2018 5th International Workshop on Compressed Sensing applied to Radar, Multimodal Sensing, and Imaging (CoSeRa)

Relevant Vector Identification using Matrix Extension for Anisotropic SFCW Radar

Yun Lu
Chair for RF Engineering
TU Dresden, Germany
Tel.: 0049 351 463 39948
Email: Yun.Lu1@tu-dresden.de

Dirk Plettemeier
Chair for RF Engineering
TU Dresden, Germany
Email: Dirk.Plettemeier@tu-dresden.de

***Abstract*—During decades microwave imaging technology has achieved remarkable progress especially by the introduction of the compressed sensing (CS), wherein the sparse modeling plays an important role. In practice, however, the sparsity is not always existing. Even for the sparse case, the signal is usually contaminated by non-sparse vectors, e.g. noise or model-mismatch. Unfortunately, the energy or variance of these non-sparse vectors are unknown, which is thus problematic for the typical sparse decoders, e.g. BPDN. Alternatively, if the sensing matrix holds the quotient property (QP), robust recovery is still possible without giving noise information. Nevertheless, matrices holding QP are usually the random matrix, e.g. Gaussian matrix. Most sensing matrices are non-Gaussian, often are ill-conditioned and anisotropic. In this paper, we will improve the noise mitigation method (NMM), which was developed in our previous work, to estimate the unknown non-sparse vectors, e.g. noise or error by model mismatch. An iterative extension method will be introduced to combat the dilemma of the basic NMM.**

I. INTRODUCTION

Sparse signal contaminated by non-sparse signal is very often to see in practice. For example, there is a noisy measurement $y \in \mathbb{C}^m$, which can be sparse represented in sensing matrix $A \in \mathbb{C}^{m\times n}$ by

$$y = Ax + e, \tag{1}$$

where x is k-sparse. If A holds the k-order null space property (NSP) [1], a good estimation of x can be obtained by applying BPDN $(\Delta_\eta^A(y))$

$$\hat{x} = \arg\min_x \|x\|_1 \quad \text{s.t.} \quad \|y - Ax\|_2 \leq \eta \tag{2}$$

and $\|e\|_2 < \eta$. The same is true for the LASSO-solution $\arg\min_x \|y - Ax\|_2^2 + \lambda\|x\|_1$, where the regularization parameter is proposed to be $\lambda \sim \sigma\|A^*y\|_\infty$. Clearly, there is problematic if neither η nor σ is known. Alternatively, if the sensing matrix A holds the QP, then BP $(\Delta_{\eta=0}^A(y))$ method can recover the sparse vector robustly without the noise information [3]. Unfortunately, this is not true for our SFCW radar setting, wherein the sensing matrix is often ill-conditioned and anisotropic.

Our proposal, NMM [2], which is based on the matrix extension E, can improve the quotient bound (QB),

$$Q_p^{AE} = \sup_{e\neq 0} \inf_{e=Au+Ev} \frac{\|(u,v)\|_p}{\|e\|_2}, \tag{3}$$

particularly for $Q_2^{AE} = \sigma_{\min}^{-1}(AE)$, such that obtaining an error bound [4]

$$\|x - \Delta_{\eta=0}^{AE}(y)\|_2 \leq \frac{1 + C\sqrt{m/k}}{\sigma_{\min}(\Theta_T)}\|e\|_2, \tag{4}$$

without deteriorating the NSP property much, and Θ_T is a square submatrix of compound matrix $\Theta = (A|E)$, where E is weighted by a proper value. More details can be found in our previous work [4].

A. *Our contribution*

Our work above is focusing on the QB property analysis, where $\Delta_{\eta=0}^{AE}(y)$ is applied. In this work, however, we do not try to improve the QB (compound RIP) as much as possible, instead a 2-stage BPDN recovery, namely

1) solving BP $\{\min\|z\|_1 \quad \text{s.t.} \quad y = \Theta z\}$ and getting a noise estimation $\eta = c\|Ez_E\|_2$ or variance σ^2;
2) applying then BPDN $\{\min\|x\|_1 \quad \text{s.t.} \quad \|y - Ax\|_2 \leq \eta\}$ or LASSO: $\{\min\|y - Ax\|_2^2 + \lambda\|x\|_1\}$.

where z_E is the recovered subvector of z corresponding the specific E space. The noise energy is estimated by $\eta = c\|Ez_E\|_2$. A similar idea can also be found in our previous work [5]. However, this method has its difficult of constructing an extension matrix with proper size. Extension matrix with too large or too small size will deteriorate its performance. In this paper, we will introduce an iterative matrix extension method to deal with this dilemma.

