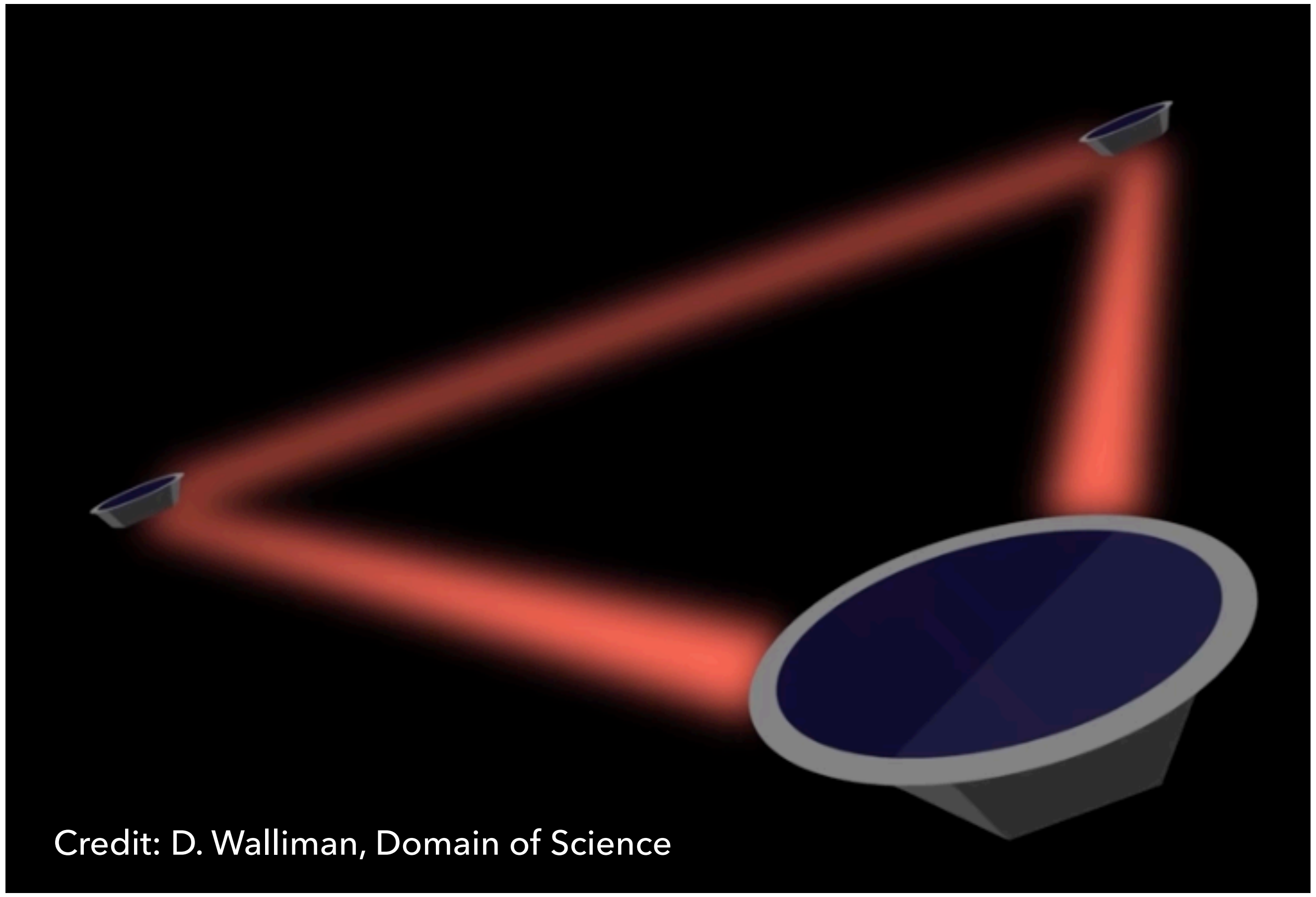
BAYESIAN INFERENCE FOR LISA DATA ANALYSIS

Quentin Baghi, CEA Saclay

Tuesday, November 15th, 2022

Rencontre du groupe de travail "méthodes d'analyse des données" du GdR oG

1. Challenges of LISA data analysis
2. Overview of used Bayesian concepts
3. Towards the future



CHALLENGES OF LISA DATA ANALYSIS

- The analysis of LISA data will be drastically **different from current ground-based detection**:
  - ✦ Numerous superimposed sources $\neq$ isolated events
  - ✦ Different time scales, larger waveform cycles observed
  - ✦ Signal-dominated measurement $\neq$ noise-dominated

- The analysis of LISA data will be drastically **different from current ground-based detection**:
  - ✦ Numerous superimposed sources ≠ isolated events
  - ✦ Different time scales, larger waveform cycles observed
  - ✦ Signal-dominated measurement ≠ noise-dominated

Research problem

- The analysis of LISA data will be drastically **different from current ground-based detection**:
  - ✦ Numerous superimposed sources ≠ isolated events
  - ✦ Different time scales, larger waveform cycles observed
  - ✦ Signal-dominated measurement ≠ noise-dominated

  Research problem

- Additional difficulties, similar to ground-based detection:
  - ✦ Stochastic noise
  - ✦ Instrumental transients (glitches)
  - ✦ Non-stationarities
  - ✦ Spectral lines
  - ✦ Data gaps

- The analysis of LISA data will be drastically **different from current ground-based detection**:
  - ✦ Numerous superimposed sources ≠ isolated events
  - ✦ Different time scales, larger waveform cycles observed
  - ✦ Signal-dominated measurement ≠ noise-dominated

  Research problem

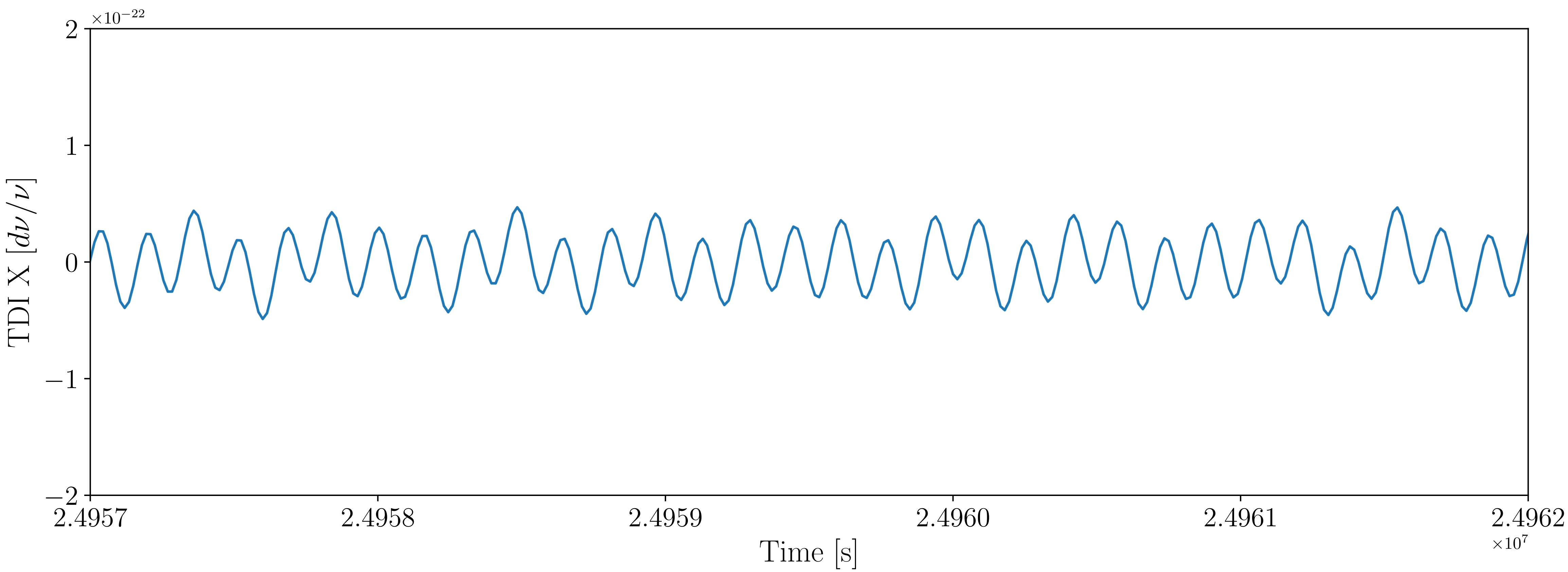
- Additional difficulties, similar to ground-based detection:
  - ✦ Stochastic noise
  - ✦ Instrumental transients (glitches)
  - ✦ Non-stationarities
  - ✦ Spectral lines
  - ✦ Data gaps

  Disturbances

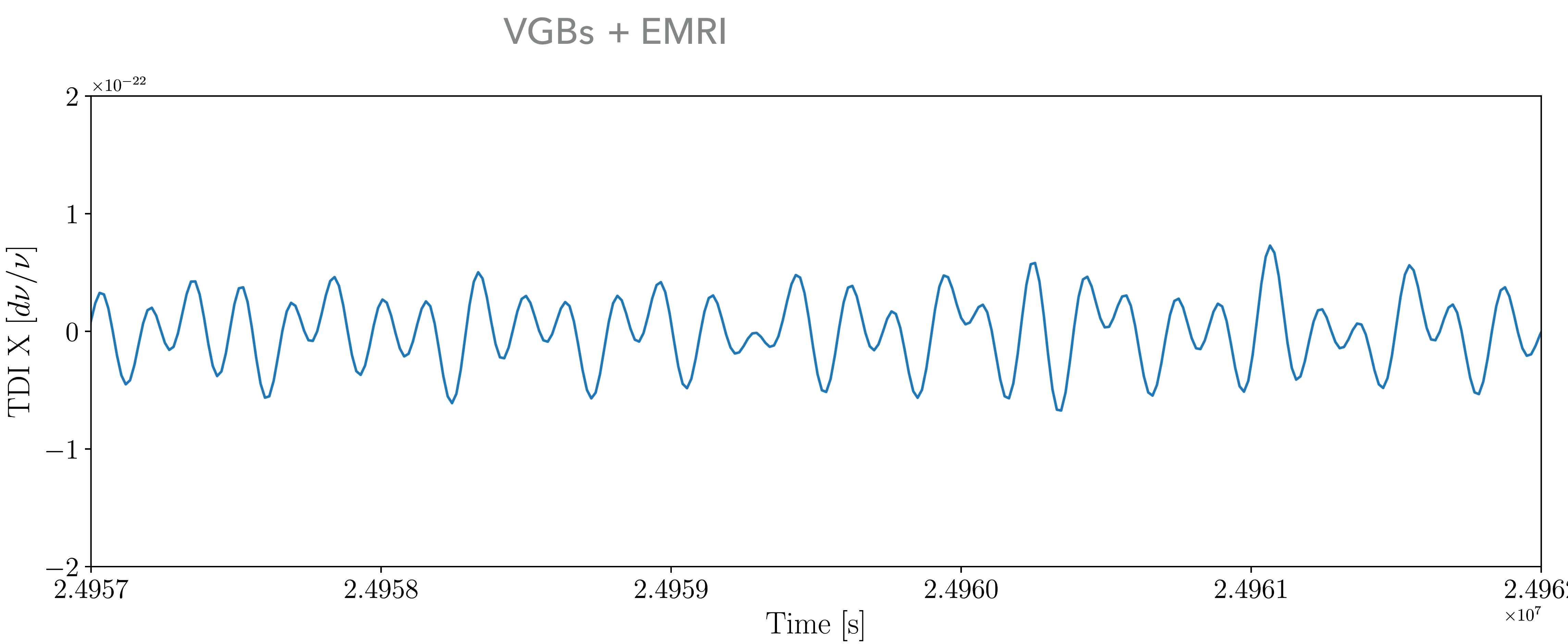
- What kind of data will LISA measure?
  - ✦ Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - ✦ Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - ✦ Dominated by laser frequency noise
  - ✦ After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)

