Applying a novel method for finding correlated features in gravitational wave detectors

2015/8/7 Fri. @Bi-Monthly Workshop(10th)
Hirotaka Yuzurihara

Background

- Correlated noises at LIGO and Virgo

Method to detect correlation

- Pearson Correlation Coefficient
- Maximum Information Coefficient

Non-linear noise model

- Virgo detector suffered from up-conversion noise before. (Now solved)

This up-conversion noise is well-modeled.

=> In this talk, using non-linearly noise model,
we evaluate the performance of each analysis methods.

Summary

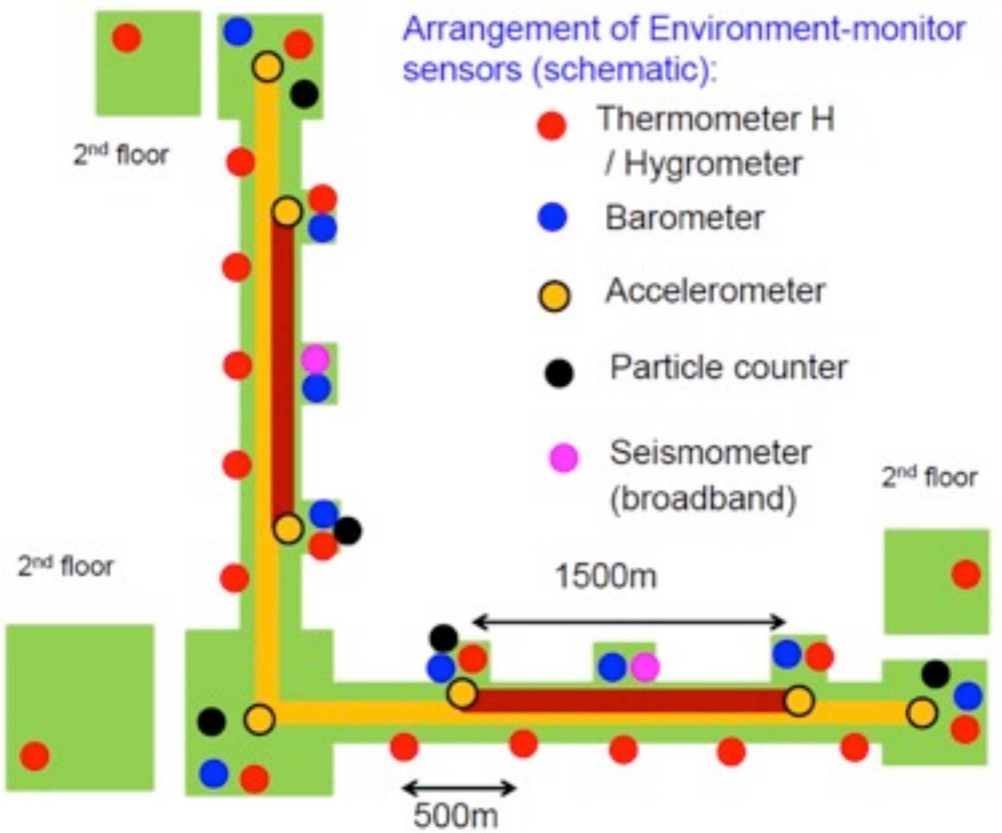
motivation - correlation analysis using environmental channels

Goal : search correlated channels between ~10000 environmental channels and finally localize noise sources

- localization of noise sources reveals noise features
- remove false trigger event generated by GW search pipeline,
  -> increase GW detection efficiency

In this talk, we define
- GW channel
  as sensitive channel to GW
- environmental channel
  as insensitive channel to GW

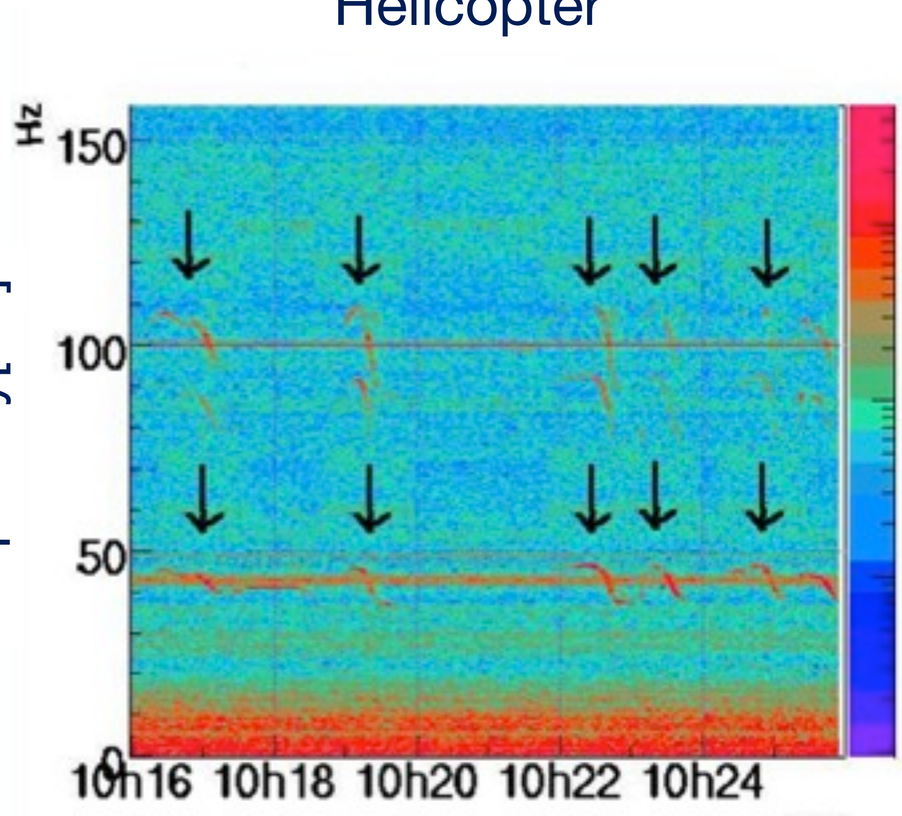
microphone, accelerometer, seismometer, thermometer, barometer, magnetometer ...



Araya (2012)

Example of linear correlation observed in LIGO and Virgo

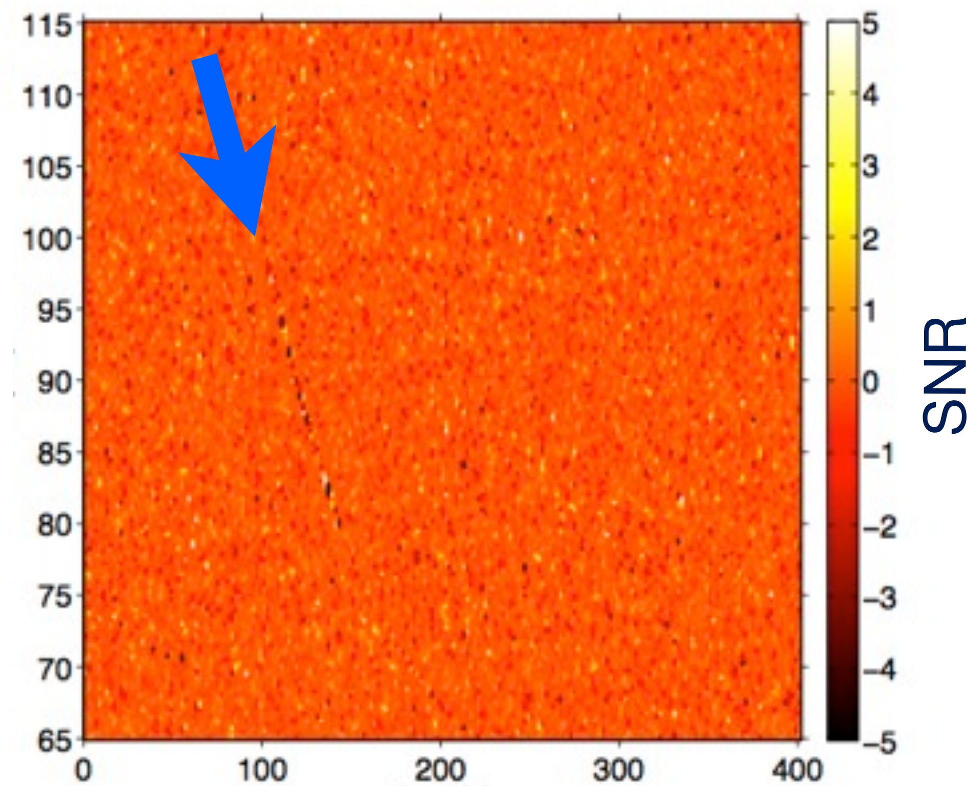
Helicopter



Coughlin et.al (2011)[arXiv 1108.1521]

time[sec]

Airplane

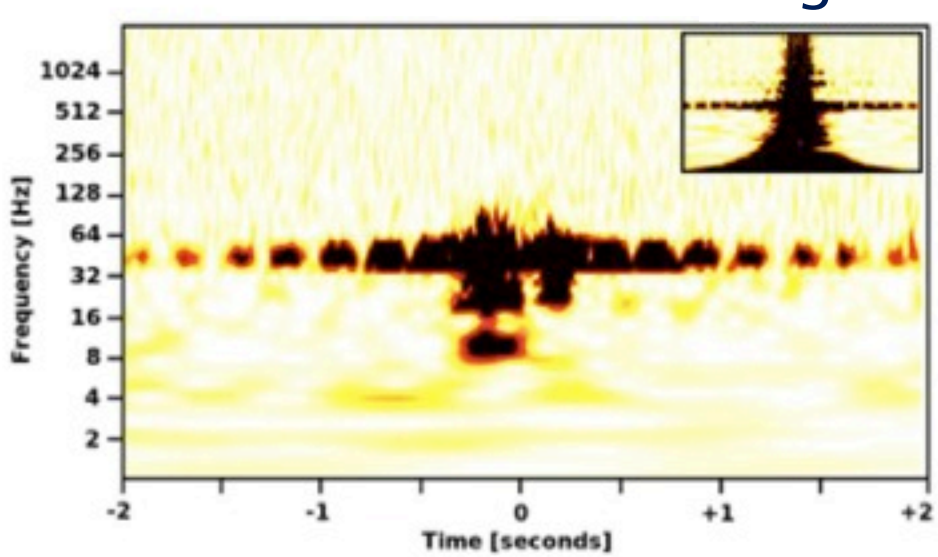


time[sec]

[LIGO-T050174]

GW channel

Magnetic sensor



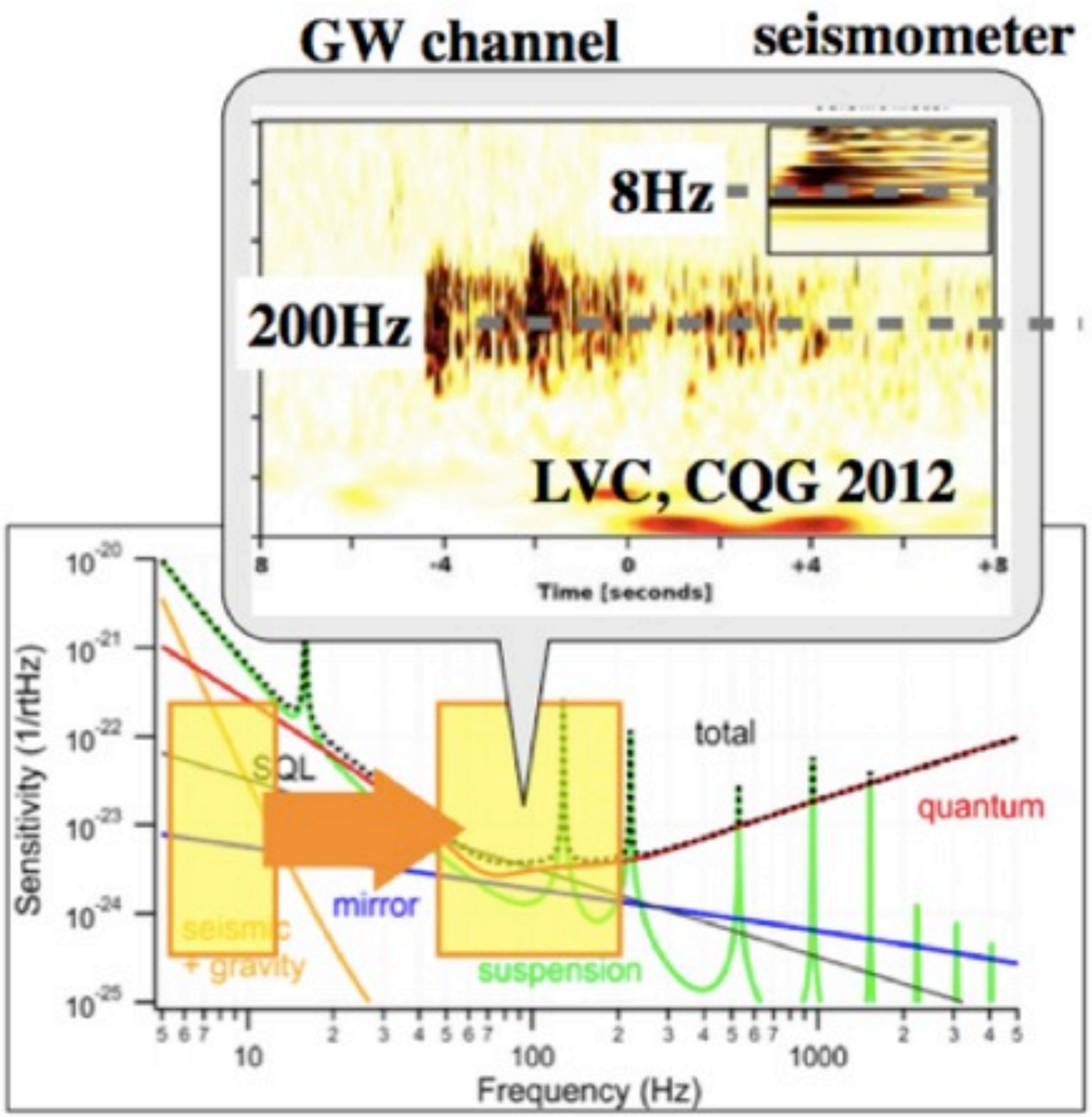
J.Asis et al. (2012) [gr-qc 1203.5613]

Example of non-linear correlation, up-converted noise

Up-converted noise:
seismic glitches will excite optical bench motion which cause scattered light noise.
In bad weather day, effect from seismic glitches is strong.
Non-linear correlation over a few Hz ~ a few hundreds Hz in GW channel was observed in detectors.

[Classical and Quantum Gravity 27, 19 (2010) 194011]

http://www.opticsinfobase.org/oe/abstract.cfm?uri=oe-20-8-8329



Sensitivity curve of KAGRA

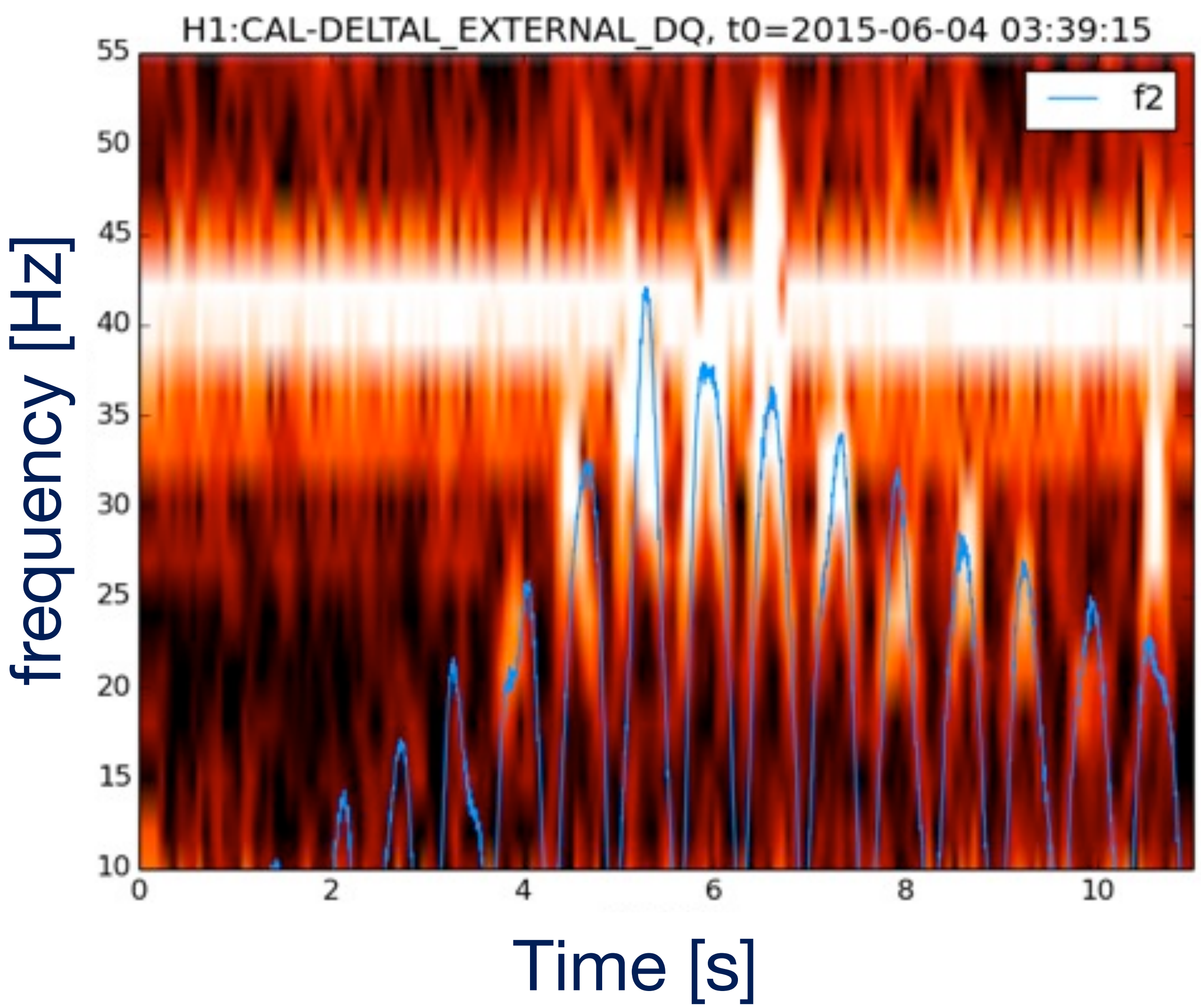
⇒ search correlation between environmental channels

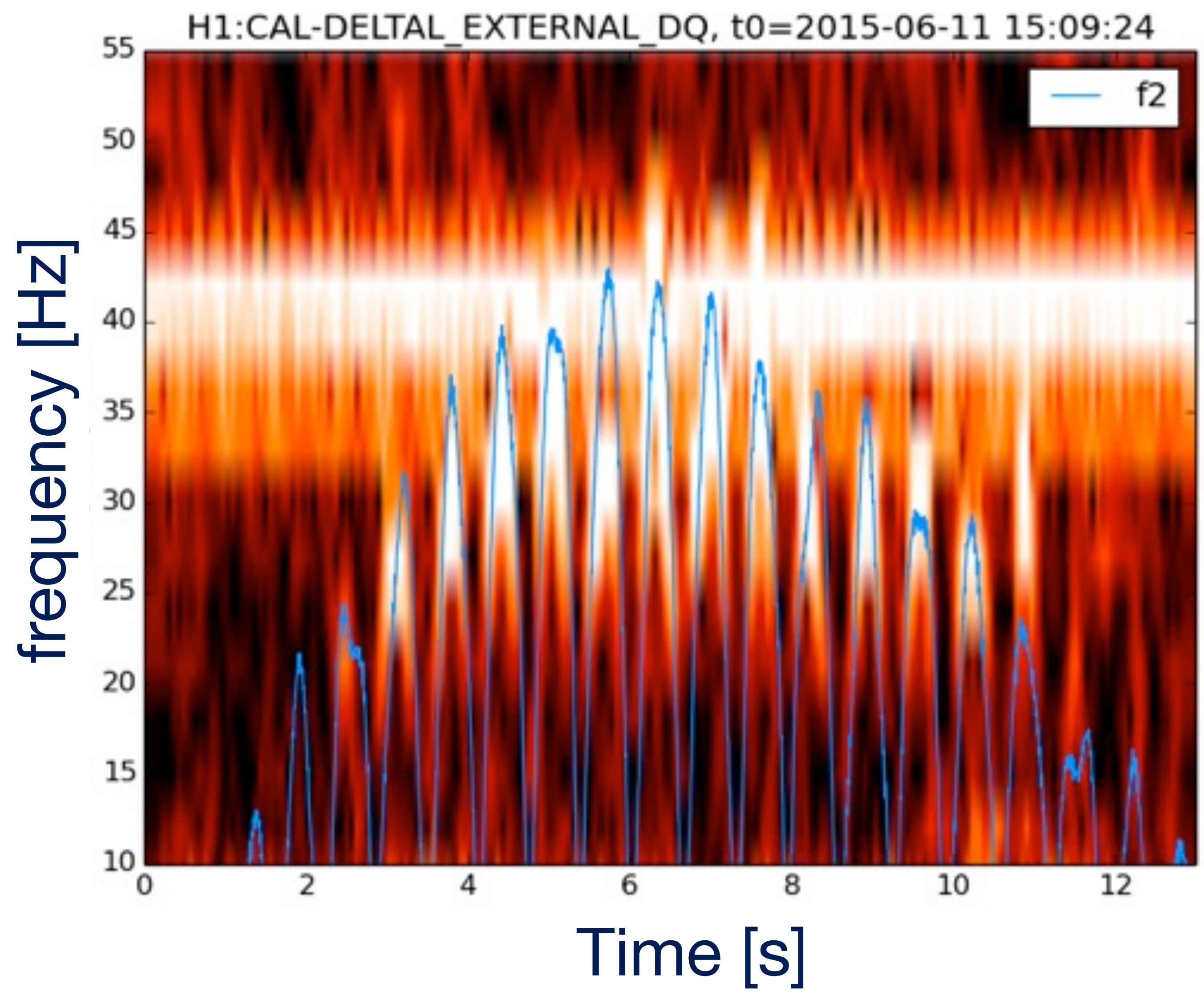
to get information of noise which effect different frequency band

Hayama (2014)
J.Asis et al. (2012) [gr-qc 1203.5613]

The latest news - LIGO Livingston face a up-conversion noise

https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=20079





- During Engineering Run of LIGO Livingston, peak at 0.8~3Hz (from seismic activity?)

