

Analysis of Fast Sparse Superposition Codes

Andrew R. Barron, and Antony Joseph,
Department of Statistics, Yale University, New Haven, CT 06520 USA
e-mail: {andrew.barron, antony.joseph}@yale.edu.

Abstract—Sparse superposition codes with a fast adaptive successive decoder for the additive white Gaussian noise channel were introduced last year by the authors, along with presentation of reliability at rates approaching capacity. The present work presents ingredients of the distributional analysis of the decoder.

I. INTRODUCTION

The framework for superposition codes is the formation of specific forms of linear combinations of a given set of vectors. We have a dictionary X_1, X_2, \dots, X_N of vectors, each with n coordinates, for which the codeword vectors take the form of superpositions $\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_N X_N$. The vectors X_j provide the terms or components of the codewords with coefficients β_j . By design, each entry of these vectors X_j is independent standard normal. The choice of codeword is conveyed through the coefficients, with sum of squares chosen to match a power requirement P . In brief, a codeword takes the form $X\beta$ where X is the matrix whose columns are the vectors X_1, X_2, \dots, X_N , with a constraint $\|\beta\|^2 = P$ imposed on the coefficient vector β . For a channel with additive white Gaussian noise (AWGN), with superposition coding, what is received is

$$Y = X\beta + \varepsilon,$$

a vector of length n , where ε is the noise vector with distribution $N(0, \sigma^2 I)$.

For a sparse superposition code the message is conveyed by a choice of L terms for which the coefficients are non-zero, with L/N a small fraction of the dictionary size. Attention is focussed on the case of partitioned codes, in which the dictionary is split into $B = N/L$ sections, with one term from each section chosen to be non-zero. Thus there are B^L codewords. With B a power of 2, the encoding from an input bit-string u_1, u_2, \dots, u_K , with $K = L \log B$, consists of partitioning the string into L substrings of length $\log B$ which index the terms from the dictionary which are chosen to be included in the codeword. Denote these indices $sent = \{j_1, j_2, \dots, j_L\}$. The coefficient value for the term from section ℓ takes the form $\beta_{j_\ell} = \sqrt{P_\ell}$ with $\sum_\ell P_\ell = P$. All terms not in $sent$ are assigned zero coefficient value.

The rate of the code is $R = (L \log B)/n$ equal to the ratio of the number of input bits to the number of uses of the channel. The capacity of the AWGN channel is $C = (1/2) \log(1 + snr)$ where $snr = P/\sigma^2$ is the signal-to-noise ratio.

These sparse superposition codes with partitioning and a fast adaptive successive decoder were introduced by the authors in [3], constituting the first practical code for the AWGN with rate arbitrarily close to capacity and error probability proven to be exponentially small.

Analogous models and algorithms have been used in sparse signal recovery and compressed sensing. Among much work that can be cited, we call attention to [9] where such sparse models are previously considered for communication for the AWGN, and a positive rate is established. However, the rate there and in other previous sparse signal recovery analyses does not approach capacity.

Our treatment in [3] also included a description of ingredients of the fit formed by the decoder, plus definition of an update function that tracts the accumulation of correct decodings for rates below capacity, as well as a statement of the reliability bounds. A fuller version of the story is in the extended manuscript [4].

The purpose of the current conference presentation is to focus on the distributional properties of components of the fit that underly that analysis. It is anticipated that this technique may be useful for examination of other related problems of iterative decoding or stepwise regression, including analysis of greedy algorithms and matching pursuit, originating in [7] and [8] and most recently studied in [10].

To review design and performance aspects of our fast decoder, to get rates approaching capacity, a variable power allocation is used. The P_ℓ is chosen to be proportional to $e^{-2C\ell/L}$, or to slight variants thereof. Such variable power originates with [5] in achieving sum-rate capacity of multi-user Gaussian channels and a recent use of it is in [6].

Our scheme provides reliable communication for rates R up to a value C_B for any $B \geq 2$. This C_B is near the capacity for B sufficiently large compared to $1 + snr$, as needed to make small expressions of the form $C/\log B$ that arise in the bounds on the gap from capacity.

The probability of more than a small fraction of section mistakes is shown to be exponentially small, with exponent of the order $(C_B - R)^2 L$ or, with worse constants, of order $(C_B - R)^2 L \sqrt{\log B}$. Here $L = nR/\log B$. This exponent is within a logarithmic factor of the optimal Shannon-Gallager form of order $(C - R)^2 n/V$ for fixed R near C .

