

Multiple Compressive Projection Measurement for Stepped Frequency Radar

Yun Lu*, Christoph Statz*, Sebastian Hegler*, Wolf-Stefan Benedix*, Valérie Ciarletti[†], Dirk Plettemeier*

*Chair for RF Engineering, Technische Universität Dresden, Germany, Email: Yun.Lu1@tu-dresden.de

[†]LATMOS/CNRS/IPSL/UVSQ, 78280 Velizy, France, Email: Valerie.Ciarletti@latmos.ipsl.fr

Abstract—In modern communication and measurement systems, signal detection and estimation play a major role. Actually, the above two terms can be considered as one issue, e.g. pure detection by densely listing all possible diversities. The penalty is however the system complexity. Up to now, a lot of work have been invested, especially the recent compressed sensing (CS) technique [1], which is a subtle mathematic application in practice and leads to a great success in signal detection both for communication and measurement, e.g. radar technique. In spite of this radical progress there are still a lot of open problems. One of them is the "noise" including background noise and non-ideal signal modelling, which is not just a problem for CS but a general difficulty for signal processing. Although there are many sophisticated recovery algorithms developed to cope with noise, the performance will be usually impacted by inaccurate noise estimation or modelling error. In this paper, we will analytically describe the multiple compressive projection measurement (m CPM¹ or MCPM) introduced in [2]. Both theoretical analysis and numerical evaluations show that m CPM is a promising measurement system.

I. INTRODUCTION

Suppose we have some target signal spanned in a particular frame Φ and provide a generative form

$$t = \Phi w + \xi, \quad (1)$$

where $t \in \mathbb{R}^n$ is the vector of targets, $\Phi \in \mathbb{R}^{n \times d}$ is a frame of d diversities that have been assumed, w is a vector of unknown weights, and ξ is Gaussian noise. The goal is to estimate w given t and Φ . Obviously, the maximum likelihood (ML) solution is very effective, if the dimension of w is smaller than that of the signal, i.e. $n > d$. In case of $n < d$, which is said to be underdetermined, a sparse w is required for a unique solution. Those problems can be handled by a general Bayesian framework [3]. An accurate result can be obtained if precise Φ and a reasonable estimation about ξ are available. Unfortunately, this is usually not the case in practice. That means, mostly we get only an estimation about Φ , denoted as $\hat{\Phi}$, as well as ξ . Thus, even without noise ξ we still cannot get the exact solution. To combat the noise ξ , which is usually non-sparse in Φ , we introduced the m CPM. By using the restricted isometry property (RIP) for both sparse and non-sparse signal as well as the multi-correlation function (MCF) [4], the noise term in (1) can be well suppressed without exact information about the noise level in the channel. This work will focus on

the general signal sensing² of stepped frequency radar (SFR) detecting sub-surface objects.

A. Organisation

The remaining of the paper is organised as follows: The Section II introduces the property of underdetermined linear systems and their solutions. Then, the theoretic detection behavior in compressive domain will be presented in Section III. In Section IV we will introduce the MCF as well as its combination with CS, termed as m CPM. The basic properties of m CPM will be discussed. After that, we apply the m CPM in SFR and formulate new measurement approach for better performance. Finally, the summary and future work are presented.

II. UNIQUE SOLUTION

Observing (1) with $n < d$, there are infinite solutions due to the fact that there are some linearly dependent diversity columns in Φ . In [5] the author gave a sparse condition in the sense of linear algebra for a unique solution, which was derived by clearly distinguishing two different unknown weights w_1 and w_2 . That means, the difference of $\delta = w_1 - w_2$ must be uniquely detectable with respect to (1). In other words, δ is not allowed to lie in the null space of Φ , i.e. $\Phi\delta = 0$. Thus, the columns controlled by δ , i.e. $\Phi_\delta \in \mathbb{R}^{n \times m_\delta}$ and $m_\delta \leq n$, have to be linearly independent. Because of the randomness of δ in practice, Φ_δ ranges over the whole Φ . Finally, the number of non-zero entries in δ , denoted as $\|\delta\|_0 := |\text{supp}(\delta)|$ and $\text{supp}(\delta) = \{j : \delta_j \neq 0\}$, have to be smaller than the *spark*, which denotes the least number of linearly dependent columns in Φ . Formally,

$$\text{spark}(\Phi) = \min_{w \neq 0} \|w\|_0 \quad \text{subject to} \quad \Phi w = 0. \quad (2)$$

