FASURA: A Scheme for Quasi-Static Massive MIMO Unsourced Random Access Channels

Michail Gkagkos, Krishna R. Narayanan, Jean-Francois Chamberland, Costas N. Georghiades
Department of Electrical and Computer Engineering, Texas A&M University

Abstract—This article considers the massive MIMO unsourced random access problem on a quasi-static Rayleigh fading channel. Given a fixed message length and a prescribed number of channel uses, the objective is to construct a coding scheme that minimizes the energy-per-bit subject to a fixed probability of error. The proposed scheme differs from other state-of-the-art schemes in that it blends activity detection, single-user coding, pilot-aided and temporary decisions-aided iterative channel estimation and decoding, minimum-mean squared error (MMSE) estimation, and successive interference cancellation (SIC). We show that an appropriate combination of these ideas can substantially outperform state-of-the-art coding schemes when the number of active users is more than 100, making this the best performing scheme known for this regime.

Index Terms—Unsourced Random Access, Massive Multi-User MIMO, random pilots and spreading, polar code.

I. INTRODUCTION

A new perspective on random access was proposed in [1] to accommodate the type of traffic generated by unattended devices. The ensuing model, coined *unsourced random access* (URA) in [2], has been widely adopted as a common task framework for emerging IoT wireless networks. A motivation behind this model is that, as the prospective user population grows, the assignment of spectral resources based on queue lengths and channel conditions becomes impractical. This is especially true when devices sporadically transmit short packets. A pragmatic alternative is to have all active devices share a same codebook. This way, the system can operate irrespective of the total user population, and target the number of active devices instead. In such situations, the decoder aims to recover the set of messages regardless of user identities.

Many conceptual approaches and candidate solutions tailored to Gaussian URA have appeared in recent years, with a majority of them confined to receivers equipped with a single antenna [2]–[8]. Many recent developments by the URA research community center on practical aspects of wireless communications, such as fading and MIMO models. For instance, Andreev et al. in [9] examine a quasi-static fading URA channel, and they propose a scheme that blends polar codes and expectation-maximization (EM) clustering. Fengler et al. in [10] combine a non-Bayesian sparse recovery algorithm with an outer code aimed at message disambiguation to facilitate MIMO URA; the performance of this algorithm improved as shown in [11]. Liu and Wang propose a receiver

that merges simultaneous orthogonal matching pursuit (S-OMP) and blind channel estimation [12]. In [13] and [14], authors consider only the user activity detection in MIMO URA problem, and propose a deep unfolding based algorithm and model-based algorithms respectively. Decurninge et al. offer an efficient solution based on a tensor construction, with good performance in some regimes [15]. On the other hand, Fengler et al. [16] construct a scheme that relies on preamble-selected pilot sequences to estimate channels and the formulate this problem as a multiple measurement vector (MMV) problem. Then polar-list decoder is employed to recover the modulated bits. Among published work, the scheme put forth by Ahmadi et al. offers the best performance [17]. It is based on an interesting construction of pilot sequences using set of orthogonal pilots. A polar code is used to protect the information bits. Other works in the MIMO URA setting includes [18], [19] [20].

Our proposed scheme extends the ideas from [2], [6] to the MIMO URA setting. Since our communication scheme incorporates random spreading, and it is designed for fading URA, we call it FASURA (fading spread unsourced random access). Like many previous contributions, it divides messages into two parts. The preamble is dedicated to the selection of operational parameters, whereas the payload is encoded using a single-user code. This mimics the architectures of [6], [8] for SISO systems, and [16] in the MIMO case. However, FASURA spreads every coded symbol before transmission to enable the computation of a linear minimum-mean squared error (LMMSE) estimate of the coded symbols based on the active spreading sequences and mitigate interference during the decoding process. We propose a new channel estimation step, which we call noisy pilot channel estimation (NOPICE). Therein, the channel is estimated using both pilots and preliminary decisions about coded symbols. This estimation technique is embedded in an iterative loop and, when parameters are picked judiciously, progressively leads to better channel estimates. With these two innovations, the proposed architecture outperforms the scheme put forth by Ahmadi et al. [17] over a range of parameters. For example, for 800 active devices, the difference in E_b/N_0 is more than 2.5 dB. This makes FASURA the state-of-the-art for MIMO URA in certain practical regimes.

