Efficient Design and Decoding of Polar Codes

Authors: Peter Trifonov

Publication: IEEE T. Comm, Nov 2012
Speaker: Jeong-Min Ryu

Short summary: Polar codes are shown to be instances of both **generalized concatenated codes** and **multilevel codes**. It is shown that the performance of a polar code can be improved by representing it as a multilevel code and applying the multistage decoding algorithm with maximum likelihood decoding of outer codes. Additional performance improvement is obtained by replacing polar outer codes with other ones with better error correction performance. In some cases this also results in complexity reduction. It is shown that Gaussian approximation for density evolution enables one to accurately predict the performance of polar codes and concatenated codes based on them.

I. Introduction

The practical performance of polar codes under the successive cancellation (SC) decoding reported up to now turns out to be worse than that of LDPC and Turbo codes.

This paper demonstrates

1) Polar codes can be efficiently constructed using Gaussian approximation for density evolution.

2) It is shown that polar codes can be treated in the framework of multilevel coding. This enables one to improve the performance of polar codes by considering them as multilevel or, equivalently, generalized concatenated (GCC) ones, and using block-wise near-maximum-likelihood decoding of outer codes. In some cases this results also in reduced decoding complexity.

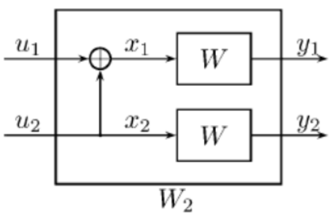
3) A simple algorithm for construction of GCC with inner polar codes.

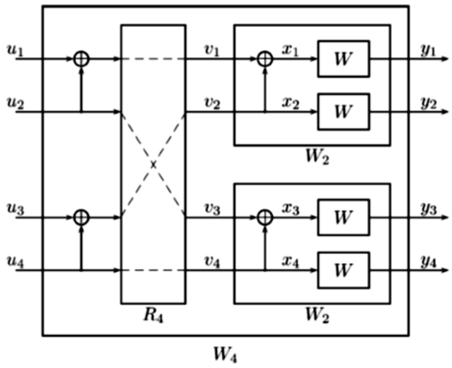
II. BACKGROUND

A. Polar codes

Consider a binary input output symmetric memoryless channel with output probability density function $W(y|x)$, $y \in Y$, $x \in \mathbb{F}_2$. It can be transformed into a vector channel given by $W_n(y_1^n | u_1^n) = W^n(y_1^n | u_1^n G_n)$, where $W^n(y_1^n | x_1^n) = \prod_{i=1}^{n} W(y_i | x_i)$, $G_n = B_s F^{\otimes s}$, $n = 2^s$, $F = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\otimes s$ denotes s-times Kronecker product of a matrix with itself, and B_s is a $2^s \times 2^s$ bit reversal permutation matrix.

For example)





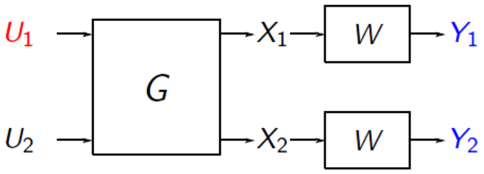
$W_2(y_1^2 | u_1^2) = W^2(y_1^2 | u_1^2 G_2)$ $\qquad$ $W_4(y_1^4 | u_1^4) = W_2(y_1^2 | u_1 \oplus u_2, u_3 \oplus u_4) W_2(y_3^4 |$

The vector channel can be further decomposed into equivalent subchannels

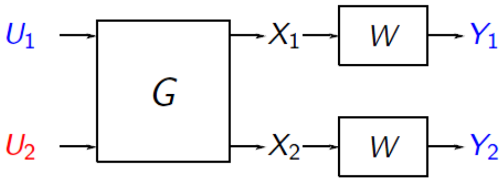
$$W_N^{(i)}(y_1^N, u_1^{i-1} | u_i) = \sum_{u_{i+1}^N} \frac{1}{2^{N-1}} W_N(y_1^N | u_1^N).$$

For example)

According to this, we can write $(W, W) \mapsto (W_2^{(1)}, W_2^{(2)})$ for any given B-DMC W.



$$W_2^{(1)}(y_1^2 | u_1) \triangleq \sum_{u_2} \frac{1}{2} W_2(y_1^2 | u_1^2) = \sum_{u_2} \frac{1}{2} W(y_1 | u_1 \oplus u_2) W(y_2 | u_2)$$



$$W_2^{(2)}(y_1^2, u_1 | u_2) \triangleq \frac{1}{2} W_2(y_1^2 | u_1^2) = \frac{1}{2} W(y_1 | u_1 \oplus u_2) W(y_2 | u_2)$$

Given y_1^n and estimates $\hat{u}_1^{i-1}$ of u_1^{i-1}, the SC decoding algorithm attempts to estimate u_i. This can be implemented by computing the following log-likelihood ratios $L_n^{(i)}\left(y_1^n, \hat{u}_1^{i-1}\right) = \log \frac{W_n^{(i)}\left(y_1^n, \hat{u}_1^{i-1} \mid u_i = 0\right)}{W_n^{(i)}\left(y_1^n, \hat{u}_1^{i-1} \mid u_i = 1\right)}$:

$$L_n^{(2i-1)}\left(y_1^n, \hat{u}_1^{2i-2}\right) = 2\tanh^{-1}\left(\tanh\left(L_{n/2}^{(i)}\left(y_1^{n/2}, \hat{u}_{1,e}^{2i-2} \oplus \hat{u}_{1,o}^{2i-2}\right)/2\right)\tanh\left(L_{n/2}^{(i)}\left(y_{n/2+1}^n, \hat{u}_{1,e}^{2i-2}\right)/2\right)\right), \quad (1)$$

$$L_n^{(2i)}\left(y_1^n, \hat{u}_1^{2i-1}\right) = L_{n/2}^{(i)}\left(y_{n/2+1}^n, \hat{u}_{1,e}^{2i-2}\right) + (-1)^{\hat{u}_{2i-1}} L_{n/2}^{(i)}\left(y_1^{n/2}, \hat{u}_{1,e}^{2i-2} \oplus \hat{u}_{1,o}^{2i-2}\right) \quad (2)$$

where $\hat{u}_{1,e}^i, \hat{u}_{1,o}^i$ are subvectors of $\hat{u}_1^i$ with even and odd indices, respectively, and $L_1^{(i)}\left(y_i\right) = \log \frac{W\left(y_i \mid 0\right)}{W\left(y_i \mid 1\right)}$.

