Phase retrieval with Bregman divergences: application to audio signal recovery

Pierre-Hugo Vial, Paul Magron, Thomas Oberlin and Cédric Févotte

CNRS, IRIT, Université de Toulouse

iTWIST - December 2, 2020

Phase retrieval

Problem statement

Recover a signal $\mathbf{x}^\star \in \mathbb{C}^L$ from nonnegative measurements $\mathbf{r} \in \mathbb{R}^K$ such that

$$\mathbf{r} \approx |\mathbf{A}\mathbf{x}^\star|^d$$

- $\mathbf{A} \in \mathbb{C}^{K \times L}$: measurement operator.
- $d = 1$ (magnitude) or 2 (power).

Common approach: nonconvex optimization

$$\min_{\mathbf{x} \in \mathbb{C}^L} \quad E(\mathbf{x}) := \| \mathbf{r} - |\mathbf{A}\mathbf{x}|^d \|_2^2$$

- Algorithms: gradient descent, alternating projections, majorization-minimization, ADMM...
- Recovery up to ambiguities.

Phase retrieval

Problem statement

Recover a signal $\mathbf{x}^\star \in \mathbb{C}^L$ from nonnegative measurements $\mathbf{r} \in \mathbb{R}^K$ such that

$$\mathbf{r} \approx |\mathbf{A}\mathbf{x}^\star|^d$$

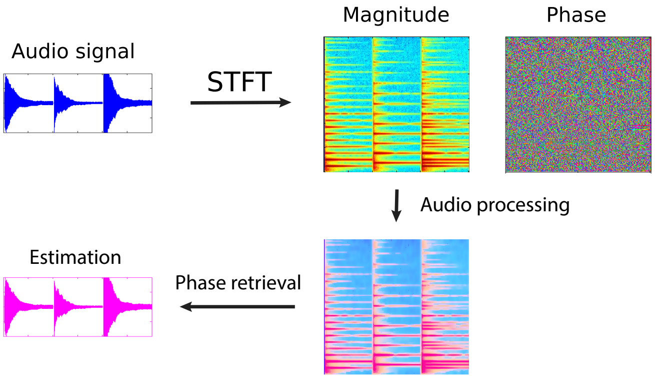
- $\mathbf{A} \in \mathbb{C}^{K \times L}$: measurement operator.
- $d = 1$ (magnitude) or 2 (power).

Common approach: nonconvex optimization

$$\min_{\mathbf{x} \in \mathbb{C}^L} \quad E(\mathbf{x}) := \| \mathbf{r} - |\mathbf{A}\mathbf{x}|^d \|_2^2$$

- Algorithms: gradient descent, alternating projections, majorization-minimization, ADMM...
- Recovery up to ambiguities.

PR for audio signal recovery

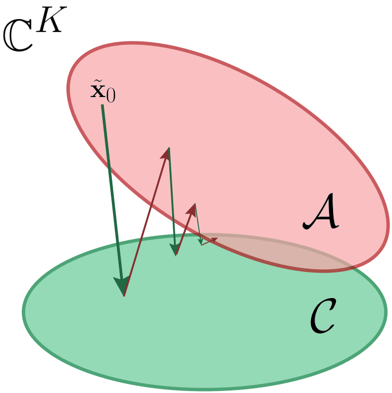


Short-time Fourier transform (STFT)

- $\mathbf{A}$: STFT operator.
- $\mathbf{A}^{\mathsf{H}}$: inverse STFT under duality conditions.

A classic algorithm: Griffin-Lim algorithm (GLA)

[Griffin and Lim, 1984]



Alternating projections

- Magnitude constraint:
  $\mathcal{A} = \left\{\tilde{\mathbf{x}} \in \mathbb{C}^K \mid |\tilde{\mathbf{x}}| = \mathbf{r}\right\}$
  - $\mathcal{P}_{\mathcal{A}}(\tilde{\mathbf{x}}) = \mathbf{r} \odot \frac{\tilde{\mathbf{x}}}{|\tilde{\mathbf{x}}|}$
- Consistency constraint:
  $\mathcal{C} = \mathrm{Im}(\mathbf{A})$
  - $\mathcal{P}_{\mathcal{C}}(\tilde{\mathbf{x}}) = \mathbf{A}\mathbf{A}^H\tilde{\mathbf{x}}$

Griffin-Lim Algorithm ($d = 1$)

- **Initialize:** ϕ_0, $\tilde{\mathbf{x}}_0 = \mathbf{r} \odot \phi_0$
- **Iterate:** $\tilde{\mathbf{x}}_{t+1} = \mathcal{P}_{\mathcal{C}}(\mathcal{P}_{\mathcal{A}}(\tilde{\mathbf{x}}_t))$

Converges to a critical point of E.