Notation: Throughout, $\mathbb{Z}_+$ and $\mathbb{C}$ refer to the non-negative integers and complex numbers, respectively. We use $[n]$ to denote $\{1, 2, \ldots, n\}$. We employ boldface lowercase **a** and boldface uppercase letters **A** to indicate vectors and matrices.

Sets are labeled with calligraphic letters, e.g., $\mathcal{A}$. We also adopt a programming-style notation with $\mathbf{A}[:,t]$ and $\mathbf{A}[k,:]$ representing the tth column and kth row of $\mathbf{A}$.

II. System Model

Consider the uplink of a wireless network with a total of K_{tot} devices out of which K devices are active ($K \ll K_{\text{tot}}$). For convenience, we label active users using integers from the set $\mathcal{K} = [K]$. All the active users share the same n complex channel uses, and each of them wishes to transmit a B-bit message to a common destination. Every user is equipped with a single antenna (the term MIMO is adopted by the URA literature for multiple active users), whereas the access point features M antennas. We consider a quasi-static Rayleigh fading model whereby channel coefficients remain fixed during the transmission of an entire codeword. Furthermore, the antennas at the access point are located far enough from one another as to create independent channel realizations.

Let $\mathbf{m}_k$ be B-bit message of user k, and $\mathbf{x}_k = \mathcal{E}(\mathbf{m}_k) \in \mathbb{C}^n$ be the encoded and modulated signal (input to the channel) corresponds to the message $\mathbf{m}_k$. Then, the received signal takes the form

$$\mathbf{Y} = \sum_{k=1}^{K_{\text{tot}}} \delta_k \mathbf{x}(\mathbf{m}_k)\mathbf{h}_k^{\mathsf{T}} + \mathbf{Z} = \sum_{k \in \mathcal{K}} \mathbf{x}(\mathbf{m}_k)\mathbf{h}_k^{\mathsf{T}} + \mathbf{Z}, \quad (1)$$

where δ_k is an indicator function that takes value one when user k is active, and zero otherwise. Vector $\mathbf{h}_k \in \mathbb{C}^M$ captures the channel coefficients between user k and the M receive antennas. The elements of $\mathbf{h}_k$ are independent complex Gaussian random variables with mean zero and unit variance, as prescribed by Rayleigh fading. Additive noise component $\mathbf{Z} \in \mathbb{C}^{n \times M}$ is a matrix with i.i.d. entries, each drawn from a circularly symmetric complex Gaussian distribution $\mathcal{CN}(0, \sigma_z^2)$. Furthermore, every transmit signal must satisfy power constraint $\|\mathbf{x}(\mathbf{m}_k)\|^2 \leq nP$; for simplicity, we assume that $P = 1$. As a result, we can define the energy per bit to noise power spectral density ratio of the system by

$$\frac{E_{\mathbf{b}}}{N_0} = \frac{\|\mathbf{x}(\mathbf{m}_k)\|^2}{B\sigma_z^2}.$$

At the access point, the decoder aims to produce a set $\hat{\mathcal{K}}$ of candidate messages with cardinality at most K. The system performance is evaluated in terms of the probability of missed detection P_{md} and probability of false alarm P_{fa},

$$\mathrm{P}_{\text{md}} = \frac{\mathbb{E}[n_{\text{ms}}]}{K} \qquad \mathrm{P}_{\text{fa}} = \mathbb{E}\left[\frac{n_{\text{fa}}}{\hat{K}}\right]$$

where n_{ms} and n_{fa} denote the number of misses and false alarms, respectively. Variable $\hat{K}$ represents the number of messages declared by the decoder. We note that the expectation is taken over the randomness of the fading process, the noise process, and the relevant algorithmic components. We define the probability of error P_{e} to be the sum of the two types of probability introduced above, $\mathrm{P}_{\text{e}} = \mathrm{P}_{\text{md}} + \mathrm{P}_{\text{fa}}$. For fixed parameters B, n, K, M, and target error probability ε, our objective is to construct a communication scheme that minimizes E_{b}/N_0 while also satisfying the constraint $\mathrm{P}_{\text{e}} \leq \varepsilon$.

III. FASURA

We proceed with the description of our proposed scheme, FASURA. We begin with an overview of the encoding process, and then we discuss our decoding strategy.

