COA 2016 中国海洋声学研讨会
IEEE/OES China Ocean Acoustics Symposium
Harbin, China | January 9-11, 2016

Kernelized Method for Compressive Beamforming

YAN Liying, ZHAO Hangfang
College of Information Science and Electronic Engineering
Zhejiang university
Hangzhou, P. R. China
hfzhao@zju.edu.cn

***Abstract*—Beamforming is a powerful processing tool for direction-of-arrival (DOA) estimation. The performance of a beamformer is limited by physical characteristics of the array in use, and searching space is often crowded in the form expressed in Euclid space. When the input space is transformed into a feature space with high dimension, e.g. infinite dimension in Hilbert space, it exhibits highly sparse characteristics. On that basis, the beamforming problem can be included in the framework of the compressive sensing (CS). In this paper, we map the sampled data into a high dimensional feature space with sparse representation induced by a kernel function. The DOA estimation is solved by using a kernel-based matching pursuit algorithm. When compared with the traditional CS, the new method improves the feature separability, and produces a better performance distinguishing coherent sources with fewer snapshots.**

Keywords—Beamforming; compressive sensing; kernel OMP

I. INTRODUCTION

Beamforming is a useful processing tool for signal enhancement or directional steering. Conventional beamforming is a spatial matched filter, which is a trade-off between width of mainlobe and level of sidelobe. It suffers from low resolution or high level of sidelobe. The minimum variance distortionless response (MVDR) beamformer uses prior information, namely the covariance matrix of interference and noise, to get highest response at the direction of the signal whilst nulling ones at interferences'. Further, the multiple signal classification (MUSIC) method utilizes the subspace decomposition and orthogonal projection concept to obtain higher resolution. These methods go beyond the resolution limit by exploiting signal construction in a cross-spectral matrix. However, their performances become poor when suffering from a mismatch problem in an uncertain environment with poor prior information. Also, they get stuck with few snapshots where cross-spectral matrix comes as rank-deficient. Moreover, these methods cannot deal very well with coherent source problems. Compressive sensing (CS) enables DOA estimation with super resolution and a requirement for fewer samples due to it exploiting sparsity structure of the received signal.

CS can reconstruct a K-sparse dominant elements signal $\mathbf{x}$, from a linear measurement $\mathbf{y}$. Fortunately, in real circumstances, the signals impinging on the sensor array exhibit highly sparse characteristics in spatial space, hence the CS method is suitable for DOA estimation. Correspondingly, many algorithms are developed for CS. The Matching Pursuit (MP) is a basic recipe in the grid search methods, whose variants come as Orthogonal Matching Pursuit (OMP), Compressive Sampling MP (CoSaMP), etc. They are all linear algorithms and need pre-constructed dictionaries. However, elements in dictionaries cannot be perfectly orthogonal due to coherence, which is caused by limited spatial sampling. Moreover, elements are dense in low dimensional space, leading to poor separability and resolution. These inherent disadvantages can lead to incorrect or biased results.

These abovementioned problems can be solved in a high-dimension feature space. It is apparent that each iteration step in OMP is checking inner products of data and candidates, and then choosing a maximum value. This reminds us that the OMP procedure can be performed via kernel methods into a higher dimensional Hilbert space. Hilbert space is a complete inner product space with infinite dimensions, which holds a better property of separability and countability [1]. A special case of Hilbert space is the Reproducing Kernel Hilbert Space (RKHS). With a reproducing kernel function, data in input space is projected into a high dimensional feature space, while the inner product in it can be calculated through the kernel function. In this paper, we take the advantage of the Kernel OMP (KOMP) algorithm to improve computational efficiency and estimation accuracy of DOA. By a careful choice of a kernel function, we map the sampled data into a Hilbert space and play the kernel trick to improve the separability within the given dictionaries. As compared with the traditional CS, the new method provides better feature separability and has enhanced the performance.

II. COMPRESSIVE BEAMFORMING

A. Sparse reconstruction

In many engineering problems, the signals to be reconstructed, or the parameters to be estimated, have a sparse characteristic in a certain space, meaning that only several dominant values occupy finite dimensions or basis. Therefore, the observed data can be recognized as a weighted linear combination of these bases. The aim of sparse signal reconstruction is to estimate these combination coefficients, which are also known as sparse coefficients [2]. This kind of problem can be easily tackled under the framework of compressive sensing, and generally formulated as following statements.

