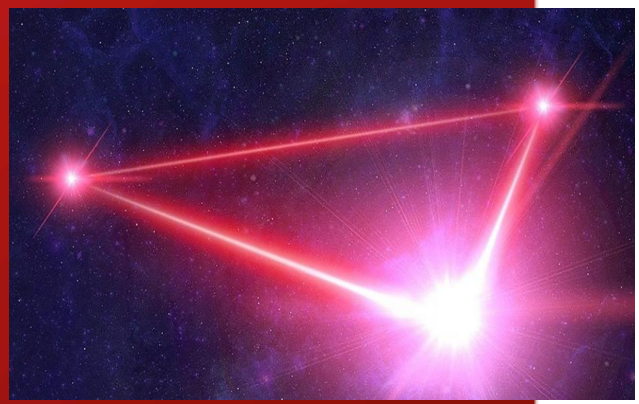
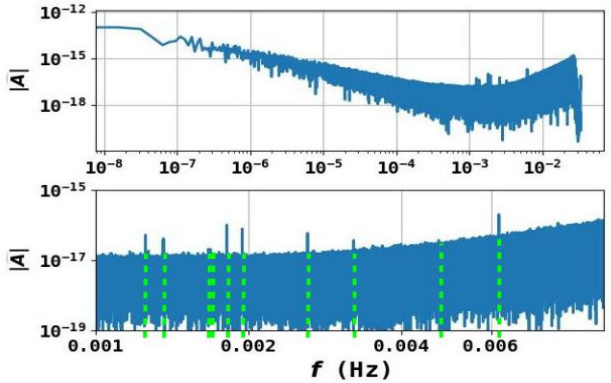
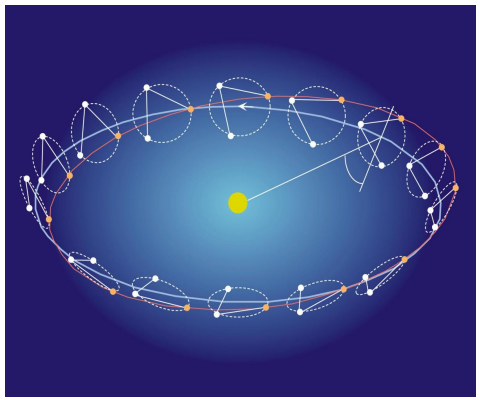
DE LA RECHERCHE À L'INDUSTRIE

www.cea.fr

Sparse data inpainting for the recovery of gapped gravitational wave signals







GDR OG Assembly | Hervé MOUTARDE

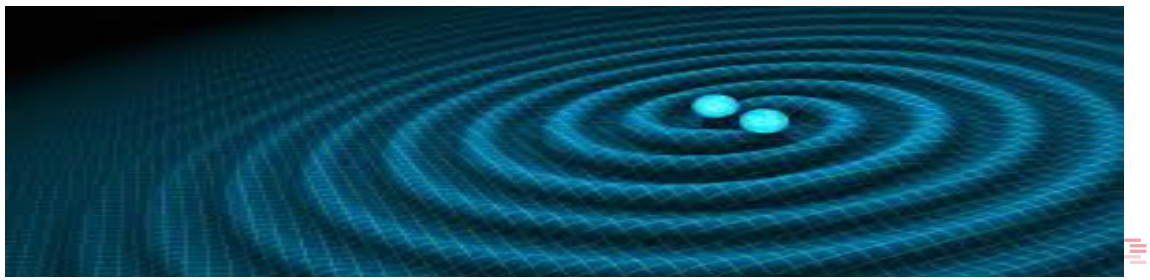
15 Oct. 2020

université PARIS-SACLAY

Galactic binaries.
Physics and data analysis aspects.

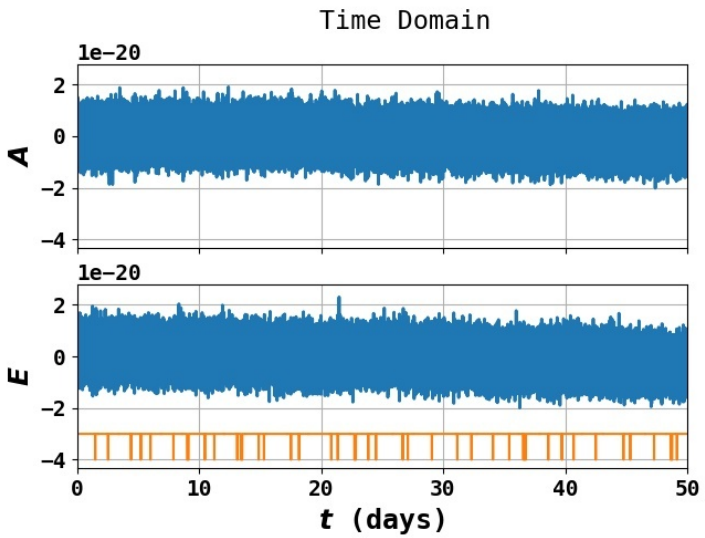
- Physics perspective:
  - Expect $\sim 10^6$ sources.
  - Aim: **characterize** a subset of $\sim 10^4$ sources.
  - **Foreground** of many other signals of physical interest.
- Data analysis perspective:
  - Characterization of individual sources.
  - Determination of a **confusion noise**.
  - Assessment of impact of **gaps**, **glitches** or **non-stationary** power spectral density.

    Baghi *et al.*, Phys. Rev. **D100**, 022003 (2019)
  - Low signal to noise ratio requires long observations.

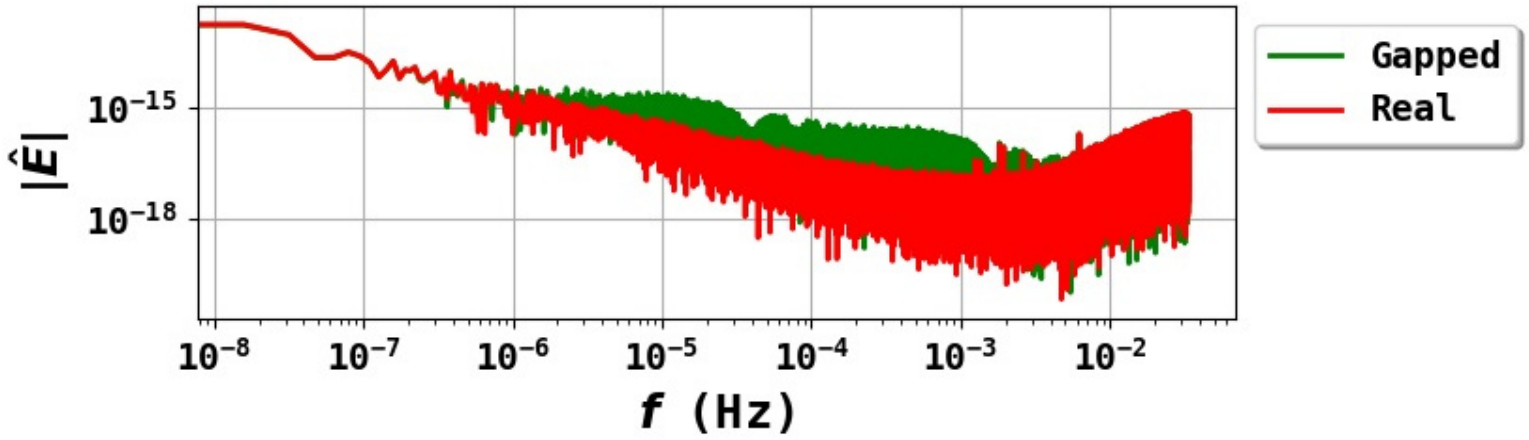


Galactic binaries from realistic LISA data.

Focus on gaps.

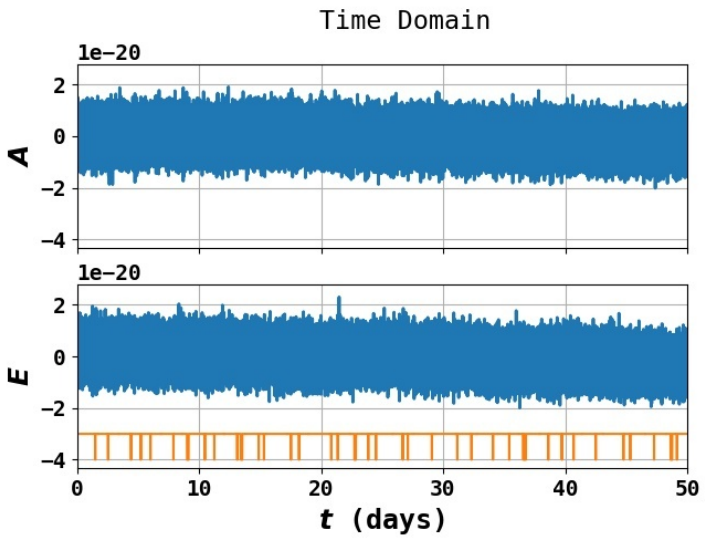


- Origins : antennas, laser frequency, orbit corrections, ...
- Both **planned** and **unplanned** gaps.
- Features : 1h / 5 days, 10 min. / day and up to 5 days (unplanned, LISA Pathfinder).

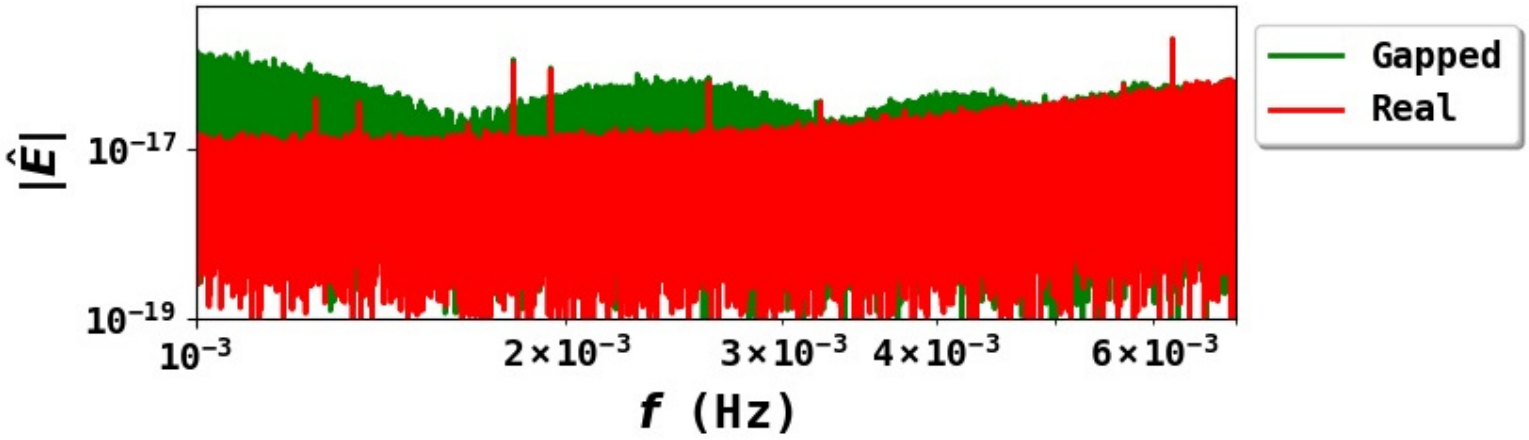


Galactic binaries from realistic LISA data.

Focus on gaps.

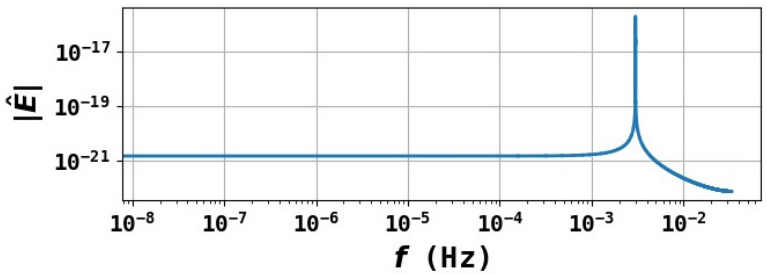


- Origins : antennas, laser frequency, orbit corrections, ...
- Both **planned** and **unplanned** gaps.
- Features : 1h / 5 days, 10 min. / day and up to 5 days (unplanned, LISA Pathfinder).

