COMPRESSED SENSING AND RADIO INTERFEROMETRY

M. Jiang[*], *J. N. Girard*[*], *J.-L. Starck*[*], *S. Corbel*[*], *C. Tasse*[†]

[*]Service d'Astrophysique, CEA Saclay,
Orme des Merisiers
91410 GIF-Sur-YVETTE, France

[†] GEPI
Observatoire de Paris-Meudon
5, rue place Jules Janssen
92190 Meudon, France

ABSTRACT

Radio interferometric imaging constitutes a strong ill-posed inverse problem. In addition, the next generation radio telescopes, such as the Low Frequency Array (LOFAR) and the Square Kilometre Array (SKA), come with an additional direction-dependent effects which impacts the image restoration. In the compressed sensing framework, we used the analysis and synthesis formulation of the problem and we solved it using proximal algorithms. A simple version of our method has been implemented within the LOFAR imager and has been validated on simulated and real LOFAR data. It demonstrated its capability to super-resolve radio sources, to provide correct photometry of point sources in a large field of view and image extended emissions with enhanced quality as compared to classical deconvolution methods. One extension of our method is to use the temporal information of the data to build a 2D-1D sparse imager enabling the detection of transient sources.

Index Terms— sparsity, compressed sensing, interferometry, imaging, transients

1. INTRODUCTION

1.1. Previous work on 2D radio imaging

The application of interferometry in radio astronomy opened a new window to observe, image and collect spatial information about extended radio sources. A radio interferometer gives a limited set of noisy Fourier samples (the visibilities) of the sky [1] inside the field of view of the instrument. The inverse Fourier Transform (FT) of those visibilities gives an approximate of the sky under a small field approximation. In addition, giant interferometers such as LOFAR [2], are no longer coplanar arrays but are defined in a 3D space and consequently, the measurements are subject to "direction-dependent" effects [3] such as the non-coplanarity impacting wide-field imaging (addressed with W-projection [4]) or variation of the beam impacting polarization (corrected with A-projection [5]). Moreover, due to the limited number of baselines, not all Fourier regions are sampled and one requires integrated measurements (in time or frequency) to increase the accuracy of the recovered sky. The problem of aperture synthesis comes down to process this incomplete Fourier map either by solving a deconvolution problem using the instrumental point spread function (PSF), such as CLEAN and its derivates (e.g. [6–8]) or by solving the inpainting problem by recovering missing information in the Fourier plane. In that scope, many teams have addressed this issue within the framework of the Compressed Sensing (CS) theory (e.g. [9] and references therein). In [9], we developed a 2D sparse radio imager working inside a LOFAR imager which corrects for A and W effects [5]. We demonstrated: i) better signal reconstruction and lower residuals compared to results obtained with classical methods, ii) high-dynamic & wide-field imaging capabilities by recovering accurate point-source flux densities ranging from 1 to 10^4 flux unit and iii) the capability to image super resolved features on real LOFAR data (containing extended sources) which were consistent with true source structures.

1.2. Transient radio imaging

Classical imagers work well with time and frequency integration by assuming a steady sky. However, another class of radio sources exists, the radio transients, the information of which are crucial for studies of "local" dynamic systems (e.g. planetary radio emissions, exoplanets) and (extra)galactic sources described by high-energy emission processes. Their detection and imaging are active fields of research in radio and a lot of work has been devoted to the development of detection pipelines (e.g. the LOFAR TRAnsient Pipeline – TRAP [10], based on fast iterative closed-loop performing calibration / imaging / source detection / catalogue cross-matching).

When the sky is steady (in the RA/DEC reference system), we rely on the time/frequency integration to improve the sampling of the visibility plane as well as the SNR. However, being variable and mostly point-like, they will suffer from the

[*] We acknowledge the financial support from the UnivEarthS Labex program of Sorbonne Paris Cité (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02) and the Physis project (H2020-LEIT-Space-COMPET)

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

imaging rate. On the one hand, when a set of successive snapshot images should theoretically enables temporal monitoring of a transient, each snapshot provides poor visibility coverage, therefore, images with low signal-to-noise ratio (SNR) due to large fraction of missing data. On the other hand, long time integration ensures a good sampling, but it will destroy the temporal information of the transient by mixing and diluting "ON" state periods with "OFF" state periods. As a result, the transient can be detected but with a large uncertainty on its time localization and temporal profile (or light curve).