Furthermore, observing δ , we will get the maximal $\|\delta\|_0$ if weights w_1 and w_2 are disjoint (extreme case). This means

$$\begin{aligned} \|\delta\|_0 &\leq \|w_1 - w_2\|_0 \\ &= \|w_1\|_0 + \|w_2\|_0 \quad \text{for} \quad w_1 \cap w_2 = \emptyset. \end{aligned} \quad (3)$$

Obviously, (3) must hold the following condition

$$\max\{\|\delta\|_0\} < \text{spark}(\Phi). \quad (4)$$

¹The special case is $m = 1$ as the conventional CS.

²It is not necessarily in sub-Nyquist framework.

Thus, for a general unknown weight w and $\|w\|_0 = K$, we have

$$\max\{\|\delta\|_0\} \leq 2K < \text{spark}(\Phi). \quad (5)$$

At last, the unique solution is given by holding the condition

$$K < \text{spark}(\Phi)/2 \quad \text{and} \quad \|w\|_0 = K. \quad (6)$$

The results in (6) can also be interpreted as: The vector δ still preserves its Euclidean distance approximately after linear projection, i.e. $\Phi\delta$. In other words, it holds

$$\|\Phi\delta\|_2^2 \approx \|I\delta\|_2^2, \quad (7)$$

where I is the identity matrix. This means, the $2K$ -th normalized singular value of Φ is not far from unit, which essentially requires that Φ_δ behaves like an orthogonal system. Similar results can also be found in Johnson-Lindenstrauss (JL) lemma [6]. The JL lemma is concerned with the following problem. Given a set of points in \mathbb{R}^d , we would like to embed these points into a lower-dimensional Euclidean space \mathbb{R}^n ($n < d$) while approximately preserving the relative distances between any two of these points. Let $\epsilon \in (0, 1)$ be given. For every set Q of $\#(Q)$ points in \mathbb{R}^d , if \hat{n} is a positive integer such that $\hat{n} > n_0 = \mathcal{O}(\ln(\#(Q))/\epsilon^2)$, there exists a Lipschitz mapping $f: \mathbb{R}^d \rightarrow \mathbb{R}^n$ such that

$$(1 - \epsilon) \leq \frac{\|f(u) - f(v)\|_{l_2^n}^2}{\|u - v\|_{l_2^d}^2} \leq (1 + \epsilon) \quad (8)$$

for all $u, v \in Q$. That means, the Lipschitz function f as an injective function and the corresponding l_2 -norm distance ratio is bounded by $1 + \epsilon$. The JL lemma leads directly to the restricted isometry property (RIP) in CS [1]: If matrix $\Phi \in \mathbb{R}^{n \times d}$ satisfies the RIP of order K with a constant $\epsilon \in (0, 1)$, such that $(1 - \epsilon)\|w\|_2^2 \leq \|\Phi w\|_2^2 \leq (1 + \epsilon)\|w\|_2^2$ holds for all $w \in \sum_K$, then w is recoverable. Usually, w can be obtained by maximum a posteriori (MAP) estimation like l_p -norm regularization and FOCUSS [8] solving:

$$\min \|\Phi w - t\|_2^2 + \lambda \|w\|_p^p \quad \text{for } 0 \leq p \leq 1. \quad (9)$$

In addition to MAP estimation, the full Bayesian approach [3], e.g. sparse Bayesian learning (SBL), which seeks for the distribution mass, improves the detection performance dramatically. The basic cost function can be given as: $\mathcal{L} = \log|\Sigma_t| + t^T \Sigma_t^{-1} t$, where $\Sigma_t = \sigma^2 I + \Phi \Gamma \Phi^T$ with hyperparameters $\Gamma = \text{diag}(\gamma)$ controlling the variance of entries in w . However, its performance is still very sensitive to inaccurate noise estimation. In the next section we will introduce the m CPM method to cope with "noise" effect.

III. COMPRESSIVE DETECTOR AND m CPM

A. Compressive Detector

To improve the detection performance in underdetermined framework, we have to examine the statistic behavior in compressive domain with respect to the original signal domain.