III. Design of Polar Codes Based on Gaussian Approximation

The main drawback of the polar code construction method based on density evolution is its high computational complexity. The most practically important case corresponds to the AWGN channel. In this scenario, $L_1^{(i)}\left(y_i\right) \sim N\left(\frac{2}{\sigma^2}, \frac{4}{\sigma^4}\right)$, provided that the all-zero codeword is transmitted.

The value given by (1)-(2) can be considered as Gaussian random variables with $\boldsymbol{D}\left[L_n^{(i)}\right] = 2\boldsymbol{E}\left[L_n^{(i)}\right]$, where $\boldsymbol{E}$ and $\boldsymbol{D}$ are the mean and variance, respectively. This enable one to compute only the expected value of $L_n^{(i)}$, drastically reducing thus the complexity. In the case of polar codes this approach reduces to

$$\boldsymbol{E}\left[L_n^{(2i-1)}\right] = \phi^{-1}\left(1 - \left(1 - \phi\left(\boldsymbol{E}\left[L_{n/2}^{(i)}\right]\right)\right)^2\right), \quad (3)$$

$$\boldsymbol{E}\left[L_n^{(2i)}\right] = 2\boldsymbol{E}\left[L_{n/2}^{(i)}\right] \quad (4)$$

where

$$\phi(x)=\begin{cases}1-\frac{1}{\sqrt{4\pi x}}\int_{-\infty}^{\infty}\tanh\frac{u}{2}e^{-\frac{(u-x)^2}{4x}}dx, & x>0\\ 1 & , \quad x=0.\end{cases}$$

The error probability for each subchannel is given by

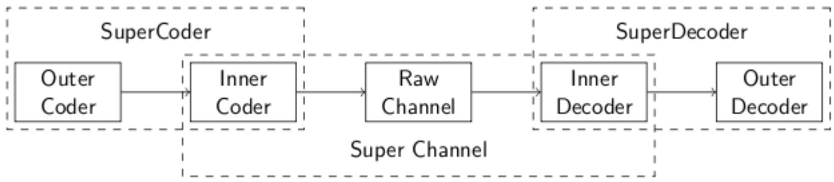
$$\pi_i \approx Q\left(\sqrt{\boldsymbol{E}\left[L_n^{(i)}\right]/2}\right),\ 1\le i\le n.$$

IV. Decomposition of Polar Codes

The overall performance of a polar code is dominated by the performance of the worst subchannel. The proposed approach avoids this problem by performing joint decoding over a number of subchannels.

A. Generalized concatenated polar codes

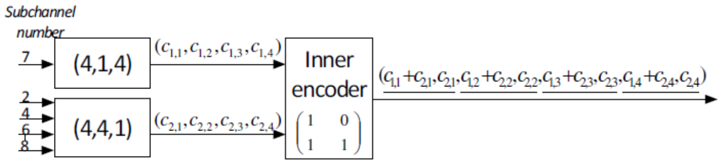
The recursive structure of polar codes enables one to consider them as GCC. Namely, the generator matrix of a polar code can be represented as $G=AF^{\otimes s}=A\left(F^{\otimes(s-l)}\otimes F^{\otimes l}\right)$, where $F=\begin{pmatrix}1&0\\1&1\end{pmatrix}$ and A is a full-rank matrix with at most one non-zero element in each column.



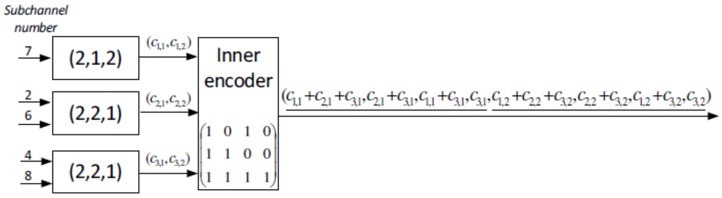
Inner code encoding: Inner codes $\mathbb{C}_i$ of length $n=2^l$ is generated by rows $i,\dots,2^l$ of matrix $B_lF^{\otimes l}$.

Outer code encoding: The generator matrix of the $(1+R(i,l))$-th outer code C_i is obtained by taking rows $1+R(j,s-l)$ of $F^{\otimes(s-l)}$, such that row $1+R\left(i2^{s-l}+j,s\right)$ of $F^{\otimes s}$ is included into the generator matrix of the original polar code, where $0\le i<2^l$, $0\le j<2^{s-l}$, and

$$R\left(\sum_{j=0}^{m-1}2^j i_j, m\right)=\sum_{j=0}^{m-1}2^j i_{m-1-j},\quad i_j\in\{0,1\}.$$



(a) $l = 1$

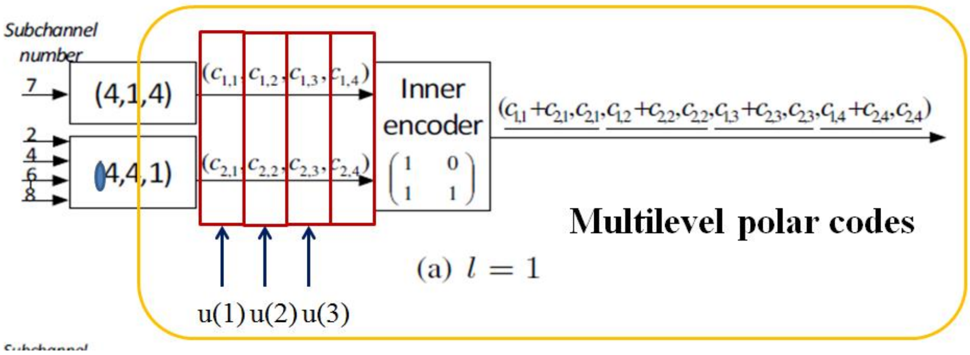


(b) $l = 2$

Fig. 3. Representation of $(8,5,2)$ polar code as GCC.