The first step of the decoder is to compute the inner product of the received string Y with each of the terms in the dictionary, and see which of these inner products are above a threshold. Such a set of inner products for a step of the decoder is performed in parallel by a computational unit, e.g. a signal-processing chip with $N = LB$ parallel accumulators, each of which has pipelined computation, so that the inner product is updated as the elements of the string arrive.

In this basic step, the terms that it decodes are those for which the test statistic is above threshold. The size of the threshold is specified to keep the fraction of incorrect terms

above threshold negligible, while allowing a moderate fraction of correct terms to be found above threshold each step. A fit is formed at the end of a step by adding the terms selected.

Additional steps are used to bring the total fraction decoded up near 1. Each step of the decoder computes updated test statistics, taking inner products of the remaining terms with a vector determined using Y and the previous fit, and sees which are above threshold.

The decoding algorithm adapts the choice of which sections of terms will be decoded on each step in accordance with which sections have a term with an inner product observed to be above threshold. Thus we have called our class of procedures *adaptive successive decoding*.

Analysis involves weighted proportions of events, which are sums across the terms of indicators of events multiplied by weights π_j . For terms j in section ℓ this weight is set to $\pi_j = P_\ell/P$. With bounded ratio of maximum to minimum power across the sections, such weighted proportions agree with unweighted proportions to within constant factors.

The analysis leads us to a function g on $[0, 1]$, called the update function, which depends on the power allocation distribution π and the parameters L , B , snr and R . The idea of this function is that if at the previous step the weighted fraction of correct detections minus the weighted fraction of false alarms is x then the corresponding weighted fraction of correct detections expected for the next step is at least $g(x)$.

For suitable rates, the function $g(x)$ stays sufficiently above x over most of $[0, 1]$ and, consequently, the decoder has a high chance of not more than a small fraction of section mistakes.

An incorrect term above threshold is a *false alarm*, while failure of the correct term to provide a statistic value above threshold after a suitable number of steps is a *failed detection*. Let $\hat{\delta}_{mis}$ refer to the failed detection rate plus the false alarm rate. It is this $\hat{\delta}_{mis}$ that our technique controls, providing a small bound δ_{mis} that holds with high probability.

A section mistake is counted as an *error* if it arises from a single incorrectly selected term. It is an *erasure* if no term is selected or more than one term is selected. Let $\hat{\delta}_{error}$ be the fraction of section errors and $\hat{\delta}_{erase}$ be the fraction of section erasures. In each section the event indicators satisfy the property that $1_{erase} + 21_{error}$ is not more than $1_{failed\ detection} + 1_{false\ alarm}$. Accordingly, $2\hat{\delta}_{error} + \hat{\delta}_{erase}$ is not more than $\hat{\delta}_{mis}$, the failed detection rate plus the false alarm rate.

Candidate subsets of terms sent could differ from each other in only a few sections. When that is so, the subsets could be difficult to distinguish, so that it would be natural to expect a few section mistakes.

To permit completion of the task of identifying the terms, arrange sufficient distance between the subsets using composition with an outer Reed-Solomon (RS) code of rate near one. The alphabet of the Reed-Solomon code of size B , a power of 2, is taken to correspond to the indices of the selected terms in each section. Suppose the likely event $\hat{\delta}_{mis} < \delta_{mis}$ holds from the output of the inner superposition code. Then the outer Reed-Solomon corrects the small fraction of remaining

mistakes so that we end up not only with small section mistake rate but also with small block error probability. If $1-\delta$ is the rate of the RS code, then the section errors and erasures can be corrected, provided $\delta_{mis} \leq \delta$. Furthermore, if R is the rate associated with our inner superposition code, then the total rate after correcting for the remaining mistakes is given by $R_{total} = (1-\delta)R$, using $\delta = \delta_{mis}$. So, if Δ is the relative rate drop from capacity of the inner code, then the relative rate drop of the composite code Δ_{total} is not more than $\delta_{mis} + \Delta$.

The end result, using our theory for the distribution of the fraction of mistakes of the superposition code, is that for suitable rate up to a value near capacity the block error probability is exponentially small.

One may regard the composite code as a superposition code in which the subsets are forced to maintain at least a certain minimal separation, so that decoding to within a certain distance from the true subset implies exact decoding.