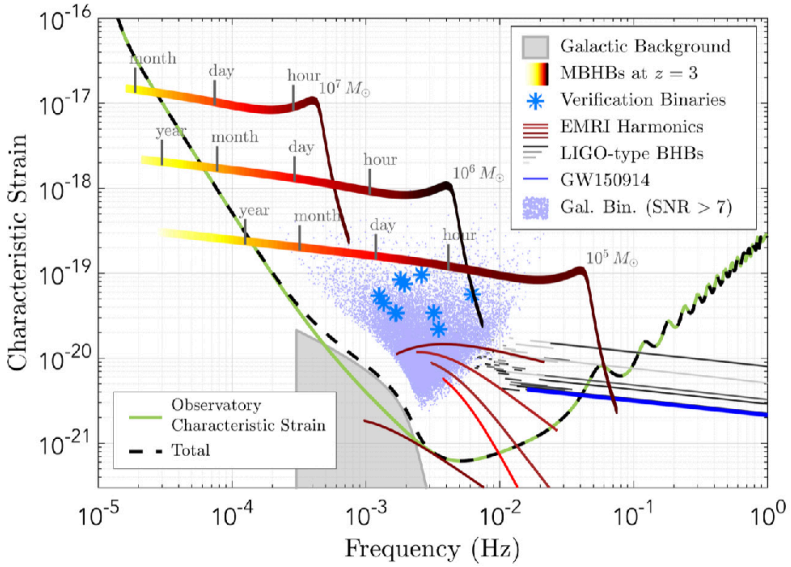


Sparse data representation.

Exploit distinctive structural features of signal and noise.



- Expected signal is **sparse in Fourier basis**.



- Instrumental noise is **Gaussian** with known power spectral density Σ.

- Detection through **non-parametric** approach.
- Possible extension to **other types of signals** by **adequate choice of basis** with sparse representation.

Ungapped data

Statistical estimator of the signal.
Construction by resolution of an optimization problem.

- Introduce:
  - Data: time series V of length N.
  - **Whitened** variables $\widehat{V} = \Sigma^{-1/2} V$ in frequency domain.
  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k] \widehat{V}[k]$.
- Signal estimator:

$$\widehat{S} =$$

Statistical estimator of the signal.
Construction by resolution of an optimization problem.

- Introduce:
  - Data: time series V of length N.
  - **Whitened** variables $\widehat{V} = \Sigma^{-1/2} V$ in frequency domain.
  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k] \widehat{V}[k]$.
- Signal estimator:

$$\widehat{S} = \underbrace{\underset{\widehat{V} \in \mathbb{C}^N}{\text{Argmin}}}_{\text{Trial signal}}$$

Statistical estimator of the signal.

Construction by resolution of an optimization problem.

- Introduce:
  - Data: time series V of length N.
  - **Whitened** variables $\widehat{V} = \Sigma^{-1/2} V$ in frequency domain.
  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k] \widehat{V}[k]$.
- Signal estimator:

$$\widehat{S} = \underbrace{\underset{\widehat{V} \in \mathbb{C}^N}{\text{Argmin}}}_{\text{Trial signal}} \left[\underbrace{\left\| \gamma \odot \widehat{V} \right\|_1}_{\text{Sparsity enforcing}} \right]$$

Statistical estimator of the signal.

Construction by resolution of an optimization problem.

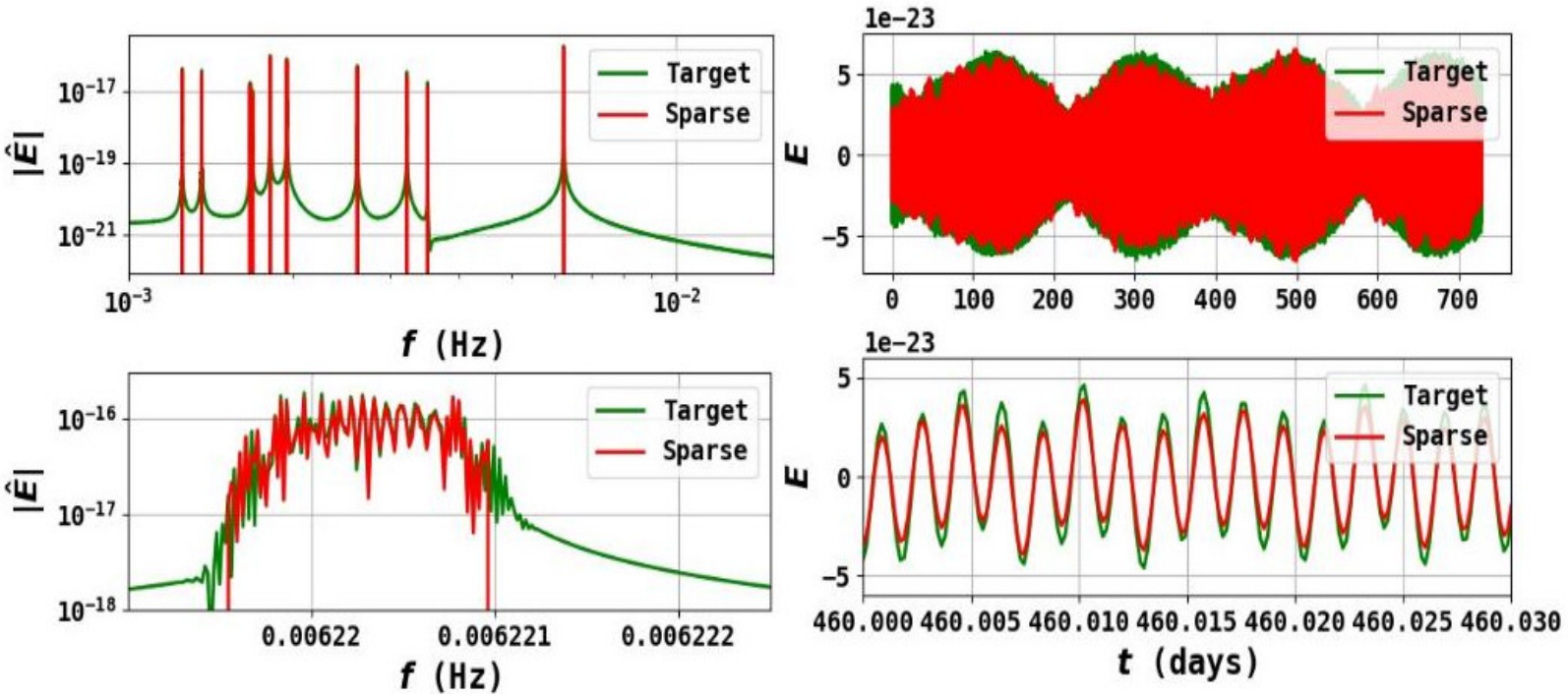
- Introduce:
  - Data: time series V of length N.
  - **Whitened** variables $\widehat{V} = \Sigma^{-1/2} V$ in frequency domain.
  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k] \widehat{V}[k]$.
- Signal estimator:

$$\widehat{S} = \underbrace{\underset{\widehat{V} \in \mathbb{C}^N}{\text{Argmin}}}_{\text{Trial signal}} \left[\underbrace{\left\| \gamma \odot \widehat{V} \right\|_1}_{\text{Sparsity enforcing}} + \underbrace{\frac{1}{2} \left\| \widehat{U} - \widehat{V} \right\|_2^2}_{\text{Data adequation}} \right]$$

- Extension to coupled TDI channels, frequency blocks, etc.

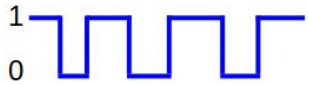
Blelly *et al.* [arXiv:2005.03696 [gr-qc]], Phys. Rev. D

Results.
Test on LDC1-3: verification galactic binaries.



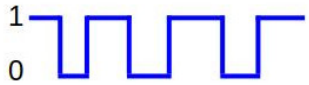
- **All peaks detected**, even for weak amplitudes.
- **No false positive signals**.
- **Accurate** phase description.
- Amplitude can be improved by adapted basis choice.

Gapped data

- Introduce:
  - Data mask M: 

  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k] \widehat{V}[k]$.
- Signal and noise estimators:

$$(S, N) =$$

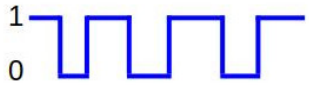
- Introduce:
  - Data mask M: 

  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k] \widehat{V}[k]$.
- Signal and noise estimators:

$$(S, N) = \underbrace{\underset{Y=M(S+N)}{\text{Argmin}}}_{\text{Data Adequacy}}$$

Statistical estimators of signal and noise.

Construction by resolution of an optimization problem.

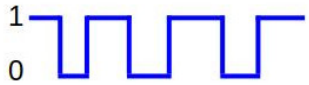
- Introduce:
  - Data mask M: 

  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k]\widehat{V}[k]$.
- Signal and noise estimators:

$$(S, N) = \underbrace{\underset{Y=M(S+N)}{\text{Argmin}}}_{\text{Data Adequacy}} \underbrace{\left[\left\| \gamma \odot \Sigma^{-1/2} \widehat{S} \right\|_{12} \right]}_{\text{Solution prior (sparsity)}}$$

Statistical estimators of signal and noise.

Construction by resolution of an optimization problem.

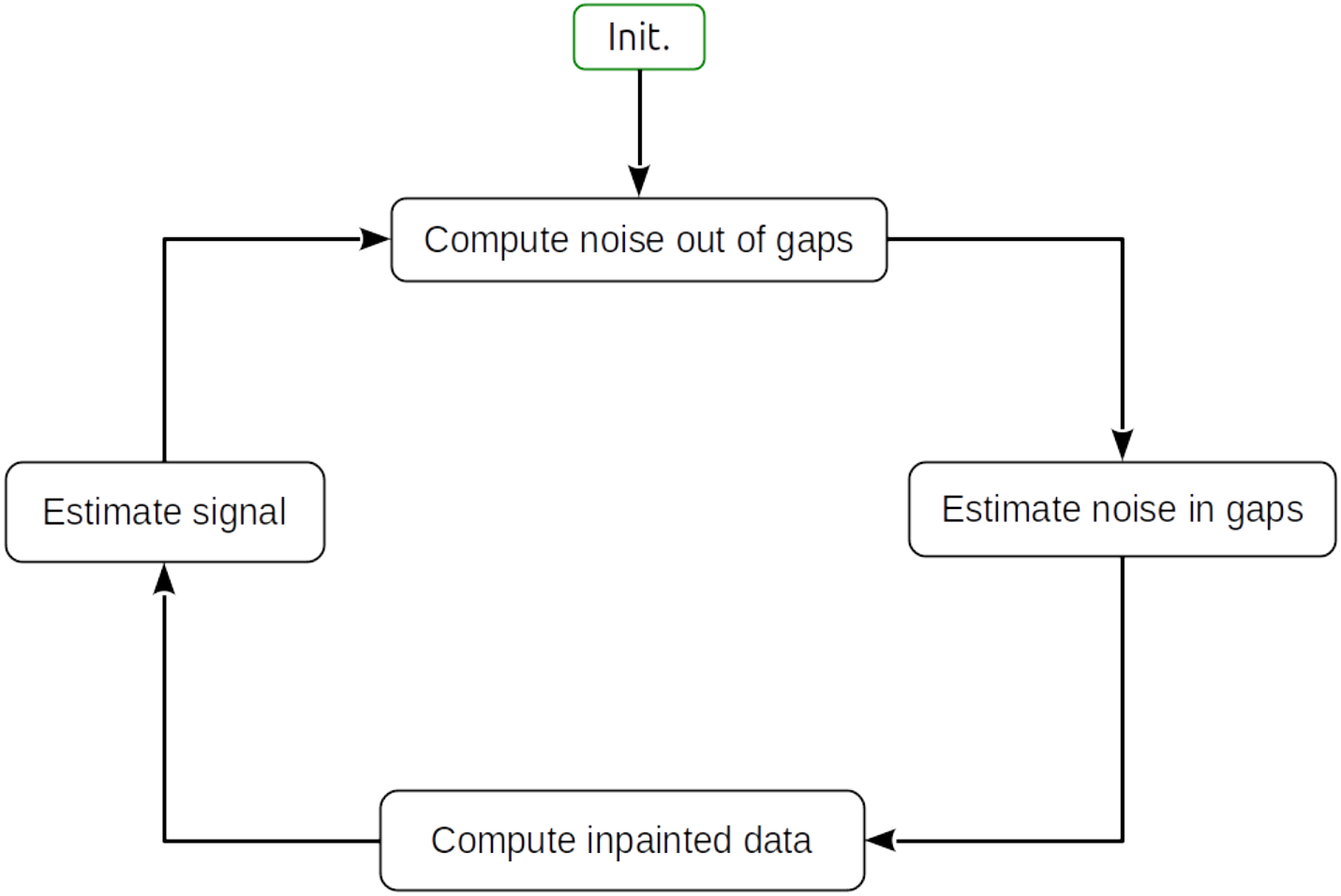
- Introduce:
  - Data mask M: 

