

MIT Open Access Articles

*A Sparsity Detection Framework for
On-Off Random Access Channels*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Fletcher, Alyson K., Sundeep Rangan, and Vivek K. Goyal. "A sparsity detection framework for on-off random access channels." Wavelets XIII. Ed. Vivek K. Goyal, Manos Papadakis, & Dimitri Van De Ville. San Diego, CA, USA: SPIE, 2009. 744607-15. © 2009 SPIE

As Published: <http://dx.doi.org/10.1117/12.824127>

Publisher: Society of Photo-optical Instrumentation Engineers

Persistent URL: <http://hdl.handle.net/1721.1/54720>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



A Sparsity Detection Framework for On–Off Random Access Channels

Alyson K. Fletcher,^a Sundeep Rangan,^b and Vivek K Goyal^c

^aUniversity of California, Berkeley

^bQualcomm Technologies

^cMassachusetts Institute of Technology

ABSTRACT

This paper considers a simple on–off random multiple access channel, where n users communicate simultaneously to a single receiver over m degrees of freedom. Each user transmits with probability λ , where typically $\lambda n < m \ll n$, and the receiver must detect which users transmitted. We show that when the codebook has i.i.d. Gaussian entries, detecting which users transmitted is mathematically equivalent to a certain sparsity detection problem considered in compressed sensing. Using recent sparsity results, we derive upper and lower bounds on the capacities of these channels. We show that common sparsity detection algorithms, such as lasso and orthogonal matching pursuit (OMP), can be used as tractable multiuser detection schemes and have significantly better performance than single-user detection. These methods do achieve some near–far resistance but—at high signal-to-noise ratios (SNRs)—may achieve capacities far below optimal maximum likelihood detection. We then present a new algorithm, called sequential OMP, that illustrates that iterative detection combined with power ordering or power shaping can significantly improve the high SNR performance. Sequential OMP is analogous to successive interference cancellation in the classic multiple access channel. Our results thereby provide insight into the roles of power control and multiuser detection on random-access signaling.

Keywords: compressed sensing, convex optimization, lasso, maximum likelihood estimation, multiple access channel, multiuser detection, orthogonal matching pursuit, power control, random matrices, single-user detection, sparsity, thresholding

1. INTRODUCTION

In wireless systems, *random access* refers to the users autonomously deciding whether 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, 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.

We consider a simple random multiple access channel where n users transmit to a single receiver, each independently with probability λ . Each user is assigned a single codeword, and when it is “on” it transmits a scalar multiple of that codeword. 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 , λ , 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 multiple access channel (MAC) in network information theory.^{1,2} In traditional MAC analysis, each user can employ a capacity-achieving code. In contrast, this type of coding is not possible with on–off signaling. In addition, “uncoded” multiuser detection results 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.^{3,4}

Our analysis is instead based on identifying a connection between the on–off random access channel and the recovery of the sparsity pattern of a signal from noisy random linear measurements. The feasibility of recovering sparse, approximately sparse, or compressible signals from a relatively small number of random linear measurements has recently been termed *compressed sensing*.^{5–7} Results in compressed sensing generally provide bounds on the ℓ^2 estimation error of a signal as a function of the number of measurements, the signal sparsity

and other factors. However, what is relevant for the random on-off multiple access channel is detecting the *positions* of the nonzero entries. This problem is addressed in several recent works.⁸⁻¹²

By exploiting recent compressed sensing results and providing an analysis of a new algorithm, we are able to provide a number of insights:

- *Performance bounds with ML detection:* Recent results provide simple upper and lower bounds on the number of degrees of freedom required to detect the users reliably assuming maximum likelihood (ML) detection.^{9,11-13} One of the consequences of these bounds is that, unlike in the classic MAC analysis, the sum rate achievable with random access signaling can be strictly less than the rate achievable with coordinated transmissions with the same total power.
- *Potential gains over single-user detection:* ML detection can be considered as a type of multiuser detection. Current commercial designs, however, almost universally use simple single-user detection. The single-user detection performance can be estimated by bounds for thresholding.^{11,14} The bounds show that ML detection offers a potentially large gain over single-user detection, particularly at high SNRs. The gap at high SNRs can be explained by a certain *multiple access interference* limit experienced by single-user detection.
- *Lasso- and OMP-based multiuser detection and near-far resistance:* ML sparsity detection is a well-known NP-hard problem.¹⁵ However, there are practical, but suboptimal, algorithms such as the orthogonal matching pursuit (OMP)¹⁶⁻¹⁹ and lasso²⁰ methods in sparse estimation that can be used for multiuser detection for on-off random access channels. In comparison to single-user detection, we show that these methods can offer improved performance when the dynamic range in received power levels is large. This near-far resistance feature is similar to that of standard MMSE multiuser detection in CDMA systems.²¹
- *Improved high SNR performance with power shaping:* While both lasso and OMP offer improvements over single-user detection, there is still a large gap in the performance of these algorithms in comparison to ML detection at high SNRs. Specifically, at high SNRs, ML achieves a fundamentally different scaling in the number of degrees of freedom required for reliable detection than that required by lasso, OMP and single-user detection.

We show, however, that when accurate power control is available, the ML scaling can be theoretically achieved with a simplified version of OMP, which we call sequential OMP (SeqOMP). The method is analogous to the classic successive interference cancellation (SIC) method for the MAC. Specifically, users are deliberately targeted at different received power levels and then detected and cancelled out in descending order of power.

While SeqOMP shows significant gains over single-user detection, for most practical problem sizes it does worse than standard OMP, even without power shaping. However, we show, at least by simulation, that power shaping can improve the performance of OMP as well.

The connection between sparsity detection methods such as OMP and the SIC technique for the MAC has also been observed in the recent work of Jin and Rao.²² A related work by Wipf and Rao also gave some empirical evidence for the benefit of power shaping when used in conjunction with sparse Bayesian learning algorithms.²³ The results in this paper make the connections between sparsity detection and the random access MAC more precise by giving conditions on the detectability of the sparsity pattern and characterizing the optimal power shaping distribution.

