Audio source separation with magnitude priors: the BEADS model

Antoine Liutkus[1], Christian Rohlfing[2], Antoine Deleforge[3]

[1] Zenith team, Inria, University of Montpellier, France

[2] RWTH, Aachen University, Germany

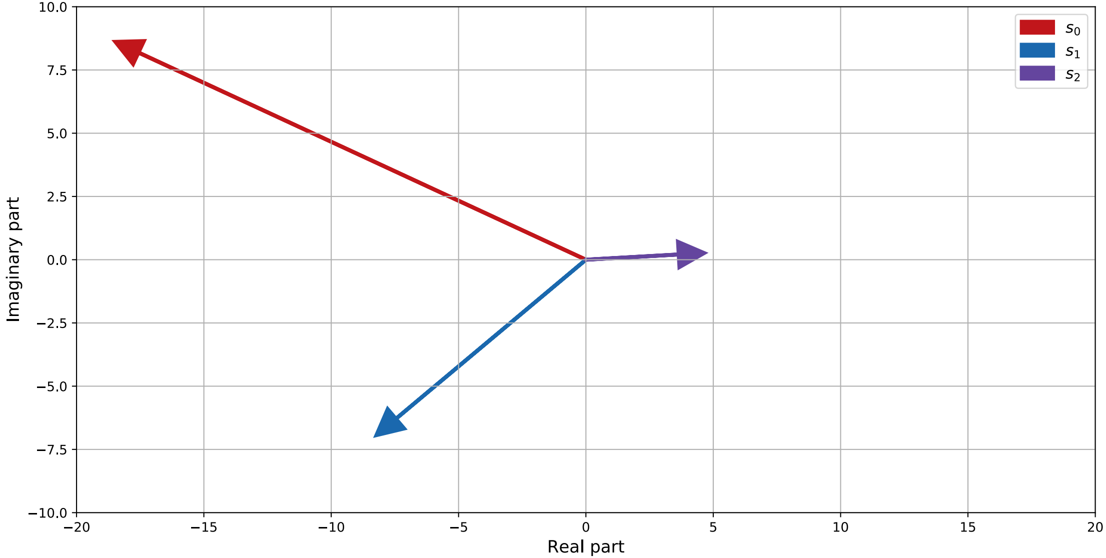
[3] Inria Rennes - Bretagne Atlantique, France

Context

Separation of complex random variables

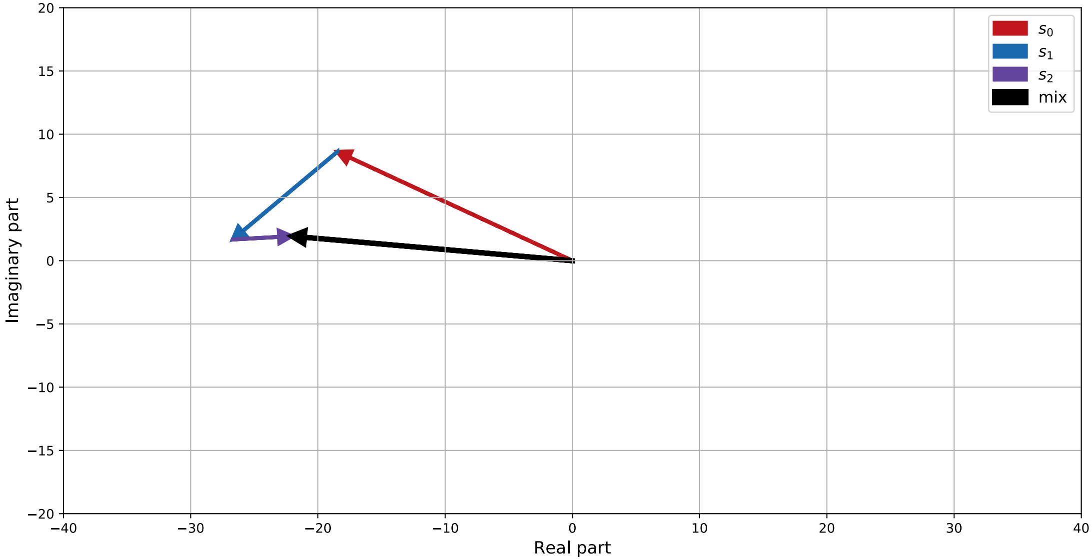
The source separation problem

For each Time-Frequency bin, the mixture is the sum of sources $x = \sum_j s_j$

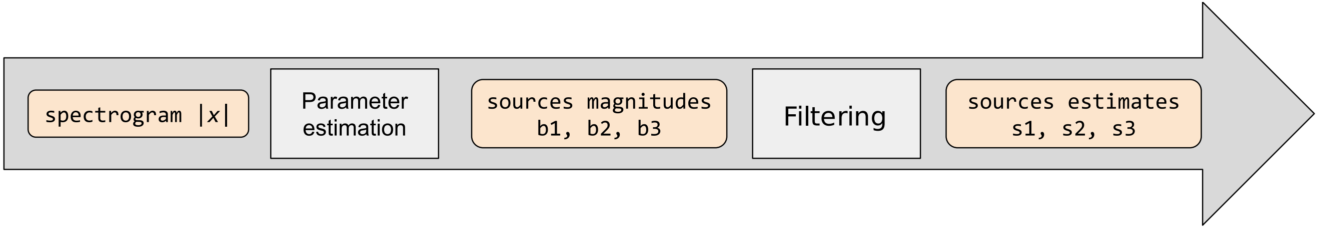


The source separation problem

For each Time-Frequency bin, the mixture is the sum of sources $x = \sum_j s_j$



Typical separation pipeline



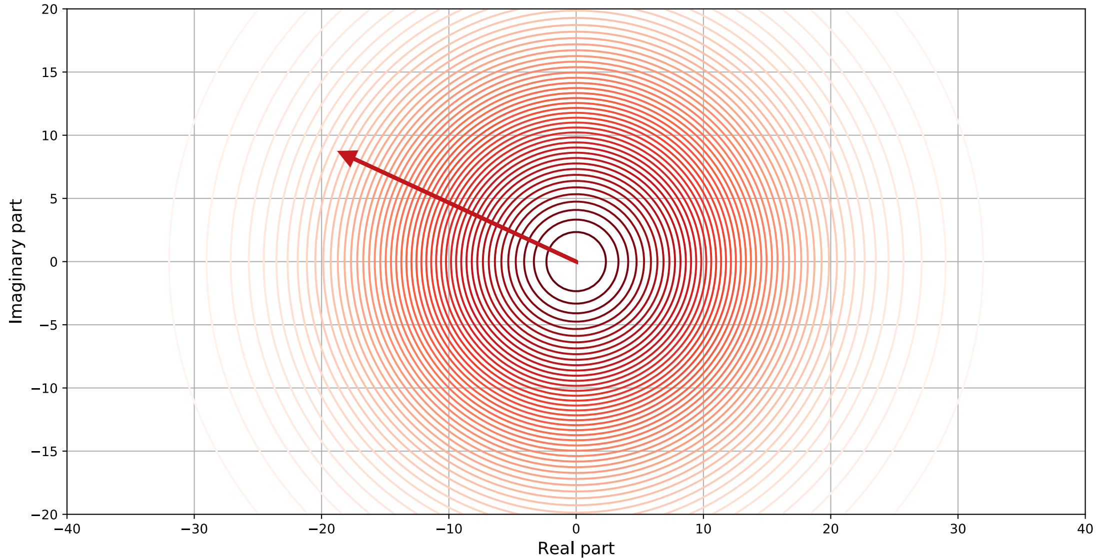
In this talk

- **Filtering** from magnitude estimates $b_j > 0$ to separated signals $s_j \in \mathbb{C}$
- Tractable model for **complex variables s_j with (approximately) known magnitude b_j**

