Data Analysis for LISA Pathfinder

lisa pathfinder

One man's gap is another man's experiment

M Hewitson for the LPF Team

Missing Data Workshop, Nice

11th May 2015

Contents

lisa pathfinder

- LPF data analysis overview
- Operations environment
- Examples of gaps or bad data segments
- Dealing with isolated glitches or data gaps
- Free-flight experiment
- Analysis methods for free-flight experiment

Science goals

lisa pathfinder

- Obtain the best geodesic motion possible
  - quietest differential acceleration of the two TMs
    - 3×10^{-14} m s^{-2} /$\sqrt{\text{Hz}}$ at 1 mHz
      - ~few pm accuracy position measurement of TM-SC, TM-TM
  - optimisation by changing system parameters
    - determine best configuration by experiments

- Develop a noise model of the system
  - allows the projection of the performance of technologies to LISA

Phases of Operations

lisa pathfinder

Launch | IOCR

Launch, LEOP, Transfer, Separation, De-spin	Commissioning	LTP Science Ops	DRS Commissioning	DRS Operations
60 days	14 days	3 months	10 days	3 months

	Day 1	Day 2	Day 3	Day 4
H1	Noise Run	Discharge	Noise Run	Discharge
H2	Noise Run	Working Point	Noise Run	Stray Potentials
H3	Noise Run	Working Point	Noise Run	Stray Potentials
H4	Sys ID	Working Point	Noise Run	Stray Potentials
H5	Sys ID		Noise Run	Stray Potentials

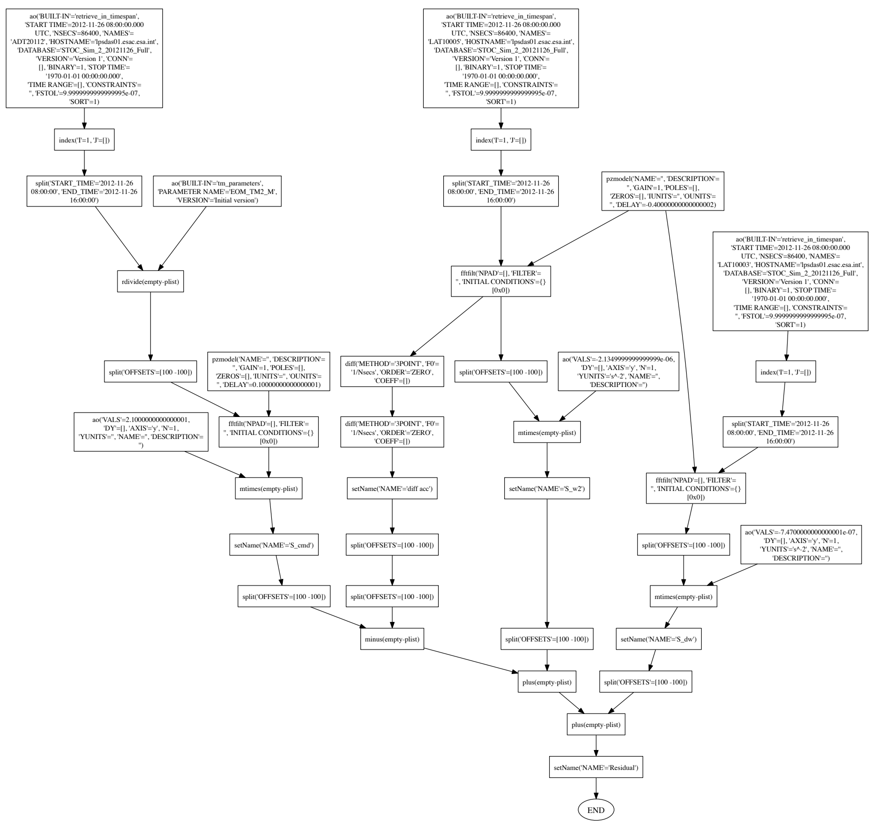
LP Data Analysis Activities

- Analysis software and infrastructure
- Defining, simulating and testing experiments
- Analysis pipelines
- End-to-end tests (including hardware wherever possible)
- Interfacing with ESA's STOC and MOC

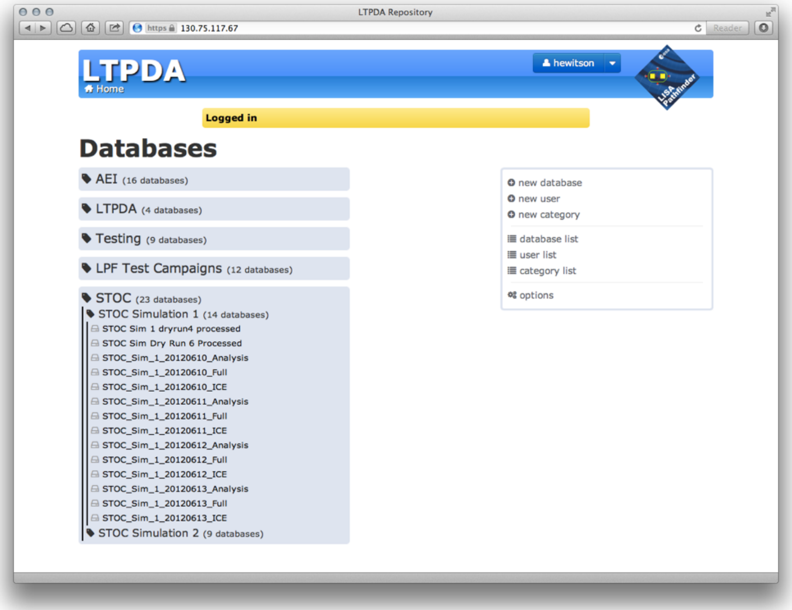
Analysis software and infrastructure

lisa pathfinder

- LTPDA Toolbox
  - MATLAB toolbox which implements an object-oriented data analysis environment
  - objects track their history so results are traceable and reproducible
  - heavily tested and documented
    - ~700 page user manual
    - ~6000 unit tests running every 3 hours
    - multiple system test campaigns
    - formal deliveries to ESA with acceptance tests



- LTPDA Repository
  - provides a centralised database structure with web interface for administration and searching
  - interface to LTPDA toolbox directly from within MATLAB to submit and retrieve objects
  - core client/server system to be used by ESA for LPF mission
  - also in heavy daily use in various labs



http://www.lisa.uni-hannover.de/ltpda/

Science Operations

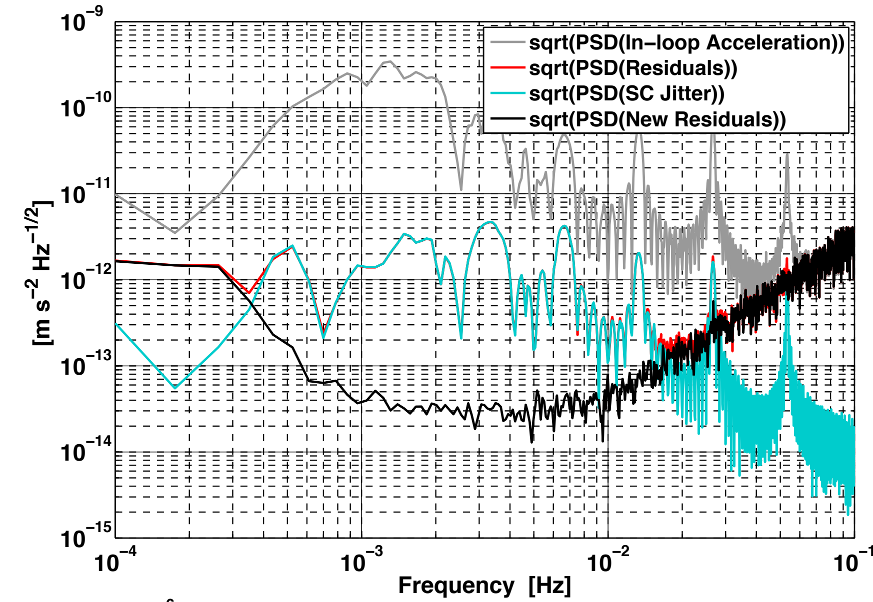
lisa pathfinder

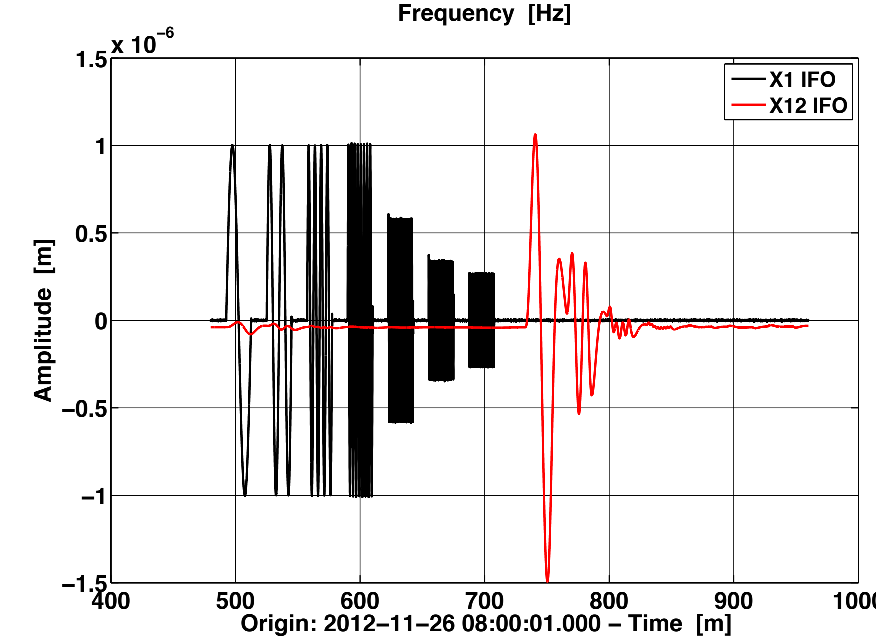
- We have 90 days in the main operational phase for LTP
- We have to plan about 1 week in advance
- We can only change the investigations that will run in about 5 days from now
- We have a large menu of possible investigations we can run
- How do we choose?
- Rough philosophy:
    - low-risk, gentle probing of the system first to gain experience and to understand the state of the system
    - move on to more invasive investigations and begin tuning the system
    - higher risk investigations are planned to be later in the operations

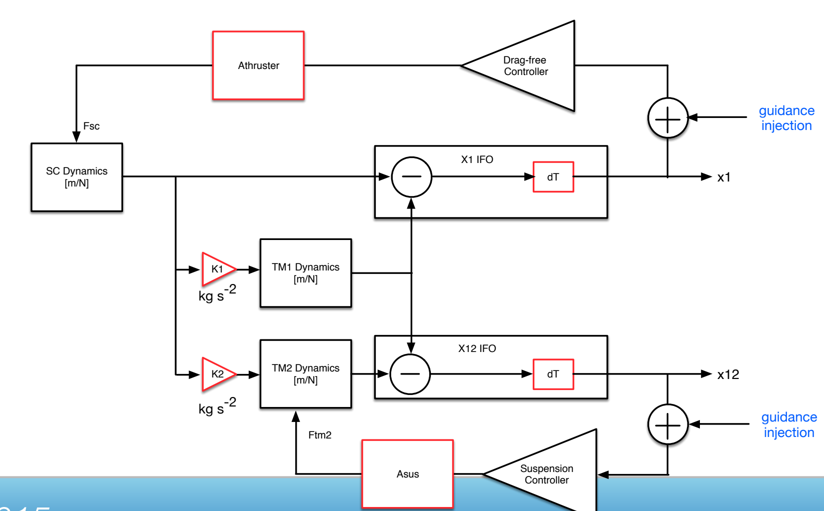
Primary analysis approaches

- Spectral estimation
  - primary science goal, assessment of residual differential acceleration
  - coherence between various signals
  - formation of noise breakdown

- Fitting
  - system identification investigations
  - signal injections
    - dynamical, thermal, magnetic, etc…







Why care about 'missing' data?

lisa pathfinder

- Stops us making spectral analysis using standard (well understood) techniques
  - something more sophisticated is needed
  - careful treatment of the data to avoid gaps/glitches
- Can cause problems during system identification experiments with injected signals
  - glitches: effective reduction of SNR
  - gaps:
    - if small, marginal loss of SNR but added complexity in processing (split into multiple experiments)
    - if large, significant loss of SNR
      - repeat experiment

