Phase recovery with PhaseCut and the wavelet transform case

Irène Waldspurger

Joint work with
Alexandre d’Aspremont and Stéphane Mallat
**Goal**: Solve the non-linear inverse problem

Reconstruct $f_0$ from $|Af_0|$

where $A$ is a linear injective operator, whose inverse is bounded and $|.|$ denotes the complex modulus

- Unicity ?
- Stability ?
- Algorithm ?
Two cases which have been studied:

- **$A$ is the Fourier transform.**
  - In dimension 1, no unicity. [Akutowicz, 1956]
  - In dimension $\geq 2$, unicity almost everywhere but no stability. [Barakat and Newsam, 1984]

- **$A$ is a gaussian random matrix with independent coefficients.**
  [Balan et al., 2006] [Candès et al., 2011b]
  - Unicity with high probability
  - Stability with high probability

Numerical results in the case of gaussian random filters are similar: $Af_0 = \{\hat{fg}_j\}_j$, for gaussian independant $g_j$. 
Is there a « middle case »?

- We want unicity and stability.
- We want A « structured ».

**Wavelet transform**

$\psi$ wavelet : $\int_{\mathbb{R}} \psi = 0$ \hspace{1cm} $\forall j \in \mathbb{Z}, \psi_j(x) = \frac{1}{2^j} \psi\left(\frac{x}{2^j}\right)$

\[ A : f \rightarrow \{ f \ast \psi_j \}_{j \in \mathbb{Z}} \]
Practical interest: Audio processing

Fig.: Spectrogram of an audio signal; musical structures are easily identifiable
Overview

1 - Reconstruction from the wavelet transform’s modulus
   - Unicity, for specific wavelets
   - Partial stability

2 - A general algorithm : PhaseCut
   - Principle and performances
   - Common points and differences with PhaseLift
The Fourier transform case

\[ \hat{f}(s) = \sum_k \hat{f}(k)e^{-\frac{2\pi is}{N}} = P_f(x_s) \]

where \( P_f(X) = \sum_k \hat{f}(k)X^k \) and \( x_s = e^{-\frac{2\pi is}{N}} \).

\[ |\hat{f}(s)|^2 = P_f(x_s)\overline{P_f} \left( \frac{1}{X_s} \right) \]

If \( \hat{f} \) is oversampled, \( |\hat{f}| \) determines \( Q(X) = P_f(X)\overline{P_f} \left( \frac{1}{X} \right) \).
Roots of $P_f(X)$:

\[ \{ z_k \} \]

Roots of $Q(X) = P_f(X) \overline{P_f} \left( \frac{1}{X} \right)$:

\[ \{ z_k \} \cup \left\{ \frac{1}{\overline{z}_k} \right\} \]

From $|\hat{f}|$, we can determine $P_f$’s roots only up to inversion with respect to the unit circle.

In the wavelet transform case, there are several circles.
Cauchy wavelets: \[ \hat{\psi}(\omega) = \omega^p e^{-\omega} 1_{\omega > 0} \]

**Fig.** Family of Cauchy wavelets, in the Fourier domain, for \( p = 5 \)
The wavelets are analytical: \( \hat{\psi}(\omega) = 0 \) if \( \omega < 0 \).

\[ \{ |f \star \psi_j| \}_j \] gives no information about \( \hat{f}(\omega) \) for \( \omega < 0 \). But we can reconstruct the positive frequencies.

**Theorem (Unicity)**

Let \( f, g \in L^2(\mathbb{R}) \) such that \( \forall j \in \mathbb{Z}, |f \star \psi_j| = |g \star \psi_j| \).

There exists \( \lambda \in \mathbb{C} \) such that \( |\lambda| = 1 \) and

\[ \hat{f}1_{\omega>0} = \lambda \hat{g}1_{\omega>0} \]

**Consequence:** If \( f \) is real and if we have an additional information about \( \lambda \), \( \{ |f \star \psi_j| \}_j \) determines \( f \).
Theorem (Unicity, discrete version)

Let \( f, g \in \mathbb{C}^n \) such that \( \forall j \in \mathbb{Z}, |f \star \psi_j| = |g \star \psi_j| \), where \( f \star \psi_j \) and \( g \star \psi_j \) are oversampled with a factor of 2.