- What kind of data will LISA measure?
  - ✦ Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - ✦ Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - ✦ Dominated by laser frequency noise
  - ✦ After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)

VGBs

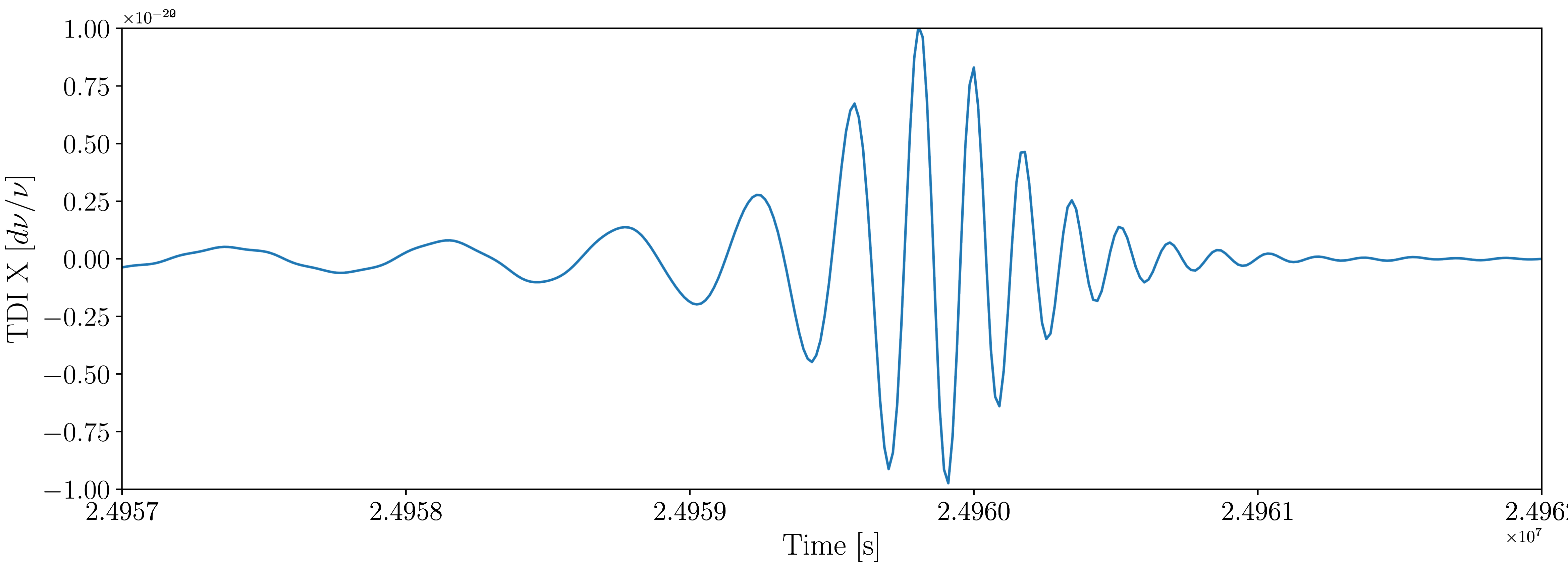


- What kind of data will LISA measure?
  - ✦ Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - ✦ Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - ✦ Dominated by laser frequency noise
  - ✦ After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)



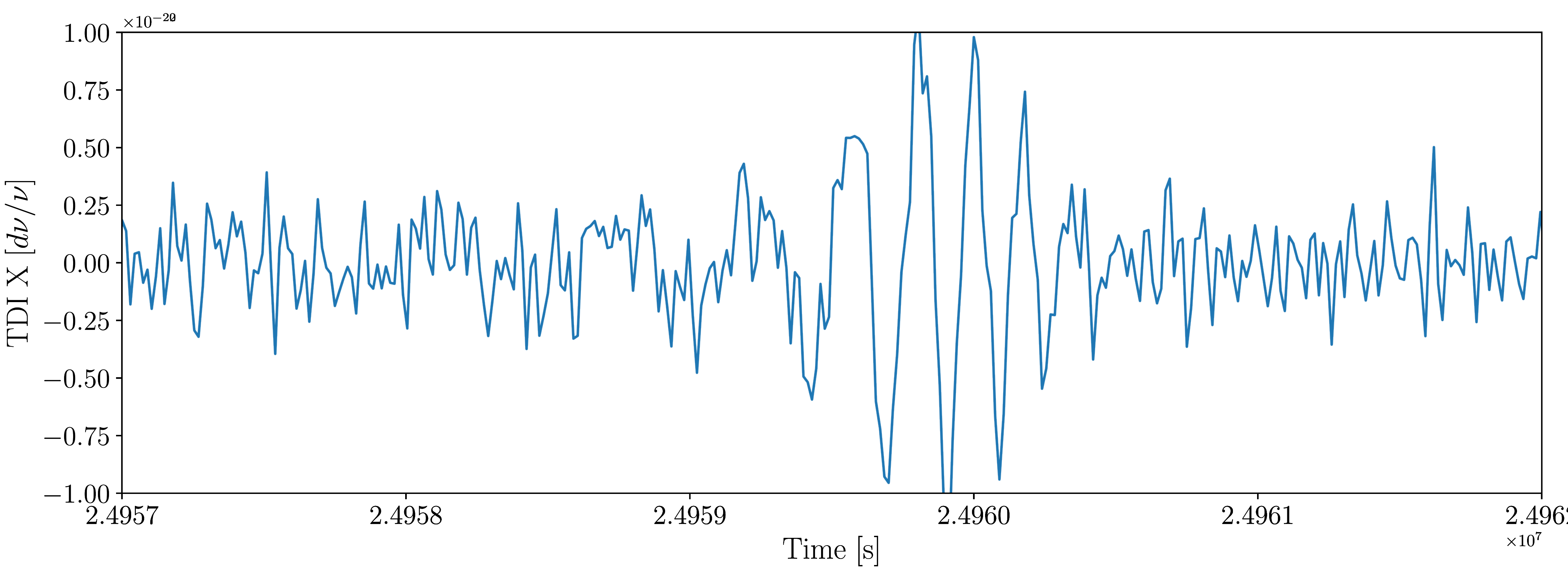
- What kind of data will LISA measure?
  - ✦ Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - ✦ Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - ✦ Dominated by laser frequency noise
  - ✦ After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)

VGBs + EMRI + MBHB



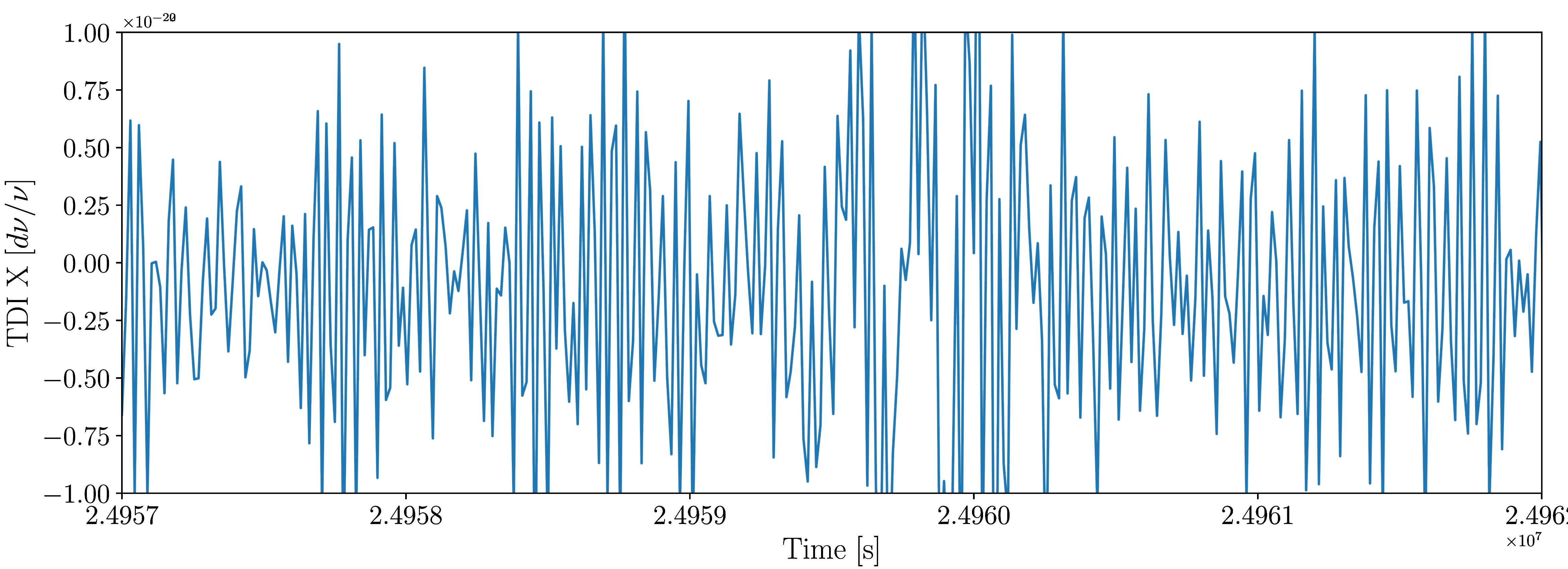
- What kind of data will LISA measure?
  - ✦ Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - ✦ Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - ✦ Dominated by laser frequency noise
  - ✦ After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)