The remainder of the paper is organized as follows. The setting is formalized in Section 2. In particular, we define all the key problem parameters. Results that can be derived from existing necessary and sufficient conditions for sparsity pattern recovery are then presented in Section 3. We will see that there is a potentially-large performance gap between single-user detection and the optimal ML detection. Existing “practical” multiuser detection techniques perform significantly better than single-user detection in that they are near-far resistant. However, their performance saturates at high SNRs, falling well short of ML detection. Section 4 presents a new

detection algorithm, sequential orthogonal matching pursuit (SeqOMP), that has near-far resistance under certain assumptions on power control. Furthermore, with optimal power shaping, it does not suffer from saturation at high SNRs. Conclusions are given in Section 5. The reader is referred to a longer document for numerical experiments, connections to MAC capacity, and a proof of the main theoretical result.²⁴

2. ON-OFF RANDOM ACCESS CHANNEL MODEL

2.1 Problem Formulation

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 λ 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 by making x_j a scalar we have assumed flat fading. The total signal at the receiver is given by

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

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

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

which is the “true” set of active users. The size of the active user set is related to the activity ratio through

$$\lambda = \frac{1}{n} |I_{\text{true}}|. \quad (3)$$

The goal of the receiver is to determine an estimate $\hat{I} = \hat{I}(\mathbf{y})$ of I_{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 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,

$$p_{\text{err}} = \Pr \left(\hat{I} \neq I_{\text{true}} \right), \quad (4)$$

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_{err} close to zero. An alternative could be to allow a nonzero constant fraction of detection errors. This may change scaling behavior considerably.^{25,26}

We identify two factors that influence the ability to detect the active user set. The first is the total SNR:

$$\text{SNR} = \mathbf{E} [\|\mathbf{A} \mathbf{x}\|^2] / \mathbf{E} [\|\mathbf{w}\|^2]. \quad (5)$$

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

$$\text{SNR} = \|\mathbf{x}\|^2. \quad (6)$$

In the case of random \mathbf{x} , this is the conditional SNR given \mathbf{x} .

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

$$\text{MAR} = \frac{\min_{j \in I_{\text{true}}} |x_j|^2}{\|\mathbf{x}\|^2 / \lambda n}. \quad (7)$$

Since I_{true} has λn 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.

One final value that will be important is the *minimum component SNR*, which, for a given \mathbf{x} , is given by

$$\text{SNR}_{\min} = \frac{1}{\mathbf{E}\|\mathbf{w}\|^2} \min_{j \in I_{\text{true}}} \mathbf{E}\|\mathbf{a}_j x_j\|^2 = \min_{j \in I_{\text{true}}} |x_j|^2, \quad (8)$$

where \mathbf{a}_j is the j th column of \mathbf{A} . The quantity SNR_{\min} has a natural interpretation: The numerator is the signal power due to the smallest nonzero component in \mathbf{x} while the denominator is the total noise power. The ratio SNR_{\min} thus represents the contribution to the SNR from the smallest nonzero component of \mathbf{x} .

The final equality in (8) is a consequence of $\mathbf{E}\|\mathbf{a}_j\|^2 = \mathbf{E}\|\mathbf{w}_j\|^2 = 1$. Observe that (6) and (7) show

$$\text{SNR}_{\min} = \min_{j \in I_{\text{true}}} |x_j|^2 = \frac{1}{\lambda n} \text{SNR} \cdot \text{MAR}. \quad (9)$$

2.2 MAR and Power Control

For wireless systems, the factor MAR in (7) has an important interpretation as a measure of the dynamic range of received power levels. With accurate power control, all users can be controlled to arrive at the same power. In this case, $\text{MAR} = 1$. However, if power control is difficult due to fading or lack of power control feedback, there can be a considerable dynamic range in the received powers from different users. In this case, some users could arrive at powers much below the average making MAR closer to zero.

One of the results in this paper is a precise quantification of the effect of MAR on the detectability of the active user set. Specifically, 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, where users with weak signals can be dominated by higher-power signals.²¹

2.3 Synchronization and Multi-Path

It is important to recognize that an implicit assumption in the above model is that the transmissions from different users are perfectly synchronized. At a minimum, the timing offsets from the users are exactly known at the receiver and there is no multipath.

Of course, in many wireless applications, exact synchronization is not possible and the receiver must estimate the timing delay of the transmission as part of the detection process. In most practical receivers, timing offsets are estimated by discretizing the delay search space, typically to a quarter or half-chip resolution. The receiver then searches over a finite set of delay hypotheses depending on the range of timing uncertainty. In the presence of multipath, the receiver could detect multiple delay hypotheses.

To model this search in the theoretical framework of this paper, we would need to model each timing shift of the codeword as a different codeword. The total number of codewords would then grow to the number of users times the number of delay hypotheses per user. While the algorithms we will present can be applied in this manner to deal with the asynchronous case, there are several theoretical issues with extending the analysis. In particular, this extended codebook would lack the independence of codewords that the simpler model has by construction. We will thus just consider only the synchronous case for the remainder of this paper.

	finite SNR · MAR	SNR · MAR → ∞
Necessary for ML	$m > \frac{1}{\text{MAR} \cdot \text{SNR}} \lambda n \log(n(1 - \lambda))$ Fletcher <i>et al.</i> ¹¹	$m > \lambda n$ (elementary)
Sufficient for sequential OMP with power shaping	$m > \frac{4}{\log(1 + \text{SNR})} \lambda n \log(n(1 - \lambda))$ From Theorem 4.1 (Section 4.5)	$m > 5\lambda n$ From Theorem 4.1 (Section 4.6)
Sufficient for lasso	(more subtle) Wainwright ²⁸	$m > \lambda n \log(n(1 - \lambda))$ Wainwright ^{8,28}
Sufficient for OMP	unknown	$m > \lambda n \log(n(1 - \lambda))$ Fletcher and Rangan ²⁷
Sufficient for single-user detection (12)	$m > \frac{4(1 + \text{SNR})}{\text{MAR} \cdot \text{SNR}} \lambda n \log(n(1 - \lambda))$ Fletcher <i>et al.</i> ¹¹	$m > \frac{4}{\text{MAR}} \lambda n \log(n(1 - \lambda))$

Table 1. 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.