A. Encoder

Active user k aims to transmit message $\mathbf{m}_k$ to the destination. Following the URA philosophy, encoding is dissociated from the identity of the user, and the sent signal only depends on the content of $\mathbf{m}_k$. As such, we describe the encoding process for a generic message $\mathbf{m}$, with the understanding that the steps are repeated by all the active devices. The encoder first splits the message into two parts $\mathbf{m}_f$ and $\mathbf{m}_s$ of lengths B_f bits and B_s bits, respectively. The two parts $\mathbf{m}_f$ and $\mathbf{m}_s$ are encoded as follows.

1) Encoding $\mathbf{m}_f$: Let $\mathcal{A} = \{\mathbf{A}_t\}_{t=1}^T$ denote a collection of $T = n_c/2$ spreading matrices indexed by time t. The coefficients of matrix $\mathbf{A}_t$, are distributed as complex Gaussian random variables, i.e. $a_{t,j} \sim \mathcal{CN}(0,1)$. The columns of matrix $\mathbf{A}_t$ can be viewed as spreading sequences of length L, which can be utilized at time t. The columns of $\mathbf{A}_t$ are normalized to have energy of L. There are $J = 2^{B_f}$ possible spreading sequences attached to every time instant. Likewise, let $\mathbf{P} \in \left\{ \pm \frac{1}{\sqrt{2}} \pm \frac{j}{\sqrt{2}} \right\}^{n_p \times J}$ be a matrix whose columns are possible pilot sequences. The selection of spreading sequences and pilots is performed using $\phi : \{0,1\}^{B_f} \to [J]$, a bijection that maps $\mathbf{m}_f$ to an index in $[J]$. Thus, when the first part of the message to be sent is $\mathbf{m}_f$, then the corresponding user employs the series of spreading sequences $\{\mathbf{A}_t[:,\phi(\mathbf{m}_f)]\}_{t=1}^T$, along with pilot sequence $\mathbf{P}[:,\phi(\mathbf{m}_f)]$. Thus, the overall encoding function for $\mathbf{m}_f$ can be summarized as $g(\mathbf{m}_f) \to \{\mathbf{A}_t[:,\phi(\mathbf{m}_f)]\}_{t=1}^T \cup \{\mathbf{P}[:,\phi(\mathbf{m}_f)]\}$.

An important aspect of the proposed scheme is that $J \ll K_{\text{tot}}$, which is critical in limiting the complexity at the decoder. We emphasize that there are no guarantees that every active user will select a distinct sequence from an orthogonal set. Rather, active users pick sequences at random from a non-orthogonal set. In fact, the URA framework makes it impossible to hand pick sequence since two devices with a common message will ultimately transmit the same signal.

2) Encoding $\mathbf{m}_s$: The second part of the message, namely $\mathbf{m}_s$, is first encoded using a cyclic redundancy check (CRC) code. The resulting codeword of length $B_c = B_s + B_{crc}$ then acts as input to an encoder for a (n_c, B_c) polar code with $n_c - B_c$ frozen bit positions. We construct at random a matrix $\mathbf{F} \in \{0,1\}^{n_c - B_c \times J}$, and the frozen bits for an active user are chosen to be the $\phi(\mathbf{m}_f)$th column of $\mathbf{F}$. The probability of a frozen bit to be one is equal to 0.5. In effect, this corresponds to each user choosing a coset of a polar code depending on $\phi(\mathbf{m}_f)$. This step provides an additional level of error detection to identify false alarms when estimating the active spreading sequences at the receiver. Suppose $\mathbf{c} \in \{0,1\}^{n_c}$ is the output of the polar encoder, and let $\mathcal{I} = \{\pi_1(\cdot), \ldots, \pi_J(\cdot)\}$ be a set of J interleavers. The polar codeword $\mathbf{c}$ is permuted using $\pi_{\phi(\mathbf{m}_f)}(\cdot)$, and the ensuing vector $\tilde{\mathbf{c}} = \pi_{\phi(\mathbf{m}_f)}(\mathbf{c})$ is modulated using QPSK to obtain vector $\mathbf{s}$ of length $T = n_c/2$.

