# A Sparsity Detection Framework for On–Off Random Access Channels

Alyson K. Fletcher
University of California, Berkeley
Email: alyson@eecs.berkeley.edu

Sundeep Rangan
Qualcomm Technologies
Email: srangan@qualcomm.com

Vivek K Goyal
Massachusetts Institute of Technology
Email: vgoyal@mit.edu

*Abstract*—This paper considers a simple on–off random multiple access channel (MAC), where $n$ users communicate simultaneously to a single receiver. Each user is assigned a single codeword which it transmits with some probability $\lambda$ over $m$ degrees of freedom. The receiver must detect which users transmitted. We show that detection for this random MAC is mathematically equivalent to a standard sparsity detection problem. Using new results in sparse estimation we are able to estimate the capacity of these channels and compare the achieved performance of various detection algorithms. The analysis provides insight into the roles of power control and multi-user detection.

## I. INTRODUCTION

In wireless systems, *random access* refers to any multi-user communication protocol where the users autonomously decide whether or not to transmit depending on their own traffic requirements and estimates of the network load. While random access is best known for its use in packet data communication in wireless local area networks (LANs), this paper considers random access for simple on–off messaging. On–off random access signaling can be used for a variety of control tasks in wireless networks such as user presence indication, initial access, scheduling requests and paging.

The limits of on–off random access signaling with multiple users are not fully understood. To this end, we consider a simple random multiple access channel where $n$ users transmit to a single receiver. Each user is assigned a single codeword which it transmits with probability $\lambda$. We wish to understand the capacity of these channels, by which we mean the total number of degrees of freedom $m$ needed to reliably detect which users transmit as a function of $n$, $\lambda$, and the channel conditions. We also wish to establish performance bounds for specific decoding algorithms.

This on–off random access channel is related to the classic multiple access channel (MAC) in network information theory [1], [2]. Unfortunately, it is difficult to apply the classic MAC channel analysis directly to the on–off random access channel under consideration here. In the classic MAC channel analysis, each user can employ a capacity achieving code. In contrast, no coding is possible with on–off signaling. In addition, "uncoded" multiuser detection results [3], [4] require that the received constellation is discrete and known at the receiver, which is not possible for non-coherent reception as is typically required for random access communication.

Our analysis is instead based on showing that the problem at the receiver of detecting the active users is mathematically equivalent to a sparsity detection problem that arises in *compressed sensing* [5]–[7]. Exploiting recent results from this field, such as those in [8], [9], provides simple bounds on the maximum achievable capacity assuming optimal detection.

Unfortunately, optimal sparsity detection is a well-known NP-hard problem [10]. Current commercial designs typically use simple single-user detection (see, for example [11] for WCDMA). However, there are well-known practical, but suboptimal, algorithms such as the popular orthogonal matching pursuit (OMP) [12]–[14] and lasso [15] methods in sparse estimation that can be used for multiuser detection for on–off random access channels. In contrast to single-user detection, we show that these methods can offer a level of robustness or insensitivity to large dynamic ranges in received power levels. This near–far resistance feature is similar to that of standard MMSE multi-user detection in CDMA [16].

However, we also show that single-user detection, as well as both lasso and OMP, all suffer from a certain *self-noise* limit at high SNRs. As a result, there is a significant potential gap in performance between these techniques and optimal maximum likelihood (ML) multiuser detection. However, we show that if accurate power control is available, this SNR limit can be reduced by using a modified version of OMP called sequential OMP (SeqOMP) in conjunction with *power shaping*. The method is analogous to the classic successive interference cancelation (SIC) method in the MAC channel.

The connection between sparsity detection methods such as OMP and the SIC technique for the MAC channel has also been observed in the recent work of Jin and Rao [17]. A related work by Wipf and Rao [18] also gave some empirical evidence for the benefit of power shaping when used in conjunction with sparse Bayesian learning algorithms. The results in this paper make the connections between sparsity detection and the random access MAC channel more precise by giving conditions on the detectability of the sparsity pattern, characterizing the optimal power shaping distribution, and contrasting the classic MAC and random access MAC capacities.

This paper summarizes the results of [19]; detailed proofs, numerical experiments, and additional context appear therein.