In the paper

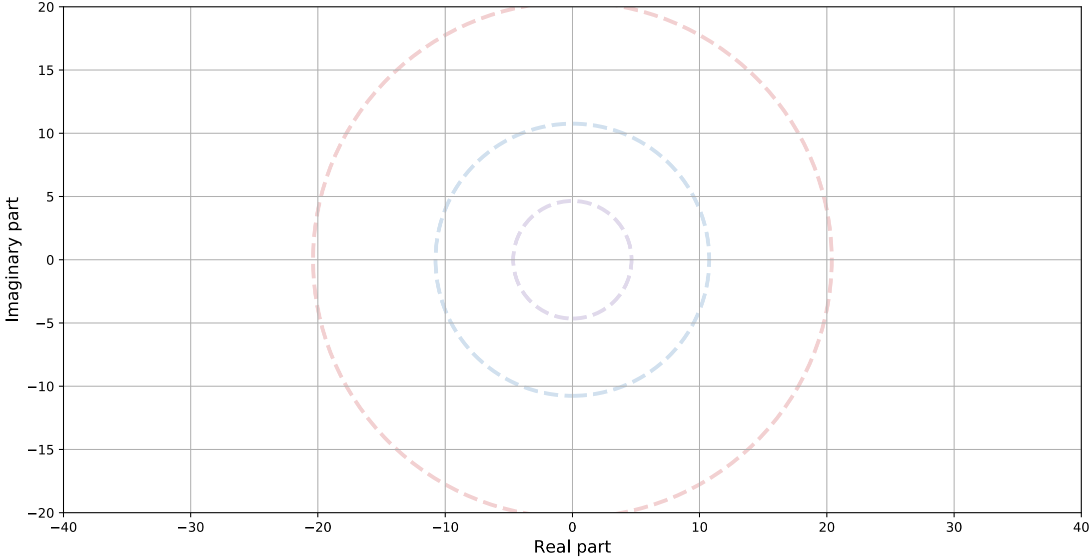
- The multichannel case
- Evaluation for audio coding

The classical Gaussian model $s_j \sim \mathcal{N}\left(0, \frac{2}{\pi} b_j^2\right)$ matches the prior $\mathbb{E}\left[|s_j|\right] = b_j$

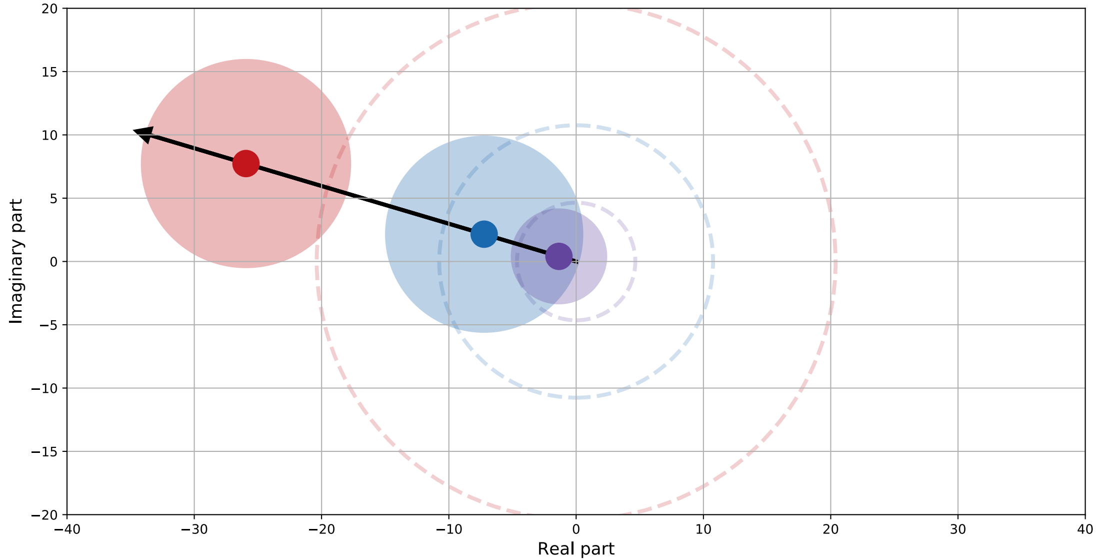


$\Rightarrow$ Highest probability mass on 0

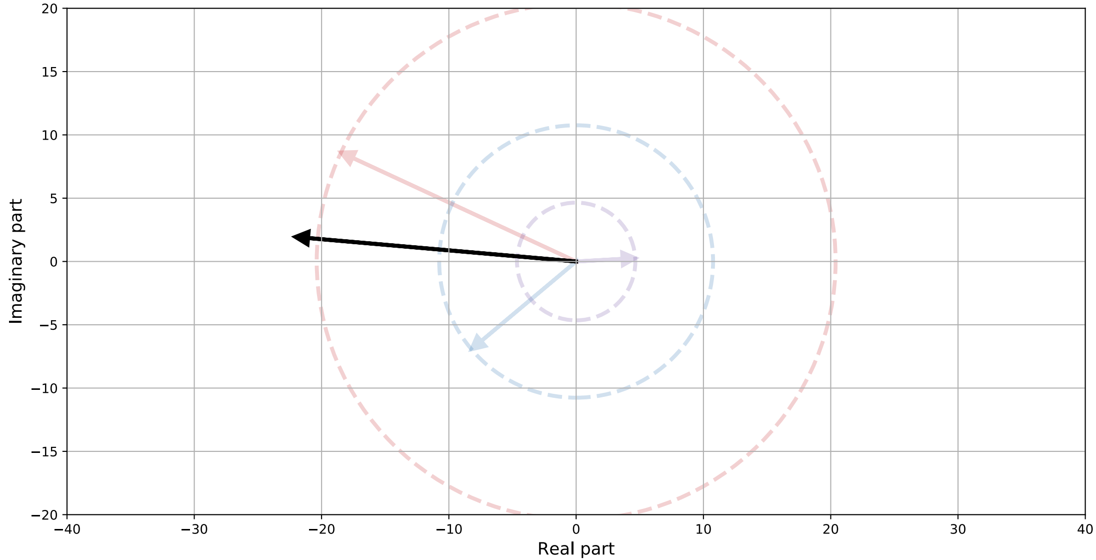
The mixture is Gaussian $x \sim \mathcal{N}\left(0, \sum_j b_j^2\right)$, sources are recovered as: $s \mid x \sim \mathcal{N}\left(\frac{b_j^2}{\sum b^2}x, b_j^2\left(1 - \frac{b_j^2}{\sum b^2}\right)\right)$