Consequently, it is difficult to use classical imagers to detect and image transient source when the temporal variability of the transient source is unknown. Therefore, a new imager, still based on CS but handling the temporality of the data, is required.

2. IMAGE RECONSTRUCTION WITH SPARSITY CONSTRAINT

2.1. Compressed sensing theory

Compressed Sensing (CS) [11], or compressive sensing is a sampling and compression theory based on the sparsity of an observed signal. According to the theory, we could go beyond the Shannon limit to capture and represent compressible signals, which can be sparsely represented in a certain dictionary $\mathbf{\Phi}$, at a rate significantly below the Nyquist rate.

The CS theory is a paradigm for finding a probably exact reconstruction in the case of an undetermined problem such as the interferometry imaging problem, as we have fewer observations than unknowns. To achieve the perfect reconstruction from few samples, the CS theory relies on two principles: sparsity and incoherence.

- Sparsity: Generally, the CS theory exploits the fact the signal can be sparse or compressive, i.e. it has an economical representation in some dictionary $\mathbf{\Phi}$. For instance, a signal $\mathbf{x}(t)$ may be not sparse in its time domain, but in some space, for example, the wavelet space, $\mathbf{x}(t)$ can be decomposed as

$$\mathbf{x} = \mathbf{\Phi}\boldsymbol{\alpha} = \sum_{i=1}^{T} \alpha[i]\varphi_i, \tag{1}$$

  with T relatively small so that $\boldsymbol{\alpha}$ is a sparse representation of $\mathbf{x}(t)$ in $\mathbf{\Phi}$.
- Incoherence: This extends the duality between time and frequency in the sense that a sparse signal in $\mathbf{\Phi}$ must be spread out in the domain where it is acquired. It means that the sensing vectors must be as different as possible from the dictionary atoms.

2.2. Inverse problem formulation

The imaging problem constitutes an ill-posed inpainting problem which can be described in its simplified form as in Eq. (2) and represented in Fig. 1:

$$\mathbf{V} = \mathbf{MFx} + \mathbf{N} \tag{2}$$

with $\mathbf{V}$, the measured visibility vector, $\mathbf{M}$ the sampling mask which accounts for the lack of information in the Fourier space, $\mathbf{F}$ the FT operator, $\mathbf{x}$ the sky, and $\mathbf{N}$ the noise. The sky $\mathbf{x}$, expressed in the direct space, is a real quantity while the noise $\mathbf{N}$ is complex as it alters both amplitude and phase of the visibility measurements.

As the number of visibility measurements is limited, the image reconstruction problem can be regarded as a visibility inpainting problem, which is an inverse problem. From the CS perspective, to best reconstruct an image of the sky $\mathbf{x}$ from its visibilities is to use sparse recovery by solving the following minimization problem derived from Eq. (2):

$$\mathbf{min}\ ||\mathbf{\Phi}^t\mathbf{x}||_1 \quad s.t. \quad ||\mathbf{V} - \mathbf{MFx}||_2^2 < \epsilon, \tag{3}$$

where $\mathbf{V}$ is the measured visibility vector, $\mathbf{M}$ the mask matrix to represent the limited measured visibilities, $\mathbf{F}$ the Fourier transform operator. The objective function to minimize is of form $||\mathbf{\Phi}^t \cdot ||_1$ where the l_1-norm (the sum of coefficients absolute values) is well known to reinforce the sparsity of the solution.

In Lagrangian form, the convex minimization problem (3) can be formulated as:

$$\min_{\mathbf{x}} ||\mathbf{V} - \mathbf{MFx}||_2^2 + k||\boldsymbol{\lambda} \odot \mathbf{\Phi}^t\mathbf{x}||_1 \tag{4}$$