PR with Bregman divergences

Optimization problem

$$\min_{\mathbf{x}\in\mathbb{C}^L} J(\mathbf{x}) := \underbrace{\mathcal{D}_\psi(\mathbf{r}\,|\,|\mathbf{A}\mathbf{x}|^d)}_{\text{"right"}} \quad \text{or} \quad \underbrace{\mathcal{D}_\psi(|\mathbf{A}\mathbf{x}|^d\,|\,\mathbf{r})}_{\text{"left"}}$$

Bregman divergences

With ψ strictly-convex and continuously-differentiable scalar function,

$$\mathcal{D}_\psi(\mathbf{y}\,|\,\mathbf{z}) := \sum_k \left[\psi(y_k) - \psi(z_k) - \psi'(z_k)(y_k - z_k)\right]$$

Motivations

- Unifying framework.
- Encompasses Quadratic, Kullback-Leibler, Itakura-Saito and β-divergences.
- Good performance in audio, e.g. in NMF.

Gradient descent and acceleration

Gradient expression

$$\nabla J(\mathbf{x}) = \frac{d}{2}\,\mathbf{A}^{\mathsf{H}}\left[|\mathbf{A}\mathbf{x}|^{d-2} \odot (\mathbf{A}\mathbf{x}) \odot \mathbf{z}\right]$$

$$\mathbf{z} = \underbrace{\psi''(|\mathbf{A}\mathbf{x}|^{d}) \odot (|\mathbf{A}\mathbf{x}|^{d} - \mathbf{r})}_{\text{“right”}} \quad \text{or} \quad \underbrace{\psi'(|\mathbf{A}\mathbf{x}|^{d}) - \psi'(\mathbf{r})}_{\text{“left”}}$$

Accelerated gradient descent

Iterate:

$$\mathbf{y}_{t+1} = \mathbf{x}_t - \mu \nabla J(\mathbf{x}_t)$$
$$\mathbf{x}_{t+1} = \mathbf{y}_{t+1} + \gamma(\mathbf{y}_{t+1} - \mathbf{y}_t)$$

- μ: step-size.
- γ: acceleration parameter.

Special cases

- GLA: $d = 1$, $\mu = 1$, $\gamma = 0$ and quadratic loss.
- Wirtinger Flow [Candès et al., 2015]: $d = 2$, $\gamma = 0$ and quadratic loss.

Gradient descent and acceleration

Gradient expression

$$\nabla J(\mathbf{x}) = \frac{d}{2}\,\mathbf{A}^{\mathsf{H}}\left[|\mathbf{Ax}|^{d-2}\odot(\mathbf{Ax})\odot\mathbf{z}\right]$$

$$\mathbf{z} = \underbrace{\psi''(|\mathbf{Ax}|^{d})\odot(|\mathbf{Ax}|^{d}-\mathbf{r})}_{\text{“right”}} \quad \text{or} \quad \underbrace{\psi'(|\mathbf{Ax}|^{d})-\psi'(\mathbf{r})}_{\text{“left”}}$$

Accelerated gradient descent

Iterate:

$$\mathbf{y}_{t+1} = \mathbf{x}_t - \mu\nabla J(\mathbf{x}_t)$$
$$\mathbf{x}_{t+1} = \mathbf{y}_{t+1} + \gamma(\mathbf{y}_{t+1} - \mathbf{y}_t)$$

- μ: step-size.
- γ: acceleration parameter.

Special cases

- GLA: $d = 1$, $\mu = 1$, $\gamma = 0$ and quadratic loss.
- Wirtinger Flow [Candès et al., 2015]: $d = 2$, $\gamma = 0$ and quadratic loss.

Gradient descent and acceleration

Gradient expression

$$\nabla J(\mathbf{x}) = \frac{d}{2}\,\mathbf{A}^{\mathsf{H}}\left[|\mathbf{A}\mathbf{x}|^{d-2}\odot(\mathbf{A}\mathbf{x})\odot\mathbf{z}\right]$$

$$\mathbf{z} = \underbrace{\psi''(|\mathbf{A}\mathbf{x}|^d)\odot(|\mathbf{A}\mathbf{x}|^d-\mathbf{r})}_{\textit{“right”}} \quad \text{or} \quad \underbrace{\psi'(|\mathbf{A}\mathbf{x}|^d)-\psi'(\mathbf{r})}_{\textit{“left”}}$$

Accelerated gradient descent

Iterate:

$$\mathbf{y}_{t+1} = \mathbf{x}_t - \mu\nabla J(\mathbf{x}_t)$$
$$\mathbf{x}_{t+1} = \mathbf{y}_{t+1} + \gamma(\mathbf{y}_{t+1}-\mathbf{y}_t)$$

- μ: step-size.
- γ: acceleration parameter.