3. PERFORMANCE WITH CURRENT SPARSITY DETECTION METHODS

The problem of detecting the active user set is precisely equivalent to a sparsity pattern recovery 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 noisy linear observations \mathbf{y} in (1).

In this section, we develop asymptotic analyses for detection of the active users based on previous results on sparsity pattern recovery. We model \mathbf{x} as deterministic, so the quantities λ , SNR, MAR and SNR_{\min} are also deterministic. Since our formulation allows simple translation of previous results, we state these translations without detailed justifications.^{8,9,11,27} Several results are here adjusted by a factor of two because we have complex, rather than real, degrees of freedom.

Our results are expressed as scaling laws on the number of degrees of freedom for *asymptotic reliable detection* of the active user set. We define this as follows:

DEFINITION 3.1. *Suppose that we are given deterministic sequences $m = m(n)$ and $\mathbf{x} = \mathbf{x}(n) \in \mathbb{C}^n$ that vary with n . For a given detection algorithm $\hat{I} = \hat{I}(\mathbf{y})$, we then define the probability of error p_{err} in (4) where the probability is taken over the randomness of the codebook \mathbf{A} and the noise vector \mathbf{w} . Given the number of degrees of freedom $m(n)$ and modulation vector $\mathbf{x}(n)$, the probability of error will then simply be a function of n . We say that the detection algorithm achieves asymptotic reliable detection when $p_{\text{err}}(n) \rightarrow 0$.*

Table 1 summarizes the results from this section and previews results from Section 4.

3.1 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 λ , \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 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 $\mathbf{y} = \mathbf{A}\mathbf{x}$ 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} .

3.2 ML Detection with Noise

Now suppose there is noise. Since \mathbf{x} is an unknown deterministic quantity, the probability of error in detecting the active user set is minimized by maximum likelihood (ML) detection. 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.⁹ The results in that work, along with the fact that $k = \lambda n$, show that there exists a constant $C > 0$ such that if

$$m \geq C \max \left\{ \frac{1}{\text{MAR} \cdot \text{SNR}} \lambda n \log(n(1 - \lambda)), \lambda n \log(1/\lambda) \right\} = C \max \left\{ \frac{1}{\text{SNR}_{\min}} \log(n(1 - \lambda)), \lambda n \log(1/\lambda) \right\} \quad (10)$$

then ML will asymptotically detect the correct active user set. The equivalence of the two expressions in (10) is due to (9). Also, Fletcher *et al.*¹¹ show that, for any $\delta > 0$, the condition

$$m \geq \frac{1 + \delta}{\text{MAR} \cdot \text{SNR}} \lambda n \log(n(1 - \lambda)) + \lambda n = \frac{1 + \delta}{\text{SNR}_{\min}} \log(n(1 - \lambda)) + \lambda n \quad (11)$$

is necessary. Observe that when $\text{SNR} \cdot \text{MAR} \rightarrow \infty$, the lower bound (11) approaches $m \geq \lambda n$, matching the noise-free case as expected.

The necessary condition for ML appears in Table 1 with smaller terms and the infinitesimal δ omitted for simplicity.

3.3 Single-User Detection

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

$$\hat{I}_{\text{SUD}} = \{ j : \rho(j) > \mu \}, \quad (12)$$

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

$$\rho(j) = \frac{|\mathbf{a}'_j \mathbf{y}|^2}{\|\mathbf{a}_j\|^2 \|\mathbf{y}\|^2}. \quad (13)$$

Single-user detection has been analyzed in the compressed sensing context in several recent works.^{11,14,29} Fletcher *et al.*¹¹ show the following result: Suppose

$$m(n) > \frac{(1 + \delta)L(\lambda, n)(1 + \text{SNR})}{\text{SNR} \cdot \text{MAR}} \lambda n = \frac{(1 + \delta)L(\lambda, n)(1 + \text{SNR})}{\text{SNR}_{\min}} \quad (14)$$

where $\delta > 0$ and

$$L(\lambda, n) = \left[\sqrt{\log(n(1 - \lambda))} + \sqrt{\log(n\lambda)} \right]^2. \quad (15)$$

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. As before, the equivalence of the two expressions in (14) is due to (9).

Comparing the sufficient condition (14) for single-user detection with the *necessary* condition (11), we see two distinct problems in single-user detection:

- *Constant offset:* The scaling (14) for single-user detection shows a factor $L(\lambda, n)$ instead of $\log((1 - \lambda)n)$ in (11). It is easily verified that, for $\lambda \in (0, 1/2)$,

$$\log((1 - \lambda)n) < L(\lambda, n) < 4 \log((1 - \lambda)n), \quad (16)$$

so this difference in factors alone could require that single-user detection use up to four times more degrees of freedom than ML for asymptotic reliable detection.

Combining the inequality (16) with (14), we see that the more stringent, but simpler, condition

$$m(n) > \frac{(1 + \delta)4(1 + \text{SNR})}{\text{SNR} \cdot \text{MAR}} \lambda n \log((1 - \lambda)n) \quad (17)$$

is also sufficient for asymptotic reliable detection with single-user detection. This simpler condition is shown in Table 1, where we have omitted the infinitesimal δ quantity to simplify the table entry.

- *Multiple access interference limit:* In addition to the $L(\lambda, n)/\log(n(1 - \lambda))$ offset, single-user detection also requires a factor of $1 + \text{SNR}$ more degrees of freedom than ML. This $1 + \text{SNR}$ factor has a natural interpretation as *multiple access interference*: When detecting any one component of the vector \mathbf{x} , single-user detection sees the energy from the other $n - 1$ components of the signal as interference. We can think of this additional noise as multiple access interference, by which we mean the interference caused from different components of the signal \mathbf{x} interfering with one another in the observed signal \mathbf{y} through the codebook matrix \mathbf{A} . This multiple access interference is distinct from the additive noise \mathbf{w} . It 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.