## II. ON–OFF RANDOM ACCESS CHANNEL MODEL

Assume that there are $n$ transmitters sharing a wireless channel to a single receiver. Each user $j$ is assigned a unique, dedicated codeword represented as an $m$-dimensional vector

$\mathbf{a}_j \in \mathbb{C}^m$, where $m$ is the total number of degrees of freedom in the channel. By degrees of freedom we simply mean the dimension of the received vector, which represents the number of samples in time or frequency depending on the modulation. In any channel use, only some fraction of the users, $\lambda \in (0, 1)$, transmit their codeword. The fraction $\lambda$ will be called the *activity ratio* and any user that transmits will be called *active*.

The signal at the receiver from each user $j$ is modeled as $x_j \mathbf{a}_j$ where $x_j$ is a complex scalar. If the user is not active, $x_j = 0$. If the user is active, $x_j$ would represent the product of the transmitted symbol and channel gain. Note that we have assumed flat fading. The total signal at the receiver is given by

$$\mathbf{y} = \sum_{j=1}^{n} \mathbf{a}_j x_j + \mathbf{w} = \mathbf{A}\mathbf{x} + \mathbf{w}, \qquad (1)$$

where $\mathbf{w} \in \mathbb{C}^m$ represents noise. The matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ is formed by codewords $\mathbf{a}_j$, $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, and will be called the *codebook*. The vector $\mathbf{x} = [x_1 \cdots x_n]^T$ will be called the *modulation vector*, and its components $\{x_j\}_{j=1}^{n}$ are referred to as the *received modulation symbols*. Note that $x_j$ succinctly encapsulates the channel gain, transmit power, and phase.

Given a modulation vector $\mathbf{x}$, define the *active user set* as

$$I_{\text{true}} = \{\, j \; : x_j \neq 0 \,\}, \qquad (2)$$

which is the "true" set of active users. The size of the active user set is related to the activity ratio through $\lambda = |I_{\text{true}}|/n$. The goal of the receiver is to determine an estimate $\hat{I} = \hat{I}(\mathbf{y})$ of $I_{\text{true}}$ based on the received noisy vector $\mathbf{y}$.

We will be interested in estimators that exploit minimal prior knowledge of the modulation vector $\mathbf{x}$ other than it being sparse. In particular, we will limit our attention to estimators that do not explicitly require *a priori* knowledge of the complex modulation symbols $x_j$. We make this assumption since the channel gain is typically unknown or uncertain at the receiver; users conducting random access communication are unlikely to be concurrently sending a persistent pilot reference. Knowledge of a discrete alphabet for each $x_j$ obviously changes the receiver's task dramatically [3]. For example, this knowledge would enable perfect cancellation of a user's signal without orthogonalization, so the number of active users could exceed the number of degrees of freedom.

We consider large random codebooks where the entries of $\mathbf{A}$ are i.i.d. complex normal $\mathcal{CN}(0, 1/m)$. We assume the noise vector is also Gaussian: $\mathbf{w} \sim \mathcal{CN}(0, (1/m)I_m)$. Given an estimator, $\hat{I} = \hat{I}(\mathbf{y})$, the probability of error,

$$p_{\text{err}} = \Pr\left(\hat{I} \neq I_{\text{true}}\right), \qquad (3)$$

is taken with respect to random codebook $\mathbf{A}$, the noise vector $\mathbf{w}$, and the statistical distribution of the modulation vector $\mathbf{x}$. We want to find estimators $\hat{I}$ that bring $p_{\text{err}}$ close to zero. An alternative could be to allow a nonzero constant fraction of detection errors. This may change scaling behavior considerably [20], [21].

We will see that two key factors influence the ability to detect the active user set. The first is the total SNR defined as

$$\text{SNR} = \frac{\mathbf{E}\|\mathbf{A}\mathbf{x}\|^2}{\mathbf{E}\|\mathbf{w}\|^2}. \qquad (4)$$

Since the components of the matrix $\mathbf{A}$ and noise vector $\mathbf{w}$ are i.i.d. $\mathcal{CN}(0, 1/m)$, it can be verified that, for deterministic $\mathbf{x}$,

$$\text{SNR} = \|\mathbf{x}\|^2. \qquad (5)$$

In the case of random $\mathbf{x}$, this is the conditional SNR given $\mathbf{x}$; we will have both deterministic and random formulations.