B. *Organization*

This paper is organized as follows: First, we will review the basic model of NMM as well as its dilemma, namely a proper extension matrix construction. Then, an extended NMM in anisotropic SFCW radar will be discussed for solving this dilemma. Finally, we give both theoretical and practical tests to show the feasibility of our proposals.

II. NOISE MITIGATED METHOD

A. *basic NMM*

Let $\Omega \subset [1\cdots m], l = |\Omega|$ and $d \in \mathbb{C}^m$. If the $m \times l$ extension matrix E has pairwise orthogonal columns, i.e., it can be written as $E = WD$ with $D = diag(d_j)_{j\in\Omega}$ and some $W^*W = Id$, we get

$$E^*E = D^*D = diag(|d_j|^2)_{j\in\Omega} \tag{5}$$

and otherwise zeros. Usually, we take $W = U_\Omega$ where $A = USV^*$ is the singular value decomposition of A , $S = diag(\sigma_1, \cdots, \sigma_m)$ are the singular values of A in decreasing order, and U and V are the corresponding unitary matrices. The set Ω denotes a set of l singular values, i.e., $l = |\Omega|$.

In [5], we indicated that by iterative successive cancellative projecting a white noise vector over A, the selected subspace in i-th iteration corresponds that space by i-th singular value of A statistically. Therefore, we set the extension matrix W to be the subset corresponding the space, denoted as U_s, by the last l smallest singular values and weighted by $\rho = \sigma_{m-l}$, where U_s is the subspace of U in $A = USV^*$.

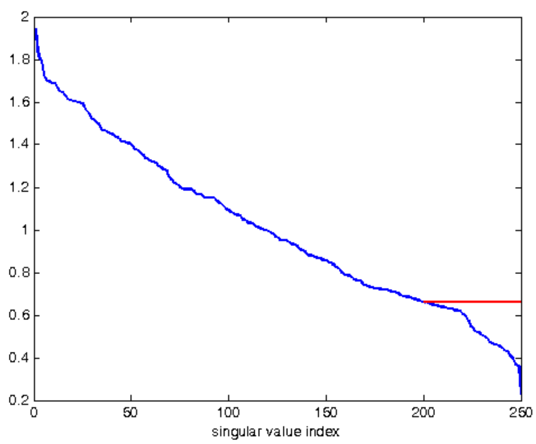

Fig. 1. For example, singular values of $A^{250\times500}$ (blue) and the extension matrix domain (red) of size $l = 50$, extension level $\frac{l}{m} = 50/250$.

If A is extended with this special matrix

$$E = \sigma_{m-l}U_{(m-l+1)\sim m} = \rho U_s, \tag{6}$$

and $\Theta = (A, E)$, then

$$\begin{aligned}
&\mathbb{E}\left\{\left\|diag\left\{\Theta^T ee^T\Theta\right\}\right\|_\infty\right\}\\
&= \left\|diag\left\{\Theta^T\mathbb{E}\left\{ee^T\right\}\Theta\right\}\right\|_\infty\\
&= \left\|diag\left\{\begin{pmatrix} A^T \\ E^T\end{pmatrix}\mathbb{E}\left\{ee^T\right\}(A,E)\right\}\right\|_\infty\\
&= \left\|diag\left\{\begin{pmatrix} A^T \\ E^T\end{pmatrix}U\Sigma U^T(A,E)\right\}\right\|_\infty
\end{aligned} \tag{7}$$