Possible causes for data gaps

lisa pathfinder

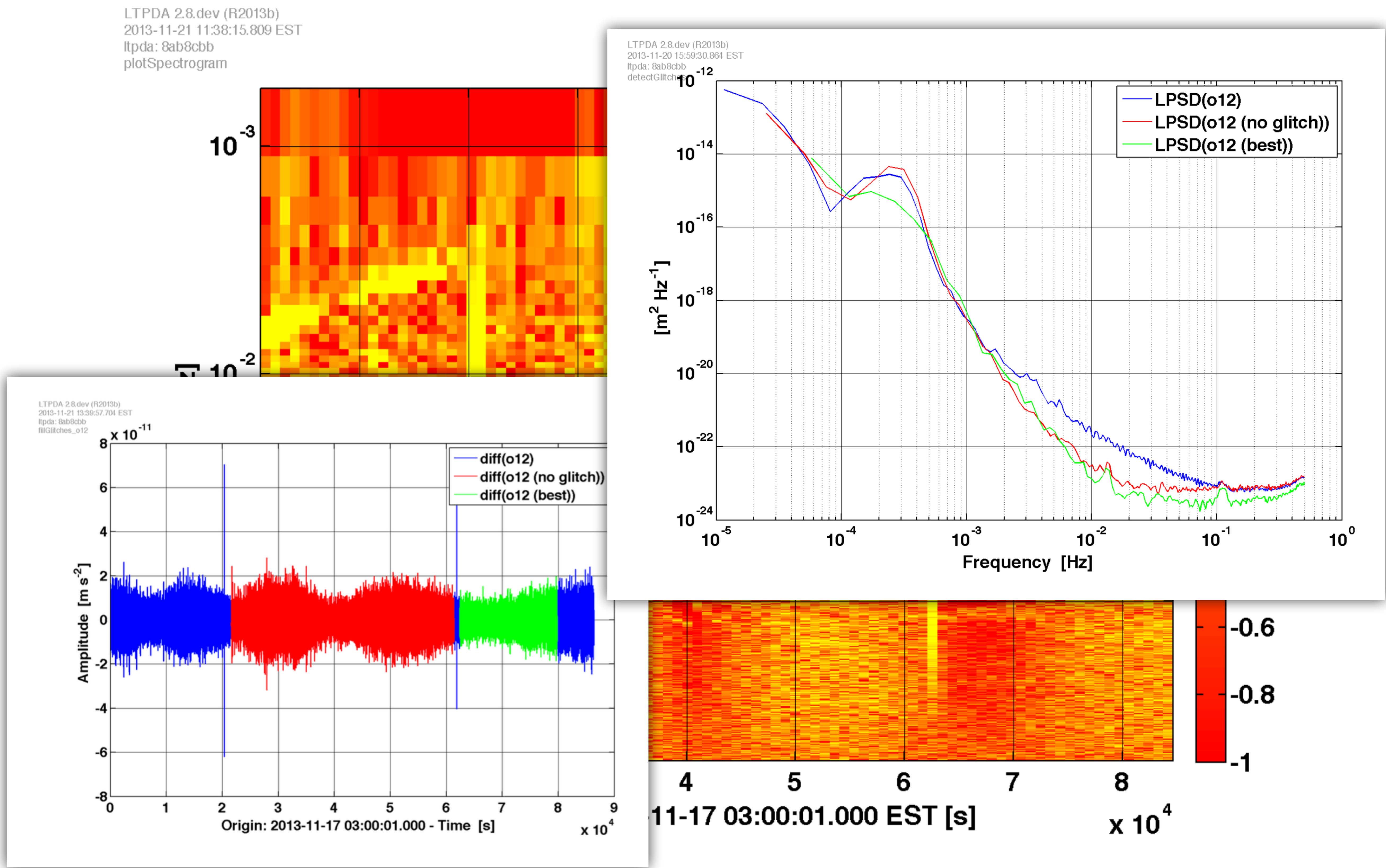
- Missing data
  - telemetry lost between satellite and ground station
    - should be rare (<1% of the time)
  - telemetry lost on-board due to errors in data handling
    - should also be rare
  - telemetry lost at MOC
    - high redundancy -> rare
- 'Bad' data
  - noisy/non-stationary
    - system configuration
    - environmental disturbances
    - …
  - glitches
    - instrumental instabilities
    - timing and clock synchronicity driven effects
    - …

Mitigation of minor data gaps

From either missing data, or unusable data:

- Ignore
  - analyse segments either side
- Interpolate
  - for simple, large and/or ‘slow’ signals, this might be ok
  - for tiny gaps (a few samples), maybe even ok for noise only data
- Fill with representative simulated noise
  - with the goal to allow standard tools to be used
- Other techniques we learn about here?!

OMS Test-campaign data



Gap filling

http://arxiv.org/abs/1412.8384

lisa pathfinder

- Fill gaps using constrained gaussian noise
  - requires a reasonable model for the spectrum of the underlying noise
- From model we generate a 2-point correlation function

$$C^{jk} = \int_{-\infty}^{+\infty} \frac{1}{2} S_y(f) e^{-2\pi f T(j-k)}.$$

- From this, we have a PDF for the noise process

$$P(\vec{y}) \propto \exp\left[-\frac{1}{2}\sum_{j,k} C_{jk} y^j y^k\right]$$

- Treat data inside and outside the gaps

Inside gaps to outside gaps

Data in gaps

$$\ln P(\vec{y}) = const - \frac{1}{2}\left[\sum^{\bar{\mathcal{G}},\bar{\mathcal{G}}'} C_{jk} y^j y^k + 2\sum^{\bar{\mathcal{G}},\mathcal{G}} C_{jk} y^j y^k + \sum^{\mathcal{G},\mathcal{G}'} C_{jk} y^j y^k\right]$$

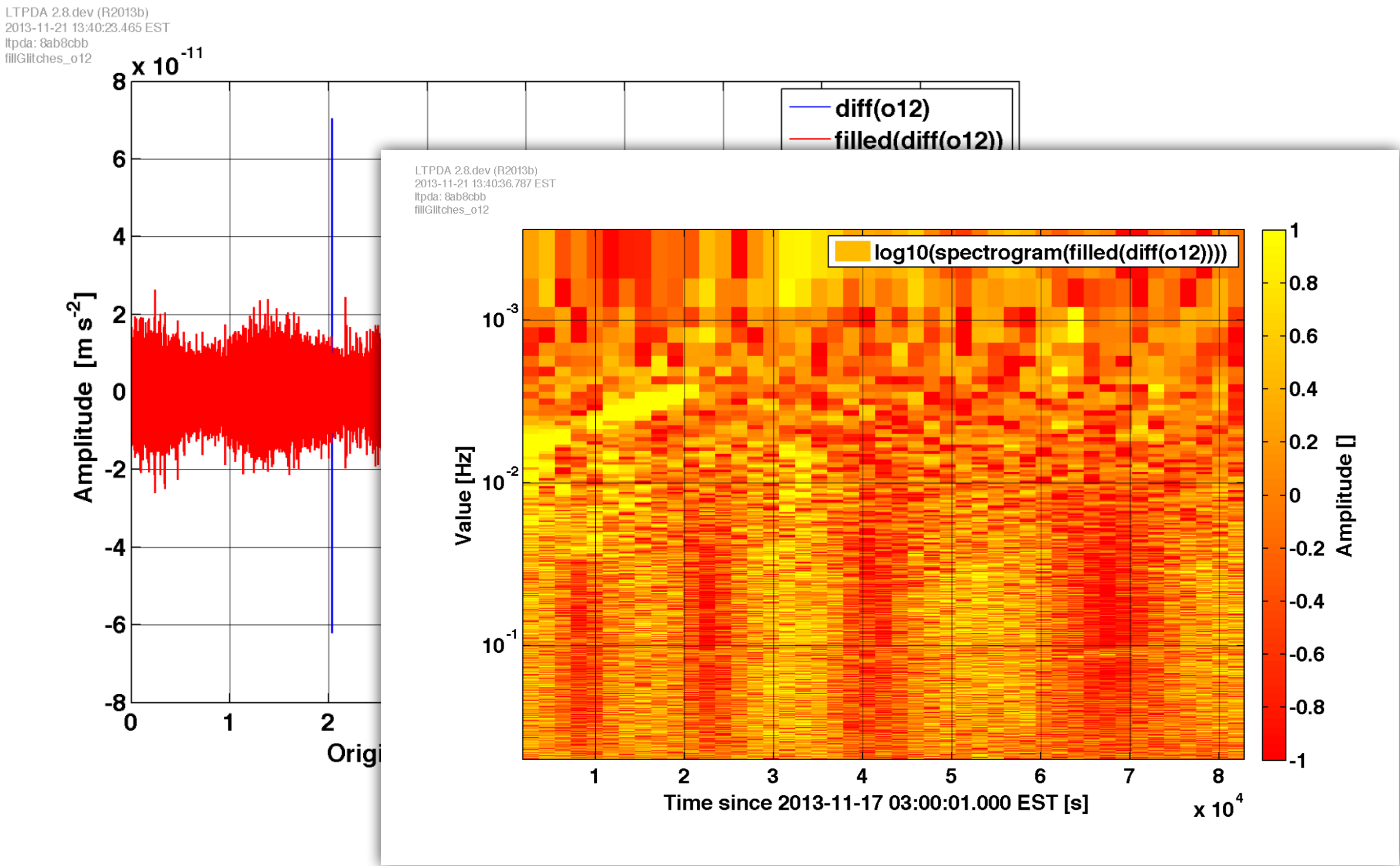
Generating missing data

We want to generate replacement data for the gaps which has the correct spectral properties and correlates 'properly' with the data we have.

Recipe:

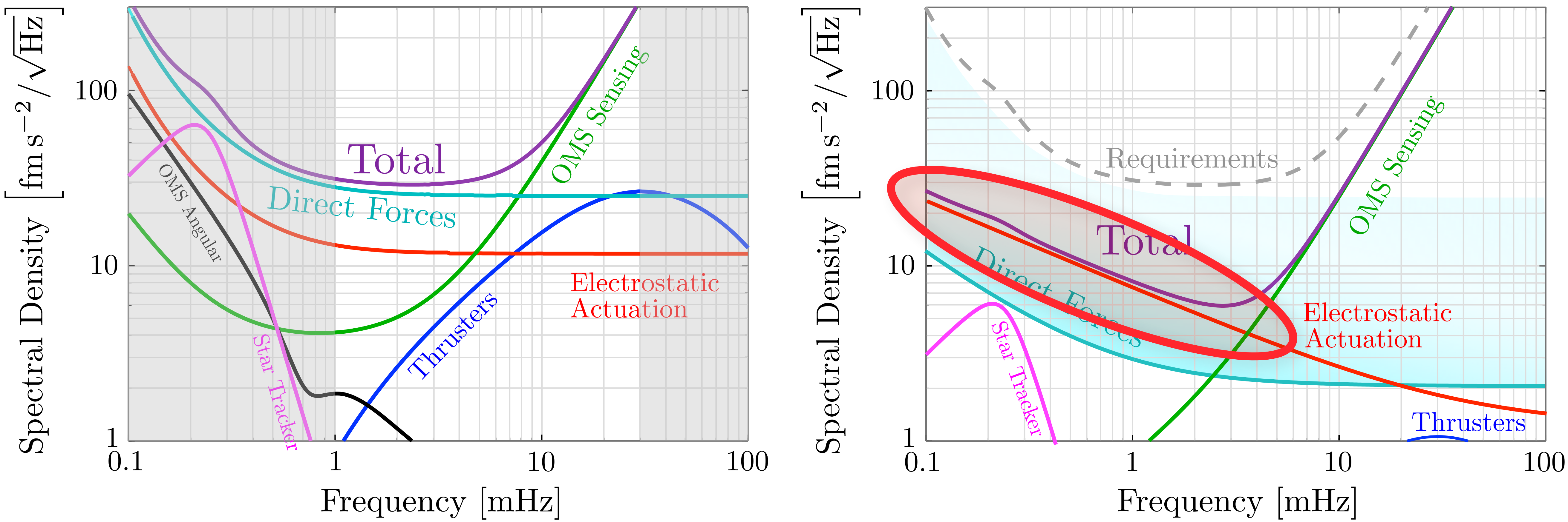
1. Estimate model for the spectral density of the data
2. Compute a 2-point correlation function from that model
3. Compute mean levels inside each gap $\Delta^j \equiv -\sum_k^{\bar{\mathcal{G}}} C^{jk}\lambda_k$
4. Draw samples from multivariate gaussian distribution with covariance C_{jk} and mean Δ^k.

Replace with constrained gaussian noise



Another source of 'gaps'

- Ones we create ourselves as a result of an experiment

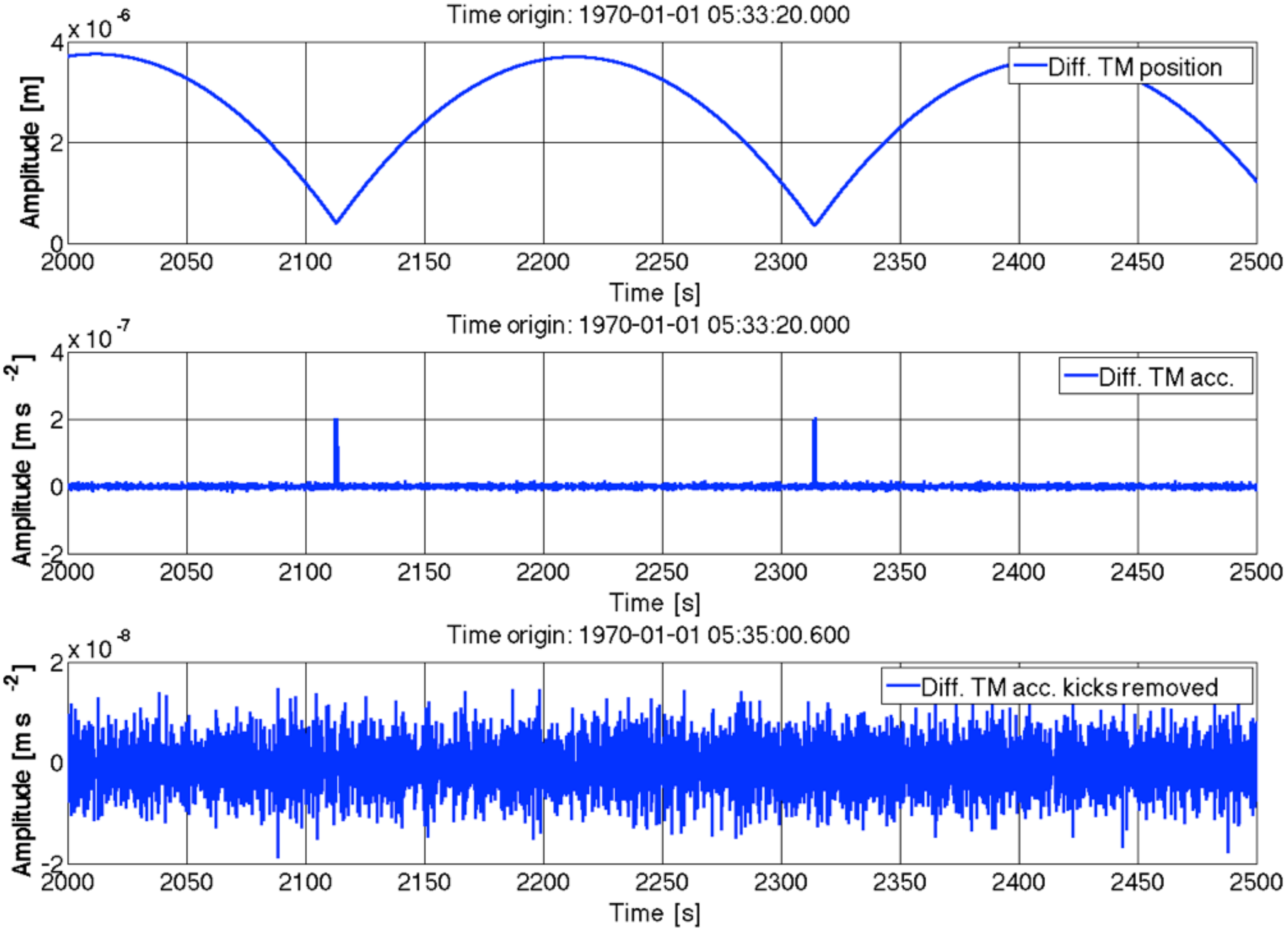


Free-flight experiment

lisa pathfinder

- Capacitive actuation may be limiting in our measurement band
- Do an experiment with the actuation off
  - must be short otherwise the TM will drift too far
  - repeat many times
- Experiment:
  - kick test-mass away
  - turn off actuation
  - let the test-mass drift (in parabola)
  - repeat