in the analysis framework, where $\boldsymbol{\lambda}$ is in vector form associated with the Lagrange multiplier which depends implicitly on ϵ of the Eq (3), and the operator $\odot$ denotes the element-by-element multiplication. However, the use of the l_1-norm, in the proximal framework, involves a soft thresholding step which has a well-known drawback of giving biased solutions [12]. This is particularly unsuitable for scientific data analysis, especially for photometry. Thus, the reweighted l_1 minimization proposed in [13] is one way to handle this issue. In addition, the wanted signal $\mathbf{x}$ is always positive, however, Eq (4) does not ensure the positivity of the solution. Therefore, the expression of these minimization problems is improved by imposing a positivity constraint and introducing a weighting vector $\mathbf{W}$, such as:

$$\min_{\mathbf{x}} ||\mathbf{V} - \mathbf{MFx}||_2^2 + k_1||\mathbf{W} \odot \boldsymbol{\lambda} \odot \mathbf{\Phi}^t\mathbf{x}||_1 + k_2 i_{\mathbb{C}^+}(\mathbf{x}) \tag{5}$$

for the analysis framework, where $i_{\mathbb{C}^+}$ denotes the indicator function in the positive set $\mathbb{C}^+$.

2.3. 2D-1D sparse representation

The data model for the 2D-1D imaging at a given frequency has two dimensions of spatial information and the remaining third dimension of the temporal information (corresponding

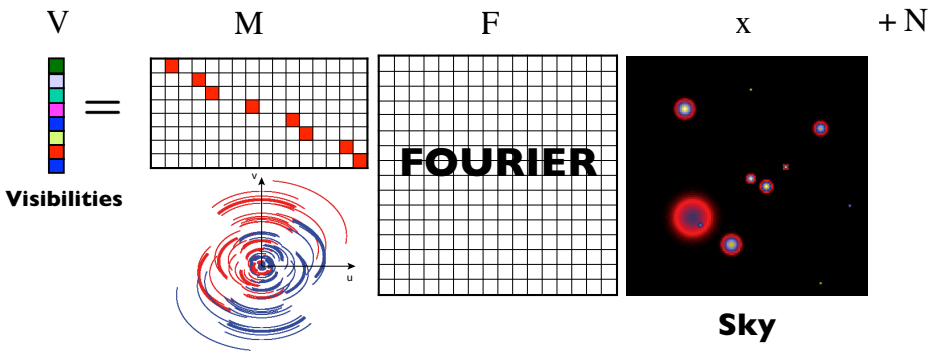

Fig. 1. Formulation of interferometric imaging as an inverse problem. $\mathbf{x}$ is the sky brightness and the signal to restore, $\mathbf{F}$ is the FT, $\mathbf{M}$ is the mask accounting for the available information (represented below in the Fourier plane), $\mathbf{N}$ is the noise which impact the visibilities measurements and $\mathbf{V}$ is the complex visibility vector measured by the interferometer.

to the sampling by the correlator). According to the CS theory, the corresponding 2D-1D sparse representation Φ in our study is vital to the final reconstruction. In 2D spatial sparse representation, it was shown that the starlet [12] or curvelet [14] dictionary were adapted to astronomical sources. Indeed, recent work of [9] has proved that the 2D reconstruction using starlet dictionary gives a better angular resolution and photometry resolution, especially for the extended sources, compared to the classical CLEAN methods. To extend to the 2D-1D sparse representation with the temporal dimension added, a direct inspiration would be to use a 3D starlet dictionary where one dimension will represent the time. However, such dictionary is not optimal as the temporal information is not correlated to the spatial information. Thus, we would like to separate 2D-spatial and 1D-temporal information. Therefore, as described in [15], an ideal wavelet function would be $\psi(x,y,t) = \psi^{(xy)}(x,y)\psi^{(t)}(t)$ where the space (xy) and time (t) are independent, and $\psi^{(xy)}$ is the spatial isotropic undecimated wavelet function (the starlet) and $\psi^{(t)}$ is a decimated wavelet function (the Haar or biorthogonal CDF 9/7 wavelets, depending on the form of the transient time profile).

2.4. Algorithms

In this section, we describe the algorithm for the reweighted l_1 minimization problem (5). We present the reweighted scheme before focusing on the complete algorithm.