B. Multilevel polar codes

In the context of polar codes, signal constellation A is given by 2^n binary n-vectors $a(u)$, which can be obtained as $a(u) = uB_l F^{\otimes l}$, $u \in GF(2)^n$, where $n = 2^l$. This constellation is recursively partitioned into subsets $A(u_1^i)$ by fixing the values of $u_1,\ldots,u_i$. The elements of u are obtained as codeword symbols of outer codes C_i of length $N = 2^{s-l}$. That is, one can construct N vectors $u^{(j)} = (c_{1,j},\ldots,c_{n,j})$, $1 \le j \le N$, where $(c_{i,1},\ldots,c_{i,N}) \in C_i$, $1 \le i \le n$ and obtain a multilevel codeword $\left(u^{(1)}B_l F^{\otimes l},\ldots,u^{(N)}B_l F^{\otimes l}\right)$.



The multilevel polar codes can be decoded by multistage decoding algorithm.

V. Concatenated Codes Based on Polar Codes

The performance of a polar code under the multistage decoding with block-wise maximum-likelihood decoding of outer codes can be improved by changing the set of frozen bits. Furthermore, if the algorithm used to perform block-wise decoding of outer codes does not take into account their structure, one can use any linear block code with suitable parameters, not necessary polar, as C_i. This enables one to employ outer codes with better error correction performance.

A. Capacity rule

The rate R_i of C_i should be chosen equal to the capacity C_i of the *i*-th subchannel of the multilevel code, which is induced by matrix $B_l F^{\otimes l}$. According to [10], one obtains

$$C_i = I\left(y_1^n; u_i \mid u_1^{i-1}\right) = E_{u_1^{i-1}}\left[C\left(A\left(u_1^{i-1}\right)\right)\right] - E_{u_1^i}\left[C\left(A\left(u_1^i\right)\right)\right]$$

where

$$C(B) = \int_{R^n} \sum_{a\in B} \frac{W^n\left(y_1^n \mid a\right)}{|B|} \log_2\left(\frac{|B| W^n\left(y_1^n \mid a\right)}{\sum_{b\in B} W^n\left(y_1^n \mid b\right)}\right) dy_1^n$$

is the capacity when using the subset B of $\mathbb{F}_2^n$ for transmission over the vector channel $W^n\left(y_1^n \mid x_1^n\right)$. In the case of binary input memoryless output symmetric channels, one can drop the expectation operator to obtain $C_i = C\left(A^{(i-1)}\right) - C\left(A^{(i)}\right)$, where $A^{(i)} = A(\underbrace{0,\dots,0}_{i\text{ times}})$.

It can be seen that the latter set is a linear block code C_i generated by $l-i$ last rows of $B_l F^{\otimes l}$. The expression can be further simplified to

$$C\left(A^{(i)}\right) = \int_{R^N} \prod_{j=1}^N W\left(y_j \mid 0\right) \log_2\left(\frac{|C_i| \prod_{j=1}^N W\left(y_j \mid 0\right)}{\sum_{b\in B} \prod_{j=1}^N W\left(y_j \mid b_j\right)}\right) dy_1^N.$$

Hence, the capacity of the *i*-th subchannel of the multilevel polar code can be computed as

$$C_i = \int_{R^N} \prod_{j=1}^N W\left(y_j \mid 0\right) \log_2\left(\frac{2 \sum_{b\in C_{i+1}} \prod_{j=1}^N W\left(y_j \mid b_j\right)}{\sum_{b\in C_i} \prod_{j=1}^N W\left(y_j \mid b_j\right)}\right) dy_1^N. \quad (5)$$

Obviously, employing this rule results in a capacity achieving concatenated code, provided that the outer codes can achieve the capacity too. However, evaluating (5) seems to be a difficult task.

B. Equal error probability rule

The probability of incorrect decoding of a binary linear block code C can be obtained as

$$p_e \leq \sum_{j=d}^{N} A_j Q\left(\sqrt{\frac{\boldsymbol{E}[L_i]}{2} j}\right)$$

where A_i are weight spectrum coefficients of code C, and d is its minimum distance. Since it is in general difficult to obtain code weight spectrum, and union bound is known to be not tight in the low-SNR region, one can use simulations to obtain a performance curve for the case of AWGN channel and some fixed (probably, non-ML) decoding algorithm, and use least squares fitting to find suitable α and δ, so that the decoding error probability is given by

$$p_e(m) \approx \alpha Q\left(\sqrt{\frac{\boldsymbol{E}[L_i]}{2}\delta}\right).$$

Assume now that the outer codes C_i are selected from some family of error-correcting codes (not necessary polar) of length N. Let K_t, D_t and $P_t(m)$ be the dimension, minimum distance and decoding error probability function for the t-th code, respectively, where m is the expected value of LLR.

Figure 4 presents a simple algorithm for construction of a generalized concatenated (multilevel) code of rate R according to the equal error probability rule. The algorithm employs the bisection method to approximately solve the equation $\sum_{i=1}^{2^l} K(i,P) = RN2^l$, where $K(i,P)$ is the maximum dimension of a code capable of achieving error probability P at the i-th subchannel. The parameter ε is a sufficiently small constant, which affects the precision of the obtained estimate for P. The code is optimized for the case of AWGN channel with noise variance σ^2. The algorithm returns the dimensions of optimal codes for each level, as well as an estimate for the decoding error probability for each code.