It can be rewritten as

$$\begin{bmatrix} diag(VS\Sigma SV^T)_\infty, & diag(VS\Sigma U^T U_s\rho)_\infty \\ diag(\rho U_s^T U\Sigma SV^T)_\infty, & diag(U_s^T U\Sigma\rho^2 U^T U_s)_\infty \end{bmatrix}$$

$$\leq \begin{bmatrix} \|VS\|_2^2\,\sigma^2, & \left\|VS\Omega^I\rho\right\|_2\sigma^2 \\ \left\|\Omega^I\rho SV^T\right\|_2\sigma^2, & \left\|\Omega^I\rho\right\|_2^2\sigma^2 \end{bmatrix}$$

where $\Omega^I = I^{m\times m}_{(m-l+1)\sim m}$ is a partial identity matrix with indexed diagonal elements from $m-l+1$ to m. Then

$$= \begin{bmatrix} \|VS\|_2^2\,, & \left\|VS_{(m-l+1)\sim m}\rho\right\|_2 \\ \left\|\rho S_{(m-l+1)\sim m}V^T\right\|_2, & \left\|\Omega^I\rho\right\|_2^2 \end{bmatrix}\sigma^2$$

$$\leq \begin{bmatrix} \|S\|_2^2\,, & \left\|S_{(m-l+1)\sim m}\rho\right\|_2 \\ \left\|\rho S_{(m-l+1)\sim m}\right\|_2, & \left\|\Omega^I\right\|_2^2\rho^2 \end{bmatrix}\sigma^2$$

since $S_{(m-l+1)\sim m} \leq \rho$, thus $S_{(m-i+1)\sim m}\rho \leq \rho^2$. Therefore,

$$\begin{aligned}
\left\|S_{1\sim(m-l)}\right\|_2^2\sigma^2 &\geq \left\|I_{(m-l+1)\sim m}\rho\right\|_2^2\sigma^2\\
&\geq \left\|S_{(m-l+1)\sim m}\rho\right\|_2\sigma^2\\
&= \left\|\rho S_{(m-l+1)\sim m}\right\|_2\sigma^2
\end{aligned} \tag{8}$$

This means, statistically the white noise successive cancellative projection will be along the subspace corresponding singular value transition in decreasing order.

Noting that by solving

$$\{\hat{z}_A, \hat{z}_E\} = \arg\min \|z\|_1 \quad \text{s.t.} \quad y = \Theta z \tag{9}$$

for noise estimation, it is necessarily to fulfill

$$\max_{e_i\subset E}\left\|A_\Lambda^\dagger e_i\right\|_1 < 1, \tag{10}$$

such that no signal component[1] will be project over the extension matrix E, where A_Λ is the support space of x. In other words, E must be less correlated with A_Λ in terms of l_1 norm.

Proof: Assuming that x is optimally spanned in A_Λ with $y = A_\Lambda x_\Lambda$. To avoid an alternative $y = E_a x_a$ from E we may use

$$\begin{aligned}
\|x_\Lambda\|_1 &= \left\|A_\Lambda^\dagger A_\Lambda x_\Lambda\right\|_1 = \left\|A_\Lambda^\dagger y\right\|_1 = \left\|A_\Lambda^\dagger E_a x_a\right\|_1\\
&\leq \max_{e_i\subset E}\left\|A_\Lambda^\dagger e_i\right\|_1\|x_a\|_1\\
&< \|x_a\|_1\,.
\end{aligned} \tag{11}$$

wherein the condition $\max_{e_i\subset E}\left\|A_\Lambda^\dagger e_i\right\|_1 < 1$ plays an important role which completes the proof. ■

Finally, we can get a noise estimation by

$$\eta = c\|E\hat{z}_E\|_2 = \frac{m}{l}\|E\hat{z}_E\|_2. \tag{12}$$

with the assumption of a white noise, i.e. $c = \frac{m}{l}$. Therefore, BPDN can be conducted now by

$$\min\|x\|_1 \quad \text{s.t.} \quad \|y - Ax\|_2 \leq \eta. \tag{13}$$

The difficulty in the basic NMM is that (10) must be always fulfilled. It is problematic in practice, since we have no idea about the signal support space. As a result, the construction of the corresponding extension matrix must be very careful. This means, if E with too large size l, there is a danger that signal components will also be caught by E. If E with too small size l, noise estimation performance will be deteriorated. The noise estimation performance of the basic NMM is shown in figure 2. One can notice that the proper extension size is signal sparsity dependent.

[1]This is also the condition for the so-called off-support knowledge, i.e. no signal will be located on E.