  - **Threshold** γ to separate signal from noise.
  - Compact notation: $\gamma \odot \widehat{V} = \sum_k \gamma[k]\widehat{V}[k]$.
- Signal and noise estimators:

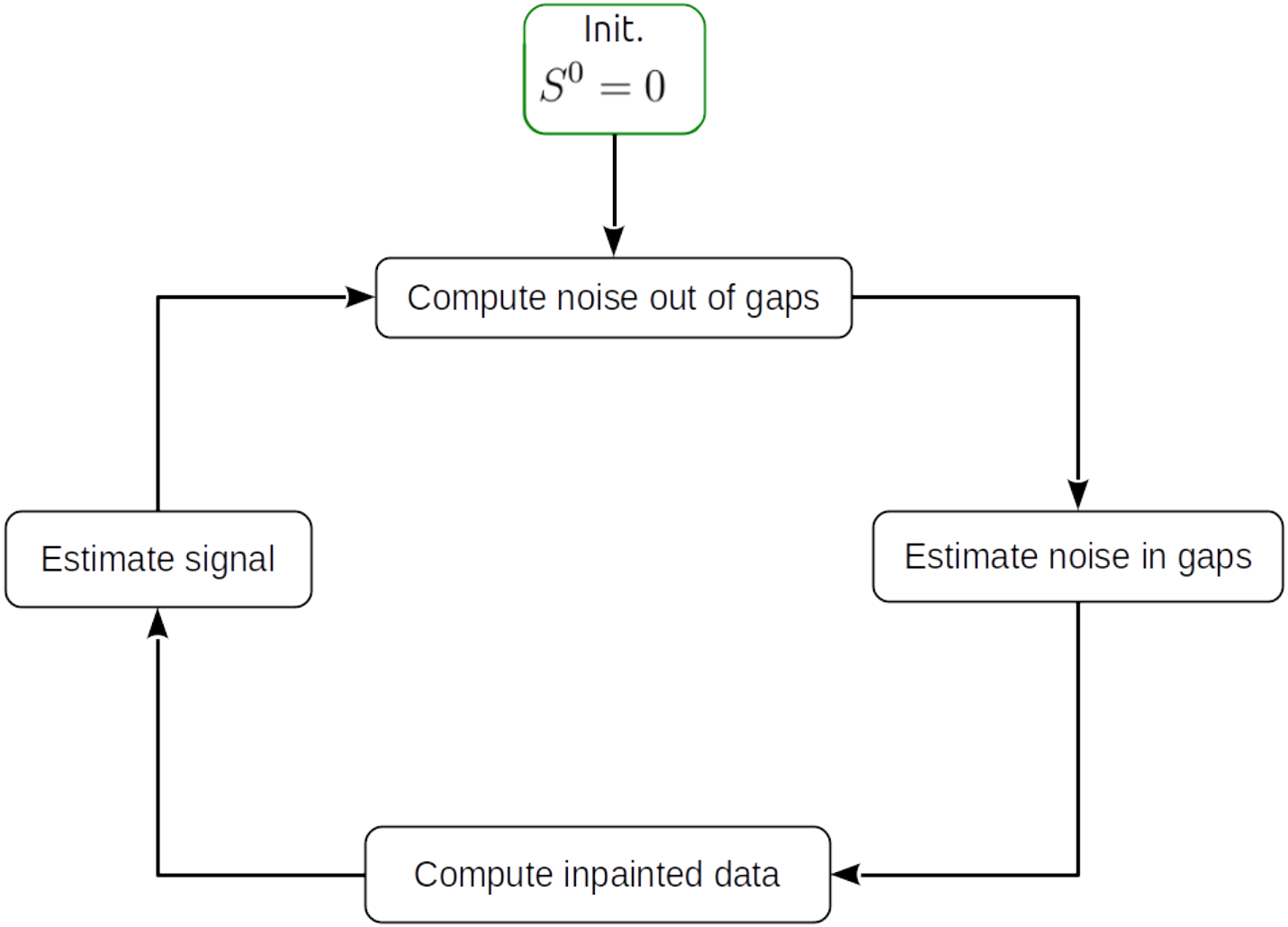
$$(S,N) = \underbrace{\underset{Y=M(S+N)}{\text{Argmin}}}_{\text{Data Adequacy}} \left[\underbrace{\left\| \gamma \odot \Sigma^{-1/2}\widehat{S} \right\|_{12}}_{\text{Solution prior (sparsity)}} + \underbrace{\frac{1}{2}\widehat{N}^T\Sigma^{-1}\widehat{N}}_{\text{Noise prior (Gaussian)}} \right]$$

- Resolution alternates between signal and noise estimations.

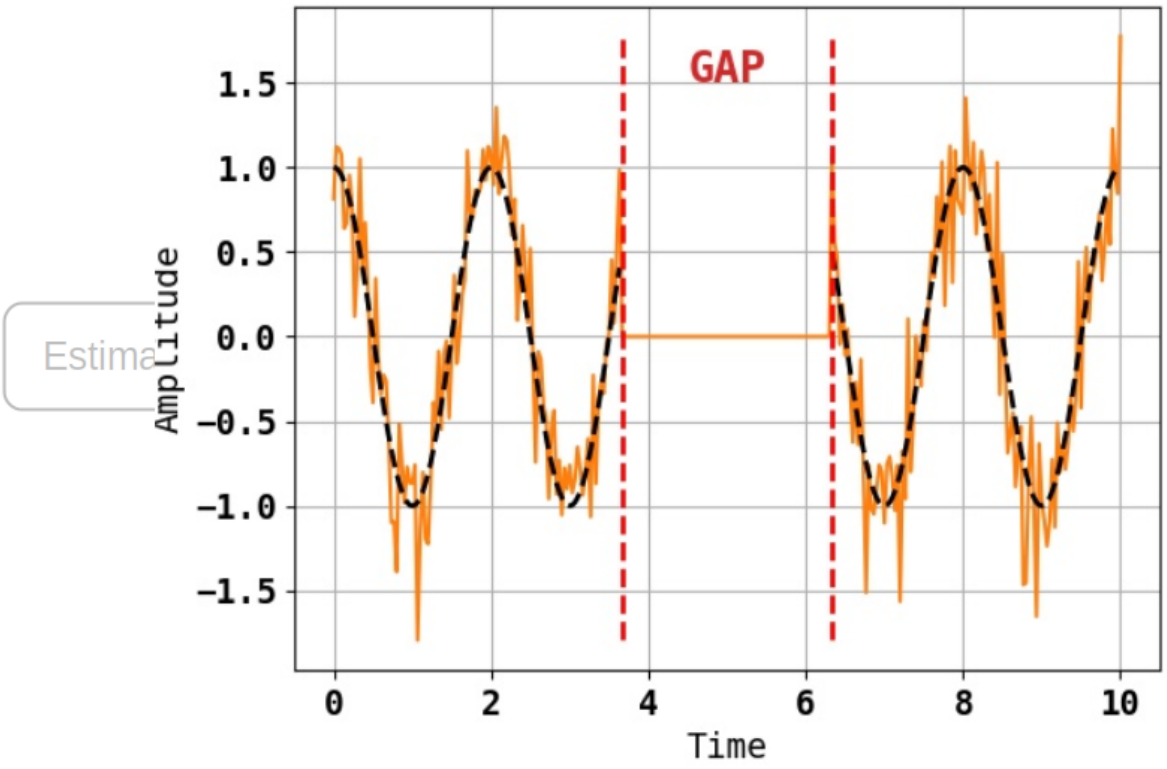
Optimization problem: algorithm.
Description of an iteration.



Optimization problem: algorithm.
Description of an iteration.