VGBs + EMRI + MBHB + Galaxy



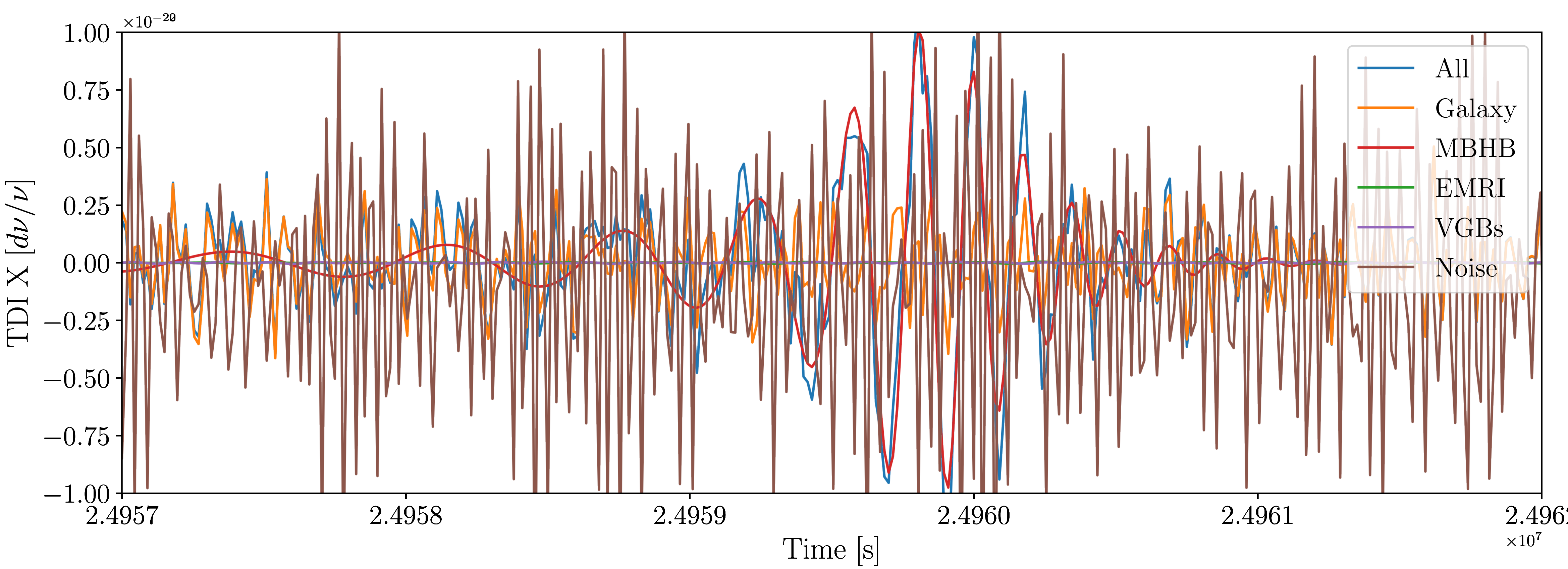
- What kind of data will LISA measure?
  - Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - Dominated by laser frequency noise
  - After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)

VGBs + EMRI + MBHB + Galaxy + noise



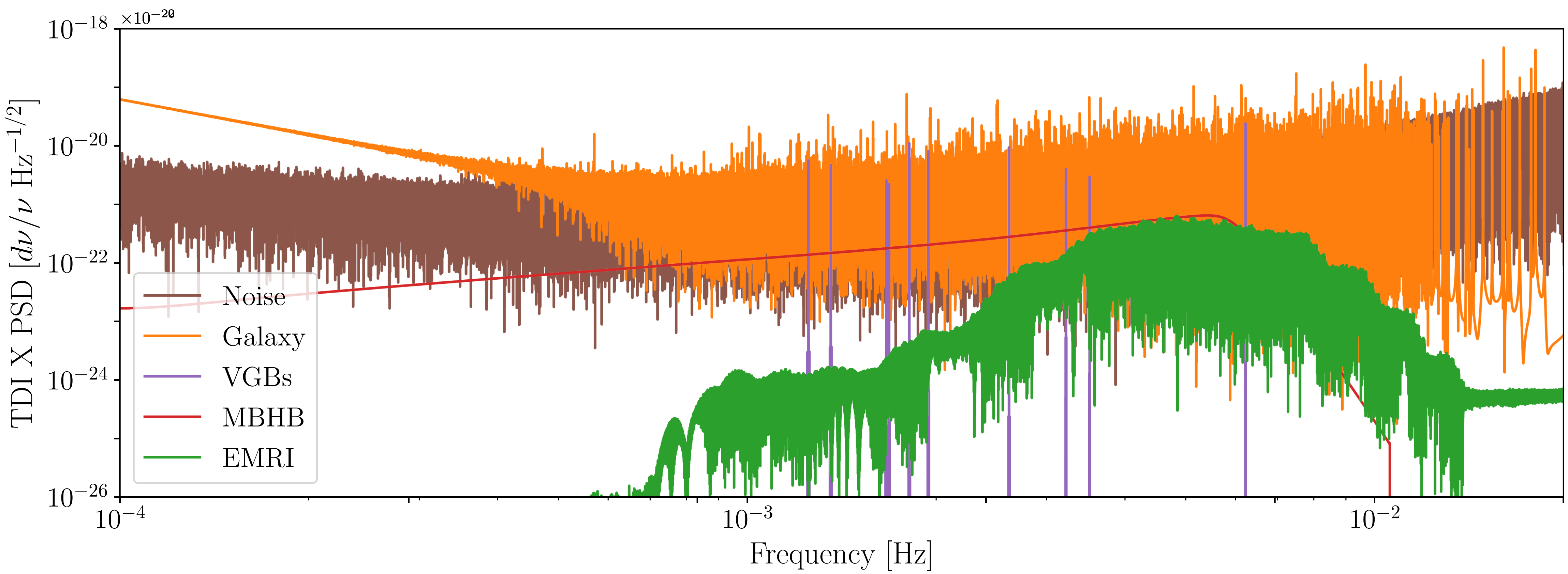
- What kind of data will LISA measure?
  - ✦ Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - ✦ Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - ✦ Dominated by laser frequency noise
  - ✦ After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)

VGBs + EMRI + MBHB + Galaxy + noise



- What kind of data will LISA measure?
  - ✦ Fractional frequency deviations (relative doppler shits) from 27 interferometers
  - ✦ Times series sampled at 4 Hz, observed over 4+ years with 89% duty cycle
  - ✦ Dominated by laser frequency noise
  - ✦ After pre-processing, obtain 3 time-delay interferometry (TDI) data streams (X, Y, Z)

VGBs + EMRI + MBHB + Galaxy + noise



- What is the strategy to analyse the data?

  ✦ Bayesian framework: probe the parameters + number of model components posterior

$$p\left(\boldsymbol{\theta}, k \,|\, \boldsymbol{d}\right) = \frac{p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) p(\boldsymbol{\theta}, k)}{p(\boldsymbol{d})}$$

- What is the strategy to analyse the data?
  - ✦ Bayesian framework: probe the parameters + number of model components posterior

$$p\left(\boldsymbol{\theta}, k \,|\, \boldsymbol{d}\right) = \frac{p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) p(\boldsymbol{\theta}, k)}{p(\boldsymbol{d})}$$

Model parameters

- What is the strategy to analyse the data?

  ✦ Bayesian framework: probe the parameters + number of model components posterior

$$p\left(\boldsymbol{\theta}, k \,|\, \boldsymbol{d}\right) = \frac{p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) p(\boldsymbol{\theta}, k)}{p(\boldsymbol{d})}$$

Model parameters

Number of model components

- What is the strategy to analyse the data?
  - ✦ Bayesian framework: probe the parameters + number of model components posterior

$$p\left(\boldsymbol{\theta}, k \,|\, \boldsymbol{d}\right) = \frac{p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) p(\boldsymbol{\theta}, k)}{p(\boldsymbol{d})}$$

Model parameters

Number of model components

Data vector. For example d=(X, Y, Z)

- What is the strategy to analyse the data?

✦ Bayesian framework: probe the parameters + number of model components posterior

$$p\left(\boldsymbol{\theta}, k \,|\, \boldsymbol{d}\right) = \frac{p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) p(\boldsymbol{\theta}, k)}{p(\boldsymbol{d})}$$

Model parameters

Number of model components

Data vector. For example d=(X, Y, Z)

✦ Define a likelihood function: e.g. Gaussian

$$p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}(\boldsymbol{\theta})|}} \exp\left\{ \left(\boldsymbol{d} - \boldsymbol{h}(\boldsymbol{\theta}, k)\right)^\dagger \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \left(\boldsymbol{d} - \boldsymbol{h}(\boldsymbol{\theta}, k)\right) \right\}$$

- What is the strategy to analyse the data?

  ✦ Bayesian framework: probe the parameters + number of model components posterior

$$p\left(\boldsymbol{\theta}, k \,|\, \boldsymbol{d}\right) = \frac{p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) p(\boldsymbol{\theta}, k)}{p(\boldsymbol{d})}$$

Model parameters

Number of model components

Data vector. For example d=(X, Y, Z)

  ✦ Define a likelihood function: e.g. Gaussian

$$p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}(\boldsymbol{\theta})|}} \exp\left\{ \left(\boldsymbol{d} - \boldsymbol{h}(\boldsymbol{\theta}, k)\right)^{\dagger} \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \left(\boldsymbol{d} - \boldsymbol{h}(\boldsymbol{\theta}, k)\right) \right\}$$

GW signals: $\boldsymbol{h}(\boldsymbol{\theta}, k) = \sum_{j=1}^{k} \boldsymbol{h}_j(\boldsymbol{\theta}_j)$

- What is the strategy to analyse the data?

  - Bayesian framework: probe the parameters + number of model components posterior

$$p\left(\boldsymbol{\theta}, k \,|\, \boldsymbol{d}\right) = \frac{p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) p(\boldsymbol{\theta}, k)}{p(\boldsymbol{d})}$$

Model parameters

Number of model components

Data vector. For example d=(X, Y, Z)

  - Define a likelihood function: e.g. Gaussian

$$p\left(\boldsymbol{d} \,|\, \boldsymbol{\theta}, k\right) = \frac{1}{\sqrt{(2\pi)^N |\boldsymbol{\Sigma}(\boldsymbol{\theta})|}} \exp\left\{ \left(\boldsymbol{d} - \boldsymbol{h}(\boldsymbol{\theta}, k)\right)^\dagger \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \left(\boldsymbol{d} - \boldsymbol{h}(\boldsymbol{\theta}, k)\right) \right\}$$

GW signals: $\boldsymbol{h}(\boldsymbol{\theta}, k) = \sum_{j=1}^{k} \boldsymbol{h}_j(\boldsymbol{\theta}_j)$

Stochastic processes: $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \sum_{i=1}^{p} \boldsymbol{\Sigma}_i(\boldsymbol{\theta}_i)$

- In practice, we need to **partition the parameter space** in several sub-problems. Example of the **blocked Gibbs** scheme:

- In practice, we need to **partition the parameter space** in several sub-problems. Example of the **blocked Gibbs** scheme:

1. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{GB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

- In practice, we need to **partition the parameter space** in several sub-problems. Example of the **blocked Gibbs** scheme:

1. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{GB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

$\downarrow$

2. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{MBHB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

- In practice, we need to **partition the parameter space** in several sub-problems. Example of the **blocked Gibbs** scheme:

1. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{GB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$
2. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{MBHB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$
3. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{EMRI}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

- In practice, we need to **partition the parameter space** in several sub-problems. Example of the **blocked Gibbs** scheme:

1. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{GB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$
2. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{MBHB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$
3. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{EMRI}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

...

K. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{noise}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

- In practice, we need to **partition the parameter space** in several sub-problems. Example of the **blocked Gibbs** scheme:

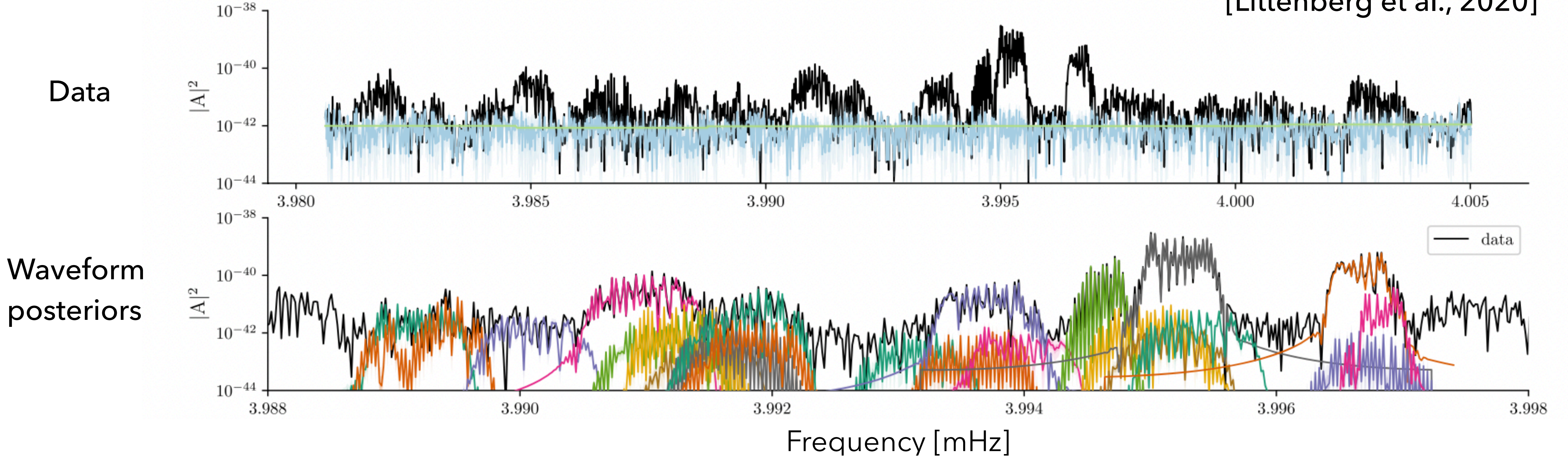
1. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{GB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$
2. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{MBHB}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$
3. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{EMRI}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

...

K. Sample for $p\left(\boldsymbol{\theta}_{\mathrm{noise}} \middle| \boldsymbol{y}, \boldsymbol{\theta}_{\mathrm{others}}\right)$

- The number of overlapping sources (especially Galactic binaries) is not know in advance
- Need to estimate the optimal number of sources

[Littenberg et al., 2020]



- Algorithm: reverse-jump Markov-chain Monte Carlo (RJMCMC)
- Allow for parallel computing by splitting the frequency-domain data into segments

- In the case of missing data points or gaps
- Example: interrupted science data due to antenna repointing
- Consequence: both the signal and the covariance become expensive to compute
- One strategy is data augmentation [Baghi et al, 2019]

$$\boldsymbol{d} = (\boldsymbol{d}_o, \boldsymbol{d}_m)$$

- In the case of missing data points or gaps
- Example: interrupted science data due to antenna repointing
- Consequence: both the signal and the covariance become expensive to compute
- One strategy is data augmentation [Baghi et al, 2019]

$$\boldsymbol{d} = (\boldsymbol{d}_o, \boldsymbol{d}_m)$$

P Sample for $p(\boldsymbol{\theta} \,|\, \boldsymbol{d}_o, \boldsymbol{d}_m)$

BAYESIAN CONCEPTS: DATA AUGMENTATION

- In the case of missing data points or gaps
- Example: interrupted science data due to antenna repointing
- Consequence: both the signal and the covariance become expensive to compute
- One strategy is data augmentation [Baghi et al, 2019]

$$\boldsymbol{d} = (\boldsymbol{d}_o, \boldsymbol{d}_m)$$

P Sample for $p(\boldsymbol{\theta} \,|\, \boldsymbol{d}_o, \boldsymbol{d}_m)$

$\downarrow$

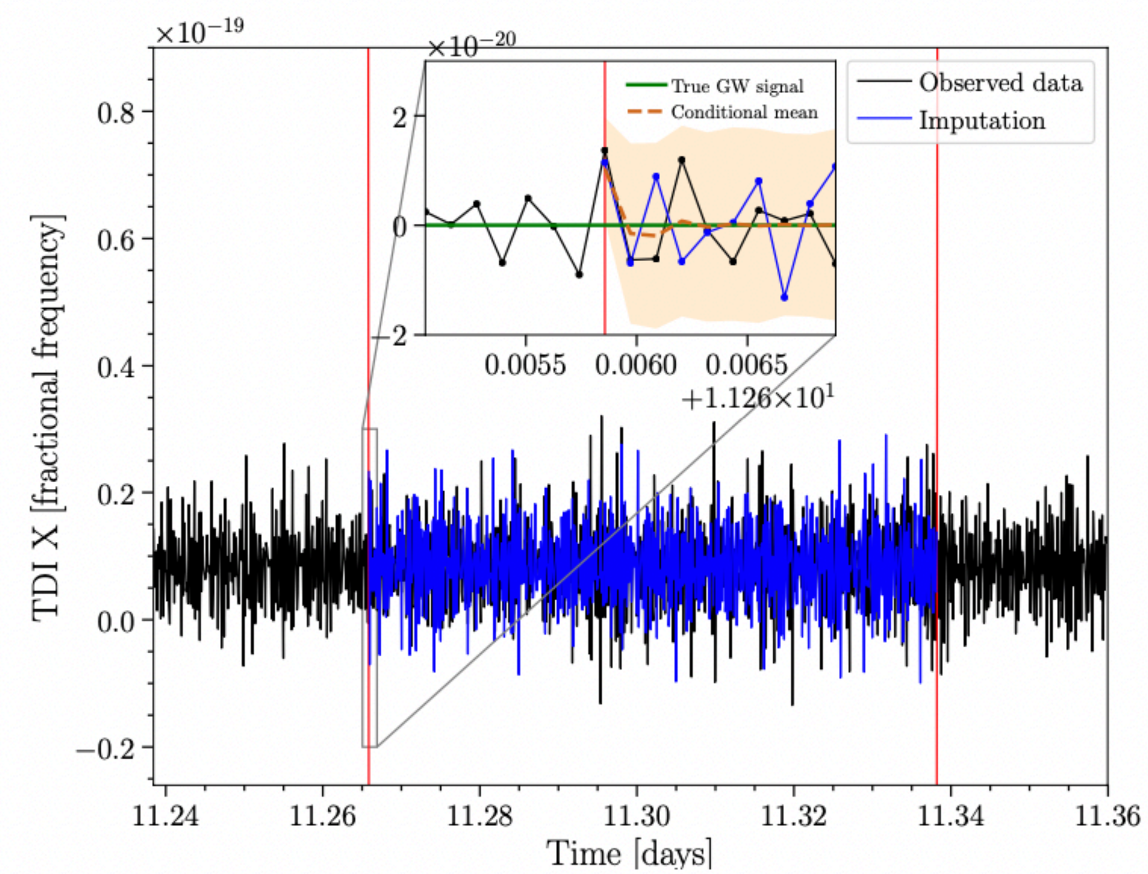
I Sample for $p(\boldsymbol{d}_m \,|\, \boldsymbol{d}_o, \boldsymbol{\theta})$

- In the case of missing data points or gaps
- Example: interrupted science data due to antenna repointing
- Consequence: both the signal and the covariance become expensive to compute
- One strategy is data augmentation [Baghi et al, 2019]

$$\boldsymbol{d} = (\boldsymbol{d}_o, \boldsymbol{d}_m)$$

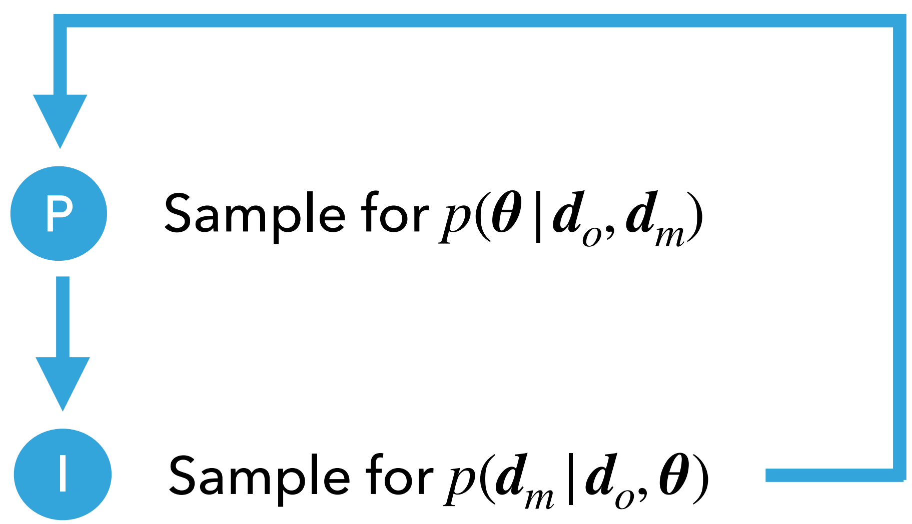
P Sample for $p(\boldsymbol{\theta} \,|\, \boldsymbol{d}_o, \boldsymbol{d}_m)$

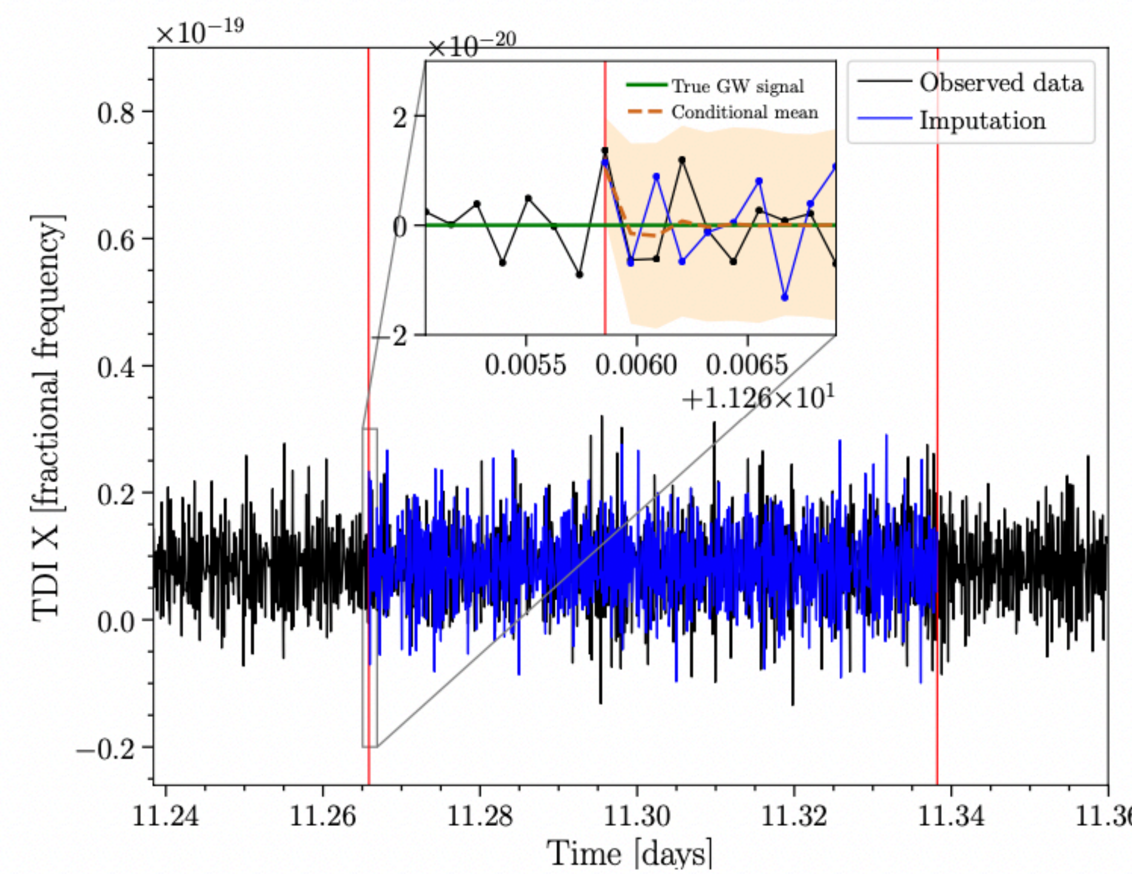
I Sample for $p(\boldsymbol{d}_m \,|\, \boldsymbol{d}_o, \boldsymbol{\theta})$