Multiple access interference results in a large performance gap at high SNRs. In particular, as $\text{SNR} \rightarrow \infty$, (14) reduces to

$$m(n) > \frac{(1 + \delta)L(\lambda, n)}{\text{MAR}} \lambda n \log((1 - \lambda)n). \quad (18)$$

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.

3.4 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,²⁰ also called basis pursuit denoising.³⁰ In the context of the random access channel, the lasso estimator would first estimate the modulation vector \mathbf{x} by solving the convex minimization

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} (\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mu \|\mathbf{x}\|_1), \quad (19)$$

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

The exact performance of lasso is not known at finite SNR. However, Wainwright⁸ has exactly characterized the conditions for lasso to work at high SNR. Specifically, if m , n and $\lambda n \rightarrow \infty$, with $\text{SNR} \cdot \text{MAR} \rightarrow \infty$, the scaling

$$m > \lambda n \log(n(1 - \lambda)) + \lambda n + 1, \quad (20)$$

is necessary and sufficient for asymptotic sparsity recovery.

Another common approach to sparsity pattern detection is the greedy OMP algorithm.^{16, 18, 19} This has been analyzed by Fletcher and Rangan in a setting with SNR approaching infinity.²⁷ They show that, when \mathbf{A} has Gaussian entries, a *sufficient* condition for asymptotic reliable recovery is

$$m > \lambda n \log(n - k), \quad (21)$$

identical to the lasso scaling law. Variants of OMP improve this scaling.^{31,32}

The conditions (20) and (21) are both shown in Table 1. As usual, the table entries are simplified by including only the leading terms.

The lasso and OMP scaling laws, (20) and (21), can be compared with the high SNR limit for the single-user detection scaling law in (18). This comparison shows the following:

- *Removal of the constant offset:* The $L(\lambda, n)$ factor in the single-user detection expression (18) is replaced by a $\log(n(1 - \lambda))$ factor in the lasso scaling law (20) and OMP scaling law (21). Similar to the discussion above, this implies that lasso and OMP could require up to 4 times fewer degrees of freedom than single-user detection.
- *Near-far resistance:* In addition, both the lasso and OMP methods do not have a dependence on MAR; thus, in the high SNR regime, they have a near-far resistance that single-user detection does not. This gain can be large when there are users whose received powers are much below the average (low MAR).

The near-far resistance of lasso and OMP is analogous to that of MMSE multiuser detection in CDMA systems.²¹ In that case, when the number of degrees of freedom m exceeds the number of users n , a decorrelating detector can null out strong users while recovering weak ones. An interesting property that we see in the random access case is that near-far resistance may be possible when $m < n$, provided that m is sufficiently greater than the number of *active users*, λn .

- *Limits at high SNR:* We also see from (20) and (21) 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 lasso and OMP have the scaling, $m = O(\lambda n \log((1 - \lambda)n))$, similar to the minimum scaling possible with single-user detection, which suffers from a multiple access interference limit.

3.5 Other Sparsity Detection Algorithms

Recent interest in compressed sensing has led to a plethora of algorithms beyond OMP and lasso. Empirical evidence suggests that the most promising algorithms for sparsity pattern detection are the sparse Bayesian learning methods developed in the machine learning community by Tipping³³ and introduced into signal processing applications by Wipf and Rao,³⁴ with related work by Schniter *et al.*³⁵ Unfortunately, a comprehensive summary of these algorithms is far beyond the scope of this paper.

Instead, we will limit our discussion to the lasso and OMP methods since these are the algorithms with the most concrete analytic results on asymptotic reliable detection. Moreover, our interest is not in finding the optimal algorithm, but merely to point out general qualitative effects such as near-far and multiple access interference limits which should be considered in evaluating any algorithm.

4. SEQUENTIAL ORTHOGONAL MATCHING PURSUIT

The analyses in the previous section suggest that ML detection may offer significant gains over the provable performance of current “practical” algorithms such as single-user detection, lasso and OMP, when the SNR is high. Specifically, as the SNR increases, the performance of these practical methods saturates at a scaling in the number of degrees of freedom that can be significantly higher than that for ML.

In this section, we show that if accurate power control is available, an OMP-like algorithm, which we call *sequential orthogonal matching pursuit* or SeqOMP, can break this barrier. Specifically, the performance of SeqOMP does not saturate at high SNR.

4.1 Algorithm

ALGORITHM 1 (SEQOMP). Given a received vector \mathbf{y} and threshold level $\mu > 0$, the algorithm produces an estimate \hat{I}_{SOMP} of the active user set with the following steps:

1. Initialize the counter $j = 1$ and set the initial active user set estimate to empty: $\hat{I}(0) = \{\emptyset\}$.
2. Compute $\mathbf{P}(j)\mathbf{a}_j$ where $\mathbf{P}(j)$ is the projection operator onto the orthogonal complement of the span of $\{\mathbf{a}_\ell, \ell \in \hat{I}(j-1)\}$.
3. Compute the correlation,

$$\rho(j) = \frac{|\mathbf{a}_j' \mathbf{P}(j)\mathbf{y}|^2}{\|\mathbf{P}(j)\mathbf{a}_j\|^2 \|\mathbf{P}(j)\mathbf{y}\|^2}. \quad (22)$$

4. If $\rho(j) > \mu$, add the index j to $\hat{I}(j-1)$. That is, $\hat{I}(j) = \hat{I}(j-1) \cup \{j\}$. Otherwise, set $\hat{I}(j) = \hat{I}(j-1)$.
5. Increment $j = j + 1$. If $j \leq n$ return to step 2.
6. The final estimate of the active user set is $\hat{I}_{\text{SOMP}} = \hat{I}(n)$.

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 except that SeqOMP passes through the data only once. For this reason, SeqOMP is actually computationally simpler than standard OMP.

As simulations in our longer document²⁴ illustrate, SeqOMP generally has much worse performance than standard OMP. It is not intended as a competitive practical alternative. Our interest in the algorithm lies in the fact that we can prove positive results for SeqOMP. Specifically, we will be able to show that this relatively poor algorithm, when used in conjunction with *power shaping*, can achieve a fundamentally better scaling at high SNRs than what has been proven is achievable with methods such as OMP.