Finally, the tth symbol of $\mathbf{s}$ is spread using the $\phi(\mathbf{m}_f)$th column of $\mathbf{A}_t$. The resulting signal $\mathbf{q}$ can be expressed as

$$\mathbf{q}(\mathbf{m}_f, \mathbf{m}_s) = \begin{bmatrix} s_1 \mathbf{a}_1^\mathsf{T} & s_2 \mathbf{a}_2^\mathsf{T} & \cdots & s_T \mathbf{a}_T^\mathsf{T} \end{bmatrix}^\mathsf{T} \quad (2)$$

where s_t is a QPSK symbol and $\mathbf{a}_t = \mathbf{A}_t[:, \phi(\mathbf{m}_f)]$ is the $\phi(\mathbf{m}_f)$th column of $\mathbf{A}_t$. The input signal to the channel is the concatenation of the pilot sequence $\mathbf{p}(\mathbf{m}_{k,f})$ and spread codeword $\mathbf{q}(\mathbf{m}_f, \mathbf{m}_s)$. Altogether, when the message of user k is $\mathbf{m}_k = (\mathbf{m}_{k,f}, \mathbf{m}_{k,s})$, the signal sent by this user is equal to

$$\mathbf{x}_k = \begin{bmatrix} \mathbf{p}^\mathsf{T}(\mathbf{m}_{k,f}) & \mathbf{q}^\mathsf{T}(\mathbf{m}_{k,f}, \mathbf{m}_{k,s}) \end{bmatrix}^\mathsf{T}$$

where $\mathbf{p}(\mathbf{m}_{k,f}) = \mathbf{P}[:, \phi(\mathbf{m}_{k,f})]$ and note that $\|\mathbf{x}_k\|^2 = n$. With this procedure, the system model of (1) can be written as the concatenation of

$$\mathbf{Y}_p = \mathbf{P}_a \mathbf{H}_a + \mathbf{Z}_p \quad \text{and} \quad \mathbf{Y}_q = \mathbf{Q}_a \mathbf{H}_a + \mathbf{Z}_q \quad (3)$$

or, in vector form,

$$\begin{bmatrix} \mathbf{Y}_p \\ \mathbf{Y}_q \end{bmatrix} = \begin{bmatrix} \mathbf{P}_a \\ \mathbf{Q}_a \end{bmatrix} \mathbf{H}_a + \begin{bmatrix} \mathbf{Z}_p \\ \mathbf{Z}_q \end{bmatrix}$$

where $\mathbf{Y}_p \in \mathbb{C}^{n_p \times M}$, $\mathbf{Y}_q \in \mathbb{C}^{TL \times M}$, and subscript a indicates sub-matrices with active columns only. That is, the kth column of $\mathbf{P}_a$ is $\mathbf{P}[:, \phi(\mathbf{m}_{k,f})]$ and the kth column of $\mathbf{Q}_a$ is $\mathbf{q}_k$.
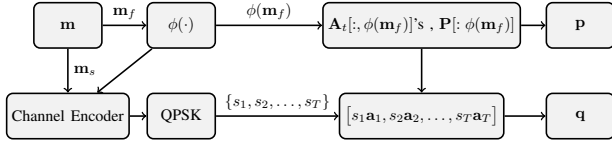

Fig. 1: This block diagram offers a synopsis of the encoder.

B. Decoder

We construct an iterative receiver that aims to recover the messages transmitted by the active users. Its main components are an energy detector, a series of MMSE estimators, and a list-polar decoder. Because of the nature of the URA, the pilot sequences cannot be assigned to the users beforehand. As a result, the first step in message recovery is to identify the active columns of $\mathbf{P}$ using an energy detector. Then MMSE estimators are utilized to estimate the channel coefficients and the QPSK symbols transmitted by the active users. Afterward, a polar list-decoder outputs the most likely messages. Finally, a successive interference canceller subtracts the estimated contribution of the decoded messages, and the procedure starts anew on the residual signal, iterating until the decoder cannot output a new message. The main building blocks of the algorithm are highlighted in Fig. 3.

1) Energy Detector: An energy detector is used to determine active spreading sequences. Mathematically, this is accomplished by correlating the received signal $\mathbf{Y}$ with the corresponding columns in $\mathbf{P}$ and $\{\mathbf{A}_t\}_{t=1}^T$ and computing the statistic λ_j given by

$$\lambda_j = \|\mathbf{P}^*[:, j]\mathbf{Y}_p\|^2 + \sum_{t=1}^T \|\mathbf{A}_t^*[:, j]\mathbf{Y}_q[\mathbf{n}_t, :]\|^2, \forall j \in [J] \quad (4)$$

where $\mathbf{n}_t = [(t-1)L + 1 : tL]$. The energy detector computes λ_j for all $j \in [J]$ and outputs the indices corresponding to the largest K values. Let $\hat{\mathcal{M}}_f$ denote the set of $\mathbf{m}_f$'s that correspond to the K largest values of λ_j's.

