

Joint Blind Source Separation and Declipping: A Geometric Approach for Time Disjoint Sources

Alastair Turl
School of Computer Science
University of Birmingham
Birmingham, UK, B15 2TT
Email: a.c.turl@cs.bham.ac.uk

Ata Kabán
School of Computer Science
University of Birmingham
Birmingham, UK, B15 2TT
Email: a.kaban@cs.bham.ac.uk

Abstract—Source separation remains a challenging problem, made even more challenging when the mixtures are distorted. We present a novel framework for the source separation from mixtures affected by clipping distortion. Our method combines ℓ_1 minimisation with knowledge about the geometry of clipped mixtures when the sources are disjoint in time. Our algorithm recovers the sources by solving a convex optimisation problem, which is constrained by the clipping geometry. Comparative evaluation experiments show a significant increase in objective recovery performance of our proposed joint method compared to sequential approaches for speech and synthetic signals.

I. INTRODUCTION

Blind source separation (BSS) is the problem of separating a set of source signals from a set of mixture signals. For example, these sources may be distinct voices but we observe microphone recordings of many voices. It is an ill-posed problem, requiring assumptions to solve it. It becomes even more challenging when the observed mixtures have been affected by clipping distortion. Existing independent component analysis (ICA) techniques are not robust to clipping [1].

To make this problem solvable, at least approximately, we shall assume that the sources are disjoint in time, meaning a maximum of one source is non-zero at a given time point. Time disjoint sources appear as straight lines in a scattergram of linear mixtures, as previously noted, and this was exploited for separating more sources than mixtures in [2]. As we shall see, this particular representation is key to the novel geometric method described in this paper.

In our approach, there is no need to assume the sources are statistically independent as required in ICA. In addition we can handle the case of increasing the number of sources and whether or not both mixtures are clipped. For simplicity of exposition, however, this paper details the core, where there are two sources ($D = 2$) and only one of the two mixtures is clipped. Section II describes the methodology, prefaced by a general overview. Section III provides a comprehensive testing of our proposed approach on both synthetic and speech data.

II. METHODOLOGY

Throughout this work we assume noiseless linear mixtures. We observe F mixtures of N time points each, denoted by $\mathbf{x} \in \mathbb{R}^{F \times N}$. Each of these F observations is a linear mixture of

D sources, denoted by $\mathbf{s} \in \mathbb{R}^{D \times N}$, obtained by a mixing matrix $\mathbf{A} \in \mathbb{R}^{F \times D}$, as shown in the equation below.

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (1)$$

Clipping occurs in practice when attempting to represent a signal that contains amplitude values that saturate the maximum amplitude permitted by a system, yet most ICA algorithms neglect this. In a digital system, input samples whose absolute value exceeds a given *clipping threshold* θ have an output limited by this threshold. Formally, denote by $x_{i,n}$ the n th sample of the i th mixture before clipping and $x'_{i,n}$ the equivalent sample after clipping.

$$x'_{i,n} = \begin{cases} \text{sgn}(x_{i,n})\theta & \text{if } |x_{i,n}| > \theta, \\ x_{i,n} & \text{if } |x_{i,n}| \leq \theta. \end{cases} \quad (2)$$

Equations (1) and (2) combined describe the model of the problem:

$$x'_{i,n} = \begin{cases} \text{sgn}\left(\sum_{j=1}^D A_{i,j}s_{j,n}\right)\theta & \text{if } \left|\sum_{j=1}^D A_{i,j}s_{j,n}\right| > \theta, \\ \sum_{j=1}^D A_{i,j}s_{j,n} & \text{if } \left|\sum_{j=1}^D A_{i,j}s_{j,n}\right| \leq \theta. \end{cases} \quad (3)$$

Now the task is to find both \mathbf{s} and \mathbf{A} .

A. Overview of proposed approach

Linear mixtures of time disjoint sources, where at most one source is non-zero at a given time index, yield straight lines in a scattergram of such mixtures [2] (Figure 1). We refer to these lines as *source lines*. Figure 1 shows an example. The left column shows the sources before mixing and the right column shows the effect of clipping the first mixture. Note that clipping distorts the source lines.

Five sequential steps were developed in order to solve the problem of estimating latent sources $\hat{\mathbf{s}}$ from clipped mixtures \mathbf{x}' as follows:

- 1) Mixing matrix estimation (Section II-B).
 - The mixing matrix A is represented by the slopes of the source lines.
- 2) Quantisation of repairable time points (Section II-C).
 - There are some points which can be trivially repaired using knowledge of the clipping geometry.
- 3) Declipping mixtures by ℓ_1 minimisation (Section II-D).