Let an N-dimension vector $\mathbf{x} \in \mathbb{C}^N$, be the unknown sparse signal. It has only K nonzero elements and K is far smaller

Financial support for this research is provided by Chinese NSF under Grants 61379008, and the Fundamental Research Funds for the Central Universities.

978-1-4673-9978-4/16/$31.00 ©2016 IEEE

than N. The measurement data $\mathbf{y} \in \mathbb{C}^M$ is a noisy linearly observation of signal $\mathbf{x}$, that $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$ .Where, $\mathbf{A}$ is a known sensing matrix with dimension of $M \times N$. It is actually a product of matrix $\Psi_{M\times N}$ and $\Phi_{N\times N}$. $\Psi_{M\times N}$ represents the measurement matrix, such as a space sampling matrix. And $\Phi_{N\times N}$ is a transform matrix such as a discrete Fourier transform (DFT) matrix [3]. Usually M is smaller than N, meaning an underdetermined problem is expected to reconstruct $\mathbf{x}$ from $\mathbf{y}$. However, the underdetermined equation may not have a unique solution. Fortunately, given a prior knowledge of the sparsity of $\mathbf{x}$, this ill problem can be turned into a minimization problem,

$$\min_{\mathbf{x}\in\mathbb{C}^N} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}, \tag{1}$$

where $\|\cdot\|_0$ denotes the ℓ_0 norm, representing the number of nonzero elements in $\mathbf{x}$. The problem in (1) is NP-hard, but can be approximately solved by greedy algorithms [4], e.g. OMP or Subspace Pursuit (SP). This is the reason why we focus on OMP algorithm. Note that (1) is also a nonconvex optimization problem. In [5], it shows that under some conditions, former nonconvex problem could be relaxed to

$$\min_{\mathbf{x}\in\mathbb{C}^N} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{2}$$

(2) is a closest convex optimization problem with fewer computations and its solution has been proved to be equivalent to the solution of (1). When dealing with noise, the constraint in (1) or (2) usually changes to

$$\min_{\mathbf{x}\in\mathbb{C}^N} \|\mathbf{x}\|_i \text{ subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \le \varepsilon, \tag{3}$$

where $i = 0$ or 1, ε is the setting limitation of error tolerance.

B. *Compressive beamforming for DOA estimation*

DOA estimation is an important part in radar or sonar system, utilizing a specific array to acquire the locations of targets of interest. In some well-known localization methods, namely: conventional beamforming, MVDR beamforming or the MUSIC method, the estimation accuracy and resolution are influenced by the geometry of the array, and snapshots used for covariance matrix calculation and direction of coherent sources. However, the truth is that the distribution of sources gets a sparse characteristic, which means they appear only at a few directions. It apparently makes sense to handle the DOA estimation under the framework of CS.

Assume that all the sources lie in the far field related to the linear array, so that the incident sound waves can be modeled as plane waves. Further, that the sound speed and the frequency of sources are accurately measured. We confine the problem as a narrowband process in 2D Cartesian coordinate and are only concerned about the direction rather than amplitude.

We partition the searching direction into an N-point grid from 0° to 180°, forming a N dimension angle vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_N]$, and use a linear receiver array with M sensors. For every particular source, the received signal of a conventional beamformer is related with the propagation delay from the source to each sensor in the array,

$$\begin{aligned} y_m(t) &= \sum_{k=1}^{K} x_k(t - \tau_m) \\ &= \sum_{k=1}^{K} x_k(t) e^{-j2\pi f \tau_m}. \end{aligned} \tag{4}$$

The second part in (4) follows a narrowband approximation. The time delay at the mth sensor τ_m is corresponding with phase shift in the frequency domain, and f is the source frequency.

