Signal Separation with Magnitude Constraints : a Phase Problem

Antoine Deleforge* and Yann Traonmilin
INRIA Rennes - Bretagne Atlantique, Campus de Beaulieu,
FR-35042 Rennes Cedex, France
Email: antoine.deleforge@inria.fr

***Abstract*—We consider the problem of estimating the phases of K mixed complex signals from a multichannel observation, when the mixing matrix and signal magnitudes are known. We compare three approaches to tackle it: a heuristic method, an alternate minimization method, and a convex relaxation into a semi-definite program. In particular, we show that the convex relaxation approach yields best results, including the potential for exact source separation in under-determined settings.**

Let M sensors record K complex signals through linear instantaneous mixing. The noisy observation $\boldsymbol{y} \in \mathbb{C}^M$ is expressed as:

$$\boldsymbol{y} = \mathbf{A}\boldsymbol{s}_0 + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{s}_0 \in \mathbb{C}^K$ is the source vector, $\boldsymbol{n} \in \mathbb{C}^M$ is the noise vector and $\mathbf{A} \in \mathbb{C}^{M\times K}$ is the mixing matrix. This model is very common is signal processing and occurs, for instance, when looking at a single time-frequency bin of the discrete short-time Fourier domain.

In this work, we suppose that we have a prior knowledge on instantaneous source magnitudes which might be availabe in the case of informed source separation [9], [10]. Under model (1), we consider the problem of estimating the phases of $\boldsymbol{s}_0 \in \mathbb{C}^K$ given strictly positive magnitudes $\boldsymbol{b} = |\boldsymbol{s}_0|$, the multichannel observation $\boldsymbol{y} \in \mathbb{C}^M$ and the mixing matrix $\mathbf{A} \in \mathbb{C}^{M\times K}$, which we assume full rank. A natural approach is to minimize the Euclidean norm of the residual:

$$\begin{aligned} \hat{\boldsymbol{s}} &= \underset{\boldsymbol{s}}{\operatorname{argmin}} \|\mathbf{A}\boldsymbol{s} - \boldsymbol{y}\|_2^2 \\ \text{s.t. } |s_k|^2 &= b_k^2,\ k = 1\dots K. \end{aligned} \tag{ΦLS}$$

We refer to this problem as *phase least-squares* (ΦLS), because without constraints, it becomes a standard least-squares problem. Least squares has infinitely many solutions in the under-determined case ($K > M$) and a unique closed-form solution $\hat{\boldsymbol{s}}_{\text{LS}} = \mathbf{A}^\dagger \boldsymbol{y}$ otherwise, where $\{\cdot\}^\dagger$ denotes the Moore-Penrose pseudo-inverse. On the other hand, (ΦLS) is an instance of quadratically constrained quadratic program (QCQP). These problems are non-convex, and solving them or even finding whether they have a solution is NP-hard in general [1]. While branch-and-bound methods exist to solve non-convex QCQPs [1], [11], they are extremely slow in practice[1]. A generally exact and efficient solution to (ΦLS) is thus most likely out-of-reach, but we propose in the following three practical approaches to tackle it.

- The **multichannel Wiener filter** (MWF) is one of the most widely used methods in signal processing [3]. (*e.g.* [4]).
- A second approach to solve (ΦLS) is by **alternated minimization** w.r.t. each coordinate s_i, *i.e.*, by coordinate descent. Since (ΦLS) is not convex, it does not generally converge to a global minimum but to a local minimum which depends on the initial guess.
- The problem can be relaxed into the following **SDP program**

$$\begin{aligned} \widehat{\mathbf{X}} &= \underset{\mathbf{X}}{\operatorname{argmin}}\ \operatorname{trace}(\mathbf{CX}) \\ \text{s.t. } \mathbf{diag}\{\mathbf{X}\} &= \widetilde{\boldsymbol{b}},\ \mathbf{X} \succeq \mathbf{0} \end{aligned} \tag{PhUnLift}$$

  where $\mathbf{C} = [\mathbf{A}, -\boldsymbol{y}]^{\text{H}}[\mathbf{A}, -\boldsymbol{y}] \in \mathbb{C}^{(K+1)^2}$, $\widetilde{\boldsymbol{b}} = [\boldsymbol{b}^{\cdot 2}, 1]^\top$ and $\mathbf{X} \in \mathbb{C}^{(K+1)^2}$. PhUnLift stands for *phase unmixing by lifting*. Following [12], we use the particularly inexpensive block-coordinate descent (BCD) method of [13] to solve it.