For example, to distinguish two hypotheses:

$$\begin{aligned} \mathcal{H}_0 &: y = M\xi \\ \mathcal{H}_1 &: y = M(x + \xi) \end{aligned} \quad (10)$$

where $M \in \mathbb{R}^{m_M \times n}$ with $m_M < n$ is a measurement matrix, $\xi \sim \mathcal{N}(0, \sigma^2 I_d)$ is i.i.d. Gaussian noise. Then, the distribution of y can be given as

$$\begin{aligned} f_0(y) &= \frac{\exp[-\frac{1}{2}y^T \Sigma_M y]}{|\sigma^2 M M^T|^{-1/2} (2\pi)^{m_M/2}} \\ f_1(y) &= \frac{\exp[-\frac{1}{2}(y - Mx)^T \Sigma_M (y - Mx)]}{|\sigma^2 M M^T|^{-1/2} (2\pi)^{m_M/2}} \end{aligned} \quad (11)$$

where $\Sigma_M = (\sigma^2 M M^T)^{-1}$. Based on the likelihood ratio test $\Lambda(y) = \frac{f_1(y)}{f_0(y)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \eta$, where η is given for particular false alarm probability $\alpha = \int_{\Lambda(y) > \eta} f_0(y) dy$, we can get an equivalent detector in a compressive domain³:

$$\hat{t} := y^T (M M^T)^{-1} M x. \quad (12)$$

It is easy to show that

$$\begin{aligned} \hat{t}_{\mathcal{H}_0} &\sim \mathcal{N}(0, \sigma^2 x^T P_M^\dagger M x) \\ \hat{t}_{\mathcal{H}_1} &\sim \mathcal{N}(x^T P_M^\dagger M x, \sigma^2 x^T P_M^\dagger M x). \end{aligned} \quad (13)$$

where $P_M^\dagger = M^T (M M^T)^{-1}$.

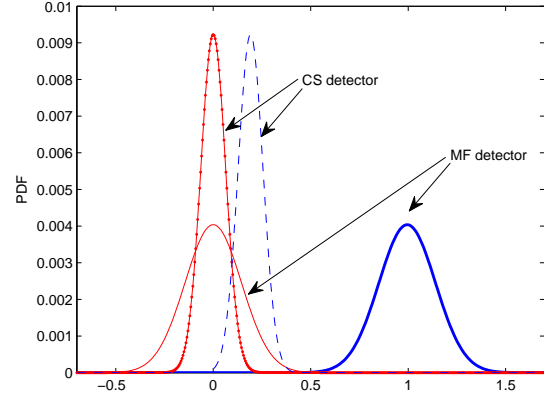


Figure 1. Distribution of hypotheses in detection domain by CS and MF.

If $x = \Phi w$ and $\Phi = I^{d \times d}$, the detection performance in compressive domain is strongly depending on M . For a compressive detector, i.e. M is a wide matrix, x must live in the row span of M with high probability, which is equivalent to require a sparse x . Furthermore, for a Gaussian matrix $M_i \in \mathbb{R}^{m_M \times n}$, $x^T P_{M_i}^\dagger M_i x$ is highly concentrated around $\frac{m_M}{n} \|x\|_2^2$. That is,

$$\mathbb{E}(x^T P_{M_i}^\dagger M_i x) = \frac{m_M}{n} \|x\|_2^2. \quad (14)$$

³ $y^T (M M^T)^{-1} M x \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \sigma^2 \log(\eta) + \frac{1}{2} x^T M^T (M M^T)^{-1} M x$. In case of $M = I$, then $y^T x \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \sigma^2 \log(\eta) + \frac{1}{2} x^T x$.

Finally, with high probability we get the detection performance [9]

$$P_D(\alpha) \approx Q\left(Q^{-1}(\alpha) - \sqrt{\hat{\text{SNR}}}\right) \quad (15)$$

with $\hat{\text{SNR}} = \frac{m_M}{n} \text{SNR}$ and $\text{SNR} = \|x\|_2^2 / \sigma^2$, and $Q(z) = \int_z^\infty \exp(-u^2/2) du$.

1) **Influence of compressive projection:** Observing (15) the detection performance will be strongly affected by the dimension ratio of M , i.e. m_M/n . In case of $m_M/n = 1$, the compressive projection is just an implementation of a traditional detector (matched filter (MF), signal based). Then, we obtain the well-known detection performance:

$$P_D(\alpha) \approx Q\left(Q^{-1}(\alpha) - \sqrt{\text{SNR}}\right). \quad (16)$$

For the case of $m_M/n < 1$, we could still get good recovery performance, if x lies in row span of M with high probability (vector based, e.g. CS, deteriorated detection performance depending on m_M/n but still detectable).