  => generate scattering events reaching from 25Hz up to 50Hz in GW channel

  => currently being investigated how this peak affect GW channel like this.

Method to search correlation

- In this study, below two method are used,

Pearson Correlation Coefficient

- efficient method to linear correlation

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

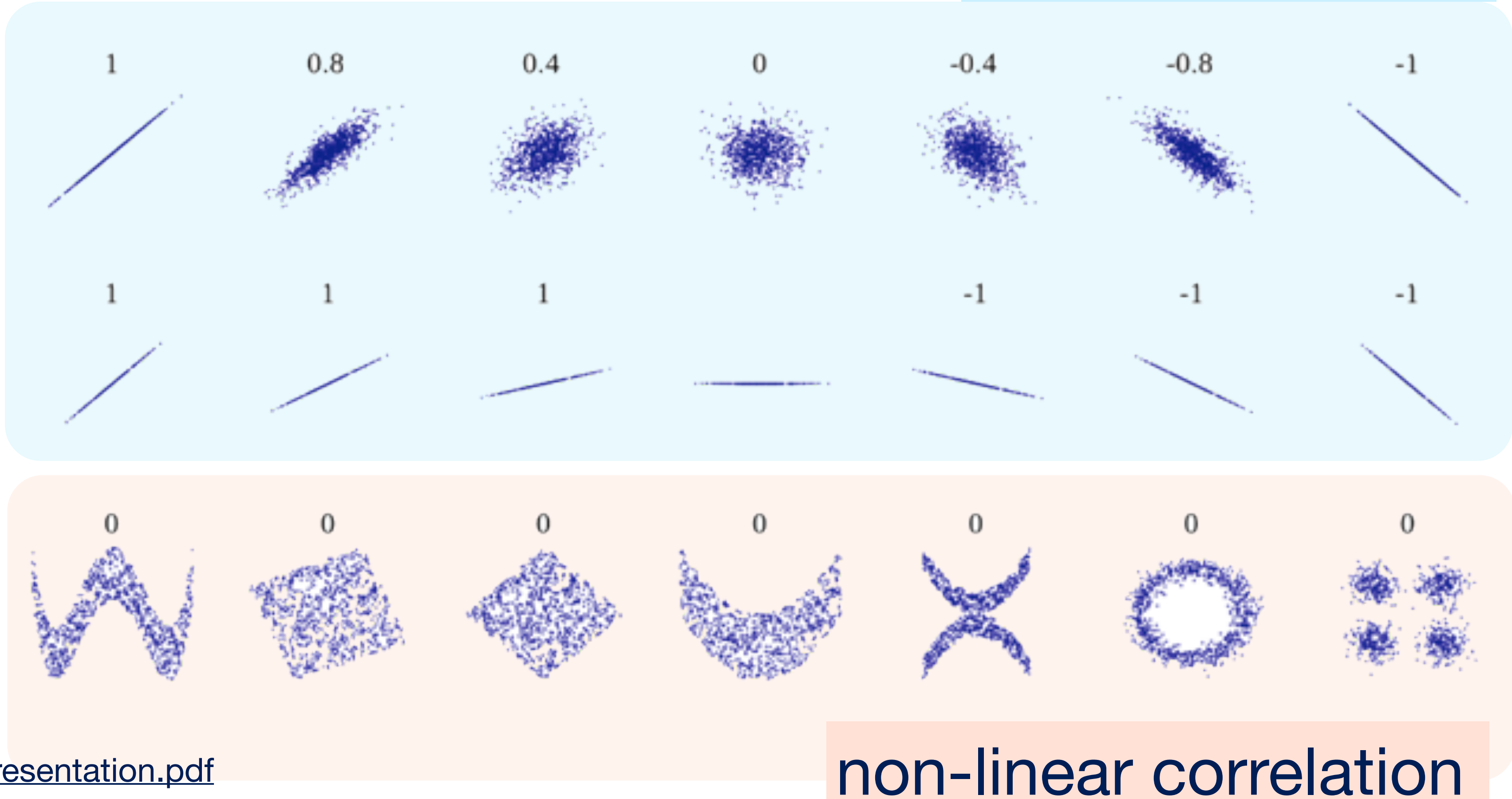
Maximum Information Coefficient (MIC)

[David N. Reshef, et al. Science 334, 1518 (2011)]

linear correlation

Linear correlation →

non-linear correlation →



non-linear correlation

http://lectures.molgen.mpg.de/algsysbio12/MINEPresentation.pdf

Maximum Information Coefficient (MIC)

[David N. Reshef, et al. Science 334, 1518 (2011)]

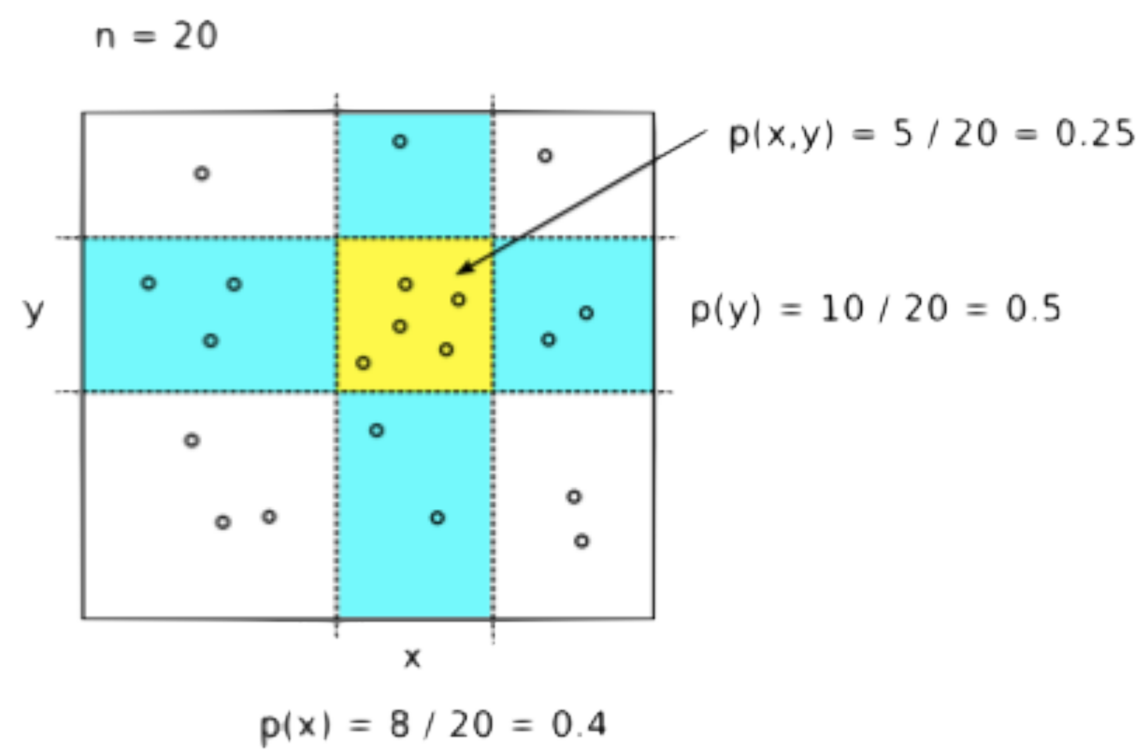
- MIC can detect both functional and non-functional dependence.

- If a relationship exists between two data, a grid can be drawn on the scatter plot of two data that partitions the data to encapsulate that relationship.

For each partitioned resolution, MIC finds grid partition placement with highest mutual information.

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \left(\frac{p(x,y)}{p(x)\,p(y)} \right)$$

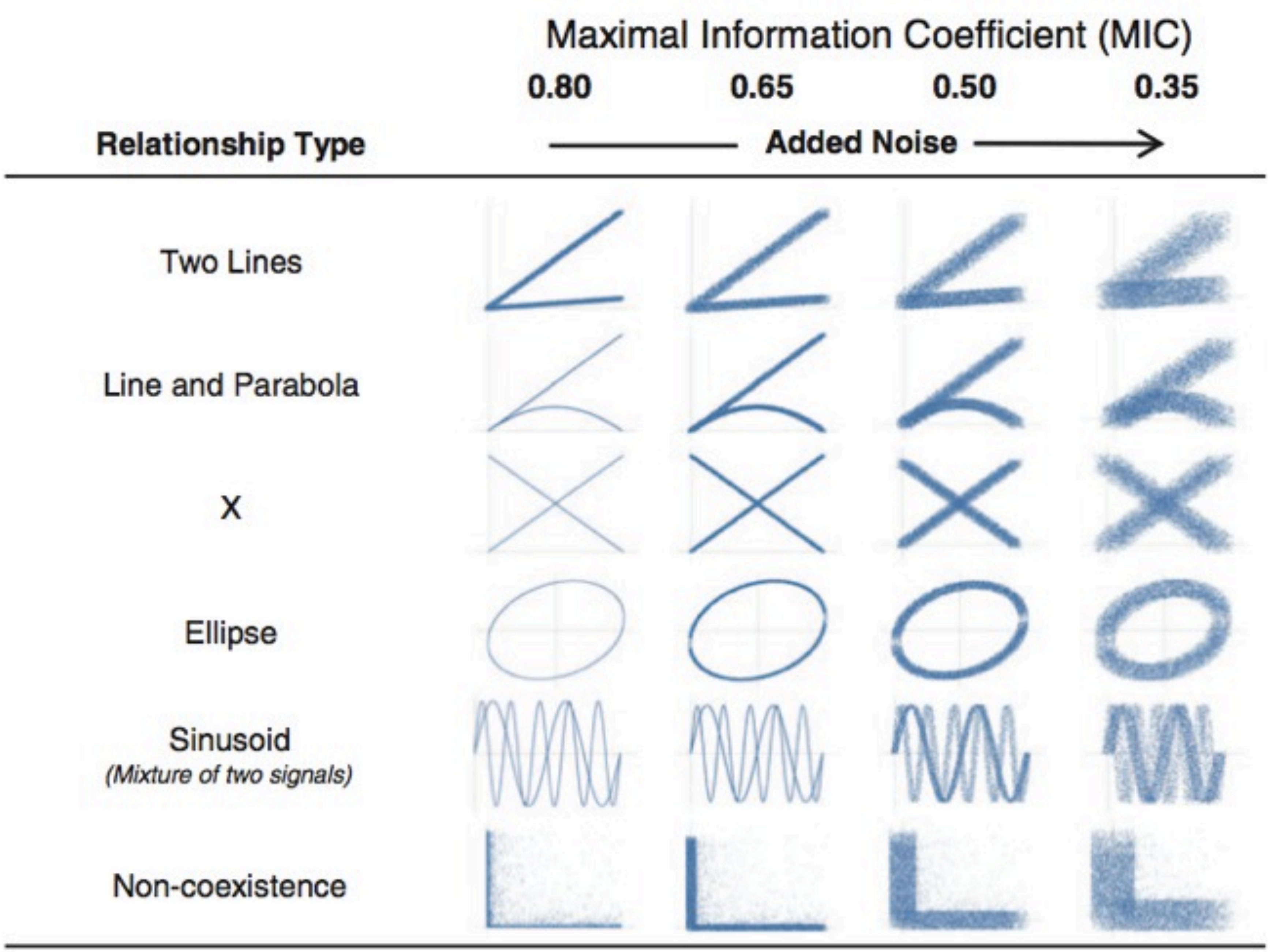
X,Y: random variables
p(x,y): joint probability distribution function
p(x), p(y): marginal probability distribution functions



http://lectures.molgen.mpg.de/algsysbio12/MINEPresentation.pdf

Which correlation MIC can search?

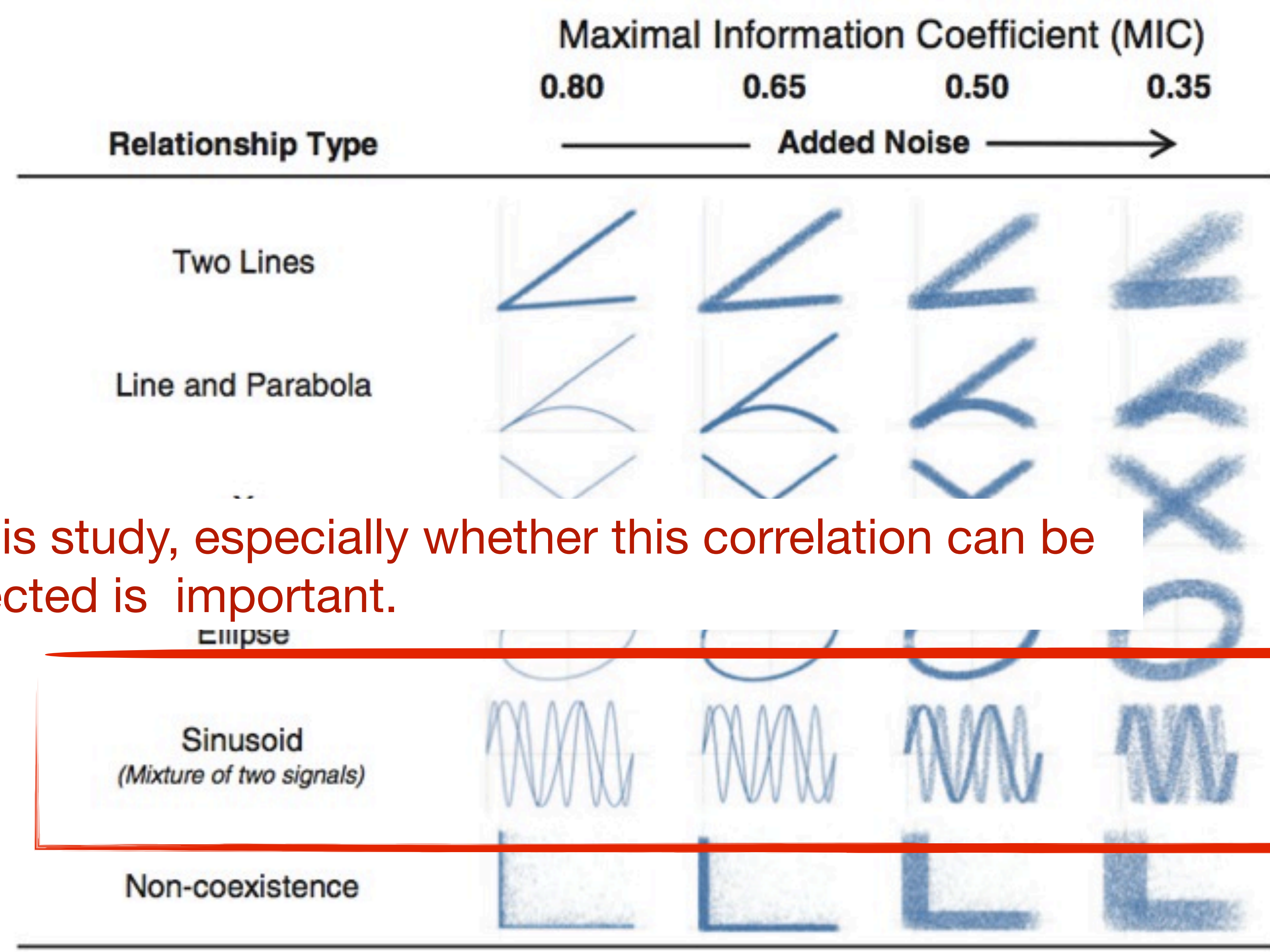
- MIC can search not only linear but also non-linear correlation.



[David N. Reshef, et al. Science 334, 1518 (2011)]

Which correlation MIC can search?

- MIC can search not only linear but also non-linear correlation.

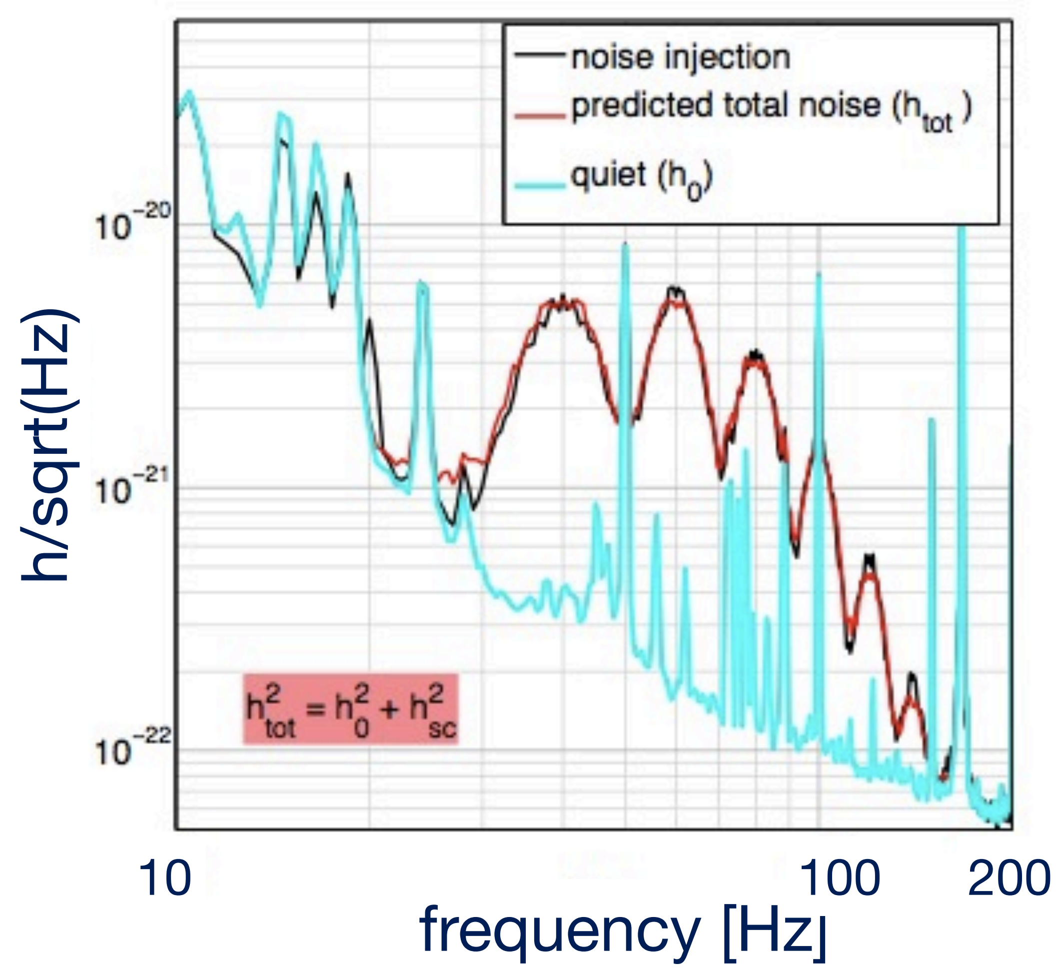


In this study, especially whether this correlation can be detected is important.