For each snapshot, with the presence of additive Gaussian noise in temporal and spatial sampling, we build the measuring equation

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{5}$$

where $\mathbf{y}$ is a complex $M \times 1$ dimensional vector, and $\mathbf{x}$ is the unknown $N \times 1$ dimensional sparse vector with K nonzero entries filling with complex signal. The ith column in $\mathbf{A}$ represents the steering vector of the source incident from θ_i,

$$\mathbf{a}(\theta_i) = \frac{1}{\sqrt{M}} \left[1, e^{-j2\pi f d_2 \cos\theta_i}, \ldots, e^{-j2\pi f d_M \cos\theta_i}\right]^T. \tag{6}$$

$d_m \cos\theta_i$ in (6) is the specific time delay τ_m, where d_m is the distance between the first and the mth sensor. T denotes transpose. In this paper, the noise is assumed to be independent and identically distributed complex Gaussian.

III. KERNEL SPARSE RECONSTRUCTION

A. *Orthogonal matching pursuit*

As stated above, a common method for solving the problem in (3) is the greedy algorithm OMP. The core idea of OMP is to pick the columns in the sensing matrix that are most correlated with the residual between the real output and estimated result, since the estimation is the linear combination of these columns. OMP has a quick rate of convergence and low computation burden. The computational complexity for a $M \times N$ sensing matrix is $O(MN)$ .Given a sensing matrix $\mathbf{A}$, an output $\mathbf{y}$ and a threshold for error ε, the main procedure of OMP comes as below:

1) Initialization: choose the residual $\mathbf{r}_0 = \mathbf{y}$. Set the index set as an empty set $\Lambda_0 = \varnothing$, so as the selected columns set $A_0 = \varnothing$. Initialize the counter $t = 1$.

2) Main iteration:

a) Calculate the correlation between the residual with each column of sensing matrix. Find the index of the maximum value and its corresponding column: $\lambda_t = \arg\max_{i=1,\ldots,N} \langle \mathbf{r}_{t-1}, \mathbf{a}_i \rangle$.

b) Update index set and selected columns set: $\Lambda_t = [\Lambda_{t-1}, \lambda_t]$, $A_t = [A_{t-1}, \mathbf{a}_{\lambda_t}]$.

c) Compute the solution of a least-square problem: $\hat{\mathbf{x}}_t = \left(A_t^* A_t\right)^{-1} A_t^* \mathbf{y}$.

d) Calculate the new residual: $\mathbf{r}_t = \mathbf{y} - A_t \hat{\mathbf{x}}_t$.

3) Stop rule:If the norm of $\mathbf{r}_t$ is smaller than the threshold ε, stop iteration, otherwise $t = t+1$ and back to *a)*.

Sometimes, the residual cannot reach the threshold. We can add some extra conditions like setting an approximate sparsity level *K*, that, if $t > K$, stop iteration and get the reconstructed signal $\hat{\mathbf{x}}$.

The success of sparse reconstruction depends on many factors, e.g. the incoherence of sensing matrix, orthogonality of bases, noise level and so on. Due to the limited number of sensors, the coherence between two adjacent columns in a sensing matrix is unavoidable, making it intractable to distinguish two adjacent sources. There are already some ideas to try to solve this in physical way, such as using random arrays [6], [7]. In our paper, we try to produce a purely mathematical solution that taking advantage of the infinite dimension of Hilbert space to enhance the separability by using the kernel trick for sparse reconstruction.

B. Kernel sparse reconstruction

Consider a mapping function $\phi(\bullet)$, it transforms the original data in $\mathbb{C}^M$ to a high dimensional space $\mathcal{H}$: $\phi(\bullet): \mathbb{C}^M \mapsto \mathcal{H}$. Apply this function to both the measured data $\mathbf{y}$ and sensing matrix $\mathbf{A}$, then we get a sparse representation in $\mathcal{H}$ that,

$$\begin{aligned} &\mathbf{y} \mapsto \phi(\mathbf{y}), \\ &\mathbf{A} \mapsto \phi(\mathbf{A}) = \left[\phi(\mathbf{a}_1), \ldots, \phi(\mathbf{a}_N)\right], \\ &\phi(\mathbf{y}) = \phi(\mathbf{A})\mathbf{x}. \end{aligned} \tag{7}$$

Thus, the constraint in (3) then changes to,

$$\min_{\mathbf{x} \in \mathbb{C}^N} \|\mathbf{x}\|_i \ \textit{subject to} \ \left\|\phi(\mathbf{y}) - \phi(\mathbf{A})\mathbf{x}\right\|_2 \le \varepsilon. \tag{8}$$