There exists \( \lambda \in \mathbb{C} \) and \( C \in \mathbb{C} \) such that \( |\lambda| = 1 \) and

\[
f = \lambda g + C
\]

The constants \( \lambda \) and \( C \) may not be removed:

\[
C \star \psi_j = 0 \Rightarrow |(\lambda g + C) \star \psi_j| = |g \star \psi_j|
\]
Proof:

\[ f \star \psi_j(s) = P_f(R_j x_s) \]

where

\[ P_f(X) = \sum_{k=0}^{N-1} \hat{f}(k) k^p X^k \]

\[ R_j = \exp(-\frac{2j}{N}) \]

\[ x_s = e^{\frac{2\pi is}{N}} \]

\( f \star \psi_j \) contains the values of \( P_f \) along the complex circle of radius \( R_j \).

For every \( j \), \( |f \star \psi_j| \) determines \( P_f \)'s roots up to inversion with respect to the circle of radius \( R_j \).
\[ \forall j \in \mathbb{Z}, \ |f \star \psi_j| = |g \star \psi_j| \]

\[ \Rightarrow P_f \text{ and } P_g \text{ have the same roots?} \]

By dividing by \( \gcd(P_f, P_g) \), we can assume that \( P_f \) and \( P_g \) have no common root.

Let us show that \( P_f \) and \( P_g \) are constants. By contradiction, we suppose that \( P_f \) has a root, \( z \).

We choose \( j_1 \neq j_2 \in \mathbb{Z} \).
\[ z \text{ (root of } P_{\ell}) \]
Wavelet transform - Unicity

\[ z \text{ (root of } P_f \text{)} \]

\[ R_{j_1}^2/z \text{ (root of } P_g \text{)} \]
\[ c = \frac{R_{j_2}^2}{R_{j_1}^2} \]

\[ z \text{ (root of } P_f) \]
\[ c = \frac{R_{j_2}^2}{R_{j_1}^2} \]

\( z \) (root of \( P_f \))

\( c^2 z \)
\[ c = \frac{R_{j_2}^2}{R_{j_1}^2} \]

\[ c^2 z \] (root of \( P_f \))

\[ c^3 z \]
Theorem (Continuity)

If \( \lim_{n \to \infty} \{ |f_n \ast \psi_j| \} = \{ h_j \} \), then

\[ \exists g \in L^2(\mathbb{R}) \text{ s.t. } h_j = |g \ast \psi_j| \]

and

\[ \lim_{n \to \infty} \lambda_n \hat{f}_n 1_{\omega > 0} = \hat{g} 1_{\omega > 0} \]
The function \( \{ |f \ast \psi_j| \}_j \rightarrow \hat{f}1_{\omega > 0} \) is continuous.

But it is not uniformly continuous:

\[
\forall \epsilon > 0, \exists f, g \text{ such that } \left\| \{ |f \ast \psi_j| \}_j - \{ |g \ast \psi_j| \}_j \right\|_2 < \epsilon \\
\left\| \hat{f}1_{\omega > 0} - \lambda \hat{g}1_{\omega > 0} \right\|_2 \geq 1, (\forall |\lambda| = 1)
\]
**Qualitative explanation**: Redundancy is not well-spread. High and low frequencies do not «communicate».
**Qualitative explanation**: Redundancy is not well-spread.

High and low frequencies do not « communicate ». 
Qualitative explanation: Redundancy is not well-spread.

High and low frequencies do not « communicate ».
Qualitative explanation : Redundancy is not well-spread. High and low frequencies do not « communicate ».
Qualitative explanation: Redundancy is not well-spread.
High and low frequencies do not « communicate ».
There are similar instabilities in the time domain.

These instabilities correspond to signals that the human ear cannot differentiate.

**Example**: $f, g$ two audio signals

\[
\frac{\|f - g\|_2}{\|f\|_2} > 100\%
\]

\[
\frac{\|\{|f \ast \psi_j\}_j - \{|g \ast \psi_j\}_j\|_2}{\|\{|f \ast \psi_j\}_j\|_2} \approx 4.5\%
\]
Which algorithm?

Problem: reconstruct $f_0$ from $|Af_0|$ ($f_0 \in \mathbb{C}^n$, $A \in \mathbb{C}^{m \times n}$)

Classical algorithms: gradient descent, alternate projections ([Gerchberg and Saxton, 1972])
The problem is not convex. They get stuck into local minima.