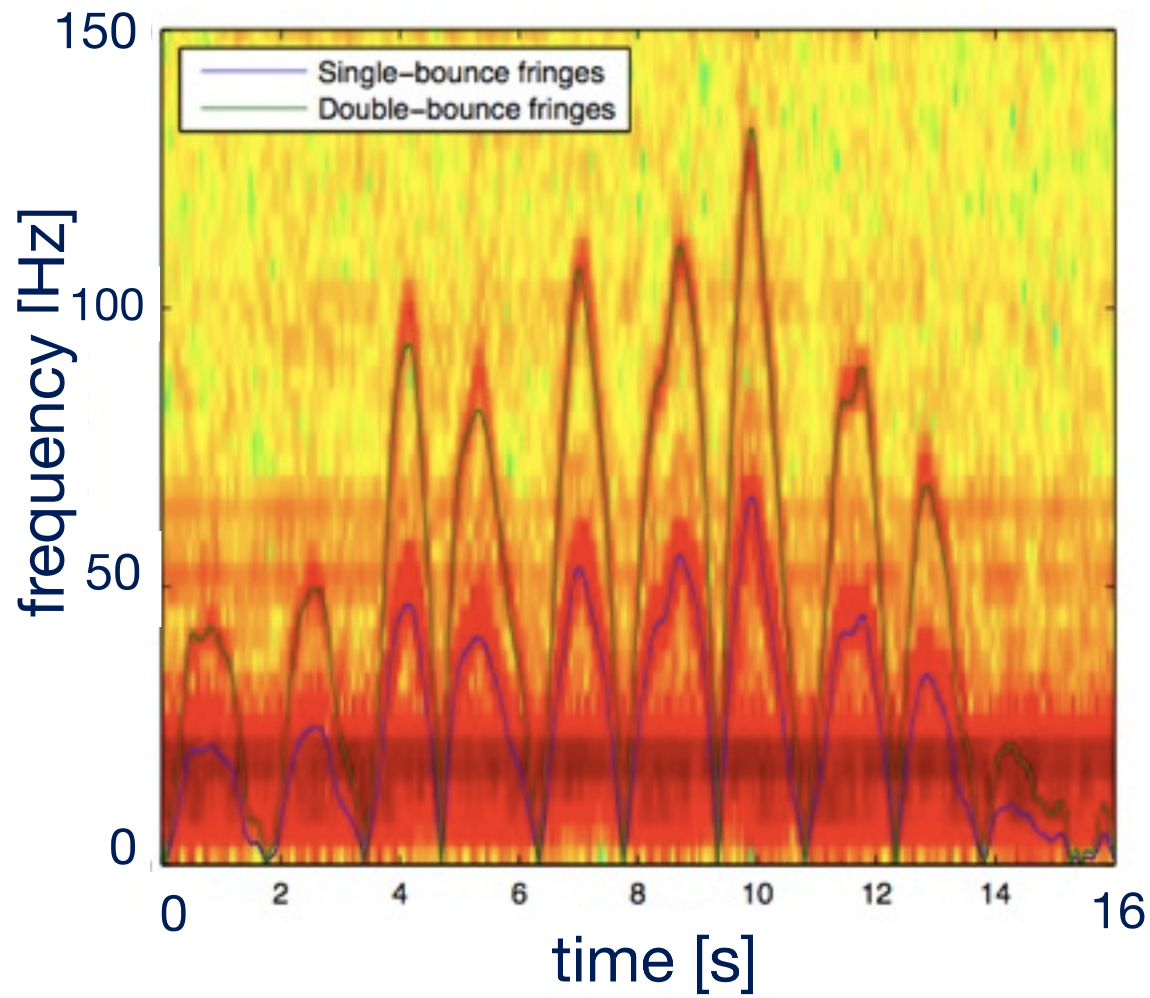
[David N. Reshef, et al. Science 334, 1518 (2011)]

Up-conversion noise observed at Virgo detector

[Classical and Quantum Gravity 27, 19 (2010) 194011]



secondary scattering light noise by mirror's vibration
The structure with many peaks becomes worse sensitivity more than 1 order.

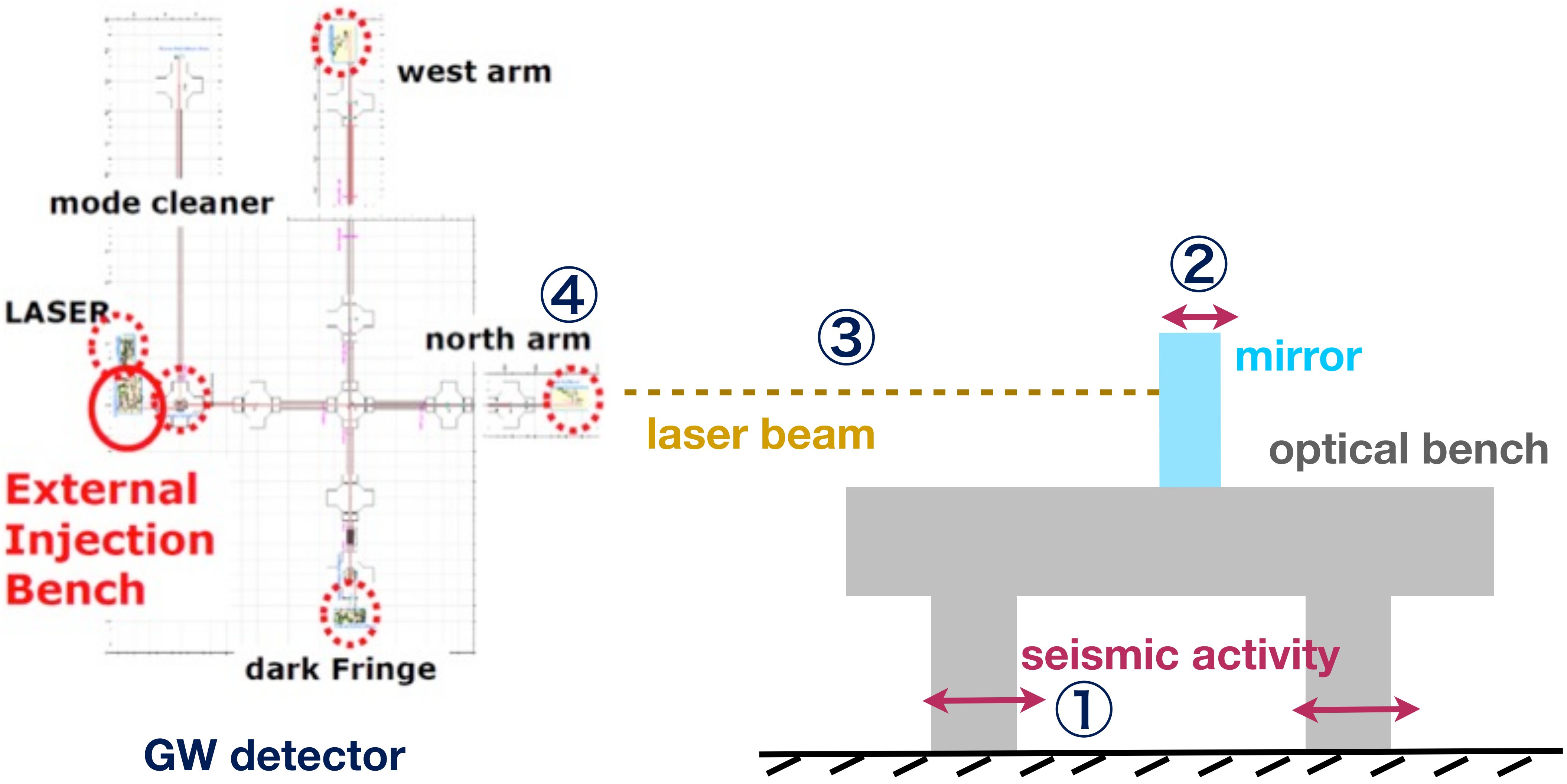


spectrogram of secondary scattering light noise

Virgo detector suffered from this noise before. Now solved
This up-conversion noise is well-modeled.
This noise model includes linear and non-linear correlation.

Mechanism of this up-conversion

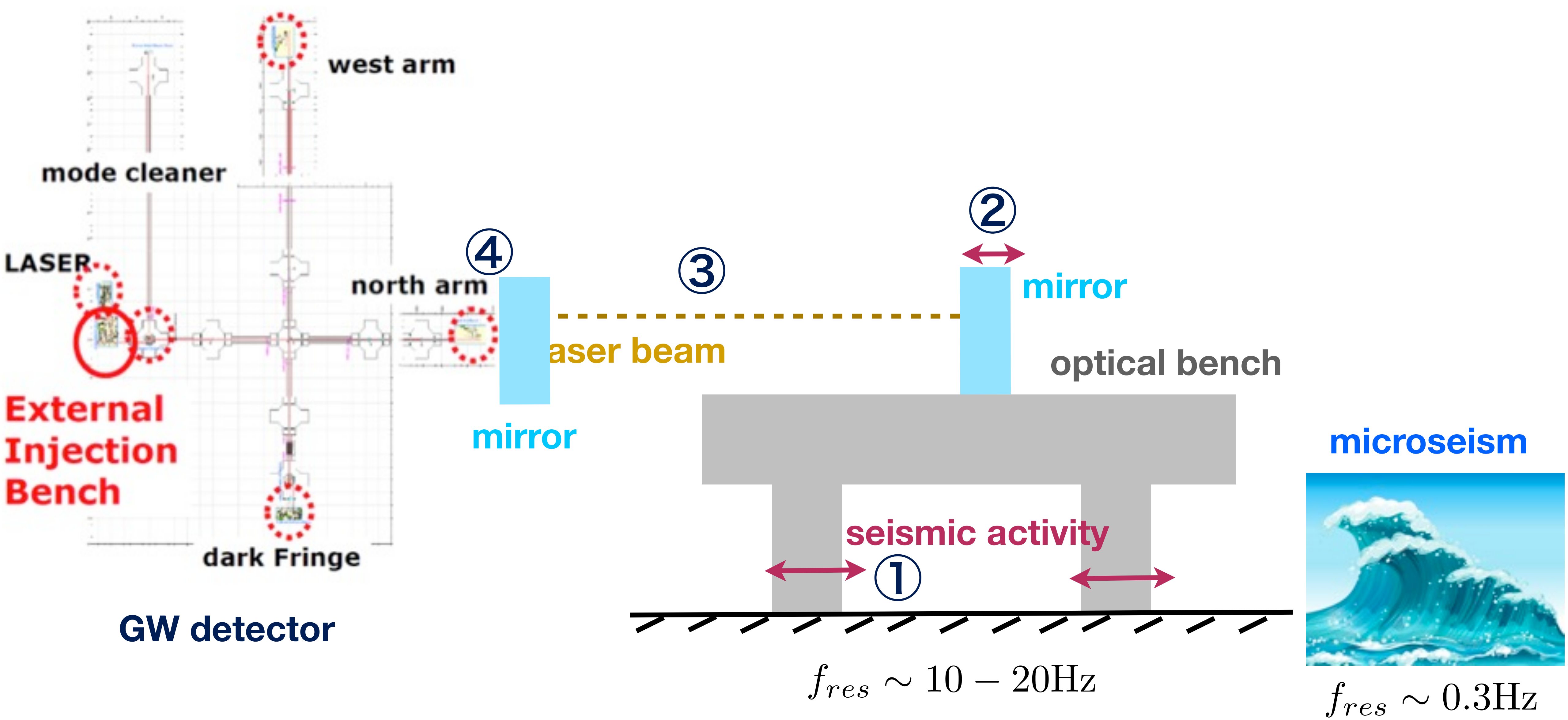
[Classical and Quantum Gravity 27, 19 (2010) 194011]



Optical system behind end-mirror controls GW detector using transmitted laser.
Sometimes accidentally transmitted laser is returned to cavity.

Mechanism of this up-conversion

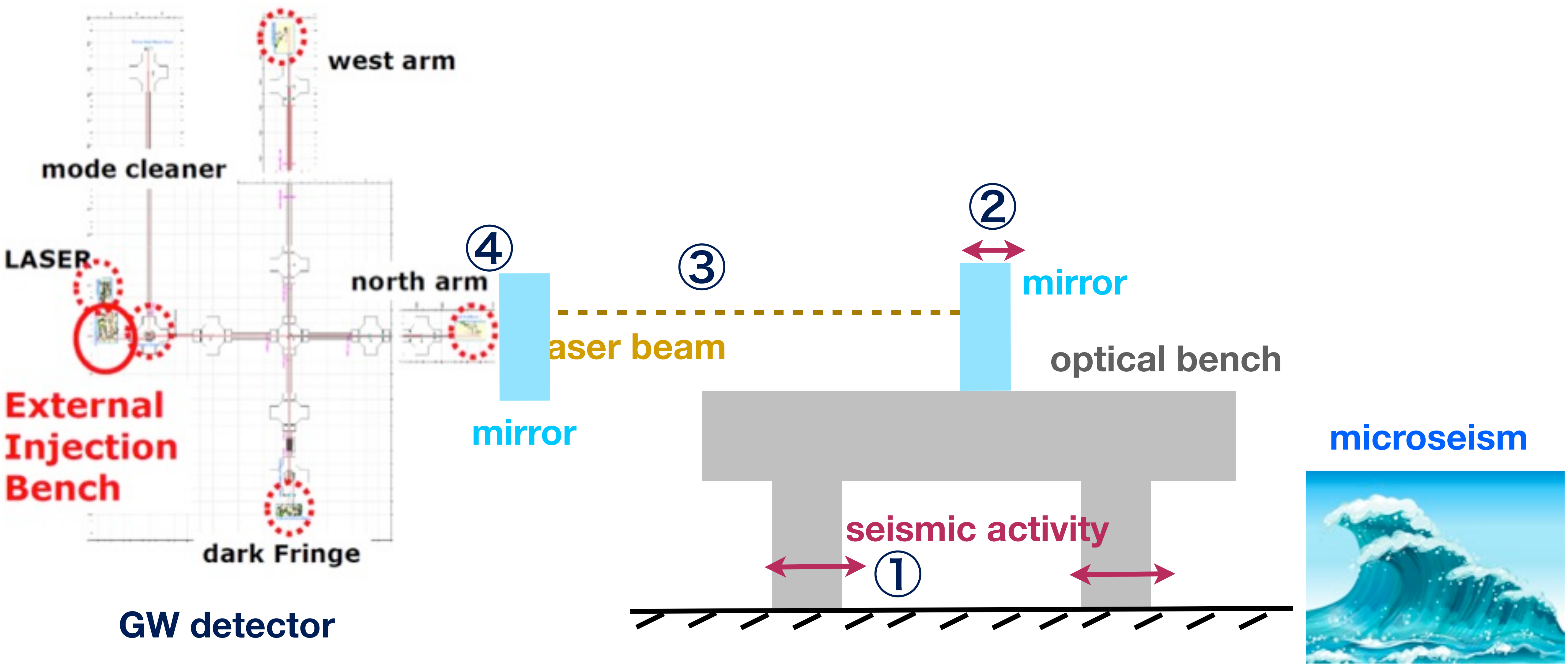
[Classical and Quantum Gravity 27, 19 (2010) 194011]



①. Strong seismic activity (such as microseism..) excite resonant motion of optical bench and generate damping motion of optical bench.

Mechanism of this up-conversion

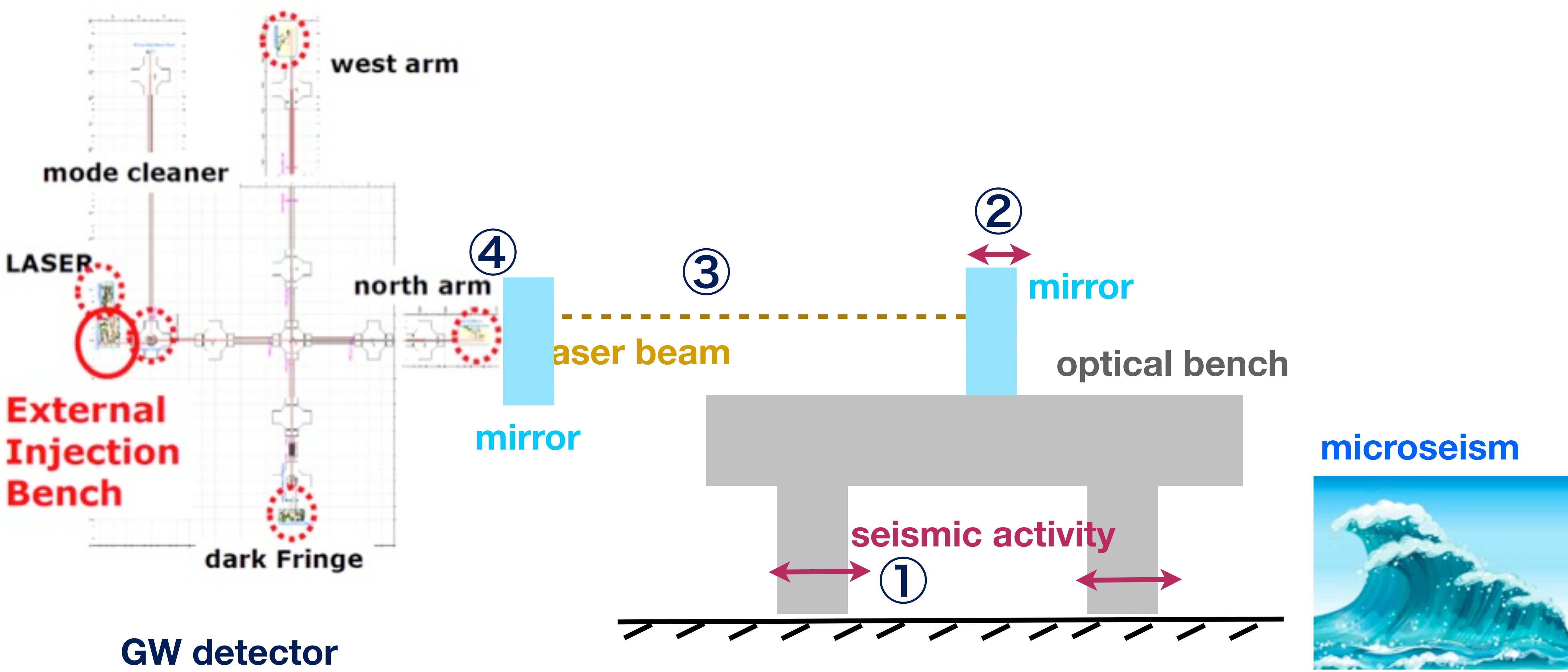
[Classical and Quantum Gravity 27, 19 (2010) 194011]



②. The motion of optical bench causes damping motion of mirror installed on optical bench