A limitation of these practical results is the slow approach of \mathcal{C}_B to the capacity \mathcal{C} , with gap of order $\mathcal{C}/\log B$. Nevertheless, a case can be made that the sparse superposition code with adaptive successive decoder is the first scheme for the Gaussian noise channel with a practical decoder proven to have exponentially small error probability at any fixed rate below capacity. Recently, in work coauthored by one of us [1], a competing method based on polar codes is adapted to the AWGN, to achieve any rate below capacity, albeit with an error probability exponentially small in \sqrt{n} rather than $n/\log n$.

As desired we turn our focus in the next section to distributional analysis of the contributions to the fit formed by the adaptive successive decoder.

The full manuscript [4] includes in addition to these matters of distributional analysis, discussion of relationship to past work, variants of the algorithm, computational analysis, proof of reliability bounds, tools for analyzing the update function $g(x)$, determination of \mathcal{C}_B , quantification of the contributions of its drop from capacity, quantification of the error exponent, and a more complete list of references.

II. FRAMEWORK FOR ANALYSIS OF THE DECODER

As we have said, from the received Y , the first step of the decoder forms for each term X_j of the dictionary, the test statistic $\mathcal{Z}_{1,j} = X_j^T Y / \|Y\|$, and compares it to a threshold $\tau = \sqrt{2 \log B} (1 + \delta_a)$. The distribution of this test statistic will be seen to be that of a standard normal plus a shift by a nearly deterministic amount, where the presence of the shift depends on whether j is one of the terms sent. The threshold is set such that very few of the terms not sent will be above threshold. Yet a positive fraction of the terms sent, determined by the size of the shift, will be above threshold and hence will be correctly decoded on this first step.

The output of the first step consists of the set dec_1 of terms above threshold and the vector $F_1 = \sum_{j \in dec_1} \sqrt{P_j} X_j$ which forms the first part of the fit. The set of terms $J_1 = J$ investigated in step 1 is the set of all columns of the dictionary, while the set $J_2 = J - dec_1$ remains for second step consideration.

In starting an adaptive Gram-Schmidt orthogonalization, we let $G_1 = Y$ and let G_2 be the part of the vector F_1 orthogonal to Y . That is $G_2 = F_1 - \hat{b}_{2,1}Y$ where $\hat{b}_{2,1} = F_1^T Y / \|Y\|^2$. Various vectors of interest have representation in terms of such orthogonal vectors. For instance the first step residual $r_1 = Y - F_1$ is a linear combination of G_1 and G_2 . Moreover, each X_j has representation as

$$X_j = \mathcal{Z}_{1,j} G_1 / \|G_1\| + \mathcal{Z}_{2,j} G_2 / \|G_2\| + V_{3,j}$$

where $\mathcal{Z}_{k,j} = X_j^T G_k / \|G_k\|$ and $V_{3,j}$ is a vector orthogonal to G_1 and G_2 . Here $\mathcal{Z}_{1,j} = X_j^T Y / \|Y\|$ is our first step test statistic.

The second step test statistic $\mathcal{Z}_{2,j}^{comb}$ is assumed to be formed from $\mathcal{Z}_{1,j}$ and $\mathcal{Z}_{2,j}$. For instance the inner product with the residuals $X_j^T r_1 / \|r_1\|$ is a linear combination of $\mathcal{Z}_{1,j}$ and $\mathcal{Z}_{2,j}$ with empirically determined weights of combination. The focus of attention in [4] is test statistics $\mathcal{Z}_{2,j}^{comb}$ that take the form $\sqrt{\lambda_1} \mathcal{Z}_{1,j} - \sqrt{\lambda_2} \mathcal{Z}_{2,j}$ with $\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_1 + \lambda_2 = 1$. The distributional properties of $\mathcal{Z}_{1,j}$ and $\mathcal{Z}_{2,j}$ are used to determine estimates of values of such λ_1 and λ_2 which will maximize the separation between the distributions for terms sent and terms not sent. These weights of combination are advocated as possibly simpler to analyze than the use of the weights associated with inner products with the residuals.

The second step decoding set dec_2 is a subset of J_2 formed in a similar manner as for the first step, by comparing $\mathcal{Z}_{2,j}^{comb}$ to a threshold. This leads to an additional contribution to the fit $F_2 = \sum_{j \in dec_2} \sqrt{P_j} X_j$, from which one forms G_3 , the part of this fit orthogonal to G_1 and G_2 , and so on.