However, $\phi(\bullet)$ is usually not explicit, so that the OMP algorithm mentioned above cannot be applied directly. In Hilbert space, there exists a kind of special function connecting the kernel function with the mapping function, known as reproducing kernel functions, $\phi(\mathbf{x}) := K(\bullet, \mathbf{x})$. Reproducing kernel functions has a reproducing property as:

$$f(\mathbf{x}) = \langle f, K(\bullet, \mathbf{x}) \rangle, \ \forall \mathbf{x} \in \mathbb{C}^M, \forall f \in \mathcal{H}. \tag{9}$$

Then, the inner product in $\mathcal{H}$ can be turned into the calculation of kernel function in data space,

$$\left\langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \right\rangle = \left\langle K(\bullet, \mathbf{x}_i), K(\bullet, \mathbf{x}_j) \right\rangle = K(\mathbf{x}_i, \mathbf{x}_j). \tag{10}$$

(10) is the well-known kernel trick, which simplifies the complex computation in a high dimensional space [8]. Since the key calculation in OMP is the inner product, all the inner product operations in the OMP algorithm can be replaced by kernel function. Therefore, the kernel sparse reconstruction in a high dimensional space is available via the kernelized greedy algorithm, or kernel orthogonal matching pursuit (KOMP). The specific transformation is given below:

$$\begin{aligned} \left\langle \phi(\mathbf{y}), \phi(\mathbf{a}_i) \right\rangle &= K(\mathbf{y}, \mathbf{a}_i), \\ \hat{\mathbf{x}}_t &= \left(\phi(A_t^*)\phi(A_t)\right)^{-1} \phi(A_t^*)\phi(\mathbf{y}) \\ &= K^{-1}(A_t, A_t) K(A_t, \mathbf{y}), \\ \left\langle \phi(\mathbf{r}_t), \phi(\mathbf{a}_i) \right\rangle &= \left\langle \phi(\mathbf{y}) - \phi(A_t)\hat{\mathbf{x}}_t, \phi(\mathbf{a}_i) \right\rangle \\ &= K(\mathbf{a}_i, \mathbf{y}) - \hat{\mathbf{x}}_t^* K(A_t, \mathbf{a}_i), \\ \left\|\phi(\mathbf{r}_t)\right\|^2 &= \left\langle \phi(\mathbf{y}) - \phi(A_t)\hat{\mathbf{x}}_t, \phi(\mathbf{y}) - \phi(A_t)\hat{\mathbf{x}}_t \right\rangle \\ &= K(\mathbf{y}, \mathbf{y}) - K(A_t, \mathbf{y})\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_t^* K(A_t, \mathbf{y}) \\ &\quad + \hat{\mathbf{x}}_t^* K(A_t, A_t)\hat{\mathbf{x}}_t. \end{aligned} \tag{11}$$

Substitute (11) into OMP algorithm and its initialization, we get the realization of KOMP.

IV. SIMULATION RESULTS

In this section, we apply this kernel version OMP, KOMP into simulated DOA estimation problems to verify its effectiveness in situation of coherent sensing matrix. KOMP's performance is in comparison with OMP. Since the accuracy of DOA estimation is the main issue, the metric of performances is the mean square error (MSE) for angle estimation. Suppose in one scenario, the number of sources is *K* and Monte Carlo of simulation is *L*, then it gives,

$$MSE = \frac{1}{LK}\sqrt{\sum_{t=1}^{L}\sum_{k=1}^{K}\left(\hat{\theta}_{k,t} - \theta\right)^2}. \tag{12}$$

Two simulations are conducted to evaluate the performance of OMP and KOMP. First, we set the localization of two sources lying in three different angle intervals, whose corresponding steering vectors in the sensing matrix are in high correlation. This is to examine the distinguishing ability of both algorithms. Second, for multiple sources, two methods are performed for comparison. Without other specific explanations, the array for simulations is a 16-element uniform linear array (ULA) with half-wavelength sensor spacing. The kernel function we choose is the common widely used Gaussian kernel function with universal approximating capability,

$$K(\mathbf{x}_1, \mathbf{x}_2) = \exp\left(-\left\|\mathbf{x}_1 - \mathbf{x}_2\right\|^2 / 2\sigma^2\right), \tag{13}$$

where σ is an important parameter to be designated, also known as kernel bandwidth or kernel size [9]. It plays a certain role of smoothing. If σ is too large, all the data projected to $\mathcal{H}$ will look similar, thus making it no different in original space. While σ is too small, the inner product will be close to zero then all data looks distinct. There is still no explicit rule for the choice of σ, requiring experience and experiments to determine its value in a certain situation.