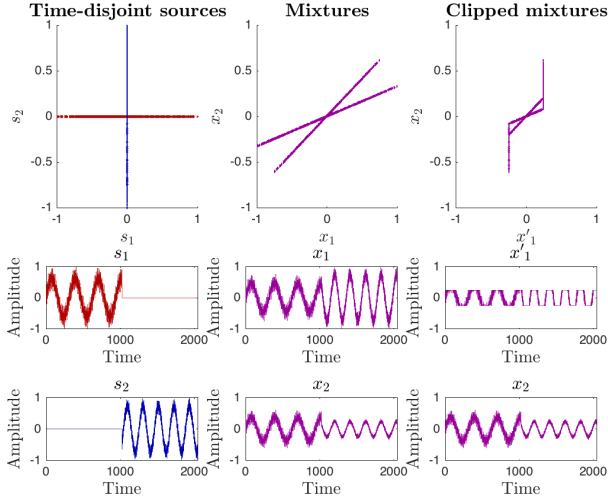


Fig. 1. Effects of clipping on one of two mixtures.

- Declipping by ℓ_1 minimisation allows information about the clipping process and mixture model to constrain the optimisation problem.
- 4) Quantisation of declipped points to the source lines (Section II-E).
 - The time points must lie on source lines because they are disjoint in time.
- 5) Source estimation (Section II-F).
 - The sources are recovered by the equation $\hat{\mathbf{s}} = \hat{\mathbf{A}}^{-1}\hat{\mathbf{x}}$.

B. Mixing matrix estimation

Since clipping distorts the source lines, first discard all clipped points, where $|x'_{1,n}| = \theta$. All remaining time points, $|x'_{1,n}| \leq \theta$, lie on a source line. Since each source line passes through the origin, each of the remaining time points $(x'_{1,n}, x'_{2,n})^T$ has enough information to recover one of the source line slopes, $\nabla_n = (x'_{2,n} - 0)/(x'_{1,n} - 0)$. We denote by M the number of unclipped time points, and calculate the slope of all of these as shown below. The symbol \odot represents the elementwise product.

$$\nabla_{\mathbf{M}} = (x'_{2,1}, \dots, x'_{2,M}) \odot \left(\frac{1}{x'_{1,1}}, \dots, \frac{1}{x'_{1,M}} \right) \quad (4)$$

The set $\nabla_{\mathbf{M}}$ contains M elements, some of which may be duplicates. We define the set $\nabla_{\mathbf{s}}$ to contain only the D unique entries of $\nabla_{\mathbf{M}}$. We make each value of the second row of the mixing matrix $\hat{A}_{2,1} \dots \hat{A}_{2,D} = 1$, which means the D elements in $\nabla_{\mathbf{s}}$ can be used directly as $\hat{A}_{1,1} \dots \hat{A}_{1,D}$, the first row of estimated mixing matrix $\hat{\mathbf{A}}$. To remove scaling indeterminacy, we normalise by dividing each element in column i of $\hat{\mathbf{A}}$ by its ℓ_2 -norm, $\hat{\mathbf{A}}_{1:2,i} = \hat{\mathbf{A}}_{1:2,i} \frac{1}{\|\hat{\mathbf{A}}_{1:2,i}\|_2}$.

C. Quantisation of repairable time points

Information from the unclipped mixture x'_2 can be used to repair some of the clipped points in x'_1 . We split the set of entries of $\nabla_{\mathbf{s}}$ into positives, denoted $\nabla_{\mathbf{s}}^+$, and negatives, denoted

$\nabla_{\mathbf{s}}^-$. Denote by ∇_+ the positives in ascending order and by ∇_- the negatives in descending order as follows:

$$\nabla_+ = \left[\frac{x'_{2,i}}{x'_{1,i}} \in \nabla_{\mathbf{s}}^+ : \frac{x'_{2,i}}{x'_{1,i}} \leq \frac{x'_{2,i+1}}{x'_{1,i+1}} \forall i = 1 \dots |\nabla_{\mathbf{s}}^+| \right] \quad (5)$$

$$\nabla_- = \left[\frac{x'_{2,i}}{x'_{1,i}} \in \nabla_{\mathbf{s}}^- : \frac{x'_{2,i}}{x'_{1,i}} \geq \frac{x'_{2,i+1}}{x'_{1,i+1}} \forall i = 1 \dots |\nabla_{\mathbf{s}}^-| \right] \quad (6)$$