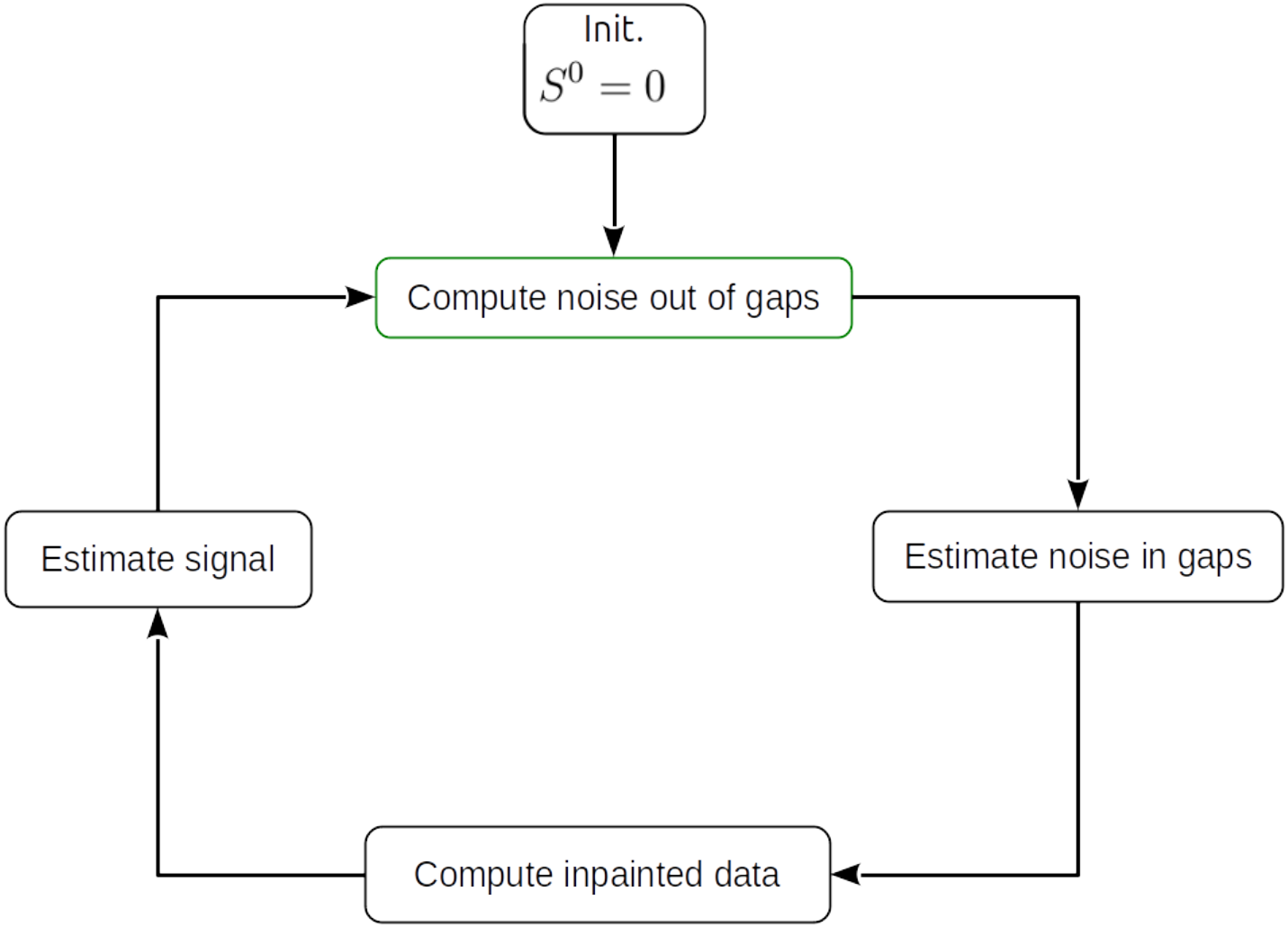


Init.
$S^0 = 0$

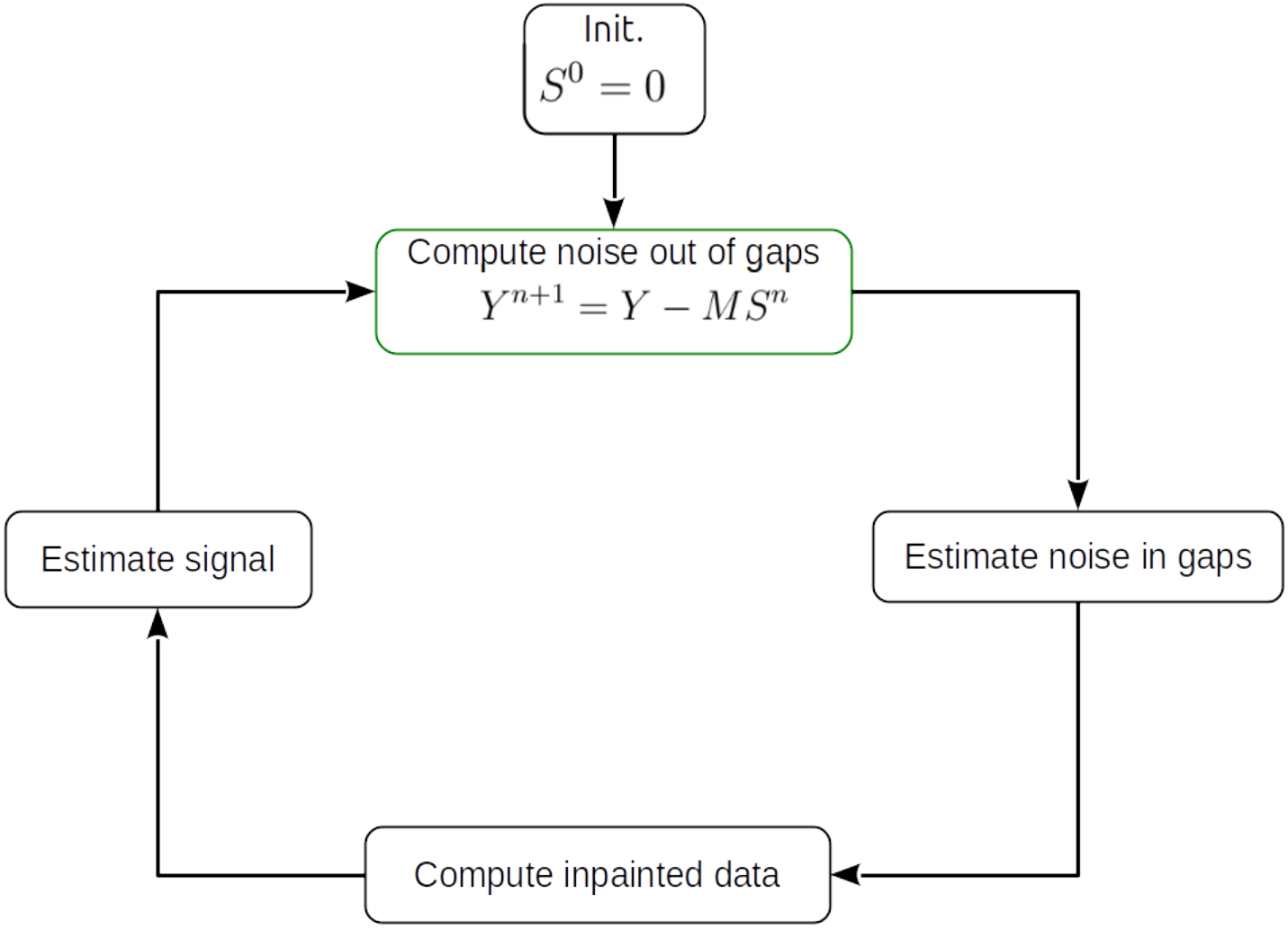


Estima

n gaps



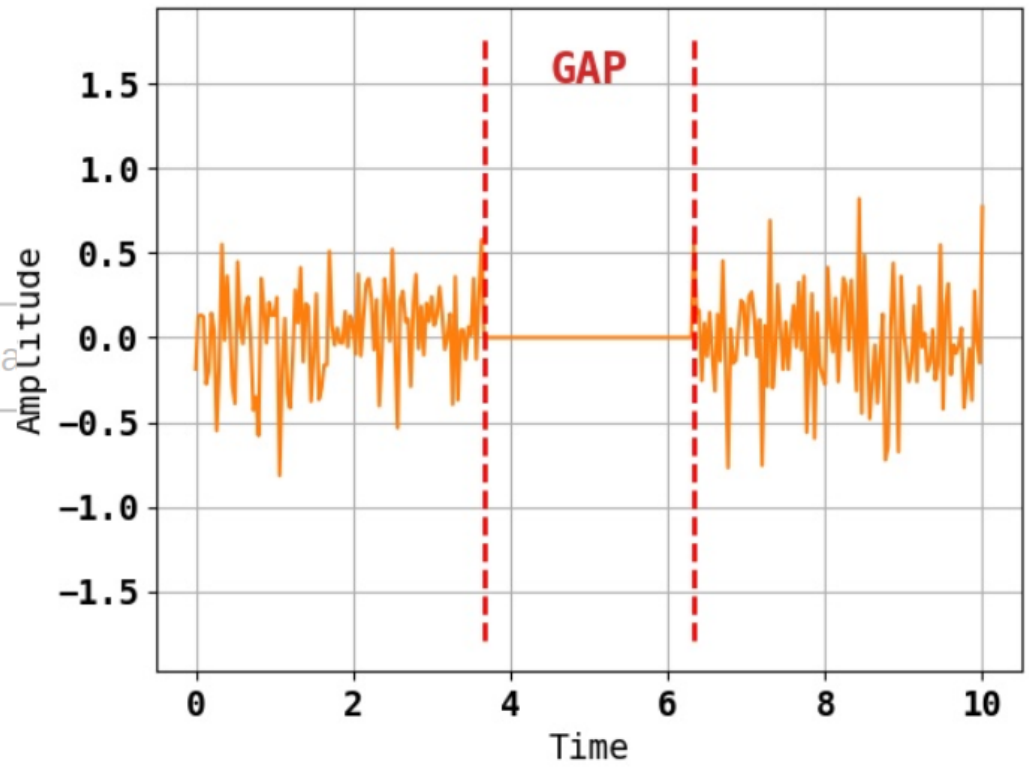
Optimization problem: algorithm.
Description of an iteration.



Optimization problem: algorithm.

Description of an iteration.