- In the case of missing data points or gaps
- Example: interrupted science data due to antenna repointing
- Consequence: both the signal and the covariance become expensive to compute
- One strategy is data augmentation [Baghi et al, 2019]

$$\boldsymbol{d} = (\boldsymbol{d}_o, \boldsymbol{d}_m)$$





- Independent component analysis to separate GW stochastic backgrounds from noise

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

$$\tilde{d} = T_h \tilde{h} + T_n \tilde{n}$$

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data $\longrightarrow$ $\tilde{d} = T_h \tilde{h} + T_n \tilde{n}$

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data $\longrightarrow$ $\tilde{d} = T_h \tilde{h} + T_n \tilde{n}$

arm + TDI response

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data → $\tilde{d} = T_h \tilde{h} + T_n \tilde{n}$

arm + TDI response

GW strain

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data → $$\tilde{d} = T_h \tilde{h} + T_n \tilde{n}$$

arm + TDI response

GW strain

TDI response

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data $\longrightarrow$

$$\tilde{d} = \underbrace{T_h}_{\text{arm + TDI response}} \underbrace{\tilde{h}}_{\text{GW strain}} + \underbrace{T_n}_{\text{TDI response}} \underbrace{\tilde{n}}_{\text{Single-link noise vector}}$$

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data $\longrightarrow$ $$\tilde{d} = T_h \tilde{h} + T_n \tilde{n}$$

arm + TDI response GW strain TDI response Single-link noise vector

$$\boldsymbol{\Sigma}(f) = \mathbf{R}_h(f)\, S_h(f, \theta_h) + \mathbf{R}_n(f)\, S_n(f, \theta_n)$$

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data → $\tilde{d} = T_h \tilde{h} + T_n \tilde{n}$

arm + TDI response | GW strain | TDI response | Single-link noise vector

$$\boldsymbol{\Sigma}(f) = \mathbf{R}_h(f)\, S_h(f, \theta_h) + \mathbf{R}_n(f)\, S_n(f, \theta_n)$$

Full 3 x 3 Michelson
TDI covariance

[Baghi et al, in prep.]

- Independent component analysis to separate GW stochastic backgrounds from noise

TDI data $\longrightarrow \tilde{d} = T_h \tilde{h} + T_n \tilde{n}$

arm + TDI response

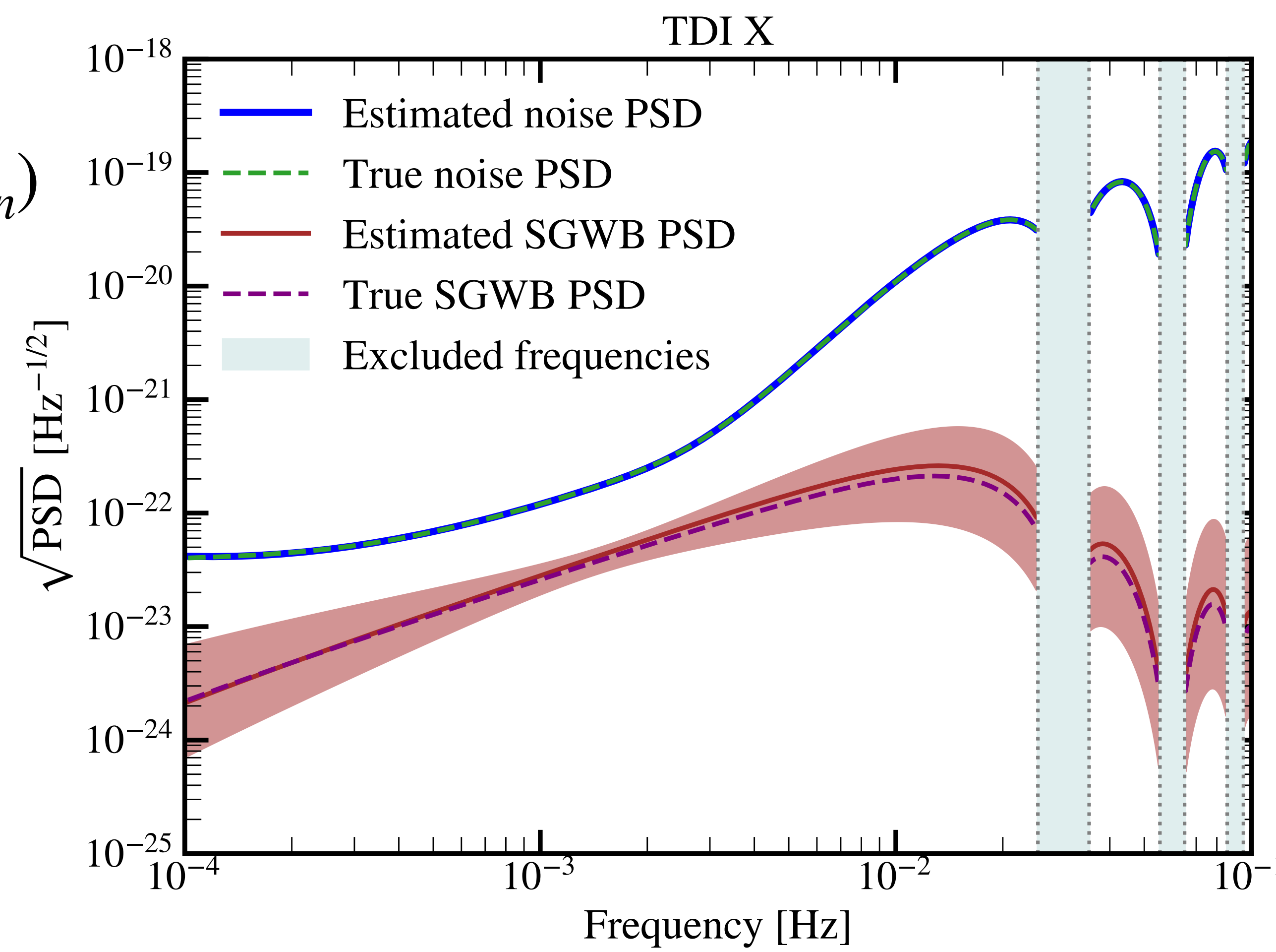
GW strain

TDI response

Single-link noise vector

$$\boldsymbol{\Sigma}(f) = \mathbf{R}_h(f)\, S_h(f, \theta_h) + \mathbf{R}_n(f)\, S_n(f, \theta_n)$$

Full 3 x 3 Michelson TDI covariance



[Baghi et al, in prep.]

- Solving the LISA global fit is a **research problem**

- Solving the LISA global fit is a **research problem**
- We use **Bayesian concepts which require acceleration**:
  - ✦ Data representation and segmentation
  - ✦ High-performance computing (14.5M CPU hrs) and GPUs [Katz et al, 2020]

CONCLUSIONS: TOWARDS THE FUTURE

- Solving the LISA global fit is a **research problem**
- We use **Bayesian concepts which require acceleration**:
  - Data representation and segmentation
  - High-performance computing (14.5M CPU hrs) and GPUs [Katz et al, 2020]
- Other techniques:
  - Reduced order modelling / heterodyned likelihood [See Sylvain Marsat's talk!]
  - Non-Bayesian: especially for detection / low-latency [Blelly et al, 2020]

CONCLUSIONS: TOWARDS THE FUTURE

- Solving the LISA global fit is a **research problem**
- We use **Bayesian concepts which require acceleration**:
  - Data representation and segmentation
  - High-performance computing (14.5M CPU hrs) and GPUs [Katz et al, 2020]
- Other techniques:
  - Reduced order modelling / heterodyned likelihood [See Sylvain Marsat's talk!]
  - Non-Bayesian: especially for detection / low-latency [Blelly et al, 2020]
- Framework for research: **the LISA Data Challenges**
  - Collaborative playground https://lisa-ldc.lal.in2p3.fr/
  - Progressively increases the number of source types in "enchiladas" + instrumental realism
  - Writing of the **LISA Data Analysis Living Reviews**:
    sources, analysis methods, acceleration algorithms, specificities of LISA, challenges

CONCLUSIONS: TOWARDS THE FUTURE

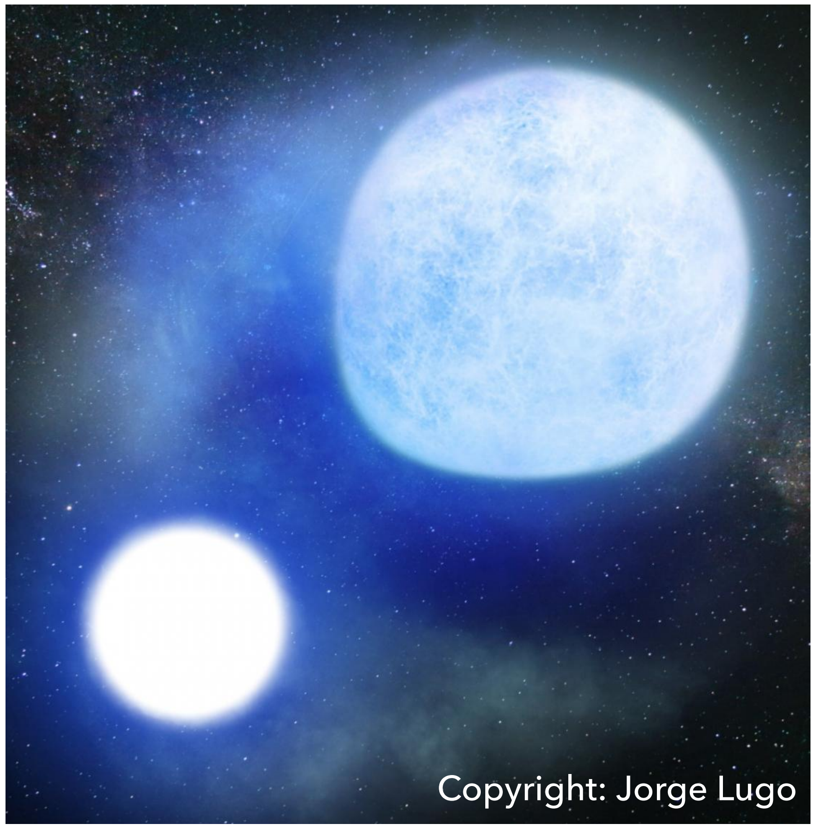
- Solving the LISA global fit is a **research problem**
- We use **Bayesian concepts which require acceleration**:
  - ✦ Data representation and segmentation
  - ✦ High-performance computing (14.5M CPU hrs) and GPUs [Katz et al, 2020]
- Other techniques:
  - ✦ Reduced order modelling / heterodyned likelihood [See Sylvain Marsat's talk!]
  - ✦ Non-Bayesian: especially for detection / low-latency [Blelly et al, 2020]
- Framework for research: **the LISA Data Challenges**
  - ✦ Collaborative playground https://lisa-ldc.lal.in2p3.fr/
  - ✦ Progressively increases the number of source types in "enchiladas" + instrumental realism
  - ✦ Writing of the **LISA Data Analysis Living Reviews**:
    sources, analysis methods, acceleration algorithms, specificities of LISA, challenges

Thank you for your attention !