In order to eliminate the false detected sources in our study, we define the weight function such as

$$w_{i,j} = f(|\alpha_{i,j}|) = \begin{cases} \frac{k\sigma_j}{|\alpha_{i,j}|} & \text{if } |\alpha_{i,j}| \geq \epsilon, \\ \frac{k\sigma_j}{\epsilon} & \text{else}, \end{cases} \tag{6}$$

to update the weights for each entity i at scale j, where ϵ is a small value constant to avoid the division by zero, σ_j is the noise standard deviation, accessible by reliable estimators such as the MAD (median of the absolute deviation), expected at scale j. k acts as a detection level in scale j. According to [13], the reweighted scheme is performed as:

1. Set the iteration count $n = 0$ and initialize $\mathbf{W}^{(0)} = \mathbf{1}$.
2. Solve the minimization problem (5) yielding a solution $\mathbf{x}^{(n)}$, and $\boldsymbol{\alpha}^{(n)}$ is obtained by $\boldsymbol{\alpha}^{(n)} = \boldsymbol{\Phi}^t \mathbf{x}^{(n)}$.
3. Update the weights by the weight function (6).
4. Terminate on convergence or when reaching the maximum number of iterations $N_{\max}$. Otherwise, increment n and go to step 2.

As for the minimization problem (5) where the ℓ_1-norm regularization term is not differentiable, a variety of proximal algorithms can be used, such as the Vu splitting method (VSM) [16], a kind of primal-dual algorithms. The main idea of the VSM is to convert the minimization problem (5) to a monotone inclusion for a bounded and existed primal-dual pair (x,u) which can be solved by a forward-backward algorithm. The summarized algorithm using reweighted scheme is presented in Algo 1, where the step τ and η are chosen under the convergence condition of $1 - \tau\eta||\Phi||^2 > \tau||\mathbf{MF}||^2/2$. More details can be found in [16]. The parameter μ in the Algo 1 is a relaxation parameter used to accelerate the algorithm. If $\mu = 1$, we are in the unrelaxed case.

3. EXPERIMENTS

We generated radio data observations by building ($32\times32\times64$) datacubes which respectively represent the two Fourier spatial frequencies and the time. We performed our signal reconstruction in the spatial domain (2D) and in the temporal domain (1D), to enable an unambiguous detection and localization of a transient source. The mask operator, which also depends on time, was generated by using a uniform random antenna distribution of 20 antennas observing at the zenith, for $\sim$2 hours (1 slice is therefore $\sim$2 min). Our time dependent sky model is constituted of a control steady source at the center of the field and a transient source with a gaussian

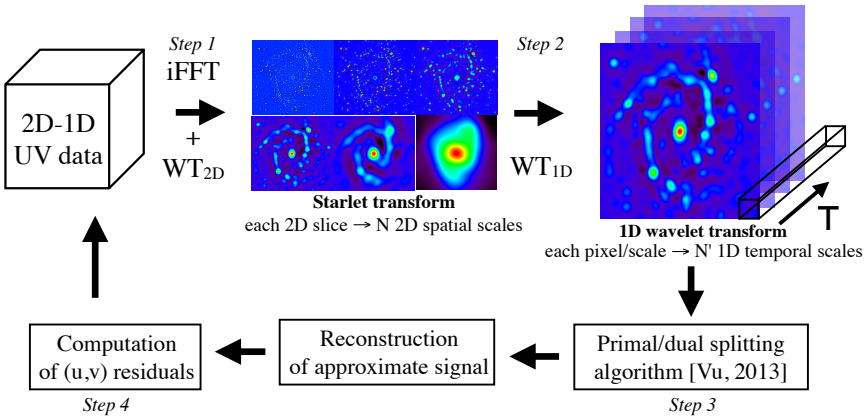

Fig. 2. Organization of the different steps of the reconstruction: i) Each 2D slice of the masked data cube undergoes inverse FT(giving a dirty image) and a 2D starlet transform giving N scales. ii) the time series of each pixel of each scales undergoes a 1D temporal wavelet transform. iii) wavelets coefficients are processed as in the Vu algorithm iv) the signal is transformed back and the residuals are computed with the data.