CodeOptimization(σ, R, N, l)

$\mathbf{E}[L_1^{(1)}] \leftarrow 2/\sigma^2$
Compute $m_i = \mathbf{E}[L_{2^l}^{(i)}], 1 \le i \le 2^l$ via (5)–(6)
$P' \leftarrow 1; P'' \leftarrow 0$
while $P' - P'' > \epsilon P'$
do $\tilde{P} \leftarrow (P' + P'')/2$
  $t_i \leftarrow \arg\max_{t: P_t(m_i) \le \tilde{P}} K_t, 1 \le i \le 2^l$
  $K \leftarrow \sum_{i=1}^{2^l} K_{t_i}$
  if $K < RN2^l$
    then $P'' \leftarrow \tilde{P}$
    else $P' = \tilde{P}$
return $(K_{t_1}, \ldots, K_{t_{2^l}}), \tilde{P}$

The SC/multistage decoder produces an error if decoding of any of the component codes is incorrect. Therefore, the overall error probability of the GCC can be computed as

$$\begin{aligned} P &= 1 - P\{C_1, \ldots, C_n\} \\ &= 1 - P\{C_1\} P\{C_2 \mid C_1\} \cdots P\{C_n \mid C_1, \ldots, C_{n-1}\} \\ &\approx 1 - \prod_{i=1}^{n} \left(1 - P_{t_i}(m_i)\right) \approx 1 - (1 - P)^n \end{aligned}$$

where C_i denotes the event of correct decoding of the outer code at the i-th level, P is the quantity computed by the above algorithm, and t_i is the index of the code selected for the i-th subchannel. This expression enables semi-analytic prediction of the performance of the concatenated code, based on the available performance results for component outer codes.

C. Decoding complexity

One can use any suitable algorithm to implement soft-decision decoding of outer codes in the GCC obtained either by decomposing a polar code, or constructed explicitly using the algorithm in Figure 4. Box-and-match algorithm is one of the most efficient methods to perform near maximum likelihood decoding of short linear block codes [20]. Its worstcase complexity for the case of (N, K) code with order t reprocessing is given by $O\left((N-K)K^t\right) = O\left(N^{t+1}\right)$, although in practice it turns out to be much more efficient. Decoding of a concatenated code of length $\nu = Nn$ involves decoding of N inner codes using the SC algorithm, and decoding of n outer codes. Therefore the overall complexity is given by $O\left(N^{t+1} n C_b + Nn \log n C_s\right)$, where C_b and C_s are some factors which reflect the cost of elementary operations performed by these algorithms.

VI. Numerical Results

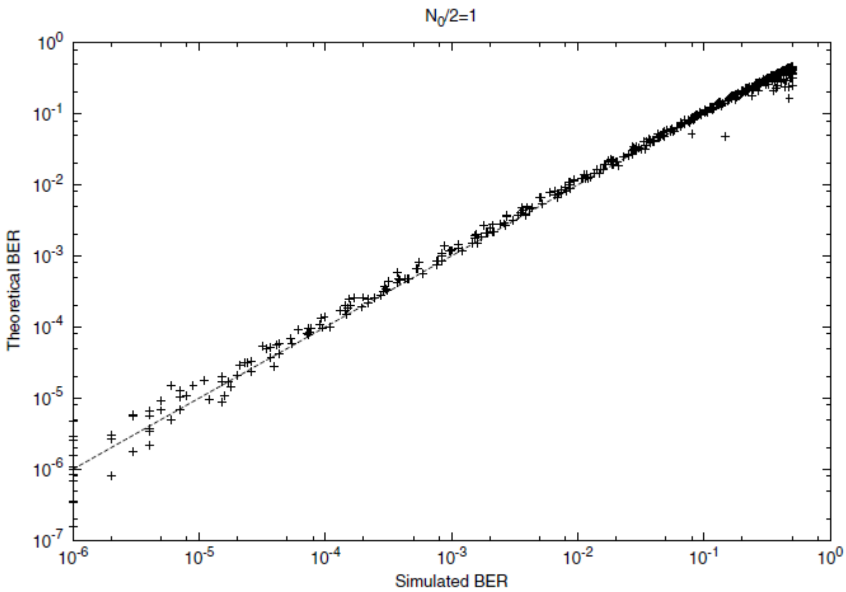


Fig. 5. Accuracy of Gaussian approximation.

Figure 5 presents simulation results illustrating the accuracy of bit error rate analysis based on the Gaussian approximation.

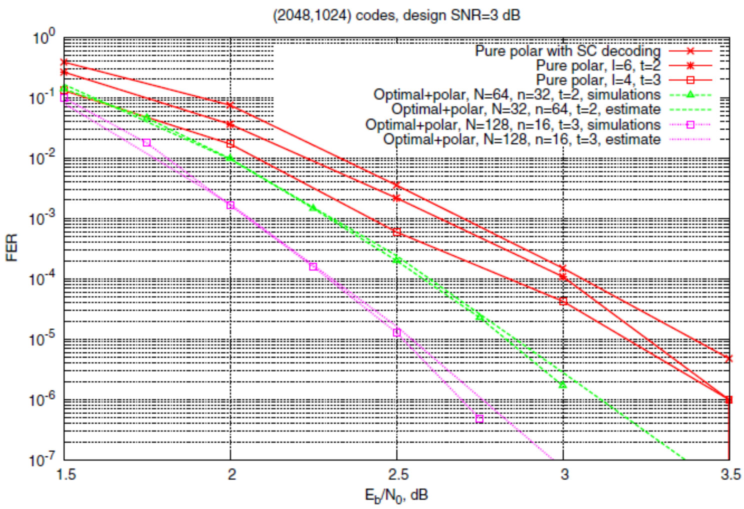


Fig. 6. Performance of polar and concatenated codes.

Figure 6 presents the performance of polar codes of length 2048 designed using the Gaussian approximation method for the case of AWGN channel with Eb/N0 = 3 dB. For multistage decoding, degree *l* decomposition of the original polar code was performed, and box-and-match algorithm with order *t* reprocessing was used for decoding of outer polar codes.

It can be seen that block-wise decoding of outer codes provides up to 0.25 dB performance gain compared to SC decoding. Higher values of N do not provide any noticeable performance improvement. The figure presents also the performance of GCC based on inner polar codes and outer optimal linear block codes with multistage decoding. It can be seen that increasing the length of outer codes provides additional 0.5 dB performance gain. This is due to much higher minimum distance of optimal codes compared to polar codes of the same length, obtained by decomposing the polar code of length Nn.