Proceed in this manner to perform the following loop of calculations, for $k \geq 2$. From the output of previous steps, there is the vectors $F_{k'}$ and $G_{k'}$ and statistics $\mathcal{Z}_{k',j}$ for $k' < k$. Plus there is a set $dec_{1,k-1} = dec_1 \cup \dots \cup dec_{k-1}$ already decoded and a set $J_k = J - dec_{1,k-1}$ of terms for us to test at step k . Determine the part G_k of F_{k-1} orthogonal to the previous $G_{k'}$ and for each j not in dec_{k-1} compute

$$\mathcal{Z}_{k,j} = X_j^T G_k / \|G_k\|$$

and a combined statistic

$$\mathcal{Z}_{k,j}^{comb} = \sqrt{\lambda_{1,k}} \mathcal{Z}_{1,j} - \sqrt{\lambda_{2,k}} \mathcal{Z}_{2,j} - \dots - \sqrt{\lambda_{k,k}} \mathcal{Z}_{k,j}.$$

To maximize the separation between the distributions that arise for terms sent and terms not sent, these positive weights are taken of the form $\lambda_{k',k} = w_{k'}/s_k$, with $w_1 = 1$, and $s_k = 1 + w_2 + \dots + w_k$, with w_k to be specified. Accordingly, the combined statistic may be computed by the update

$$\mathcal{Z}_{k,j}^{comb} = \sqrt{1 - \lambda_k} \mathcal{Z}_{k-1,j}^{comb} - \sqrt{\lambda_k} \mathcal{Z}_{k,j},$$

where $\lambda_k = w_k/s_k$. For terms j in J_k these statistics are compared to a threshold, leading to events $\{\mathcal{Z}_{k,j}^{comb} \geq \tau\}$. As quantified by an analysis of the distribution of the $\mathcal{Z}_{k,j}$, for successive steps there is an increasing separation between the distribution for terms j sent and the others.

With dec_k taken to be $\{j \in J_k : \mathcal{Z}_{k,j}^{comb} \geq \tau_k\}$ the output of step k consists of the vectors G_k and $F_k = \sum_{j \in dec_k} \sqrt{P_j} X_j$

and the statistics $\mathcal{Z}_{k,j}$. Moreover one updates to the set of decoded terms $dec_{1,k} = dec_{k-1} \cup dec_k$ and the set $J_{k+1} = J_k - dec_k$ of terms remaining for consideration. This completes the actions of step k of the loop.

On each step k we decode a substantial part of what remains, because of growth of the mean separation between terms sent and the others.

The algorithm stops when either L terms have been decoded or no terms from J_k are found to have statistic above threshold, so that F_k is zero and the statistics would remain thereafter unchanged. Analytical stopping conditions are also available. Our best bounds on the closeness of the rate target \mathcal{C}_B to capacity occur with a total number of steps m not more than an integer part of $2 + snr \log B$.

Up to step k , the total fit fit_k is $\sum_{j \in dec_{1,k}} \sqrt{P_j} X_j$ which is the sum of the pieces from each step $fit_k = F_1 + F_2 + \dots + F_k$.

As to the part G_k of F_{k-1} orthogonal to previous $G_{k'}$, take advantage of two ways to view it, one emphasizing computation and the other analysis.

For computation of G_k from F_{k-1} apply a standard Gram-Schmidt step. The G_1, G_2, \dots, G_{k-1} are orthogonal vectors, so the parts of F_{k-1} in these directions are $\hat{b}_{k,k'} G_{k'}$ with $\hat{b}_{k,k'} = F_{k-1}^T G_{k'} / \|G_{k'}\|^2$ for $k' < k$, where if $\|G_{k'}\| = 0$ use $\hat{b}_{k,k'} = 0$. Accordingly, the new G_k is computed by $G_k = F_{k-1} - \sum_{k'=1}^{k-1} \hat{b}_{k,k'} G_{k'}$. This computation entails the n -fold sums of products $F_{k-1}^T G_{k'}$ for determination of the $\hat{b}_{k,k'}$. Then from this computed G_k we obtain the inner products with the X_j to yield $\mathcal{Z}_{k,j} = X_j^T G_k / \|G_k\|$ for j in J_k .