A. *Two coherent sources localization*

To demonstrate a better accuracy and resolution of KOMP in compressive beamforming, two coherent sources with same frequency of 1000Hz and propagation speed is 1500m/s. The angle interval of sources is selected from the coherence of the sensing matrix. For a uniform linear array, Fig. 1 shows the coherence of a sensing matrix. We pick three representative cases that: (I) [40°, 136°] with coherent value of 0.03, (II) [70°, 75°] with coherent value of 0.42, and (III) [8°, 165°] with coherent value of 0.79, respectively.

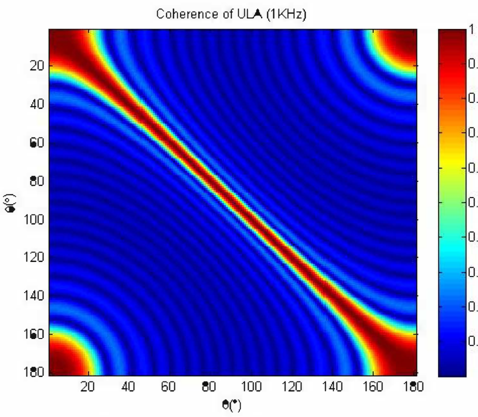


Fig.1. Coherence of sensing matrix of a 16-sesonr ULA with half-wavelength spacing of 1kHz sources

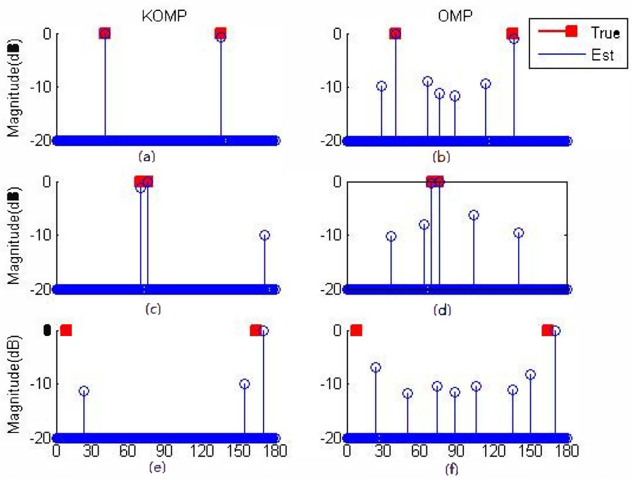


Fig.2. Compressive beamforming results of KOMP and OMP with SNR 20dB. The first column is KOMP results the other is OMP. Two sources are detected by the ULA in Fig.1. and located respectively at (a)(b) [40°, 136°], (c)(d) [70°, 75°]and (e)(f) [8°, 165°]

The beamforming results are shown in Fig.2. The first column represents the results of KOMP while the second one is the outputs of OMP, both under the additive white Gaussian noise environment with SNR 20dB and the grid for reconstruction is [0°:1°:180°]. The locations and coherent values respectively are [40°, 136°] (0.03), [70°, 75°] (0.42) and [8°, 165°] (0.79) from top to bottom. Apparently, both performances are getting worse when the coherent value becomes larger. However, the results clearly indicate that KOMP has fewer estimation errors than OMP, no matter what the coherent value is. Accordingly the misjudgment rate is reduced. As for accuracy, it seems as though there is no significant difference between two methods from Fig.2. Therefore, we repeat each simulation for 1000 times Monte Carlo under the same condition to calculate the mean square error in (12).