Mechanism of this up-conversion

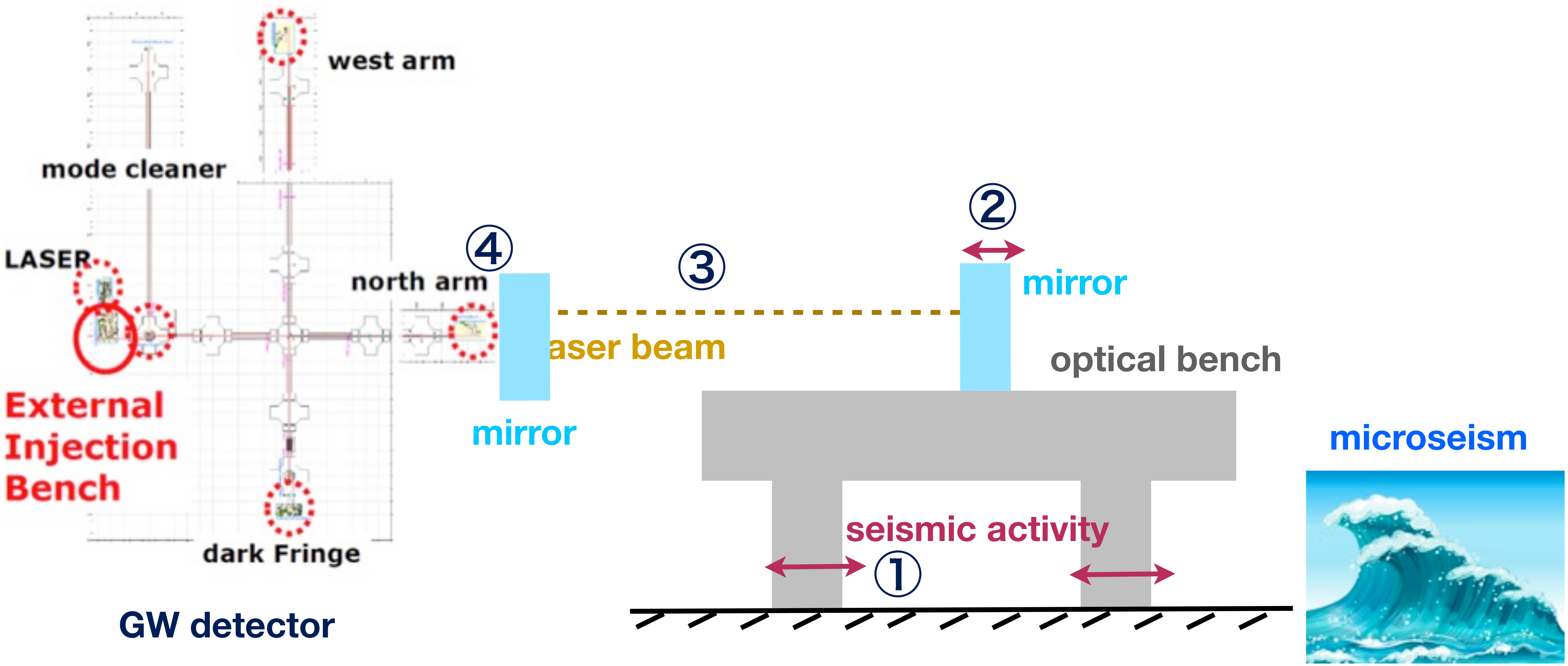
[Classical and Quantum Gravity 27, 19 (2010) 194011]



③. Time variation of optical path length between end-mirror of cavity and mirror on optical bench because of damping motion of mirror on optical bench

Mechanism of this up-conversion

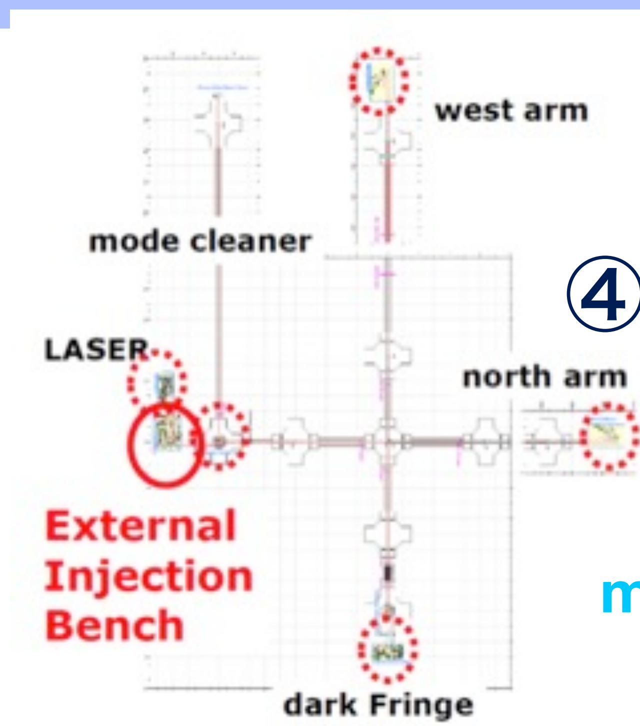
[Classical and Quantum Gravity 27, 19 (2010) 194011]

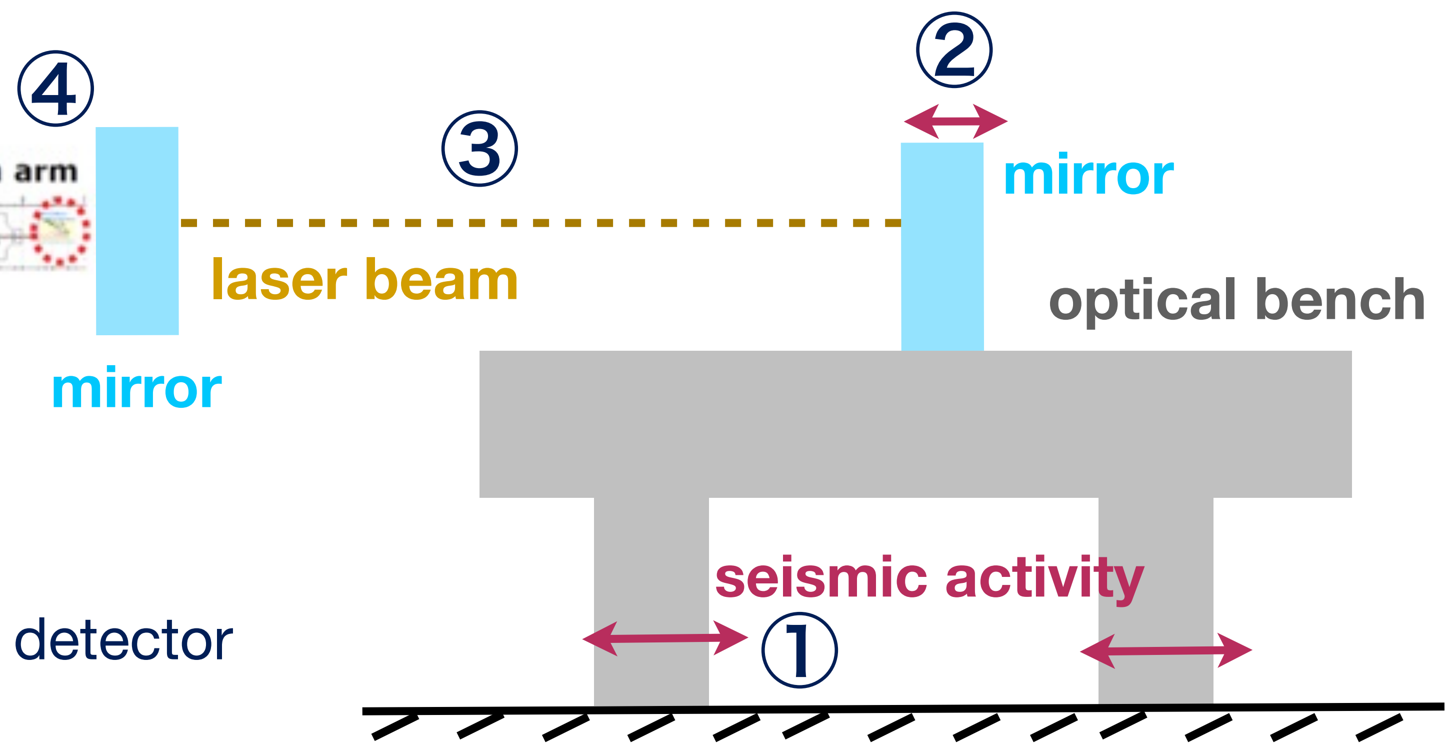


④. After modulated laser is returned to cavity, modulated laser will be noise source because of different phase.

Mechanism of this up-conversion

[Classical and Quantum Gravity 27, 19 (2010) 194011]





Optical system behind end-mirror controls GW detector using transmitted laser.
Sometimes accidentally transmitted laser is returned to cavity.

①. Strong seismic activity (such as microseism) excite resonant motion of optical bench and generate damping motion of optical bench.

②. damping motion of mirror installed on optical bench

③. time variation of optical path length between end-mirror and mirror on optical bench because of damping motion of mirror on optical bench

④. After modulated laser is returned to cavity, modulated laser will be noise source because of different phase.

About noise model

[Classical and Quantum Gravity 27, 19 (2010) 194011]

up-conversion noise

$$h_{sc}(t) = G \cdot \sin\left(\frac{4\pi}{\lambda}(x_0 + \delta x_{sc(t)})\right)$$

x_0 : distance between end mirror and reflector

$\delta x_{sc}(t)$: displacement of mirror by seismic activity

G : parameter depending on interferometer $(G = 5 \times 10^{-20})$
λ : laser wavelength (1064 [nm])

GW channel : $s(t) = h_{sc}(t) + n(t)$

n(t) : detector fundamental noise
(Virgo sensitivity is used. Assuming gaussian and stationary noise)

displacement of mirror excited by seismic activity

$$\delta x_{sc}(t) = A_m \sin(2\pi f_m t) \exp(-t/\tau) + n_{seis}(t)$$

A_m : amplitude of mirror's displacement

$f_m = 15[\mathrm{Hz}]$: resonant frequency of optical bench

$\tau = 0.1[\mathrm{sec}]$: damping time
(estimated from Virgo paper)

$n_{seis}(t)$: background motion of mirror,
Assuming gaussian and stationary noise and S(f) = 10^{-8}[m/sqrtHz]

About noise model

[Classical and Quantum Gravity 27, 19 (2010) 194011]

up-conversion noise

$$h_{sc}(t) = G \cdot \sin\left(\frac{4\pi}{\lambda}(x_0 + \delta x_{sc(t)})\right)$$

x_0 : distance between end mirror and reflector

$\delta x_{sc}(t)$: displacement of mirror by seismic activity

G : parameter depending on interferometer $(G = 5 \times 10^{-20})$

λ : laser wavelength (1064 [nm])

Whether correlation is linear or non-linear depends on this term.

$\delta x_{sc}(t) << \frac{\pi}{\lambda} \simeq 10^{-7}$ => linear correlation

$\delta x_{sc}(t) >> 10^{-7}$ => non-linear correlation

About noise model

[Classical and Quantum Gravity 27, 19 (2010) 194011]

up-conversion noise

$$h_{sc}(t) = G \cdot \sin\left(\frac{4\pi}{\lambda}(x_0 + \delta x_{sc(t)})\right)$$

x_0 : distance between end mirror and reflector

$\delta x_{sc}(t)$: displacement of mirror by seismic activity

We can observe two channels

G : parameter depending on interferometer $(G = 5 \times 10^{-20})$
λ : laser wavelength (1064 [nm])

GW channel : $s(t) = h_{sc}(t) + n(t)$

n(t) : detector fundamental noise
(Virgo sensitivity is used. Assuming gaussian and stationary noise)

displacement of mirror excited by seismic activity

$$\delta x_{sc}(t) = A_m \sin(2\pi f_m t)\exp(-t/\tau) + n_{seis}(t)$$

Following this noise model, simulation noise is generated.
=> Using correlation analysis methods (Pearson and MIC),

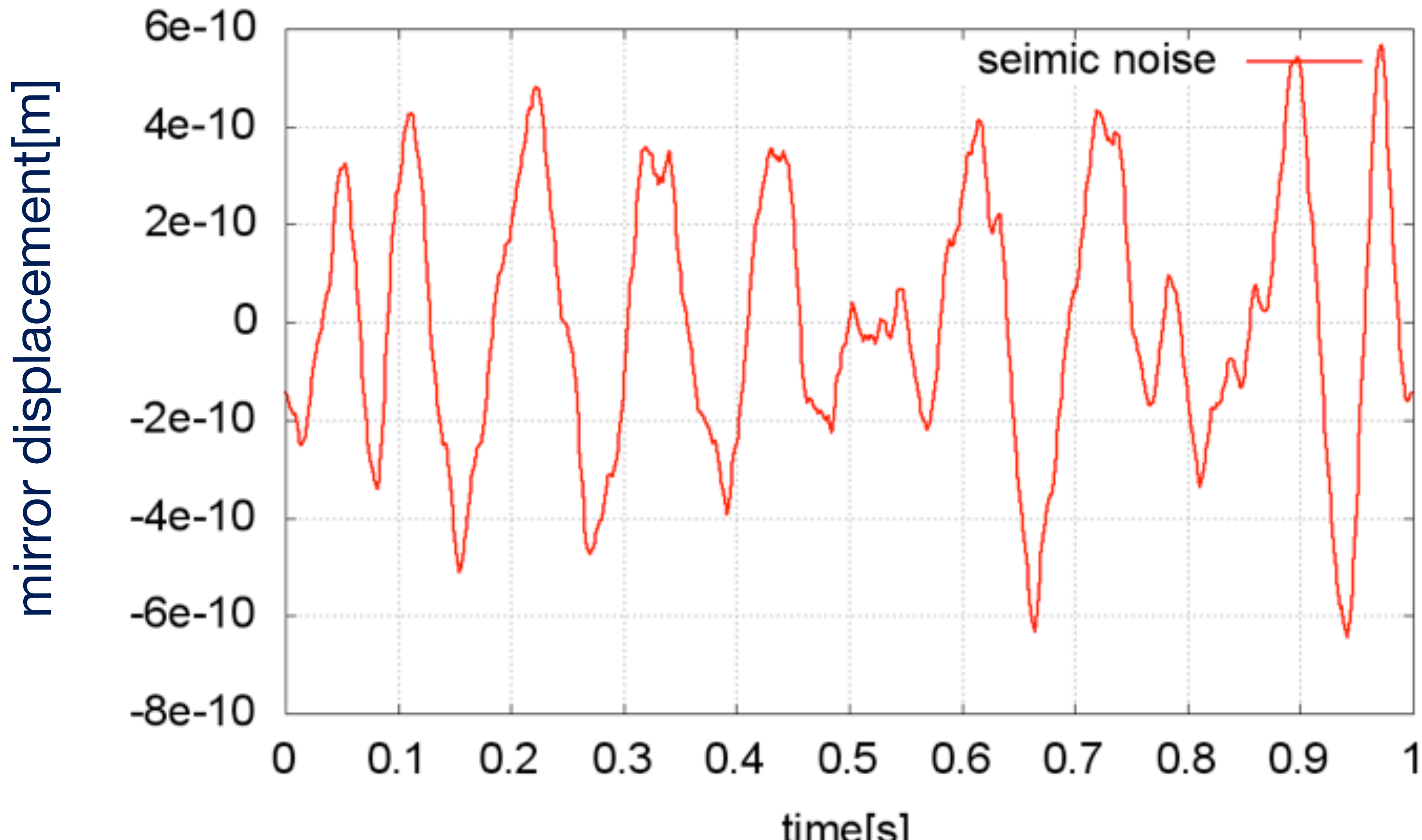
We check the performance of analysis methods for non-linear noise.

Mirror displacement generated by simulation

simulation condition : data duration = 1[sec], sampling rate = 1024[Hz]

In the case of mirror displacement $A_m = 0[\mathrm{m}]$

=> no-correlation case
We think this case as background.

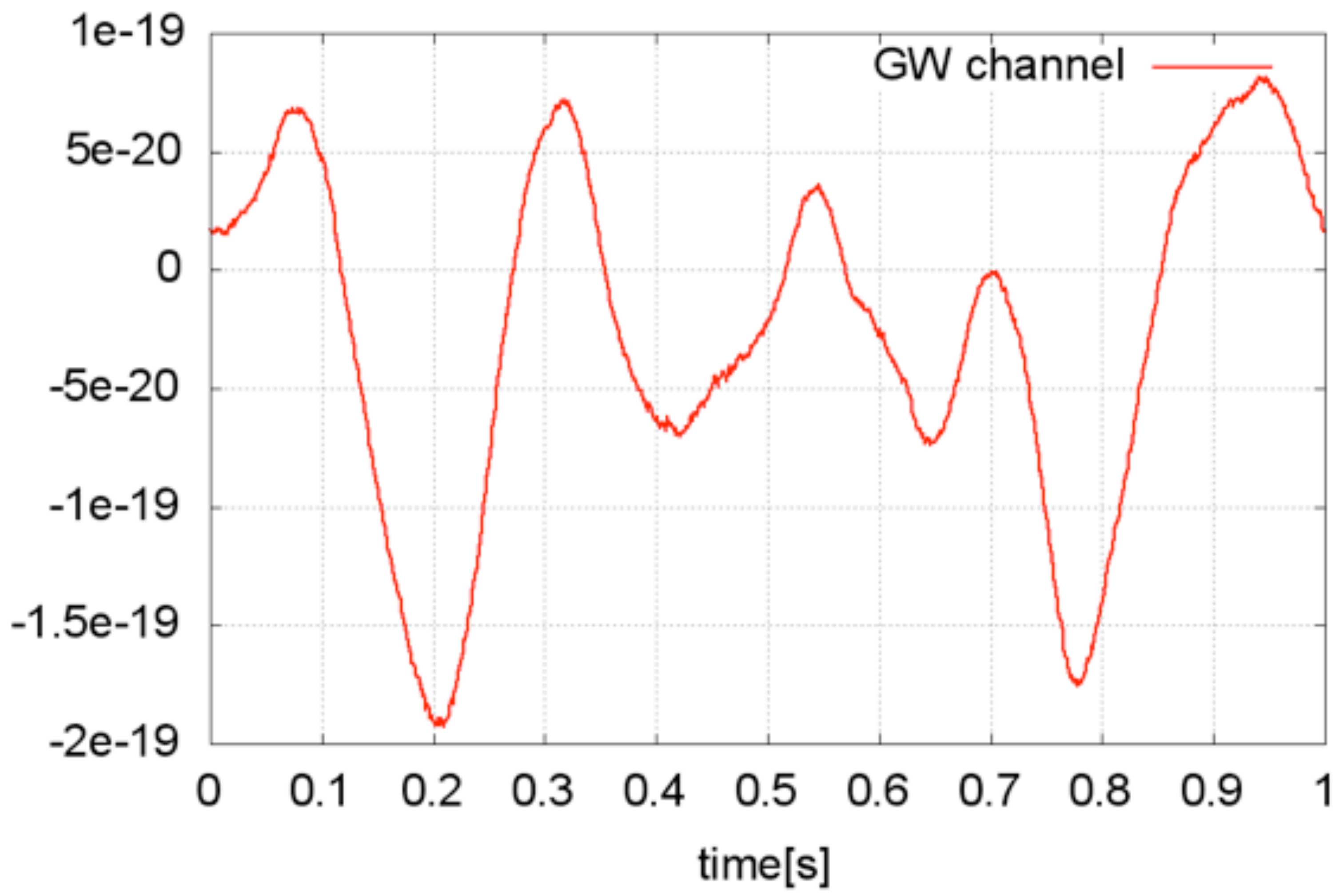


GW channel generated by simulation

In the case of mirror displacement $A_m = 0[\mathrm{m}]$

GW channel

$$s(t) = h_{sc}(t) + n(t)$$

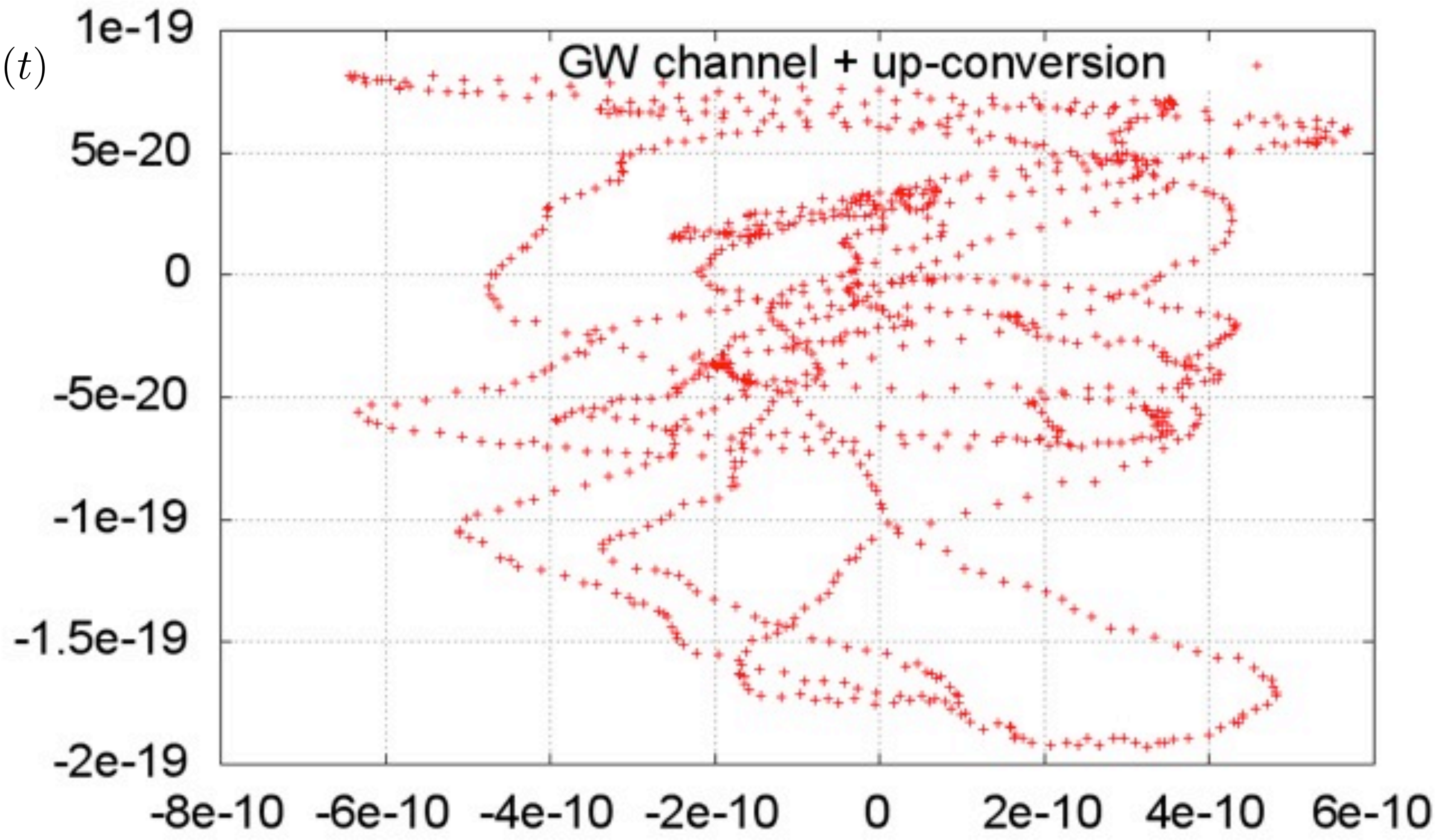


Scatter plot of mirror displacement and GW channel

In the case of mirror displacement $A_m = 0[\mathrm{m}]$

GW channel

$s(t) = h_{sc}(t) + n(t)$



displacement of mirror[m]