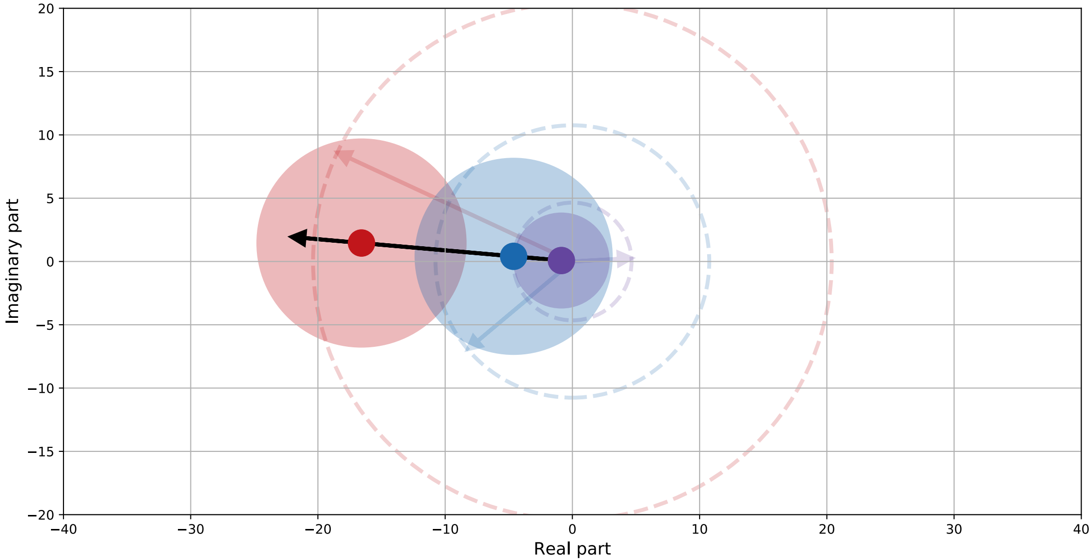
The mixture is Gaussian $x \sim \mathcal{N}\left(0, \sum_j b_j^2\right)$, sources are recovered as: $s \mid x \sim \mathcal{N}\left(\frac{b_j^2}{\sum b^2} x, b_j^2 \left(1 - \frac{b_j^2}{\sum b^2}\right)\right)$



The mixture is Gaussian $x \sim \mathcal{N}\left(0, \sum_j b_j^2\right)$, sources are recovered as: $s \mid x \sim \mathcal{N}\left(\frac{b_j^2}{\sum b^2}x, b_j^2\left(1 - \frac{b_j^2}{\sum b^2}\right)\right)$



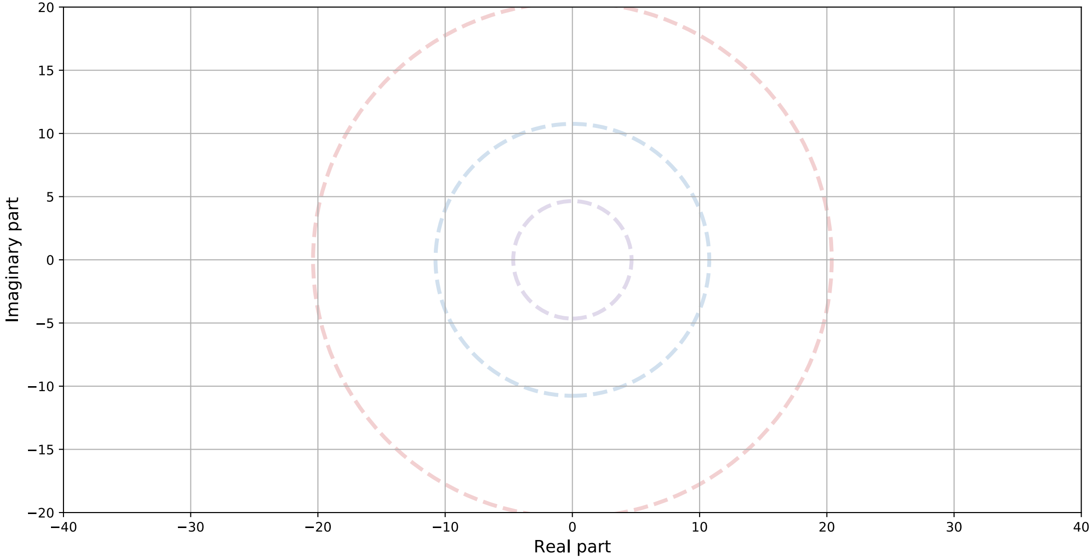
The mixture is Gaussian $x \sim \mathcal{N}\left(0, \sum_j b_j^2\right)$, sources are recovered as: $s \mid x \sim \mathcal{N}\left(\frac{b_j^2}{\sum b^2} x, b_j^2 \left(1 - \frac{b_j^2}{\sum b^2}\right)\right)$



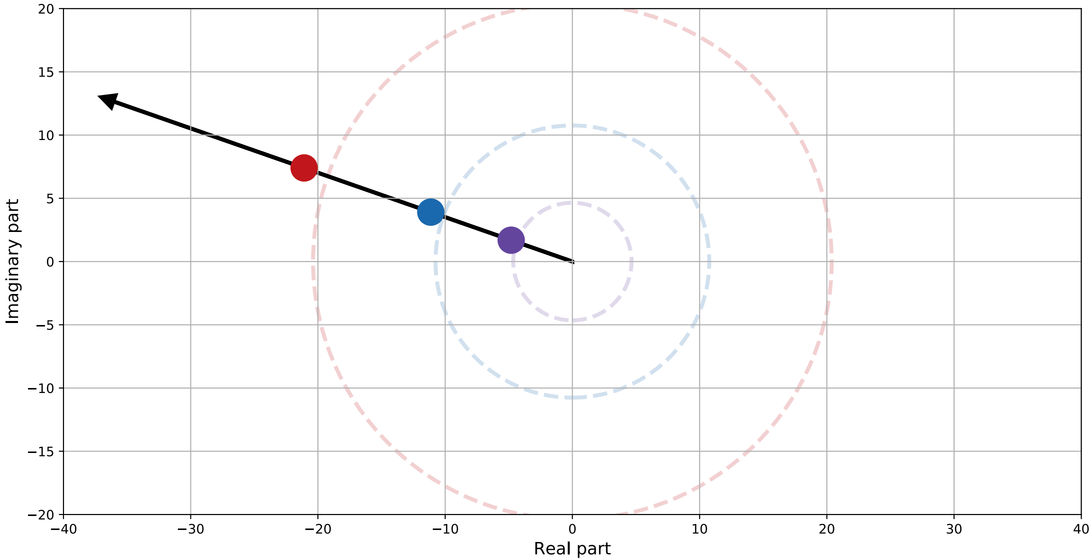
$\Rightarrow$ Aligned estimated sources, magnitudes inconsistent with prior

$\Rightarrow$ Uncertainty independent of the mixture

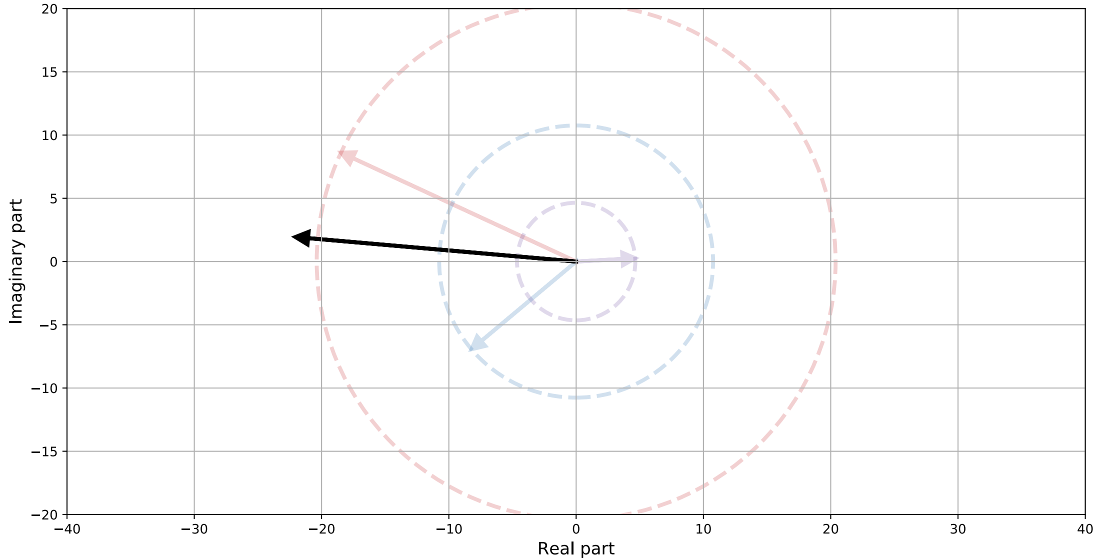
Another classical solution: magnitude ratios: $\hat{s}_j = \frac{b_j}{\sum b} x$