Solution: convexification techniques ([Chai, Moscoso, and Papanicolaou, 2011],[Candès, Eldar, Strohmer, and Voroninski, 2011a])
Outline of the section

1 - Review of *PhaseLift*

2 - Principle of *PhaseCut*

3 - Theoretical equivalence between *Weak PhaseLift* and *PhaseCut*, without noise

4 - Computational differences

5 - Noisy case
**PhaseLift**

Instead of reconstructing \( f_0 \), one reconstructs \( X = f_0 f_0^* \).

Find \( X \) s.t.
- Rank \( X = 1 \)
- \( X \succeq 0 \)
- \( \text{diag} \left( AXA^* \right) = \text{diag} \left( (Af_0)(Af_0)^* \right) = |Af_0|^2 \)

Minimize \( \text{Tr} \ X \) s.t.
- \( X \succeq 0 \)
- \( \text{diag} \left( AXA^* \right) = |Af_0|^2 \) (PhaseLift)
Properties of PhaseLift:
[Candès, Strohmer, and Voroninski, 2011b]
[Candès and Li, 2012] [Demanet and Hand, 2012]

If $A$ is a gaussian random matrix and $m \geq Cn$, then, with high probability,
- For every $f_0$, the solution $X$ of PhaseLift is unique.
- PhaseLift is stable to noise.

Same properties hold without minimizing the trace.

Find $X$
s.t. $X \succeq 0$ \hspace{1cm} (Weak PhaseLift)
$\text{diag}(AXA^*) = |Af_0|^2$
**PhaseCut** : We reconstruct the phases.

We want to reconstruct $Y = (Af_0)(Af_0)^*$. 

Find $Y$ 

s.t. Rank $Y = 1$

$Y \succeq 0$

$\text{diag}(Y) = \text{diag}((Af_0)(Af_0)^*) = |Af_0|^2$

Range $Y \subset \text{Range } A$

Range $Y \subset \text{Range } A \iff (\text{Id} - AA^\dagger)Y = 0$

$\iff \text{Tr}((\text{Id} - AA^\dagger)Y) = 0$
Find $Y$

s.t.  

$\text{Rank } Y = 1$

$Y \succeq 0$

$\text{diag}(Y) = |Af_0|^2$

$\text{Tr}((\text{Id} - AA^\dagger)Y) = 0$

This is PhaseCut. [d’Aspremont, Mallat, and Waldspurger, 2012]
Find \( Y \)

s.t. \( \text{Rank} \ Y = 1 \)

\( Y \succeq 0 \)

\( \text{diag} \ (Y) = |Af_0|^2 \)

\( \min \ \text{Tr} \ ((\text{Id} - AA^\dagger)Y) = 0 \)

This is PhaseCut. [d'Aspremont, Mallat, and Waldspurger, 2012]
Find $Y$
\[\text{s.t.} \quad \text{Rank } Y = 1\]
$Y \succeq 0$
$\text{diag}(Y) = |A f_0|^2$
\[\min \quad \text{Tr} ((I - AA^\dagger) Y) = 0\]

We approximate this problem by:

\[\text{Minimize} \quad \text{Tr} ((I - AA^\dagger) Y)\]
\[\text{s.t.} \quad Y \succeq 0\]
$\text{diag}(Y) = |A f_0|^2$

This is **PhaseCut**. [d’Aspremont, Mallat, and Waldspurger, 2012]
PhaseCut is an instance of the max-cut problem.

\[
\begin{align*}
\text{Minimize} & \quad \text{Tr} (WU) \\
\text{s.t.} & \quad U \succeq 0 \\
& \quad \text{diag} (U) = 1
\end{align*}
\]

(Maximum Cut)

\[
\begin{align*}
\min \quad & \text{Tr} ((\text{Id} - AA^\dagger) Y) \\
\text{s.t.} & \quad Y \succeq 0 \\
& \quad \text{diag} (Y) = |A f_0|^2
\end{align*}
\]

where \( B = \text{diag} (|A f_0|) \)

\[
\begin{align*}
\min \quad & \text{Tr} ((\text{Id} - AA^\dagger) B V B) \\
\text{s.t.} & \quad V \succeq 0 \\
& \quad \text{diag} (V) = 1
\end{align*}
\]
Theorem

If $A$ is invertible, $X$ is a solution of Weak PhaseLift if and only if $AXA^*$ is a solution of PhaseCut.