BACKUP SLIDES

What makes gravitational noise in the milihertz band?

What makes gravitational noise in the milihertz band?

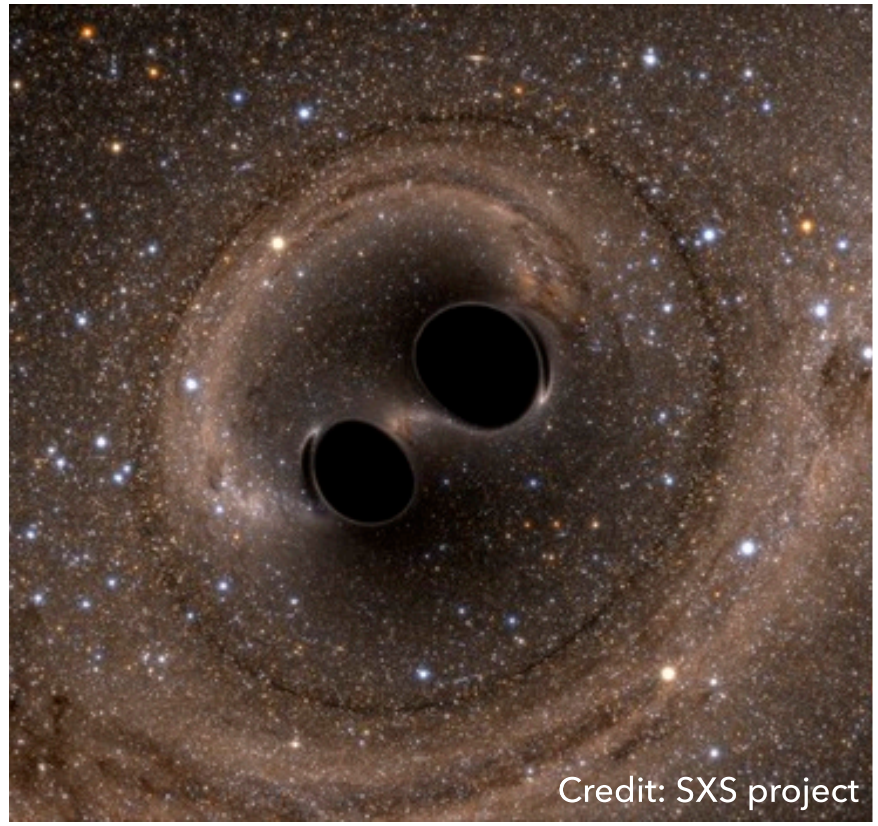


In our galaxy: pairs of orbiting white dwarfs

What makes gravitational noise in the milihertz band?

In our galaxy: pairs of orbiting white dwarfs

What makes gravitational noise in the milihertz band?



In our galaxy: pairs of orbiting white dwarfs

One billion light-years away: collision of black holes

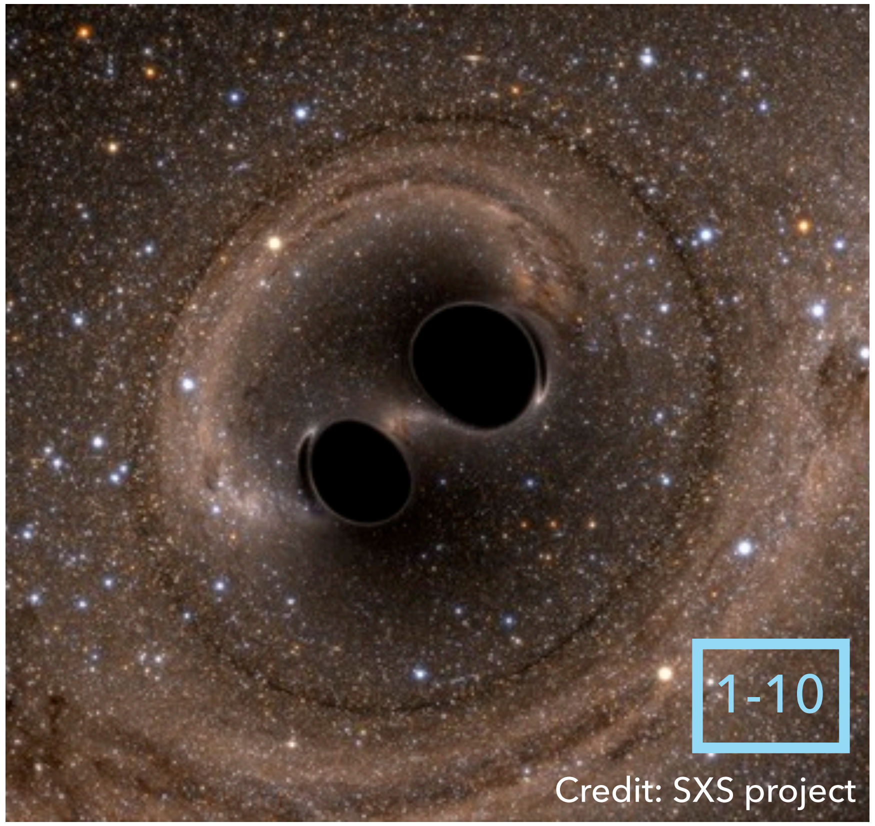
What makes gravitational noise in the milihertz band?

In our galaxy: pairs of orbiting white dwarfs

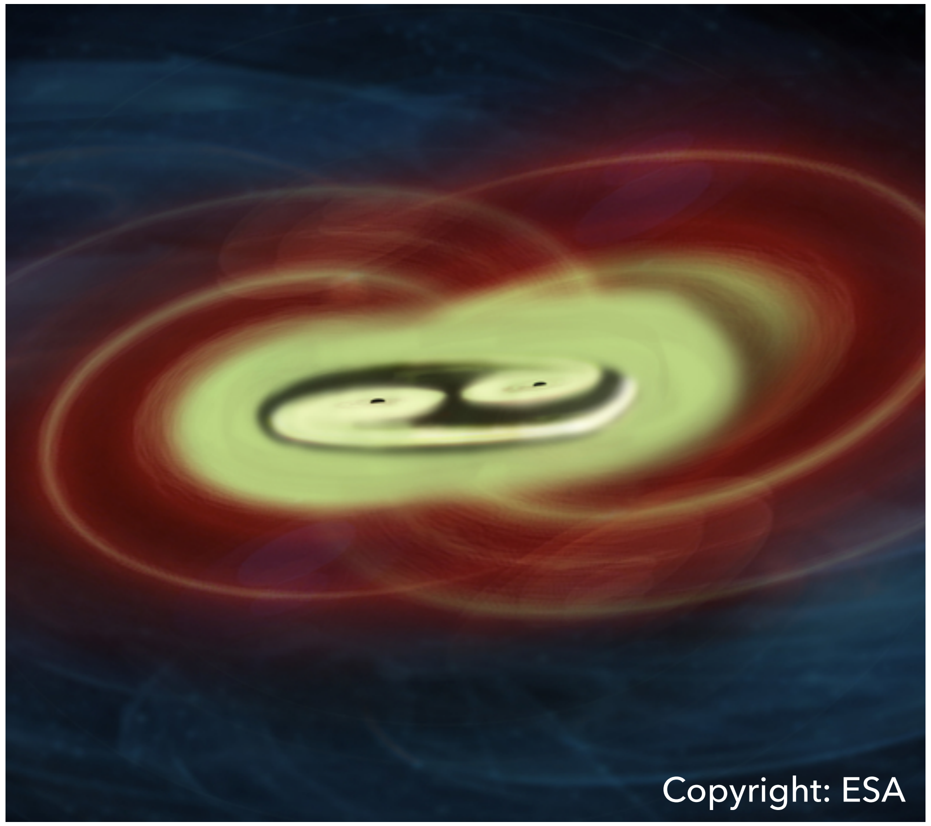
One billion light-years away: collision of black holes

What makes gravitational noise in the milihertz band?

In our galaxy: pairs of orbiting white dwarfs



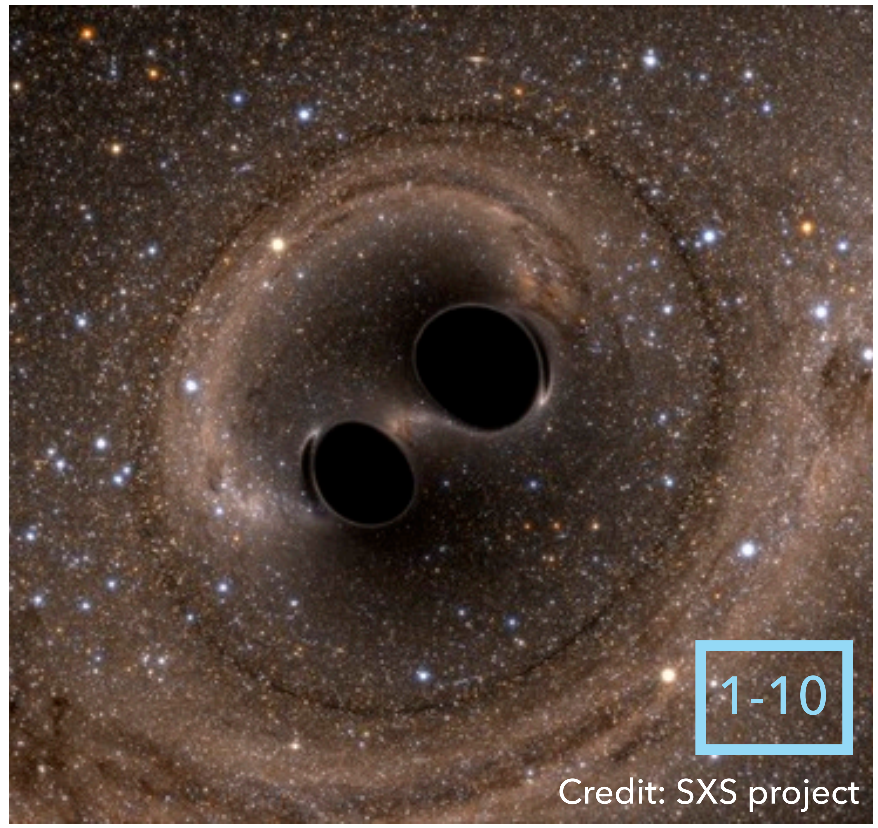
One billion light-years away: collision of black holes