2) **Advantage of compressive projection:** Compressive projection is not always negative for information detection. As we know, the basic feasible solution to (9) is sparse, i.e. $\|w\|_0 \leq n$. However, in some cases like strong noise and coherent signals, the resulting solution still cannot avoid radical over-fitting. Although it can be partly compensated by choosing proper λ , the performance remains too sensitive to inaccurate λ . Besides, one is not allowed to set the λ too large (otherwise w disappears). Alternatively, observe the sparsity bound of solution from (9), it is limited by the row dimension n of Φ . In other words, we can reduce the row dimension to be an appropriate value in need for avoiding radical over-fitting regardless of recovery algorithms. Mathematically,

$$\min \|M(\Phi w - t)\|_2^2 + \lambda \|w\|_p^p \quad \text{for } 0 \leq p \leq 1. \quad (17)$$

If x lies in the row span of M in high probability, the obtained sparser solution (local optimal solution) usually includes weights w or at least part of them (in critical situation). This sparsity controlling, i.e. $\|w\|_0 \leq m_M < n$ with $M \in \mathbb{R}^{m_M \times n}$ and $\Phi \in \mathbb{R}^{n \times d}$, excludes the over-fitting in unfavorable situation dramatically, since over-fitting usually needs particular amount of non-zero entries in solution. Thus, we expect the compressive detector having better performance against over-fitting than that of traditional detector in particular circumstance, namely requiring that reduction of n will not result in severe detection performance deterioration⁴. Yet, a solution without strong over-fitting is not enough for stable detection. Recall that the sparser solution by compressive detector has possibly only part of weights w . To cope with this problem we will introduce the principle of multi-correlation function (MCF).

⁴The row dimension of matrix gives an upper bound of sparsity of feasible solution regardless of recovery algorithms. To control sparsity one can also limit the number of iterations of sparse recovery algorithm (however, it sometimes hampers the convergence). An alternative way for over-fitting controlling is noise mitigation by column dimension extension. In this paper we mainly consider the case of compressive projection, i.e. row dimension reduction.

B. MCF

The MCF is basically an extension of the MF for achieving better correlation properties. The prime idea of MCF is based on the delay and multiply (DAM) property of m-code, which was later extended for generation of Gold codes. The DAM of m-code as well as Gold codes indicates that a transform of one m-code to its other phase delay or a transform of one Gold code to other family members can be realized by applying the DAM operation. Later, this property was applied to other Galois field (GF)-based codes. More information can be found in [10]. A parallel combination of several different correlation functions can result in much better correlation property than $1/\sqrt{N}$ given by the Sarwate lower bound, where N is the sequence length. Thus, it is very favorable for signal detection. Formally, the MCF can be given as

$$C(s) = \frac{1}{m} \sum_{k=1}^m F_k = \frac{1}{mN} \sum_{k=1}^m \sum_{i=0}^{N-1} \mathcal{I}_k(u_i) \mathcal{I}_k^*(v_{i+s}), \quad (18)$$

where m is the number of combinations, u and v are the original input sequences and the reference sequences, respectively. The code transform is given by the transform function \mathcal{I} , e.g. DAM operation as above. The term F_k in (18) could be any kind of "correlation" process or more generally recovery algorithm. The objective of combination of m parallel results from F_k is achieving coherent combination or collecting the partial result from single F_k with respect to information and non-coherent combination of "noise" such that an even better detection scenario.

C. mCPM

The introduction of MCF into compressive detector, which is termed as mCPM here, could be very promising in our particular case and can also facilitate parameter adjustment with respect to information recovery in practice. Basically, mCPM consists of two steps: **i**) the first phase is actually a normal *compressive projection and recovery* (CPR) process by algorithm promoting sparse solution just as (17); **ii**) and the second phase is iteratively updating (combining) sparse solutions from each CPR by different compressive projections, e.g. different measurement matrix M_i , i.e. $\Theta_i = M_i \Phi$, i denotes the i -th iteration. Due to the underdetermined property $y = Mx = \Theta w$ with $\Theta \in \mathbb{R}^{n \times d}$ and $n < d$ as well as Θ holding the RIP with respect to w in high probability, the sparse solution \hat{w} by single CPR includes w with high probability while it provides randomness of noise in terms of their amplitudes and positions, since noise is usually non-sparse. Therefore, a combination of a series of \hat{w}_i

$$w_g = \sum_{i=1}^m \hat{w}_i \quad (19)$$

can result in coherent combination of information components w and non-coherent combination of noise and thus a favorable detection scenario. Finally, the two close distribution modes of CS detector in Fig. 1 would be pulled apart from each other as well as their variance can be decreased depending

on M and m . A schematic illustration of m CPM is presented in Fig. 2. As a result, the radical over-fitting of the solution