Simulation



Time origin: 1970-01-01 05:33:20.000
x 10 -6
Amplitude [m]
Diff. TM position
0
2
4
2000
2050
2100
2150
2200
2250
2300
2350
2400
2450
2500
Time [s]
Time origin: 1970-01-01 05:33:20.000
x 10 -7
Amplitude [m s -2]
Diff. TM acc.
-2
0
2
4
Time [s]
Time origin: 1970-01-01 05:35:00.600
x 10 -8
Amplitude [m s -2]
Diff. TM acc. kicks removed
-2
0
2
Time [s]

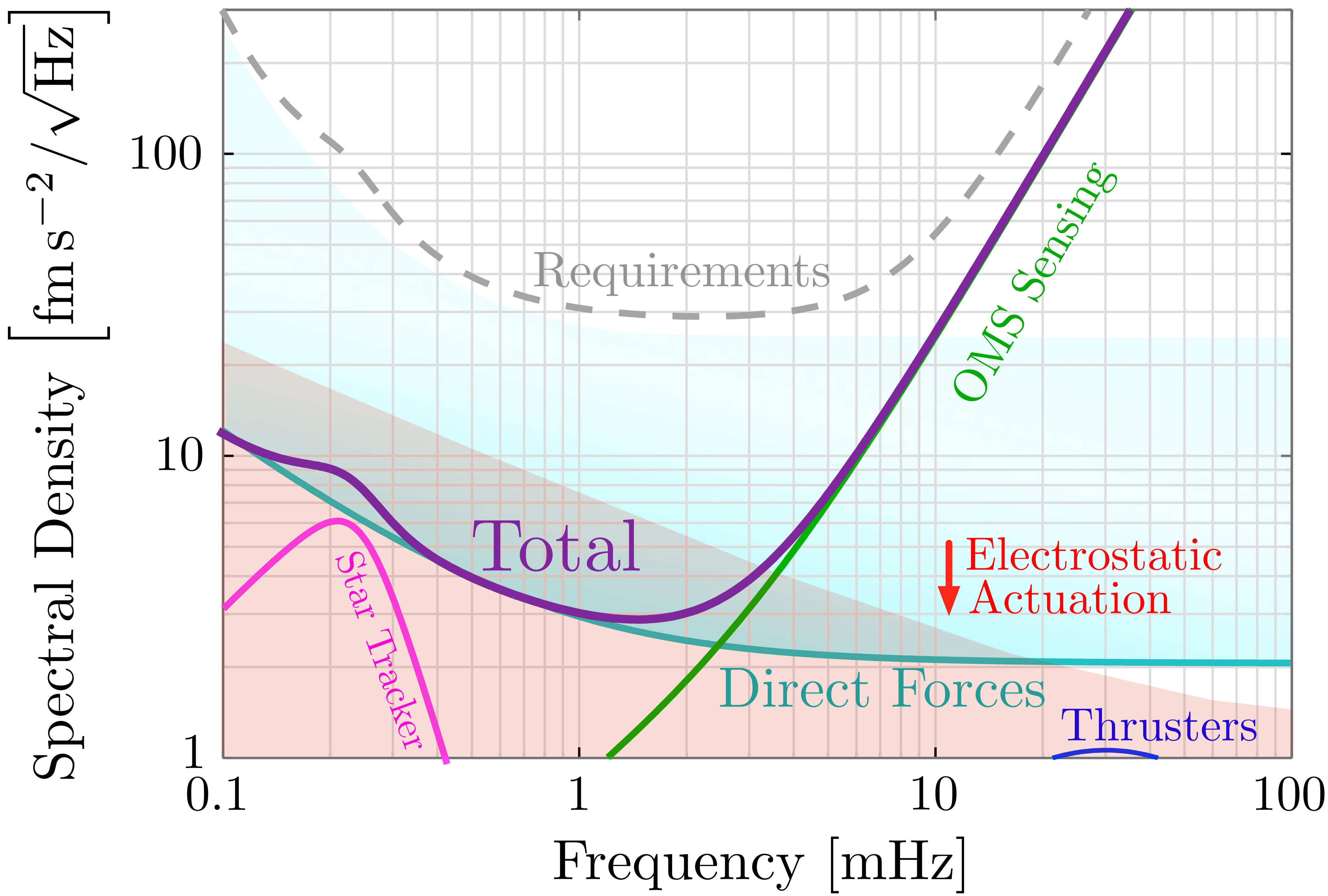
Data Analysis

lisa pathfinder

- Need to analyse the multiple short (200s) drift segments to estimate the spectrum at 1mHz
- One approach: window the data, and proceed with normal PSD estimate
- Alternative methods include replacing the noise data during the kicks with noise of same spectral content as drift phases to allow standard spectral estimation techniques

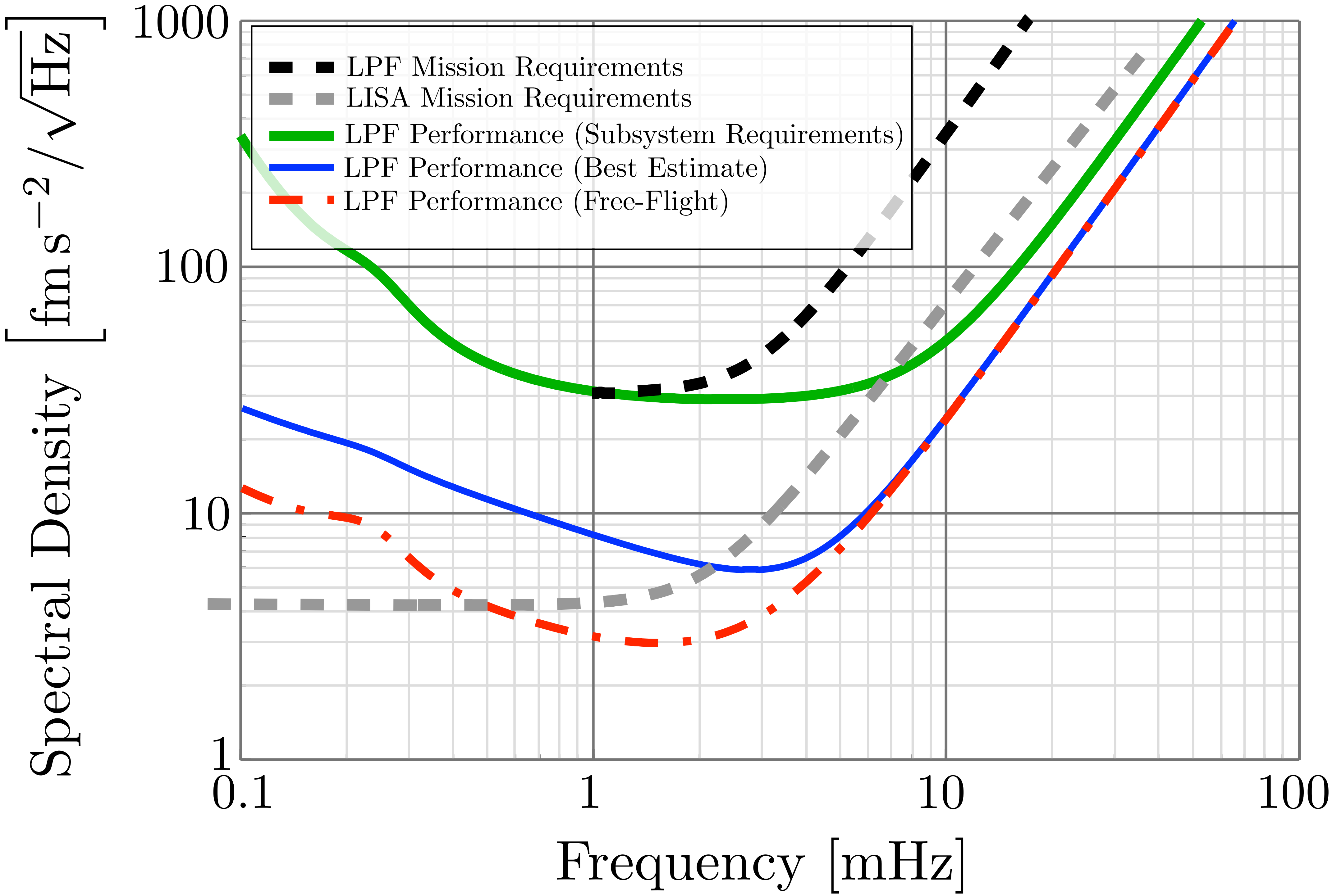
Best estimate for free-flight experiment