We define $[\nabla_+]_i$ as the i th element of ∇_+ , the list of positive slopes in ascending order, and $[\nabla_-]_i$ as the i th element of ∇_- , the list of negative slopes in descending order. Denote by p_i the value of x'_2 when the source line with positive slope $[\nabla_+]_i$ intersects the clipping threshold ($x'_1 = \theta$). Denote by q_i the same for the source line with negative slope $[\nabla_-]_i$.

$$p_i = [\nabla_+]_i \theta \quad (7)$$

$$q_i = [\nabla_-]_i \theta \quad (8)$$

A time point with index n can be repaired if it satisfies two conditions:

- 1) It is clipped in x'_1 , the first mixture ($|x'_{1,n}| = \theta$).
- 2) In x'_2 , it lies between the intersection points p_1 and p_2 or q_1 and q_2 (the two least steep positive slopes and negative slopes, respectively).

If a time point satisfies both conditions, it must belong to the least steep source line. Figure 2, which shows the mixtures before clipping has occurred, shows why this is true. When we observe clipped data, we know that a point $(x'_{1,n}, x'_{2,n})^T$ on the dashed line between p_1 and p_2 is repairable. We also know that the original point $(x_{1,n}, x_{2,n})^T$ was located on a horizontal straight line, because only the first mixture coordinate $x_{1,n}$ is unknown ($x_{2,n} = x'_{2,n}$). The only source line that intersects the partially bounded region $|x_{1,n}| \geq \theta \wedge p_1 \leq |x_{2,n}| < p_2$ is the least steep, having slope $[\nabla_+]_1$. This region is shaded in Figure 2.

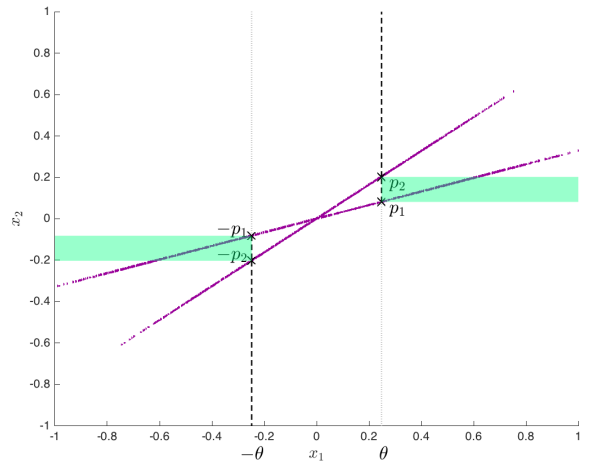


Fig. 2. Possible x_2 values for the first source at clipped points.

With knowledge of one coordinate $x'_{2,n}$, the slope $[\nabla_+]_1$ and that all source lines pass through the origin, we can calculate $x'_{1,n}$ geometrically for points in the region $|x_{1,n}| \geq \theta \wedge p_1 \leq$

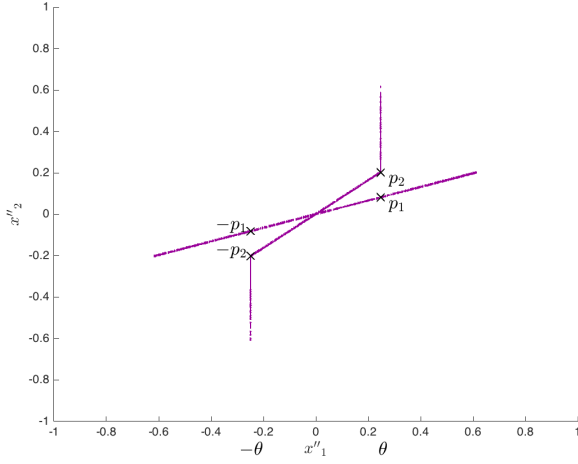


Fig. 3. Clipped mixtures after repairing repairable points.

$|x_{2,n}| < p_2$. Note that this also applies to points in the region $|x_{1,n}| \geq \theta \wedge q_1 \geq |x_{2,n}| > q_2$ for reasons of symmetry:

$$x''_{1,n} = \begin{cases} x'_{2,n}/[\nabla_+]_1 & \text{if } |x'_{1,n}| \geq \theta \wedge p_1 \leq |x'_{2,n}| < p_2, \\ x'_{2,n}/[\nabla_-]_1 & \text{if } |x'_{1,n}| \geq \theta \wedge q_1 \geq |x'_{2,n}| > q_2, \\ x'_{1,n} & \text{otherwise.} \end{cases} \quad (9)$$

$$x''_{2,n} = x'_{2,n} \quad (10)$$

Figure 3 shows the result of this repair for our example.