Experiments and Results: We compare the efficiency of MWF, NMWF (Wiener filtering followed by a normalization to correct magnitudes), PhUnAlt and PhUnLift on the task of estimating the phases of $\boldsymbol{s}_0$ given an M-channel mixture $\boldsymbol{y} = \mathbf{A}\boldsymbol{s}_0 + \boldsymbol{n}$. Three initializations are considered for PhUnAlt: random phases (PhUnAlt), the output of NMWF (NMWF+) or the output of PhUnLift (PhUnLift+). Moreover, a *brute-force* approach (PhUnAlt*5) is considered, which picks the PhUnAlt estimate with smallest residual out of 5 randomly initialized runs. Fig. 1(a)(b)(c) shows the mean relative error as a function of M for different K (both determined and under-determined), under low-noise conditions (SNR = 60dB). PhUnLift and PhUnLift+ achieve the best results in all cases. This is further illustrated in Fig. 1(d) and Fig. 1(e), which show the probability of exact reconstruction of PhUnLift and PhUnAlt for different values of (M, K). 100% exact recovery seems possible with PhUnLift+ in a number of under-determined cases where PhUnAlt only achieves around 80%. Fig. 1(f) illustrates that the PhUnLift reconstructon error is proportional to the noise level when $K \leq M$. This is also true for MWF, NMWF and NMWF+, but not for PhUnAlt due to local-minima. In the under-determined setting showed in Fig. 1(g), stability to noise is less obvious. PhUnLift and PhUnLift+ perform best, closely followed by PhUnAlt*5. In general, the fair results obtained with PhUnAlt*5 suggests that the number of local-minima is often not too high, making multiple initialization of PhUnAlt a feasible approach.

We also conduct an informed speech separation task using random 1 second utterances from the TIMIT dataset [8]. The results in terms of signal-to-distortion-ratios (SDRs) calculated with [6] for each considered method are showed in table I (*Rand* means random phases with correct magnitudes). PhUnLift and PhUnLift+ outperform the other methods in under-determined settings, while MWF, NMWF, NMWF+, PhUnLift and PhUnLift+ performs similarly for $K = M$.

Related work and perspectives.: The considered problem of phase unmixing is related but not to be confused with the problem of *phase retrieval*, which has triggered considerable research interest over the past 30 years [5], [7] and has recently regained momentum thanks to novel methodologies [2], [12]. In the former, only the magnitudes of $\boldsymbol{y}$ are observed while $\boldsymbol{s}_0$ is completely unknown. However the similar SDP formulation opens a line of research for theoretical recovery guarantees on the phase unmixing problem.

[1] Solving one instance of (ΦLS) using the Matlab version of BARON [11] on a regular laptop takes over a minute with $M = 2$, $K = 3$.