4.2 Sequential OMP Performance

The analysis in Section 3 was based on deterministic vectors \mathbf{x} . To characterize the SeqOMP performance, it is simpler to use a partially-random model where the active user set is random while the received modulation signal power $|x_j|^2$, conditioned on user j being active, remains deterministic. We reuse the notation λ because its meaning remains almost the same.

We assume that each user is active with some probability $\lambda \in (0, 1)$, which we now call the *activity probability*. The activities of different users are assumed to be independent. Thus, unlike in Section 3, λn represents the *average* number of users that are active, as opposed to the actual number.

Let p_j denote the received modulation symbol power

$$p_j = |x_j|^2, \quad (23)$$

conditional that user j is active. We will call the set $\{p_j\}_{j=1}^n$ the *power profile*, which we will treat as a deterministic quantity. Since each user transmits with a probability λ , the total average SNR is given by

$$\text{SNR} = \lambda \sum_{j=1}^n p_j. \quad (24)$$

This factor is also deterministic.

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

$$\gamma = \min_{\ell=1, \dots, n} p_\ell / \hat{\sigma}^2(\ell), \quad (25)$$

where $\hat{\sigma}^2(\ell)$ is given by

$$\hat{\sigma}^2(\ell) = 1 + \lambda \sum_{j=\ell+1}^n p_j. \quad (26)$$

The parameters γ and $\hat{\sigma}^2(\ell)$ have simple interpretations: Suppose that the SeqOMP algorithm has correctly decoded all the users for $j < \ell$. Then, in detecting the ℓ th user, the receiver sees the noise \mathbf{w} with power $\mathbf{E}\|\mathbf{w}\|^2 = 1$ and, for each user $j > \ell$, an interference power p_j with probability λ . Hence, $\hat{\sigma}^2(\ell)$ is the total average interference power seen when detecting ℓ th user, assuming perfect cancellation. Since user ℓ arrives at a power p_ℓ , the ratio $p_\ell/\hat{\sigma}^2(\ell)$ in (25) represents the average SINR seen by user ℓ . The value γ is the minimum SINR over all n users.

THEOREM 4.1. *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 $m - \lambda n$, λn and $(1 - \lambda)n \rightarrow \infty$, and $\gamma \rightarrow 0$. Also, assume the sequence of power profiles satisfies the limit*

$$\lim_{n \rightarrow \infty} \max_{i=1, \dots, n-1} \log(n) \hat{\sigma}^{-4}(i) \sum_{j>i}^n p_j^2 = 0. \quad (27)$$

Finally, assume that for all n ,

$$m \geq \frac{(1 + \delta)L(n, \lambda)}{\gamma} + \lambda n, \quad (28)$$

for some $\delta > 0$ and $L(n, \lambda)$ defined in (15). Then, there exists a sequence of thresholds, $\mu = \mu(n)$, such that SeqOMP will achieve asymptotic reliable detection of the active user set in that

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

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 provides a simple sufficient condition on the number of degrees of freedom as a function of the SINR γ , activity probability λ and number of users n . The condition (27) is somewhat technical, but is satisfied in the cases that interest us. The remainder of this section will discuss some of the implications of this theorem.

4.3 Near-Far Resistance with Known Power Ordering

First, 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 $p_j \rightarrow \infty$ for all j , then it can be verified that $\gamma > 1/(\lambda n)$. In this case, the sufficiency of the scaling (28) shows that

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

is sufficient for asymptotic reliable detection. This is identical to the lasso performance except for the factor $L(\lambda, n)/\log((1 - \lambda)n)$, which lies in $(0, 4)$ for $\lambda \in (0, 1/2)$. In particular, the minimum number of degrees of freedom does not depend on MAR; therefore, similar to lasso and OMP, SeqOMP can theoretically detect users even when they are much below the average power.

With SeqOMP, simply knowing the order of powers is sufficient to achieve near-far resistance when the SNR is sufficiently high. Unlike for single-user detection, unequal received powers do not hurt the performance of SeqOMP, as long as the order of the powers are known at the receiver. The feasibility of knowing the power ordering is addressed in Section 4.9 below. We will now look at the effect of the power profile on the performance.

4.4 Performance with Constant Power

Consider the case when all the powers p_j are equal. To satisfy the constraint (24), the constant power level must be $p_j = \text{SNR}/(\lambda n)$. From (25), the minimum SINR is $\gamma = \gamma_{\text{const}}$, where

$$\gamma_{\text{const}} = \frac{\text{SNR}}{\lambda(n + (n - 1)\text{SNR})} \approx \frac{\text{SNR}}{\lambda n(1 + \text{SNR})}, \quad (29)$$

and the approximation holds for large n .

It can be verified that the constant power profile satisfies the technical condition (27) provided λ is bounded away from zero and the SNR does not grow too fast. Specifically, the SNR must satisfy $\text{SNR} = o(n/\log(n))$. In this case, we can substitute $\gamma = \gamma_{\text{const}}$ in (28) to obtain the condition

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

for asymptotic reliable detection. The condition is precisely the condition for single-user detection in (17) with $\text{MAR} = 1$ and an additional λn term.

Thus, for a constant power profile, Theorem 4.1 does not show any benefit in using SeqOMP.

4.5 Optimal Power Shaping

The constant power profile, however, is not optimal. 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 γ in (25) for a given total SNR constraint (24). It is easily verified that any power profile p_j maximizing the SINR γ in (25) will satisfy

$$p_\ell = \gamma \left(1 + \lambda \sum_{j=\ell+1}^n p_j \right) \quad (30)$$

for all $\ell = 1, \dots, n$. The solution to (30) and (24) is given by

$$p_\ell = \gamma(1 + \lambda\gamma)^{n-\ell}, \quad (31)$$

where $\gamma = \gamma_{\text{opt}}$ is the SINR,

$$\gamma_{\text{opt}} = \lambda^{-1} \left[(1 + \text{SNR})^{1/n} - 1 \right] \approx (\lambda n)^{-1} \log(1 + \text{SNR}). \quad (32)$$

Here, the approximation holds for large n . Again, some algebra shows that, when λ is bounded away from zero, the power profile p_j in (31) will satisfy the technical condition (27) when $\log(1 + \text{SNR}) = o(n/\log(n))$.