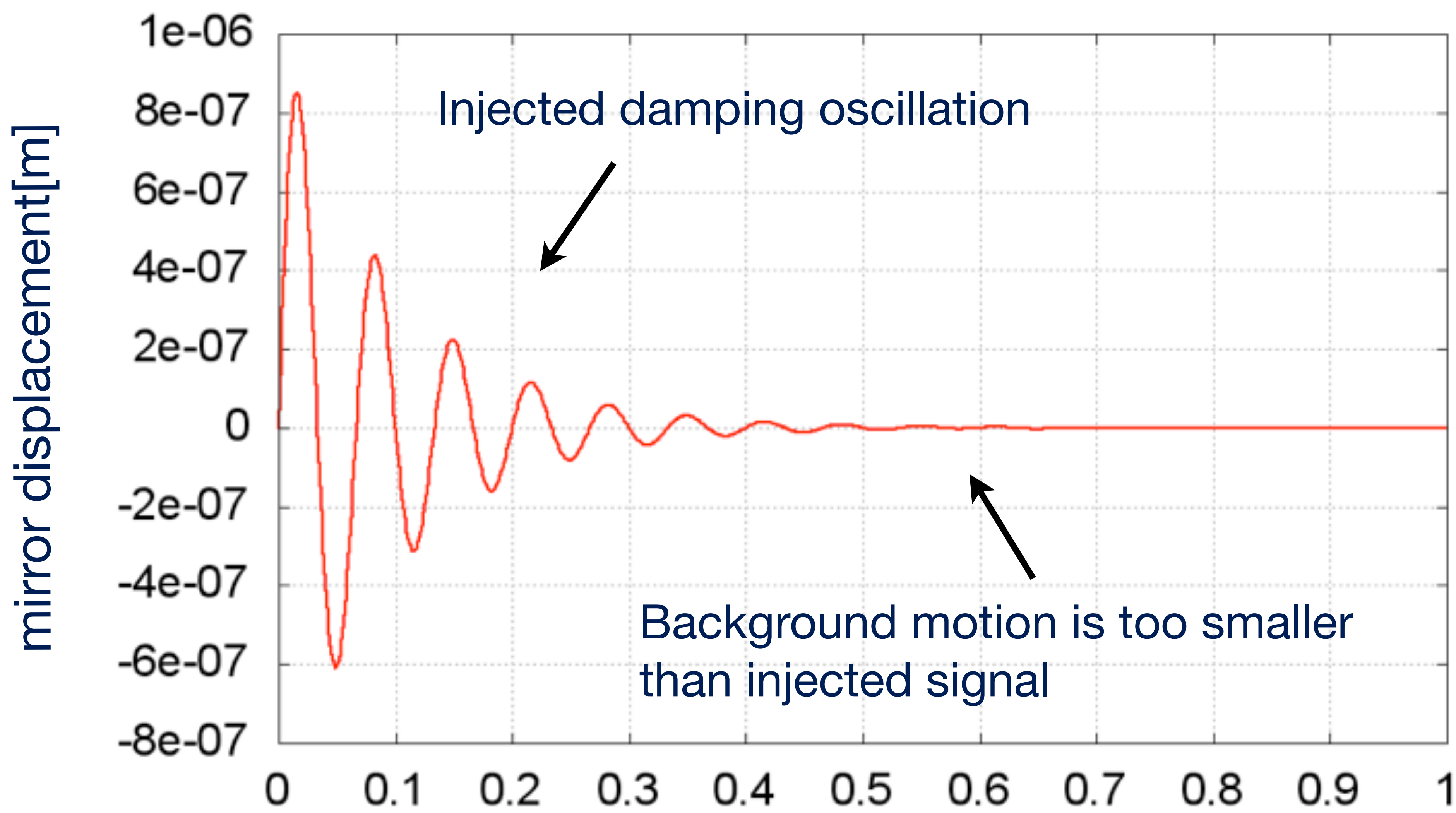
This scatter plot looks no-correlation.

Mirror displacement generated by simulation

In the case of mirror displacement $A_m = 10^{-6}[\mathrm{m}]$

$\delta x_{sc}(t) >> 10^{-7}$

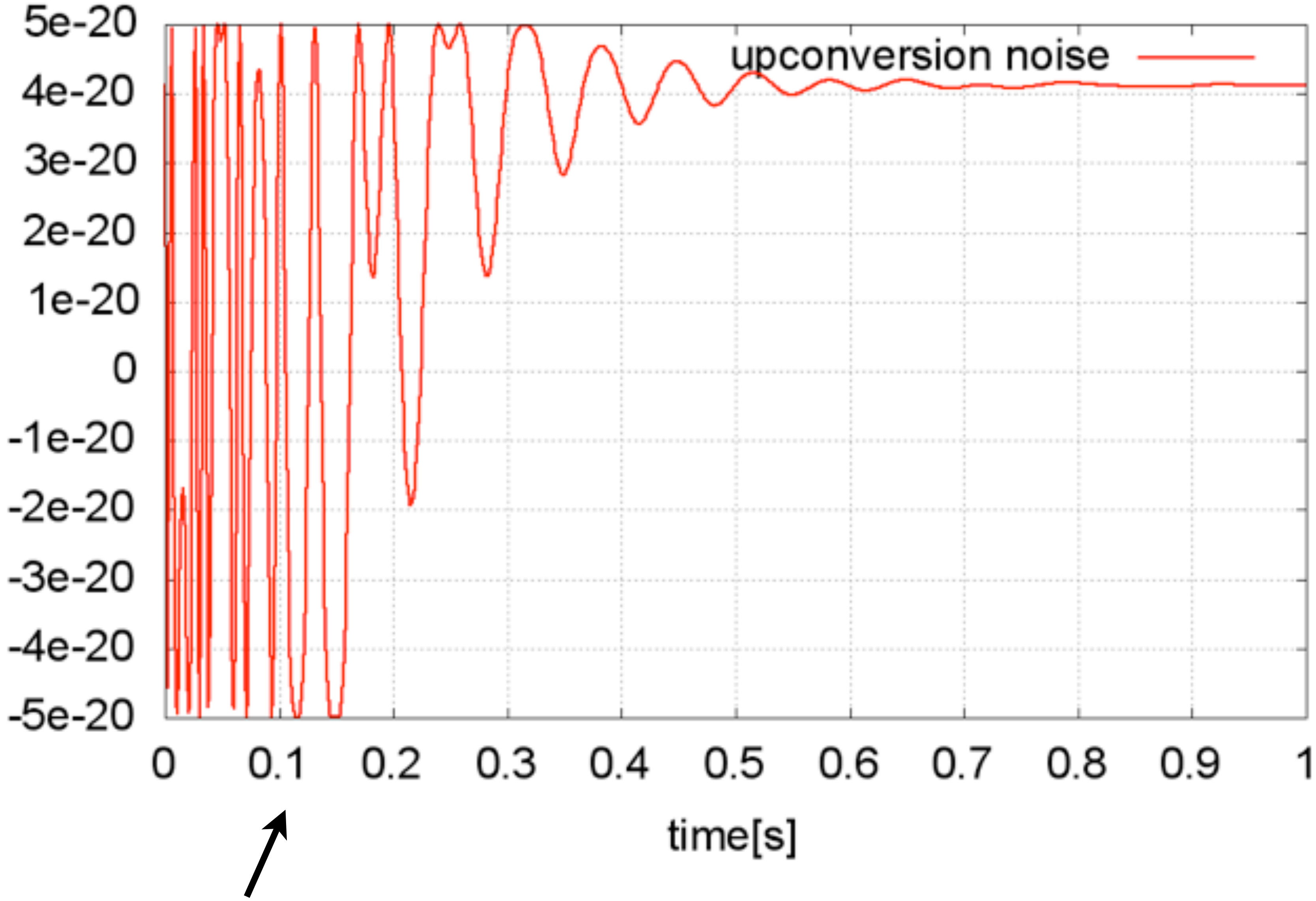
=> non-linear correlation case



Scattering light noise generated by simulation

In the case of mirror displacement $A_m = 10^{-6}[\mathrm{m}]$

scattering light noise $h_{sc}(t)$



Frequency changes with time.

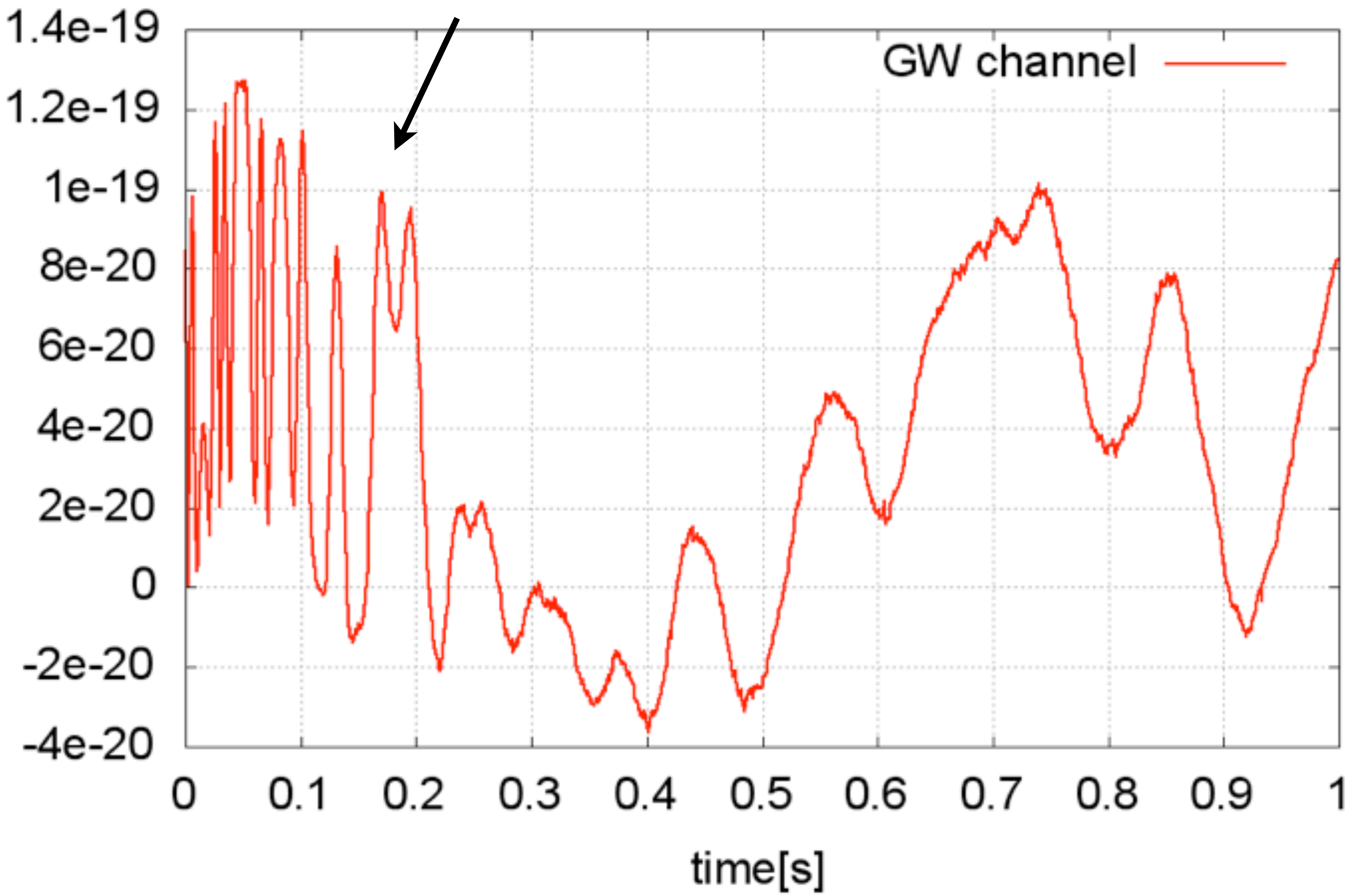
GW channel generated by simulation

In the case of mirror displacement $A_m = 10^{-6}[\mathrm{m}]$

GW channel

$s(t) = h_{sc}(t) + n(t)$

Effect from injected damping oscillation is still remain.

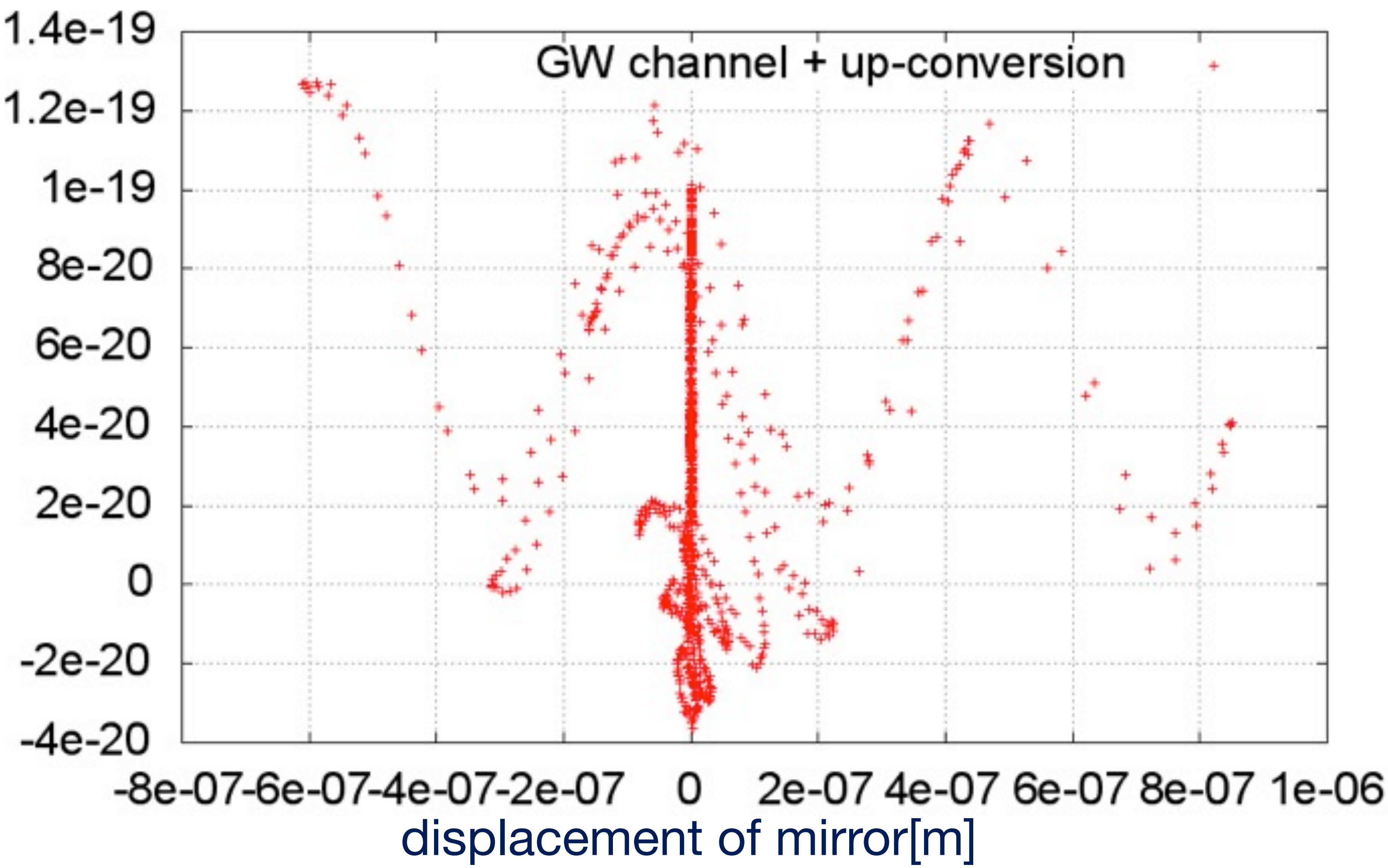


Scatter plot of mirror displacement and GW channel

In the case of mirror displacement $A_m = 10^{-6}[\mathrm{m}]$

GW channel

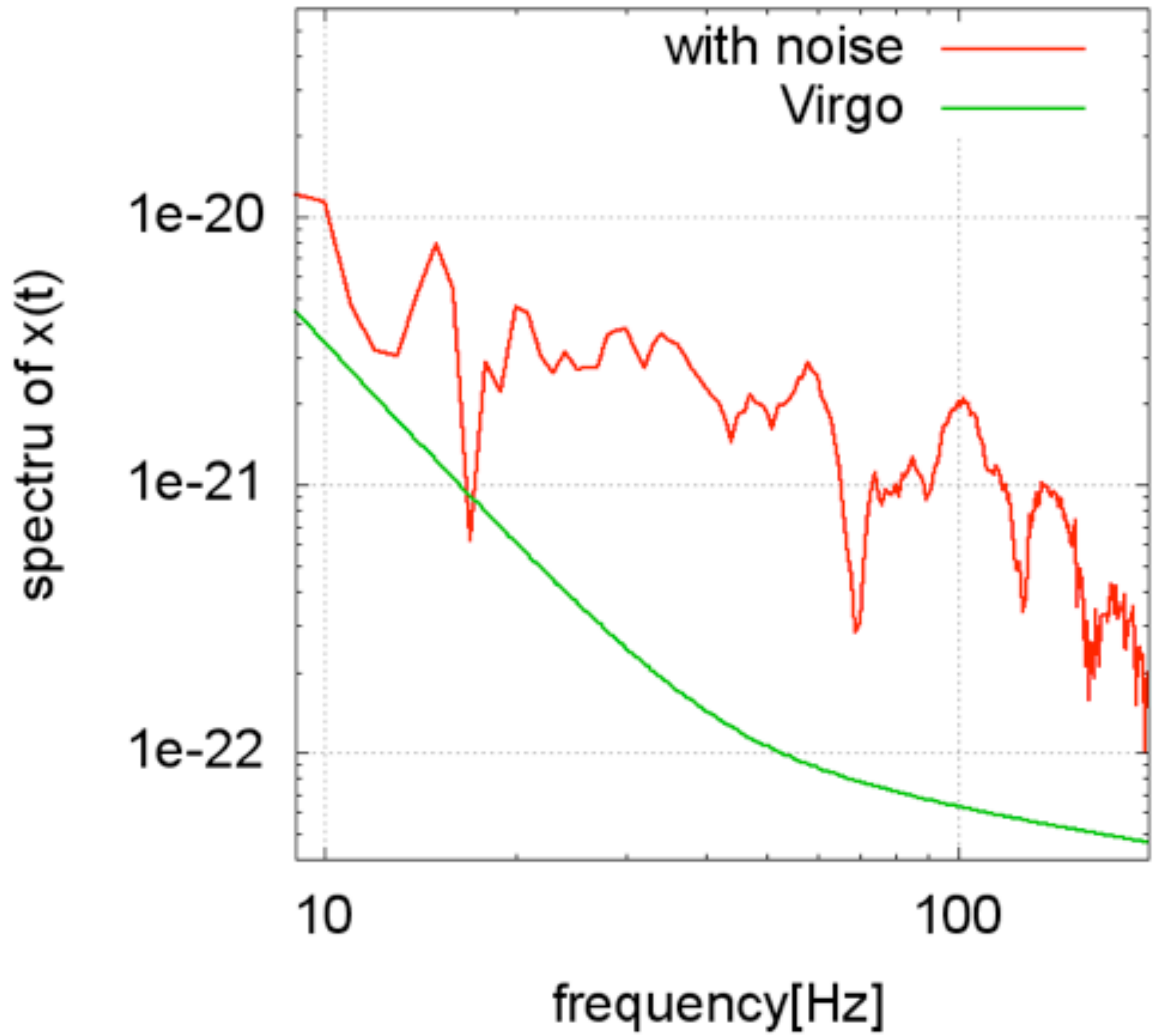
$s(t) = h_{sc}(t) + n(t)$



There is structure of sine-shape which is cased by injected dumping oscillation. => non-linear correlation