After this step, the active columns of $\mathbf{P}$ and $\{\mathbf{A}_t\}_{t=1}^T$, the values of the frozen positions, and the interleavers are considered known for the purpose of the algorithmic progression. Thus, the decoder can move on to the recovery of the second part of every message.

2) Channel Estimation: The channel coefficients have to be estimated before proceeding with the symbol estimation step. An MMSE filter can be derived to estimate the SIMO channels of the candidate users included in $\hat{\mathcal{K}}$. Using active pilots from the first part of the received signal $\mathbf{Y}_p$, as in (3), we obtain

$$\mathbf{W}_1 = \left(\mathbf{I}_{\hat{K}} + \frac{\hat{\mathbf{P}}^* \hat{\mathbf{P}}}{\sigma_z^2} \right)^{-1} \frac{\hat{\mathbf{P}}^*}{\sigma_z^2}$$

where $\hat{\mathbf{P}} = \mathbf{P}[:, \phi(\hat{\mathcal{M}}_f)]$. Since the filter is independent of the antenna index, the estimated channel coefficients between the $\hat{K} = |\hat{\mathcal{K}}|$ users and the M antennas are taken to be

$$\hat{\mathbf{H}} = \mathbf{W}_1 \mathbf{Y}_p \quad (5)$$

where $\hat{\mathbf{H}}$ is a $\hat{K} \times M$ matrix containing all the estimated coefficients.

3) Symbol Estimation: Since the active columns of $\mathbf{A}_t$'s and the channel coefficients have been estimated, the next step is to recover the QPSK symbols. We can write the received signal (of length L) for each symbol t as follows

$$\mathbf{Y}_q[\mathbf{n}_t, m] = \mathbf{A}_t \operatorname{diag}(\mathbf{r}_t)\mathbf{H}[:, m] + \mathbf{Z}[\mathbf{n}_t, m]$$
$$= \mathbf{A}_t \operatorname{diag}(\mathbf{H}[:, m]) \mathbf{r}_t + \mathbf{Z}[\mathbf{n}_t, m]$$

where $\mathbf{r}_t = (s_{t,1}, s_{t,2}, \ldots, s_{t,K})$ are the symbols of the users at time t. By stacking the columns of $\mathbf{Y}_q$, we can obtain the expression below

$$\underbrace{\begin{bmatrix} \mathbf{Y}_q[\mathbf{n}_t, 1] \\ \vdots \\ \mathbf{Y}_q[\mathbf{n}_t, M] \end{bmatrix}}_{\mathsf{y}_t \in \mathbb{C}^{LM \times 1}} = \underbrace{\begin{bmatrix} \mathbf{A}_t \operatorname{diag}(\mathbf{H}[:, 1]) \\ \vdots \\ \mathbf{A}_t \operatorname{diag}(\mathbf{H}[:, M]) \end{bmatrix}}_{\mathsf{B}_t \in \mathbb{C}^{LM \times K}} \underbrace{\mathbf{r}_t}_{K \times 1} + \underbrace{\begin{bmatrix} \mathbf{Z}[\mathbf{n}_t, 1] \\ \vdots \\ \mathbf{Z}[\mathbf{n}_t, M] \end{bmatrix}}_{\mathsf{z}_t \in \mathbb{C}^{LM \times 1}}.$$

As a consequence, we can apply an LMMSE estimator to the vectorized received signal

$$\mathsf{y}_t = \mathsf{B}_t \mathbf{r}_t + \mathsf{z}_t. \quad (6)$$

Since active spreading sequences and channel coefficients are not known a priori, we can modify (6) to account for the estimation inaccuracies. However, disregarding the interference term, the performance of our scheme essentially remains unaffected, and the LMMSE filter takes the form

$$\mathsf{W}_t = \left(\mathsf{I}_{\hat{K}} + \frac{\hat{\mathsf{B}}_t^* \hat{\mathsf{B}}_t}{\sigma_z^2} \right)^{-1} \frac{\hat{\mathsf{B}}_t^*}{\sigma_z^2}.$$

We stress that the set of spreading sequence changes over time, therefore an LMMSE filter has to be computed for every

$t \in [T]$. The estimated symbols of the active users at time t are given by

$$\hat{\mathbf{r}}_t = \mathtt{W}_t \mathbf{y}_t, \quad \forall \quad t \in [T].$$

4) NOPICE: One of the salient features of this scheme is the estimation of the channel using, not only the original pilots, but also the temporary coded decisions.