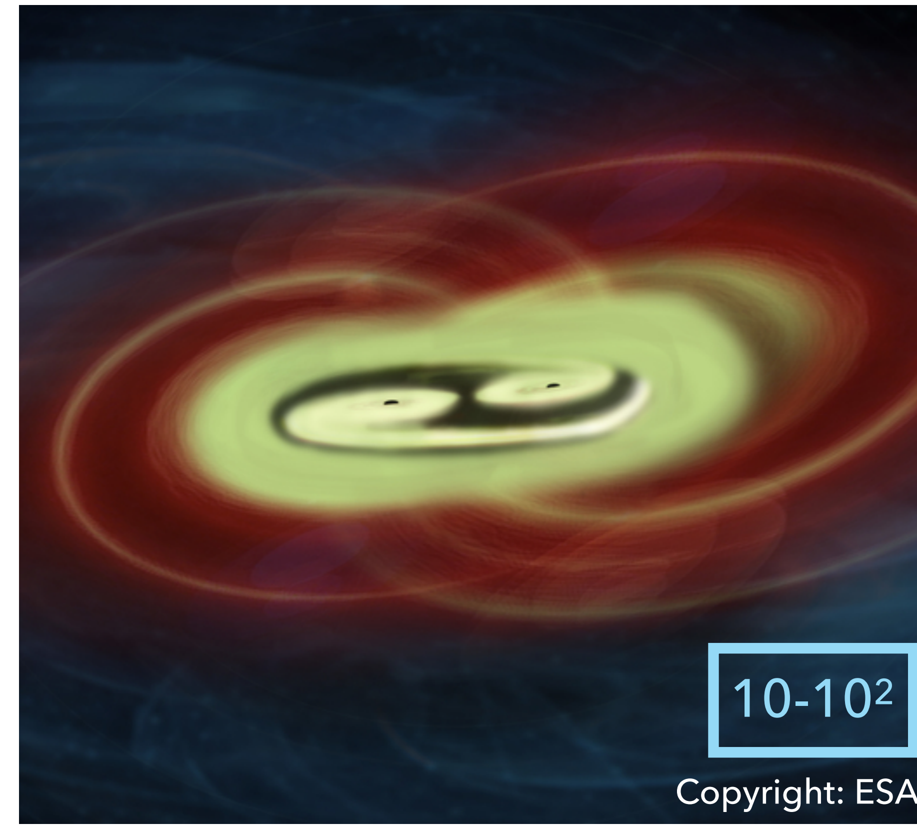
Possibly farther away: merging supermassive black holes

What makes gravitational noise in the milihertz band?

In our galaxy: pairs of orbiting white dwarfs



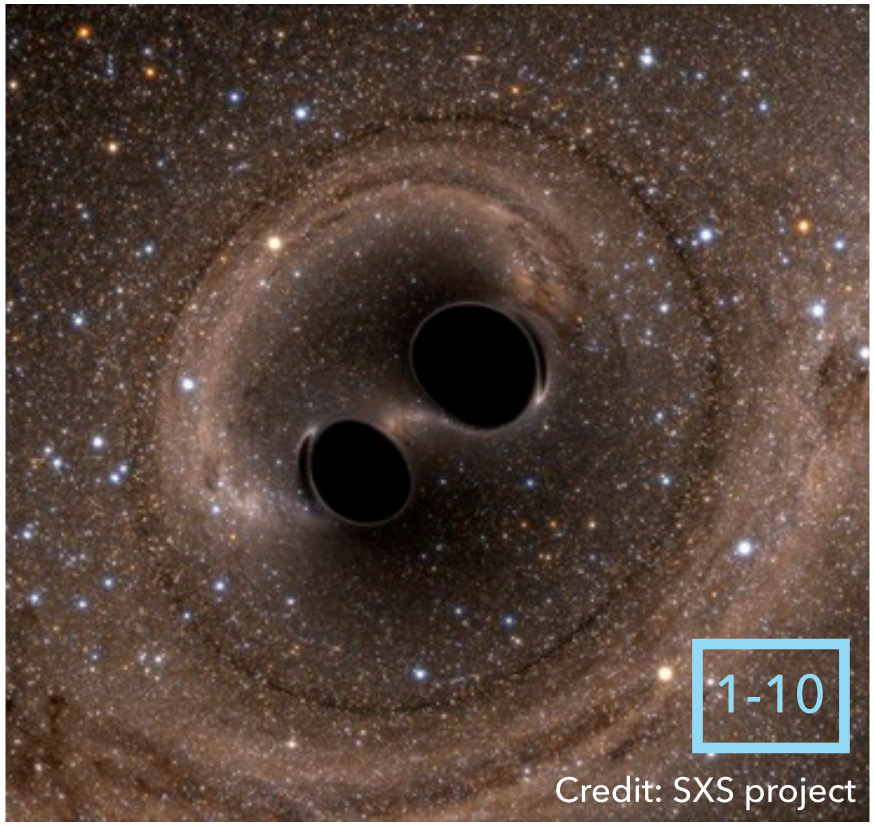
One billion light-years away: collision of black holes



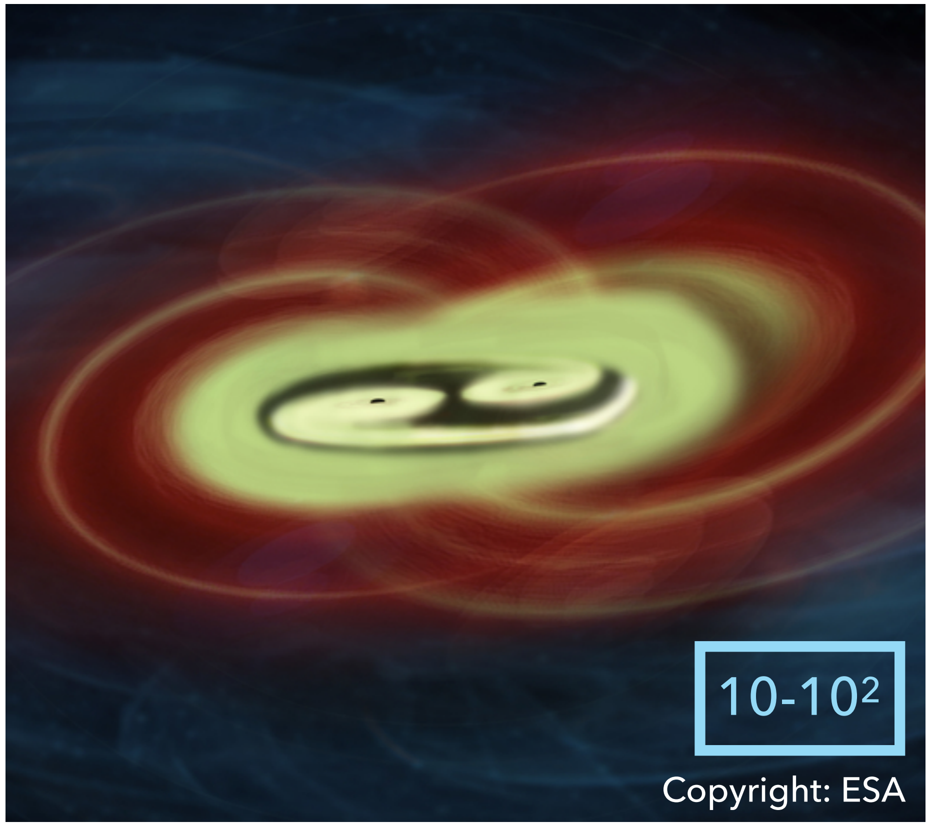
Possibly farther away: merging supermassive black holes

What makes gravitational noise in the milihertz band?

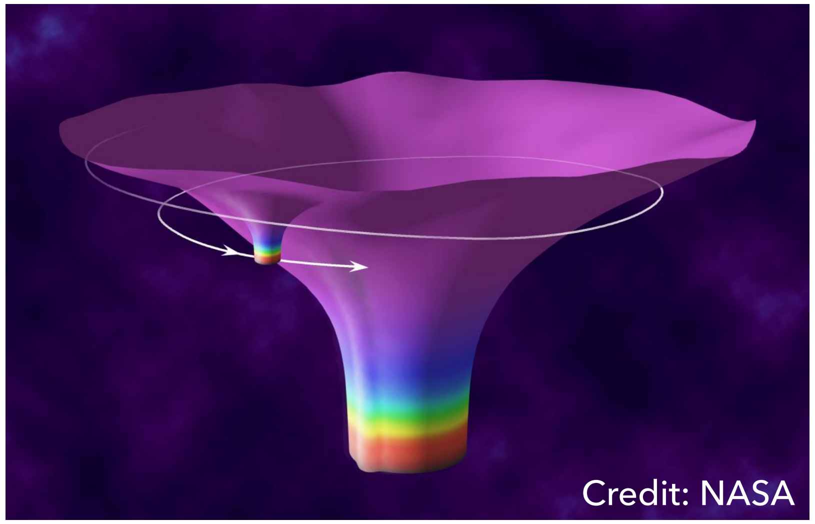
In our galaxy: pairs of orbiting white dwarfs



One billion light-years away: collision of black holes



Possibly farther away: merging supermassive black holes

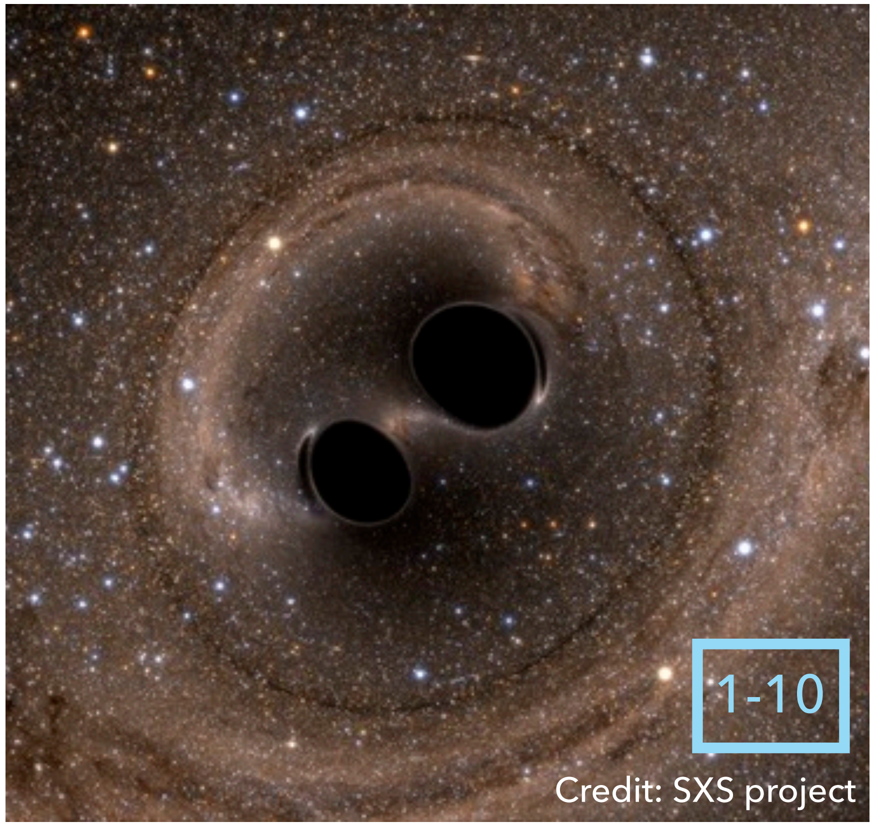


◀ Extreme-mass-ratio inspirals: a smaller compact object orbiting a supermassive black hole

What makes gravitational noise in the milihertz band?