For analysis, look at what happens to the representation of the individual terms. Each term X_j for $j \in J_{k-1}$ has the decomposition

$$X_j = \mathcal{Z}_{1,j} \frac{G_1}{\|G_1\|} + \mathcal{Z}_{2,j} \frac{G_2}{\|G_2\|} + \dots + \mathcal{Z}_{k-1,j} \frac{G_{k-1}}{\|G_{k-1}\|} + V_{k,j},$$

where $V_{k,j}$ is the part of X_j orthogonal to G_1, G_2, \dots, G_{k-1} . Since $F_{k-1} = \sum_{j \in dec_{k-1}} \sqrt{P_j} X_j$ it follows that G_k has the representation $G_k = \sum_{j \in dec_{k-1}} \sqrt{P_j} V_{k,j}$, from which $\mathcal{Z}_{k,j} = V_{k,j}^T G_k / \|G_k\|$, with which the above representation of X_j updates. With the initialization $V_{0,j} = X_j$, these $V_{k+1,j}$ may be thought of as iteratively obtained from the corresponding vectors at the previous step, that is, $V_{k+1,j}$ is $V_{k,j} - \mathcal{Z}_{k,j} G_k / \|G_k\|$. These V do not actually need to be computed, nor do we need to compute its components detailed below, but this representation of the terms X_j is used in the proof of the distributional properties of the $\mathcal{Z}_{k,j}$.

The $\pi_j = P_j/P$ sum to 1 across j in *sent*, and sum to $B-1$ across j in *other*. Define in general $\hat{q}_k = \sum_{j \in sent \cap dec_k} \pi_j$ for the step k correct detections and $\hat{f}_k = \sum_{j \in other \cap dec_k} \pi_j$ for the false alarms.

The total weighted fraction of correct detections up to step k is $\hat{q}_k^{tot} = \sum_{j \in sent \cap dec_{1,k}} \pi_j$ which may be written as the sum $\hat{q}_1 + \hat{q}_2 + \dots + \hat{q}_k$. This total may be regarded as the same as the π weighted measure of the union $\sum_j \pi_j 1_{\{\mathcal{H}_{1,j} \cup \dots \cup \mathcal{H}_{k,j}\}}$.

Likewise the weighted count of false alarms $\hat{f}_k^{tot} = \sum_{j \in other \cap dec_{1,k}} \pi_j$ may be written as $\hat{f}_1 + \hat{f}_2 + \dots + \hat{f}_k$ which may be expressed as $\sum_j \pi_j 1_{\{\mathcal{H}_{1,j} \cup \dots \cup \mathcal{H}_{k,j}\}}$.

III. DISTRIBUTIONAL ANALYSIS OF THE DECODER:

In this section we describe the distributional properties of the random variables $\mathcal{Z}_k = (\mathcal{Z}_{k,j} : j \in J_k)$ for each $k = 1, 2, \dots, n$. In particular we show for each k that $\mathcal{Z}_{k,j}$ are location shifted normal random variables with variance near one for $j \in \text{sent} \cap J_k$ and are independent standard normal random variables for $j \in \text{other} \cap J_k$.

Define

$$C_{j,R} = \pi_j L\nu/(2R),$$

where $\pi_j = P_j/P$ and $\nu = \nu_1 = P/(\sigma^2 + P)$. Likewise define $C_{j,R,B} = (C_{j,R}) 2 \log B$, which is also $C'_{j,R} \tau^2$ with $C'_{j,R} = C_{j,R}/(1 + \delta_a)^2$. It also simplifies to $C_{j,R,B} = n \pi_j \nu$.

Each step of the decoder will derive benefit from looking at all sections. Nevertheless, it will be seen that $C_{j,R}$ needs to be at least 1 for some sections, to get the decoder to begin successfully, and then not taper too rapidly to allow decodings to accumulate on successive steps.

For the case of power P_j proportional to $e^{-2C\ell/L}$, we have $\pi_j = e^{-2C(\ell-1)/L} (1 - e^{-2C/L}) / (1 - e^{-2C})$ for each j in section ℓ , for ℓ from 1 to L . Define $\tilde{C} = (L/2)[1 - e^{-2C/L}]$, which is essentially identical to C , for L large compared to C . Then for j in section ℓ we have that

$$C_{j,R} = (\tilde{C}/R) e^{-2C(\ell-1)/L}.$$

For rates R not more than C , this $C_{j,R}$ is at least 1 in some sections and it tapers at the fastest rate at which we can still accumulate decoding successes.

We now are in a position to give the lemma for the distribution of \mathcal{Z}_1 . Recall $J_1 = J$ is the set of all N indices.

Lemma 1: *Distributional analysis of the first step.* For each $j \in J_1$, the statistic $\mathcal{Z}_{1,j}$ can be represented as

$$\sqrt{C_{j,R,B}} [\mathcal{X}_n/\sqrt{n}] 1_{j \text{ sent}} + \mathcal{Z}_{1,j},$$

where $\mathcal{Z}_1 = (\mathcal{Z}_{1,j} : j \in J_1)$ is multivariate normal $N(0, \Sigma_1)$ and $\mathcal{X}_n^2 = \|Y\|^2/\sigma_Y^2$ is a Chi-square n random variable that is independent of \mathcal{Z}_1 . Here recall that $\sigma_Y^2 = P + \sigma^2$ is the variance of each coordinate of Y .