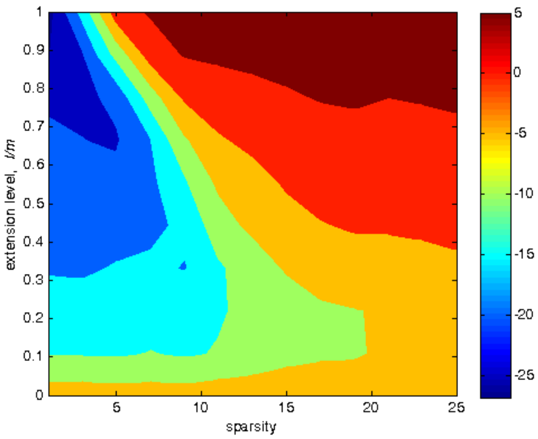

Fig. 2. MSE (dB) of noise estimation $|\frac{\|\hat{e}\|_2 - \|e\|_2}{\|e\|_2}|$, $A^{60\times 120}$

B. extended NMM

In this subsection, we will introduce an iterative extension method to deal with the dilemma in the basic NMM such that an extension with optimal size is possible.

First of all, I give a short review of how the basic SFCW works. Let's take the 1D distance measurement as an example. We collect the data in the frequency domain. Assuming that there is a complex[2] signal plane wave of frequency f_1 as transmit signal

$$x_s(t) = e^{j(2\pi f_1 t + \theta)}. \quad (14)$$

The effective reflected (Hilbert transformed) signal from one object with time delay τ is

$$x_r(t) = \alpha_1 e^{j[2\pi f_1 (t-\tau)+\theta]}. \quad (15)$$

Note that, in a linear time-invariant system, the input frequency f_1 has not changed, only the amplitude and the phase angle of the sinusoid has been changed by the system. Thus, the receiver collects the signal, multiplies by the complex conjugate replica of the transmitted signal, yielding

$$y(f_1) = \alpha_1 e^{j[2\pi f_1(t-\tau)+\theta]} e^{-j(2\pi f_1 t + \theta)} = \alpha_1 e^{-j2\pi f_1 \tau}. \quad (16)$$

Thus, for multiple frequencies and multiple objects system there is

$$y(f_i) = \sum_{k=1}^{N} \alpha_{i,k} e^{-j2\pi f_i \tau_k}, \quad (17)$$

where $f_i \in [f_1, f_2, \cdots, f_n]$ and $k \in [1, N]$ is the index of objects.

Clearly, τ_k is a continuous parameter. In practice, however, we estimate τ_k in a discrete domain and assume that $\tau_k \in [0, \Delta\tau, 2\Delta\tau, \cdots, (n-1)\Delta\tau]$. Therefore, $y(f_i)$ can be spanned in the following domain, denoted as A_0

$$\begin{bmatrix} 1, & e^{-j2\pi f_1 \Delta\tau}, & e^{-j2\pi f_1 2\Delta\tau}, & \cdots & e^{-j2\pi f_1 (n-1)\Delta\tau} \\ 1, & e^{-j2\pi f_2 \Delta\tau}, & e^{-j2\pi f_2 2\Delta\tau}, & \cdots & e^{-j2\pi f_2 (n-1)\Delta\tau} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1, & e^{-j2\pi f_n \Delta\tau}, & e^{-j2\pi f_n 2\Delta\tau}, & \cdots & e^{-j2\pi f_n (n-1)\Delta\tau} \end{bmatrix}. \quad (18)$$

Since the delay τ_k does not live in the absolute phase of the received signal, one can then replace the frequencies $[f_1, f_2, \cdots, f_n]$ by $[0, \Delta f, \cdots, (n-1)\Delta f]$. Finally, A_0 can be rewritten as

$$A_0 = \begin{bmatrix} 1, & 1, & 1, & \cdots & 1 \\ 1, & W, & W^2, & \cdots & W^{(n-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1, & W^{n-1}, & W^{2(n-1)}, & \cdots & W^{(n-1)(n-1)} \end{bmatrix}, \quad (19)$$

where $W = e^{-j2\pi\Delta f \Delta\tau}$. To avoid the ambiguity of the phase rotation, we limit the phase within $[0, 2\pi)$. If there are n samples, then

$$A_0 = \begin{bmatrix} 1, & 1, & 1, & \cdots & 1 \\ 1, & W_0, & W_0^2, & \cdots & W_0^{(n-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1, & W_0^{n-1}, & W_0^{2(n-1)}, & \cdots & W_0^{(n-1)(n-1)} \end{bmatrix} \quad (20)$$

and $W_0 = e^{-j2\pi(\frac{1}{n})}$.