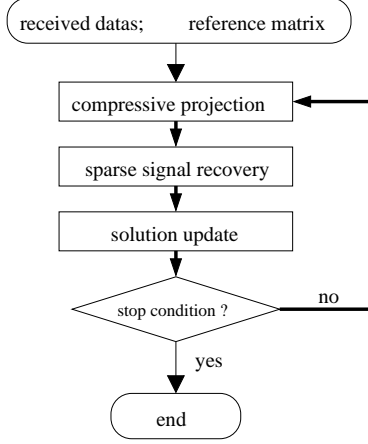


Figure 2. m CPM

can be controlled by compressive recovery while its estimation deterioration is compensated by MCF principle. Fig. 3 shows

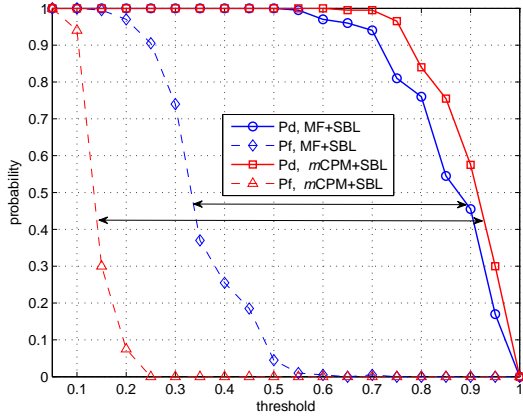


Figure 3. SNR=3dB, $\|w\|_0 = 2$, MF with $\Theta \in \mathbb{R}^{200 \times 200}$, m CPM with $\Theta_i \in \mathbb{R}^{50 \times 200}$ and $m = 10$

the false alarm probability (Pf) and detection probability (Pd) depending on different detection thresholds by normalizing the values within $[0, 1]$. Indeed, lines of Pf and Pd by m CPM are more favorable than that by MF.

Obviously, the performance of m CPM is strongly depending on the compressive projections. An effective way for non-coherent combination of noise term in (19) requires that all compressive projection matrix M_i should be less correlated with each other. Otherwise, the recovered noise term from two different iterations are similar and will also be coherently combined by (19), i.e. yield no contribution for distance expansion between signal term and noise term. Fig. 4 presents the estimation error $e = \|w_g - w\|_2 / \|w\|_2$ by m CPM with $m = 20$ depending on row dimension of M . In case of small row dimension, the estimation error is relatively large, since the resulting reference matrix Θ does not hold the RIP condition with respect to w . For relatively large row dimension of M it also provides increased estimation error. The reason,

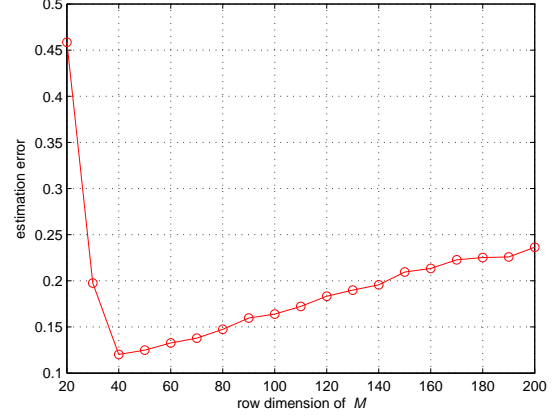


Figure 4. Sparsity $\|w\|_0 = 4$, SNR=10dB, $\Theta = M\Phi$ and $\Phi \in \mathbb{R}^{200 \times 200}$