The power profile (31) is exponentially decreasing in the index order ℓ . Thus, users early in the detection sequence are allocated exponentially higher power than users later in the sequence. This allocation insures that early users have sufficient power to overcome the interference from all the users later in the detection sequence that are not yet cancelled. This power shaping is analogous to the optimal power allocations in the classic MAC when using a SIC receiver.²

The ratio of the optimal SINR γ_{opt} to the SINR with a constant power profile γ_{const} is given by

$$\frac{\gamma_{\text{opt}}}{\gamma_{\text{const}}} = \frac{(1 + \text{SNR}) \log(1 + \text{SNR})}{\text{SNR}}.$$

This ratio represents the potential increase in SINR with exponential power shaping relative to the SINR with equal power for all users. The ratio increases with SNR and can be large when the SNR is high. For example, when $\text{SNR} = 10$ dB, $\gamma_{\text{opt}}/\gamma_{\text{const}} \approx 2.6$. When $\text{SNR} = 20$ dB, the gain is even higher at $\gamma_{\text{opt}}/\gamma_{\text{const}} \approx 4.7$.

Based on Theorem 4.1, this gain in SINR will result in a proportional decrease in the minimum number of degrees of freedom. Specifically, if we substitute the SINR γ_{opt} in (32) into (28), we see that that the condition

$$m \geq \frac{(1 + \delta)L(n, \lambda)}{\log(1 + \text{SNR})} \lambda n + \lambda n \quad (33)$$

is sufficient for SeqOMP to achieve asymptotic reliable detection of the active users, when the users use exponential power shaping (31).

As before, if $\lambda < 1/2$ then $L(n, \lambda) < 4 \log(n(1 - \lambda))$ and the sufficient condition (33) can be simplified to

$$m \geq \frac{4(1 + \delta) \log(n(1 - \lambda))}{\log(1 + \text{SNR})} \lambda n + \lambda n, \quad (34)$$

the leading term of which appears in Table 1 with the δ omitted.

4.6 SNR Saturation

As discussed earlier, a major problem with both single-user detection and lasso multiuser detection was that their performance “saturates” with high SNR. That is, even as the SNR scales to infinity, the minimum number of degrees of freedom scales as $m = O(\lambda n \log((1 - \lambda)n))$. In contrast, optimal ML detection can achieve a scaling $m = O(\lambda n)$ when the SNR is sufficiently high.

An important consequence of (33) is that SeqOMP with exponential power shaping can overcome this bound. Specifically, if we take the scaling of $\text{SNR} = \Theta(\lambda n)$ in (34) and assume that λ is bounded away from zero we see that asymptotically, SeqOMP requires only $m \geq 5\lambda n$ degrees of freedom. In this way, unlike single-user and lasso detection, SeqOMP is able to obtain the scaling $m = O(\lambda n)$ when the $\text{SNR} \rightarrow \infty$.

4.7 Power Shaping with Sparse Bayesian Learning

The fact that power shaping can provide benefits when combined with certain iterative detection algorithms confirms the observations in the work of Wipf and Rao.²³ That work considers signal detection with a certain sparse Bayesian learning (SBL) algorithm. They show the following result: Suppose \mathbf{x} has k non-zero components and p_i , $i = 1, 2, \dots, k$, is the power of the i th largest component. Then, for a given codebook matrix \mathbf{A} , there exist constants $\nu_i > 1$ such that if

$$p_i \geq \nu_i p_{i-1}, \quad (35)$$

the SBL algorithm will correctly detect the sparsity pattern of \mathbf{x} .

The condition (35) shows that a certain growth in the powers can guarantee correct detection. The parameters ν_i however depend in some complex manner on the matrix \mathbf{A} , so the appropriate growth is difficult to compute. They also provide strong empirical evidence that shaping the power with certain profiles can greatly reduce the number of degrees of freedom needed.

The results in this paper add to Wipf and Rao’s observations showing that growth in the powers can also assist sequential OMP. Moreover, for the SeqOMP case, we can explicitly derive the optimal power profile for certain large random codebook matrices.

This is not to say that SeqOMP is better than SBL. In fact, empirical results of Wipf and Rao suggest that SBL will outperform OMP, which will in turn do better than SeqOMP.³⁴ As we have stressed before, the point here of analyzing SeqOMP is that we can easily derive concrete analytic results. These results may provide guidance for more sophisticated algorithms.

4.8 Robust Power Shaping

The above analysis shows certain benefits of SeqOMP used in conjunction with power shaping. However, these gains are theoretically only possible at infinite block lengths. Unfortunately, when the block length is finite, power shaping can actually reduce the performance.

The problem is that when an active user is not detected in SeqOMP, the user’s energy is not cancelled out and remains as interference for all subsequent users in the detection sequence. With power shaping, users early in the detection sequence have much higher power than users later in the sequence, so missing an early user can make the detection of subsequent users difficult. At infinite block lengths, the probability of missing an active user can be driven to zero. But, at finite block lengths, the probability of missing an active user early in the sequence will always be nonzero, and therefore a potential problem with power shaping.

Agrawal *et al.* observed a similar problem when SIC is used in a CDMA uplink.³⁶ To mitigate the problem, they proposed to adjust the power allocations to make them more robust to decoding errors early in the decoding sequence. The same technique, which we will call *robust power shaping*, can be applied to the SeqOMP as follows.