D. Declipping by ℓ_1 minimisation

At this stage, we attempt to reconstruct the clipped points that were not repairable in Section II-C. Our method uses a *sparse representation* of the source signals \mathbf{s} . Sparse representations have increasingly been used in research in the field of signal processing to solve problems such as declipping [3] and audio inpainting [4]. There are a number of benefits in using sparse representations, namely the ability to represent a signal with a small number of non-zero elements. The flexibility to choose a basis tailored to the type of signal is another advantage. We assume that source signals $\hat{\mathbf{s}}$ can be sufficiently represented by a small number of non-zero coefficients \mathbf{r} in the basis matrix Ψ like so:

$$\hat{\mathbf{s}} = \mathbf{r}^T \Psi^T \quad (11)$$

We choose the Discrete Cosine Transform (DCT) operator matrix, commonly used in signal processing, as Ψ for our example. It is likely that a sparser basis could be chosen, however the DCT will be sufficient to demonstrate our approach. We then formulate a convex optimisation problem that will give us a sparse solution with few non-zero elements. We use ℓ_1 minimisation because the number of non-zero elements, known as the ℓ_0 “norm”, is not convex. The task is then to find a solution for the linear system below, where sparse representations $\mathbf{r} \in \mathbb{R}^{N \times D}$:

$$\begin{aligned} & \text{minimize} \quad \sum_{i=1}^D \|\mathbf{r}_i\|_1 \\ & \text{subject to} \quad \hat{\mathbf{A}} \mathbf{r}^T \Psi^T = \mathbf{x}'' \end{aligned} \quad (12)$$

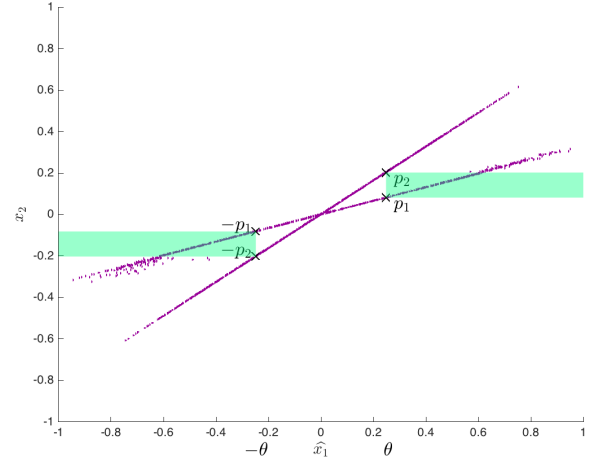


Fig. 4. Clipped mixtures after ℓ_1 minimisation repair.

The above optimisation problem is constrained by the mixture model $\mathbf{x}'' = \hat{\mathbf{A}} \hat{\mathbf{s}} = \hat{\mathbf{A}} \mathbf{r}^T \Psi^T$. We know, however, that some time points in x'_1 were unrepairable in Section II-C and should be ignored by this constraint. We instead use a separate constraint for x'_1 and unclipped mixture x'_2 . A masking matrix \mathbf{C}_u is constructed, consisting of the \mathbf{I}_N rows corresponding to indices of all unclipped and repairable points in x'_1 , which can be used to isolate these points. Similar masking matrices, \mathbf{C}_+ and \mathbf{C}_- , contain \mathbf{I}_N rows that correspond, respectively, to the positively clipped ($x'_{1,n} \geq \theta$) and negatively clipped ($x'_{1,n} \leq -\theta$) points in x'_1 . This pair of masking matrices is used to form constraints, similar to those developed in [3], that ensure the sparse sources \mathbf{r} obey the clipping model in Equation (2). The final form of the optimisation problem, where $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_2$ refer to the first and second rows of the matrix $\hat{\mathbf{A}}$ respectively, is as follows:

$$\begin{aligned} & \text{minimize} \quad \sum_{i=1}^D \|\mathbf{r}_i\|_1 \\ & \text{subject to} \quad \hat{\mathbf{A}}_1 \mathbf{r}^T \Psi^T \mathbf{C}_u^T = x'_1 \mathbf{C}_u^T, \\ & \quad \hat{\mathbf{A}}_2 \mathbf{r}^T \Psi^T = x'_2, \\ & \quad \hat{\mathbf{A}}_1 \mathbf{r}^T \Psi^T \mathbf{C}_-^T \leq -\theta, \\ & \quad \hat{\mathbf{A}}_1 \mathbf{r}^T \Psi^T \mathbf{C}_+^T \geq \theta. \end{aligned} \quad (13)$$