Special cases

- GLA: $d=1$, $\mu=1$, $\gamma=0$ and quadratic loss.
- Wirtinger Flow [Candès et al., 2015]: $d=2$, $\gamma=0$ and quadratic loss.

Experimental protocol

Data

10 speech samples (TIMIT).

Scenarios

- Exact spectrograms.
- Modified spectrograms: simulation of non-consistency by adding Gaussian white noise and Wiener filtering.

Evaluation

- Short-term objective intelligibility (STOI): assessing perceptual intelligibility in the time-domain.
- Signal-to-noise ratio (SNR) improvement: assessing quality in the time-domain.

Experimental protocol

Data

10 speech samples (TIMIT).

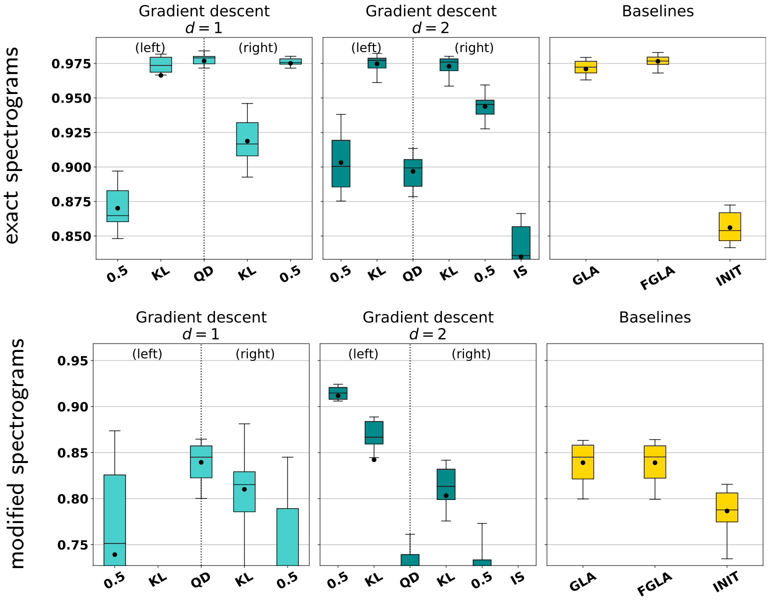
Scenarios

- Exact spectrograms.
- Modified spectrograms: simulation of non-consistency by adding Gaussian white noise and Wiener filtering.

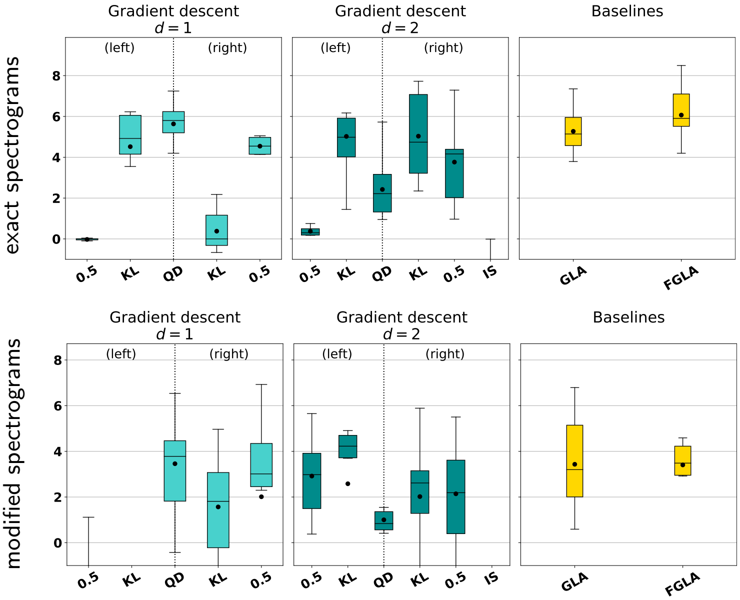
Evaluation

- Short-term objective intelligibility (STOI): assessing perceptual intelligibility in the time-domain.
- Signal-to-noise ratio (SNR) improvement: assessing quality in the time-domain.

Results: STOI



Results: SNR improvement



Phase retrieval with Bregman divergences

Conclusion

- New formulation of PR with Bregman divergences.
- Optimization with gradient descent.
- Promising performances in the presence of high degradation.

Extended work

- ADMM algorithm.
- Long paper under revision (available on arXiv).

Thank you for your attention!

Phase retrieval with Bregman divergences

Conclusion

- New formulation of PR with Bregman divergences.
- Optimization with gradient descent.
- Promising performances in the presence of high degradation.

Extended work

- ADMM algorithm.
- Long paper under revision (available on arXiv).

Thank you for your attention!