In the condition (30), it is assumed that all the energy of users with index $j < \ell$ have been correctly detected and subtracted. But, following Agrawal *et al.*, suppose that on average some fraction $\theta \in [0, 1]$ of the energy of users early in the detection sequence is not cancelled out due to missed detections. We will call θ the *leakage fraction*. With nonzero leakage, the condition (30) would be replaced by

$$p_\ell = \gamma \left(1 + \theta \lambda \sum_{j=1}^{\ell-1} p_j + \lambda \sum_{j=\ell+1}^n p_j \right). \quad (36)$$

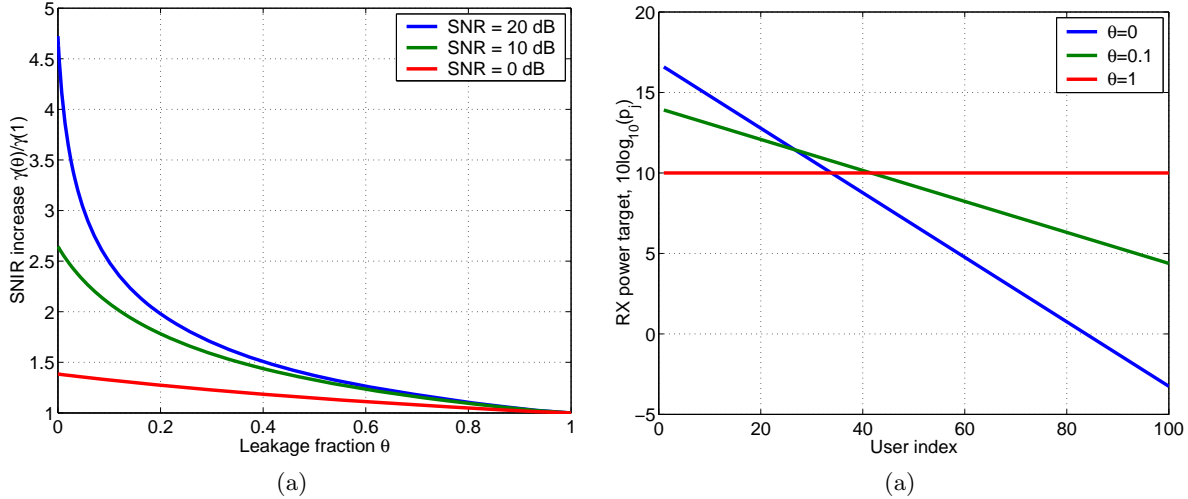


Figure 1. (a) Increase in SINR, $\gamma(\theta)/\gamma_{\text{const}} = \gamma(\theta)/\gamma(1)$, as a function of the leakage fraction θ . (b) Robust power shaping: Power profiles for various leakage values θ . For all the curves, the number of users is $n = 100$, SNR = 20 dB, and activity probability is $\lambda = 0.1$.

For given SNR, θ and λ , the linear equations (24) and (36) can be solved to obtain the optimal power profile, given by

$$p_j = \frac{(1 - \theta)\gamma}{1 + \lambda\theta\gamma} \left(\frac{1 + \lambda\gamma}{1 + \lambda\theta\gamma} \right)^{n-j}, \quad (37)$$

where $\gamma = \gamma(\theta)$ is the optimal SINR given by

$$\gamma(\theta) = \frac{1}{\lambda(1 - \theta)} \left[\left(\frac{1 + \text{SNR}}{1 + \theta \text{SNR}} \right)^{1/n} - 1 \right] \approx \frac{1}{\lambda n(1 - \theta)} \log \left(\frac{1 + \text{SNR}}{1 + \theta \text{SNR}} \right). \quad (38)$$

The approximation here is valid for large n .

Fig. 1(a) plots the SINR, $\gamma(\theta)$, as a function of the leakage fraction θ . The SINR is plotted relative to γ_{const} in (29), which is the SINR that one obtains with a constant power profile. The increase in SINR is maximized when the leakage fraction, $\theta = 0$. When $\theta = 0$, $\gamma(\theta) = \gamma_{\text{exp}}$, the SINR (32) for the exponential power shaping. This is the optimal SINR, but assumes that there are no missed detections.

As the leakage fraction θ is increased, the SINR $\gamma(\theta)$ decreases, which is price for the robustness to missed detections. In the limit as $\theta \rightarrow 1$, the optimal power profile, p_j in (37), approaches a constant and the corresponding SINR converges to γ_{const} . However, even at a reasonable leakage fraction, say $\theta = 0.1$, the SINR $\gamma(\theta)$ can still be significantly larger than γ_{const} .

It is illustrative to actually look at the optimal power profiles as a function of θ . Fig. 1(b) plots the optimal power profile, p_j in (37), for leakage values of $\theta = 0, 0.1$ and 1 . In the plot, $n = 100$, SNR = 20 dB, and $\lambda = 0.1$. It can be seen that when $\theta = 0$, there is a large range of almost 20 dB in the target receive powers from the first to last user. While this power profile is optimal when there are no missed detections, the power allocations can be very damaging if an active user is missed. In an extreme case, for example, if the first user is active but not detected and not cancelled it will cause an interference level 20 dB above the signal level of the last user. As the leakage fraction θ is increased, the range of powers is decreased, which improves the robustness to missed detection at the expense of reduced SINR.

4.9 Practical Power Control Considerations

In the original description of the problem in Section 2, we said that we would restrict our attention to estimators that do not require *a priori* knowledge of the modulation vector \mathbf{x} . However, although SeqOMP does not require

knowledge at the receiver of the channel phases, the above analysis shows that knowledge of the *order* of the conditional received powers is necessary to achieve near-far resistance. Additionally, eliminating the multiple access interference limit requires that powers are explicitly targeted to a certain profile.

The use of power control for on-off random access communication requires some justification. On-off random access signaling is most likely to be used when the users do not already have some ongoing communication. For example, in cellular systems, it is used for initial access or requests to transmit. If the users were already transmitting, the one bit could be embedded in the other communication and on-off random access signaling would not be needed. Consequently, fast feedback power control would likely not be available for such on-off random access transmissions since the users are not likely to have a continuous transmission to measure the received power. Thus, in practice, power control is likely achievable only by open-loop methods. Open-loop power control is used for example in cellular systems where each mobile estimates the path loss in the downlink and adjusts its access power appropriately in the uplink. Open-loop power control is most accurate when the uplink and downlink are time-division duplexed (TDD) in the same band.