There is also an equivalence between PhaseLift and a modified version of PhaseCut.
PhaseCut - Computational differences

$n$: size of $f_0$  
$nJ$: number of measurements  
$\epsilon$: error over the objective value

Matrices in PhaseCut are larger ($nJ \times nJ$ instead of $n \times n$) but the max-cut formulation allow simplifications.

- Block coordinate method [Wen et al., 2009]  
  Complexity bound unknown; low cost per iteration  
  For PhaseCut only; to test!

- First-order solvers
  
  $\text{PhaseLift}: O\left(\frac{Jn^3}{\epsilon}\right)$  
  $\text{PhaseCut}: O\left(\frac{J^3n^3}{\epsilon}\right)$

- Interior-point solvers
  
  $\text{PhaseLift}: O\left(J^2n^{4.5}\log\frac{1}{\epsilon}\right)$  
  $\text{PhaseCut}: O\left(J^{3.5}n^{3.5}\log\frac{1}{\epsilon}\right)$
The optimal algorithm depends upon the necessary precision and thus upon $A$ and $f_0$.

**Examples for 64-sized signals:**

- 6 gaussian random filters (stable) :
  
  $\epsilon \sim 0.2\eta$

- Wavelet transform, real signal, 4 wavelets (less stable) :
  
  $\epsilon \sim 10^{-3}\eta$ or $10^{-2}\eta$

$\eta$ : relative $L^2$-norm error $\frac{||f_0-f_{\text{rec}}||_2}{||f_0||_2}$. 
Performances

- Three classes of 128-sized signals

\[\text{(Gaussian white noise) (Sum of sinusoids) (Scan-line of an image)}\]

**Fig.:** Real parts of sample test signals

- Three matrices \(A\):
  - Oversampled Fourier transform
  - 4 gaussian random filters
  - Wavelet transform, 4 Cauchy wavelets
<table>
<thead>
<tr>
<th></th>
<th>Fourier</th>
<th>Random Filters</th>
<th>Wavelets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gerchberg-Saxton</td>
<td>5</td>
<td>49</td>
<td>0</td>
</tr>
<tr>
<td><em>PhaseLift</em> with reweighting</td>
<td>3</td>
<td>100</td>
<td>62</td>
</tr>
<tr>
<td><em>PhaseCut</em></td>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

**Tab.**: Percentage of perfect reconstruction over 300 test signals; each algorithm is followed by some steps of a classical Gerchberg-Saxton algorithm.

For *PhaseLift*, we performed 10 steps of reweighting with 1000 iterations per step. We used TFOCS ([Becker et al., 2012]).

In the wavelet transform case, convergence is not reached for *PhaseLift*. 
The measurements are noisy:

\[ b = |Af_0| + b_{\text{noise}} \]

- If the reconstruction problem is stable to noise, we want to reconstruct \( f_{\text{rec}} \approx f_0 \).
- If the reconstruction problem is not stable to noise, we want to reconstruct \( f_{\text{rec}} \) such that \( |Af_{\text{rec}}| \approx |Af_0| \).
There are theoretical results for the case of a stable reconstruction problem.

One can define a notion of **C-stability**.

To be $C$-stable $\Rightarrow \|f_{\text{rec}} - f_0\|_2 \leq C\|b_{\text{noise}}\|_2$

**Theorem**

*If Weak PhaseLift is $C$-stable in $f_0$, then PhaseCut is $(2C + 2\sqrt{2} + 1)$-stable in $f_0$."

When $|Af_0|$ is **sparse**, numerical results indicate that the converse is not true and that *PhaseCut* is more stable. In practice, sparsity occurs very often.
Fig. 5: Reconstruction errors versus amount of noise for \textit{PhaseLift} and \textit{PhaseCut}, both in decimal logarithmic scale, for three types of signals: Gaussian white noise, sinusoids and scan-line of an image.
- *PhaseCut* has good performances and is stable.
- It works only for small signals.
- For larger signals:
  - Multi-scale version?
  - Block coordinate descent?


E. J. Candès and X. Li. Solving quadratic equations via phaselift when there are about as many equations as unknowns. Preprint, 2012.