Algorithm 1: Analysis reconstruction using VSM

Data: Visibility $\mathbf{V}$; Mask $\mathbf{M}$
Result: Reconstructed image $\mathbf{x}$

Initialize
$(\mathbf{x}^{(0)}, \mathbf{u}^{(0)}), \mathbf{W}^{(0)} = \mathbf{1}, \tau > 0, \eta > 0, \mu \in]0, 1]$;
for $n = 0$ **to** $N_{max} - 1$ **do**
  $\mathbf{p}^{(n+1)} =$
  $\mathrm{Proj}_{\mathbb{C}^+}(\mathbf{x}^{(n)} - \tau\Phi\mathbf{u}^{(n)} + \tau(\mathbf{MF})^*(\mathbf{V} - \mathbf{MF}\mathbf{x}^{(n)}))$;
  $\mathbf{q}^{(n+1)} =$
  $(\mathbf{Id} - \mathrm{ST}_{\boldsymbol{\lambda}\odot\mathbf{W}})(\mathbf{u}^{(n)} + \eta\Phi^T(2\mathbf{p}^{(n+1)} - \mathbf{x}^{(n)}))$;
  $(\mathbf{x}^{(n+1)}, \mathbf{u}^{(n+1)}) =$
  $\mu(\mathbf{p}^{(n+1)}, \mathbf{q}^{(n+1)}) + (1 - \mu)(\mathbf{x}^n, \mathbf{u}^n)$;
  $\boldsymbol{\alpha}^{(n)} = \boldsymbol{\Phi}^t\mathbf{x}^{(n)}$;
  Update $\mathbf{W}$ by $w_{i,j}^{(n+1)} = f(|\alpha_{i,j}^{(n)}|)$;
end
return $\mathbf{x}^{(N_{\max})}$

light curve (FWHM = 10 min located at time slice T=24). Both sources have the same flux density of 10 arbitrary unit. We took the FT of the sky and applied the mask cube on the visibility cube after adding white gaussian noise with various magnitude ($\sigma = 0.0, 0.5, 1.0, 1.5$ flux unit). We show on Fig. 3 (left), characteristic dirty cube slices during the transient "OFF" state (first row) and its "ON" state (second row). When the noise level is high, the transient source is indiscernible from background artifacts.

Figure 3 (right) illustrates the profile of the transient source from the sky model cube (dash line), the dirty cube (red line) and the reconstructed cube obtained by the Vu algorithm (green) described in Sect. 2.4. With no additional noise (but with the sampling noise due to missing data), the CS reconstruction shows very small bias in flux density ($\sim 10^{-5}$ Flux unit relative error) as compared to the dirty profile. As the gaussian noise level increases, the flux density degrades in flux while keeping an approximate accurate light curve. For $\sigma = 1.5$, the reconstructed flux density has a similar bias as that of the dirty image but the shape of the light curve is still accurate in time, allowing the transient localization and further data processing around this date. The flux of the steady source (not shown) was also affected by the increasing level of noise.

4. DISCUSSION & CONCLUSION

CS offers a sound framework for solving the imaging problem in radio interferometry. In previous studies, such as [9], We have shown that simple algorithms and implementation can outperform classical tools used for deconvolution. In this work, we present an extension of our imager, solving the inpainting problem on data containing radio transients. It handles the third dimension being the temporal information in the data. We implemented the Vu algorithm and used a combination of 2D and 1D wavelets dictionaries to perform the reconstruction. We presented preliminary results based on simulated data cubes containing both steady and transient sources. The reconstructed transient light curve is accurate and is relatively robust to an increasing level on noise injected in the data. A full comparison of the imager performance with classical deconvolution methods is underway. A "time-agile" imager, which enables the detection of radio transients, may have a strong impact in radio astronomy for transient studies which will be addressed with next generation of interferometers, such as LOFAR and SKA.