After solving the optimisation problem, the following equation yields an estimate of the declipped mixture \hat{x}_1 :

$$\hat{x}_1 = \hat{\mathbf{A}}_1 \mathbf{r}^T \Psi^T \quad (14)$$

To solve the problem in Equation (13) we used CVX, a package for specifying and solving convex problems [5], [6]. The results are shown in Figure 4.

E. Quantisation of declipped points to the source lines

The mixture points must lie on known source lines, which is not a guaranteed outcome of the optimisation problem. For this reason, we need to quantise the points to the source lines. Since the second mixture x_2 was not clipped, we should only adjust \hat{x}_1 . In the scattergram, this quantisation translates to horizontal adjustments of the repaired points. Each repaired point is

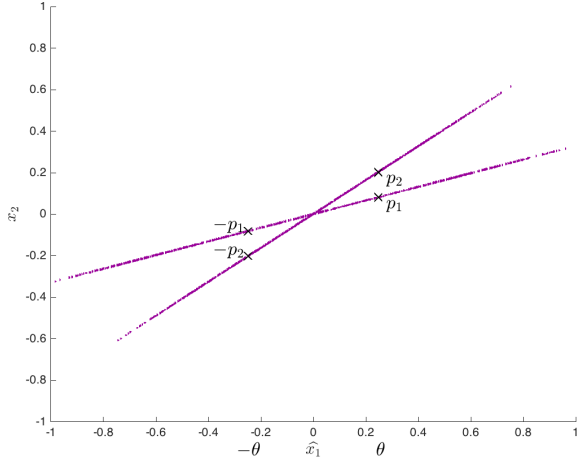


Fig. 5. Quantisation of mixtures repaired by ℓ_1 minimisation.

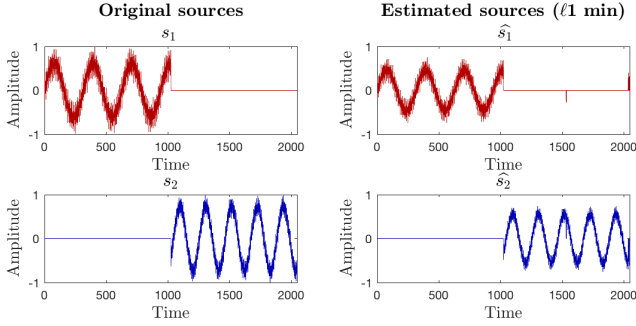


Fig. 6. Sources estimated from clipped mixtures by mixing matrix inverse $\hat{\mathbf{A}}^{-1}$.

quantised to the closest source line i using the following pair of equations:

$$l_n = \arg \min_i \left| \hat{x}_{1,n} - \frac{x_{2,n}}{\nabla_{s_i}} \right| \quad (15)$$

$$\hat{x}_{1,n} = \frac{x_{2,n}}{\nabla_{s_{l_n}}} \quad (16)$$

The results of quantising in this manner for our example are shown in Figure 5.

F. Source estimation

The final step is the estimation of the sources $\hat{\mathbf{s}}$. Since we have two mixtures, source recovery for two sources is trivial due the square mixing matrix. The estimated sources $\hat{\mathbf{s}}$ can be obtained, using the estimated mixtures $\hat{\mathbf{x}}$ and mixing matrix $\hat{\mathbf{A}}$, by the equation below:

$$\hat{\mathbf{s}} = \hat{\mathbf{A}}^{-1} \hat{\mathbf{x}} \quad (17)$$

Figure 6 shows the sources estimated for our example.

G. Sequential approach

In order to test our novel ideas of using knowledge of the mixture model to quantise repairable points (Section II-C) and inform the optimisation constraints (Section II-D), we omit the former step and modify the latter. The simplified approach is as follows:

- 1) Mixing matrix estimation (unchanged from Section II-B).
- 2) Declipping mixtures by ℓ_1 minimisation.
- 3) Quantisation of declipped points to the source lines (unchanged from Section II-E).
- 4) Source estimation (unchanged from Section II-F).

