & z[1] & 0 & 0 & \cdots & 0 & 0\\ z\left[-\frac{1}{M}\right] & z[0] & \cdots & z\left[\frac{M-1}{M}\right] & z[1] & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & 0 & z[-1] & \cdots & z\left[-\frac{1}{M}\right] & z[0]\end{bmatrix}. \tag{7}$$

In Eq. (6), $\boldsymbol{u}=\left[0,...,0,1\right]^T\in\mathbb{R}^{M\times 1}$ is the zero-inserting vector and $\boldsymbol{I}_N$ denotes the $N\times N$ identity matrix. Furthermore, $\otimes$ represents the Kronecker product and $\mathrm{diag}(\boldsymbol{a})$ is a diagonal matrix with the diagonal specified by $\boldsymbol{a}$. In particular, $M=1$ refers to the case of Nyquist sampling rate.

Similar to Eq. (6), the received oversampled noise samples at the n_r-th receive antenna $\mathbf{y}_{n_r}^n\in\mathbb{C}^{MN\times 1}$ is

$$\boldsymbol{y}_{n_r}^n=\boldsymbol{F}\left(\boldsymbol{I}_N\otimes\boldsymbol{u}\right)\boldsymbol{w}_{n_r}, \tag{8}$$

where $\boldsymbol{F}\in\mathbb{R}^{MN\times MN}$ is the Toeplitz matrix constituted by $m\left[\frac{k}{M}\right]$ with the form similar to Eq. (7), and $\boldsymbol{w}_{n_r}\sim\mathcal{CN}\left(0_{N\times 1},\sigma^2\boldsymbol{I}_N\right)$ represents the complex Gaussian random variables with zero mean and variance σ^2.

Combining Eqs. (6) and (8), the received oversampled samples at the n_r-th receive antenna $\mathbf{y}_{n_r}\in\mathbb{C}^{MN\times 1}$ are

$$\boldsymbol{y}_{n_r}=\boldsymbol{y}_{n_r}^d+\boldsymbol{y}_{n_r}^n=$$
$$\boldsymbol{Z}\left(\boldsymbol{I}_N\otimes\boldsymbol{u}\right)\sum_{l=1}^{L}h_{n_r,l}g_{l,n_t}\mathrm{diag}\left(\boldsymbol{s}_l\right)\boldsymbol{x}_{n_t}+\boldsymbol{F}\left(\boldsymbol{I}_N\otimes\boldsymbol{u}\right)\boldsymbol{w}_{n_r}. \tag{9}$$

Defining $\boldsymbol{Z}''\triangleq\boldsymbol{Z}\left(\boldsymbol{I}_N\otimes\boldsymbol{u}\right)$, $\boldsymbol{F}''\triangleq\boldsymbol{F}\left(\boldsymbol{I}_N\otimes\boldsymbol{u}\right)$, and stacking the n_r-th received oversampled samples on top of the previous receive antenna, Eq. (9) is extended to

$$\boldsymbol{y}=\left[\boldsymbol{y}_1{}^T,\cdots,\boldsymbol{y}_{N_r}{}^T\right]^T=\boldsymbol{Z}'\mathrm{vec}\left\{\boldsymbol{A}^T\right\}+\boldsymbol{F}'\mathrm{vec}\left\{\boldsymbol{W}^T\right\}, \tag{10}$$

where $\boldsymbol{Z}'\triangleq\boldsymbol{I}_{N_r}\otimes\boldsymbol{Z}''$, $\boldsymbol{F}'\triangleq\boldsymbol{I}_{N_r}\otimes\boldsymbol{F}''$ and $\mathbf{W}=\left[\mathbf{w}_1^T;...;\mathbf{w}_{N_r}^T\right]\in\mathbb{C}^{N_r\times N}$. Consider all the users $\boldsymbol{A}\triangleq\boldsymbol{H}\left(\boldsymbol{S}\odot(\mathbf{GX})\right)\in\mathbb{C}^{N_r\times N}$, where $\boldsymbol{G}\in\mathbb{C}^{L\times N_t}$ and $\boldsymbol{H}\in\mathbb{C}^{N_r\times L}$ are the channel matrices from the users to the RIS and from the RIS to the BS, respectively; $\boldsymbol{X}=\left[\boldsymbol{x}_1^T;...;\boldsymbol{x}_{N_t}^T\right]\in\mathbb{C}^{N_t\times N}$ and $\boldsymbol{S}=\left[\boldsymbol{s}_1^T;...;\boldsymbol{s}_L^T\right]\in\mathbb{C}^{L\times N}$. $\mathrm{vec}\{\boldsymbol{A}\}$ denotes the operation of vectorization by stacking the columns of $\boldsymbol{A}$ sequentially on top of one another.

Let $\mathcal{Q}(\cdot)$ represent the 1-bit quantization function. The resulting quantized signal $\boldsymbol{y}_{\mathcal{Q}}$ is given by

$$\boldsymbol{y}_{\mathcal{Q}}=\mathcal{Q}\left(\boldsymbol{y}\right)=\frac{1}{\sqrt{2}}\left(\mathrm{sign}\left(\boldsymbol{y}^{\Re}\right)+j\mathrm{sign}\left(\boldsymbol{y}^{\Im}\right)\right), \tag{11}$$

where $\mathrm{sign}(\cdot)$ is the signum function, and $j=\sqrt{-1}$. $(\cdot)^{\Re}$ and $(\cdot)^{\Im}$ denote the real and imaginary parts, respectively.

The problem to be solved is estimating the channel parameters in $\boldsymbol{G}$ and $\boldsymbol{H}$. The minimum mean square error (MMSE) estimators of the channel matrices are given by

$$\min_{\hat{G}}E\left\{\left\|\boldsymbol{G}-\hat{\boldsymbol{G}}\right\|_F^2\right\}\ \text{and}\ \min_{\hat{H}}E\left\{\left\|\boldsymbol{H}-\hat{\boldsymbol{H}}\right\|_F^2\right\}. \tag{12}$$

Based on the Bayes' rule, the closed-form solutions of Eq. (12) are given by $E\{\boldsymbol{G}|\boldsymbol{y}_{\mathcal{Q}}\}$ and $E\{\boldsymbol{H}|\boldsymbol{y}_{\mathcal{Q}}\}$. For avoiding the high computational complexity of calculating $p\left(\boldsymbol{G}|\boldsymbol{y}_{\mathcal{Q}}\right)$ and $p\left(\boldsymbol{H}|\boldsymbol{y}_{\mathcal{Q}}\right)$ in the marginalization of $p\left(\boldsymbol{G},\boldsymbol{H}|\boldsymbol{y}_{\mathcal{Q}}\right)$, a practical message-passing based algorithm is presented in the next section.