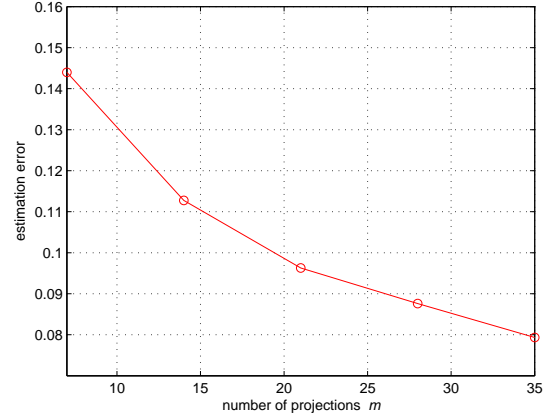


Figure 5. Reference matrix for m CPM $\Theta_i \in \mathbb{R}^{30 \times 200}$, sparsity $\|w\|_0 = 2$, SNR=10dB, the recovery algorithm is SBL, estimated noise factor is 10^{-2} .

as aforementioned, is that except the signal components a lot of recovered noise components also exhibit independence of random M . This means, there are many noise components also live in the row span of the resulting reference matrix Θ , i.e. inefficiency of non-coherent noise combination. Besides, we

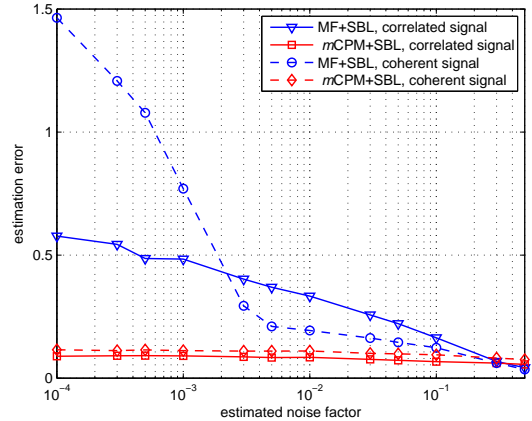


Figure 6. Reference matrix for m CPM $\Theta_i \in \mathbb{R}^{30 \times 200}$, $m = 30$ and for MF $\Theta \in \mathbb{R}^{200 \times 200}$, sparsity $\|w\|_0 = 2$, SNR=10dB.

can freely determine the number of projections m in m CPM for particular performance (see. Fig. 5). Setting a relatively large estimated noise factor the error by MF decreases fast, however, is not recommended, which is equivalent to set large λ in (9) and possibly results in the loss of information. This problem can be solved very well by m CPM, since m CPM still provides relatively small estimation error at low estimated noise factor (see Fig. 6).

IV. m CPM FOR STEPPED FREQUENCY RADAR (SFR)

A. SFR using CS

In the SFR radar [11], it observes the scene with a discrete set of frequencies and synthesizes the impulse in the frequency domain, and brings advantage of better accuracy. They pointed out that CS can be done in frequency domain by randomly measuring all Fourier coefficients. Thus, for rough detection it requires only a small amount of frequency measurements, which can first reduce the measurement time, and second save energy such that long-term activity is possible.

In this paper, we collect the frequency measurements from real circumstance and devices and process them directly in frequency domain. The connection to time domain for ranging information is simply the IFFT transform. Our objective vector, which indicates the ranging information, is w_t . Its corresponding vector in frequency domain is $w_f = Fw_t$, where $F \in \mathbb{C}^{d \times d}$ is the FFT matrix and the dimension d determines the required resolution in time domain. The practical measurement for SFR is basically conducted in frequency domain and the number of frequency measurement points n is usually less than d . Thus, the actual obtained frequency

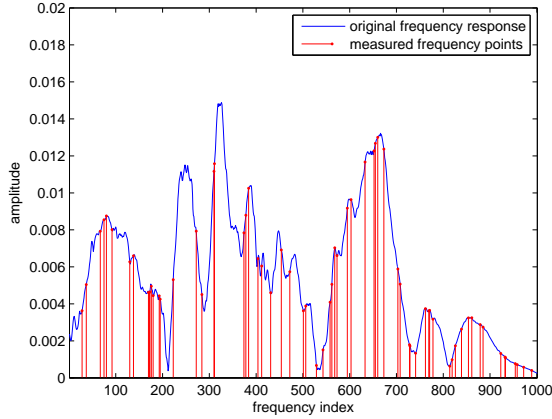


Figure 7. One practical example for randomly sub-sampling only 67 measurements in frequency domain, there are totally 1001 frequencies range from 500MHz to 3GHz.