The second term is the *minimum-to-average ratio*

$$\text{MAR} = \frac{\min_{j \in I_{\text{true}}} |x_j|^2}{\|\mathbf{x}\|^2/\lambda n}. \qquad (6)$$

Since $I_{\text{true}}$ has $\lambda n$ elements, $\|\mathbf{x}\|^2/\lambda n$ is the average of $\{|x_j|^2 \mid j \in I_{\text{true}}\}$. Therefore, $\text{MAR} \in (0, 1]$ with the upper limit occurring when all the nonzero entries of $\mathbf{x}$ have the same magnitude. MAR is a deterministic quantity when $\mathbf{x}$ is deterministic and a random variable otherwise.

The factor MAR in (6) is the ratio of the minimum to average received power and is thus related to the dynamic range amongst users. We will show that low MAR can make reliable detection significantly more difficult for certain algorithms. The problem is analogous to the well-known *near–far effect* in CDMA systems [16], where users with weak signals can be dominated by higher-power signals.

## III. CURRENT SPARSITY DETECTION METHODS

The problem of detecting the active user set is precisely equivalent to a sparsity detection problem. To see this, note that the modulation vector $\mathbf{x}$ is *sparse* with nonzero components only in positions corresponding to the active users. The problem at the receiver is to detect these nonzero positions in $\mathbf{x}$ from linear noisy observations $\mathbf{y}$ in (1).

Table I summarizes scaling laws for various detection algorithms. We discuss details below.

### A. Optimal Detection with No Noise

To understand the limits of detection, it is useful to first consider the minimum number of degrees of freedom when there is no noise. Since the activity ratio is $\lambda$, $\mathbf{x}$ will have $k = \lambda n$ nonzero components. For a lower bound on the minimum number of degrees of freedom needed for reliable detection, suppose that the receiver knows the number of active users $k$ as side information.

With no noise, the received vector is $\mathbf{y} = \mathbf{A}\mathbf{x}$, which will belong to one of $J = \binom{n}{k}$ subspaces spanned by $k$ columns of $\mathbf{A}$. If $m > k$, then these subspaces will be distinct with probability 1. Thus, an exhaustive search through the subspaces will reveal which subspace $\mathbf{y}$ belongs to and thus determine the active user set. This shows that with no noise and no computational limits, having more degrees of freedom than active users is sufficient for asymptotic reliable detection.

Conversely, if no prior information is known at the receiver other than $\mathbf{x}$ being $k$-sparse, then having more degrees of

| | finite SNR · MAR | SNR · MAR $\to \infty$ |
|---|---|---|
| Necessary for ML | $m > \frac{1}{\text{MAR·SNR}}\lambda n \log(n(1-\lambda))$ <br> Fletcher *et al.* [9, Thm. 1] | $m > \lambda n$ <br> (elementary) |
| Sufficient for ML | $m > \frac{C}{\text{MAR·SNR}}\lambda n \log(n(1-\lambda))$ <br> Wainwright [8] | $m > \lambda n$ <br> (elementary) |
| Sufficient for sequential OMP with power shaping | $m > \frac{4}{\log(1+\text{SNR})}\lambda n \log(n(1-\lambda))$ <br> **From Theorem 1** | $m > 5\lambda n$ <br> **From Theorem 1** |
| Necessary and sufficient for lasso | unknown (expression to the right is necessary) | $m > \lambda n \log(n(1-\lambda))$ <br> Wainwright [22] |
| Sufficient for OMP | unknown | $m > 2\lambda n \log(n)$ <br> Tropp and Gilbert [23] |
| Sufficient for single-user detection (9) | $m > \frac{4(1+\text{SNR})}{\text{MAR·SNR}}\lambda n \log(n(1-\lambda))$ <br> Fletcher *et al.* [9, Thm. 2] | $m > \frac{4}{\text{MAR}}\lambda n \log(n(1-\lambda))$ |

TABLE I

SUMMARY OF RESULTS ON DEGREE OF FREEDOM SCALINGS FOR ASYMPTOTIC RELIABLE DETECTION FOR VARIOUS DETECTION ALGORITHMS. ONLY LEADING TERMS ARE SHOWN. SEE BODY FOR DEFINITIONS AND ADDITIONAL TECHNICAL LIMITATIONS.

freedom than active users is also necessary. If $m \leq k = \lambda n$, then for almost all codebooks $\mathbf{A}$, any $k$ columns of $\mathbf{A}$ span $\mathbb{C}^m$. Consequently, any received vector $\mathbf{y} = \mathbf{Ax}$ is consistent with any $k$ users transmitting. Thus, the active user set cannot be determined without further prior information on the modulation vector $\mathbf{x}$.