lisa pathfinder



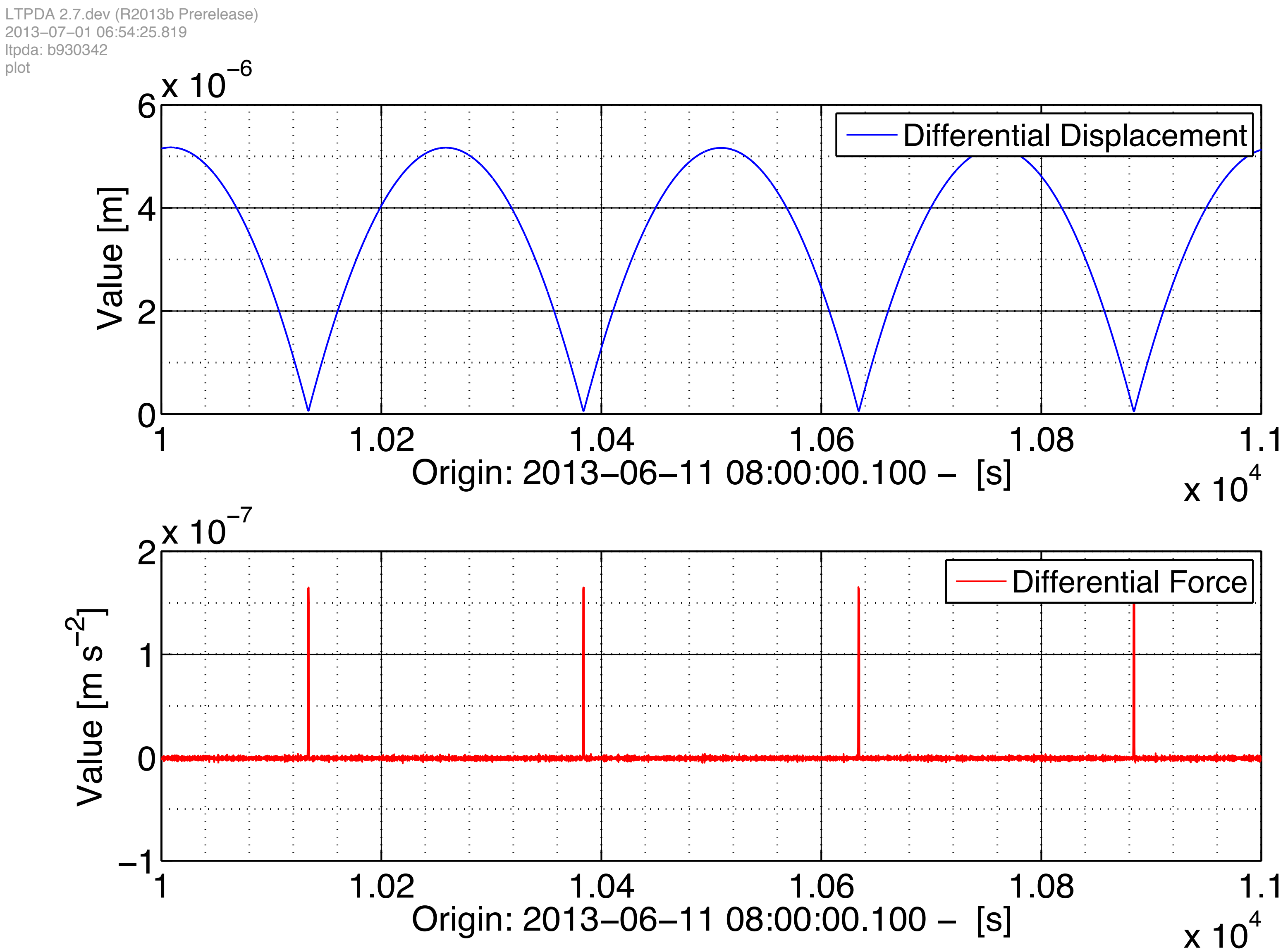
Towards LISA

lisa pathfinder

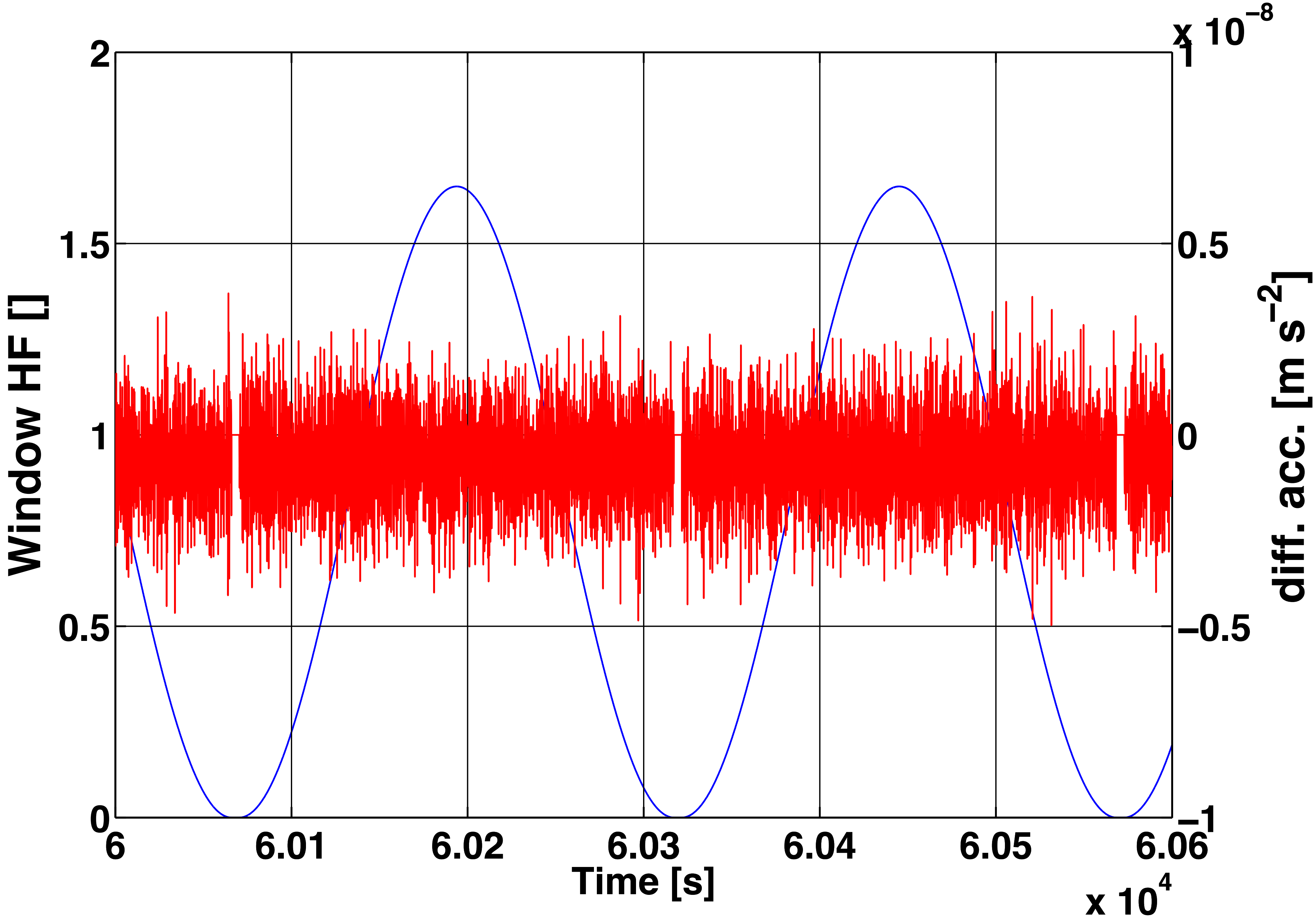


Suppress gaps: “windowing”

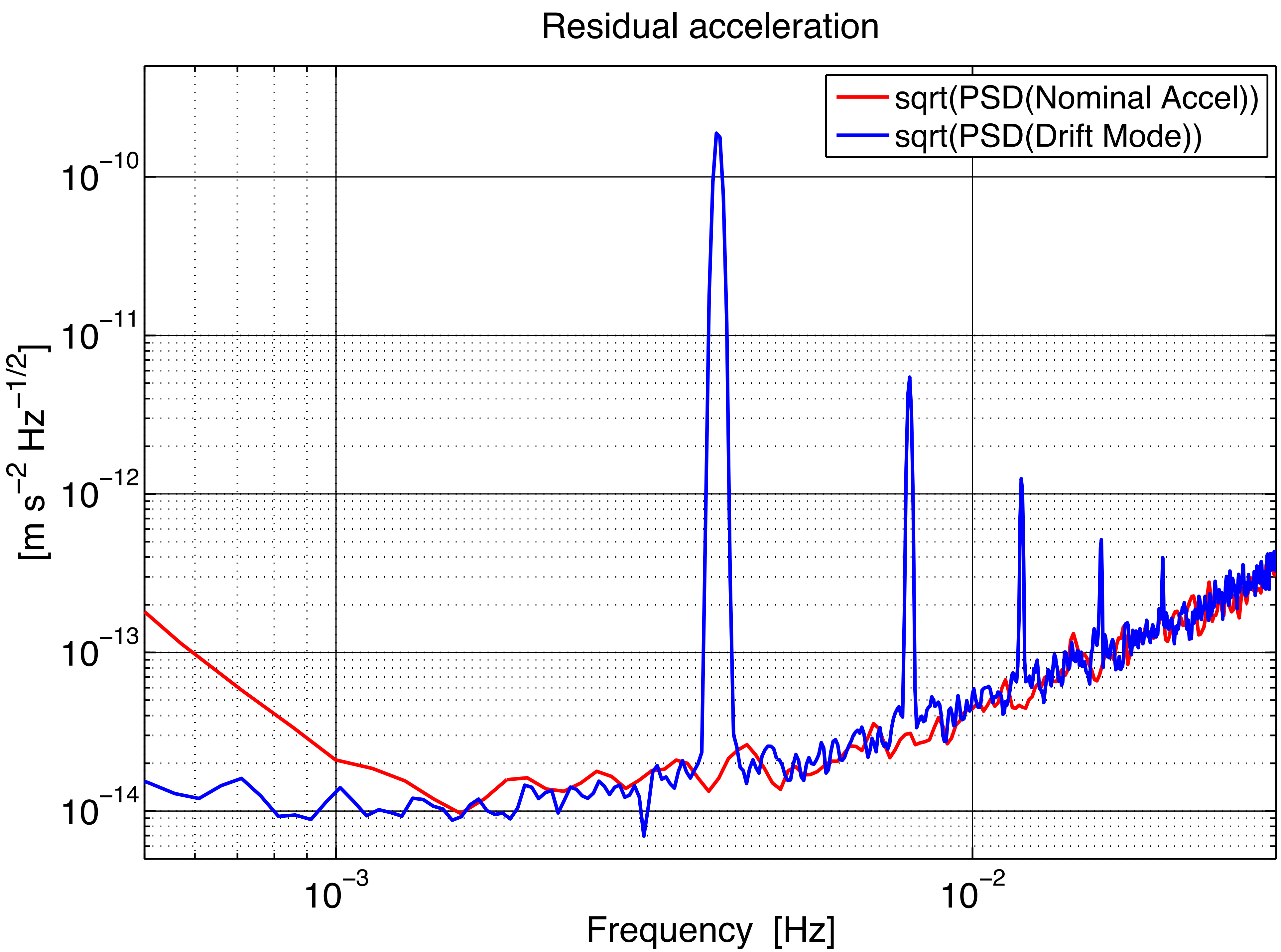
- For regular gaps, we can suppress the contribution by windowing
- Take second derivative of position:



Apply windows



Take spectrum

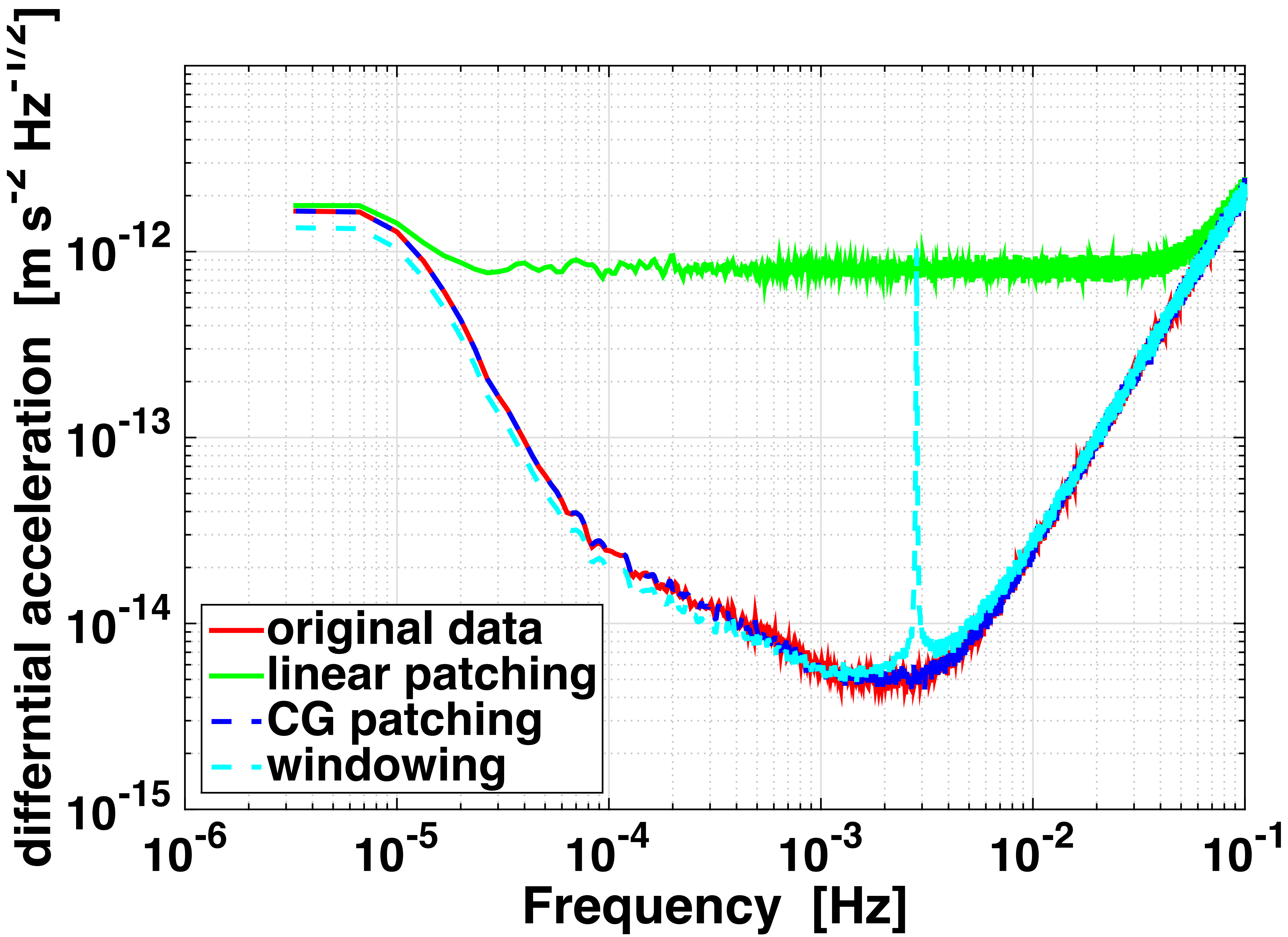


Fill gaps with constrained gaussian noise

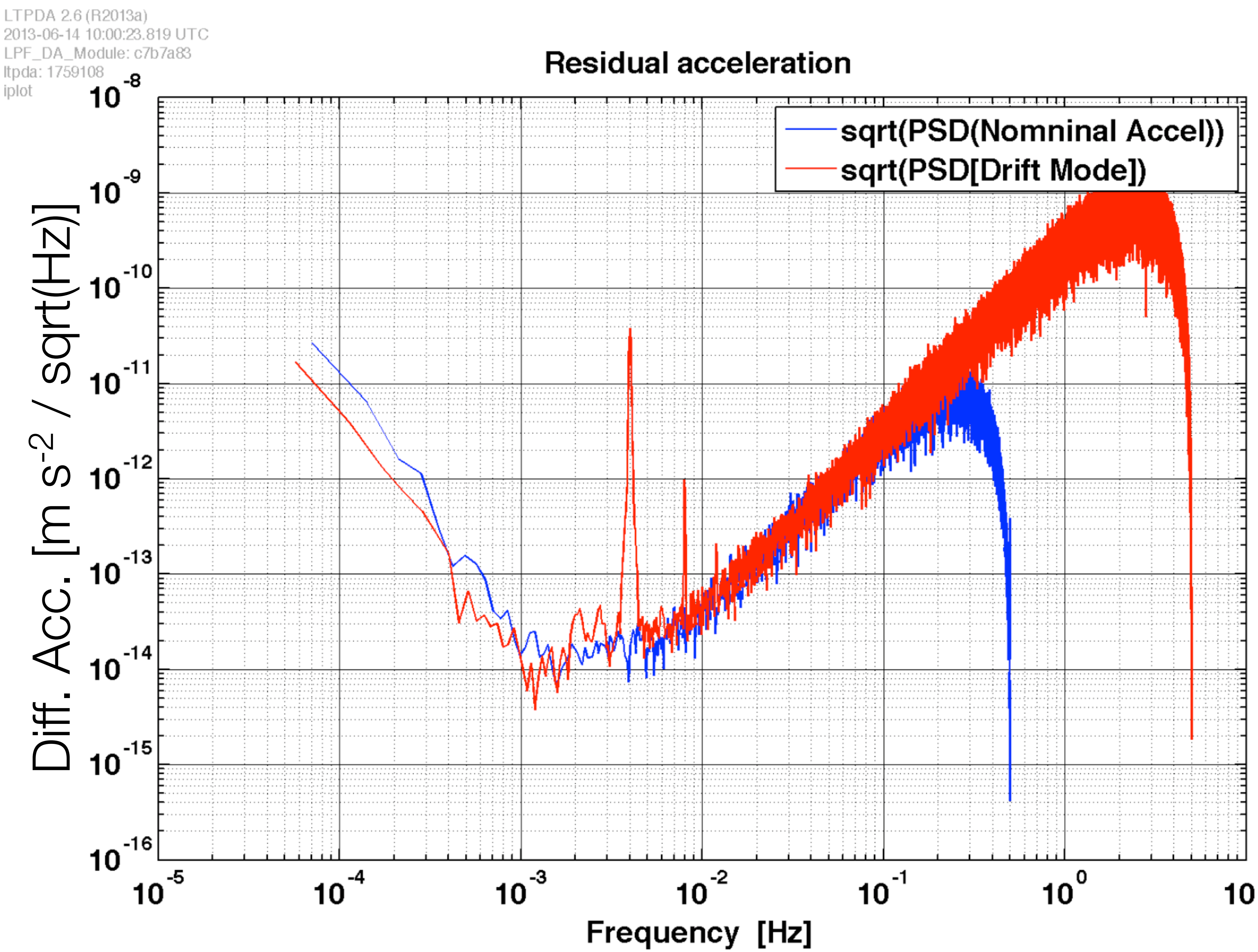
lisa pathfinder

- Again, start from a model spectrum
  - fit to data from before free-flight experiment
- Identify each ‘gap’ corresponding to the noisy force kicks
- Generate data which has the correct spectral characteristics
  - correct gaussian noise within the gaps
  - correct correlations on longer timescales

Simulated LPF noise with gaps



Simulated LPF free-flight data



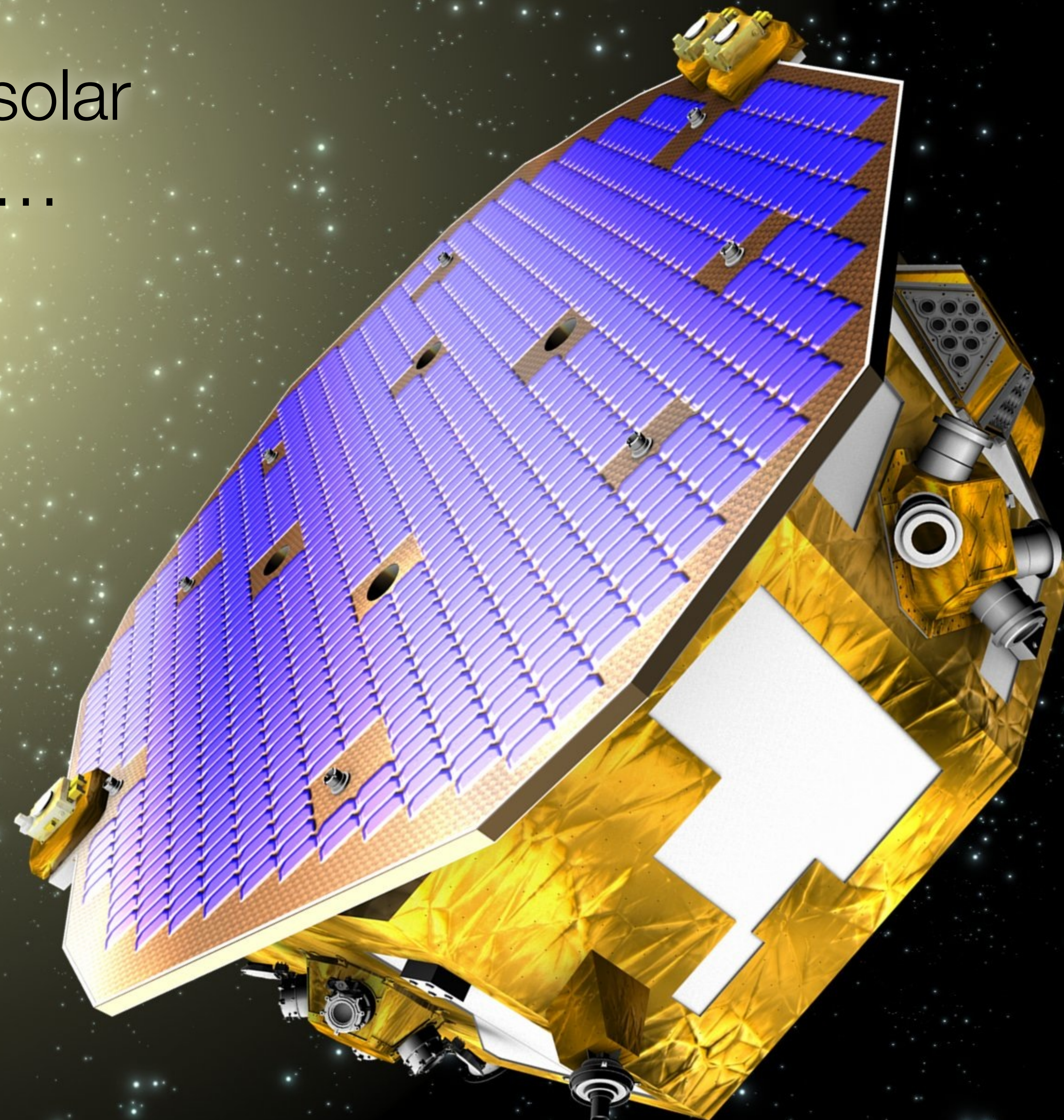
But…

- The devil is in the details
- On few gaps this works well
  - accumulated error is small?
- On more realistic simulated data, or on real data from a torsion pendulum experiment, the results are less good.
  - aliasing issues?
  - incorrect modelling of the deterministic free-flight?
  - something else we didn't think of?

Summary

- Gap filling with constrained gaussian noise promises good results
  - details need to be worked out
  - too much ‘tuning’ needed at the moment
  - need to keep error accumulation under control
- It is likely that we will have some gaps and/or periods of bad data in LPF
  - mission lifetime is short, so all data is valuable
  - need to handle these cases

Coming soon to a solar system near you…

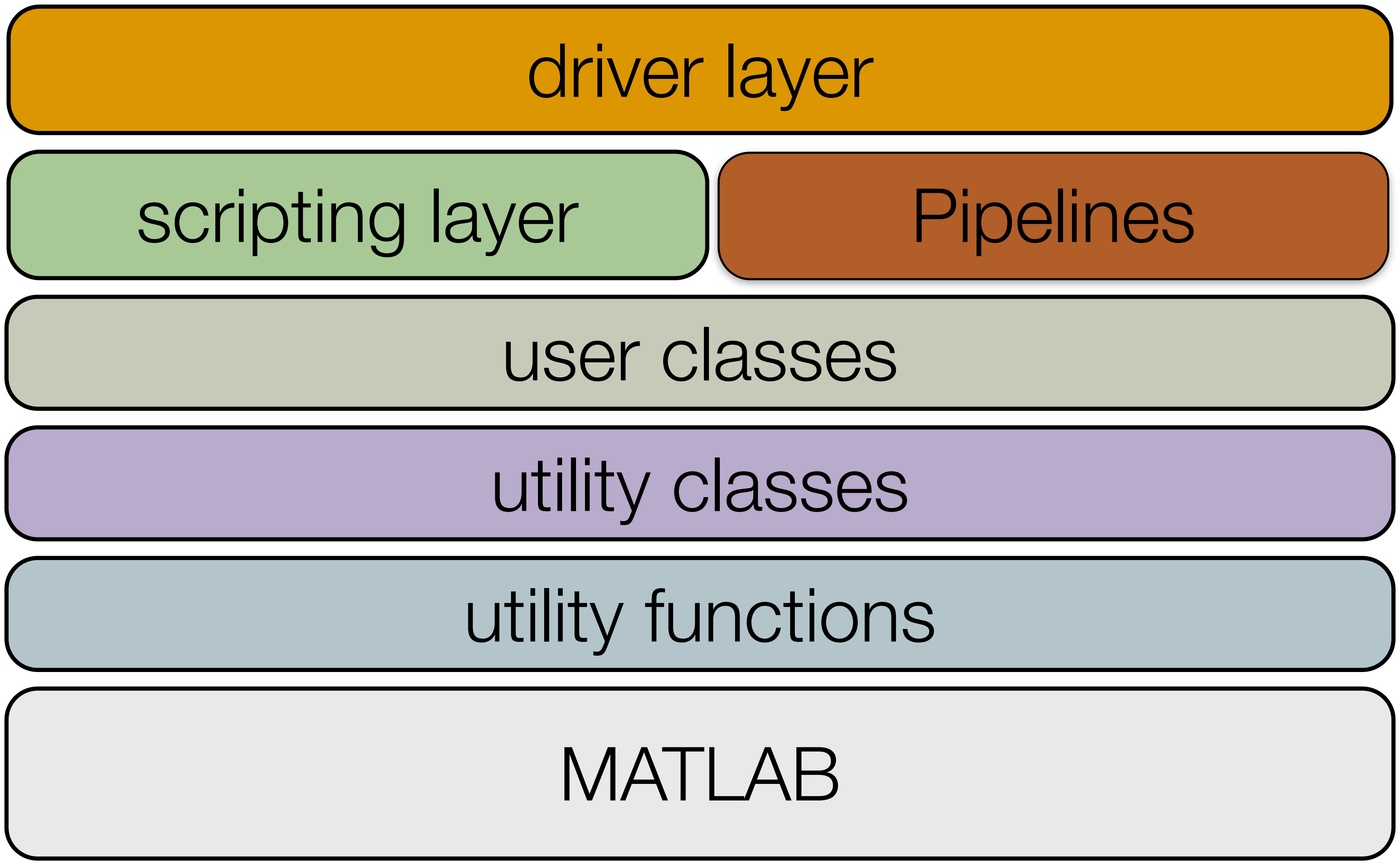


Data Analysis Toolbox: LTPDA

- Main requirement:
  - provide a data analysis environment suitable for analysing the experiment data in 'real-time'

- Additional requirements:
  - toolkit of fairly standard instrument characterisation algorithms
  - results should be reproducible and accountable
  - reduce testing by building on a commercial product
  - ease of use (usable by non-programming experts)
  - simple data access for multiple users

System Architecture



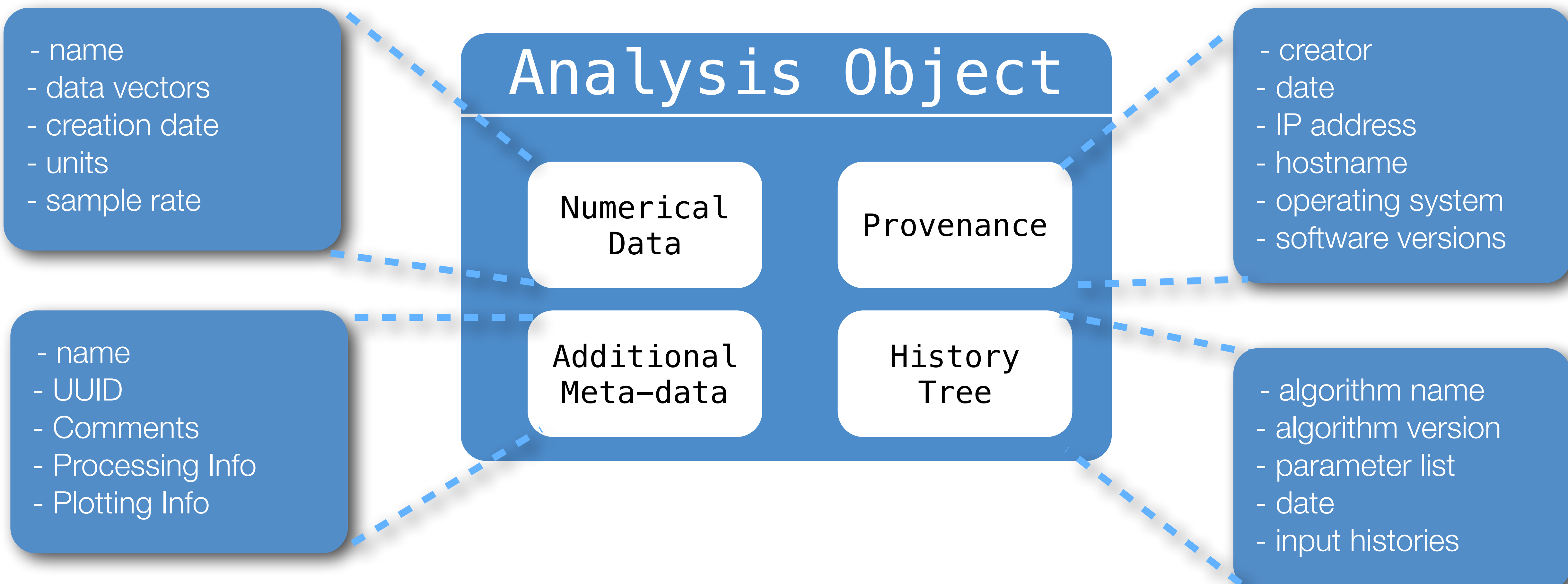
Objects, objects everywhere...