CS Journal Club, Apr. 18, 2013

Aliasing-Free Wideband Beamforming Using Sparse Signal Representation

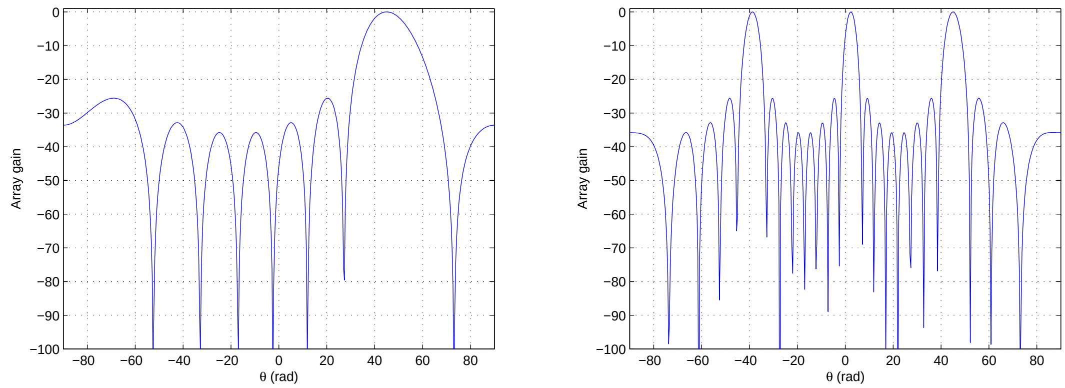
Authors: Zijian Tang, Gerrit Balcquiere, and Geert Leus
Journal: IEEE Trans. on. Sign. Proc. July, 2011.
Presenter: J. Oliver

Abstract

This paper considers the use of sparse signal representation for the wideband direction of arrival (DOA) or angle of arrival estimation problem. In particular, this paper discusses about the two ambiguities, namely, spatial and algebraic aliasing that arise in wideband-DOA. The authors of the paper suggest procedures to avoid the aliasing using multiple measurement vector and multiple dictionaries.

Introduction and Background

- A beamformer is a processor used in conjunction with an array of sensors to provide spatial filtering. The sensor array collects spatial samples of propagating wave fields, which can be processed by the beamformer.
- The objective of a beamformer is to estimate the signal arriving from a desired direction in the presence of noise and interfering signals. A beamformer thus performs spatial filtering to separate signals that have overlapping frequency content but originate from different spatial locations.
- Estimating the spatial locations (or directions) is a well-known problem in array signal processing.
- Three major DOA estimation techniques are 1. Classical methods (Delay-sum beamformer, MVDR) 2. Subspace methods (MUSIC, ESPRIT) 3. ML-based methods
- This paper discusses about beamforming and in particular wide-band beamforming.
- DOA estimation by beamforming can be subjected to ambiguity called spatial aliasing [1].
- Spatial aliasing occurs when the spacing, d, between the sensors is larger than half of the apparent wavelength, that is, $d > \lambda/2$ (See Fig. 1)
- We note from the figures, the resolution increases as d increases, but spatial aliasing also increases.
- This paper discusses how to avoid spatial aliasing (if there is any) in a wideband setting.
- Various authors [2-5] have studied sparse representation (SSR) for narrowband DOA estimation in various contexts. In [2], CS is applied to reduce the ADC sampling rate, in [3,4] it is used to improve angle resolution, in [5] it is used to reduce hardware complexity. All these works assume spatial aliasing is not present.
- However, in SSR based methods aliasing (or ambiguity) comes not only from spatial aliasing, but also from the over-completeness of the dictionary (algebraic aliasing).
- This paper discusses, how to avoid both spatial and algebraic aliasing.



(a) Array tuned to 45 deg with $d = 0.5\lambda$

(b) Array tuned to 45 deg with $d = 1.5\lambda$

Figure 1: Illustration of spatial aliasing

- In summary, the spatial aliasing can be avoided by using multiple dictionaries and the robustness to algebraic alaising can be obtained by using multiple measurement vectors.

Data model

- A uniform linear array (ULA) comprised of N channels indexed by which are equally spaced on a line with spacing d. It receives signals radiated from Q point sources.
- The signal at each channel after time-sampling is partitioned into P segments, where for each segment, K frequency subbands are computed by e.g., a filter bank or the discrete Fourier transform (DFT).
- Let $S_{q,k}(p)$ denote the kth subband (frequency) coefficient computed for the pth segment of the signal that is radiated from the Qth target; similarly, let $y_{n,k}(p)$ denote the kth subband (frequency) coefficient for the pth segment of the signal received at the nth channel.
- With narrow-band assumption, the received signal at the nth sensor at the kth DFT bin is given [1] by

$$y_{n,k}(p) = \sum_{q=0}^{Q-1} e^{j2\pi f_k \frac{d}{c} n \sin\theta_{m_q}} \; S_{q,k}(p) \tag{1}$$

- The aim of this paper is to estimate the target DOAs $\{\theta_0, \theta_1, \cdots, \theta_{Q-1}\}$
- The matrix-vector form of Eqn. (1) is

$$\boldsymbol{y}_{k,p} = \sum_{q=0}^{Q-1} \boldsymbol{a}_{k,m_q} \; S_{q,k}(p) = A_k \boldsymbol{S}_{k,p} \tag{2}$$

where $\boldsymbol{a}_{k,m_q} = \left[1, e^{j2\pi f_k \frac{d}{c} 1 \sin\theta_{m_q}}, \cdots, e^{j2\pi f_k \frac{d}{c}(N-1)\sin\theta_{m_q}}\right]^T$ is called array response vector and A_k steering matrix.

Assumption 1: The array response vectors corresponding to different targets are mutually independent.