The covariance matrix Σ_1 can be expressed as $I - b_1 b_1^T$, where $b_1 = (b_{1,1}, b_{1,2}, \dots, b_{1,N})$ with each $b_{1,j} = \beta_j/\sigma_Y$.

Proof of Lemma 1: Recall that the X_j for j in J are independent $N(0, I)$ random vectors and that $Y = \sum_j \beta_j X_j + \varepsilon$, where the sum of squares of the β_j is equal to \bar{P} .

Consider first the decomposition of each random vector X_j of the dictionary into a vector in the direction of the received Y and a vector U_j uncorrelated with Y . That is

$$X_j = b_{1,j} Y/\sigma_Y + U_j,$$

where the coefficient is $b_{1,j} = \mathbb{E}[X_{i,j} Y_i]/\sigma_Y = \beta_j/\sigma_Y$, which indeed makes each coordinate of U_j uncorrelated with each coordinate of Y . These coefficients collect into a vector $b_1 = \beta/\sigma_Y$ in \mathbb{R}^N . In the case that the magnitude of the non-zero coefficients is $\sqrt{P/\bar{L}}$, then for terms $j \text{ sent}$, the square of $b_{1,j}$

is equal to ν/L . In any case, since $\|\beta\|^2 = P$, the sum of squares $\|b_1\|^2$ is equal to

$$\nu = P/(\sigma^2 + P).$$

The subscript 1 on first step quantities such as $b_{1,j}$ is to distinguish them from corresponding values that will arise on the subsequent steps. Likewise, $U_j = U_{1,j}$.

These first step vectors $U_j = X_j - b_{1,j} Y/\sigma_Y$ along with Y are linear combinations of joint normal random variables and so are also joint normal, with zero correlation implying that Y is independent of the collection of U_j . The independence of Y and U_j facilitates development of distributional properties of the $U_j^T Y$. For these purposes we need the characteristics of the joint distribution of the U_j across terms j (clearly there is independence for distinct time indices i).

The coordinates of U_j and $U_{j'}$ have mean zero and expected product $1_{\{j=j'\}} - b_{1,j} b_{1,j'}$. For the constant power allocation case, this value is $1 - (\nu/L) 1_{j \text{ sent}}$ when $j' = j$ and it is a small covariance $-(\nu/L) 1_{j,j' \text{ sent}}$ when $j' \neq j$. In general, the covariances $(\mathbb{E}[U_{i,j} U_{j,j'}] : j, j' \in J)$ organize into a matrix

$$\Sigma_1 = \Sigma = I - \Delta = I - b_1 b_1^T.$$

For any constant vector $\alpha \neq 0$, consider $U_j^T \alpha / \|\alpha\|$. Its joint normal distribution across terms j is the same for any such α . Specifically, it is a normal $N(0, \Sigma)$, with mean zero and the indicated covariances.

Likewise define the random variables $Z_j = U_j^T Y / \|Y\|$, also denoted $\mathcal{Z}_{1,j}$ when making explicit that it is for the first step. Jointly across j , these Z_j have the normal $N(0, \Sigma)$ distribution, independent of Y . Indeed, since the U_j are independent of Y , when we condition on $Y = \alpha$ we get the same $N(0, \Sigma)$ distribution, and since this conditional distribution does not depend on Y , it is the unconditional distribution as well.

Where this gets us is revealed via the representation of the inner product $X_j^T Y$ as $b_{1,j} \|Y\|^2/\sigma_Y + U_j^T Y$, which can be written as

$$X_j^T Y = \beta_j \frac{\|Y\|^2}{\sigma_Y^2} + \|Y\| Z_j.$$

This identifies the distribution of the $X_j^T Y$ as that obtained as a mixture of the normal Z_j with scale and location shifts determined by an independent random variable $\mathcal{X}_n^2 = \|Y\|^2/\sigma_Y^2$, distributed as Chi-square with n degrees of freedom.