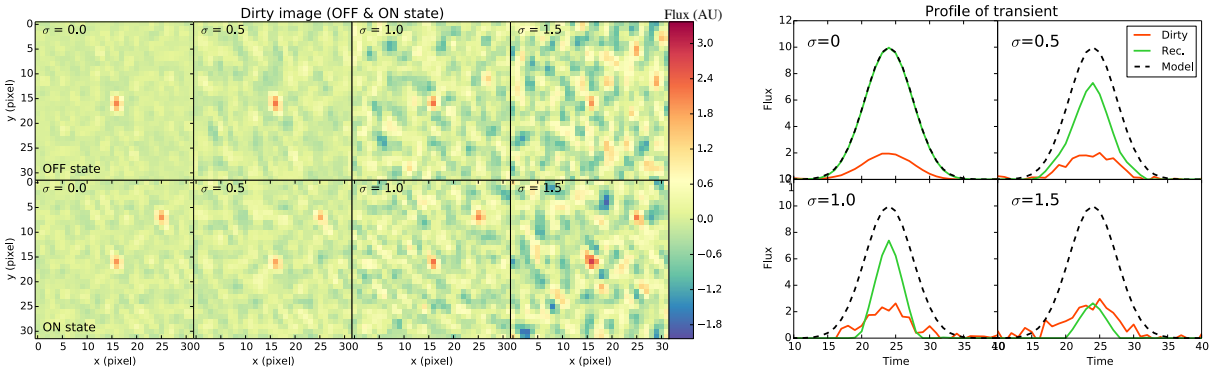

Fig. 3. (left) Dirty images from the benchmark data cube containing two point sources: a central steady source (x=15,y=15) and a transient source (x=24,y=6). First row corresponds to the OFF state (T=10) and second row to the ON state (T=25) at the maximum of the transient. Each column corresponds to various level of additive gaussian noise with $\sigma = 0, 0.5, 1.0, 1.5$ arbitrary flux units. (right) Time profiles at the spatial location of the transient source from the sky model (dash line), the dirty cube (red) and the reconstructed cube (green), for various levels of additive gaussian noise.

REFERENCES

[1] T. L. Wilson, K. Rohlfs, and S. Hüttemeister, *Tools of Radio Astronomy*, Springer-Verlag, 2009.

[2] M. P. van Haarlem, M. W. Wise, A. W. Gunst, and et al., "LOFAR: The LOw-Frequency ARray," *Astronomy and Astrophysics*, vol. 556, pp. A2, Aug. 2013.

[3] C. Tasse, G. van Diepen, and et al. van der Tol, "LOFAR calibration and wide-field imaging," *Comptes Rendus Physique*, vol. 13, pp. 28–32, Jan. 2012.

[4] T. J. Cornwell, K. Golap, and S. Bhatnagar, "The Noncoplanar Baselines Effect in Radio Interferometry: The W-Projection Algorithm," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 647–657, Nov. 2008.

[5] C. Tasse, S. van der Tol, J. van Zwieten, G. van Diepen, and S. Bhatnagar, "Applying full polarization A-Projection to very wide field of view instruments: An imager for LOFAR," *Astronomy and Astrophysics*, vol. 553, pp. A105, May 2013.

[6] J. A. Högbom, "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines," *Astron. Astrophys.*, vol. 15, pp. 417, June 1974.

[7] F. R. Schwab, "Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry," *Astronomical Journal*, vol. 89, pp. 1076–1081, July 1984.

[8] U. Rau and T. J. Cornwell, "A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry," *Astronomy and Astrophysics*, vol. 532, pp. A71, Aug. 2011.

[9] H. Garsden, J. N. Girard, J. L. Starck, S. Corbel, C. Tasse, A. Woiselle, and McKean et al., "LOFAR sparse image reconstruction," *Astron. Astrophys.*, vol. 575, pp. A90, Mar. 2015.

[10] J. et al. Swinbank, "The LOFAR Transients Pipeline," *Astronomy and Computing*, submitted.

[11] E.J. Candes and M.B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, March 2008.

[12] J.-L. Starck, F. Murtagh, and M.J. Fadili, *Sparse Image and Signal Processing*, Cambridge University Press, 2010.

[13] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing Sparsity by Reweighted L1 Minimization," *ArXiv:0711.1612*, Nov. 2007.

[14] J. L. Starck, D. L. Donoho, and E. J. Candès, "Astronomical image representation by the curvelet transform," *Astronomy and Astrophysics*, vol. 398, pp. 785–800, Feb. 2003.

[15] J.-L. Starck, J. M. Fadili, S. Digel, B. Zhang, and J. Chiang, "Source detection using a 3D sparse representation: application to the Fermi gamma-ray space telescope," *Astronomy and Astrophysics*, vol. 504, pp. 641–652, Sept. 2009.

[16] Bng Công Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," *Advances in Computational Mathematics*, vol. 38, no. 3, pp. 667–681, 2013.