vector is $w_{fr}^{n \times 1} = R w_f^{d \times 1}$, where $R = [I^{n \times n}, 0^{n \times (d-n)}]$. By using CS, we can furthermore reduce the number of frequency measurements per projection matrix M . Therefore, an underdetermined linear system can be constructed as

$$\hat{w}_{fr} = M \cdot R \cdot F \cdot w_t = \Theta w_t, \quad (20)$$

where M determines which frequency point should be active (see Fig. 7). Finally, one needs to seek w_t by solving

$\min\{\lambda \|w_t\|_p^p + \|\hat{w}_{fr} - \Theta w_t\|_2\}$ with $0 \leq p \leq 1$. This non-

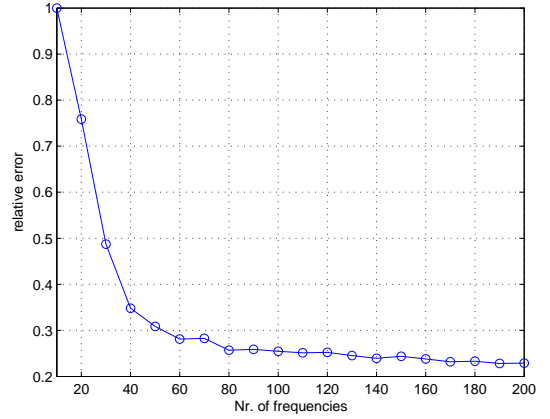


Figure 8. Relative error normed on the recovery error at sampling 10 frequencies.

convex optimization can be realized by FOCUSS or SBL approach. By given prior knowledge that the most reflections are clustered together we can limit the column dimension of Θ for better performance. Fig. 8 presents the relative recovery error by different number of frequency measurements. We can notice that it requires only about 50 frequencies for rough detection. In case of frequency number less than 30 the error increases dramatically. In practice, the SFR can work with following principle: For rough detection the radar can randomly (if necessary) and quickly measure just a few number of frequency points within a defined frequency range as presented in Fig. 7; In case of fine detection, the radar usually tries to collect as many frequency measurements as possible.

B. Full frequency measurement by m CPM

However, how to collect these frequency measurements plays a major role for stable and accurate information recovery. According to the conventional collection scheme, all defined frequency points will be measured once and processed. This approach, however, suffers from non-ideal signal modelling and inaccurate noise estimation especially for recovery algorithms, which are very sensitive to those effects (see Fig. 9).

Alternatively, the radar can work in m CPM mode, i.e. frequency measurements are collected partly, randomly and iteratively. Results from each iteration will be combined. This work mode, as discussed above, provides very stable recovery performance in a non-ideal scenario and facilitates parameter adjustment (see Fig. 10). The results in Fig. 9 and 10 are in the case that the influence of antenna pattern has not been calibrated. Nevertheless, the performance by m CPM is still well and better than the conventional work mode.

V. SUMMARY

In this paper we investigate the stepped frequency radar signal acquisition and processing by CS (or by generalized

m CPM). SFR using CS is a very promising approach for long-term activity. By using m CPM⁵, which can facilitate the fine measurement in practice, it provides even better performance than the conventional work mode. Basically, this work is not considering the feasibility and advantages of CS, rather a stable detection and estimation in noisy situation. The above results shows that even in full frequency mode, the m CPM mode is recommended.

As future work, we will focus on better energy distribution in frequency domain (frequency coding) such that giving better correlation properties in particular area in time domain. The results above are just based on the random selection principle. An optimal selection is still an open problem and also problem dependent. Further research direction is the noise mitigation (NM) by reference matrix extension. Generally [12],

$$\min \|M(\Psi\tilde{w} - t)\|_2^2 + \lambda \|\tilde{w}\|_p^p \quad \text{for } 0 \leq p \leq 1. \quad (21)$$

where $\Psi = [\Phi, R]$ and M^6 , R could be random matrix. The final solution is given by pruning the \tilde{w} , i.e. $w \propto \tilde{w}_\Phi$, where \tilde{w}_Φ is the subvector related with Φ . The NM method is expected to be more stable than row dimension reduction, since the change of row dimension is directly proportional to the recoverable sparsity, while the variation of column dimension is effected by a logarithm factor. More details will be discussed in our forthcoming paper.

Acknowledgment: WISDOM development and this research were supported by funding from the French space agency CNES and from the German space agency DLR.