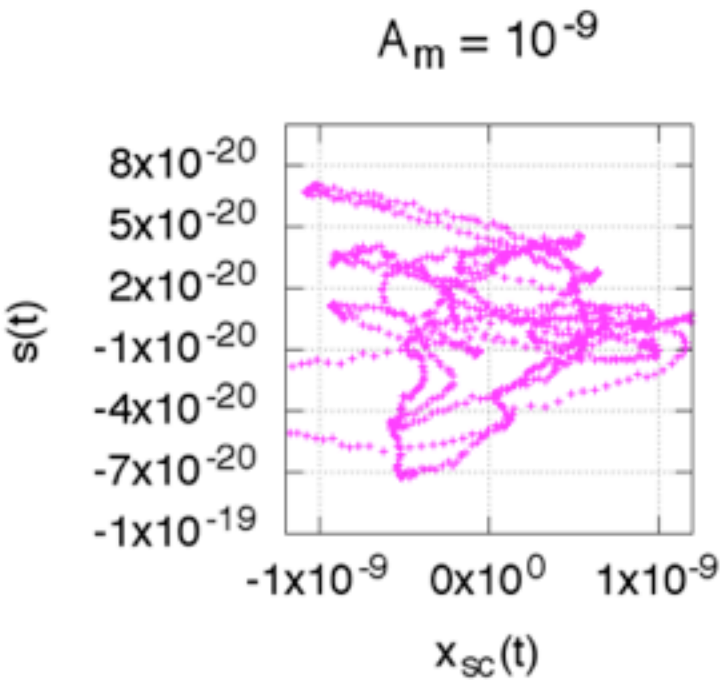
Spectrum of GW channel generated by simulation

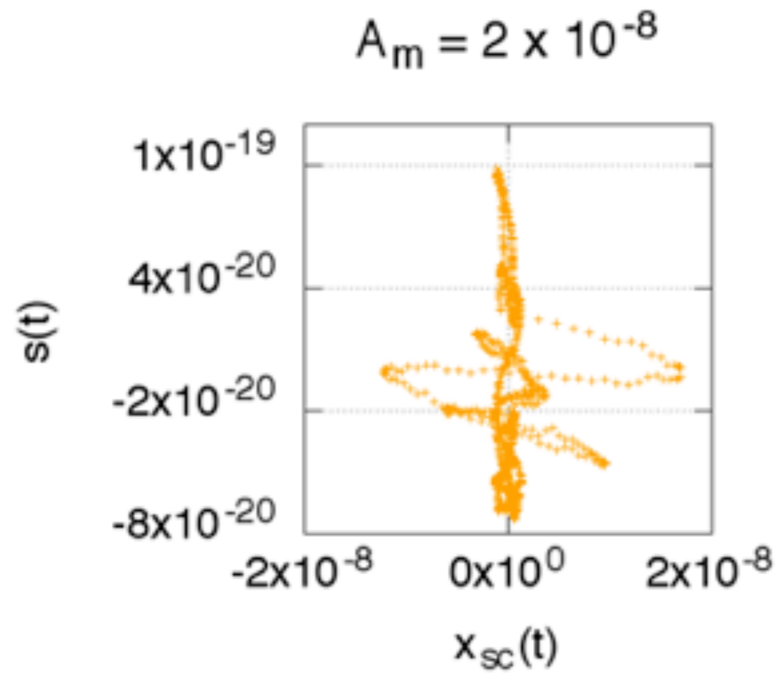
In the case of mirror displacement $A_m = 10^{-6}[\mathrm{m}]$

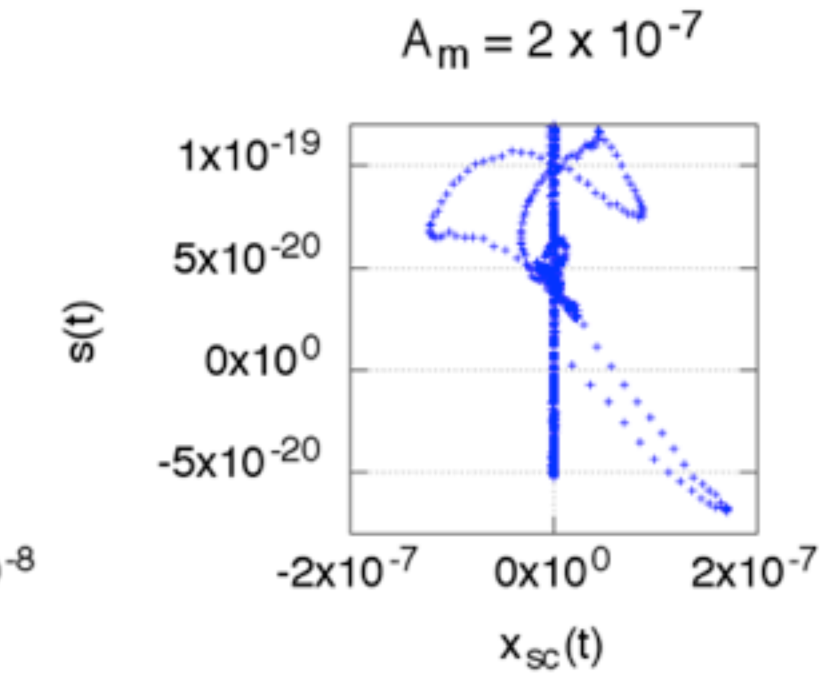


Red line is detector noise when up-conversion noise happened.
Comparing with stable detector noise(Green line), sensitivity becomes worth.

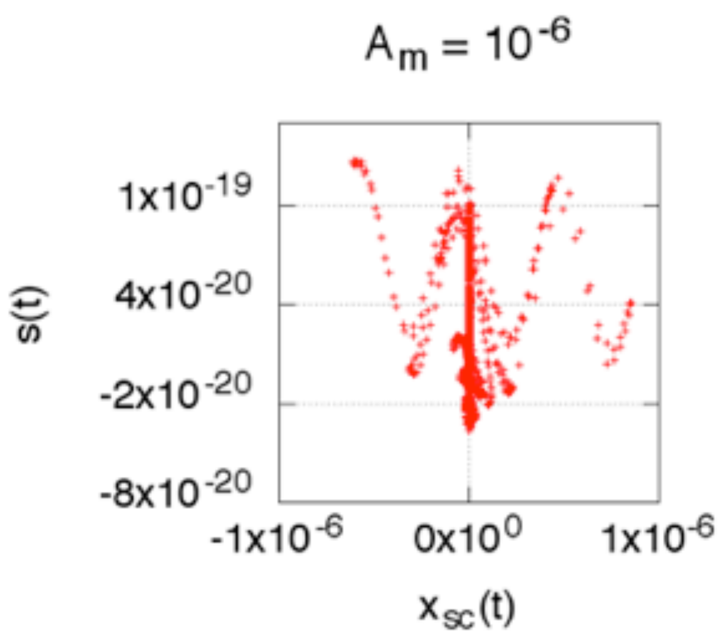
Displacement of mirror and scatter plot

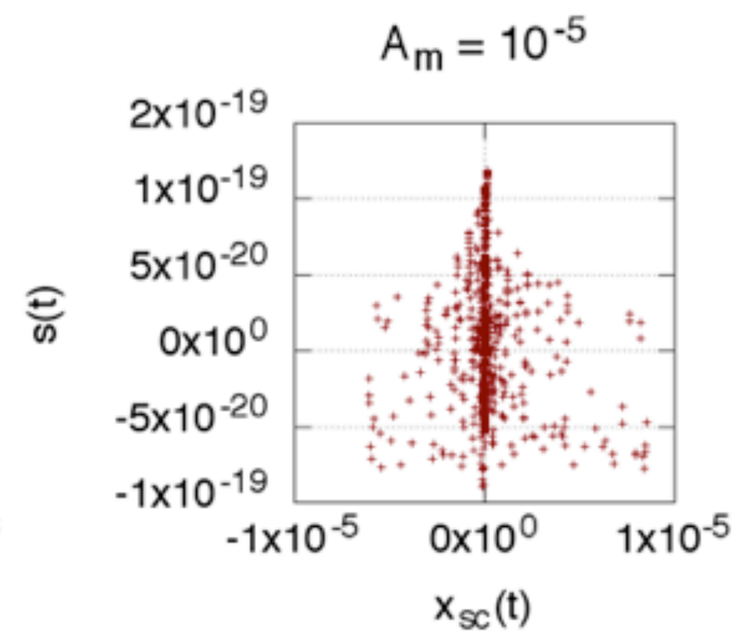






Weak Seismic Activity



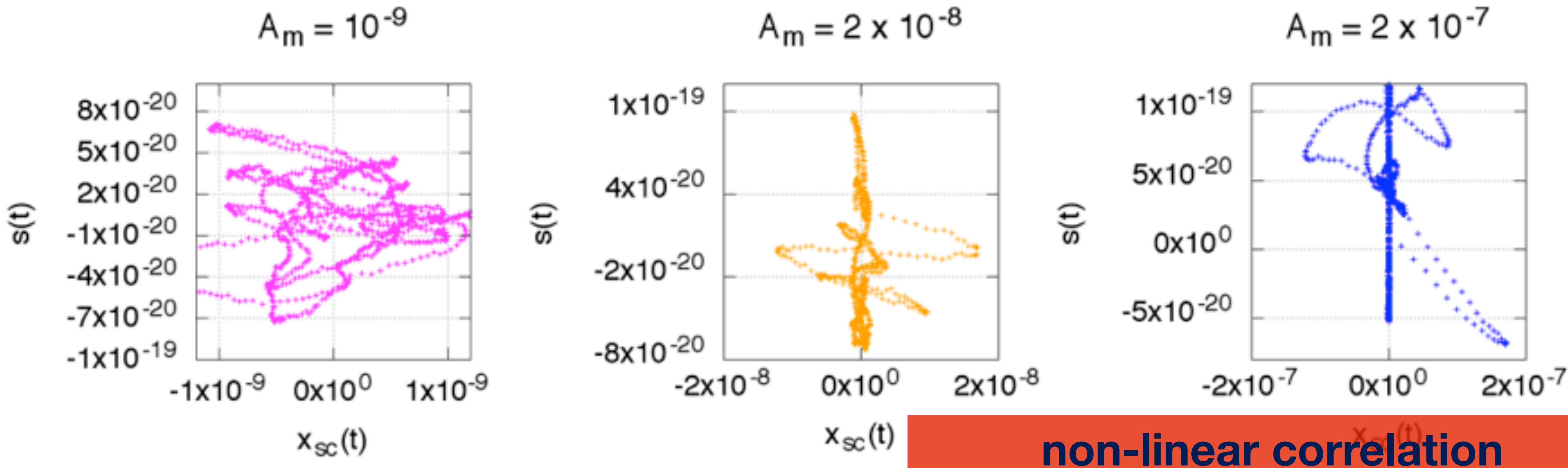


x-axis : displacement of mirror
y-axis : gravitational wave channel

Strong Seismic Activity

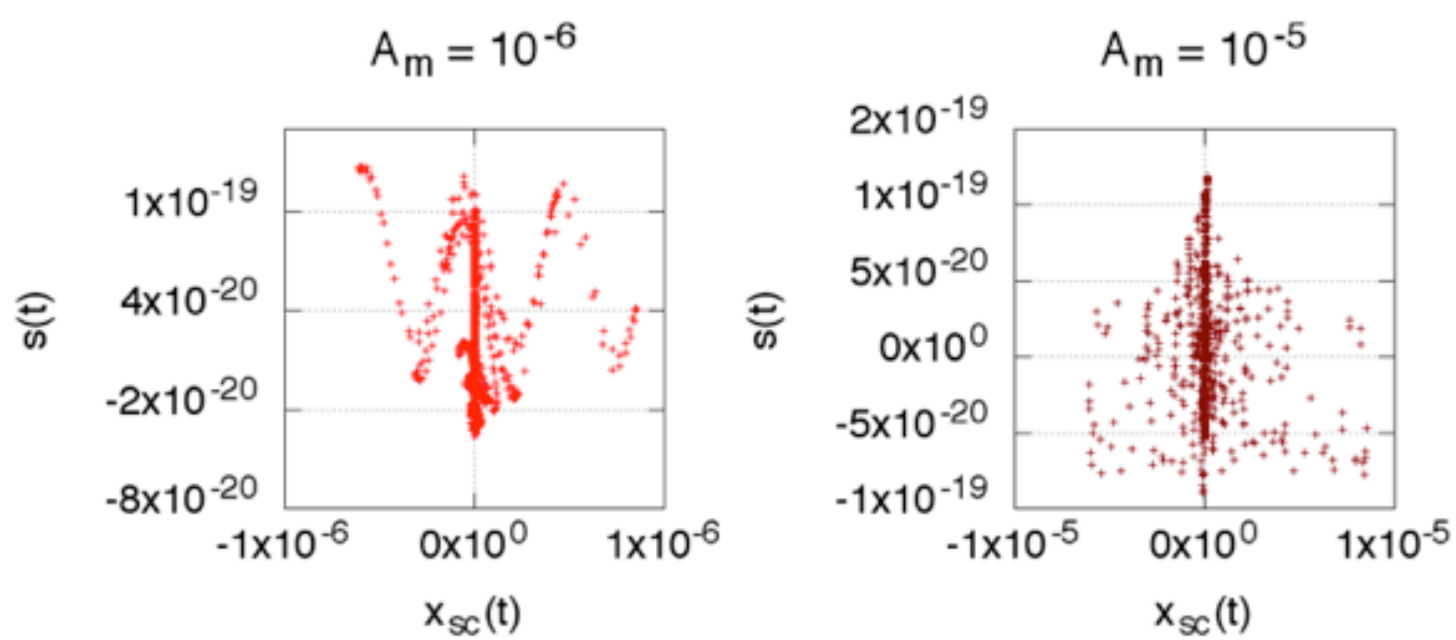
Classification of scatter plot

no correlation **linear correlation**



non-linear correlation (sine-shape)

Weak Seismic Activity

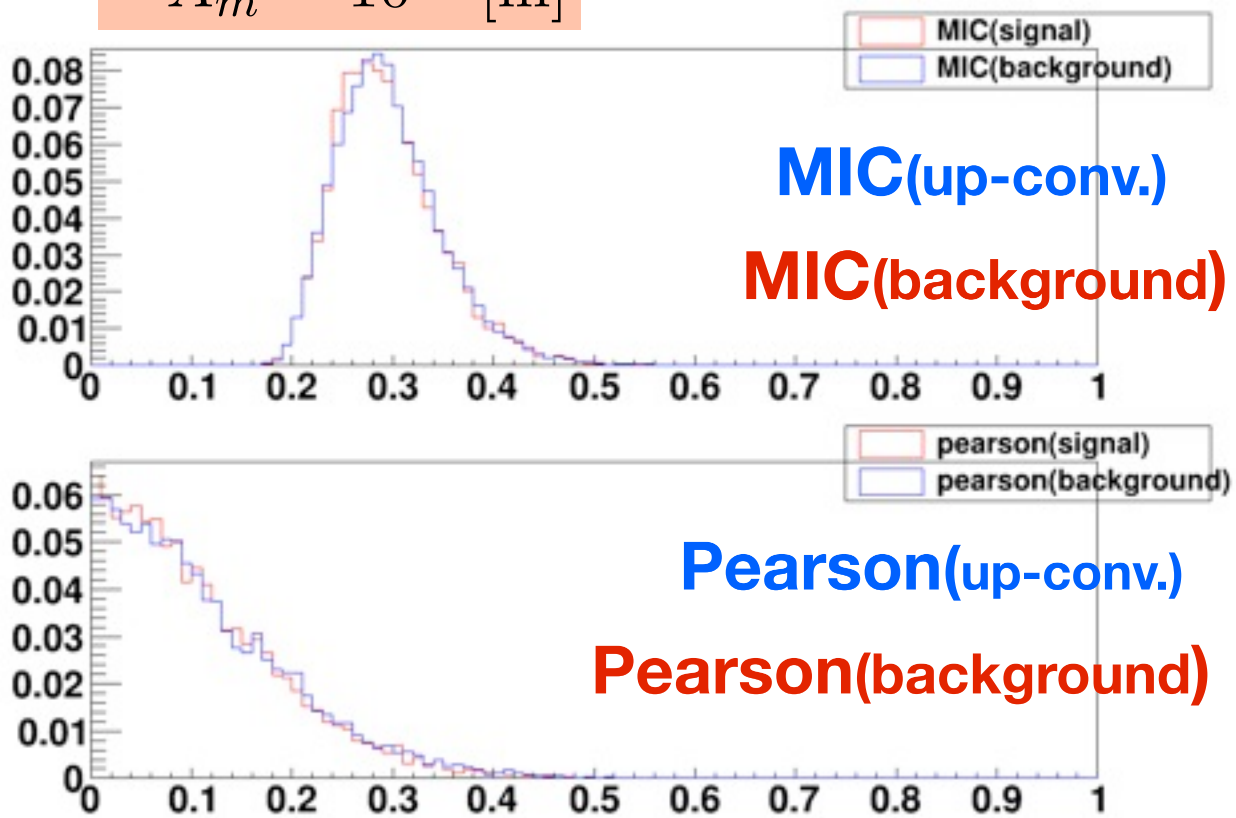


x-axis : displacement of mirror
y-axis : gravitational wave channel

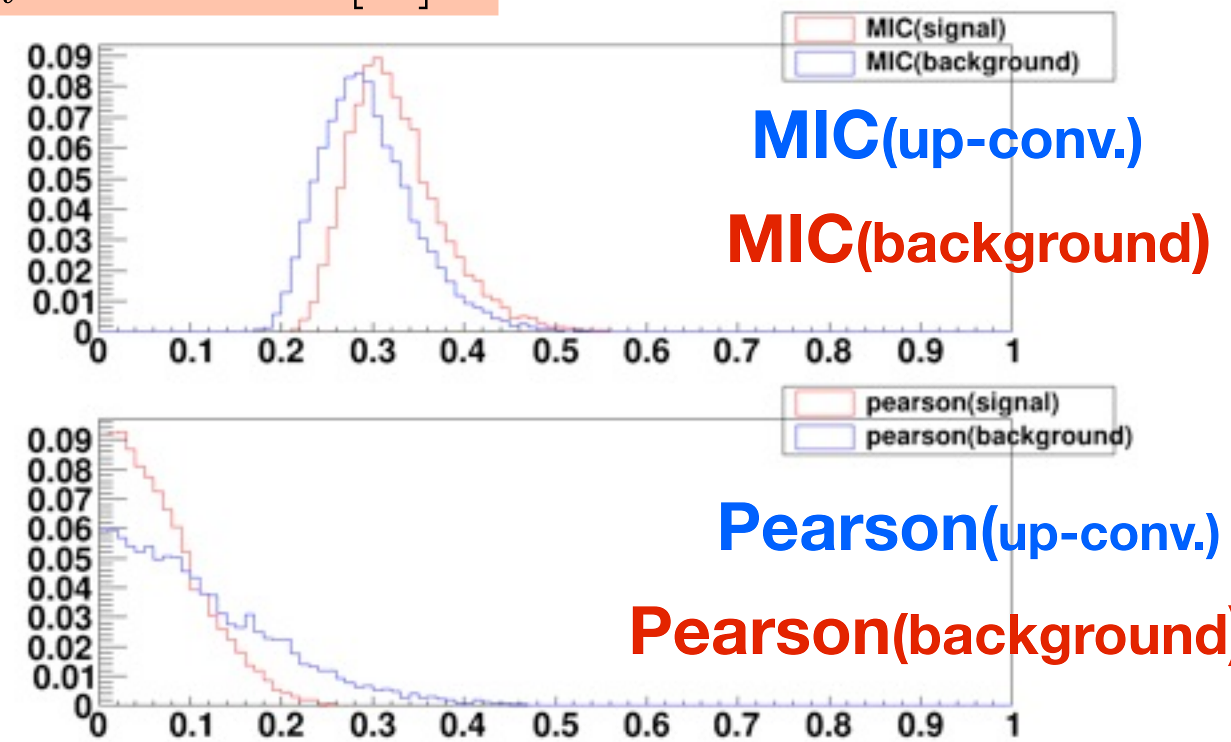
Strong Seismic Activity

Histogram of correlation value calculated by simulation(10000 trials)