Classical beamforming

- Classical beamforming (in this paper, delay-sum beamformer) sets the beamformer coefficients corresponding to a single target angle.
- For example, if the beamformer wants to listen to angle $\theta = 30\,\text{deg}$, then it sets its coefficient vector as $\left[1, e^{j2\pi f_k \frac{d}{c} 1 \sin\frac{\pi}{6}}, \cdots, e^{j2\pi f_k \frac{d}{c}(N-1)\sin\frac{\pi}{6}}\right]^T$ and forms the product $\boldsymbol{a}_{k,m}^H \boldsymbol{y}_{k,p}$
- The angle domain is divided into M points $\Theta = \{\theta_0, \cdots, \theta_{M-1}\}$
- In many applications, such as sonar, a range (time)-bearing(angle) image is desired which can be made by repeating the above procedure for all the subbands; the outputs are then combined and transformed back into the time domain by means of e.g., an inverse Fourier transform.
- In the end, the signal for the pth segment at the mth angle in the range-bearing image $I(p,m)$ can be computed as

$$I(p,m) = \left| \frac{1}{N} \sum_{k=0}^{K-1} \boldsymbol{a}_{k,m}^H \boldsymbol{y}_{q,k}(p) e^{j2\pi f_k p} \right|^2 \tag{3}$$

- $I(p,m)$ can be interpreted as the power of the output of a spatial-temporal filter steered to the direction θ_m. The DOAs are estimated by seeking those θ_m whose corresponding values in $\sum_{p=0}^{P-1} I(p,m)$ are the largest.
- The delay-sum beamformer is subject to spatial aliasing. That is, when the spacing d is larger than apparent wavelength, it is possible to find another $\theta_{m'} \neq \theta_m$ such that for an arbitrary integer j

$$f_k \frac{d}{c} \sin\theta_m = f_k \frac{d}{c} \sin\theta_{m'} + j \tag{4}$$

  holds and thus $\boldsymbol{a}_{k,m} = \boldsymbol{a}_{k,m'}$, which gives multiple peaks in the range-bearing image $I(p,m)$.

DOA estimation via SSR

Problem formulation

- Divide the whole angle search range into a fine grid $\Theta = \{\theta_0, \theta_1, \cdots, \theta_{M-1}\}$.
- Each θ_m corresponds to a certain array response vector $\boldsymbol{a}_{k,m}$, which depends on f_k.
- Construct $N \times M$ steering matrix $A_k = [\boldsymbol{a}_{k,0}, \cdots, \boldsymbol{a}_{k,M-1}]$ (dictionary)
- Assumption 2: The DOAs of the targets $\{\theta_{m_0}, \theta_{m_1}, \cdots, \theta_{m_{Q-1}}\} \in \Theta \quad \Omega = \{m_0, m_1, \cdots, m_{Q-1}\}$
- Data model : $\quad \boldsymbol{y}_{k,p} = A_k \boldsymbol{x}_{k,p} \quad \boldsymbol{x}_{k,p}$ is a Q-sparse signal $(Q < N)$
- We have P such snapshots (measurement vector), then we can form $Y_k = A_k X_k \quad Y_k$ is $N \times P$, A_k is $N \times M$ and X_k is $M \times P$
- Assume that the DOAs during the span of P snapshots remain unchanged, then the columns of X_k share a common sparsity.
- Let $\mathcal{R}(A)$ denote an operation that collects the indexes of all the nonzero rows of a matrix A. $\mathcal{R}(X_k) = \Omega$ and $|\mathcal{R}(X_k)| = Q$.
- With these notations, we can formulate the sparse recovery problem as

$$\min_{\widehat{X}_k} \; |\mathcal{R}(\widehat{X}_k)| \quad \textit{subject to} \quad Y_k = A_k X_k \tag{5}$$

Aliasing Suppression

- As mentioned earlier, spatial aliasing occurs if d is larger than half of the apparent wavelength, which leads to similar columns in the steering matrix
- In classical beamforming, the DOAs are sought by steering a beamformer to different potential angles.
- However, the SSR-based method recovers X_k first and then estimates the DOAs by locating the rows of X_k that contain dominant entries.
- The over-completeness of the SSR dictionary gives rise to non-unique solutions and thus ambiguity in DOA estimation, which is termed as algebraic aliasing.
- Algebraic aliasing is essentially related to the "goodness" of the sensing matrices (steering matrix) for the DOA recovery.

Proposition 1: Under Assumption 1, if the number of targets Q and channels N satisfy

$$N > 2Q - rank(Y_k) \tag{6}$$

then the SSR-based method will not suffer from algebraic aliasing.

Proof: Algebraic aliasing will not exist if we can find unique solution $\widehat{X}_k$ satisfying $Y_k = A_k \widehat{X}_k$. This is only possible if the Kruskal-rank of A_k is larger than $2Q - rank(Y_k)$ [6, Theorem 2.4]. Since A_k is a Vandermonde matrix, whose Kruskal-rank is equal to its rank, N.

Kruskal-rank (or k-rank) of a matrix A is defined as the largest integer r for which every set of r columns of A is linearly independent.

Remarks:

- If $rank(Y_k) = 1$ ($P = 1$), then $Q < N/2$, that is we can discriminate at most $N/2$ targets.
- On the other side, $rank(Y_k) \leq rank(X_k) \leq Q$, suggests that $Q < N$.
- Thus, using multiple measurement vectors the authors argue that it is possible to counter the algebraic aliasing.

So far, we have concentrated on the data model for a single frequency f_m. We can obtain different measurements and different dictionaries $A_k \neq A_l$ if we use different frequency $f_k \neq f_l$. We will next show that using multiple dictionaries enables us to eliminate spatial aliasing.

- Let Γ_k denote the support of all possible DOA solutions for the k-th dictionary
  $\Gamma_k = \{\mathcal{R}(\widehat{X}_k^{(0)}), \mathcal{R}(\widehat{X}_k^{(1)}), \cdots\}$
- Spatial aliasing is frequency-dependent, which means that for different center frequencies, the resulting ambiguity will not (completely) overlap. Therefore, we can imagine that if we solve Eqn. (5) for several frequencies: $f_0, f_1, \cdots, f_{K-1}$ and combine the solutions in a judicious way, the ambiguity due to spatial aliasing will at least be reduced.