3 Markovian Cascaded Channel Estimation

3.1 Markov Chain for Oversampling

Due to the banded structure of the matrix $\boldsymbol{Z}$ in Eq. (7), we develop a Markov chain, where every two adjacent columns in $\boldsymbol{A}$ are combined together for defining one state $\boldsymbol{b}_n$ as

$$\boldsymbol{b}_1=\mathrm{vec}\left\{\left[0_{N_r\times 1},\boldsymbol{a}_1\right]^T\right\},...,\boldsymbol{b}_N=\mathrm{vec}\left\{\left[\boldsymbol{a}_{N-1},\boldsymbol{a}_N\right]^T\right\}, \tag{13}$$

and each state is with the mean and covariance:

$$\hat{\boldsymbol{b}}_n = \mathrm{vec}\left\{\left[\hat{\boldsymbol{a}}_{n-1}, \hat{\boldsymbol{a}}_n\right]^T\right\} \text{ and}$$
$$\boldsymbol{V}_n^b = \mathrm{diag}\left(\mathrm{vec}\left\{\left[\boldsymbol{v}^a{}_{n-1}, \boldsymbol{v}_n^a\right]^T\right\}\right), \tag{14}$$

with $\boldsymbol{a}_n \in \mathbb{C}^{N_r \times 1}$ being the n-th column of $\boldsymbol{A}$. The transition function from the current state to the next is

$$\boldsymbol{b}_{n+1} = \boldsymbol{T}\boldsymbol{b}_n + \mathrm{vec}\left\{\left[0_{N_r \times 1}, \boldsymbol{a}_{n+1}\right]^T\right\}, \tag{15}$$

where $\boldsymbol{T} = \boldsymbol{I}_{N_r} \otimes [0,1;\, 0,0]$ and the conditional probability is given by

$$p\left(\boldsymbol{b}_{n+1} | \boldsymbol{b}_n, \boldsymbol{a}_{n+1}\right) = \\ \mathcal{CN}(\boldsymbol{b}_{n+1};\, \boldsymbol{T}\hat{\boldsymbol{b}}_n + \mathrm{vec}\left\{\left[0_{N_r \times 1}, \hat{\boldsymbol{a}}_{n+1}\right]^T\right\}, \\ \mathrm{diag}\left(\mathrm{vec}\left\{\left[0_{N_r \times 1}, \boldsymbol{v}_{n+1}^a\right]^T\right\}\right)). \tag{16}$$

The system model of Eq. (11) can be decomposed by using the Bussgang theorem[15]:

$$\boldsymbol{y}_Q = \boldsymbol{K}\boldsymbol{Z}'\mathrm{vec}\left\{\boldsymbol{A}^T\right\} + \boldsymbol{w}' \quad \text{with} \quad \boldsymbol{K} = \sqrt{\frac{2}{\pi}}\,\mathrm{diag}\left(\boldsymbol{C}_y\right)^{-\frac{1}{2}}, \tag{17}$$

where $\boldsymbol{w}' \triangleq \boldsymbol{K}\boldsymbol{F}'\mathrm{vec}\{\boldsymbol{W}\} + \boldsymbol{n}_q$ is the equivalent noise containing the filtered noise and the quantization noise, assumed to follow the Gaussian distribution with zero mean and covariance $\boldsymbol{V}^{w'} = \sigma_n^2 \boldsymbol{K}\boldsymbol{F}\boldsymbol{F}^H\boldsymbol{K}^H + \boldsymbol{C}_{n_q}$. And $\mathrm{diag}\left(\boldsymbol{C}_y\right)$ is the diagonal matrix with the diagonal specified by the diagonal of $\boldsymbol{C}_y$. The received quantized signal $\boldsymbol{y}_{Q_n}$ at the state $\boldsymbol{b}_n$ is then calculated as

$$\boldsymbol{y}_{Q_n} = \boldsymbol{D}\mathrm{vec}\left\{\left[\boldsymbol{a}_{n-1}, \boldsymbol{a}_n\right]^T\right\} + \boldsymbol{w}''_n = \boldsymbol{D}\boldsymbol{b}_n + \boldsymbol{w}''_n, \tag{18}$$

where $\boldsymbol{D} \triangleq \left(\boldsymbol{I}_{N_r} \otimes [0_{M \times M} \quad \boldsymbol{I}_M]\right)\boldsymbol{K}_n \boldsymbol{Z}'_n$; $\boldsymbol{K}_n \in \mathbb{C}^{2N_rM \times 2N_rM}$, $\boldsymbol{Z}'_n \in \mathbb{C}^{2N_rM \times 2N_r}$ and $\boldsymbol{w}''_n \in \mathbb{C}^{2N_rM \times 1}$ represent the corresponding submatrices of $\boldsymbol{K}$, $\boldsymbol{Z}'$ and $\boldsymbol{w}'$, respectively. The prior probability $p\left(\boldsymbol{y}_{Q_n} | \boldsymbol{b}_n\right)$ is given by

$$p\left(\boldsymbol{y}_{Q_n} | \boldsymbol{b}_n\right) = \mathcal{CN}\left(\boldsymbol{y}_{Q_n};\, \boldsymbol{D}\hat{\boldsymbol{b}}_n, \boldsymbol{V}_n^{w''}\right), \tag{19}$$

where $\boldsymbol{V}_n^{w''}$ is the covariance of $\boldsymbol{w}''_n$ in Eq. (18).

3.2 Factor Graph Representation

With Eqs. (16) and (19), the posterior probability $p\left(\boldsymbol{b}, \boldsymbol{A} | \boldsymbol{y}_Q\right)$ is calculated as

$$p\left(\boldsymbol{b}, \boldsymbol{A} | \boldsymbol{y}_Q\right) = \frac{1}{p\left(\boldsymbol{y}_Q\right)} \prod_{n=1}^{N} p\left(\boldsymbol{y}_{Q_n} | \boldsymbol{b}_n\right) p\left(\boldsymbol{b}_n | \boldsymbol{b}_{n-1}, \boldsymbol{a}_n\right). \tag{20}$$