### B. ML Detection with Noise

Now suppose there is noise. For a lower bound on the degrees of freedom needed for reliable detection, suppose the receiver has, as side information, the number of active users $k$. In this case, we can consider the maximum likelihood (ML) detector, which is the optimal detector without any prior on $\mathbf{x}$ other than $\mathbf{x}$ being $k$-sparse. Since the noise $\mathbf{w}$ is Gaussian, the ML detector finds the $k$-dimensional subspace spanned by $k$ columns of $\mathbf{A}$ containing the maximum energy of $\mathbf{y}$.

The ML estimator was first analyzed by Wainwright [8]. The results in that work, along with the fact that $k = \lambda n$, show that there exists a constant $C > 0$ such that if

$$m \geq C \max\left\{ \frac{1}{\text{MAR} \cdot \text{SNR}}\lambda n \log(n(1-\lambda)), -\lambda n \log(\lambda) \right\} \tag{7}$$

then ML will asymptotically detect the correct active user set. Also, [9, Thm. 1] shows that, for any $\delta > 0$, the condition

$$m \geq \frac{1+\delta}{\text{MAR} \cdot \text{SNR}}\lambda n \log(n(1-\lambda)) + \lambda n, \tag{8}$$

is necessary.[1]

### C. Single User Detection

The most common and simple method to detect the active user set is a single-user detection estimator of the form

$$\hat{I}_{\text{SUD}} = \{ j \,:\, \rho(j) > \mu \}, \tag{9}$$

[1]Note that here and elsewhere, there is a factor of two difference between the real and complex cases.

where $\mu > 0$ is a threshold parameter and $\rho(j)$ is the correlation coefficient:

$$\rho(j) = \frac{|\mathbf{a}'_j \mathbf{y}|^2}{\|\mathbf{a}_j\|^2 \|\mathbf{y}\|^2}. \tag{10}$$

Single-user detection has been analyzed in the compressed sensing context in [9], [24], [25]. A small modification of [9, Thm. 2] shows the following result: Suppose,

$$m(n) > \frac{(1+\delta)L(\lambda, n)(1+\text{SNR})}{\text{SNR} \cdot \text{MAR}}\lambda n, \tag{11}$$

where $\delta > 0$ and

$$L(\lambda, n) = \left[ \sqrt{\log(n(1-\lambda))} + \sqrt{\log(n\lambda)} \right]^2. \tag{12}$$

Then there exists a sequence of detection thresholds $\mu = \mu(n)$ such that single-user detection achieves asymptotic reliable detection of the active user set.

Comparing the sufficient condition (11) for single-user detection with the *necessary* condition (8), we see two distinct problems in single-user detection:

- *Constant offset:* The scaling (11) for single-user detection shows a factor $L(\lambda, n)$ instead of $\log((1-\lambda)n)$ in (8). It is easily verified that, for $\lambda \in (0, 1/2)$, $L(\lambda, n)/\log((1-\lambda)n) < 4$, so the more stringent, but simpler, condition

$$m(n) > \frac{(1+\delta)4(1+\text{SNR})}{\text{SNR} \cdot \text{MAR}}\lambda n \log((1-\lambda)n) \tag{13}$$

  is also sufficient.

- *Self-noise limit:* Single-user detection has an additional $1+\text{SNR}$ term in the number of degrees of freedom. When detecting any one component of $x$, single-user detection sees the energy from the rest of the signal as interference. This "self-noise" increases the effective noise by a factor of $1 + \text{SNR}$, which results in a proportional increase in the minimum number of degrees of freedom.