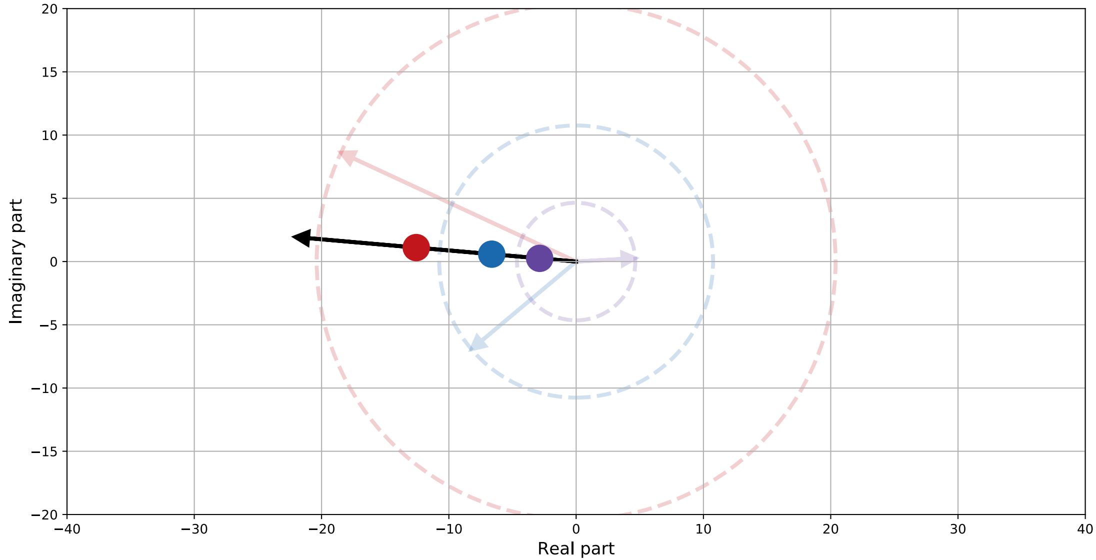
Another classical solution: magnitude ratios: $\hat{s}_j = \frac{b_j}{\sum b} x$



Another classical solution: magnitude ratios: $\hat{s}_j = \frac{b_j}{\sum b} x$



Another classical solution: magnitude ratios: $\hat{s}_j = \frac{b_j}{\sum b} x$



$\Rightarrow$ Still estimating aligned sources rather than complying with the magnitude prior

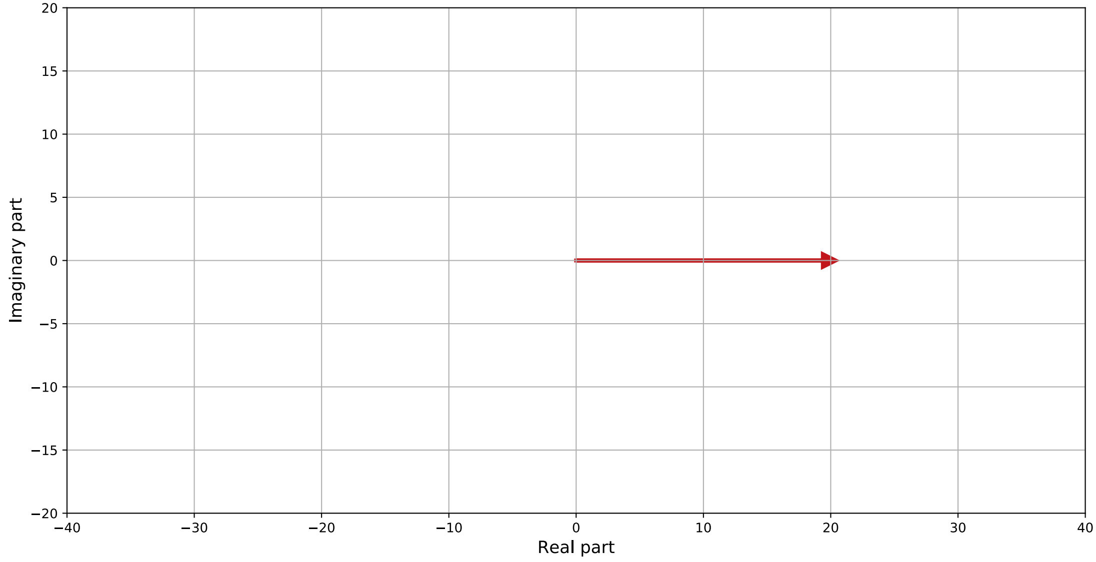
$\Rightarrow$ No tractable uncertainty

An ideal model

The donut-shaped distribution

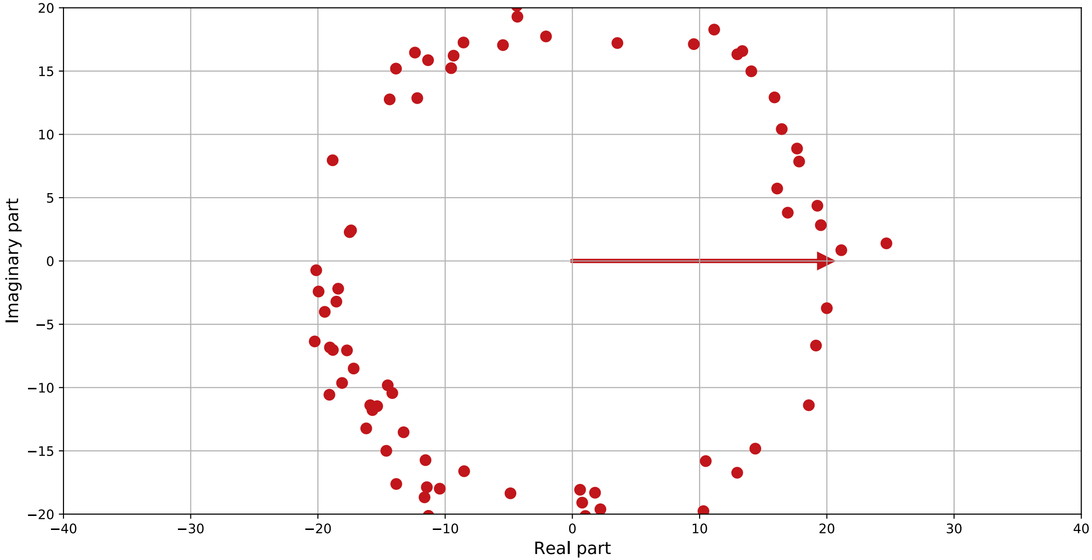
Objective

What do we want of a probabilistic model for a complex random variable with (approximately) known magnitude?