Definition 1. *(Temporary Coded Decisions) Once step 3 above is complete, the symbols across time corresponding to the same user are aggregated in the form of noisy codewords. These signals are then passed to a polar list-decoder. (The channel decoder block is explained in the next section.) The output of this decoder block is a set of most likely messages. These messages are subsequently re-encoded and modulated. We call the outcome of this process* temporary coded decisions.

To produce such temporary coded decisions for all users, the first three blocks in NOPICE have to be a polar decoder, an encoder, and a modulator. Let $\hat{\mathbf{s}}_k = [\hat{s}_{1,k} \ \hat{s}_{2,k} \ \dots \ \hat{s}_{T,k}]$ be the temporary coded decisions of user k, where $\hat{s}_{t,k}$ is the tth symbol of the user k. Because we know the pilots, the spreading sequences, and the symbols candidates for this user, we can construct its channel input,

$$\hat{\mathbf{x}}_k = \begin{bmatrix} \mathbf{p}^{\mathsf{T}}(\hat{\mathbf{m}}_{k,f}) & \mathbf{q}^{\mathsf{T}}(\hat{\mathbf{m}}_{k,f}, \hat{\mathbf{m}}_{k,s}) \end{bmatrix}^{\mathsf{T}}$$

where $\hat{\mathbf{q}}_k$ can be constructed as (2), by replacing the true values with the candidate symbols. Our strategy is to have $\hat{\mathbf{x}}_k$'s act as a known signal while re-estimate the channel coefficients. We stress that some of the symbols in $\hat{\mathbf{s}}_k$ may be erroneous. Nevertheless, they are used to re-estimate the channel in our algorithm, with the hope that most candidate symbols are correct. Consider the received signal at antenna m,

$$\mathbf{Y}[:, m] = \hat{\mathbf{X}}\mathbf{H}[:, m] + \left(\mathbf{X} - \hat{\mathbf{X}} \right) \mathbf{H}[:, m] + \mathbf{Z}[:, m].$$

By ignoring the interference term, we can apply the following MMSE-like filter to estimate the channel coefficients,

$$\mathbf{W}_2 = \hat{\mathbf{X}}^* \left(\hat{\mathbf{X}}^* \hat{\mathbf{X}} + \sigma_z^2 \mathbf{I}_{\hat{K}} \right)^{-1}. \tag{7}$$

Once more, the MMSE filter is independent of the antenna index and, hence, $\mathbf{W}_2$ can be applied directly to observation $\mathbf{Y}$. At this stage, the updated channel estimates are used to re-estimate the QPSK symbols by the procedure described before. Figure 2 shows a block diagram of this process.
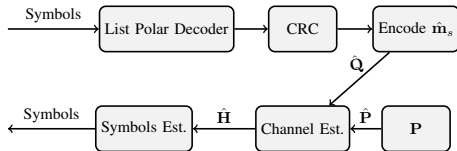

Fig. 2: This block diagram highlights the main functionalities of NOPICE. The channel is estimated using pilots and estimated symbols, although the latter may include errors.

5) Channel Decoder: This block consists of a single-user list polar decoder, and a CRC validation step. We describe the process for user k, with the understanding that these step are reproduced for all the candidate users in $\hat{\mathcal{K}}$. We note that, as a consequence, part of this procedure can be parallelized in a straightforward manner. As mentioned before, the symbol estimates and the frozen values $\mathbf{F}[:, \phi(\hat{\mathbf{m}}_f)]$ are passed to the decoder. After completion, the polar decoder provides a list of n_L likely messages. CRC validation is applied to elements of this list, and consistent messages that meet the CRC structure are retained. The most likely message within the pruned list is returned by this block.