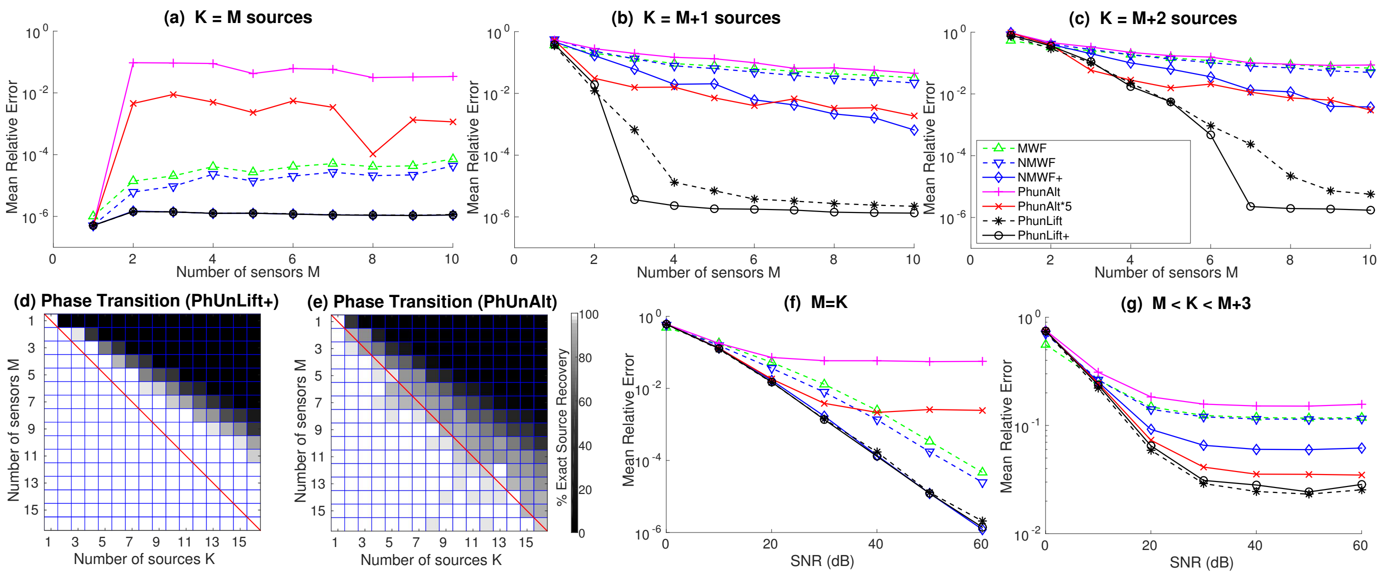

Fig. 1. (a)-(c) : Mean relative error for a fixed SNR of 60dB. (d)-(e) Probability of exact reconstruction with PhUnLift+ and PhUnAlt (noiseless). (f)-(g): Robustness to noise in determined cases and under-determined cases ($M = 2 \ldots 10$).

$M, K \rightarrow$	$2, 2$	$2, 3$	$2, 4$	$4, 4$	$4, 5$	$4, 6$
Input	0.49	-2.70	-4.66	-4.51	-5.76	-7.27
Rand	-6.44	-6.28	-4.99	-4.41	-4.46	-5.28
MWF	**59.6**	21.7	17.0	**58.6**	27.9	25.0
NMWF	**59.9**	21.6	16.9	**59.4**	28.6	25.9
NMWF+	**59.9**	22.8	19.2	**59.7**	34.1	31.6
PhUnAlt	22.7	17.6	15.3	25.8	21.5	22.0
PhUnAlt*5	43.5	35.4	**21.5**	47.7	37.3	33.6
PhUnLift	**59.9**	37.6	**22.3**	**59.0**	57.3	40.9
PhUnLift+	**59.9**	**39.6**	**21.2**	**58.0**	**59.3**	**44.8**

TABLE I
MEAN SDR (DB) FOR 1-SECOND M-CHANNEL MIXTURES OF K SPEECH SOURCES. MEANS ARE OVER THE K SOURCES FOR EACH MIXTURE.

REFERENCES

[1] C. Audet, P. Hansen, B. Jaumard, and G. Savard. A branch and cut algorithm for nonconvex quadratically constrained quadratic programming. *Mathematical Programming*, 87(1):131–152, 2000.

[2] E. J. Candes, T. Strohmer, and V. Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.

[3] I. Cohen, J. Benesty, and S. Gannot. *Speech processing in modern communication: challenges and perspectives*, volume 3. Springer Science & Business Media, 2009.

[4] N. Q. Duong, E. Vincent, and R. Gribonval. Under-determined reverberant audio source separation using a full-rank spatial covariance model. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(7):1830–1840, 2010.

[5] Y. C. Eldar, N. Hammen, and D. G. Mixon. recent advances in phase retrieval. *IEEE Signal ProcessIng Magazine*, page 158, 2016.

[6] C. Févotte, R. Gribonval, and E. Vincent. Bss_eval toolbox user guide–revision 2.0. 2005.

[7] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 21(15):2758–2769, 1982.

[8] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, and D. S. Pallett. The DARPA TIMIT acoustic-phonetic continuous speech corpus CD-ROM. Technical Report NISTIR 4930, National Institute of Standards and Technology, Gaithersburg, MD, 1993.

[9] A. Ozerov, A. Liutkus, R. Badeau, and G. Richard. Informed source separation: source coding meets source separation. In *Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 257–260. IEEE, 2011.

[10] C. Rohlfing, J. M. Becker, and M. Wien. Nmf-based informed source separation. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 474–478. IEEE, 2016.

[11] N. V. Sahinidis. Baron: A general purpose global optimization software package. *Journal of global optimization*, 8(2):201–205, 1996.

[12] I. Waldspurger, A. dAspremont, and S. Mallat. Phase recovery, maxcut and complex semidefinite programming. *Mathematical Programming*, 149(1-2):47–81, 2015.

[13] Z. Wen, D. Goldfarb, and W. Yin. Alternating direction augmented lagrangian methods for semidefinite programming. *Mathematical Programming Computation*, 2(3-4):203–230, 2010.