The optimisation problem to declip the clipped mixture x'_1 is shown below, where $w \in \mathbb{R}^N$ is a sparse representation of the first mixture:

$$\begin{aligned} & \text{minimize} \quad \|w\|_1 \\ & \text{subject to} \quad w^T \Psi^T \mathbf{C}_u^T = x'_1 \mathbf{C}_u^T, \\ & \quad \quad \quad w^T \Psi^T \mathbf{C}_-^T \leq -\theta, \\ & \quad \quad \quad w^T \Psi^T \mathbf{C}_+^T \geq \theta. \end{aligned} \quad (18)$$

H. Extensions

Our approach can be extended to handle an increasing number of sources ($D > 2$). For this case, the optimisation step (Section II-D) and the source estimation step (Section II-F) are altered. There exist some points for which we know, from the geometry, what the source label is not. This information allows new constraints for the optimisation problem. Source estimation cannot be performed using Equation (17) because the mixing matrix is not square. A different approach is employed, which involves rotating the mixtures to obtain the sources directly.

We can also handle the more challenging case where both mixtures are clipped. For this case, the quantisation of trivial points step (Section II-C) and the optimisation step (Section II-D) are altered. Points clipped in both mixtures cannot be trivially repaired. Points clipped in only one mixture are quantised using information from the unclipped mixture, similar to Section II-C. If only one source line has a positive slope, then we know the source label for any clipped point having a positive slope. The same is true for negative slopes. This knowledge informs additional constraints in the optimisation problem.

III. EXPERIMENTS

A series of experiments were designed to evaluate our proposed method and two others:

- 1) The joint source separation and declipping algorithm, as described in Section II.
- 2) The sequential approach, as described in Section II-G.
- 3) FastICA [7] - one of the most well known and popular source separation algorithms.

For FastICA, we first declip the mixtures \mathbf{x}' as in Equation (18) from Section II-G. The declipped mixtures are then processed by the FastICA code from [8].

Our hypotheses were that:

- 1) A larger proportion of clipped samples will lead to worse performance
 - More of the data will need to be approximated by the convex optimisation, increasing the chance of errors.

- 2) The joint algorithm will outperform the sequential approach in general
 - Performing the source separation and declipping tasks jointly will avoid the propagation of errors.
- 3) FastICA will perform worse than our approaches
 - Our approaches exploit knowledge of the clipping geometry and time disjointness assumption, which FastICA does not.
- 4) Speech sources will yield worse performance than sine wave sources
 - The DCT basis, which consists of sinusoids, will enable a sparser representation of sine wave sources than speech sources.
- 5) Gaussian sources will yield worse performance than sine wave sources
 - For the same reason given for hypothesis 4.

A. Preparing the data

Three types of source signal are used:

- 1) *Speech*
The audio files used are from [9] and available under the Creative Commons Attribution-NonCommercial-ShareAlike 2.0 (<http://creativecommons.org/licenses/by-nc-sa/2.0/>) license. The authors are Hendrik Kayser and Jorn Anemuller for the original data and Valentin Emiya for the modifications of each file. No changes were made.
 s_1 is the samples from 3900 + 1 to 3900 + N of 'targetSrc.wav'. s_2 is the samples from 3900 + 1 to 3900 + N of 'interfSrc1.wav'.
- 2) *Synthetic (sine wave mixtures)*
 $s_{1,n} = 3 \sin(.02n) + .8 \sin(3.5n) + .8 \sin(4.1n) + .5 \sin(4.5n)$
 $s_{2,n} = 3.4 \sin(.03n) + .5 \sin(2.2n) + .6 \sin(1.4n) + .2 \sin(3.7n)$
- 3) *Synthetic (Gaussian noise)*
 $s_{1,n} \sim \mathcal{N}(0, 0.5)$
 $s_{2,n} \sim \mathcal{N}(0, 0.5)$

Each source file is made disjoint in time in the following manner:

$$s_{1,n} = 0 \quad \forall n = \frac{N}{2} + 1 \dots N$$

$$s_{2,n} = 0 \quad \forall n = 1 \dots \frac{N}{2}$$

The time-disjoint sources are then normalised:

$$s_{1,n} = s_{1,n} / \max |s_1|$$

$$s_{2,n} = s_{2,n} / \max |s_2|$$

The mixing matrix is generated $\mathbf{A} \stackrel{iid}{\sim} \mathcal{U}(-1, 1)$. If \mathbf{A} is singular, it is discarded and generated again. The mixing matrix is scaled $\mathbf{A} = \mathbf{A} / \max |\mathbf{A}|$. The mixtures \mathbf{x} are generated as in Equation 1, clipping only mixture x_1 , and scaled $x_{1:2,n} = x_{1:2,n} / \max |\mathbf{x}|$. The mixtures \mathbf{x} are processed as described by Equation (2) to give \mathbf{x}' . Five clipping thresholds θ are selected, such that 10%, 20%, 30%, 40% and 50% of the samples of the sample are clipped each time.