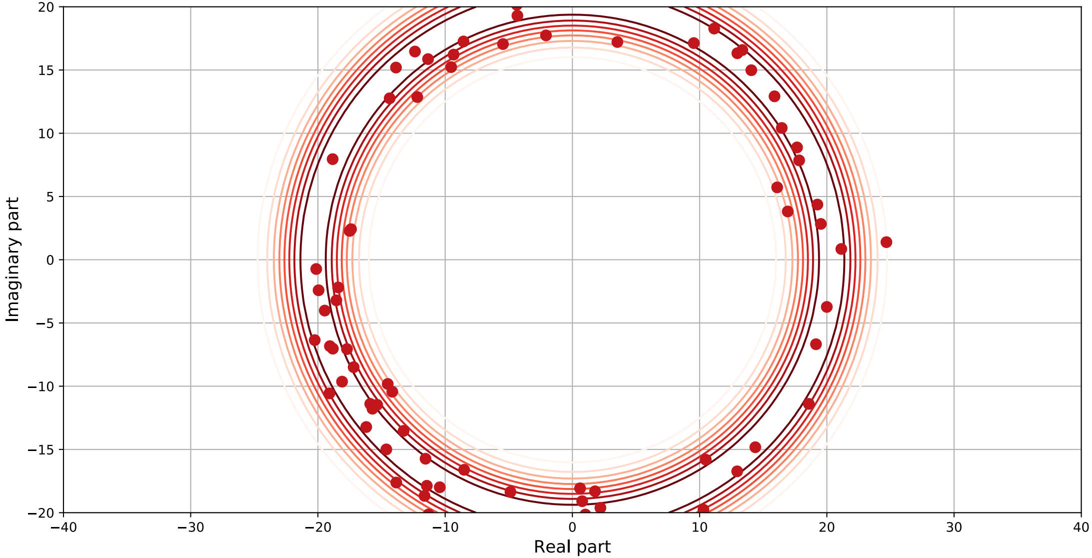
Objective

What do we want of a probabilistic model for a complex random variable with (approximately) known magnitude?

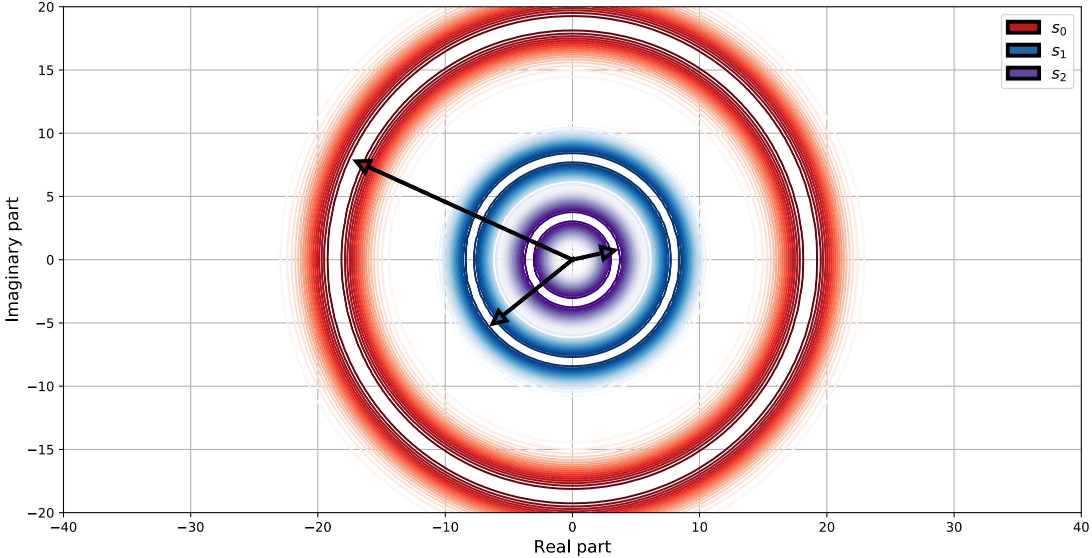


Objective

What do we want of a probabilistic model for a complex random variable with (approximately) known magnitude?



The Donut distribution for modeling the sources

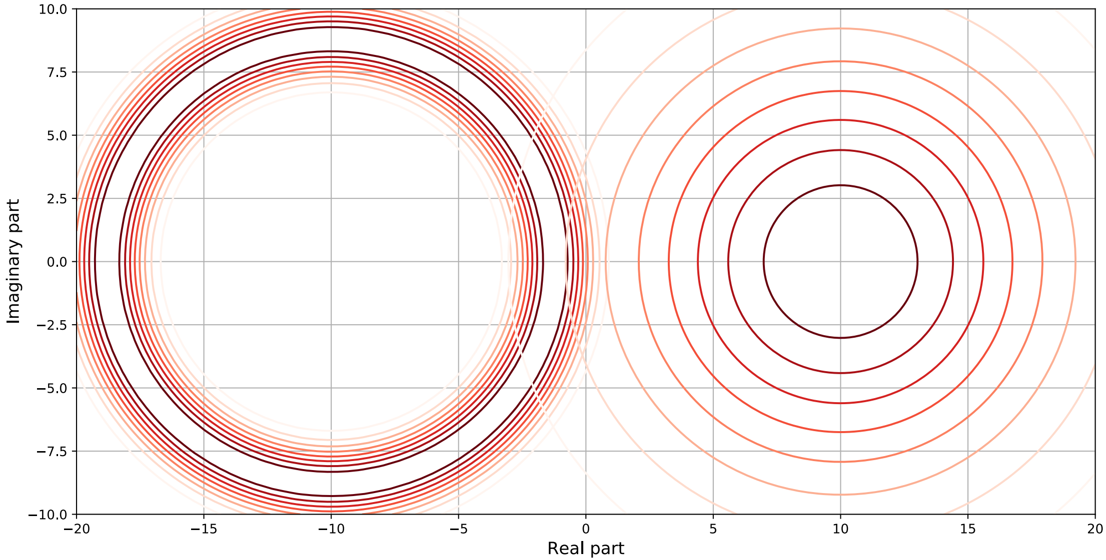


⇒ No model for the sum of donut variables

⇒ No easy way for separation: $\mathbb{P}\left[s \mid x\right]$ non tractable

Contribution

BEADS Bayesian Expansion to Approximate the Donut Shape

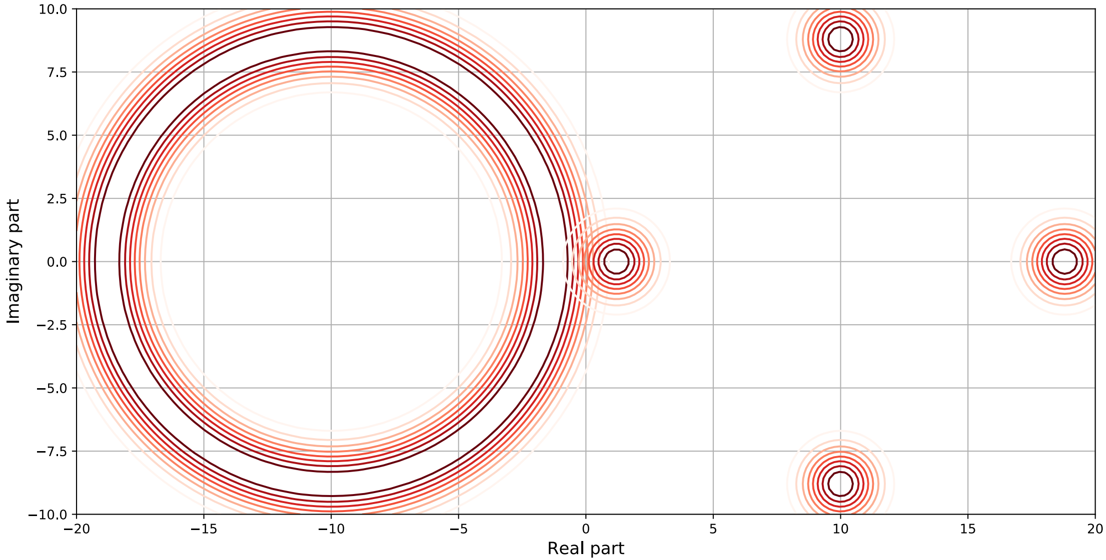


Sources distribution as a Gaussian Mixture Model: $P\left[s_j\right] = \sum_c \pi[c]\mathcal{N}\left(s_j \mid b_j\omega^c, \sigma_j\right)$