Divide through by $\|Y\|$ to normalize these inner products to a helpful scale and to simplify the distribution of the result to be only that of a location mixture of normals. The resulting random variables $\mathcal{Z}_{1,j} = X_j^T Y / \|Y\|$ take the form

$$\mathcal{Z}_{1,j} = \sqrt{n} b_{1,j} |Y|/\sigma_Y + Z_j,$$

where $|Y|/\sigma_Y = \mathcal{X}_n/\sqrt{n}$ is near 1. Note that $\sqrt{n} b_{1,j} = \sqrt{n} \beta_j/\sigma_Y$ which is $\sqrt{n \pi_j \nu}$ or $\sqrt{C_{j,R,B}}$. This completes the proof of Lemma 1.

We next present the corresponding result for $k \geq 2$. The distribution of $\mathcal{Z}_{k,j}$ with $j \in J_k$ can also be expressed in a manner similar to that above, that is, each $\mathcal{Z}_{k,j}$ can be

expressed as a normal random variable $Z_{k,j}$ plus a location shift depending on whether j is in *sent* or not.

We maintain the pattern used in Lemma 1 and use the calligraphic font $\mathcal{Z}_{k,j}$ to denote the test statistics that incorporate the shift for j in *sent* and the standard font $Z_{k,j}$ to denote their counterpart mean zero normal random variables before the shift.

We provide in the lemma below characterization of the sequence of conditional distributions of the $Z_k = (Z_{k,j} : j \in J_k)$ and $\|G_k\|$, given \mathcal{F}_{k-1} , for $k = 1, 2, \dots, n$, where

$$\mathcal{F}_{k-1} = (\|G_{k'}\|, Z_{k'} : k' = 1, \dots, k-1).$$

This determines also the distribution of $\mathcal{Z}_k = (Z_{k,j} : j \in J_k)$ conditional on \mathcal{F}_{k-1} . Initializing with the distribution of \mathcal{Z}_1 derived in Lemma 1, we provide the conditional distributions for all $2 \leq k \leq n$. The algorithm will stop at a k much less than n . The J_k is never empty because we decode at most L , so there must always be at least $(B-1)L$ remaining. For an index set which may depend on the conditioning variables, we let $N_{J_k}(0, \Sigma)$ denote a mean zero multivariate normal distribution with index set J_k and the indicated covariance matrix.

Lemma 2: Distributional analysis for subsequent steps. For $k \geq 2$, given \mathcal{F}_{k-1} , the conditional distribution $\mathbb{P}_{Z_k, J_k | \mathcal{F}_{k-1}}$ of $Z_{k, J_k} = (Z_{k,j} : j \in J_k)$ is normal $N_{J_k}(0, \Sigma_k)$; the random variable $\mathcal{X}_{d_k}^2 = \|G_k\|^2 / \sigma_k^2$ is a Chi-square distributed, with $d_k = n - k + 1$ degrees of freedom, conditionally independent of the Z_k , where σ_k^2 are positive values depending on \mathcal{F}_{k-1} ; and, moreover, $\mathcal{Z}_{k,j}$ has the representation

$$-\sqrt{\hat{w}_k C_{j,R,B}} [\mathcal{X}_{d_k} / \sqrt{n}] 1_{j \text{ sent}} + Z_{k,j}.$$

The shift increment from this statistic is $\hat{w}_k = \hat{s}_k - \hat{s}_{k-1}$, which are increments of a series with total

$$1 + \hat{w}_2 + \dots + \hat{w}_k = \hat{s}_k = \frac{1}{1 - (\hat{q}_1^{adj} + \dots + \hat{q}_{k-1}^{adj}) \nu}$$

where $\hat{q}_j^{adj} = \hat{q}_j / (1 + \hat{f}_j / \hat{q}_j)$, obtained from weighted fractions of correct detections and false alarms on previous steps. Here $\hat{s}_1 = \hat{w}_1 = 1$. The covariance Σ_k has the representation

$$\Sigma_k = I - \delta_k \delta_k^T = I - \nu_k \beta \beta^T / P$$

where $\nu_k = \hat{s}_k \nu$. That is $(\Sigma_k)_{j,j'} = 1_{j=j'} - \delta_{k,j} \delta_{k,j'}$, for j, j' in J_k , where the vector δ_k is in the direction β , with $\delta_{k,j} = \sqrt{\nu_k P_j / P} 1_{j \text{ sent}}$ for j in J_k .

The proof of this lemma follows the same pattern as the proof of Lemma 1. For details of the proof see [4].