Considering the effect of mis-estimates, we have to suppose the number of sources (K = 2 in this simulation) is known. Only the first two largest signals are selected. Specially, we set a threshold which is 1/10 of the biggest signal amplitudes, since Fig.2(e) and (f) show that the estimated signal amplitudes for the rest are around -10dB, as compared to the largest one. An estimate is regarded as an effective one only when it is bigger than the threshold. MSE in different situations are given in Table.1. Results in 0.03 coherent value have distinct difference in an order of magnitude. As for the second case, MSE of KOMP is a little higher in the permitted range than OMP. In last simulation, KOMP achieves a comparable accuracy compared with OMP although it is in the same order of magnitude. Taking into the account the accuracy and the number of mis-estimates, KOMP is much better than OMP.

TABLE I. MSE OF KOMP AND OMP IN DIFFERNET SITUATION

	Location of sources and correlation		
	[40°,136°], *0.03*	[70°,75°] ,*0.42*	[8°,165°] ,*0.79*
KOMP	0.0056	0.0299	0.9081
OMP	0.015	0.0281	1.4826

B. *Multiple sources localization*

It is common sense that a N-element array has N-1 degrees of freedom (DOF) [10]. In normal conditions, only a number of N/2 DOF is available for a conventional beamforer [11]. While the kernel method maps data into an infinite Hilbert space, the DOF of the array should be enlarged. We first let 9 sources located randomly at the DOAs of 40°, 55°, 70°, 80°, 90°, 100°, 120°, 135° and 145° with the other conditions unchanged for comparison between KOMP and OMP. Fig.3 indicates that KOMP and OMP both could recognize the DOAs of all true sources. In addition, the KOMP does not have any mis-estimate, while OMP still suffers from false source problem. Furthermore, there are some mis-estimates with a level of about -5dB, making them easily recognized as a source.

For further research, we simulate the same 9 sources in different SNR and add a little disturbance in the array location to test these two methods performance in uncertain and noisy environment. The range of SNR is [-20:5:25]dB. And the disturbance we added satisfies uniform distribution with the upper bound are respectively $[0:0.05:0.2]\lambda$, where λ is the wavelength. The half wavelength is 0.75m, which is the ideal spacing between sensors. So for the worst situation, the error of array position is nearly 15cm when the disturbance is about 0.2 wavelength. Again, we repeat each scenario for 1000 Monte Carlo simulations. Considering that both methods may mis-estimate under high noise background with uncalibrated array element position, we take average hit ratio into account

978-1-4673-9978-4/16/$31.00 ©2016 IEEE

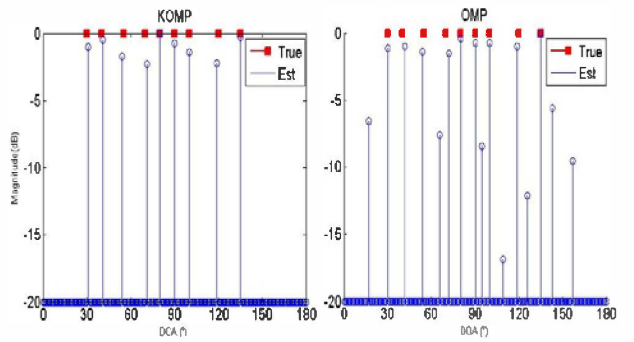
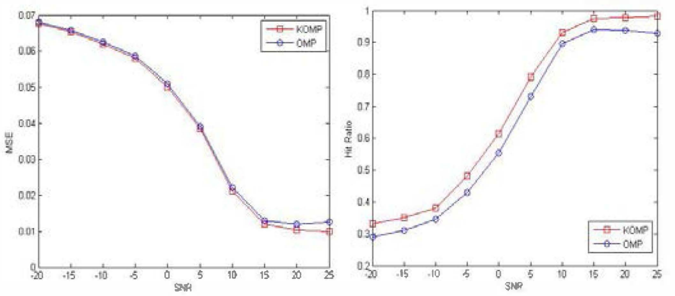


Fig.3. Estimation results of 9 sources located at $40^{\circ}, 55^{\circ}, 70^{\circ}, 80^{\circ}, 90^{\circ}, 100^{\circ}, 120^{\circ}, 135^{\circ}, 145^{\circ}$, $f = 1\text{kHz}$, $\text{SNR} = 20\text{dB}$, a 16-elemnts ULA with half-wavelength spacing.