Defining $\boldsymbol{C} \triangleq (\boldsymbol{S} \odot \boldsymbol{G}\boldsymbol{X}) \in \mathbb{C}^{L \times N}$ in Eq. (10), the joint posterior probability can be further factored as

$$p\left(\boldsymbol{b}, \boldsymbol{A}, \boldsymbol{G}, \boldsymbol{H}, \boldsymbol{C} | \boldsymbol{y}_Q\right) = \\ \frac{1}{p\left(\boldsymbol{y}_Q\right)} p\left(\boldsymbol{y}_Q | \boldsymbol{b}\right) p(\boldsymbol{b} | \boldsymbol{A}) p(\boldsymbol{A} | \boldsymbol{H}, \boldsymbol{C}) p(\boldsymbol{C} | \boldsymbol{G}) p(\boldsymbol{H}) p(\boldsymbol{G}) = \\ \frac{1}{p\left(\boldsymbol{y}_Q\right)} \left(\prod_{n=1}^{N} p\left(\boldsymbol{y}_{Q_n} | \boldsymbol{b}_n\right) p\left(\boldsymbol{b}_n | \boldsymbol{b}_{n-1}, \boldsymbol{a}_n\right)\right) \left(\prod_{n_r=1}^{N_r} \prod_{n=1}^{N} p\left(a_{n_r,n} | \boldsymbol{h}_{n_r}, \boldsymbol{c}_n\right)\right) \\ \left(\prod_{l=1}^{L} \prod_{n=1}^{N} p\left(c_{l,n} | \boldsymbol{g}_l\right)\right) \left(\prod_{n_r=1}^{N_r} \prod_{l=1}^{L} p\left(h_{n_r,l}\right)\right) \left(\prod_{l=1}^{L} \prod_{n_t=1}^{N_t} p\left(g_{l,n_t}\right)\right), \tag{21}$$

where the first bracket is from the Markov chain and the rests belong to the bilinear structure described in Ref. [13]. Some involved probabilities are

$$p\left(a_{n_r,n} | \boldsymbol{h}_{n_r}, \boldsymbol{c}_n\right) = \delta\left(a_{n_r,n} - \boldsymbol{h}_{n_r}^T \boldsymbol{c}_n\right) \quad \text{and}$$
$$p\left(c_{l,n} | \boldsymbol{g}_l\right) = \delta\left(c_{l,n} - s_{l,n} \boldsymbol{g}_l^T \boldsymbol{x}_n\right), \tag{22}$$

where $\boldsymbol{h}_{n_r} \in \mathbb{C}^{L \times 1}$ and $\boldsymbol{g}_l \in \mathbb{C}^{N_t \times 1}$ are the n_r-th and l-th row of $\boldsymbol{H}$ and $\boldsymbol{G}$, respectively; $\boldsymbol{c}_n \in \mathbb{C}^{L \times 1}$ is the n-th column of $\boldsymbol{C}$. The second formula in Eq. (22) comes from the definition of $\boldsymbol{C}$, where $c_{l,n}$ and $s_{l,n}$ are the (l, n)-th element of $\boldsymbol{C}$ and $\boldsymbol{S}$, respectively, and $\boldsymbol{x}_n$ is the n-th column of $\boldsymbol{X}$. The factor graph representation of Eq. (21) is shown in Fig. 2, where the hollow circles and the solid squares represent the variable nodes and the factor nodes, respectively. The message passing in the first part will be described in the next subsection, while the illustration of the second part can be found in Ref. [8].

3.3 Message Passing in Markov Chain

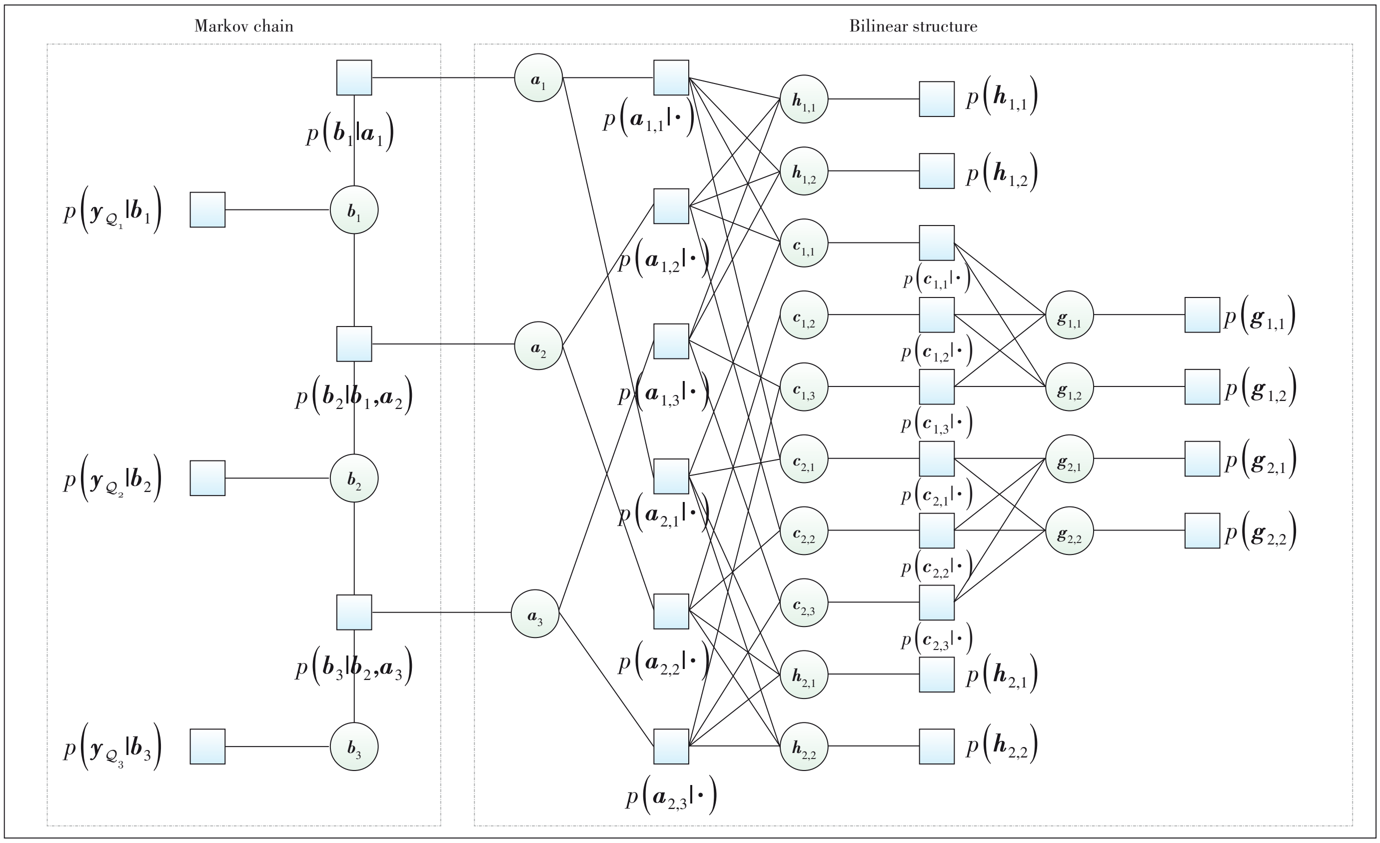
In the following, the approximate message passing in the Markov chain is derived according to the sum-product rule. The linear operator $\boldsymbol{K}$ described in Eq. (17) involves the covariance of unquantized signal $\boldsymbol{y}$ in Eq. (10), calculated as

$$\boldsymbol{C}_y = \boldsymbol{Z}'\mathrm{diag}\left(\mathrm{vec}\left\{\boldsymbol{V}^{A\,T}\right\}\right)\boldsymbol{Z}'^H + \sigma^2 \boldsymbol{F}'\boldsymbol{F}'^H. \tag{23}$$