To summarize its use, it yields weights $\lambda_{k'} = \hat{w}_{k'} / \hat{s}_k$ that provide the ideal combination $\mathcal{Z}_{k,j}^{comb} = \sum_{k'=1}^k \sqrt{\lambda_{k'}} \mathcal{Z}_{k',j}$. In the analysis the chi-square random variables divided by their degrees of freedom are bounded by constants near 1 (excepting events of exponentially small probability) and this combined test statistic takes the approximate form

$$\sqrt{\hat{s}_k C_{j,R,B}} 1_{j \text{ sent}} + Z_{k,j}^{comb},$$

where $Z_{k,j}^{comb} = \sum_{k'=1}^k \sqrt{\lambda_{k'}} \mathcal{Z}_{k',j}$. Likewise, it is inductively demonstrated that, with very high probability, the \hat{s}_k in the

lemma above are not less than a sequence of deterministic values $s_k = 1 / (1 - \nu x_{k-1})$ with $x_{k-1} = q_{1,k-1}^{adj}$, where the $q_{1,k}$ is a target correct detection rate obtained from k fold application of the function g as specified below.

This analysis allows for the $\lambda_{k'}$ to be replaced by corresponding deterministic $w_{k'} / s_k$ for $k' = 1, 2, \dots, k$ where $w_k = s_k - s_{k-1}$, with which $Z_{k,j}^{comb}$ is jointly normal across terms j . Using the covariance from the above lemma, analysis of the joint normal density of the $Z_{k,j}$, shows that it has a bounded ratio with respect to the independent standard normal joint density. Accordingly, for determination of events of exponentially small probability, there is no loss to analyze the $Z_{k,j}^{comb}$ as being independent standard normal. With the distribution shifted for j in *sent* as given above, the probability of $Z_{k,j}^{comb}$ being above threshold is given by

$$g_j(x) = \Phi\left(\left(\sqrt{C'_{j,R} / (1 - \nu x)} - 1\right)\tau\right),$$

evaluated at $x = x_{k-1}$. It yields the expected weighted fraction above threshold at step k given by $g(x) = \sum_{j \text{ sent}} \pi_j g_j(x)$.

To examine the behavior of this update function $g(x)$, consider what it would be, denoted $g_{ideal}(x)$, in the case that each $g_j(x)$ is replaced by a corresponding indicator $1_{\{C'_{j,R} / (1 - \nu x) > 1\}}$. The use of the recommended variable power allocation with rate $R < C'$ with $C' = \tilde{C} / (1 + \delta_a)^2$, leads to $g_{ideal}(x)$ being at least the minimum of $x C' / R$ and 1, which stays above x for most of the interval $[0, 1]$. An approximate odd symmetry of the difference between $g_j(x)$ and its indicator approximation, leads to $g(x)$ behaving similarly, also above x for most of the interval $[0, 1]$.

This is the way in which the distributional properties of the inner product test statistics determine the ingredients of the analysis detailed in [4].

REFERENCES

- [1] E. Abbe and A. Barron, "Polar Codes for the AWGN." *Proc. IEEE Intern. Symp. Inform. Theory*, St. Petersburg, Russia, 2011.
- [2] A.R. Barron A. Joseph, "Least squares superposition coding of moderate dictionary size, reliable at rates up to channel capacity," Full manuscript submitted to the *IEEE Trans. Inform. Theory*. Conference version in *Proc. IEEE Intern. Symp. Inform. Theory*, Austin, TX, June 13-18, 2010.
- [3] A.R. Barron and A. Joseph, "Toward fast reliable communication at rates near capacity with Gaussian noise," *Proc. IEEE Intern. Symp. Inform. Theory*, Austin, TX, June 13-18, 2010.
- [4] A.R. Barron and A. Joseph, "Sparse superposition codes: Fast and reliable at rates approaching capacity with Gaussian noise," Extended manuscript of this submission, 2011. Available from www.stat.yale.edu/~arb4/publications.html
- [5] T.M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol.18, pp.2-14, 1972.
- [6] A.K. Fletcher, S. Rangan, and V.K. Goyal, "On-Off Random Access Channels: A Compressed Sensing Framework." ArXiv:0903.1022v2.
- [7] L. Jones, "A simple lemma for optimization in a Hilbert space, with application to projection pursuit and neural net training," *Annals of Statistics*, vol.20, pp.608-613, 1992.
- [8] S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries." *IEEE Trans. Signal Proc.*, vol.41, pp.3397-3415, 1993.
- [9] J. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inform. Theory*, vol.52, no.3, pp. 1030-1051, Mar. 2006.
- [10] T. Zhang, "Adaptive forward-backward greedy algorithm for learning sparse representations," Technical report, Statistics Department, Rutgers University. 2011.