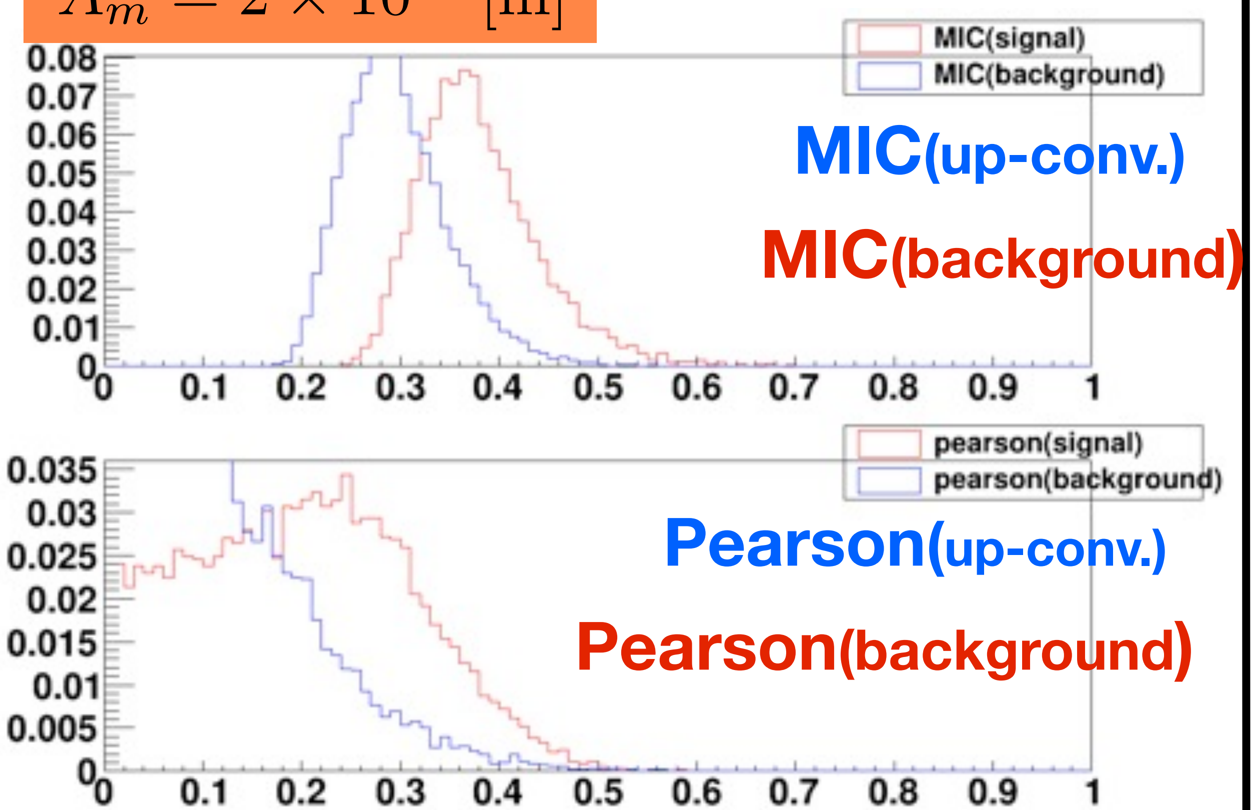
$A_m = 10^{-9}[\mathrm{m}]$



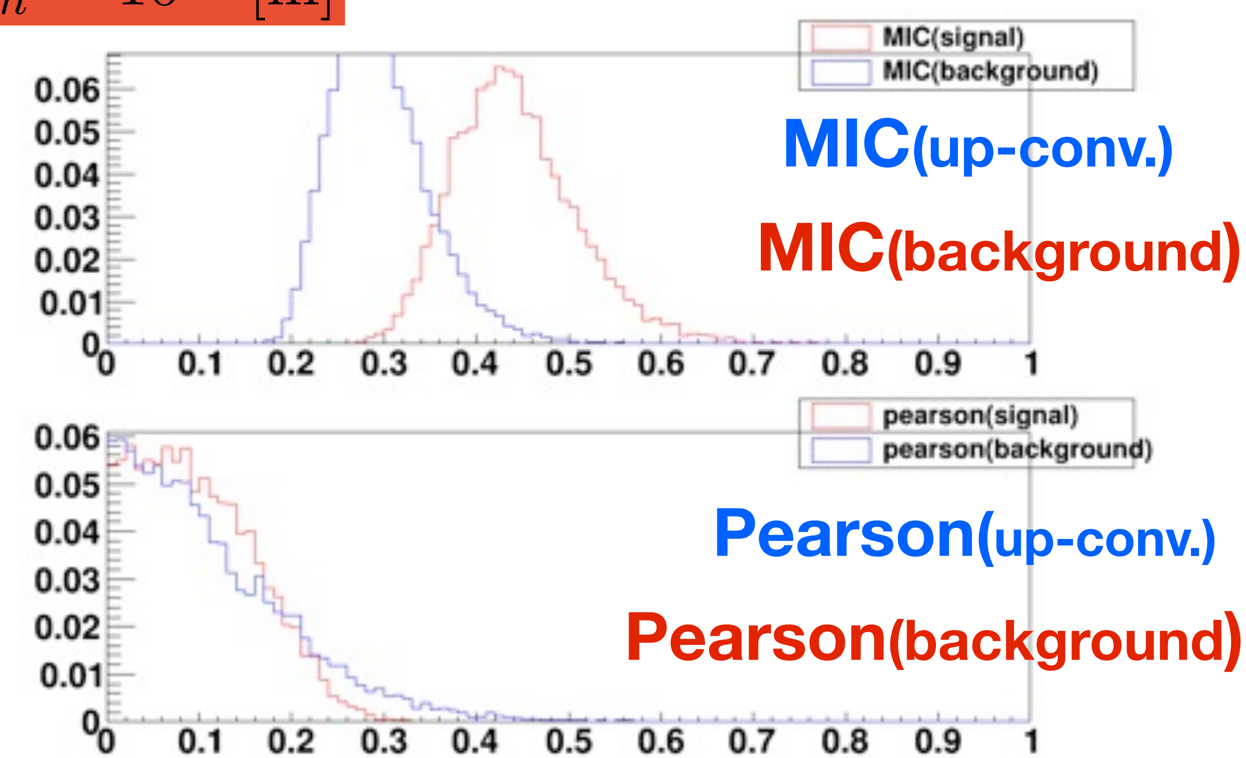
$A_m = 2 \times 10^{-8}[\mathrm{m}]$



$A_m = 2 \times 10^{-7}[\mathrm{m}]$



$A_m = 10^{-6}[\mathrm{m}]$



Calculation of ROC curve

For each histogram, we calculate false alarm probability(FAP) and detection efficiency at each threshold x_{th},

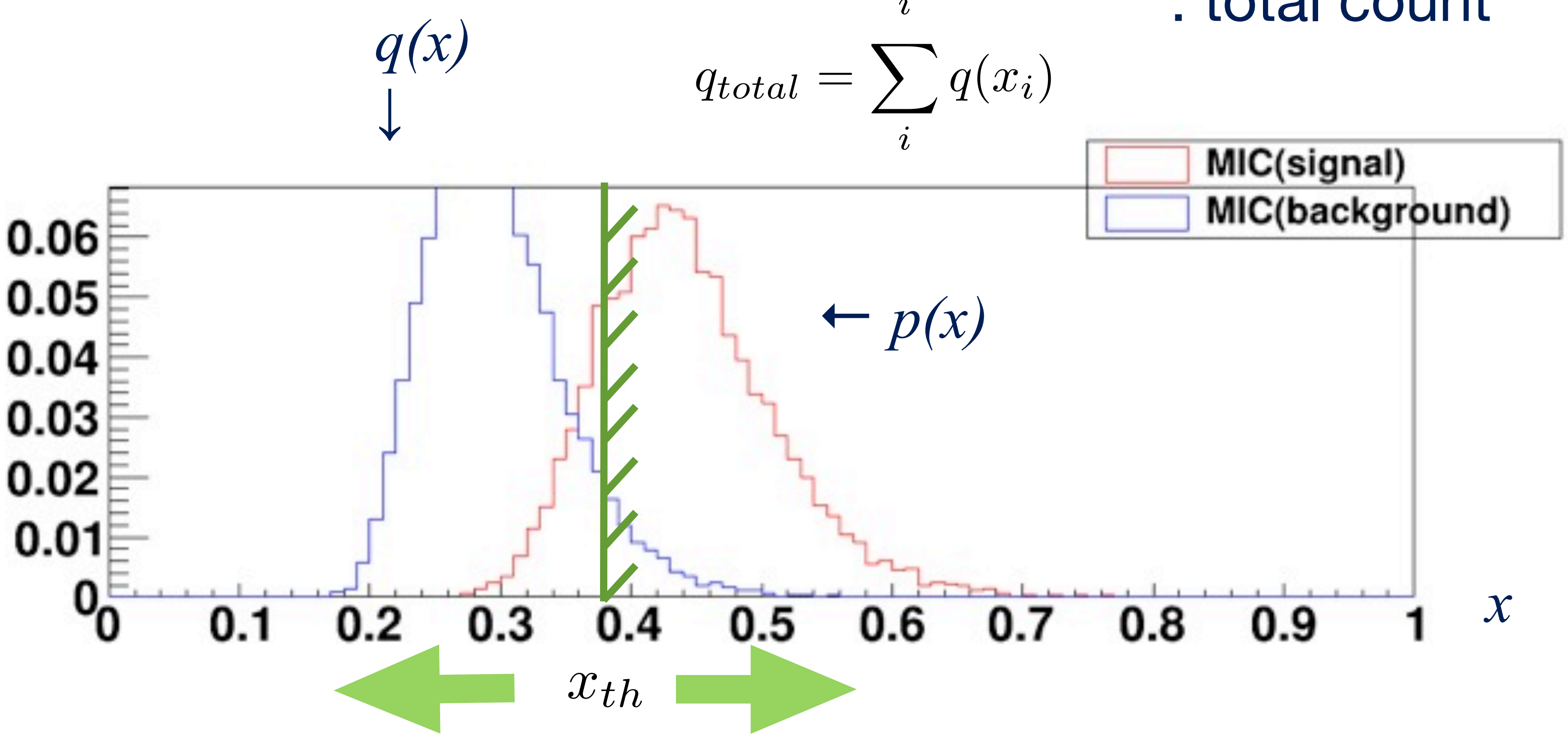
$$\text{false alarm probability}(x_{th}) = \sum_{x>x_{th}} p(x_i)/p_{total}$$

$$\text{efficiency}(x_{th}) = \sum_{x>x_{th}} q(x_i)/q_{total}$$

$$p_{total} = \sum_i p(x_i)$$

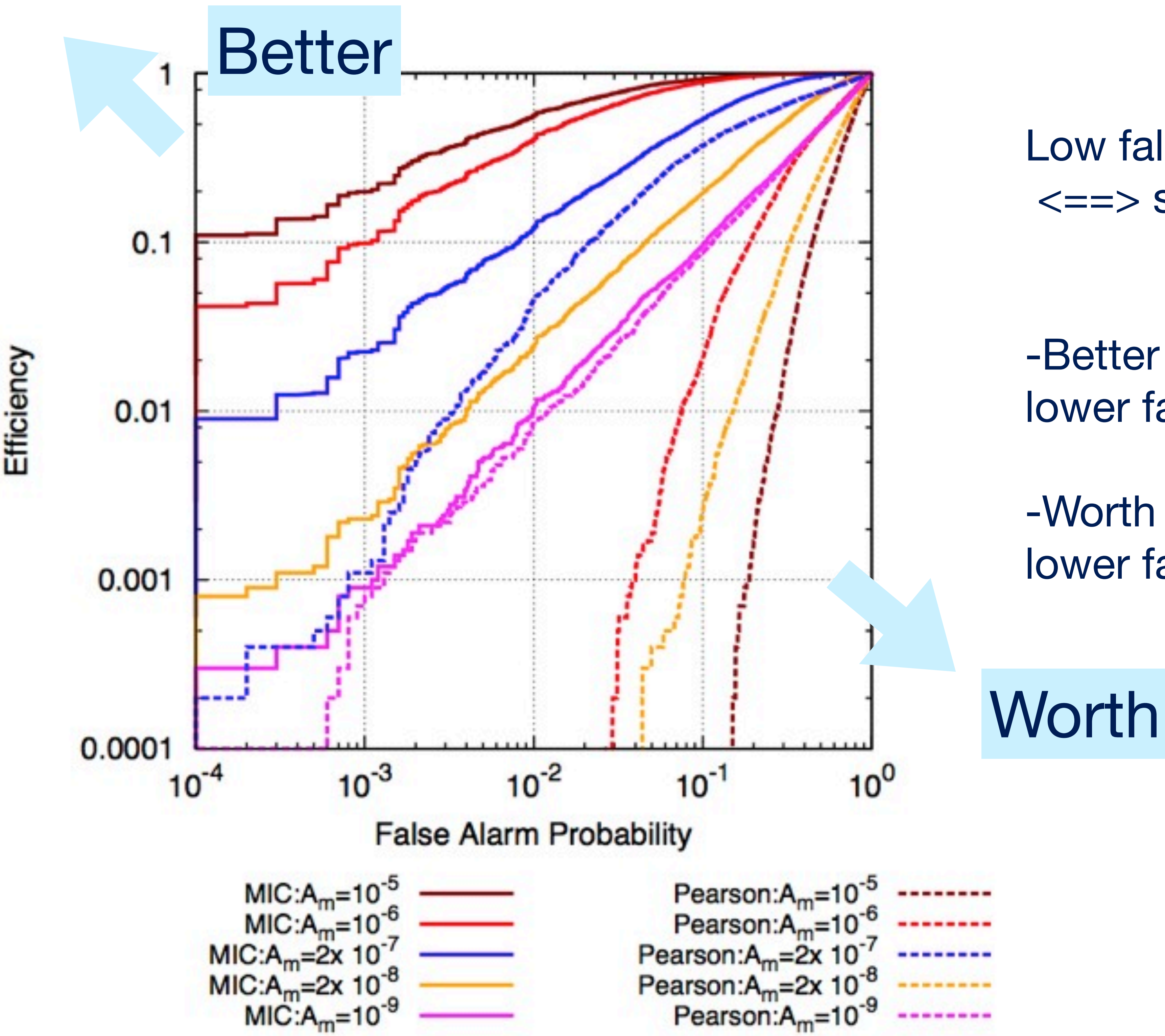
$$q_{total} = \sum_i q(x_i)$$

: total count



Using obtained FAP and efficiency, Receiver Operating Characteristic(ROC) curve can be calculated.

Evaluated performance of analysis methods - ROC curve

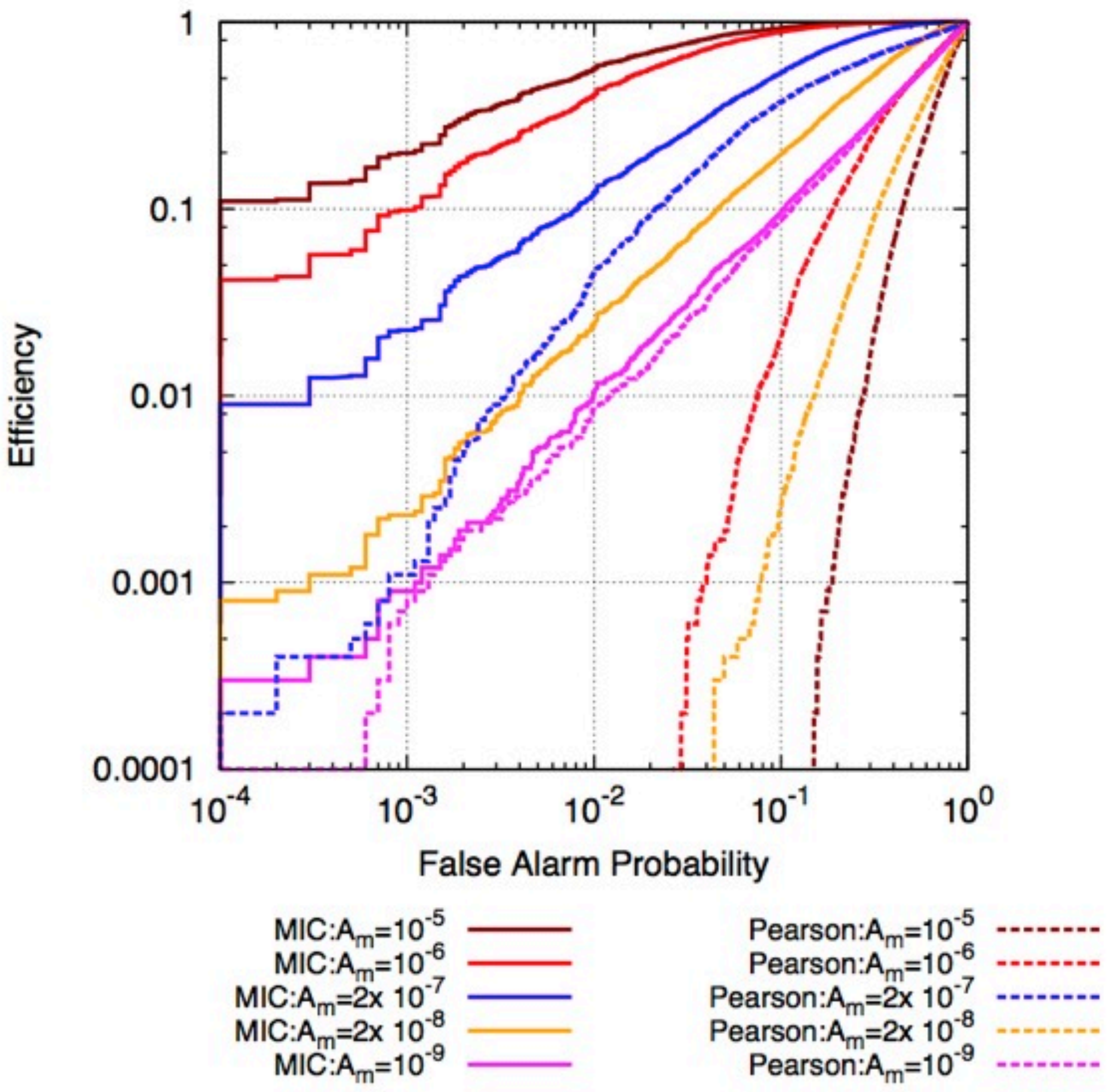


Low false alarm probability
<==> strict threshold

-Better curve is high efficiency with lower false alarm probability.

-Worth curve is low efficiency with lower false alarm probability.

Evaluated performance of analysis methods - ROC curve



- In this study, for non-linearly noise (Am>10^{-6}), MIC has better efficiency than Pearson.
For example, under FAP=0.01, efficiency of MIC is 0.41~0.56 and Pearson is less than 0.0001.

- As increasing optical bench displacement, efficiency of MIC increases, because shape of scatter plot change from no correlation to linear correlation and finally non-linear correlation.