During the message updates, the covariance of $\boldsymbol{A}$ at the i-th iteration is $\boldsymbol{V}^A(i)$, and Eq. (23) can be rewritten as

$$\boldsymbol{C}_y(i) = \boldsymbol{Z}'\mathrm{diag}\left(\mathrm{vec}\left\{\boldsymbol{V}^{A\,T}(i)\right\}\right)\boldsymbol{Z}'^H + \sigma^2 \boldsymbol{F}'\boldsymbol{F}'^H. \tag{24}$$

The covariance of the quantization noise at the i-th iteration is



▲**Figure 2. Factor graph representation of Eq. (21) for an example with $M = N_r = N_t = L = 2$ and $N = 3$**

$$C_{n_q}(i) = C_Q(i) - K(i) C_y(i) K(i)^H, \tag{25}$$

where $C_Q(i)$ is calculated from Ref. [16] as

$$C_Q(i) = \frac{2}{\pi}\left(\arcsin\left(K(i) C_y^{\Re}(i) K(i)^H\right) + j\arcsin\left(K(i) C_y^{\Im}(i) K(i)^H\right)\right). \tag{26}$$

3.3.1 Downward AMP

With Eq. (19), the message $m_{b_n \to p\left(b_{n+1}|b_n, a_{n+1}\right)}$ at the i-th iteration can be approximated as Gaussian distribution with the following mean and variance:

$$\hat{b}_n^{Down}(i) = \widehat{b'}_n^{Down}(i) + V'^{Down}_n(i) D^H R\left(y_{Q_n} - D\widehat{b'}_n^{Down}(i)\right), \tag{27}$$

$$V_n^{Down}(i) = V'^{Down}_n(i) - V'^{Down}_n(i) D^H R D V'^{Down}_n(i), \tag{28}$$

where $R \triangleq \left(V_n^{w''} + D V'^{Down}_n(i) D^H\right)^{-1}$. Moreover, with Eq. (15), the message from the factor node $p\left(b_{n+1}|b_n, a_{n+1}\right)$ to the variable node b_{n+1} at the i-th iteration is given by

$$m_{p\left(b_{n+1}|b_n, a_{n+1}\right) \to b_{n+1}} = \mathcal{CN}\left(b_{n+1}; \widehat{b'}_{n+1}^{Down}(i), V'^{Down}_{n+1}(i)\right), \tag{29}$$

$$\widehat{b'}_{n+1}^{Down}(i) = T\hat{b}_n^{Down}(i) + \mathrm{vec}\left\{\left[0_{N_r \times 1}, \hat{a}_{n+1}(i)\right]^T\right\}, \tag{30}$$

$$V'^{Down}_{n+1}(i) = T V_n^{Down}(i) T^H + \mathrm{diag}\left(\mathrm{vec}\left\{\left[0_{N_r \times 1}, v_{n+1}^{a}(i)\right]^T\right\}\right). \tag{31}$$

Especially, when $n = 1$, Eqs. (30) and (31) are reduced to

$$\widehat{b'}_1^{Down}(i) = \hat{b}_1(i) \quad \text{and} \quad V'^{Down}_1(i) = V_1^b(i). \tag{32}$$

3.3.2 Upward AMP

Similar to Eqs. (27) and (28), the message $m_{b_n \to p\left(b_n|b_{n-1}, a_n\right)}$ at the i-th iteration is approximated as Gaussian distribution with the following mean and variance

$$\hat{b}_n^{Up}(i) = \widehat{b'}_n^{Up}(i) + V'^{Up}_n(i) D^H R\left(y_{Q_n} - D\widehat{b'}_n^{Up}(i)\right), \tag{33}$$

$$V_n^{Up}(i) = V'^{Up}_n(i) - V'^{Up}_n(i) D^H R D V'^{Up}_n(i), \tag{34}$$

where $R \triangleq \left(V_n^{w''} + D V'^{Up}_n(i) D^H\right)^{-1}$. Moreover, with the inverse of Eq. (15), the message from the factor node $p\left(b_n|b_{n-1}, a_n\right)$ to the variable node b_{n-1} at the i-th iteration is given by

$$m_{p\left(\boldsymbol{b}_n|\boldsymbol{b}_{n-1},\boldsymbol{a}_n\right)\rightarrow \boldsymbol{b}_{n-1}} = \mathcal{CN}\left(\boldsymbol{b}_{n-1};\widehat{\boldsymbol{b}'}^{Up}_{n-1}(i),\boldsymbol{V'}^{Up}_{n-1}(i)\right), \tag{35}$$

where

$$\widehat{\boldsymbol{b}'}^{Up}_{n-1}(i) = \boldsymbol{T}^{\dagger}\hat{\boldsymbol{b}}^{Up}_{n}(i) + \mathrm{vec}\left\{\left[\hat{\boldsymbol{a}}_{n-2}(i),0_{N_r\times 1}\right]^T\right\}, \tag{36}$$

$$\boldsymbol{V'}^{Up}_{n-1}(i) = \boldsymbol{T}^{\dagger}\boldsymbol{V}^{Up}_{n}(i)\left(\boldsymbol{T}^H\right)^{\dagger} + \mathrm{diag}\left(\mathrm{vec}\left\{\left[\boldsymbol{v}^{a}_{n-2}(i),0_{N_r\times 1}\right]^T\right\}\right), \tag{37}$$

where $(\cdot)^{\dagger}$ denotes the operation of pseudo-inverse. Especially, when $n = N$, Eqs. (36) and (37) are reduced to

$$\widehat{\boldsymbol{b}'}^{Up}_{N}(i) = \hat{\boldsymbol{b}}_N(i) \quad \text{and} \quad \boldsymbol{V'}^{Up}_{N}(i) = \boldsymbol{V}^{b}_{N}(i). \tag{38}$$