Theorem 1: With Proposition 1 met, if there exist at least two dictionaries, whose corresponding frequencies, say f_k and f_l , satisfy

$$0 < |f_k - f_l| < \frac{c}{2d} \tag{7}$$

then the intersection of the solution support related to different dictionaries will contain exclusively the target DOAs, i.e.,

$$\bigcap_k \Gamma_k = \Omega \tag{8}$$

Proof: With proposition 1 satisfied, we can exclude the ambiguity due to algebraic aliasing and need to focus only on spatial aliasing. Let us proceed with a counter-example. Suppose θ_m is one of the target angles and $\theta_m \neq \theta_{m'}$ is spatial aliasing contained in both dictionaries corresponding to f_k and f_l , which implies that $\{\theta_m, \theta_{m'}\}$ belongs to both Γ_k and Γ_l. In accordance with Eqn. (4) we then have

$$f_k\frac{d}{c}\sin\theta_m - f_k\frac{d}{c}\sin\theta_{m'} = j_1$$

$$f_l\frac{d}{c}\sin\theta_m - f_l\frac{d}{c}\sin\theta_{m'} = j_2$$

$$f_k\frac{d}{c}\sin\theta_m - f_k\frac{d}{c}\sin\theta_{m'} - f_l\frac{d}{c}\sin\theta_m + f_l\frac{d}{c}\sin\theta_{m'} = j_1 - j_2 = j_3$$

where j_1, j_2, j_3 are integers and j_1, j_2 are not equal to 0. Using trignometric identities, the above equations can be written as

$$-2f_k\frac{d}{c}\sin\frac{\theta_m - \theta_{m'}}{2}\cos\frac{\theta_m + \theta_{m'}}{2} = j_1 \tag{9}$$

$$-2f_l\frac{d}{c}\sin\frac{\theta_m - \theta_{m'}}{2}\cos\frac{\theta_m + \theta_{m'}}{2} = j_2 \tag{10}$$

$$-2(f_k - f_l)\frac{d}{c}\sin\frac{\theta_m - \theta_{m'}}{2}\cos\frac{\theta_m + \theta_{m'}}{2} = j3 \tag{11}$$

Since $0 < |f_k - f_l| < \frac{c}{2d}$, it is only possible for Eqn. (11) to hold if the integer $j_3 = j_1 - j_2 = 0$. On the other hand, from the above equations we know that j_1 and j_2 are not zero and they cannot be equal. Therefore, the angle $\theta_{m'}$ cannot be contained simultaneously in Γ_k and Γ_l, which concludes the proof. A judicious choice of frequencies can not only prevent spatial aliasing, but also enhance the performance in a noisy environment.

Aliasing-Free SSR Recovery

Based on the analysis in the previous section, the authors formulate the following multi-dictionary (MD) joint optimization problem with the joint-sparsity constraint:

$$\min_{\widehat{X}_k} \ |\mathcal{R}(\widehat{X}_k)| \quad for\ k = 0, 1, \cdots, K-1$$

$$subject\ to \quad Y_k = A_k X_k, \quad and\ \ \mathcal{R}(\widehat{X}_k) = \mathcal{R}(\widehat{X}_l)\ for\ k \neq l \tag{12}$$

whose solution will be free from any ambiguity under Theorem 1.

The authors have not proposed any new algorithm. They have used OMP in their simulations.

Numerical Examples

1. The authors demonstrate their approach using synthetic and real data.
2. For both cases, they considered ULA with $N = 16$ hydrophones, with a spacing of $d = 0.06\ m$. The speed of the signal wave is assumed to be $c = 1500\ m/s$
3. In the synthetic data they consider two sinusoids $Q = 2$ with frequencies $f_0 = 25$ kHz and $f_1 = 35$ kHz. The DOA are $\{35°, 39°\}$. The search gird is defined as $\Theta = \{-90°, -89.75°, \cdots, 90°\}$. With P=100 snapshots, each dictionary has a dimension of 16×720

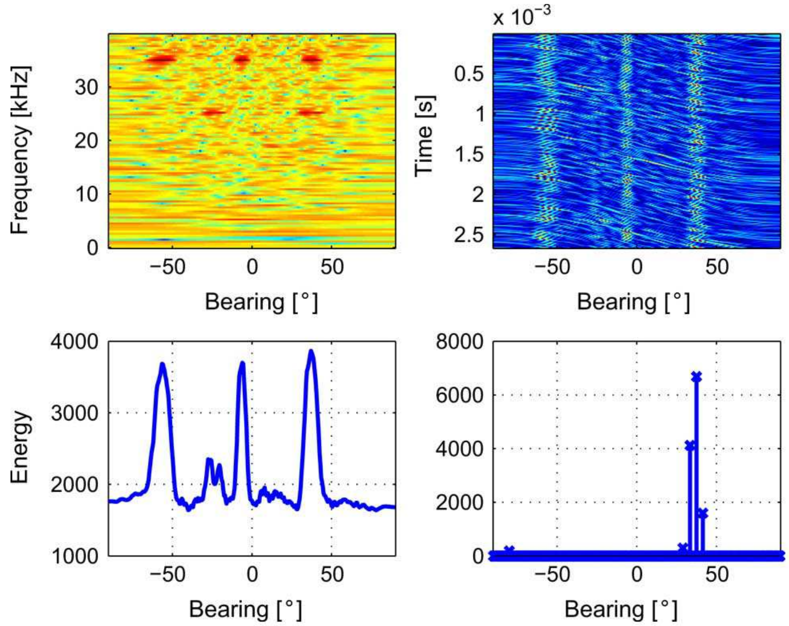


Figure 2: Comparison of classical beamforming with the proposed method. Upper-left subplot: the frequency-bearing image after beamforming; upper-right subplot: the time-bearing image after beamforming; lower-left subplot: the integrated energy of the time-bearing image; lower-right subplot: the result yielded by the proposed method.