Remark 2. *(List-Decoder and NOPICE) Under normal operation, the polar/CRC decoder returns the most likely consistent message. Yet, it is possible that the list contains no codewords that fulfill their CRC constraints. In this case, the most likely and, necessarily, inconsistent message is returned. The hope is that, although the message is guaranteed to be wrong, a portion of the encoded symbols can still be correct.*

6) SIC: The final step in this composite iterative algorithm seeks to mitigate interference through successive interference cancellation (SIC). This is accomplished by calculating the estimated channel input $\hat{\mathbf{x}}$ for all the users whose most likely messages met the CRC validation process. Then a channel estimation similarly to (7) is done. Finally, we subtract their contribution from the received signal and pass the residual to the energy detector.

Remark 3. *(Residual) When the first iteration has passed, all the building blocks of the receiver use the residual. Nevertheless, SIC uses the received signal and produces a new residual for the next iteration.*
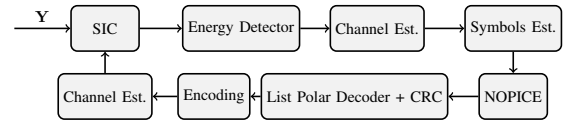

Fig. 3: This notional diagram outlines the message recovery process at the receiver.

IV. SIMULATION RESULTS

We compare FASURA, (the source code is available at https://github.com/EngProjects/mMTC), with the state-of-the-art scheme proposed by Ahmadi and Duman [17], and other two schemes proposed by Fengler et al. [16] and Decurninge et al. [15] to assess its performance. To ensure a fair comparison between the two URA communication schemes, we pick parameters for our system that match their reported implementation. Specifically, we choose $B = 100$ message bits, $n = 3200$ complex channel uses, and a target probability of error $\mathrm{P}_e \leq 0.05$. Beyond these constraints, the other parameters for FASURA are $n_p = 896$, $L = 9$, $n_c = 512$, $n_L = 64$ and $J = 2^{16}$. We randomly generate $T = \frac{n_c}{2} = 256$ spreading sequence matrices and one pilot matrix, as discussed

in Section III. For $K = 100$, we use 12 CRC bits; whereas when $K > 100$, we employ 16 CRC bits. The number of antennas at the base station is set to $M = 50$.

Figure 4 plots the performance of FASURA, along with that of the communication scheme by Ahmadi et al. found in [17]. Also, to motivate the use of the NOPICE block, we report the performance of FASURA with and without the NOPICE channel estimation technique. For the operational parameters studied and a user population exceeding 100 active devices, FASURA outperforms the scheme proposed in [17]. Interestingly, as the number of active users grows, the gap between the two schemes widens. For example, when the number of users goes from 100 to 800, the gap increases from 0.95 dB to 2.5 dB. Furthermore, the presence of the NOPICE block seems to uniformly improve the performance of FASURA. Our proposed scheme also substantially outperforms the pilot-based [16] and tensor based modulation scheme in [15]. We should mention, however, that these benefits in terms of E_b/N_0 come at the expense of additional computations.

V. Conclusion

This article considers the massive MIMO unsourced random access problem on a quasi-static Rayleigh fading channel. We propose a novel communication scheme called FASURA that outperforms existing schemes when $K \geq 100$. We explore a different way to perform channel estimation, which we call NOPICE, whereby after temporary decoding decisions have been made, the channel is re-estimated assuming that the decoded messages are accurate. Also, through numerical simulations, we conduct a comparative study of FASURA, and the schemes in [15], [16] and [17]. The proposed scheme outperforms the orthogonal pilots scheme recently published in [17], when the number of users is more than 100. Finally, it seems that the approach adopted within the NOPICE block leads to uniformly better performance, with gains of 0.4 dB on average.
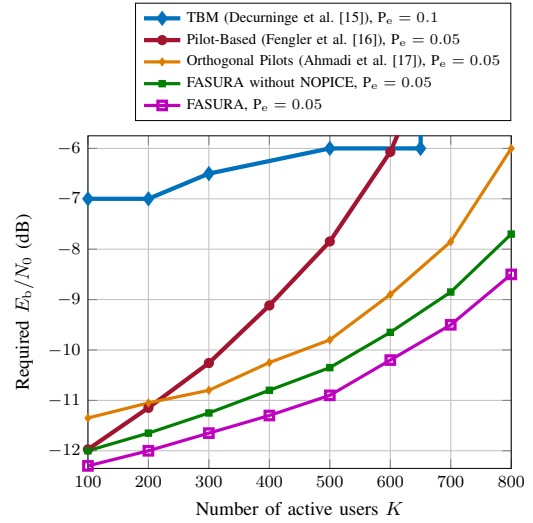

Fig. 4: This plot compares the performance of the proposed scheme to the performance of previously published schemes when $M = 50$, $n = 3200$ and $P_e \leq 0.05$.