Clearly, A_0 in (20) is a DFT matrix and the received $y(f)$ can be given by

$$\begin{aligned} y(f) &= A_0 \xi(\tau) \\ &= FT\{\xi(\tau)\} = \bar{\xi} \odot FT\{h(\tau)\} \\ &= Ah(\tau) \end{aligned} \quad (21)$$

where ξ is the object response in the time domain and $\bar{\xi}$ is its frequency response, $h(\tau)$ is the channel impulse response. To note is that $\bar{\xi}$ is usually not flat due to e.g. the influence of the antenna, and therefore anisotropic (e.g. see figure 1).

By applying the basic NMM, one can get a robust estimation of the channel impulse response $h(\tau)$. Namely,

$$\begin{aligned} (\Delta^{AE}_{\eta=0}(y)) : \{\hat{z}_E\} &= \min \|z\|_1 \quad \text{s.t.} \quad y = (A|E)z \\ \hat{\eta} &= \frac{m}{l}\|E\hat{z}_E\|_2, \\ (\Delta^{A}_{\hat{\eta}}(y)) : \{\hat{h}\} &= \min \|h\|_1 \quad \text{s.t.} \quad \|y - Ah\|_2 \le \hat{\eta}. \end{aligned} \quad (22)$$

Nevertheless, the selection of an extension matrix E with a reasonable size is difficult.

1) Iterative extension: Recall (10) it is clear that E must be less correlated with signal space especially the support space A_Λ. Generally, an extension with relative high dimension can achieve more accurate noise estimation, yet it will be penalized by capturing the signal as well.

Regarding the noise[3] estimation in (12), the expected value of η is less dependent on the size of l, if (10) is fulfilled. We

[2]In practice, it is only the real part as the radiation signal in the free space. The main reason one would choose to work with complex exponential form of plane waves is that complex exponentials are often algebraically easier to handle than the trigonometric sines and cosines. Specifically, the angle-addition rules are extremely simple for exponentials.

[3]If the additive noise is white.

denote noise estimation by E of size l_i as

$$\eta_{l_i} = \frac{m}{l_i}\|E_{l_i}\hat{z}_E\|_2 \approx m\bar{e}_w = m\frac{\|e\|_2}{m}, \tag{23}$$

where $\bar{e}_w$ is the expectation of $\|e\|_2$ on each orthogonal coordinate. For given SNR and different measurement realizations we have

$$\mathcal{E}_j\{\eta_{l_i}^{(j)}\} = \mathcal{E}_j\{\|e^{(j)}\|_2\}, \tag{24}$$

where j is the realization index, $\mathcal{E}$ is the expectation operator. At the same time, there is

$$\mathcal{E}_{l_i}\{\eta_{l_i}\} = \|e\|_2, \tag{25}$$

over different parallel estimations and $l_i \in [1, m]$.

Therefore, if the signal x is sparse, by iteratively increasing the value of l_i and controlling the variance of $\{\eta_{l_i}\}$, i.e. $var_{l_i}\{\eta_{l_i}\}$, we can get a reasonable size for extension matrix without capturing much signal by E. In our case, we set the extension update criterion

$$\frac{\eta_{l_i}}{\eta_{l_1}} \leq v, \tag{26}$$

where v is a parameter for estimation accuracy. This means, if $\frac{\eta_{l_i}}{\eta_{l_1}}$ remains relative small, the recovered vector $\hat{z}_E$ does not contain signal in high probability. In iterations, the extension level l_i/m will be increased, if (26) is fulfilled. An overview of the extended NMM is follows:

1) $\{\hat{z}_E\} = \Delta_{\eta=0}^{AE_{l_1}}(y), \frac{m}{l_1} = 0.05, \eta_{l_1} = \frac{m}{l_1}\|E_{l_1}\hat{z}_E\|_2$,
2) loop l_i $(i \geq 1)$,
   a) $\frac{m}{l_i} = \frac{m}{l_i} + \Delta(l/m)$,
   b) $\{\hat{z}_E\} = \Delta_{\eta=0}^{AE_{l_i}}(y)$,
   c) $\eta_{l_i} = \frac{m}{l_i}\|E_{l_i}\hat{z}_E\|_2$,

   if $\frac{\eta_{l_i}}{\eta_{l_1}} > v$, end loop;
3) $\{\hat{h}\} = \Delta_{\hat{\eta}_{l_{i-1}}}^{A}(y)$