3.3.3 Backward AMP

Combining with $m_{\boldsymbol{b}_{n-1}\rightarrow p\left(\boldsymbol{b}_n|\boldsymbol{b}_{n-1},\boldsymbol{a}_n\right)}$ and $m_{\boldsymbol{b}_n\rightarrow p\left(\boldsymbol{b}_n|\boldsymbol{b}_{n-1},\boldsymbol{a}_n\right)}$, the message from the factor node $p\left(\boldsymbol{b}_n|\boldsymbol{b}_{n-1},\boldsymbol{a}_n\right)$ back to the variable node $\boldsymbol{a}_n$ is

$$m_{p\left(\boldsymbol{b}_n|\boldsymbol{b}_{n-1},\boldsymbol{a}_n\right)\rightarrow \boldsymbol{a}_n} = \int_{\boldsymbol{a}_{n-1}}\mathcal{CN}\left(\boldsymbol{b}_n;\boldsymbol{T}\hat{\boldsymbol{b}}^{Down}_{n}(i),\boldsymbol{T}\boldsymbol{V}^{Down}_{n}(i)\boldsymbol{T}^H\right)\mathcal{CN}\left(\boldsymbol{b}_n;\hat{\boldsymbol{b}}^{Up}_{n}(i),\boldsymbol{V}^{Up}_{n}(i)\right), \tag{39}$$

where the multiplication of two Gaussian functions is another Gaussian function[3].

4 Further Discussions

The overall Markovian cascaded channel estimation algorithm is presented in Algorithm 1. In lines 3 – 6, the estimates of the each mean and variance in the matrix product $\boldsymbol{A} = \boldsymbol{HC}$ are calculated. The message passing in the Markov chain is illustrated in lines 7 – 9. Subsequently, the residual $q_{n_r,n}$ and the inverse residual variance $u_{n_r,n}$ are calculated in lines 10 – 11. In lines 12 – 13, these residual terms are used to compute $v'^{h}_{n_r,l}$ and $\hat{h}'_{n_r,l}$, which can be interpreted as an observation of $\hat{h}_{n_r,l}$ under an AWGN channel with zero mean and a variance of $v'^{h}_{n_r,l}$. The posterior means and variances of each elements in $\boldsymbol{H}$ are estimated in lines 14 and 15. Similarly, the posterior means and variances of each elements in the auxiliary matrix $\boldsymbol{C}$ are estimated in lines 16 – 21; the same is performed for $\boldsymbol{G}$ in lines 22 – 27.

In this paper, a damping method is applied to improve the robustness of the proposed algorithm. Specifically, in each iteration the posterior means and variances of $\boldsymbol{H}$ (lines 14 – 15), $\boldsymbol{C}$ (lines 20 – 21) and $\boldsymbol{G}$ (lines 26 – 27) are updated by using a linear combination of the current and previous updates. For example, the updates of the posterior mean $\hat{h}_{n_r,l}$ and variance $v^{h}_{n_r,l}$ in lines 14 – 15 of Algorithm 1 are replaced by

$$\hat{h}_{n_r,l}(i+1) = (1-\beta)\hat{h}_{n_r,l}(i) + \beta\hat{h}_{n_r,l}(i+1), \tag{40}$$

$$v^{h}_{n_r,l}(i+1) = (1-\beta)v^{h}_{n_r,l}(i) + \beta v^{h}_{n_r,l}(i+1), \tag{41}$$

where $\beta \in [0,1]$ is the damping factor. In our work, β is chosen within $[0.2,0.4]$.

We now give a brief discussion on the computational complexity of the proposed algorithm. Note that the total algorithm is separated into two parts, the Markov chain and the bilinear structure. We thus sketch the respective complexity as follows. First, the complexity of the Markov chain is dominated by basic matrix multiplications in Eqs. (27), (28), (33), (34) and (39), requiring $\mathcal{O}\left(\left(MN_r\right)^3\right)$ flops per iteration. Second, the complexity of the bilinear structure is dominated by the calculations in Lines 3 – 5, 10 – 11, 12 – 13, 16 – 17 and 22 – 25, requiring $\mathcal{O}\left(\left(N_r + N_t\right)LN\right)$ flops per iteration. Finally, the complexity of the proposed algorithm is $I_{\max}\mathcal{O}\left(\left(MN_r\right)^3 + \left(N_r + N_t\right)LN\right)$, where $I_{\max}$ is the maximum number of the iterations.

Algorithm 1. Markovian cascaded channel estimation

Input: $\boldsymbol{y}_{\mathcal{Q}}$, $\boldsymbol{x}$, $\boldsymbol{S}$, σ^2, prior distributions $p(\boldsymbol{G})$, $p(\boldsymbol{H})$

Output: $\hat{\boldsymbol{G}}$, $\hat{\boldsymbol{H}}$

1: **Initialization:** $\forall n_r,l,n$: $\hat{h}_{n_r,l}(1)$, $v^{c}_{l,n}(1)$, $v^{h}_{n_r,l}(1)$, $\hat{c}_{l,n}(1)$, $\eta_{l,n}(0) = 0$, $u_{n_r,n}(0) = 0$;

2: **for** $\boldsymbol{i}$**=1,…,** $\boldsymbol{I}_{\max}$ **do**

3: $\forall n_r,n$: $v'^{p}_{n_r,n}(i) = \sum_l \left|\hat{h}_{n_r,l}(i)\right|^2 v^{c}_{l,n}(i) + v^{h}_{n_r,l}(i)\left|\hat{c}_{l,n}(i)\right|^2$

4: $\forall n_r,n$: $\hat{p}'_{n_r,n}(i) = \sum_l \hat{h}_{n_r,l}(i)\hat{c}_{l,n}(i)$

5: $\forall n_r,n$: $v^{p}_{n_r,n}(i) = v'^{p}_{n_r,n}(i) + \sum_l v^{h}_{n_r,l}(i)v^{c}_{l,n}(i)$

6: $\forall n_r,n$: $\hat{p}_{n_r,n}(i) = \hat{p}'_{n_r,n}(i) - u_{n_r,n}(i-1)v'^{p}_{n_r,n}(i)$

7: Calculate the adaptive linear operator and quantization noise via Eqs. (24) and (25) using $v^{p}_{n_r,n}(i)$ and $\hat{p}_{n_r,n}(i)$

8: Calculate the downward and upward messages at teach check node in the Markov chain via Eqs. (27), (28), (33), and (34)