References

[1] Y. Polyanskiy, "A perspective on massive random-access," in *IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 2523–2527.

[2] A. Vem, K. R. Narayanan, J. Cheng, and J.-F. Chamberland, "A user-independent serial interference cancellation based coding scheme for the unsourced random access gaussian channel," in *IEEE Information Theory Workshop (ITW)*, 2017, pp. 121–125.

[3] V. K. Amalladinne, J.-F. Chamberland, and K. R. Narayanan, "A coded compressed sensing scheme for unsourced multiple access," *IEEE Transactions on Information Theory*, vol. 66, no. 10, pp. 6509–6533, 2020.

[4] A. Fengler, P. Jung, and G. Caire, "SPARCs and AMP for unsourced random access," in *IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2843–2847.

[5] K. Andreev, P. Rybin, and A. Frolov, "Unsourced random access based on list recoverable codes correcting t errors," in *IEEE Information Theory Workshop (ITW)*, 2021, pp. 1–6.

[6] A. K. Pradhan, V. K. Amalladinne, K. R. Narayanan, and J.-F. Chamberland, "Polar coding and random spreading for unsourced multiple access," in *IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[7] A. K. Pradhan, Amalladinne V. K., K. R. Narayanan, and J.-F. Chamberland, "LDPC codes with soft interference cancellation for uncoordinated unsourced multiple access," *arXiv preprint arXiv:2105.13985*, 2021.

[8] E. Marshakov, G. Balitskiy, K. Andreev, and A. Frolov, "A polar code based unsourced random access for the gaussian MAC," in *IEEE Vehicular Technology Conference (VTC)*, 2019, pp. 1–5.

[9] K. Andreev, E. Marshakov, and A. Frolov, "A polar code based TIN-SIC scheme for the unsourced random access in the quasi-static fading MAC," in *IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 3019–3024.

[10] A. Fengler, S. Haghighatshoar, P. Jung, and G. Caire, "Non-bayesian activity detection, large-scale fading coefficient estimation, and unsourced random access with a massive MIMO receiver," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 2925–2951, 2021.

[11] V. K Amalladinne, J. R. Ebert, J.-F. Chamberland, and K. R Narayanan, "An enhanced decoding algorithm for coded compressed sensing with applications to unsourced random access," *Sensors*, vol. 22, no. 2, pp. 676, 2022.

[12] J. Liu and X. Wang, "Sparsity-exploiting blind receiver algorithms for unsourced multiple access in MIMO and massive MIMO channels," *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8055–8067, 2021.

[13] A. P. Sabulal and S. Bhashyam, "Joint sparse recovery using deep unfolding with application to massive random access," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 5050–5054.

[14] Y. Cheng, L. Liu, and L. Ping, "Orthogonal AMP for massive access in channels with spatial and temporal correlations," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 726–740, 2021.

[15] A. Decurninge, I. Land, and M. Guillaud, "Tensor-based modulation for unsourced massive random access," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 552–556, 2021.

[16] A. Fengler, O. Musa, P. Jung, and G. Caire, "Pilot-based unsourced random access with a massive MIMO receiver, interference cancellation, and power control," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2022.

[17] M. J. Ahmadi and T. M. Duman, "Unsourced random access with a massive MIMO receiver using multiple stages of orthogonal pilots," *arXiv preprint arXiv:2202.01477*, 2022.

[18] J. Liu and X. Wang, "Unsourced multiple access based on sparse tanner graph – efficient decoding, analysis and optimization," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2022.

[19] C. R. Srivatsa and C. R. Murthy, "On the impact of channel estimation on the design and analysis of IRSA based systems," *arXiv preprint arXiv:2112.07242*, 2021.

[20] V. Shyianov, F. Bellili, A. Mezghani, and E. Hossain, "Massive unsourced random access based on uncoupled compressive sensing: Another blessing of massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 820–834, 2021.