This self-noise results in a large performance gap at high SNRs. In particular, as $\mathsf{SNR} \to \infty$, (11) reduces to

$$m(n) > \frac{(1+\delta)L(\lambda,n)}{\mathsf{MAR}} \lambda n \log((1-\lambda)n). \quad (14)$$

In contrast, ML may be able to succeed with a scaling $m = O(\lambda n)$ for high SNRs, which is fundamentally better than the $m = \Omega(\lambda n \log((1-\lambda)n)$ required by single-user detection.

### D. Lasso and OMP Estimation

While ML has clear advantages over single-user detection, it is not computationally feasible. However, one practical method used in sparse signal estimation is the *lasso* estimator [15], also called basis pursuit denoising [26]. In the context of the random access channel, the lasso estimator would first estimate the modulation vector $\mathbf{x}$ by solving the convex minimization

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left( \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mu\|\mathbf{x}\|_1 \right), \quad (15)$$

where $\mu > 0$ is an algorithm parameter that "encourages" sparsity in the solution $\widehat{\mathbf{x}}$. The nonzero components of $\widehat{\mathbf{x}}$ can then be used as an estimate of the active user set.

The exact performance of lasso is not known at finite SNR. However, Wainwright [22] has exactly characterized the conditions for lasso to work at high SNR. Specifically, if $m$, $n$ and $\lambda n \to \infty$, with $\mathsf{SNR} \cdot \mathsf{MAR} \to \infty$, the scaling

$$m > \lambda n \log(n(1-\lambda)) + \lambda n + 1, \quad (16)$$

is necessary and sufficient for asymptotic sparsity recovery.

Another common approach to sparsity pattern detection is the greedy OMP algorithm [12]–[14]. This has been analyzed by Tropp and Gilbert [23] in a setting with no noise. They show that, when $\mathbf{A}$ has Gaussian entries, a *sufficient* condition for asymptotic reliable recovery is

$$m > 2\lambda n \log(n) + C\lambda n, \quad (17)$$

for some constant $C > 0$, similar to the lasso scaling law. A variant of OMP, called CoSaMP [27] improves upon this scaling.

The lasso and OMP scaling laws, (16) and (17), can be compared with the high SNR limit for the single-user detection scaling law in (14). We see two important benefits of lasso and OMP: First, the $L(n,k)$ term for single-user detection is replaced by $\log(n(1-\lambda))$ for lasso or $2\log(n)$ for OMP, providing a gain of up to 4 for lasso and 2 for OMP.

Second, and more importantly, both lasso and OMP do not have a dependence on MAR as long as $\mathsf{SNR} \cdot \mathsf{MAR} \to \infty$; thus, in the high SNR regime, they have a near–far resistance that single-user detection does not.

On the other hand, we also see from (16) and (17) that both lasso and OMP are unable to achieve the scaling $m = O(\lambda n)$ that may be achievable with ML at high SNR. Instead, both have the scaling $m = O(\lambda n \log((1-\lambda)n))$ and thus also suffer from the self-noise limit.

## IV. SEQUENTIAL ORTHOGONAL MATCHING PURSUIT

To show that the self-noise barrier can be broken with a practical algorithm when accurate power control is available, we propose a new algorithm, which we call *sequential orthogonal matching pursuit* or SeqOMP. The algorithm works in $n$ iterations indexed by $j = 1, \ldots, n$, producing a sequence of estimates of the active user set, $\hat{I}(j)$, in each iteration. The initial estimate $\hat{I}(0)$ is taken as empty. At iteration $j$, the algorithm computes the correlation

$$\rho(j) = \frac{|\mathbf{a}_j'\mathbf{P}(j)\mathbf{y}|^2}{\|\mathbf{P}(j)\mathbf{a}_j\|^2\|\mathbf{P}(j)\mathbf{y}\|^2},$$

where $\mathbf{P}(j)$ is the projection onto the orthogonal complement of the previously detected vectors: $\{\mathbf{a}_\ell, \ell \in \hat{I}(j-1)\}$. If the correlation $\rho(j)$ exceeds a threshold level $\mu$, the index is added so: $\hat{I}(j) = \hat{I}(j-1) \cup \{j\}$. Otherwise, $\hat{I}(j) = \hat{I}(j-1)$. The final estimate is $\hat{I}_{\mathrm{SOMP}} = \hat{I}(n)$.