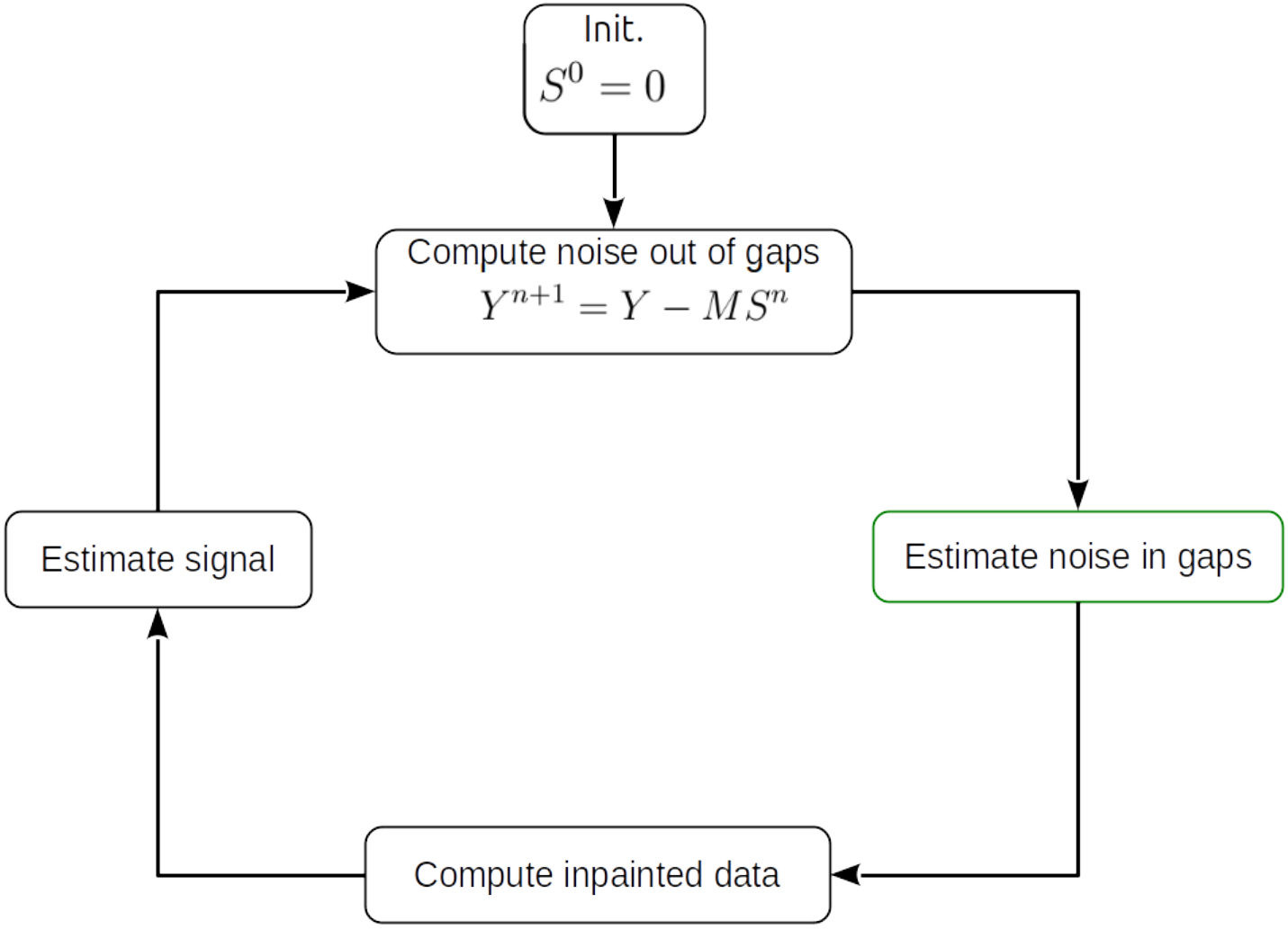
Init.
$S^0 = 0$

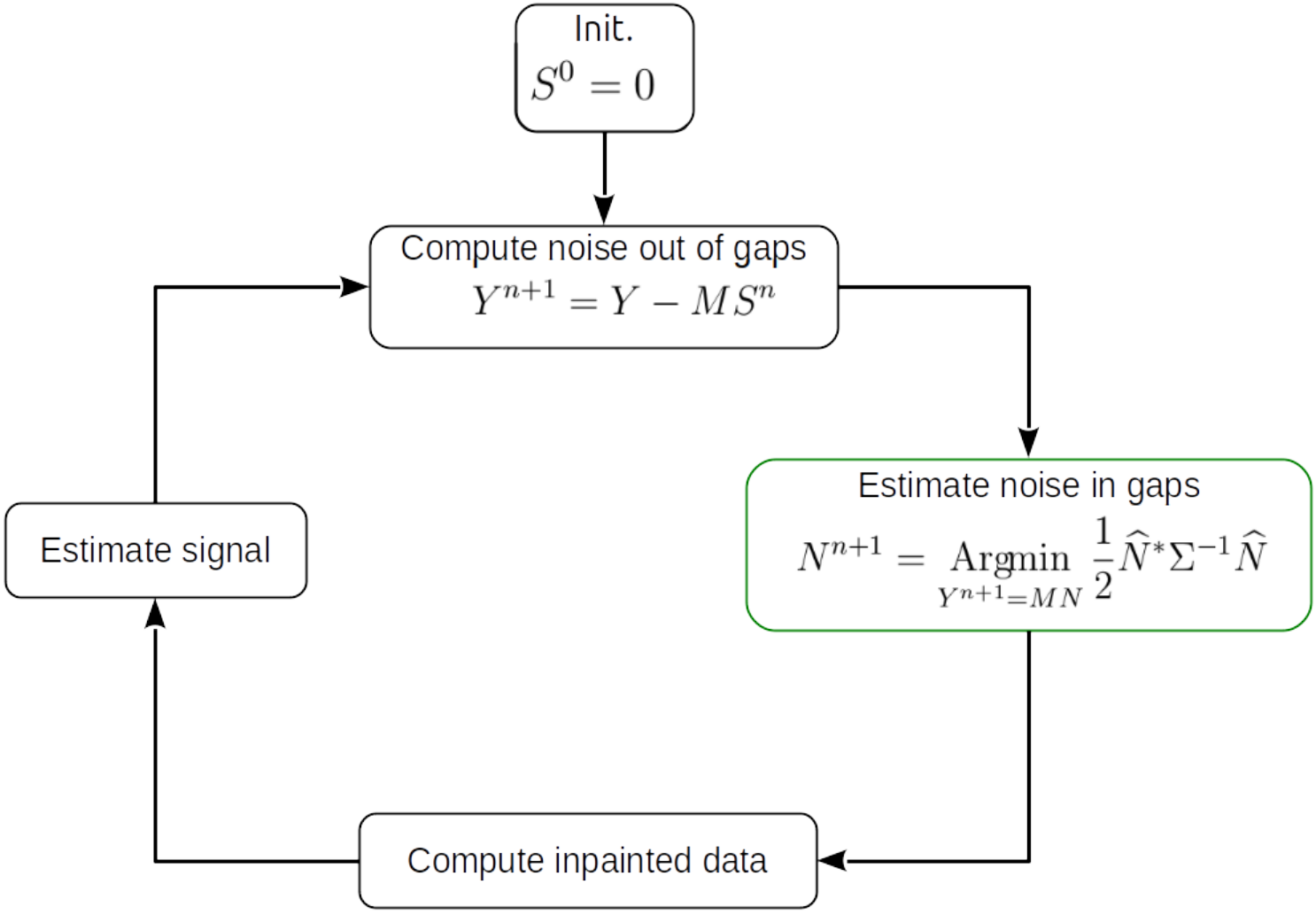


Estima

n gaps

Optimization problem: algorithm.
Description of an iteration.

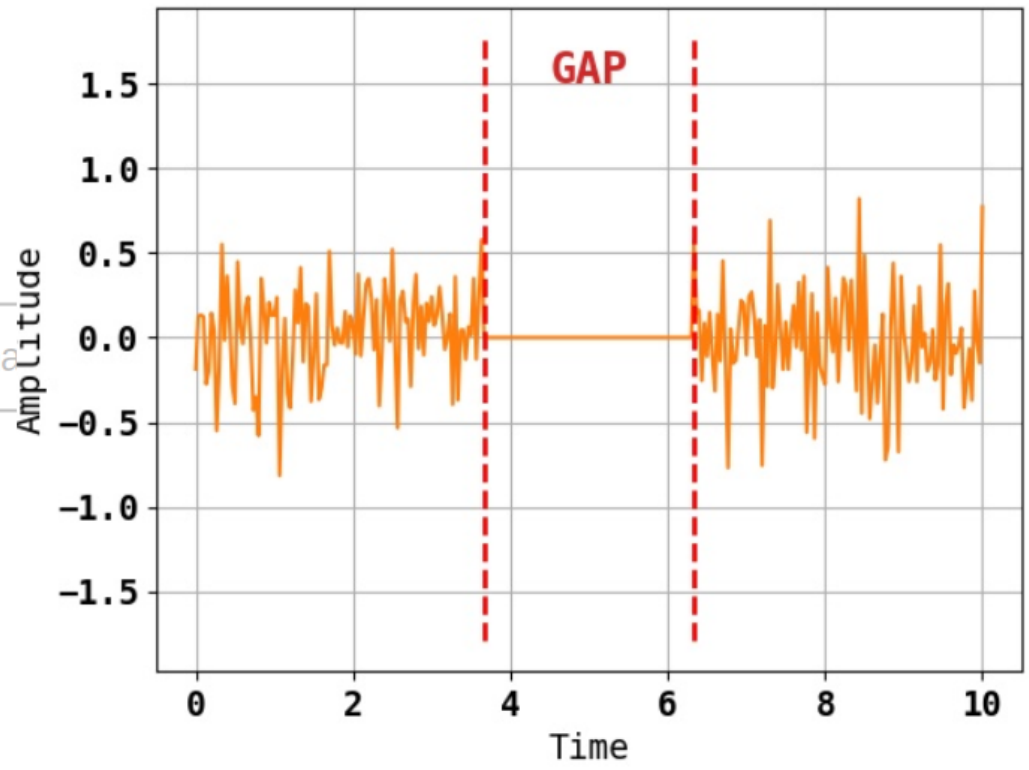




Optimization problem: algorithm.

Description of an iteration.

Init.
$S^0 = 0$



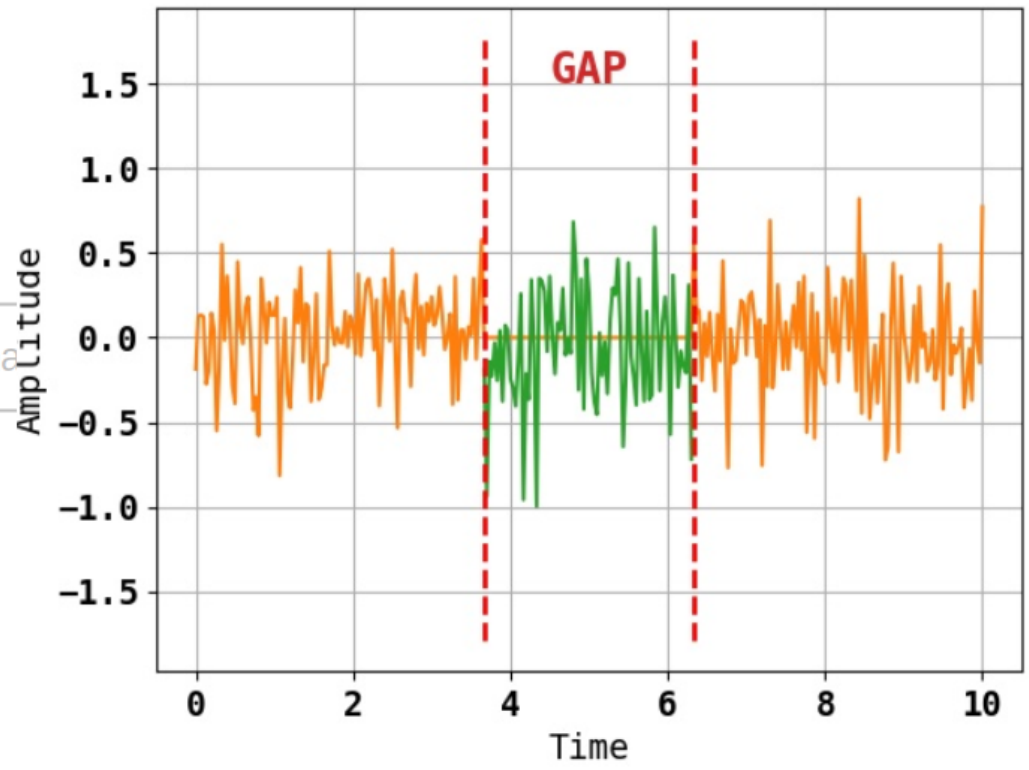
Estima

n gaps

Optimization problem: algorithm.

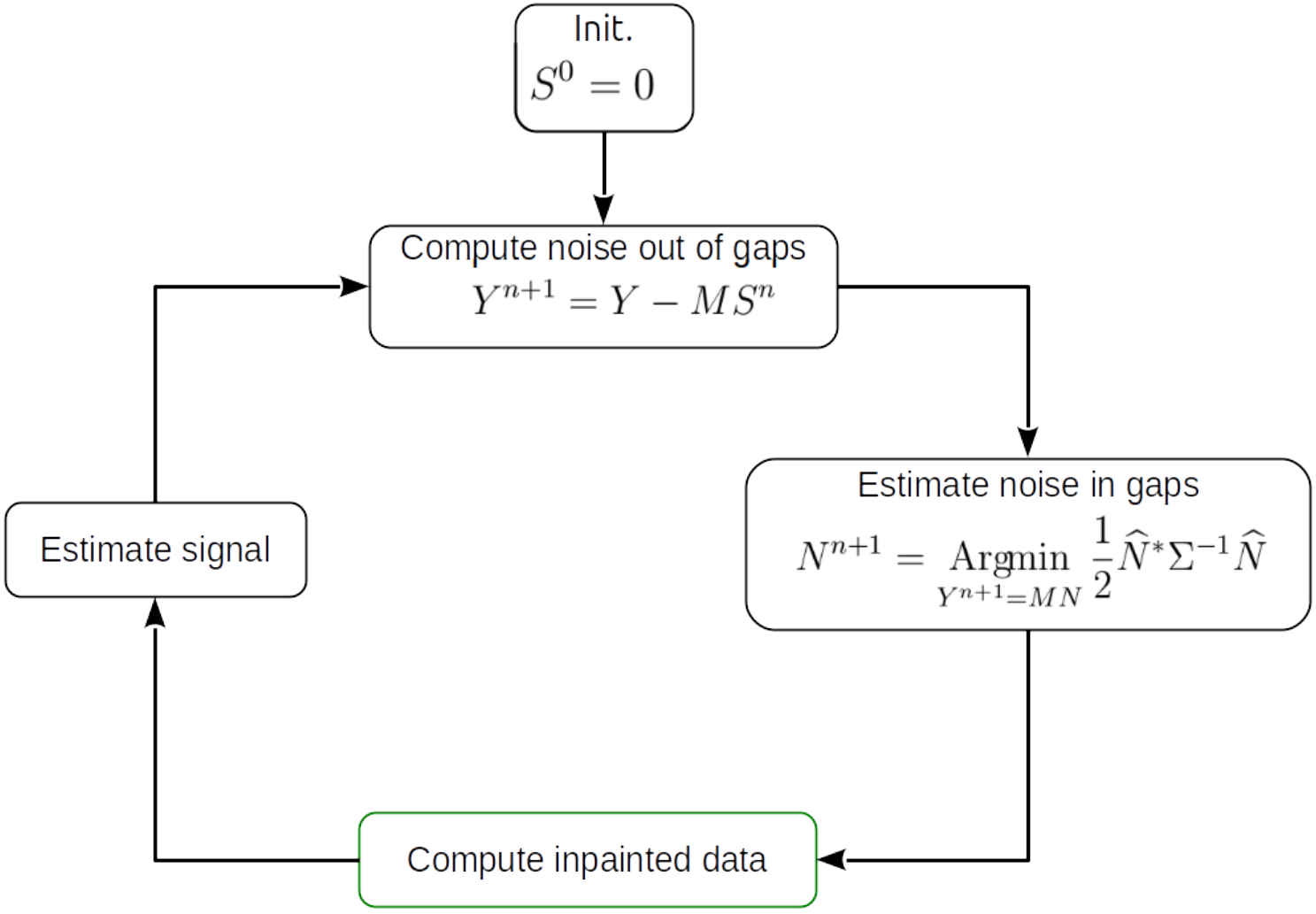
Description of an iteration.