9: Calculate the backward messages via Eq. (39) and ob-

3. $\mathcal{CN}\left(\boldsymbol{x};\boldsymbol{m}_1,\boldsymbol{V}_1\right)\mathcal{CN}\left(\boldsymbol{x};\boldsymbol{m}_2,\boldsymbol{V}_2\right) \propto \mathcal{CN}(\boldsymbol{x};\boldsymbol{m},\boldsymbol{V})$, where $\boldsymbol{V} = \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1}$ and $\boldsymbol{m} = \boldsymbol{V}\left(\boldsymbol{V}_1^{-1}\boldsymbol{m}_1 + \boldsymbol{V}_2^{-1}\boldsymbol{m}_2\right)$.

tain $v^a_{n_r,n}(i)$ and $\hat{a}_{n_r,n}(i)$

$\forall n_r,n$: $q_{n_r,n}(i) = \left(1 - \frac{v^a_{n_r,n}(i)}{v^p_{n_r,n}(i)}\right)/v^p_{n_r,n}(i)$

$\forall n_r,n$: $u_{n_r,n}(i) = \left(\hat{a}_{n_r,n}(i) - \hat{p}_{n_r,n}(i)\right)/v^p_{n_r,n}(i)$

Update $\boldsymbol{H}$:

$\forall n_r,l$: $v'^h_{n_r,l}(i) = 1/\left(\sum_n q_{n_r,n}(i)\left|\hat{c}_{l,n}(i)\right|^2\right)$

$\forall n_r,l$: $\hat{h}'_{n_r,l}(i) = \hat{h}_{n_r,l}(i)\left(1 - v'^h_{n_r,l}(i)\sum_n q_{n_r,n}(i)v^c_{l,n}(i)\right) + v'^h_{n_r,l}(i)\sum_n u_{n_r,n}(i)\hat{c}^*_{l,n}(i)$

$\forall n_r,l$: $\hat{h}_{n_r,l}(i+1) = E\{h_{n_r,l}|\hat{h}'_{n_r,l}(i), v'^h_{n_r,l}(i)\}$

$\forall n_r,l$: $v^h_{n_r,l}(i+1) = Var\{h_{n_r,l}|\hat{h}'_{n_r,l}(i), v'^h_{n_r,l}(i)\}$

Update $\boldsymbol{C}$:

$\forall l,n$: $v'^c_{l,n}(i) = s_{l,n}/\left(\sum_{n_r} q_{n_r,n}(i)\left|\hat{h}_{n_r,l}(i)\right|^2\right)$

$\forall l,n$: $\hat{c}'_{l,n}(i) = s_{l,n}\left(\hat{c}_{l,n}(i)\left(1 - v'^c_{l,n}(i)\sum_{n_r} q_{n_r,n}(i)v^h_{n_r,l}(i)\right) + v'^c_{l,n}(i)\sum_{n_r} u_{n_r,n}(i)\hat{h}^*_{n_r,l}(i)\right)$

$\forall l,n$: $v''^c_{l,n}(i) = \sum_{n_t} v^g_{l,n_t}(i)\left|x_{n_t,n}\right|^2$

$\forall l,n$: $\hat{c}''_{l,n}(i) = \sum_{n_t} \hat{g}_{l,n_t}(i)x_{n_t,n} - \eta_{l,n}(i-1)\sum_{n_t} v^g_{l,n_t}(i)\left|x_{n_t,n}\right|^2$

$\forall l,n$: $\hat{c}_{l,n}(i) = E\{c_{l,n}|\hat{c}'_{l,n}(i), v'^c_{l,n}(i), \hat{c}''_{l,n}(i), v''^c_{l,n}(i)\}$

$\forall l,n$: $v^c_{l,n}(i) = Var\{c_{l,n}|\hat{c}'_{l,n}(i), v'^c_{l,n}(i), \hat{c}''_{l,n}(i), v''^c_{l,n}(i)\}$

Update $\boldsymbol{G}$:

$\forall l,n$: $\eta_{l,n}(i) = s_{l,n}\left(\hat{c}_{l,n}(i) - \hat{c}''_{l,n}(i)\right)/v''^c_{l,n}(i)$

$\forall l,n$: $v^\eta_{l,n}(i) = s_{l,n}\left(1 - v^c_{l,n}(i)/v''^c_{l,n}(i)\right)/v''^c_{l,n}(i)$

$\forall l,n_t$: $v'^g_{l,n_t}(i) = 1/\left(\sum_n v^\eta_{l,n}(i)\left|x_{n_t,n}\right|^2\right)$

$\forall l,n_t$: $\hat{g}'_{l,n_t}(i) = \hat{g}_{l,n_t}(i) + v'^g_{l,n_t}(i)\sum_n \eta_{l,n}(i)x^*_{n_t,n}$

$\forall l,n_t$: $\hat{g}_{l,n_t}(i+1) = E\{g_{l,n_t}|\hat{g}'_{l,n_t}(i), v'^g_{l,n_t}(i)\}$

$\forall l,n_t$: $v^g_{l,n_t}(i+1) = Var\{g_{l,n_t}|\hat{g}'_{l,n_t}(i), v'^g_{l,n_t}(i)\}$

if stopping criterion is met **then**

break

end if

end for

5 Numerical Results

In this section, the uplink of a RIS aided 1-bit massive MIMO system with $N_r = 64$, $N_t = 8$ and $L = 128$ is considered. The $m(t)$ and $p(t)$ are normalized root-raised-cosine (RRC) filters with a roll-off factor of 0.8. The channel is assumed to experience Rayleigh block fading. The simulation results presented here are obtained by averaging over 100 independent realizations of the channel matrices, noise and pilots. The signal-to-noise ratio (SNR) is defined as $10\log\left(\frac{\rho L N_t}{\sigma^2}\right)$, where ρ is the sparsity level of $\boldsymbol{S}$[4] and is set as 0.3. The pilot length is 200. For the correlated channel, the channel matrices are modeled as

$$\boldsymbol{H} = \boldsymbol{R}_{rr}^{\frac{1}{2}}\boldsymbol{H}'\boldsymbol{R}_{r}^{\frac{1}{2}} \quad \text{and} \quad \boldsymbol{G} = \boldsymbol{G}'\boldsymbol{R}_{rl}^{\frac{1}{2}}, \tag{42}$$

where the elements of $\boldsymbol{H}'$ and $\boldsymbol{G}'$ are independent and identically distributed (i. i. d.) complex Gaussian random variables with zero mean and unit variance. The matrix $\boldsymbol{R}_r$ denotes the receive correlation matrix with the following form:

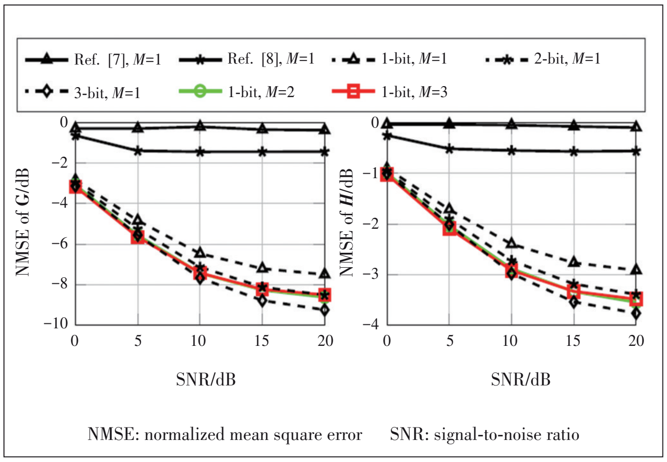
$$\boldsymbol{R}_r = \begin{bmatrix} 1 & \rho & \cdots & \rho^{(N_r-1)} \\ \rho^* & 1 & \cdots & \rho^{(N_r-2)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{*(N_r-1)} & \rho^{*(N_r-2)} & \cdots & 1 \end{bmatrix}, \tag{43}$$

where ρ is the correlation index of neighboring antennas at the BS and set as $0.4 + 0.3j$. ($|\rho| = 0$ represents an uncorrelated scenario and $|\rho| = 1$ implies a fully correlated scenario.) The matrices $\boldsymbol{R}_{rr}$ and $\boldsymbol{R}_{rl}$ have the same form as Eq. (43), where ρ is set as $0.2 + 0.5j$ and $0.1 + 0.2j$ at the RIS, respectively.

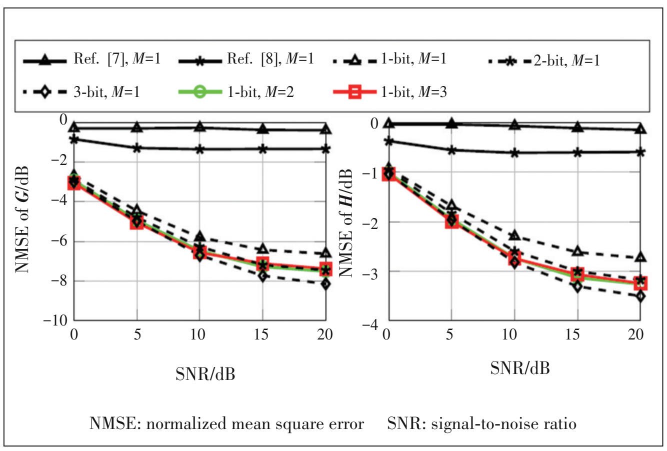
Figs. 3 and 4 show the NMSEs of $\boldsymbol{G}$ and $\boldsymbol{H}$ under uncorrelated and correlated channels, respectively. For the multi-bit systems sampling at the Nyquist rate, the calculation of the posterior probability $p\left(\boldsymbol{A}|\boldsymbol{y}_q\right)$ is referred to as Eqs. (23) and (24) in Ref. [17]. From the figures, it can be seen that our proposed 1-bit multi-fold oversampled systems outperform 1-bit systems sampling at the Nyquist rate, and can even approach the performance of 2-bit systems sampling at the Nyquist rate. Another observation is that after 2-fold oversampling, the performance of 1-bit systems goes into the saturation field, and no further gain can be achieved. This reveals the performance limit of the proposed method. Furthermore, we have also compared the performance of Refs. [7] and [8] as references, where the resolutions of the ADCs are changed to 1-bit. From the results, the performance of the references is worse than that of our proposed method, since they do not consider the impact of 1-bit quantization and the advantages of oversampling.

The advantage of 1-bit ADCs is that they do not require au-

4. In simulations, the phases of the RIS are set to zeros, and the matrix $\boldsymbol{S}$ is generated as a 0 – 1 random matrix.



▲**Figure 3. NMSE of *G* and *H* comparisons between the multi-bit systems with Nyquist rate (*M*=1) and the 1-bit systems with multi-fold oversampling under uncorrelated channel**



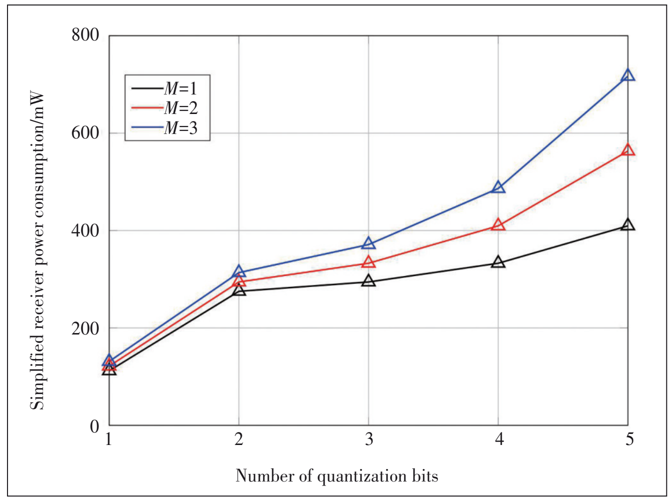
▲**Figure 4. NMSE of *G* and *H* comparisons between the multi-bit systems with Nyquist rate (*M*=1) and the 1-bit systems with multi-fold oversampling under correlated channel**

tomatic gain control (AGC), and can be replaced by simple LAs. Fig. 5 shows the simplified[5] receiver power consumption as a function of the quantization bits and the oversampling factor M, which is calculated as

$$P_{\text{simplified}} = 2N_{\text{r}}\left(cP_{\text{AGC}} + (1-c)P_{\text{LA}} + \text{FOM} \times Mf_{\text{Nyquist}}2^{q}\right), \quad (44)$$

where P_{AGC} and P_{LA} denote the power consumption of AGC and LA, respectively; q is the quantization bits and f_{Nyquist} is the Nyquist-sampling rate; c is chosen as 0 for 1-bit system and 1 for systems with multi-bit. Numerical parameters[18] are $P_{\text{AGC}} = 2$ mW, $P_{\text{LA}} = 0.8$ mW, $f_{\text{Nyquist}} = 2.5$ GHz and figures-of-merit (FOM) = 15 fJ. From the results, it can be seen that 1-bit systems consume much less power than multi-bit systems with either the Nyquist rate or the oversampling rate.

5. The power consumption of shared components in different systems is neglected.



▲**Figure 5. Simplified power consumption at the receiver as a function of the number of quantization bits**

6 Conclusions

In this work, we propose RIS aided 1-bit massive MIMO systems with oversampling at the receiver. The aim of oversampling is to compensate for the performance loss due to the coarse quantization. A Markovian cascaded channel estimation algorithm is developed for such systems. Simulation results have shown good performance gains of the proposed oversampled system, which can achieve the same performance of the corresponding 2-bit system sampling at the Nyquist rate while consuming less power at the receiver.