The SeqOMP algorithm can be thought of as an iterative version of single-user detection with the difference that, after an active user is detected, subsequent correlations are performed only in the orthogonal complement to the detected codeword. The method is identical to the standard OMP algorithm of [12]–[14], except that SeqOMP passes through the data only once. For this reason, SeqOMP is actually computationally simpler than standard OMP.

To understand the performance of SeqOMP, let $p_j$ denote the received modulation symbol power $p_j = |x_j|^2$, conditional that user $j$ is active. Since each user transmits with a probability $\lambda$, the total SNR is given by

$$\mathsf{SNR} = \lambda \sum_{j=1}^n p_j. \quad (18)$$

Given a power profile, a key parameter in estimating the performance of SeqOMP is what we will call the *minimum signal-to-interference and noise ratio (SINR)* defined as

$$\gamma = \min_{\ell=1,\ldots,n} p_\ell/\widehat{\sigma}^2(\ell), \quad \widehat{\sigma}^2(\ell) = 1 + \lambda\sum_{j=\ell+1}^n p_j \quad (19)$$

The ratio in the right-hand side of (19) represents the SINR seen when detecting user $\ell$ assuming users $j < \ell$ have been canceled perfectly. Our main result is the following:

*Theorem 1 ( [19]):* Let $\lambda = \lambda(n)$, $m = m(n)$ and the power profile $\{p_j\}_{j=1}^n = \{p_j(n)\}_{j=1}^n$, be deterministic quantities that all vary with $n$ satisfying the limits $m - \lambda n$, $\lambda n$ and $(1-\lambda)n \to \infty$, and $\gamma \to 0$. Also, assume the sequence of power profiles satisfies the limit

$$\lim_{n\to\infty} \max_{i=1,\ldots,n-1} \log(n)\widehat{\sigma}^{-4}(i) \sum_{j>i}^n p_j^2 = 0. \quad (20)$$

Finally, assume that for all $n$,

$$m \geq (1+\delta)L(n,\lambda)/\gamma + \lambda n, \quad (21)$$

for some $\delta > 0$ and $L(n,\lambda)$ defined in (12). Then, there exists a sequence of thresholds, $\mu = \mu(n)$, such that SeqOMP will achieve asymptotic reliable detection of the active user set:

$$p_{\mathrm{err}} = \Pr\left(\hat{I}_{\mathrm{SOMP}} \neq I_{\mathrm{true}}\right) \to 0,$$

where the probability is taken over the randomness in the activities of the users, the codebook $\mathbf{A}$, and the noise $\mathbf{w}$. The sequence of threshold levels can be selected independent of the sequence of power profiles.

The theorem shows the following properties of SeqOMP:

- *Near-far resistance with known power ordering and high SNR:* Suppose that the power ordering $p_j$ is known at the receiver so the receiver can detect the users in order of decreasing power. If, in addition, the SNRs of all the users go to infinity so that $p_j \to \infty$ for all $j$, then it can be verified that $\gamma > 1/(\lambda n)$. In this case, the scaling (21) reduces to

$$m \geq (1 + \delta)\lambda n L(n, \lambda) + \lambda n,$$

which is identical to the lasso performance except for the constant offset $L(\lambda, n)/\log((1 - \lambda)n)$. In particular, the minimum number of degrees of freedom does not depend on MAR and therefore, similar to lasso and OMP, SeqOMP is able to achieve some level of near–far resistance.

- *Reduction of the SNR barrier with power control:* Now suppose that accurate power control is feasible so that the receive power levels $p_j$ can be set by the receiver. In this case, we can maximize the SINR $\gamma$ in (19) for a given total SNR constraint (18). For large $n$, the optimal value is $\gamma = (1/\lambda n)\log(1 + \mathsf{SNR})$. Substituting this value into (21) results in a scaling

$$m \geq \frac{1 + \delta}{\log(1 + \mathsf{SNR})}\lambda n L(n, \lambda) + \lambda n.$$

An important property of this scaling is that when $\mathsf{SNR} = \Omega(\lambda n)$, the number of degrees of freedom scales as $m = O(\lambda n)$, which matches the scaling of ML at high SNRs. This scaling improves the high SNR scaling of $m = O(\lambda n \log((1 - \lambda n))$ achieved by single-user detection, lasso and OMP. In this sense, SeqOMP is a "practical" algorithm that can break the SNR barrier.