Init.
$S^0 = 0$

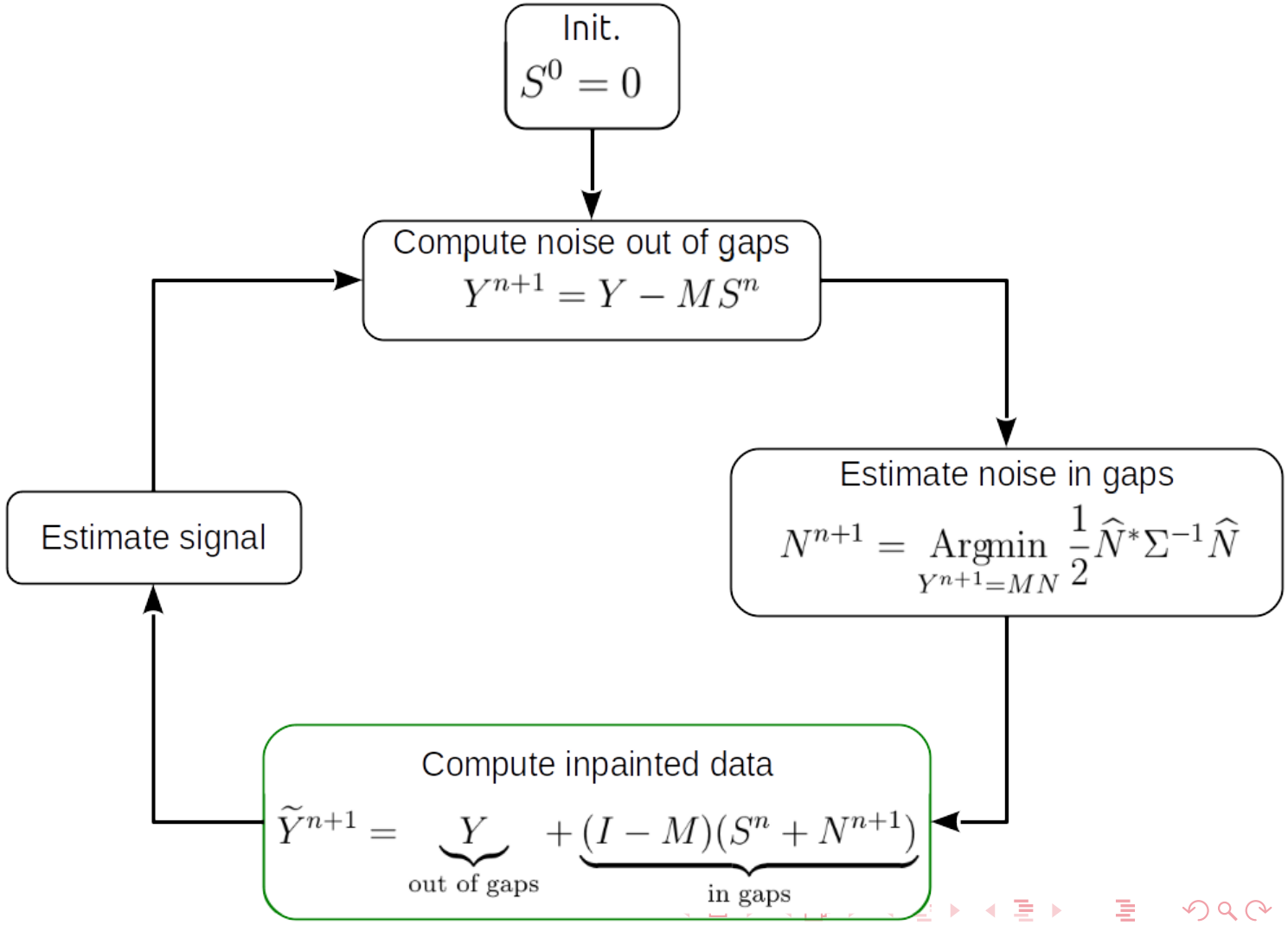


Estima

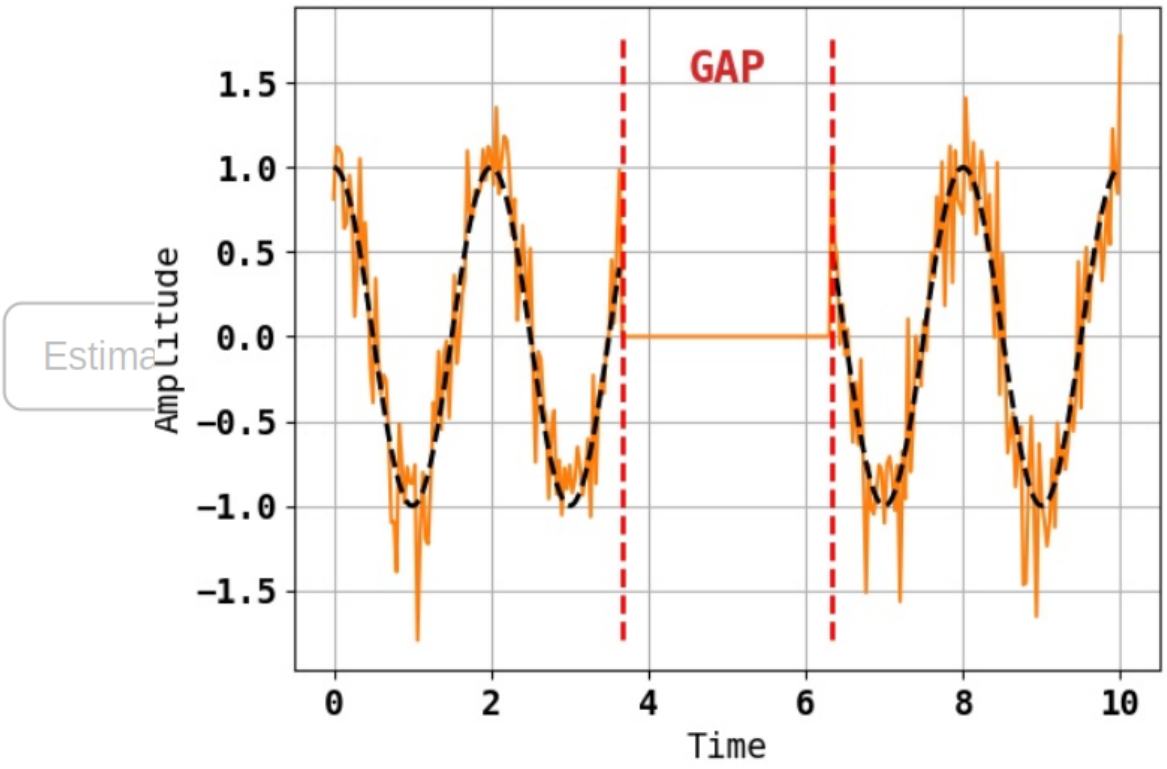
gaps
$\widehat{V}^* \Sigma^{-1} \widehat{N}$



Optimization problem: algorithm.
Description of an iteration.



Init.
$S^0 = 0$

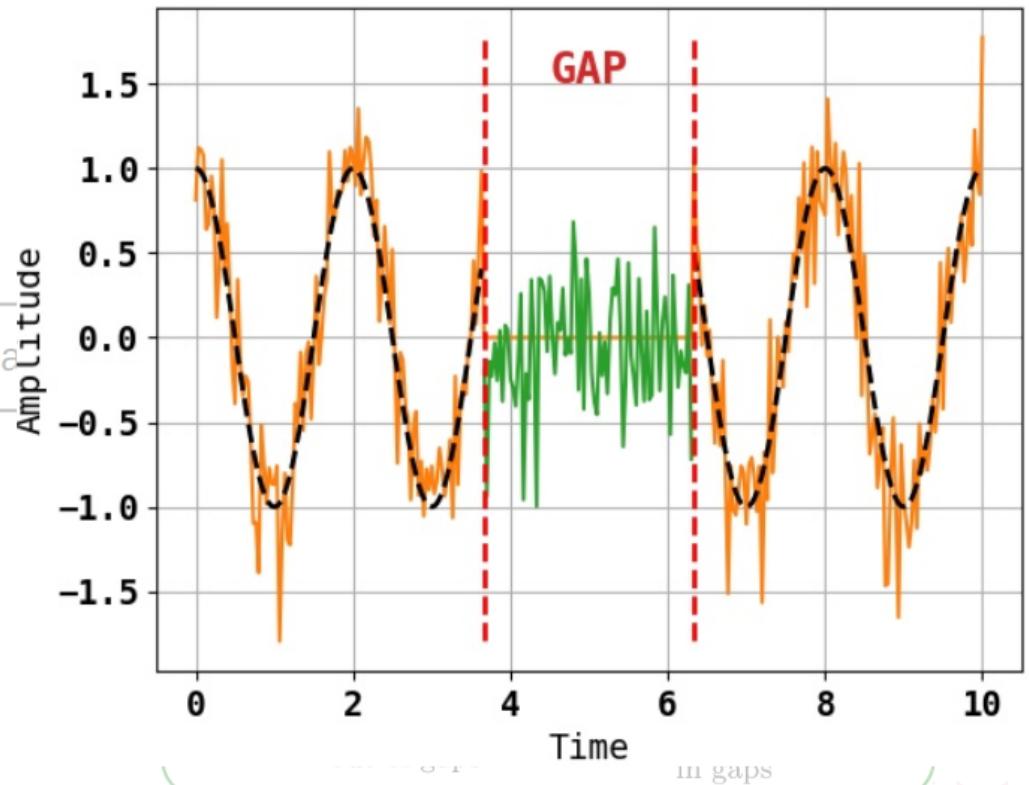


Estima

n gaps

Optimization problem: algorithm.
Description of an iteration.