lisa pathfinder

- LTPDA offers an object-oriented data analysis framework
  - we encapsulate (describe) different data analysis concepts with classes
  - users instantiate (build) these classes to get ltpda objects
  - DA algorithms are methods (functions) of these different classes
  - users act on the objects using the methods

Example: Analysis Objects

- We want to encapsulate the concept of an analysis result
  - avoid storing images, text files, documents, etc



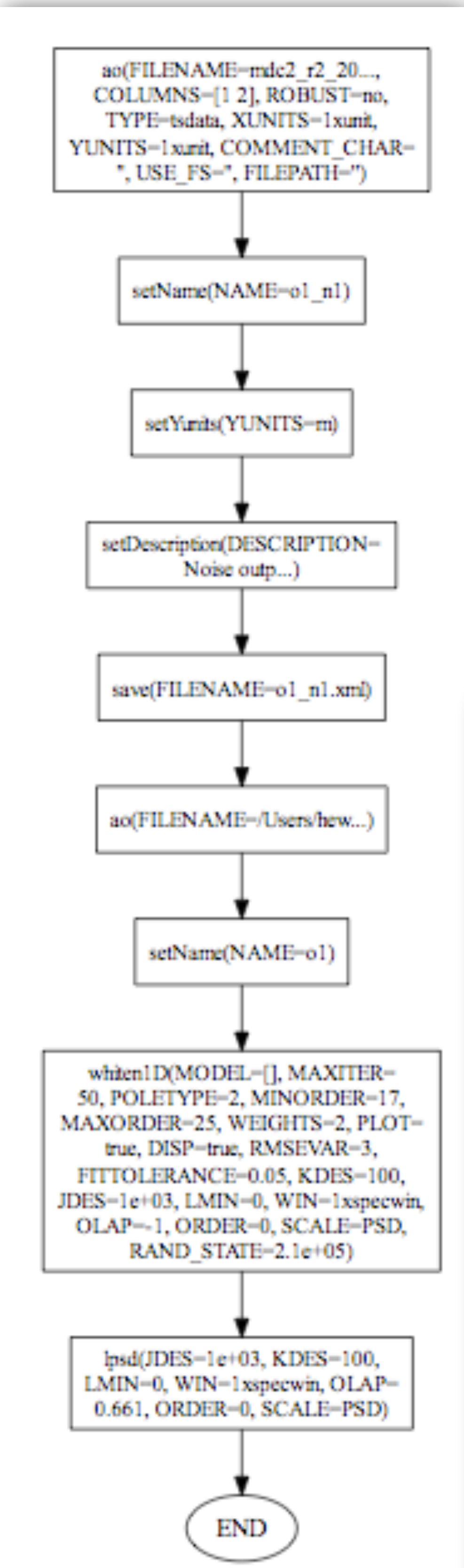
AO Algorithms

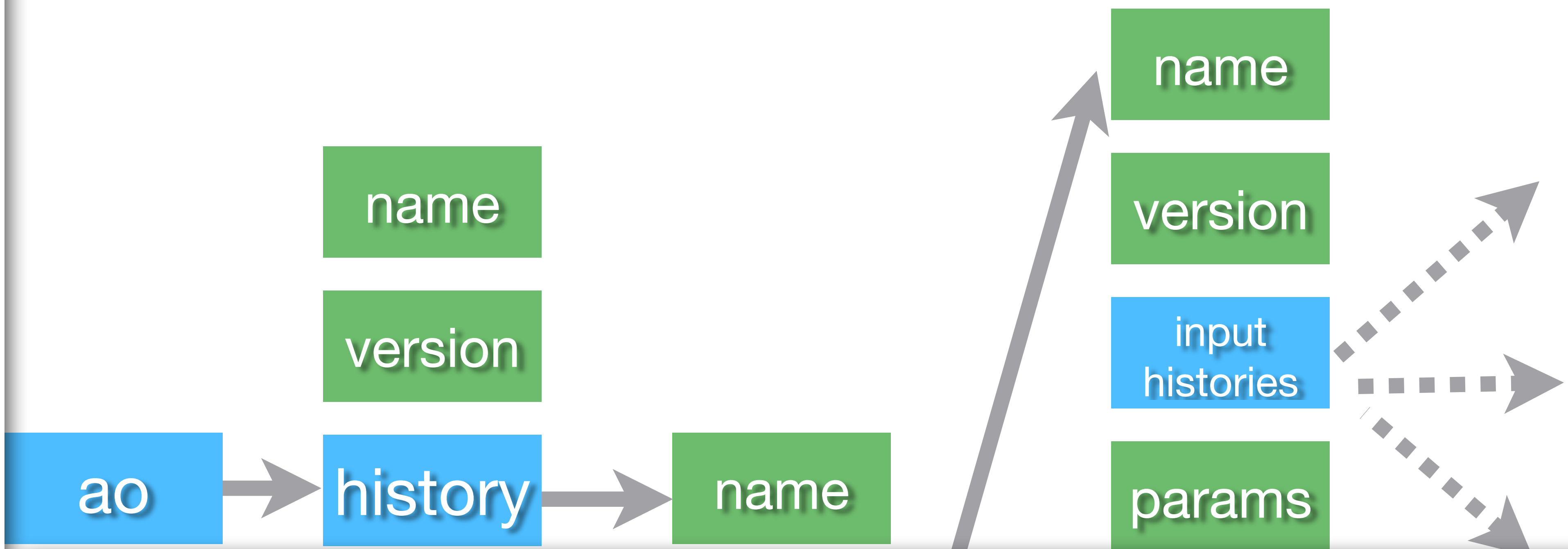
lisa pathfinder

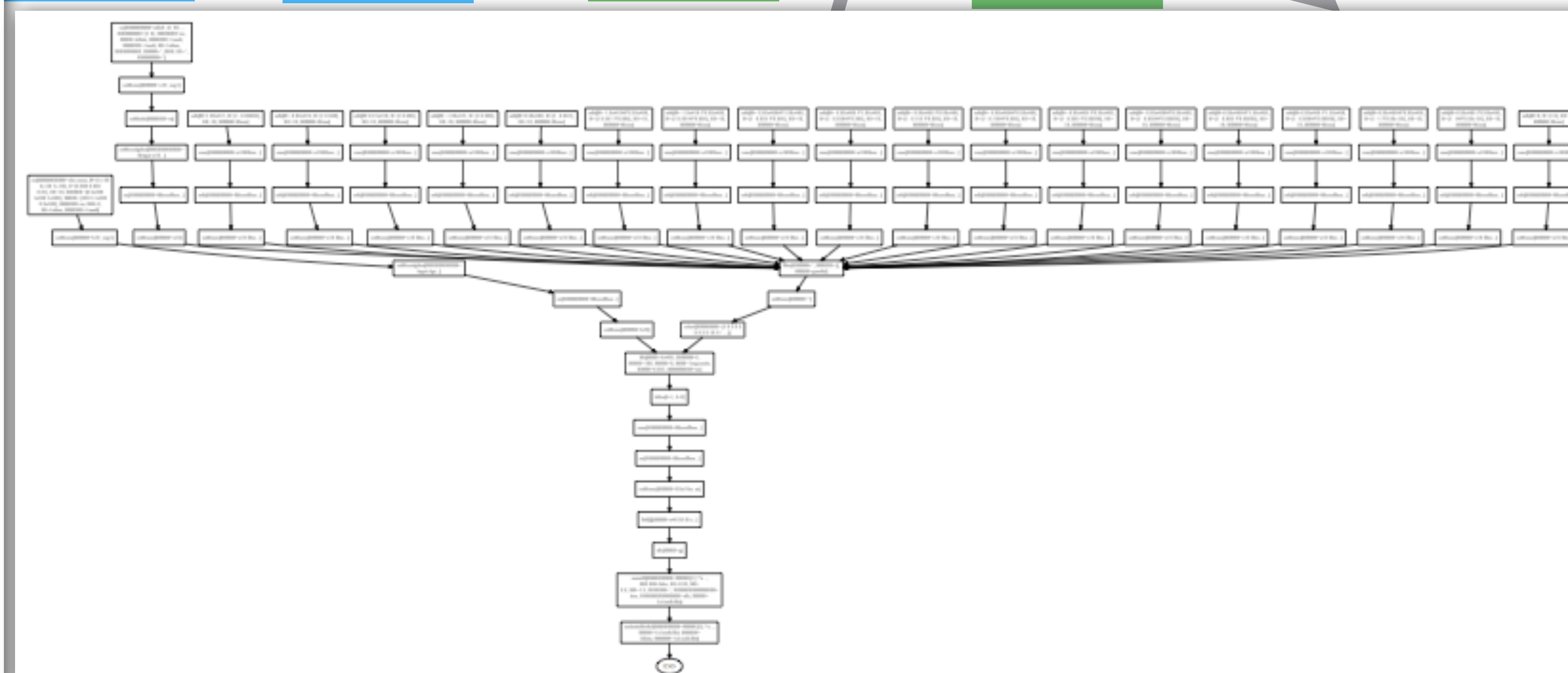
```
>> methods ao

Contents             detrend              hist_gauss           mdc1_ifo2acc_inloop  rotate               std
abs                  dft                  hypot                mdc1_ifo2cont_utn    round                straightLineFit
acos                 diag                 ifft                 mdc1_ifo2control     sDomainFit           string
angle                diff                 imag                 mdc1_x2acc           save                 submit
ao                   display              index                mean                 scale                sum
asin                 dopplercorr          integrate            median               scatterData          sumjoin
atan                 downsample           interp               min                  search               svd
atan2                dropduplicates       interpmissing        minus                select               svd_fit
bilinfit             dsmean               inv                  mode                 setDescription       t0
bin_data             dx                   iplot                mpower               setDx                table
bsubmit              dy                   iplotyy              mrdivide             setDy                tan
buildWhitener1D      eig                  isprop               mtimes               setFs                tdfit
cat                  eq                   isvalid              ne                   setMdlfile           tfe
char                 eqmotion             join                 noisegen1D           setName              timeaverage
cohere               evaluateModel        lcohere              noisegen2D           setPlotinfo          timedomainfit
complex              exp                  lcpsd                norm                 setProcinfo          times
compute              export               le                   normdist             setT0                timeshift
confint              fft                  len                  nsecs                setUUID              transpose
conj                 fftfilt              linSubtract          offset               setX                 type
consolidate          filtSubtract         lincom               optSubtraction       setXY                uminus
conv                 filter               linedetect           phase                setXunits            unwrap
convert              filtfilt             linfit               plot                 setY                 update
copy                 find                 lisovfit             plus                 setYunits            upsample
corr                 firwhiten            ln                   polyfit              setZ                 validate
cos                  fixfs                log                  polynomfit           sign                 var
cov                  fngen                log10                power                simplifyYunits       viewHistory
cpsd                 fromProcinfo         lpsd                 psd                  sin                  whiten1D
crbound              fs                   lscov                psdconf              sineParams           whiten2D
created              gapfilling           lt                   pwelch               smallvector_lincom   x
creator              gapfillingoptim      ltfe                 quasiSweptSine       smallvectorfit       xcorr
csvexport            ge                   ltp_ifo2acc          rdivide              smoother             xfit
ctranspose           get                  max                  real                 sort                 xunits
curvefit             getdof               mcmc                 rebuild              spectrogram          y
delay                gnuplot              md5                  removeVal            spikecleaning        yunits
delayEstimate        gt                   mdc1_cont2act_utn    report               split                zDomainFit
demux                heterodyne           mdc1_ifo2acc_fd      resample             spsd                 zeropad
det                  hist                 mdc1_ifo2acc_fd_utn  rms                  sqrt
```

Don’t touch me with that AO, I don’t know where it’s been!