- At Am=2x10^{-7}, Pearson has highest efficiency
(under FAP=0.01, efficiency~0.047)
because of linear correlation.

Summary

- Many linear and non-linear correlated noises are observed in LIGO and Virgo.
  These noise limit detector sensitivity.
  Correlation analysis in this study will localize the noise source.
  As a result, false triggers generated GW search pipeline can be removed and increase detection efficiency.

- In KAGRA’s commissioning and observation phase, localization of correlated noise is very important.

- We explained up-conversion noise of Virgo detector which is caused by strong seismic activity. This up-conversion noise is well-modeled.
  Following this noise model, simulation noise is generated and analyzed with two correlation analysis method(Pearson and MIC).
  We showed that, for non-linear noise, MIC has better efficiency than Pearson. Assuming FAP=0.01, efficiency of MIC is 0.41~0.56 and Pearson is less than 0.0001.

Virgo’s typical seismic activity

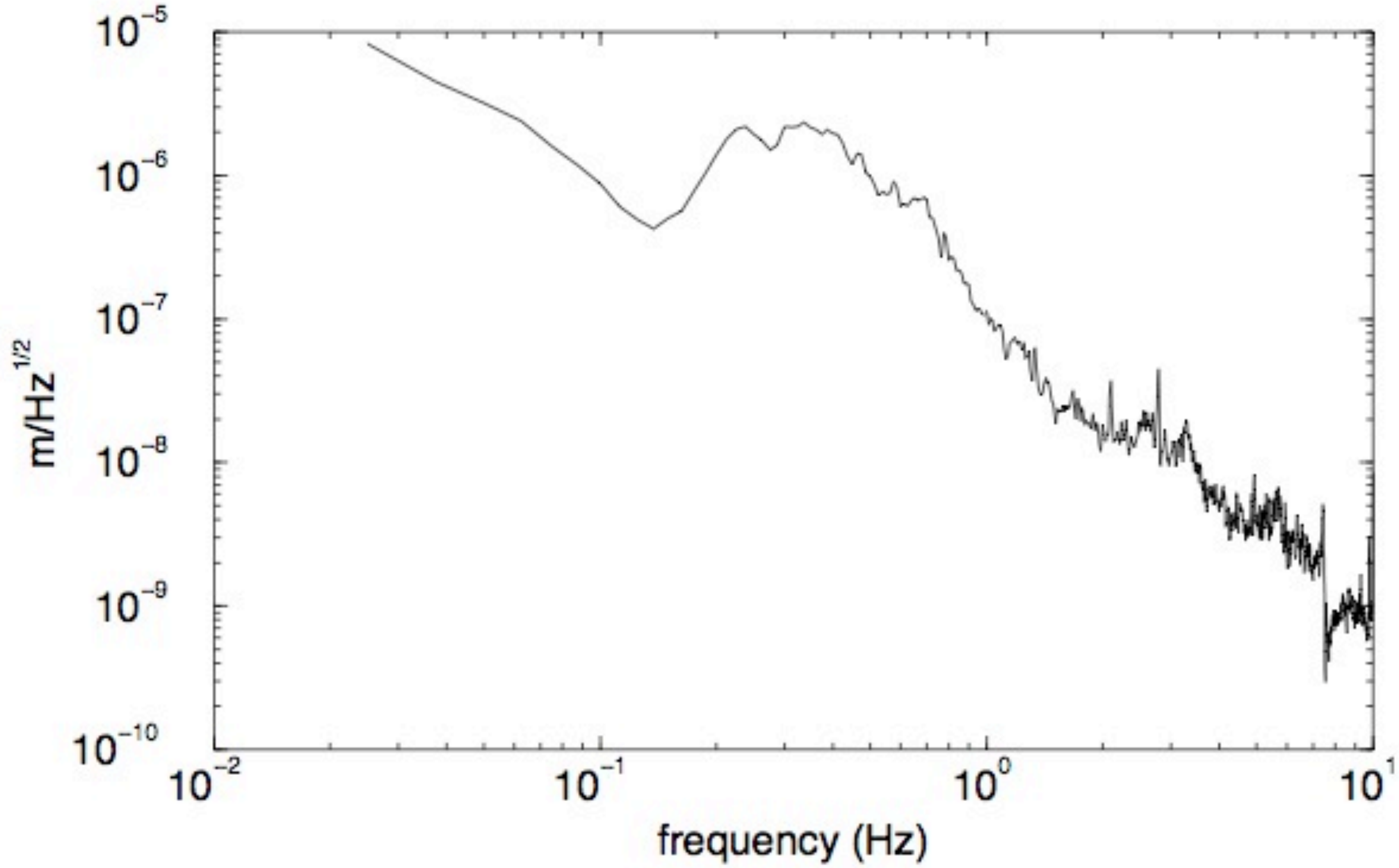
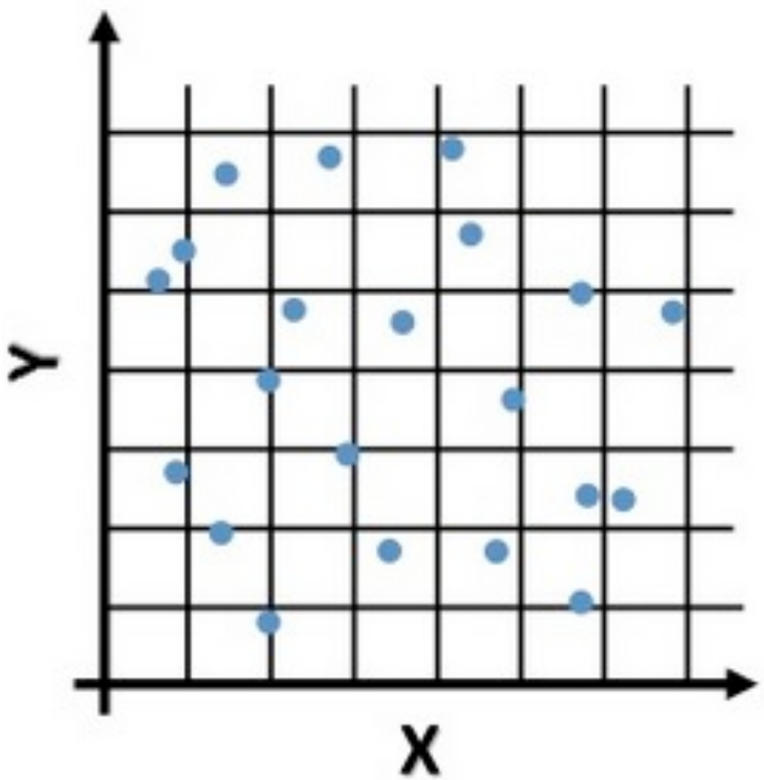
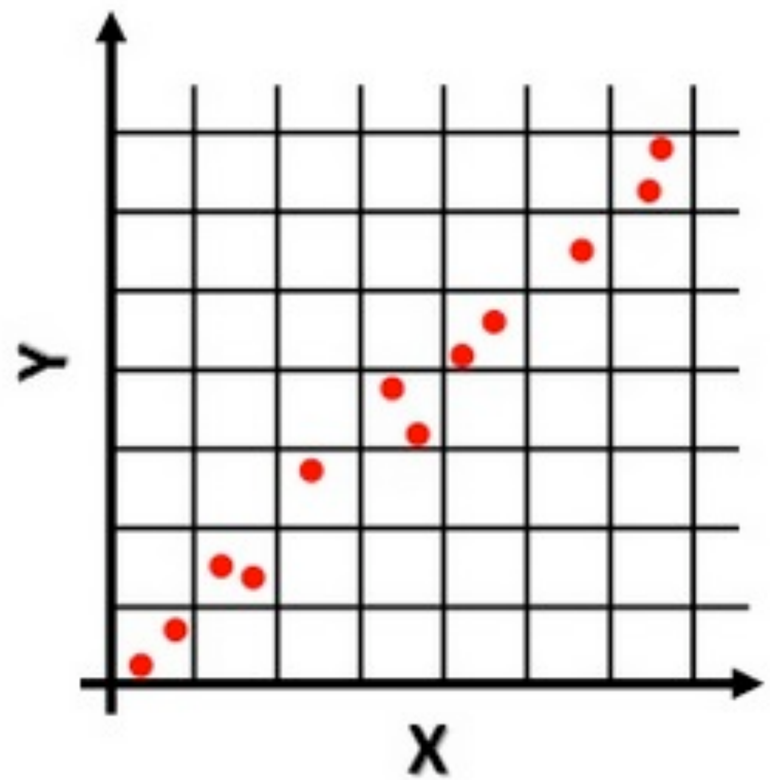


Figure 2.5: *Linear spectrum of the horizontal seismic displacement measured at Pisa INFN laboratories on a Sunday (low human activity, $\xi \sim 0.1$ while in weekdays $\xi \sim 0.3 \div 0.5$). The curve shows a rough $1/f^2$ slope above 1 Hz. The broad peak at 0.14 Hz is found all over the world and is due to the oceans activity.*

Maximum Information Coefficient (MIC)



相関なし

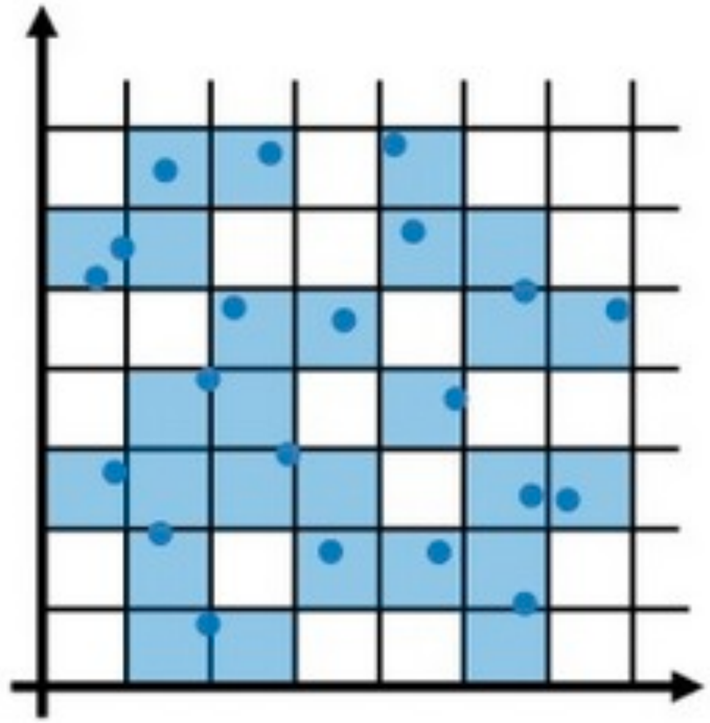


相関あり

MIC「もし2つの変数間に相関があるならデータを要約するように
データを分割するグリッドが引けるはず」

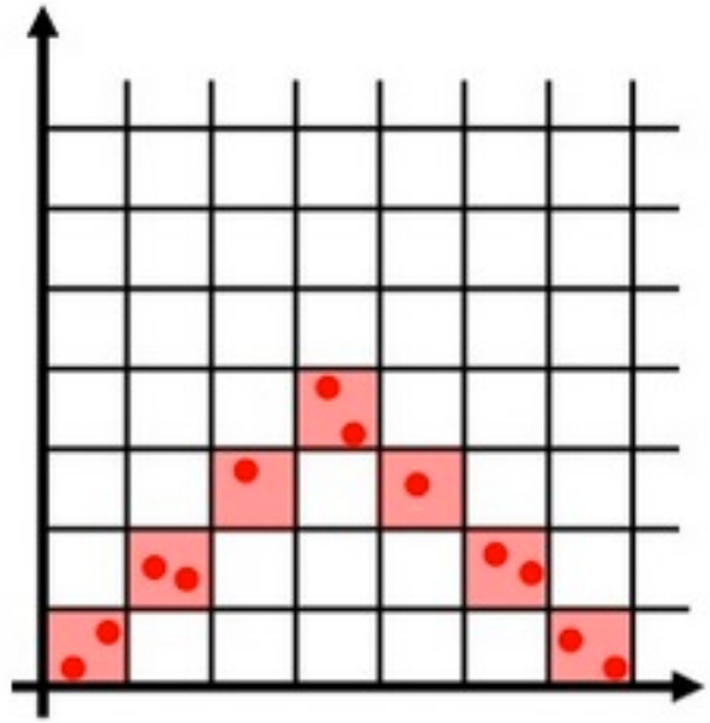
http://www.slideshare.net/logics-of-blue/mic-31810194

Maximum Information Coefficient (MIC)



マス目には関係なし

=> 相関なし



7つのマス目にデータが集中している

=> 相関あり

・定量的に表すために各マスに対して、相互情報量を計算する

$$I(x,y) = \sum\sum p(x,y)\log\frac{p(x,y)}{p(x)p(y)}$$

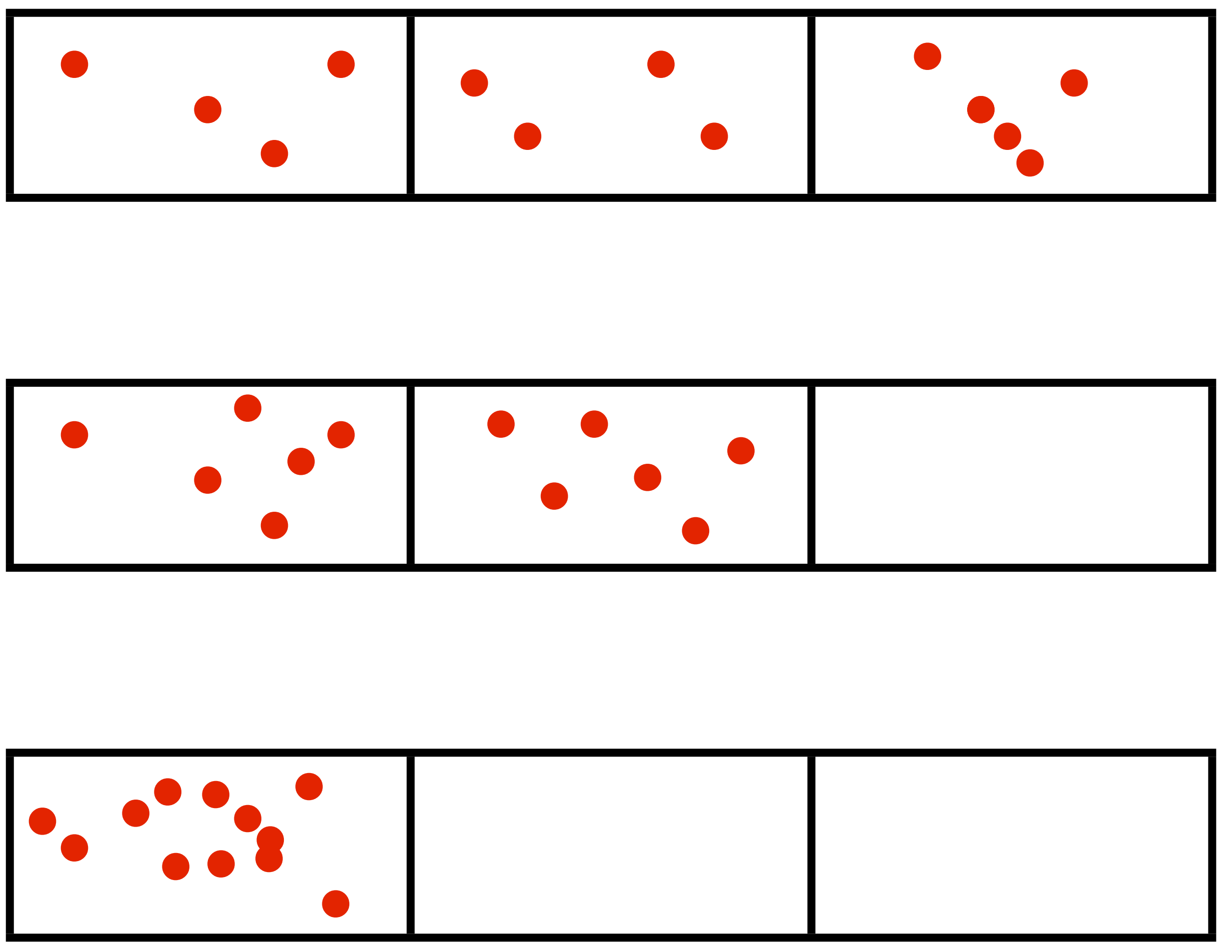
p(x), p(y) : x, yの確率密度関数

p(x, y) : x, yの同時確率密度関数

http://www.slideshare.net/logics-of-blue/mic-31810194

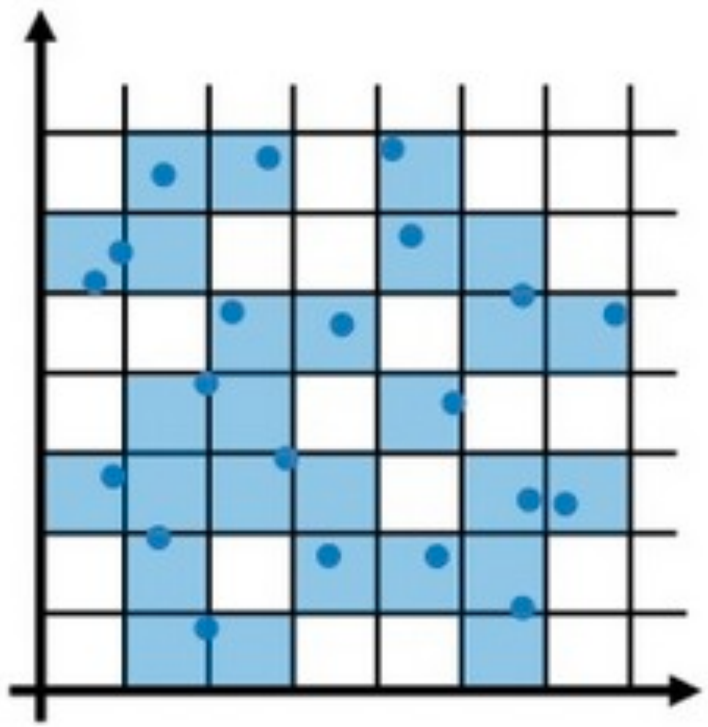
Maximum Information Coefficient (MIC)

相互情報量は不確実性がどれだけ少ないかを定量的に表す



	不確実性	相互情報量
①	大	小
②	中	中
③	小	大

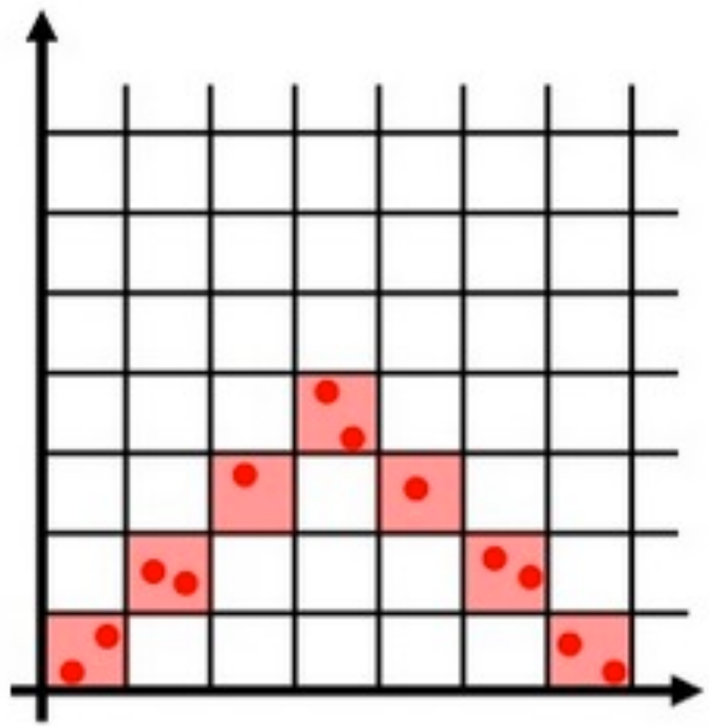
Maximum Information Coefficient (MIC)



マス目には関係なし

=> 相互情報量 小さい

=> 相関なし



7つのマス目にデータが集中している

=> 相互情報量 大きい

=> 相関あり

MICのgrid

- データをgridで各領域に分ける
- 領域ごとに相互情報量を求め、その合計が最大となるgridの組み合わせを探す
  => MICの出力
- $n_x \times n_y < n^{0.6}$ の範囲で