B. Implementation and frame-based processing

The method proposed in Section II was implemented in MATLAB version 2016b. The length N of the mixtures and sources was fixed as $N = 2048$. Each experiment was repeated 50 times. For the optimisation problem in Section II-D, it is necessary to process the audio in short distinct frames sequentially and concatenate the results to form the declipped mixtures $\hat{\mathbf{x}}$, as in [3]. This is because of the short-time stationarity of audio signals, but has the added advantage of greatly decreasing processing time by making the basis matrix Ψ smaller. For our implementation, the mixtures \mathbf{x}'' are split into uniform-windowed non-overlapping frames of length $n = 256$.

C. Measuring performance

Recovery performance cannot be measured by evaluating the estimated mixing matrix \mathbf{A} as the estimated sources $\hat{\mathbf{s}}$ are also dependent on the declipping method used. For this reason, performance is measured by comparing the sources \mathbf{s} and their estimates $\hat{\mathbf{s}}$, using the following equation, cited in [10]:

$$\mathbf{D} := \min_{\epsilon = \pm 1} \left\| \frac{\hat{s}_i}{\|\hat{s}_i\|} - \epsilon \frac{s_i}{\|s_i\|} \right\|_2^2 \quad (19)$$

The final problem to solve is the permutation indeterminacy of the sources. We use the Hungarian algorithm, as implemented for MATLAB in [11]. Each source s_i is compared with each source estimate \hat{s}_j and the results are stored in a $D \times D$ matrix. The algorithm processes this matrix to determine the lowest performance cost and the corresponding estimated source to true source assignments.

D. Results

The following plots show D performance results for sine wave sources (Figure 7), Gaussian sources (Figure 8) and speech sources (Figure 9). Note that smaller is better and the error bar shows the standard error. The number above the error bar shows the number of times FastICA did not converge to a solution.

All of our results showed a trend of worsening performance as the proportion of clipped samples increased. This trend was not apparent in the FastICA approach with speech and sine wave sources. In these cases, however, FastICA performance was very poor at all clipping levels. Indeed, FastICA yielded the worst performance of all three methods in general, as predicted by our hypotheses. Our proposed joint methodology consistently yielded better performance with sine wave sources than the other source types, with very good performance (> 0.1) even when 50% of the data were clipped. This is likely in part due to our use of DCT as the sparse basis, which is known to be a good basis for sinusoid-based signals. Conversely, speech sources yielded worse performance than other source types. It is likely that this performance would be improved by repeating the experiments but choosing a basis better suited to speech signals.

In all of our experiments, our proposed joint approach yielded better performance than the sequential approach. It

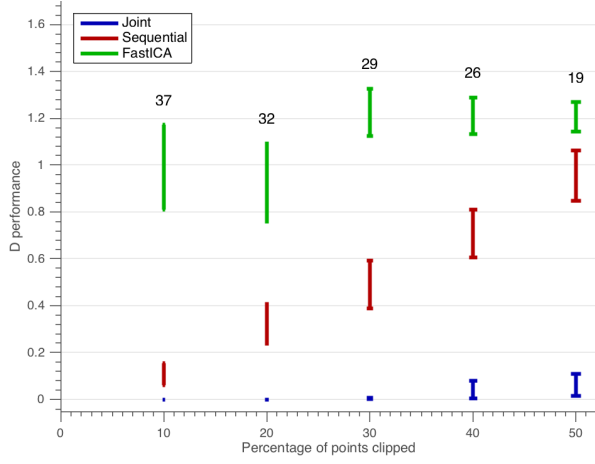


Fig. 7. D performance results for 2 sources, 1 clipped mixture, $n = 256$, $N = 2048$, 50 repetitions (Sine wave sources).