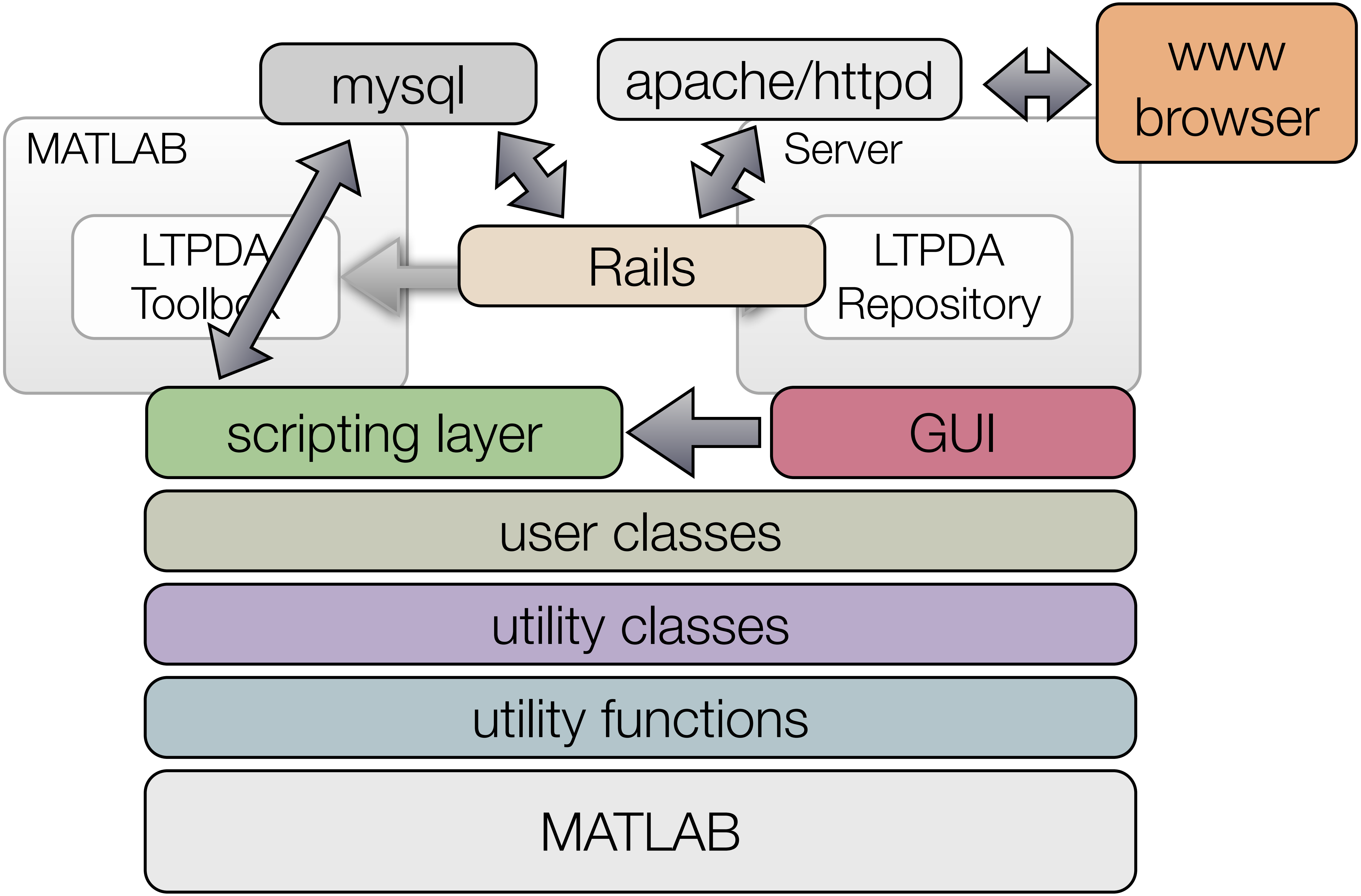
lisa pathfinder



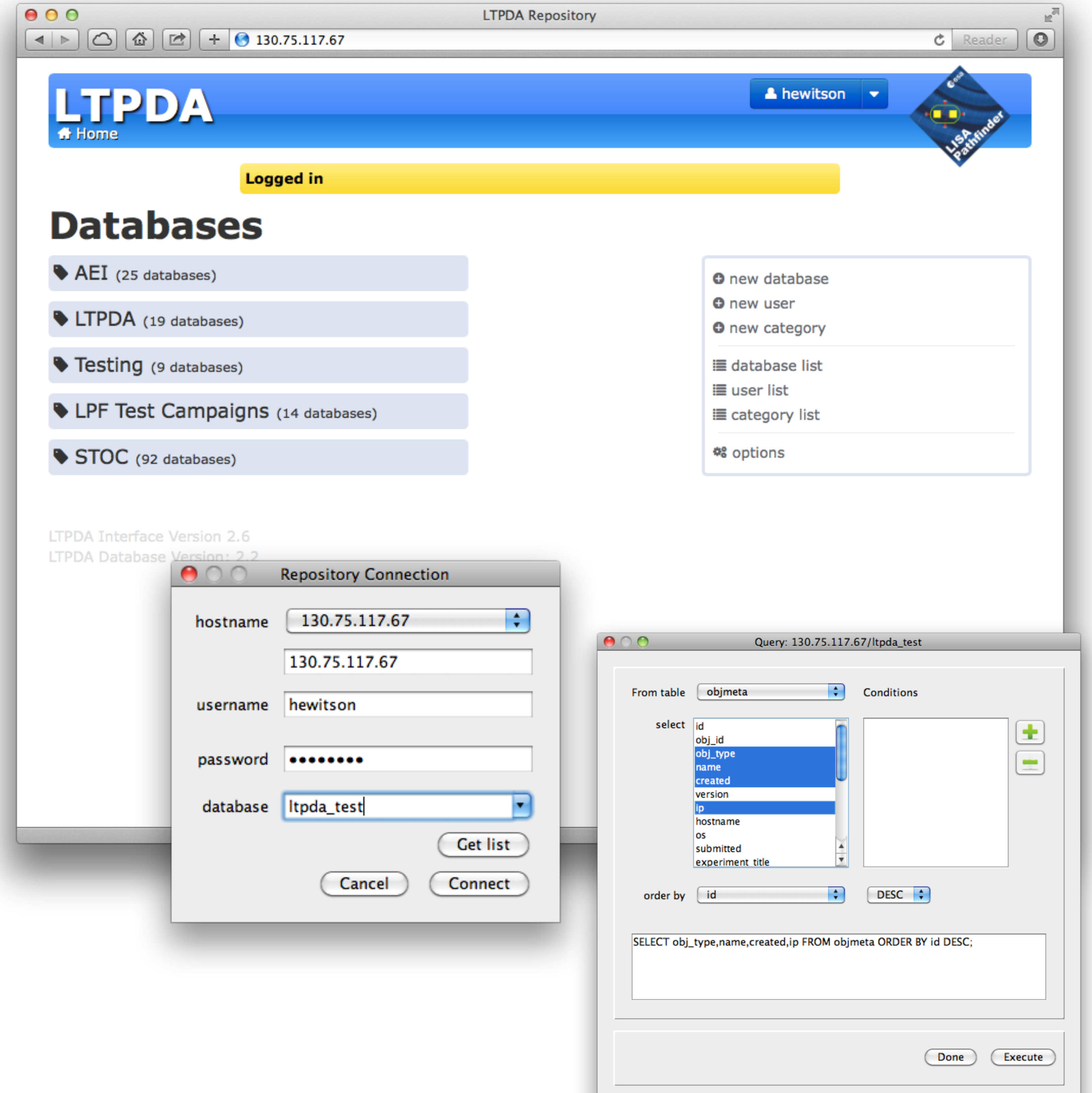




Where can I put my AO?



LTPDA Repository



Manage User Accounts

Manage Databases

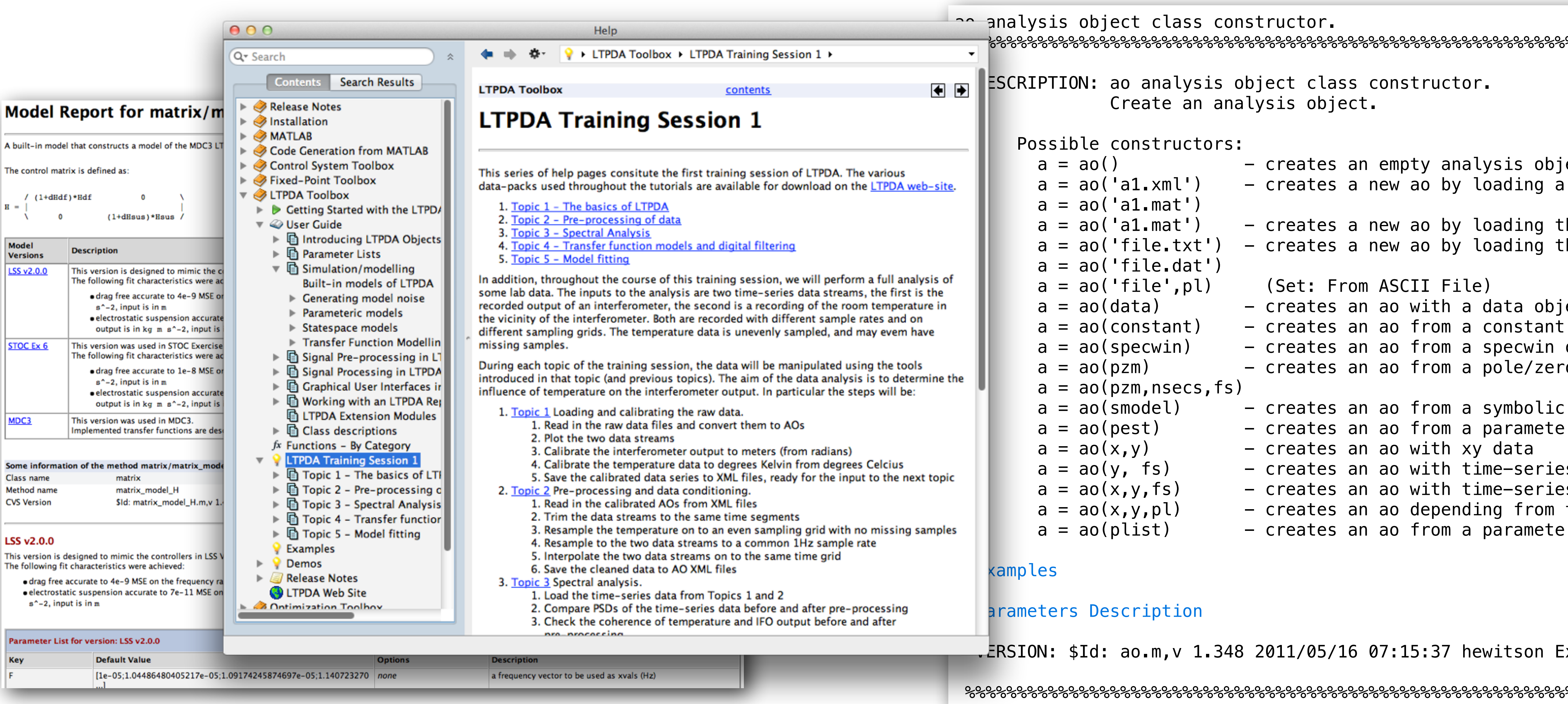
Browse/Query Databases

Submit Objects

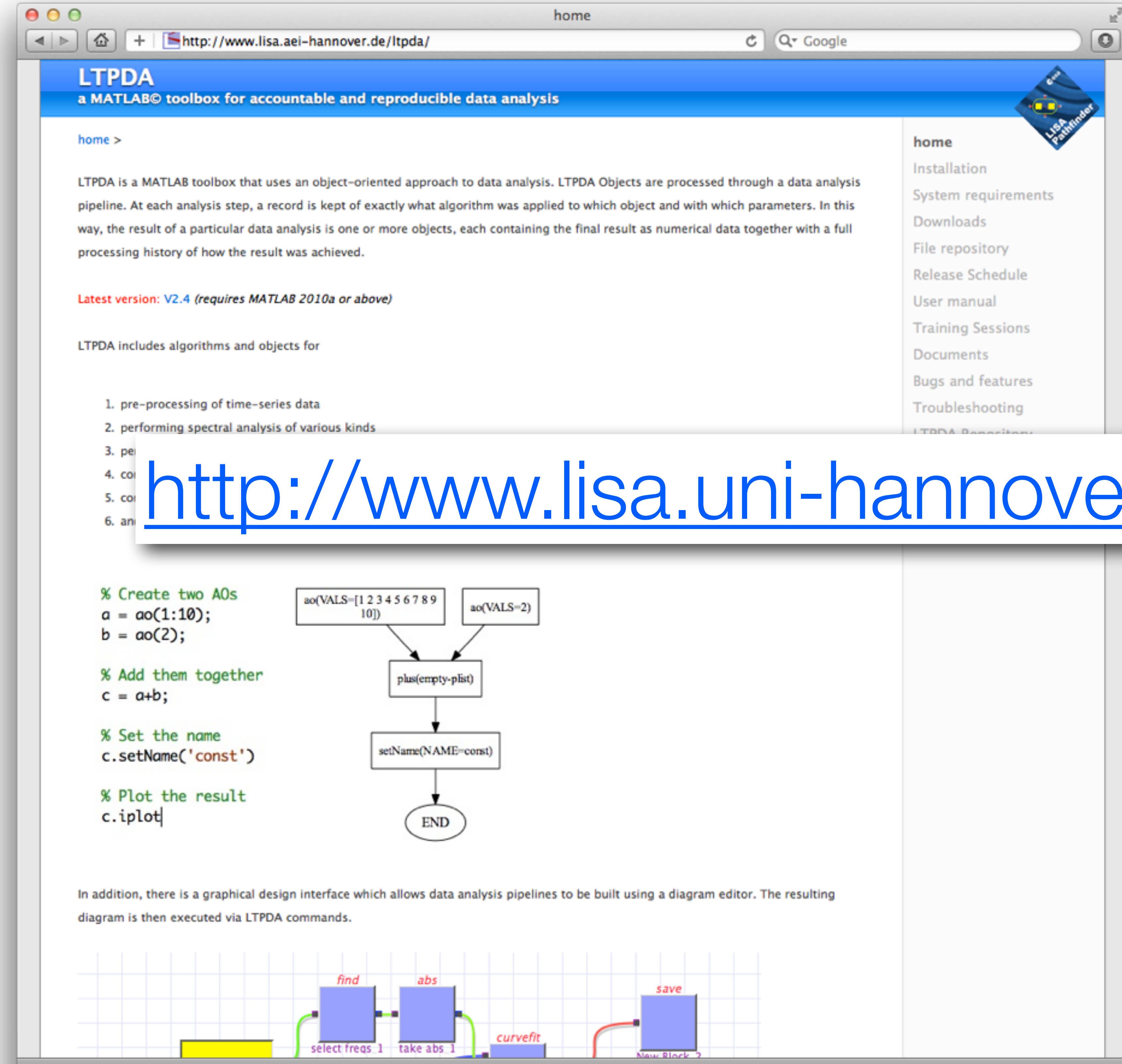
Retrieve Objects

Documentation

- Documentation
  - user-manual ~700 pages
  - ~4500 documented m-files (methods/functions)