Init.
$S^0 = 0$



Estima

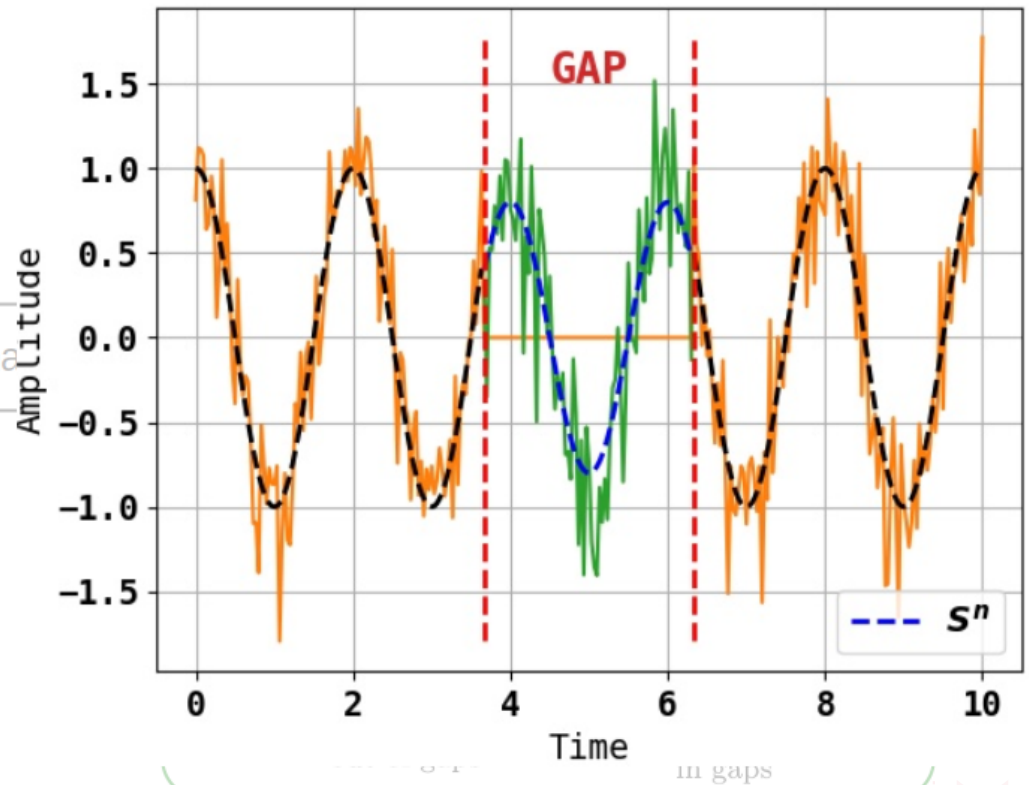
gaps
$\widehat{N}^* \Sigma^{-1} \widehat{N}$

in gaps

Optimization problem: algorithm.

Description of an iteration.

Init.
$S^0 = 0$

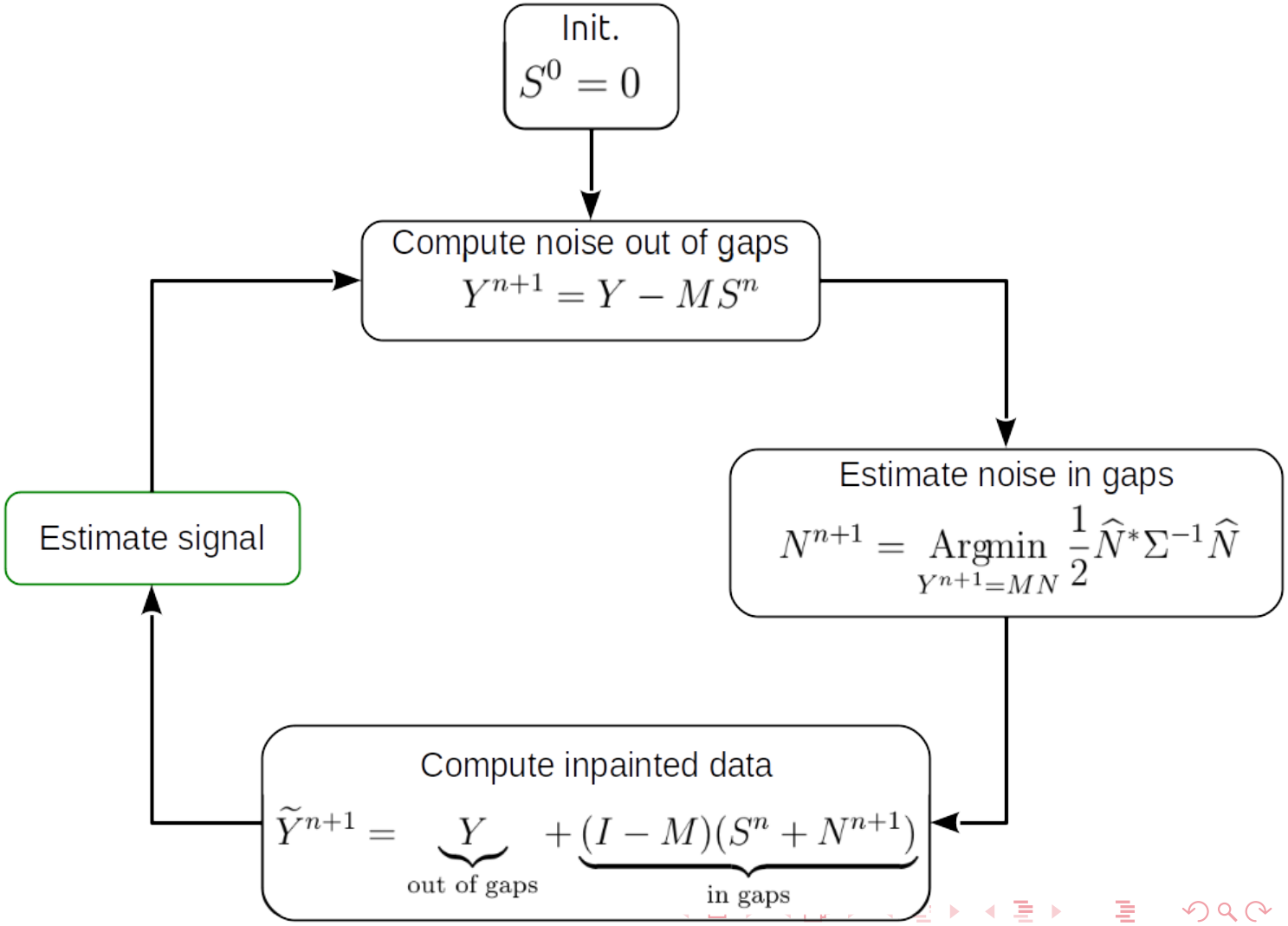


Estima

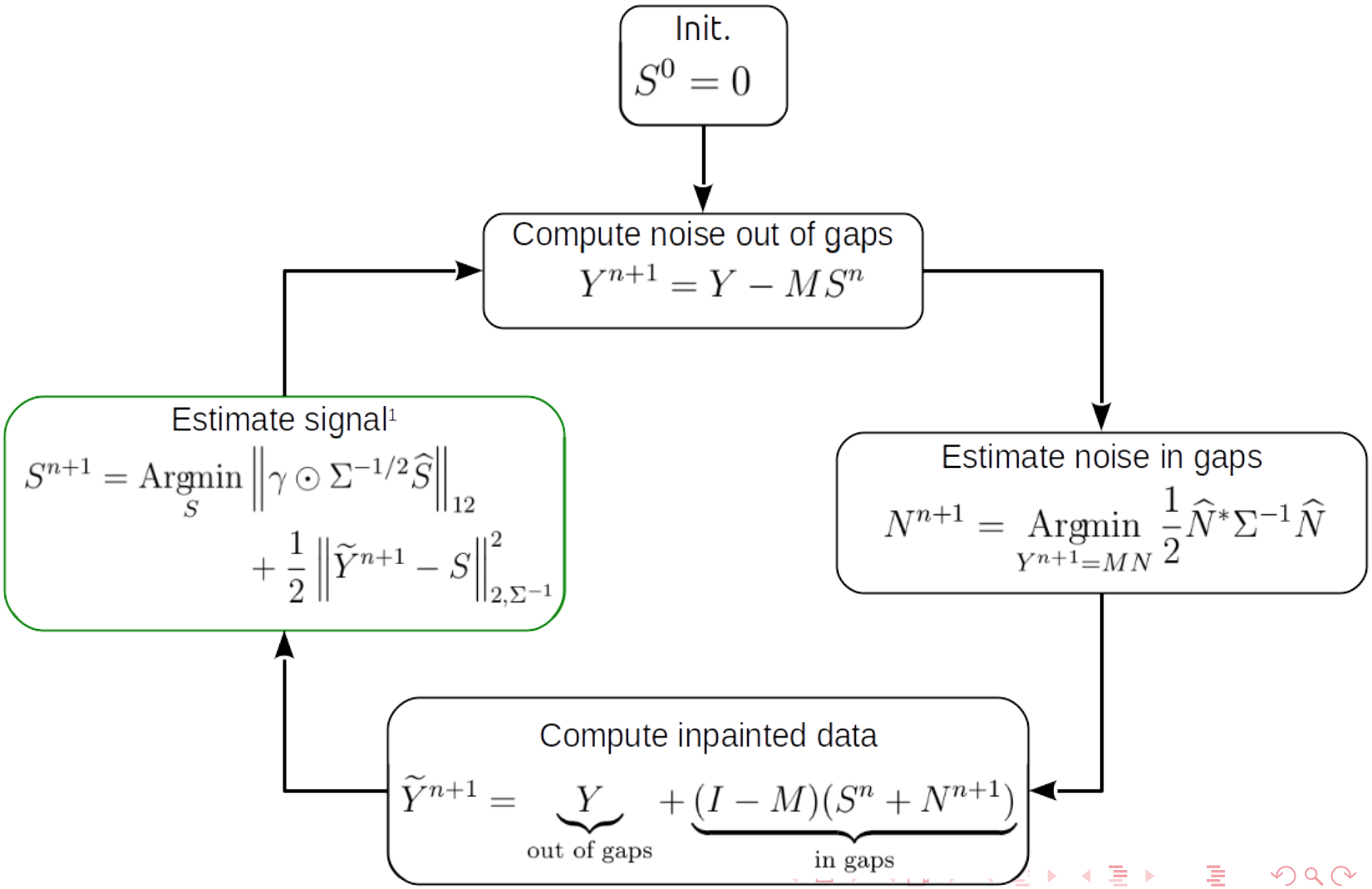
gaps
$\widehat{N}^*\Sigma^{-1}\widehat{N}$

in gaps

Optimization problem: algorithm.
Description of an iteration.



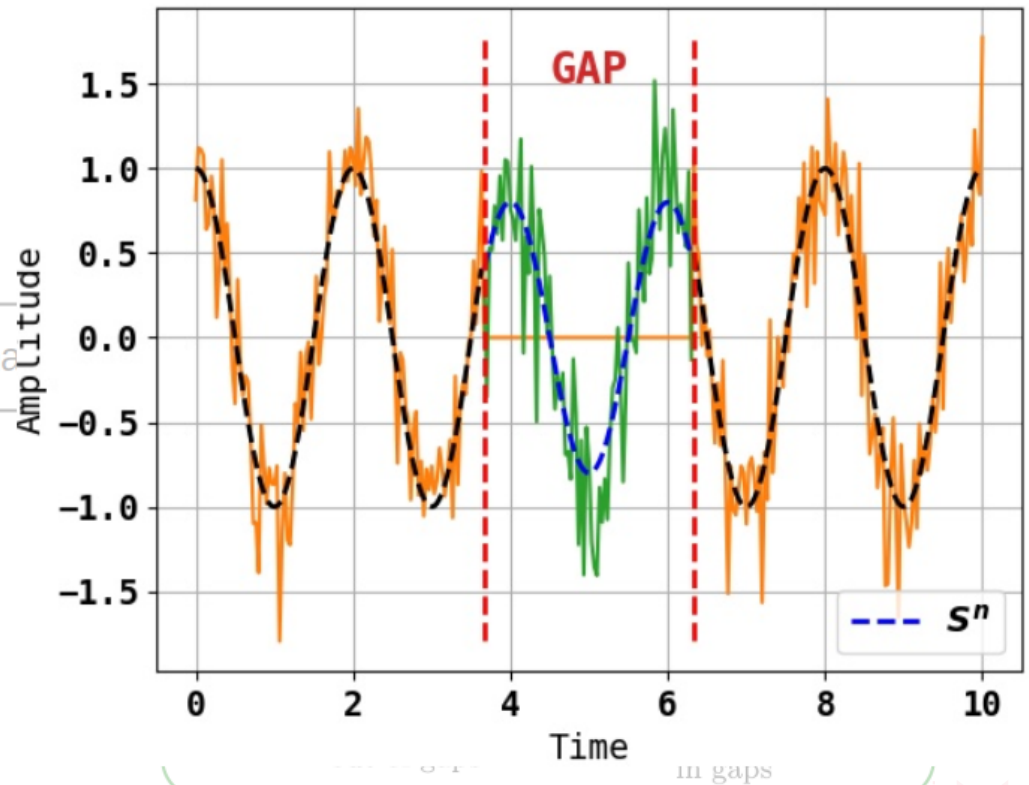
Optimization problem: algorithm.
Description of an iteration.