III. Numerical Results

A. theoretical tests

As for the basic NMM, we also give some theoretical evaluations for the extended NMM. Assuming a Gaussian matrix $A^{m\times n}$ of size 60×120, the received signal y, which is spanned in A, is added with white noise of SNR=20dB. For non-sparse vector estimation, the extension matrix E starts with a small extension level $l/m = 0.05$ (to fulfill the condition in (10) in high probability) and is updated with step size $\Delta(l/m) = 0.1$ and $v = 2$. In this manner, the corresponding extension level is determined automatically. The performance is given by 500 realizations[4] for each sparsity.

In figure 3, one can notice that the extended NMM outperforms the basic NMM much for the sparse case. Regarding the non-sparse case, there will be signal components also recovered on the extension matrix. In this case, the non-sparse vectors are dominated by signal. Good MSE with respect to noise estimation does not indicate a better performance of relevant vector estimation. Noting that both basic NMM and extended NMM provide a confidence level for a stable relevant vector recovery. The difference lies that the extended NMM is more conservative, i.e. consider more signal energy as non-sparse vectors. Finally, its estimated error with respect to noise is worse, yet for the relevant vector identification it is in turn more favorable (stable).

[4]Each time with a new Gaussian matrix A.

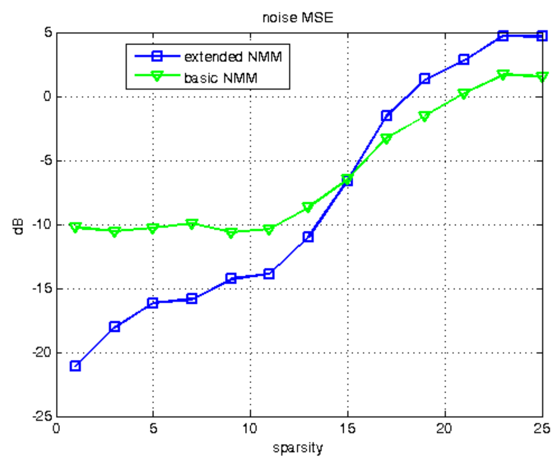

Fig. 3. MSE of noise estimation, $|\frac{\|\hat{e}\|_2 - \|e\|_2}{\|e\|_2}|$

In figure 4, there is the corresponding the average extension level in terms of the signal sparsity. The extension level in the sparse case is higher than that for the less sparse case. This means that signal components distribution in the singular value domain is less uniform. Moreover, signal are more located over directions with larger singular values.

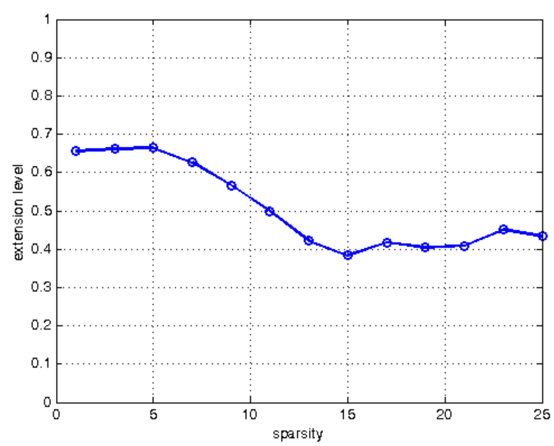

Fig. 4. Average extension level l/m

B. field tests

In this subsection, we will show some practical tests. A scenario is shown in figure 5, which is measured by SFCW principle. There are totally 1001 frequency points at range [0.5, 3] GHz with scan step size of 5 cm in the azimuth direction. Due to the fact that our main scenario is located within particular zone in the time domain, instead of the whole vector of dimension 1001 we select only a small fraction, i.e. dimension of 39, in the time domain. Besides, over-sampling is applied. As a result, the corresponding sensing matrix is of size $A^{39\times 94}$.

Regarding the non-sparse signal in this real scenario there are not only background noise existing, but a lot of compressible signal components, model mis-match components etc..