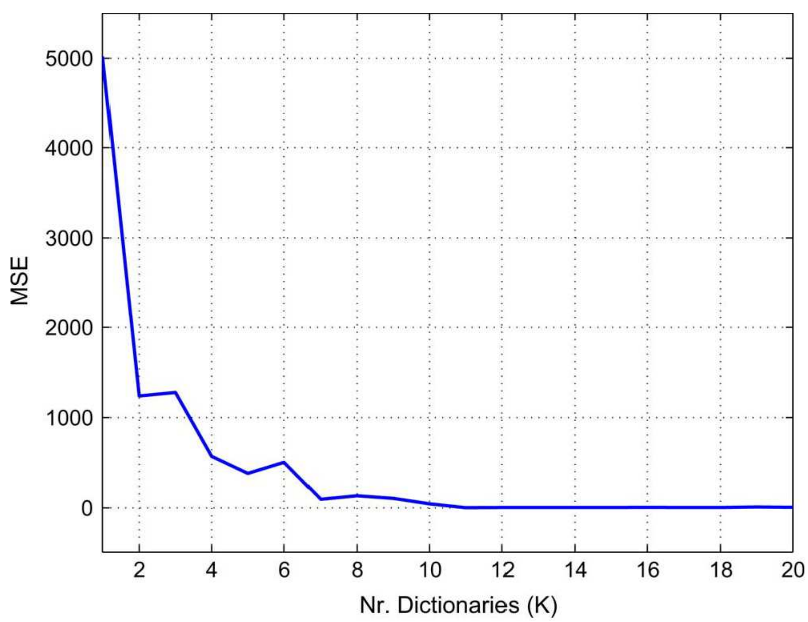


Figure 3: MSE performance against the number of utilized frequencies

4. They halt OMP after 5 iterations. MSE $= \frac{1}{N_s}\sum_{q=0}^{Q-1}(\widehat{\theta}_m - \theta_m)$

1. In the real data experiment, the direction of the divers has to be estimated based on their exhaling sound. Two divers who are 150 m away from the hydrophone are considered. The received signals are from 52° and 60°, respectively.

2. We can see that the frequencies lower than 10 kHz are completely useless for DOA estimation: the diver signal is subdued by the ambient noise dominated by the ship traffic in the harbor.

3. In the midfrequency range (between 10 and 12.5kHz), where the hydrophone array is not subject to aliasing, there is a strong interference signal at a direction around −40°, which possibly comes from a

departing ship blowing the horn.

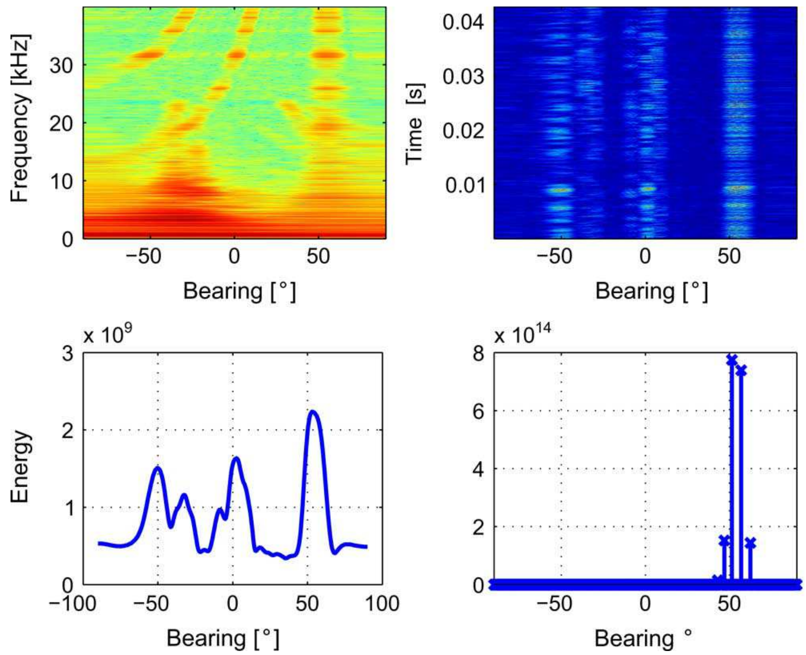


Figure 4: Comparison of classical beamforming with the proposed method for the diver signal. Upper-left subplot: the frequency-bearing image after beamforming; upper-right subplot: the time-bearing image after beamforming (only signals above 25 kHz are taken); lower-left subplot: the integrated energy of the time-bearing image; lower-right subplot: the result yielded by the proposed method.

In Summary, the authors have applied sparse signal reconstruction for DOA estimation (for ULA). They formed an MD optimization problem with joint sparsity constraints. They show how to avoid ambiguities (spatial and algebraic) by using multiple dictionaries and multiple measurement vectors, respectively. They have demonstrated their findings through synthetic and real-life examples.

References

[1] H. L. van Trees, Optimum Array Processing (Detection, Estimation and Modulation Theory, Part IV). New York: Wiley, 2002.

[2] V. Cevher, A. C. Grbz, J. H. McClellan, and R. Chellappa, Compressive wireless arrays for bearing estimation of sparse sources in angle domain, in Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP), Mar. 2008, pp. 24972500.

[3] J. J. Fuchs, On the application of the global matched filter to DOA estimation with uniform circular arrays, IEEE Trans. Signal Process., vol. 49, no. 4, pp. 702709, Apr. 2001.

[4] D. Malioutov, M. Cetin, and A. S. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Trans. Signal Process., vol. 53, no. 8, pp. 30103022, Aug. 2005.

[5] Y. Wang, G. Leus, and A. Pandharipande, Direction estimation using compressive sampling array processing, presented at the IEEE Workshop on Statistical Signal Processing, Cardiff, U.K., 2009.

[6] J. Chen and X. Huo, Theoretical results on sparse representations of multiple-measurement vectors, IEEE Trans. Signal Process., vol. 54, no. 12, pp. 46344643, Dec. 2006.

[7] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, Distributed compressive sensing, 2009. Available: http://arxiv.org/abs/0901.3403