*Practical Considerations:* The use of power control requires some justification. On–off random access is most likely used when users are transmitting intermittently, so fast closed-loop power control is not likely available. However, in cellular systems, mobiles can achieve some degree of power control by estimating the path loss in the downlink and adjusting the uplink transmit power accordingly. Also, SeqOMP does not require any knowledge of the channel phase. Moreover, the effect of small power control errors may not be large since they result only in a suboptimal detection sequence and interference profile. The SeqOMP algorithm does not assume the receive power is known in the cancellation.

## V. CONCLUSIONS

Sparse signal detection is a valuable framework for understanding multiple access on–off random signaling. Results can provide simple capacity estimates and clarify the role of power control and multiuser detection. Combining these ideas with concepts from the classic MAC such as SIC and power shaping, we have presented a new algorithm, SeqOMP, that has certain advantages at high SNR. We believe that new methods—including iterative decoding and MMSE detection—can improve these results further.

## REFERENCES

[1] R. Ahlswede, "Multi-way communication channels," in *Proc. IEEE Int. Symp. Inform. Th.*, Sep. 1971, pp. 23–52.
[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 1991.
[3] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channel," *IEEE Trans. Inform. Th.*, 32(1):85–96, 1986.
[4] S. Verdu, *Multiuser Detection*. Cambridge University Press, 1998.
[5] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Th.*, 52(2):489–509, 2006.
[6] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Th.*, 52(4):1289–1306, 2006.
[7] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inform. Th.*, 52(12):5406–5425, 2006.
[8] M. J. Wainwright, "Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting," Univ. of California, Berkeley, Dept. of Statistics, Tech. Rep. 725, Jan. 2007.
[9] A. K. Fletcher, S. Rangan, and V. K. Goyal, "Necessary and sufficient conditions on sparsity pattern recovery," arXiv:0804.1839v1, Apr. 2008.
[10] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Computing*, 24(2):227–234, 1995.
[11] L. Qiu, Y. Huang, and J. Zhu, "Fast acquisition scheme and implementation of PRACH in WCDMA system," in *Proc. IEEE Veh. Tech. Conf.*, Oct. 2001, pp. 1701–1705.
[12] S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *Int. J. Control*, 50(5):1873–1896, 1989.
[13] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Asilomar Conf. Sig., Sys., Comput.*, 1993, pp. 40–44.
[14] G. Davis, S. Mallat, and Z. Zhang, "Adaptive time-frequency decomposition," *Optical Eng.*, 37(7):2183–2191, 1994.
[15] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Stat. Soc., Ser. B*, 58(1):267–288, 1996.
[16] R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Comm.*, 38(4):496–508, 1990.
[17] Y. Jin and B. Rao, "Performance limits of matching pursuit algorithms," in *Proc. IEEE Int. Symp. Inform. Th.*, Jun. 2008, pp. 2444–2448.
[18] D. Wipf and B. Rao, "Comparing the effects of different weight distributions on finding sparse representations," in *Proc. Neural Information Process. Syst.*, Dec. 2006.
[19] A. K. Fletcher, S. Rangan, and V. K. Goyal, "On–off random access channels: A compressed sensing framework," arXiv:0903.1022v1, Mar. 2009.
[20] M. Akçakaya and V. Tarokh, "Shannon theoretic limits on noisy compressive sampling," arXiv:0711.0366v1, Nov. 2007.
[21] G. Reeves, "Sparse signal sampling using noisy linear projections," Univ. of California, Berkeley, Tech. Rep. UCB/EECS-2008-3, Jan. 2008.
[22] M. J. Wainwright, "Sharp thresholds for high-dimensional and noisy recovery of sparsity," Univ. of California, Berkeley, Dept. of Statistics, Tech. Rep., May 2006, arXiv:math.ST/0605740 v1 30 May 2006.
[23] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Th.*, 53(12):4655–4666, 2007.
[24] M. F. Duarte, S. Sarvotham, D. Baron, W. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Conf. Rec. Asilomar Conf. on Sig., Sys. & Computers*, Nov. 2005.
[25] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Inform. Th.*, 54(5):2210–2219, 2008.
[26] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comp.*, 20(1):33–61, 1999.
[27] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," arXiv:0803.2392v2, Apr. 2008.