$\Rightarrow$ Only two parameters: b_j and σ_j

Contribution

BEADS Bayesian Expansion to Approximate the Donut Shape

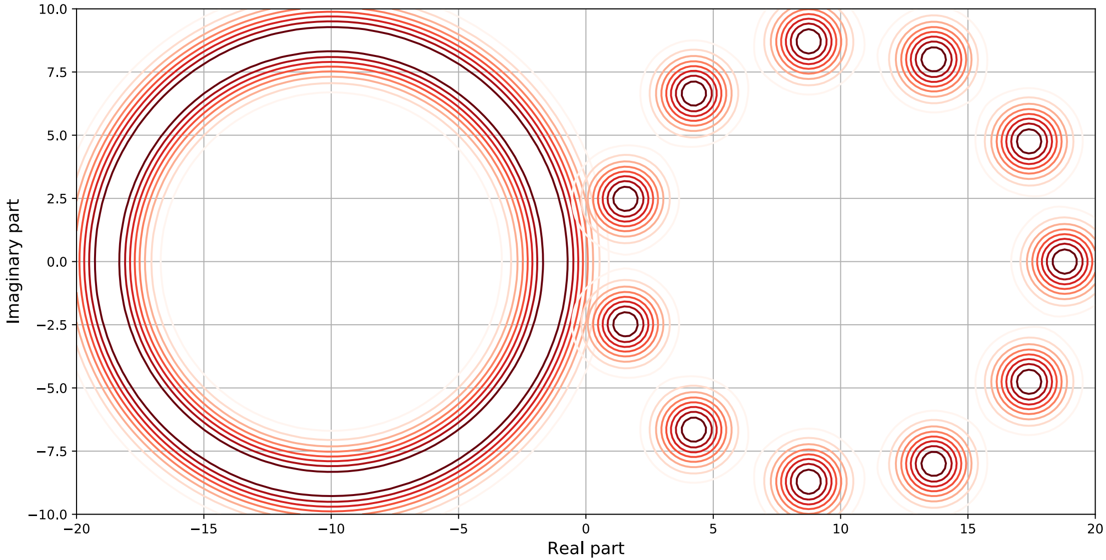


Sources distribution as a Gaussian Mixture Model: $P\left[s_j\right] = \sum_c \pi[c]\mathcal{N}\left(s_j \mid b_j\omega^c, \sigma_j\right)$

$\Rightarrow$ Only two parameters: b_j and σ_j

Contribution

BEADS Bayesian Expansion to Approximate the Donut Shape

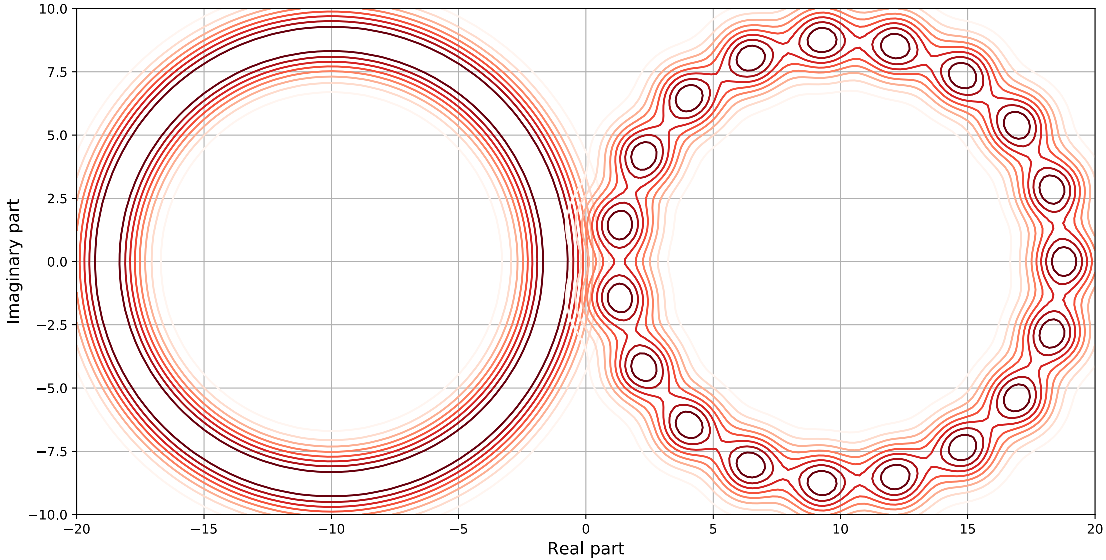


Sources distribution as a Gaussian Mixture Model: $P\left[s_j\right] = \sum_c \pi[c]\mathcal{N}\left(s_j \mid b_j\omega^c, \sigma_j\right)$

$\Rightarrow$ Only two parameters: b_j and σ_j

Contribution

BEADS Bayesian Expansion to Approximate the Donut Shape

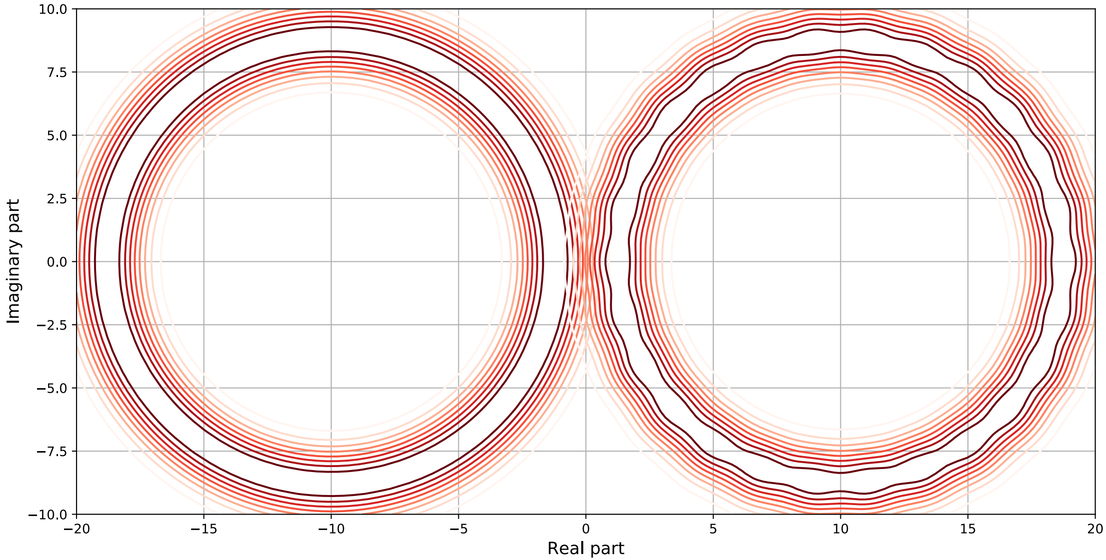


Sources distribution as a Gaussian Mixture Model: $P\left[s_j\right] = \sum_c \pi[c]\mathcal{N}\left(s_j \mid b_j\omega^c, \sigma_j\right)$