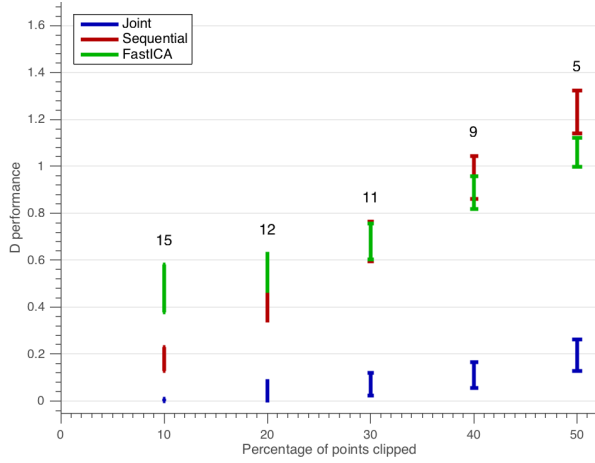


Fig. 8. D performance results for 2 sources, 1 clipped mixture, $n = 256$, $N = 2048$, 50 repetitions (Gaussian sources).

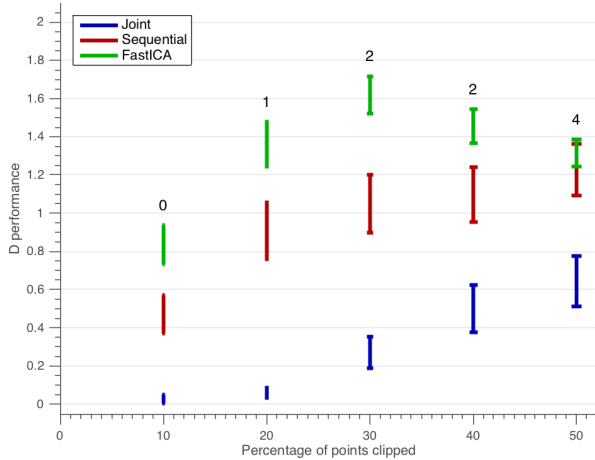


Fig. 9. D performance results for 2 sources, 1 clipped mixture, $n = 256$, $N = 2048$, 50 repetitions (Speech sources).

is likely that performance is improved by our novel idea to constrain the optimisation problem using knowledge of the mixing matrix and the clipped mixture geometry. Another factor is the quantisation of some points as described in Section II-C, which aids the optimisation problem by reducing the number of unknown values. The improved performance is less noticeable only at low clipping levels (10%) with sine wave sources (Figure 7). This is likely because it is the easiest case in our experiments: it has the smallest proportion of unknown values and the advantage that we used the DCT, in which sinusoids are very sparse, as our basis.

IV. CONCLUSION

This paper has presented a novel framework for the separation of sources from mixtures affected by clipping distortion, which combined ℓ_1 minimisation with knowledge about the geometry of clipped mixtures. An algorithm was developed to recover the sources by solving a convex optimisation problem, with constraints derived from the clipping geometry under the assumption that the sources are disjoint in time. Comparative evaluation experiments showed a significant increase in objective recovery performance of our proposed joint method compared to sequential approaches for speech and synthetic signals.

REFERENCES

- [1] A. C. Turl, "Robustness of ICA methods for clipping-distorted mixtures," MSc thesis, University of Birmingham, 2015.
- [2] M. D. Plumbley, T. Blumensath, L. Daudet, R. Gribonval, and M. E. Davies, "Sparse Representations in Audio and Music: From Coding to Source Separation," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 995–1005, 2010.
- [3] B. Defraene, N. Mansour, S. De Hertogh, T. van Waterschoot, M. Diehl, and M. Moonen, "Declipping of Audio Signals Using Perceptual Compressed Sensing," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 12, pp. 2627–2637, 2013.
- [4] V. Mach, "Inpainting of missing audio signal samples," Ph.D. dissertation, Brno, 2016.
- [5] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.
- [6] —, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110.
- [7] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4-5, pp. 411–430, 2000.
- [8] H. Gävert, J. Hurri, J. Särelä, and A. Hyvärinen, "FastICA for MATLAB." [Online]. Available: <http://research.ics.aalto.fi/ica/fastica/>
- [9] V. Emiya, E. Vincent, N. Harlander, and V. Hohmann, "Subjective and objective quality assessment of audio source separation," no. 19 (7), pp. 2046–2057, 2011.
- [10] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Transactions on Audio, Speech and Language Processing*, 2006.
- [11] Y. Cao, "Hungarian algorithm for linear assignment problems (v2.3)," 2011. [Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/20652-hungarian-algorithm-for-linear-assignment-problems-v2-3-?>