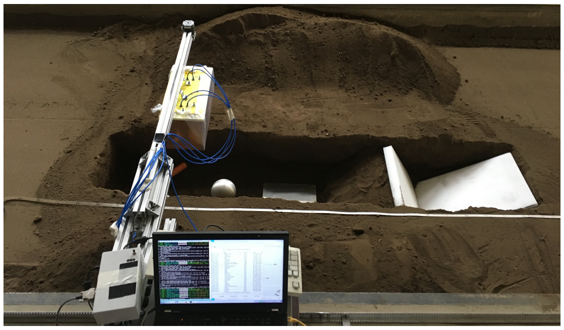

Fig. 5. Scenario (man-made dihedral by soil, metallic plate, sphere, and rods)

In figure 6, there is the raw image, where both relevant and non-sparse signal are existing. After a non-sparse signal estimation by the extended NMM, the relevant information by this scenario is automatically highlighted (see figure 7) and a lot of sidelobes as well as irrelevant signal are disappeared. Noting that the suppressed "irrelevant" components could still be reasonable information. However, these components are not stable with respect to the applied inverse system. Thus, they are considered as irrelevant (or non-stable) signal.

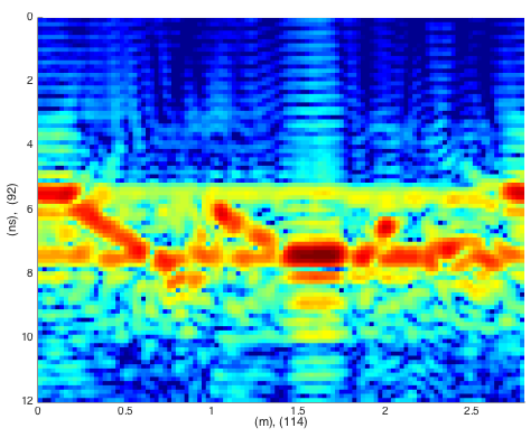

Fig. 6. Raw image

IV. Conclusion

In this paper, we introduced an extended NMM method for stably recovering relevant signal in SFCW radar. It can solve the practical dilemma for extension matrix construction. The basic idea of NMM is constructing a subspace, which is less correlated with signal (support) space, to sense the non-sparse signal, e.g. noise. A stable BPDN is then possible based on the estimated noise energy. The difficulty of the basic NMM is a reasonable and stable extension matrix construction.

In our extended NMM, we use an iterative extension principle, which is based on the ergodic property of the noise,

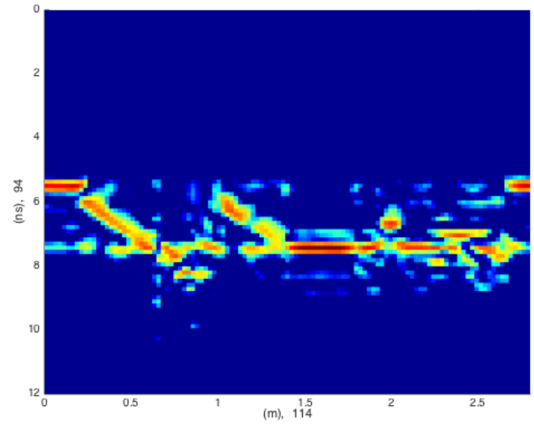

Fig. 7. Extended NMM

to get a reasonable extension automatically, i.e. achieve more accurate noise estimation while keeping signal projection on extension matrix small. So far, our research focus is on the uniformly distributed non-sparse vector, e.g. white noise, in the singular value domain. The question of how to combat non-sparse vectors with non-uniform distributed property is still open, and it will be discussed in our future work.

References

[1] S. Foucart; H. Rauhut, "A mathematical introduction to compressive sensing", ser. Applied and Numerical Harmonic Analysis. New York, NY: Springer New York, 2013.

[2] Y. Lu; C. Statz; V. Ciarletti; D. Plettemeier, "Noise mitigated compressive sensing for radar applications", *Radar Symposium (IRS), 2014 15th International*, 2015.

[3] P. Wojtaszczyk, "Stability and instance optimality for Gaussian measurements in compressed sensing," Foundations of Computational Mathematics, vol. 10, no. 1, pp. 113, 2010.

[4] C. Herrmann; Y Lu; C. Scheunert; P. Jung, "Improving Robustness for Anisotropic Sparse Recovery using Matrix Extensions", *Workshop on Smart Antennas, WSA 2018*, 2018.

[5] Y. Lu; D. Plettemeier, "Robust recovery for radar applications by noise mitigated compressed sensing", *IRS 2015*, 2015.