Optimization problem: algorithm.

Description of an iteration.

Init.
$S^0 = 0$



Estima

gaps
$\widehat{N}^*\Sigma^{-1}\widehat{N}$

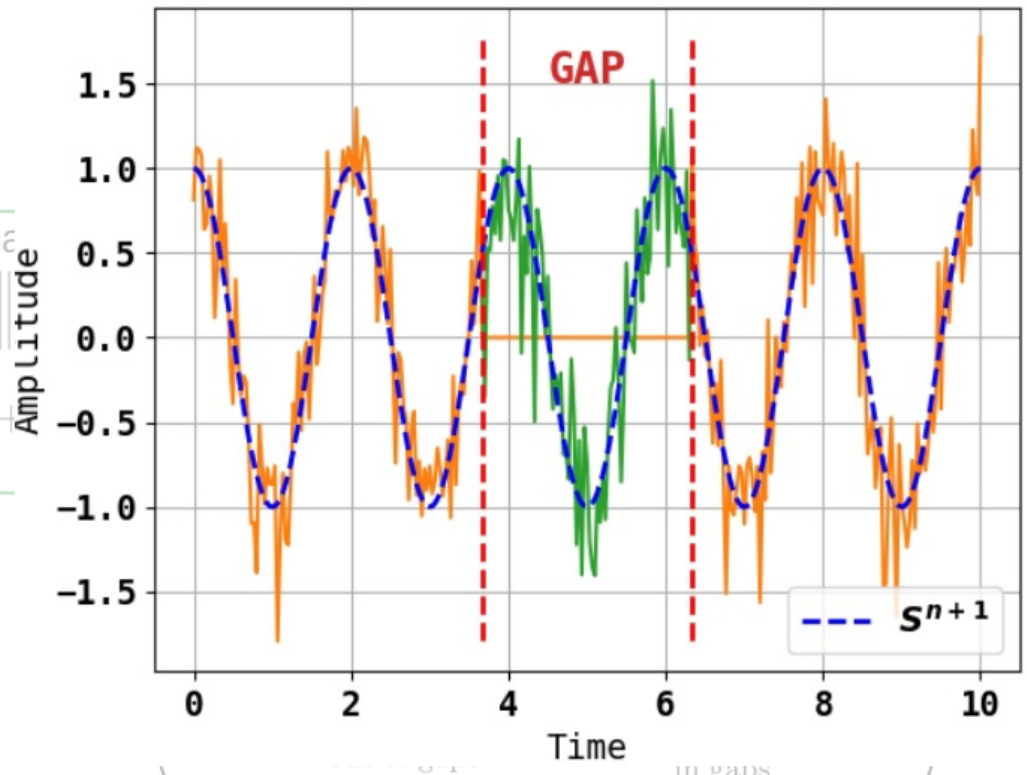
in gaps

Optimization problem: algorithm.
Description of an iteration.

Init.
$S^0 = 0$

Estima
$S^{n+1} = \underset{S}{\text{Argmin}}$

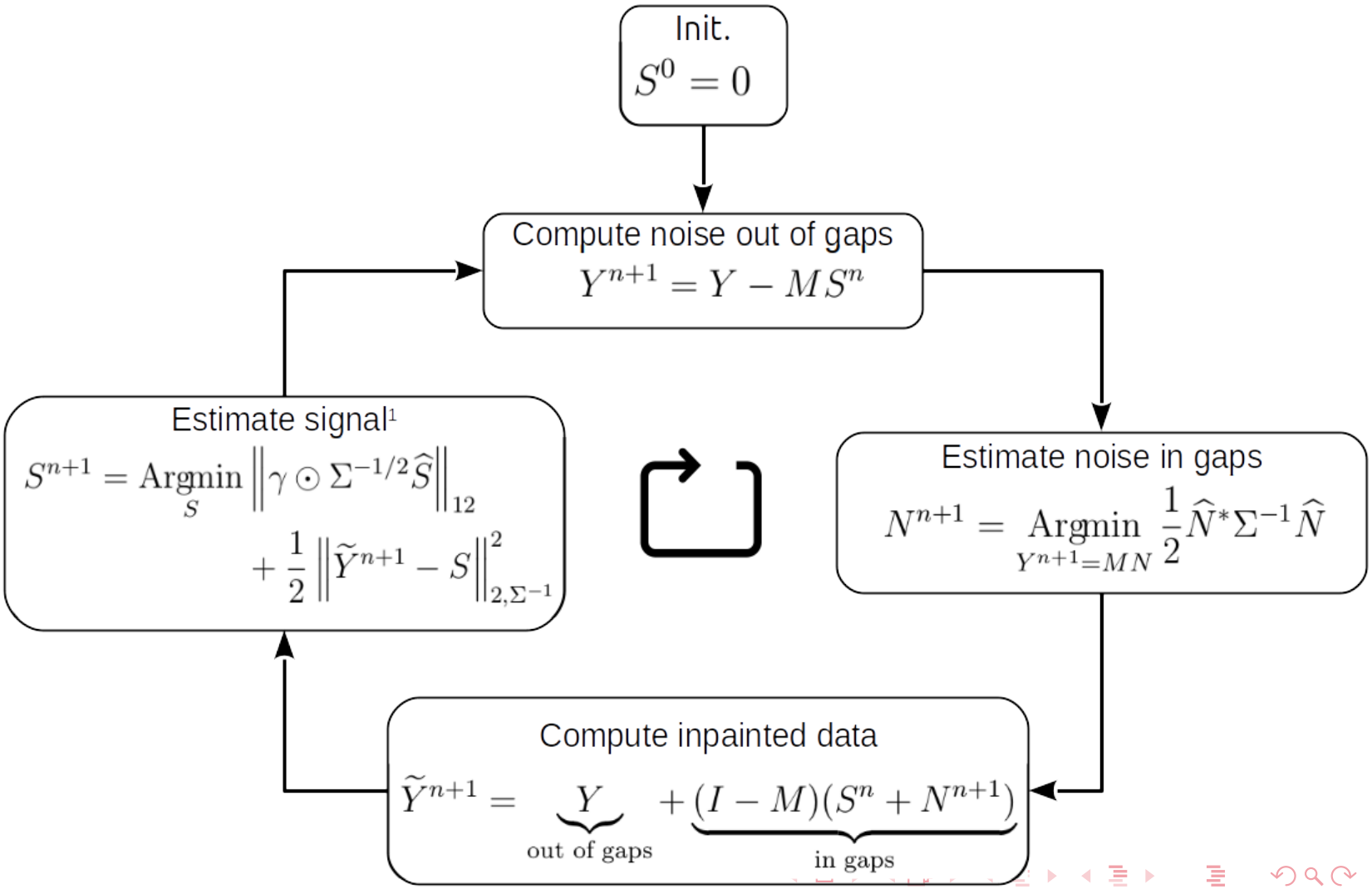
gaps
$\widehat{V}^* \Sigma^{-1} \widehat{N}$



in gaps

Optimization problem: algorithm.

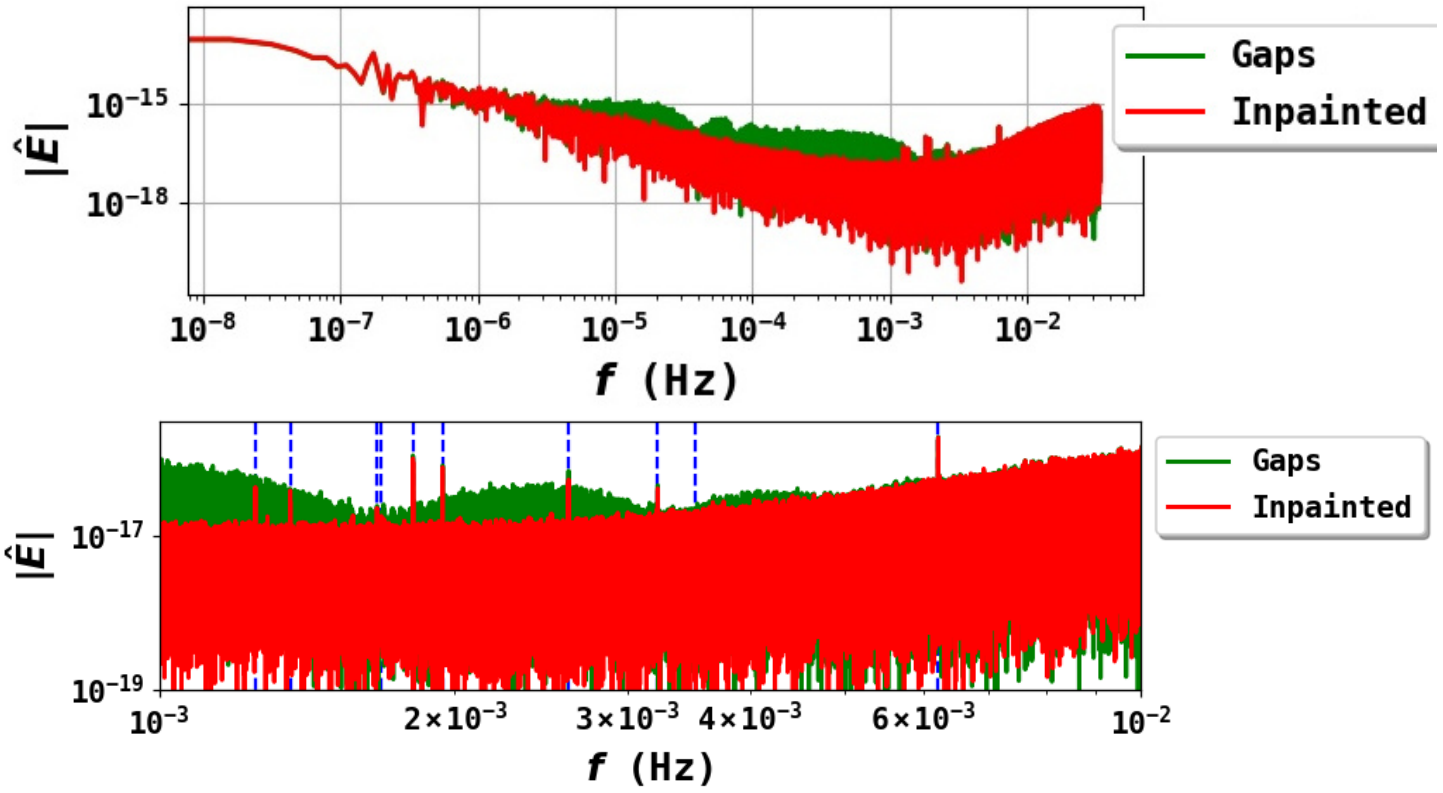
Description of an iteration.



Results.

Test on LDC1-3: verification galactic binaries.

- Gaussian noise.
- Gaps: 7h/2 weeks (planned), 10 min / 1 day (unplanned).
- **Flattened** effective noise power spectral density and **limited** leakage between signal and noise.



Conclusion

Conclusion.

Development and evaluation of a new non-parametric method.

- Tests with both **gapped** and **ungapped** data.
- **Efficient reconstruction** of the signal in the gaps.
- **Fast**: $\sim$ 1 hour for 2 years of data on a laptop.
- Works as well with a **large number of separated sources**.
- **Flexible framework** which could (in principle) be adapted to:
  - Estimation of power spectral density.
  - Other types of sources.
  - Sources of different types.

Commissariat à l'énergie atomique et aux énergies alternatives
Centre de Saclay | 91191 Gif-sur-Yvette Cedex
T. +33(0)1 69 08 73 88 | F. +33(0)1 69 08 75 84

DRF
Irfu
DPhN

Etablissement public à caractère industriel et commercial | R.C.S. Paris B 775 685 019