REFERENCES

- [1] E. Candes; J. Romberg; T. Tao "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information" IEEE trans. inform theory, vol. 52, no. 2, pp. 489-509, Feb. 2006.
- [2] Y. Lu; C. Statz; S. Hegler; D. Plettmeier "Signal Sensing by Multiple Compressive Projection Measurement", *IRS 2013*, Dresden.
- [3] S. H. Ji; Y. Xue; L. Carin "Bayesian compressive sensing". *Signal Processing, IEEE Transactions on*, 56(6):2346-2356, June 2008.
- [4] Y. Lu; A. Finger "Novel multi-correlation differential detection for improving detection performance in DSSS," *International symposium on spread spectrum techniques and applications (ISSSTA)*, Taichung, Taiwan, Oktober 2010.
- [5] Donoho, D.L. and Elad, Michael "Optimally Sparse Representation from Overcomplete Dictionaries via ℓ_1 -norm minimization". *Proc. Natl. Acad. Sci. USA* March 4, 2003 100 5, 2197-2002, 2002.
- [6] W. B. Johnson; J. Lindenstrauss "Extensions of lipschitz mappings into a hilbert space," *In conference in modern analysis and probability*, vol. 26 of Contemp. Mathematic, pp. 189-206, 1984.
- [7] J. Tropp; A. Gilbert; M. Strauss "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit", *Signal Processing*, vol. 86, pp. 572-588, April 2006.
- [8] I. Barrodale; F. Roberts "Applications of mathematical programming to lp approximation", *Symposium on Nonlinear Programming*, pp. 447-464, May 1970.
- [9] L. L. Scharf "Statistical Signal Processing: Detection, Estimation, and Time Series Analysis", Addison-Wesley, 1991.
- [10] Y. Lu; A. Finger "Method and apparatus for the differential multiple correlation in broadband ultrasonic location systems," (Patent) Nr.: EP10177353.9, 2010.

⁵It mainly focuses on stability by modifying the cost function.

⁶ M can be identity matrix or random matrix.

- [11] A. B. Suksmono; E. Bharata; A. A. Lestari "A compressive SFCW-GPR system". in *proc. of 12th Int Conf. on Ground Penet Radar*, Birmingham, pp. 16-19, 2008.
- [12] Y. Lu, C. Statz, M. Mütze, S. Hegler, D. Plettmeier "Noise Mitigation and Extraction of Scatters in Acoustic Imaging by Compressed Sensing", *XVIII-th International Seminar/Workshop direct and inverse problems of electromagnetic and acoustic wave theory*, Lviv, Ukraine, 2013.

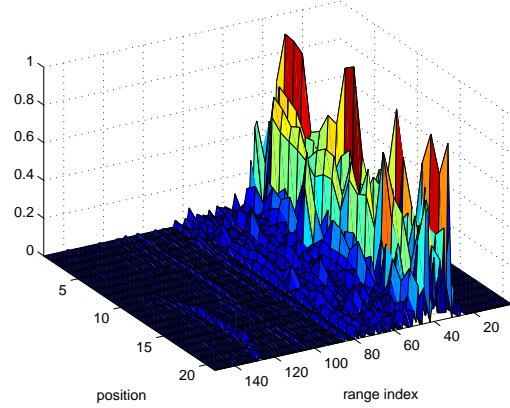


Figure 9. Recovery by full-sampling (1001 frequency points), MF+SBL, estimated noise factor is 0.5×10^{-1} . Due to non-ideal noise estimation the performance is moderate

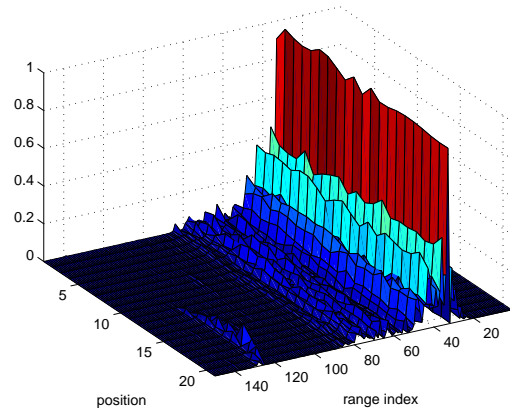


Figure 10. Recovery by randomly sub-sampling (average frequency points for each iteration is 80), m CPM+SBL with $m = 30$, estimated noise factor is 0.5×10^{-1} . Recovery condition was same as in Fig. 9.