$\Rightarrow$ Only two parameters: b_j and σ_j

Contribution

BEADS Bayesian Expansion to Approximate the Donut Shape

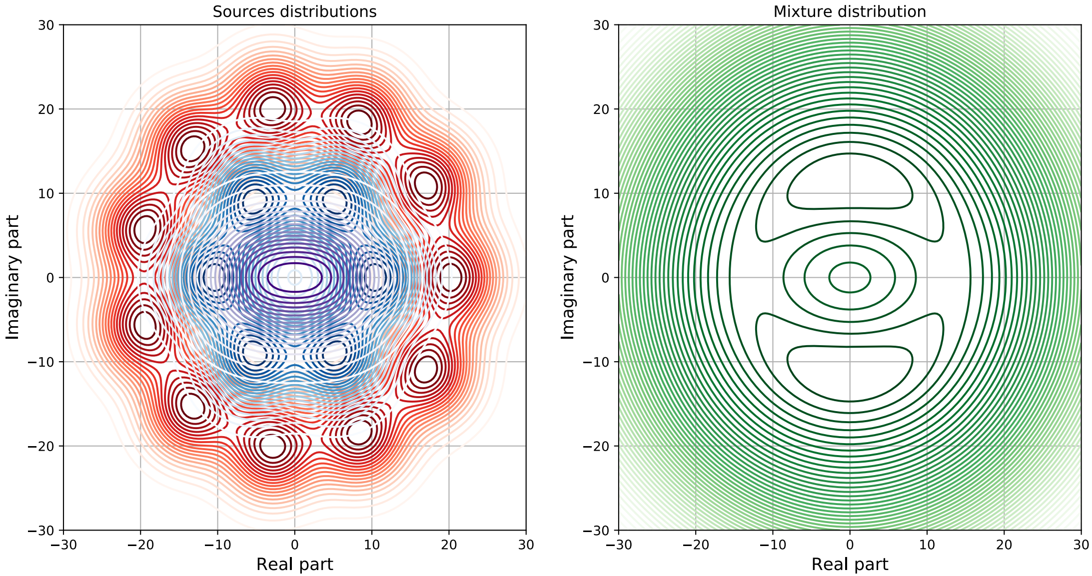


Sources distribution as a Gaussian Mixture Model: $P\left[s_j\right] = \sum_c \pi[c]\mathcal{N}\left(s_j \mid b_j\omega^c, \sigma_j\right)$