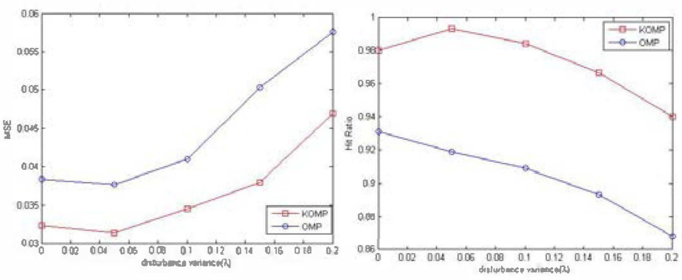
as another performance metric besides MSE. Hit ratio describes the percentage of valid estimates . In this simulation, a valid estimation is adopted when the estimation is within 2° from real source direction.

Fig.4 represents the curve of MSE and hit ratio changing with SNR. In general, MSE is decreasing when SNR becomes higher, while the hit ratio is increasing for both methods. The difference between the two methods in MSE is subtle, but KOMP is a littler lower (nearly 10^{-3}) than OMP. While the hit ratio shows much more explicitly that KOMP always outperforms OMP by about 4% to 6%. Results about their performances in uncalibrated array element position are shown in Fig.5. As we can see, KOMP and OMP both get worse with disturbance getting bigger. But still, KOMP has lower MSE and higher hit ratio than OMP and the difference is obvious this time.



(a) (b)

Fig.4. (a) MSE and (b) hit ratio of DOAs estimation for 9 sources via SNR. Monte Carlo of simulation is 1000, $f = 1\text{kHz}$, a 16-elemnts ULA with half-wavelength spacing.



(a) (b)

Fig.5. (a) MSE and (b) hit ratio of DOAs estimation for 9 sources via disturbance in array location satisfied uniform distribution. Monte Carlo of simulation is 1000, $f = 1\text{kHz}$, $\text{SNR} = 20\text{dB}$, a 16-elemnts ULA with half-wavelength spacing.

V. CONCLUSION

In this paper, we worked on a DOA estimation problem under the CS framework and Hilbert space. We map measured data and sensing matrix into high dimensional RKHS, then employ a kernel sparse reconstruction algorithm, KOMP method, to solve this sparse reconstruction problem. We conducted different simulations and verified that KOMP gives a much more excellent performance as compared to OMP. KOMP is more accurate even in the situation where the sensing matrix shows a strong coherence. In a multiple source localization scenario, KOMP can estimate the majority of these with less probability of error than OMP, even in low SNR or with uncertain array element position.

REFERENCES

[1] R. A. Kennedy, P.Sadeghi, Hilbert space methods in signal processing Cambridge Uk, 2013.

[2] H. Li, Y. Gao, J. Sun, "Fast Kernel Sparse Representation," Digital Image Computing Techniques and Applications (DICTA), 2011 International Conference on. IEEE, 2011, pp.72-77.

[3] A. Xenaki, P.Gerstoft, K.Mosegaard, "Compressive beamforming," Journal of the Acoustical Society of America, vol.136(1), pp.260-271, 2014.

[4] Y.Chen, N. M. Nasrabadi, T. D.Tran, "Hyperspectral image classification via kernel sparse representation," Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE, 2011, pp.1233-1236.

[5] E. Cand` es, J. Romberg, and T. Tao, "Stable signalrecovery from incomplete and inaccurate measurements," Communications on Pure and Applied Mathematics, vol.59, pp. 1207–1223, 2006

[6] C. F. Gaumond and G. F. Edelmann, "Sparse array design using statistical restricted isometry property," J. Acoust. Soc. Am. vol.134,pp.191–197,2013.

[7] E. J. Candes, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," Appl. Comput. Harmon.Anal.vol.31, pp.59–73,2011.

[8] B.Scholkopf and A. J.Smola, Learning With Kernels. MIT press, USA ,2001.

[9] W.Liu, J.C.Principe, S.Haykin, Kernel Adaptive Filtering: A Comprehensive Introduction, Neural Networks (IJCNN), The 2013 International Joint Conference on. IEEE, 2011.

[10] H. L. Van Trees, Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory. New York: Wiley, 2002.

[11] J. Li, P.Stoica, MIMO radar with colocated antennas, Signal Processing Magazine, IEEE, vol.24(5), pp.106-114,2007.