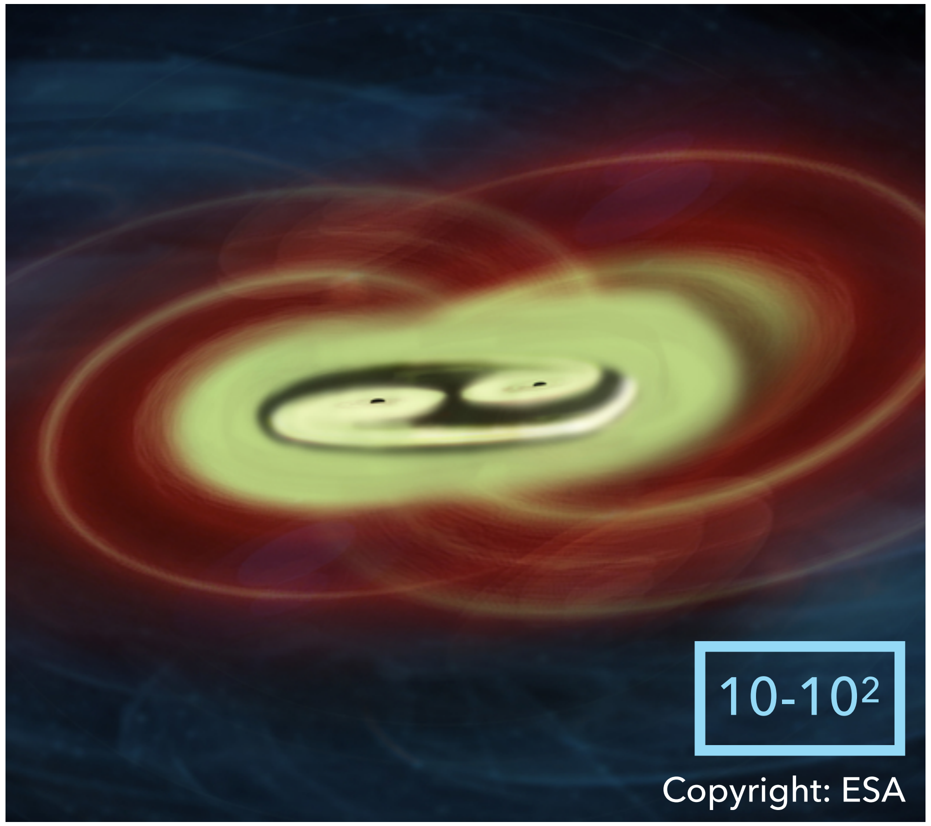
Copyright: Jorge Lugo

In our galaxy: pairs of orbiting white dwarfs



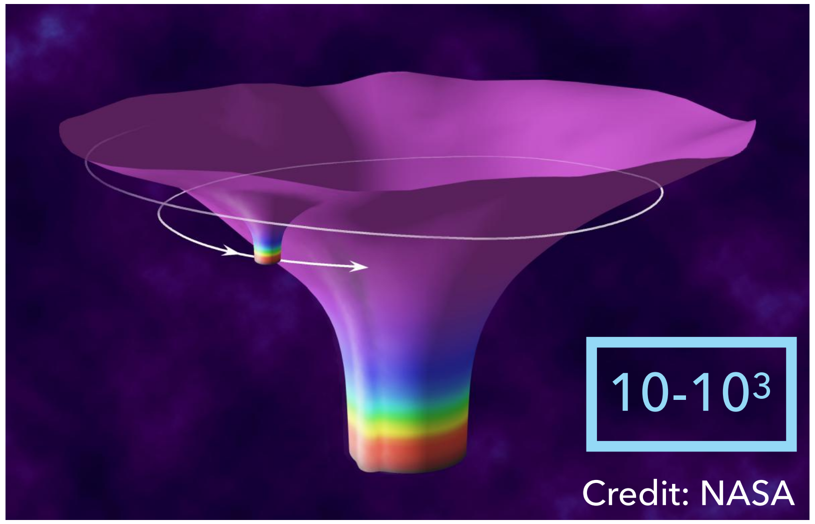
Credit: SXS project

One billion light-years away: collision of black holes



Copyright: ESA

Possibly farther away: merging supermassive black holes



Credit: NASA

◀ Extreme-mass-ratio inspirals: a smaller compact object orbiting a supermassive black hole

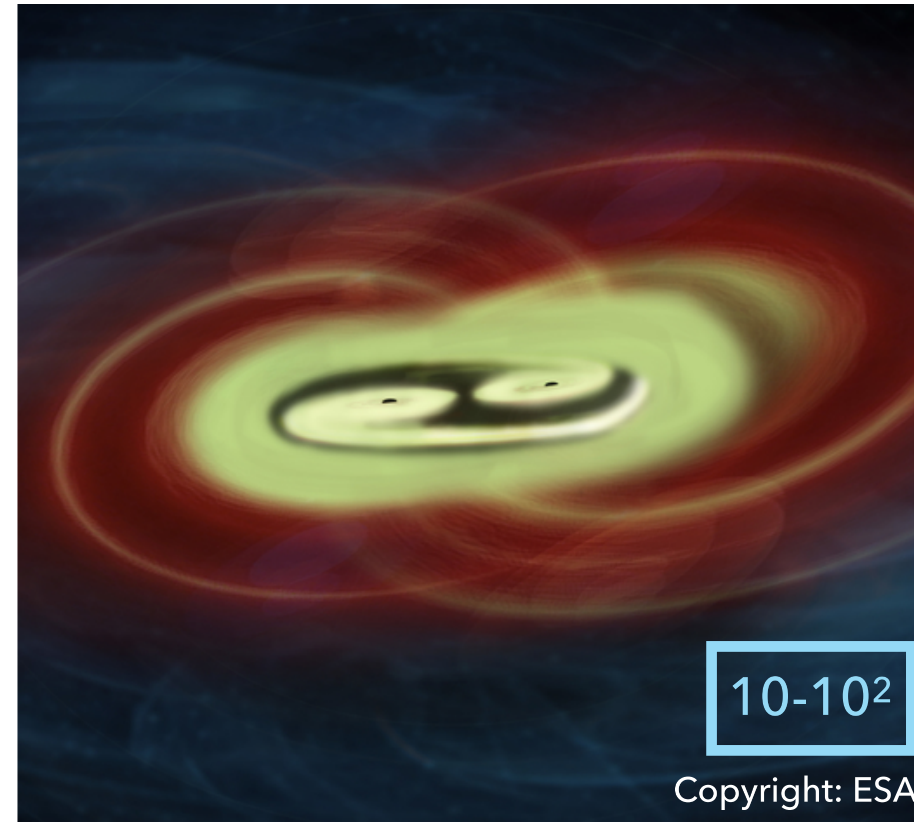
What makes gravitational noise in the milihertz band?

Copyright: Jorge Lugo

In our galaxy: pairs of orbiting white dwarfs

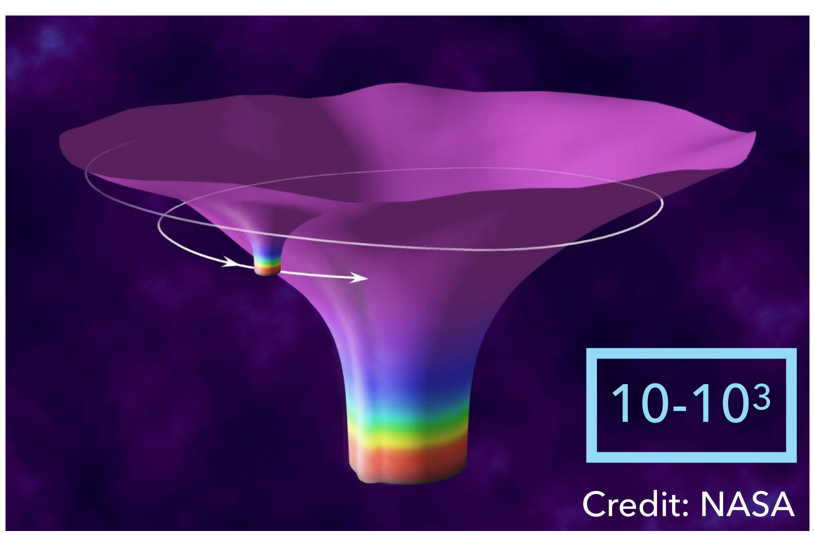
Credit: SXS project

One billion light-years away: collision of black holes



Copyright: ESA

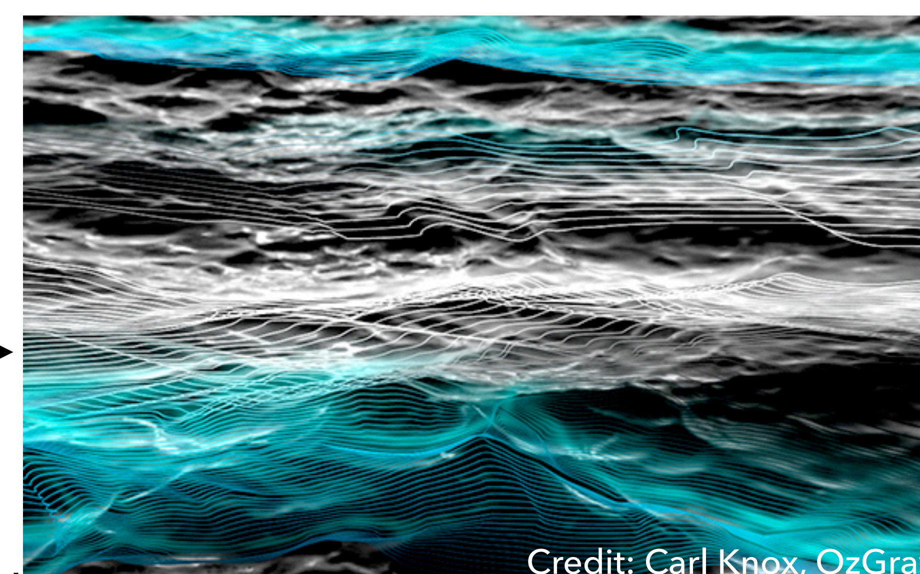
Possibly farther away: merging supermassive black holes



Credit: NASA

◀ Extreme-mass-ratio inspirals: a smaller compact object orbiting a supermassive black hole

In the entire universe: a cosmic gravitational wave background? ▶


Credit: Carl Knox, OzGrav