5. CONCLUSIONS

Sparse signal detection is a valuable framework for understanding on-off random access signaling. Results can provide simple capacity estimates and clarify the role of power control and multiuser detection. Methods such as OMP and lasso, which are widely used in sparse signal detection problems, can be applied as multiuser detection methods for on-off random access channels. Analysis shows that these methods may offer improved near-far resistance over single-user detection in high SNRs. Optimal ML detection may theoretically offer further gains in the high SNR regime, but is not computationally feasible. However, some gains at high SNR may be practically achievable through power shaping and SIC-like techniques such as OMP.

ACKNOWLEDGMENTS

The authors thank Martin Vetterli for his support, wisdom, and encouragement. The authors also thank Gerhard Kramer for helpful comments on an early draft of their related longer manuscript.²⁴ This work was supported in part by NSF CAREER Grant CCF-643836.

REFERENCES

1. R. Ahlswede, "Multi-way communication channels," in *Proc. IEEE Int. Symp. Inform. Th.*, pp. 23–52, (Armenian S.S.R.), Sept. 1971.
2. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.
3. S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channel," *IEEE Trans. Inform. Theory* **32**, pp. 85–96, Jan. 1986.
4. S. Verdú, *Multiuser Detection*, Cambridge University Press, New York, 1998.
5. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory* **52**, pp. 489–509, Feb. 2006.
6. D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory* **52**, pp. 1289–1306, Apr. 2006.
7. E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inform. Theory* **52**, pp. 5406–5425, Dec. 2006.
8. M. J. Wainwright, "Sharp thresholds for high-dimensional and noisy recovery of sparsity," tech. rep., Univ. of California, Berkeley, Dept. of Statistics, May 2006. arXiv:math.ST/0605740 v1 30 May 2006.
9. M. J. Wainwright, "Information-theoretic limits on sparsity recovery in the high-dimensional and noisy setting," Tech. Rep. 725, Univ. of California, Berkeley, Dept. of Statistics, Jan. 2007.
10. J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inform. Theory* **53**, pp. 4655–4666, Dec. 2007.
11. A. K. Fletcher, S. Rangan, and V. K. Goyal, "Necessary and sufficient conditions for sparsity pattern recovery," *IEEE Trans. Inform. Theory* **2009**, Dec. 2009. To appear. Original submission available online.³⁷
12. S. Aeron, M. Zhao, and V. Saligrama, "Information theoretic bounds to performance of compressed sensing and sensor networks." arXiv:0804.3439v4 [cs.IT]., July 2009.

13. W. Wang, M. J. Wainwright, and K. Ramchandran, "Information-theoretic limits on sparse signal recovery: Dense versus sparse measurement matrices." arXiv:0806.0604v1 [math.ST]., June 2008.
14. H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Inform. Theory* **54**, pp. 2210–2219, May 2008.
15. B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Computing* **24**, pp. 227–234, Apr. 1995.
16. S. Chen, S. A. Billings, and W. Luo, "Orthogonal least squares methods and their application to non-linear system identification," *Int. J. Control* **50**, pp. 1873–1896, Nov. 1989.
17. S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.* **41**, pp. 3397–3415, Dec. 1993.
18. Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Conf. Rec. 27th Asilomar Conf. Sig., Sys., & Comput.*, **1**, pp. 40–44, (Pacific Grove, CA), Nov. 1993.
19. G. Davis, S. Mallat, and Z. Zhang, "Adaptive time-frequency decomposition," *Optical Eng.* **37**, pp. 2183–2191, July 1994.
20. R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Stat. Soc., Ser. B* **58**(1), pp. 267–288, 1996.
21. R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Comm.* **38**, pp. 496–508, Apr. 1990.
22. Y. Jin and B. Rao, "Performance limits of matching pursuit algorithms," in *Proc. IEEE Int. Symp. Inform. Th.*, pp. 2444–2448, (Toronto, Canada), June 2008.
23. D. Wipf and B. Rao, "Comparing the effects of different weight distributions on finding sparse representations," in *Proc. Neural Information Process. Syst.*, (Vancouver, Canada), Dec. 2006.
24. A. K. Fletcher, S. Rangan, and V. K. Goyal, "On-off random access channels: A compressed sensing framework." arXiv:0903.1022v1 [cs.IT]., Mar. 2009.
25. M. Akçakaya and V. Tarokh, "Shannon theoretic limits on noisy compressive sampling." arXiv:0711.0366v1 [cs.IT]., Nov. 2007.
26. G. Reeves, "Sparse signal sampling using noisy linear projections," Tech. Rep. UCB/EECS-2008-3, Univ. of California, Berkeley, Dept. of Elec. Eng. and Comp. Sci., Jan. 2008.
27. A. K. Fletcher and S. Rangan, "On recovering sparse signals via orthogonal matching pursuit." Preprint, July 2009.
28. M. J. Wainwright, "Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1 -constrained quadratic programming (lasso)," *IEEE Trans. Inform. Theory* **55**, pp. 2183–2202, May 2009.
29. 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*, (Pacific Grove, CA), Nov. 2005.
30. S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comp.* **20**(1), pp. 33–61, 1999.
31. D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harm. Anal.* **26**, pp. 301–321, July 2008.
32. W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing: Closing the gap between performance and complexity." arXiv:0803.0811v1 [cs.NA]., Mar. 2008.
33. M. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Machine Learning Research* **1**, pp. 211–244, Sept. 2001.
34. D. Wipf and B. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.* **52**, pp. 2153–2164, Aug. 2004.
35. P. Schniter, L. C. Potter, and J. Ziniel, "Fast Bayesian matching pursuit: Model uncertainty and parameter estimation for sparse linear models," *IEEE Trans. Signal Process.*, submitted Aug. 2008.
36. A. Agrawal, J. G. Andrews, J. M. Cioffi, and T. Meng, "Iterative power control for imperfect successive interference cancellation," *IEEE Trans. Wireless Comm.* **4**, pp. 878–884, May 2005.
37. A. K. Fletcher, S. Rangan, and V. K. Goyal, "Necessary and sufficient conditions on sparsity pattern recovery." arXiv:0804.1839v1 [cs.IT]., Apr. 2008.