$\Rightarrow$ Only two parameters: b_j and σ_j

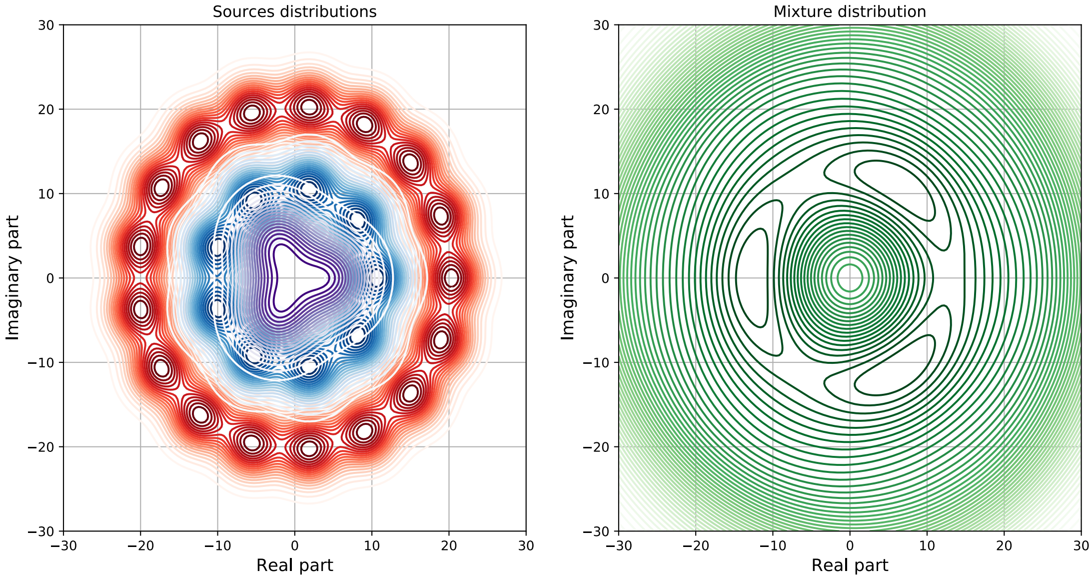
Summing beads random variables

BEADS model for the sources ⇒ Gaussian Mixture Model for the mixture



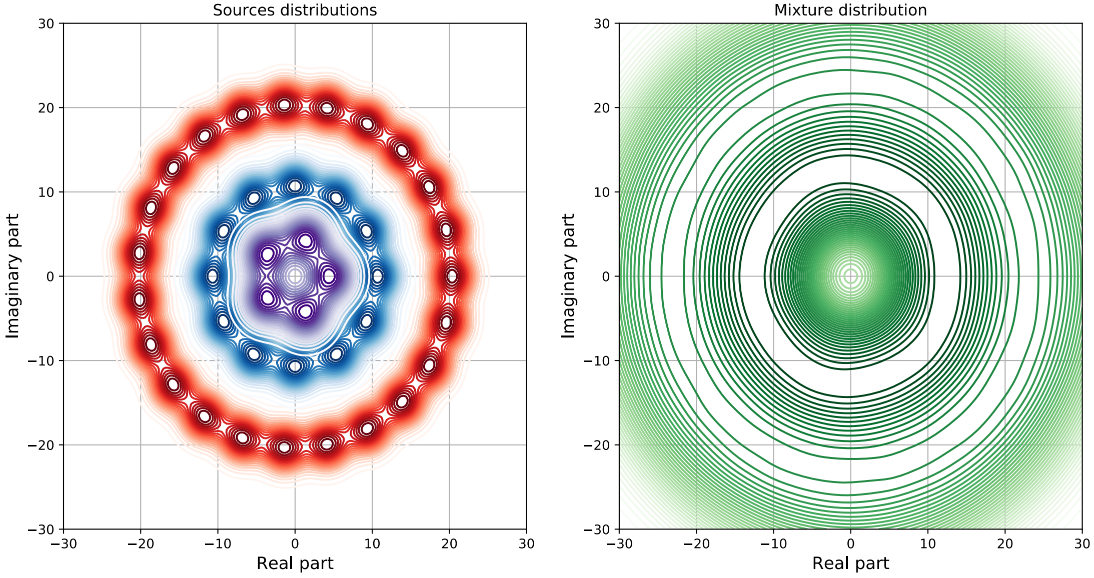
Summing beads random variables

BEADS model for the sources $\Rightarrow$ Gaussian Mixture Model for the mixture

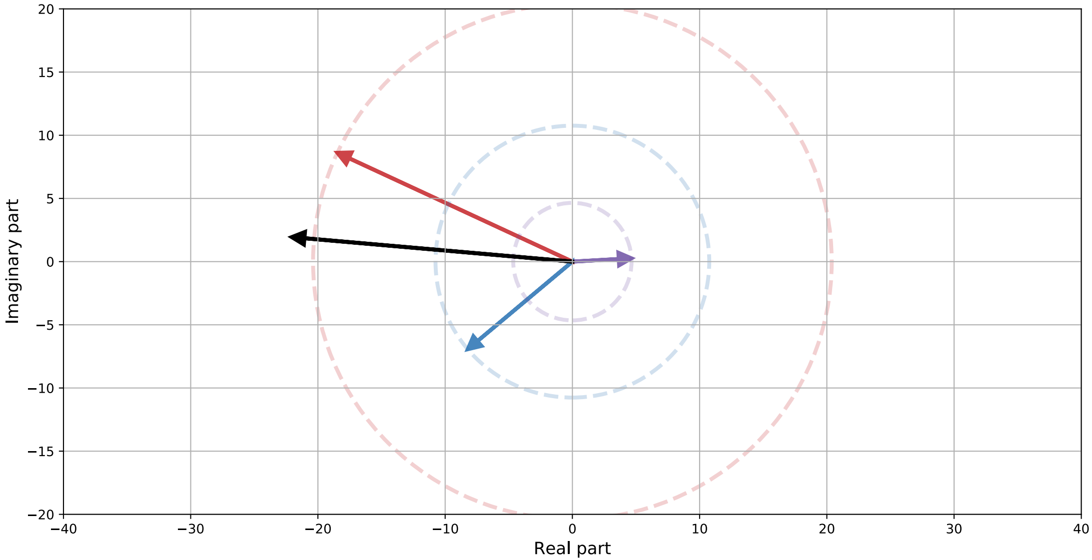


Summing beads random variables

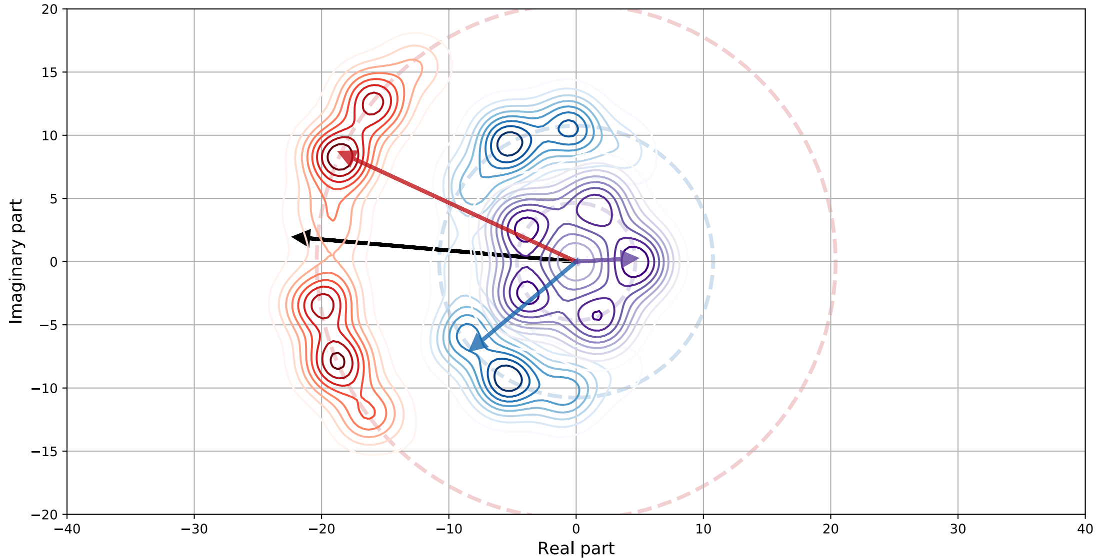
BEADS model for the sources $\Rightarrow$ Gaussian Mixture Model for the mixture



The sources are estimated through Bayes theorem as $s \mid x = \sum_c \pi(c \mid x)\mathcal{N}(s \mid \mu_{c|x}, \sigma_{|x})$



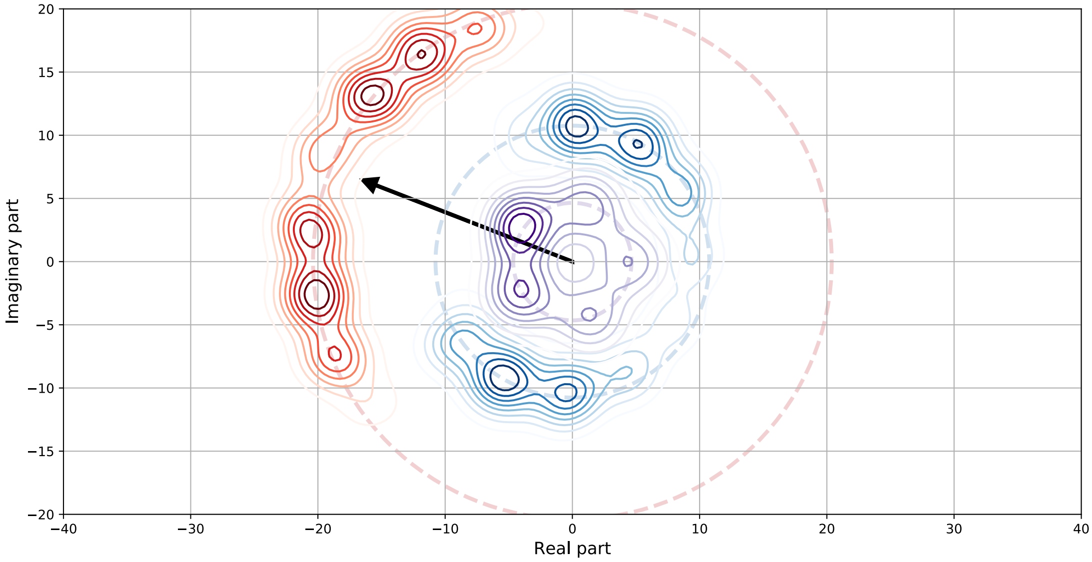
The sources are estimated through Bayes theorem as $s \mid x = \sum_c \pi(c \mid x)\mathcal{N}(s \mid \mu_{c|x}, \sigma_{|x})$



⇒ Posterior is tractable, estimates consistent with the magnitude prior

⇒ Uncertainty is mix-dependent

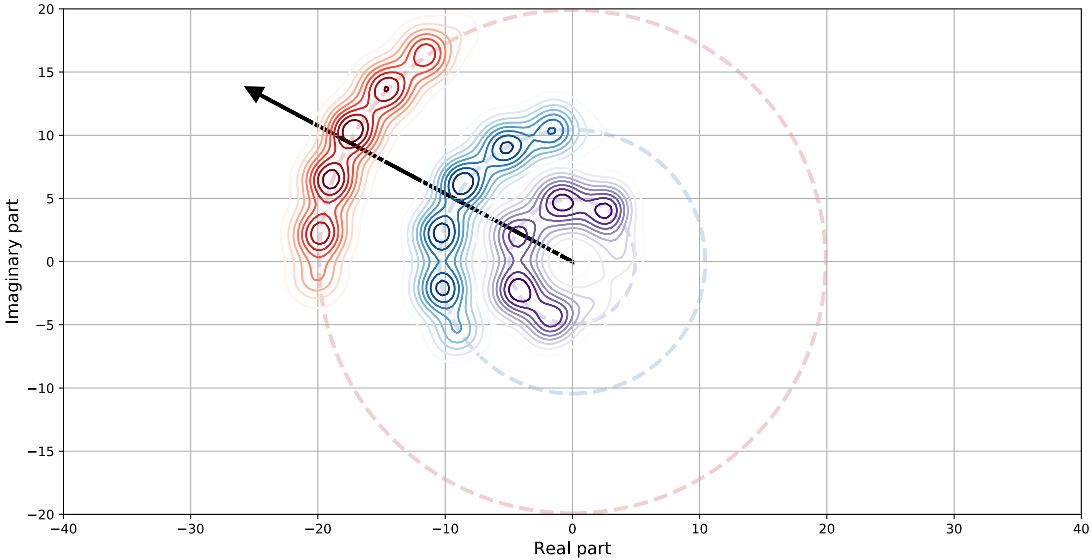
The sources are estimated through Bayes theorem as $s \mid x = \sum_c \pi(c \mid x)\mathcal{N}(s \mid \mu_{c|x}, \sigma_{|x})$



⇒ Posterior is tractable, estimates consistent with the magnitude prior

⇒ Uncertainty is mix-dependent

The sources are estimated through Bayes theorem as $s \mid x = \sum_c \pi(c \mid x)\mathcal{N}(s \mid \mu_{c|x}, \sigma_{|x})$



⇒ Posterior is tractable, estimates consistent with the magnitude prior

⇒ Uncertainty is mix-dependent

Conclusion: The beads model

Core advantages

- Complex random variables with approximately known magnitudes
- Sums of beads sources is a GMM
- Separation is easy as GMM inference

To go further

- Generalizes easily to multichannel
- Shared variances for the beads $\Rightarrow$ computational savings

Source code for this presentation

https://github.com/aliutkus/beads-presentation