References

[1] LARSSON E G, EDFORS O, TUFVESSON F, et al. Massive MIMO for next generation wireless systems [J]. IEEE communications magazine, 2014, 52(2): 186 – 195. DOI: 10.1109/MCOM.2014.6736761
[2] LIASKOS C, NIE S, TSIOLIARIDOU A, et al. A new wireless communication paradigm through software-controlled metasurfaces [J]. IEEE communications magazine, 2018, 56(9): 162 – 169. DOI: 10.1109/MCOM.2018.1700659
[3] JAMALI V, TULINO A M, FISCHER G, et al. Intelligent surface-aided transmitter architectures for millimeter-wave ultra massive MIMO systems [J]. IEEE open journal of the communications society, 2020, 2: 144 – 167. DOI: 10.1109/OJCOMS.2020.3048063
[4] TAHA A, ALRABEIAH M, ALKHATEEB A. Enabling large intelligent surfaces with compressive sensing and deep learning [J]. IEEE access, 2021, 9: 44304 – 44321. DOI: 10.1109/ACCESS.2021.3064073
[5] ARDAH K, GHEREKHLOO S, DE ALMEIDA A L F, et al. TRICE: A channel estimation framework for RIS-aided millimeter-wave MIMO systems [J]. IEEE signal processing letters, 2021, 28: 513 – 517. DOI: 10.1109/

LSP.2021.3059363

[6] HE Z Q, YUAN X J. Cascaded channel estimation for large intelligent metasurface assisted massive MIMO [J]. IEEE wireless communications letters, 2020, 9 (2): 210 – 214. DOI: 10.1109/LWC.2019.2948632

[7] WANG P L, FANG J, DUAN H P, et al. Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems [J]. IEEE signal processing letters, 2020, 27: 905 – 909. DOI: 10.1109/LSP.2020.2998357

[8] HE Z Q, LIU H, YUAN X J, et al. Semi-blind cascaded channel estimation for reconfigurable intelligent surface aided massive MIMO [EB/OL]. [2021-11-15]. https://arxiv.org/abs/2101.07315

[9] WALDEN R H. Analog-to-digital converter survey and analysis [J]. IEEE journal on selected areas in communications, 1999, 17(4): 539 – 550. DOI: 10.1109/49.761034

[10] MO J H, HEATH R W. Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information [J]. IEEE transactions on signal processing, 2015, 63(20): 5498 – 5512. DOI: 10.1109/TSP.2015.2455527

[11] LI Y Z, TAO C, SECO-GRANADOS G, et al. Channel estimation and performance analysis of one-bit massive MIMO systems [J]. IEEE transactions on signal processing, 2017, 65(15): 4075 – 4089. DOI: 10.1109/TSP.2017.2706179

[12] SHAO Z C, DE LAMARE R C, LANDAU L T N. Iterative detection and decoding for large-scale multiple-antenna systems with 1-bit ADCs [J]. IEEE wireless communications letters, 2018, 7(3): 476 – 479. DOI: 10.1109/LWC.2017.2787159

[13] SHAO Z C, LANDAU L T N, DE LAMARE R C. Channel estimation for large-scale multiple-antenna systems using 1-bit ADCs and oversampling [J]. IEEE access, 2020, 8: 85243 – 85256. DOI: 10.1109/ACCESS.2020.2992246

[14] SHAO Z C, LANDAU L T N, DE LAMARE R C. Dynamic oversampling for 1-bit ADCs in large-scale multiple-antenna systems [J]. IEEE transactions on communications, 2021, 69(5): 3423 – 3435. DOI: 10.1109/TCOMM.2021.3059303

[15] BUSSGANG J J. Crosscorrelation functions of amplitude-distorted Gaussian signals [R]. Cambridge, USA: Massachusetts Institute of Technology, 1952

[16] JACOVITTI G, NERI A. Estimation of the autocorrelation function of complex Gaussian stationary processes by amplitude clipped signals [J]. IEEE transactions on information theory, 1994, 40(1): 239 – 245. DOI: 10.1109/18.272490

[17] WEN C K, WANG C J, JIN S, et al. Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs [J]. IEEE transactions on signal processing, 2016, 64(10): 2541 – 2556. DOI: 10.1109/TSP.2015.2508786

[18] ROTH K, NOSSEK J A. Achievable rate and energy efficiency of hybrid and digital beamforming receivers with low resolution ADC [J]. IEEE journal on selected areas in communications, 2017, 35(9): 2056 – 2068. DOI: 10.1109/JSAC.2017.2720398

Biographies

SHAO Zhichao received the B.S. degree in information engineering from Xidian University, China in 2012, the M.S. degree in electrical engineering from Technische Universität Dresden, Germany in 2016, and the Ph.D. degree in electrical engineering from the Pontifical Catholic University of Rio de Janeiro, Brazil in 2020. Currently, he is a post-doctoral researcher with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China (UESTC). His research interests lie in communications and signal processing.

YAN Wenjing received the B.S. degree from Chongqing University, China in 2018, and the M.S. degree from University of Electronic Science and Technology of China (UESTC) in 2021. She is currently pursuing the Ph.D. degree with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology (HKUST), China. Her contribution to the work in this paper was done when she was with UESTC.

YUAN Xiaojun (xjyuan@uestc.edu.cn) received the Ph.D. degree in electrical engineering from the City University of Hong Kong, China in 2009. From 2009 to 2011, he was a research fellow with the Department of Electronic Engineering, City University of Hong Kong. He was also a visiting scholar with the Department of Electrical Engineering, University of Hawaii at Manoa, USA in spring and summer 2009, as well as in the same period of 2010. From 2011 to 2014, he was a research assistant professor with the Institute of Network Coding, The Chinese University of Hong Kong, China. From 2014 to 2017, he was an assistant professor with the School of Information Science and Technology, Shanghai Tech University, China. He is currently a State-Specially-Recruited Professor with the Center for Intelligent Networking and Communications, University of Electronic Science and Technology of China. His research interests include signal processing, machine learning, and wireless communications, including but not limited to multi-antenna and cooperative communications, sparse and structured signal recovery, Bayesian approximate inference, and network coding. He has published more than 180 peer-reviewed research articles in the leading international journals and conferences in the related areas. He has served on a number of technical programs for international conferences. He was a co-recipient of the Best Paper Award of IEEE International Conference on Communications (ICC) 2014 and a co-recipient of the Best Journal Paper Award of IEEE Technical Committee on Green Communications and Computing (TCGCC) 2017. He has been an editor of *IEEE Transactions on Communications* since 2017 and *IEEE Transactions on Wireless Communications* since 2018.