LTPDA



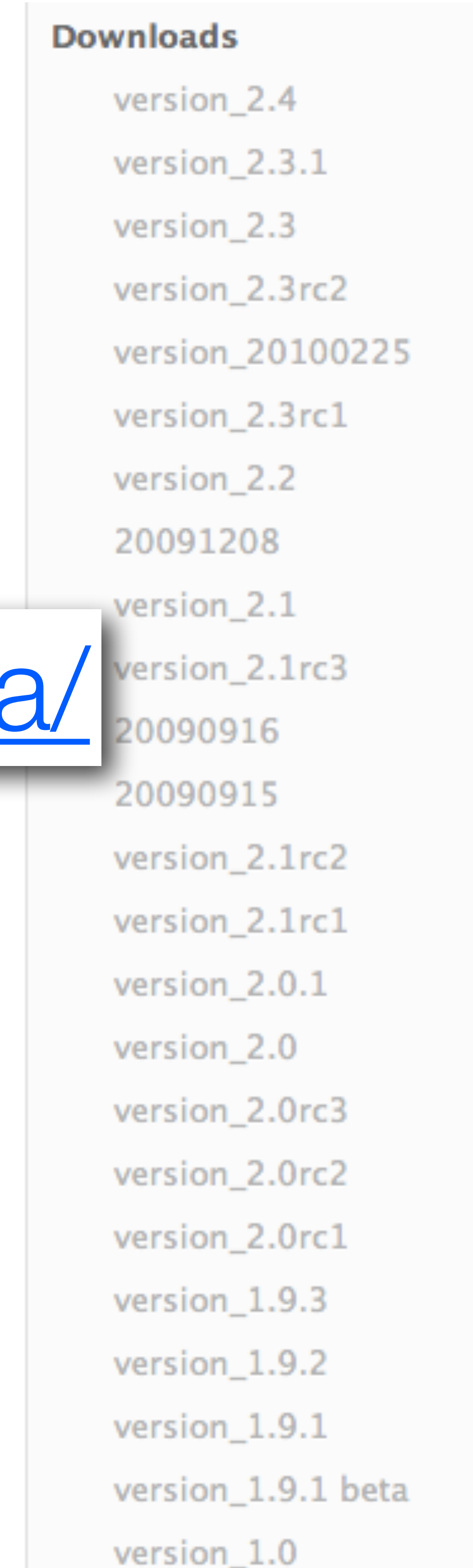
3

veries to

ining

t labs

http://www.lisa.uni-hannover.de/ltpda/



Analysis Pipelines

lisa pathfinder

- An infrastructure in which to develop 'golden' analyses for the STOC
- We need at least one pipeline for *every* planned investigation
- Key features:
  - capture blocks of analysis in a controlled way
  - arrange blocks into pipelines
  - maintain flexibility regarding
    - ordering of pipeline steps
    - configuration of each step
  - allow for easy controlled changes during operations

Phases of Operations

lisa pathfinder

Launch

IOCR

Launch, LEOP, Transfer, Separation, De-spin	Commissioning	LTP Science Ops	DRS Commissioning	DRS Operations
60 days	14 days	3 months	10 days	3 months

	Day 1	Day 2	Day 3	Day 4
H1	Noise Run	Discharge	Noise Run	Discharge
H2		Working Point		Stray Potentials
H3				
H4	Sys ID			
H5				

Week 1: Gentle Probing

lisa pathfinder

- The first two weeks are all about gathering information and gaining experience
- This is our first interaction with the system
- Focus on:
  - noise runs
  - first tests of signal injection (system identification)
  - getting a handle on the charge rate and discharging

CE1	Charge estimate TM1
CE2	Charge estimate TM2
FD1	Fast Discharge TM1
FD2	Fast Discharge TM2

	Hour																							
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	Noise run in Sci 1.2																							
2	CE1		CE2		Noise run in Sci 1.2																			
3	CE1		CE2		Sys ID (low amp)								Noise run in Sci 1.2											
4	CE1		CE2		Working point scan (x,y,z), both TMs																			
5	CE1		CE2		Cross-talk investigations, low amplitude																			
6	CE1		CE2		Noise run in Sci 1.2																			
7	Station Keeping												Transition Acc3 -> Sci 1.2								FD1		FD2	

What is a noise run?

lisa pathfinder

- Enter nominal science mode (DFACS mode Sci. 1.2)
  - SC following TM1
  - TM2 following TM1
- Put the system in the 'best' state we know
  - discharged TMs
  - optimal dc compensation voltages
  - best test-mass working point for OMS and GRS
  - …
- Take data for, e.g., 10 hours

Estimating Residual Differential Acceleration

lisa pathfinder

- Understanding the purity of the free-fall we achieve, and what limits it, requires us to assess the residual forces acting on the TMs
  - what's left when we subtract the forces we can account for?
- We compute the relative acceleration of the two TMs based on the observed relative position

$$g_{\mathrm{res}} = \ddot{x}_{12}$$

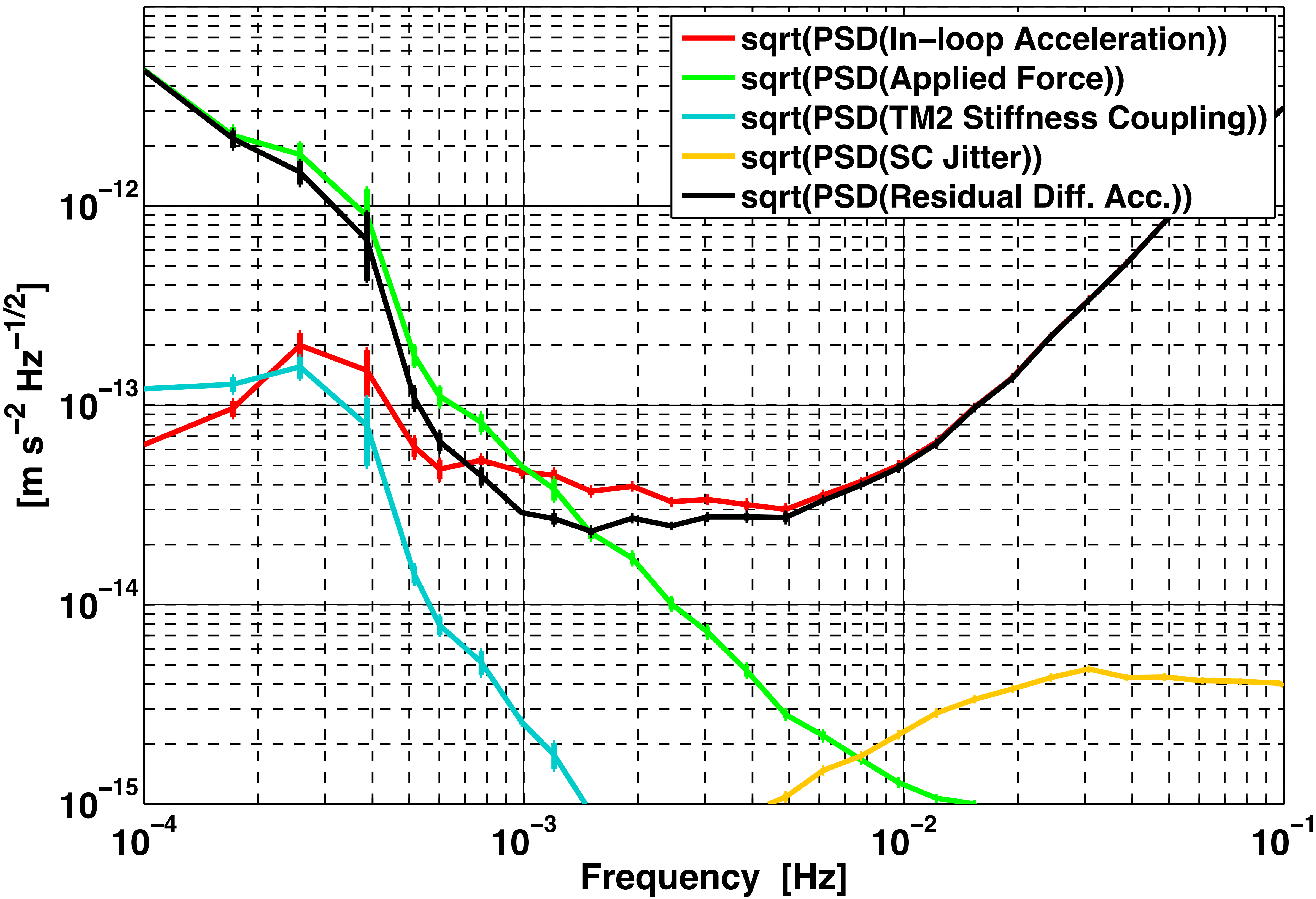
- Try to account for the contributions of g_res that we know
  - applied control forces
  - couplings due to force gradients

$$\begin{aligned} g_{\mathrm{res}} = \; & \ddot{x}_{12} \\ & -g_{\mathrm{control}} \\ & -\omega_{\Delta}^{2} x_1 - \omega_2^2 x_{12} \end{aligned}$$

Analysing the data

1. Download the time-series
2. Assemble the current best estimate of the required system parameters
   - actuator gains, delays, stiffnesses, …
3. Form linear combination of the time-series
   - with delays, and filtering as necessary
4. Take spectrum of the residuals

The contributions

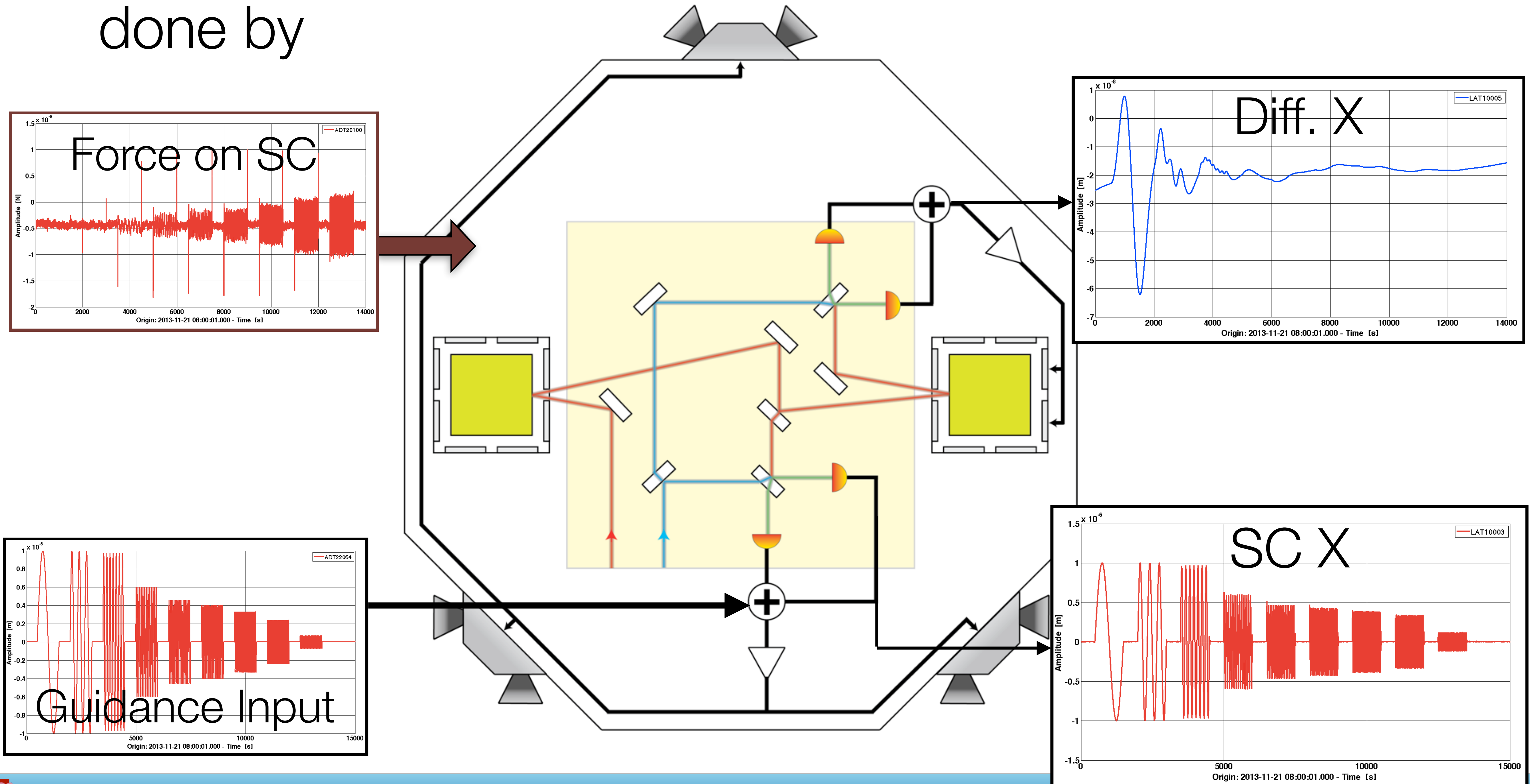


System Identification

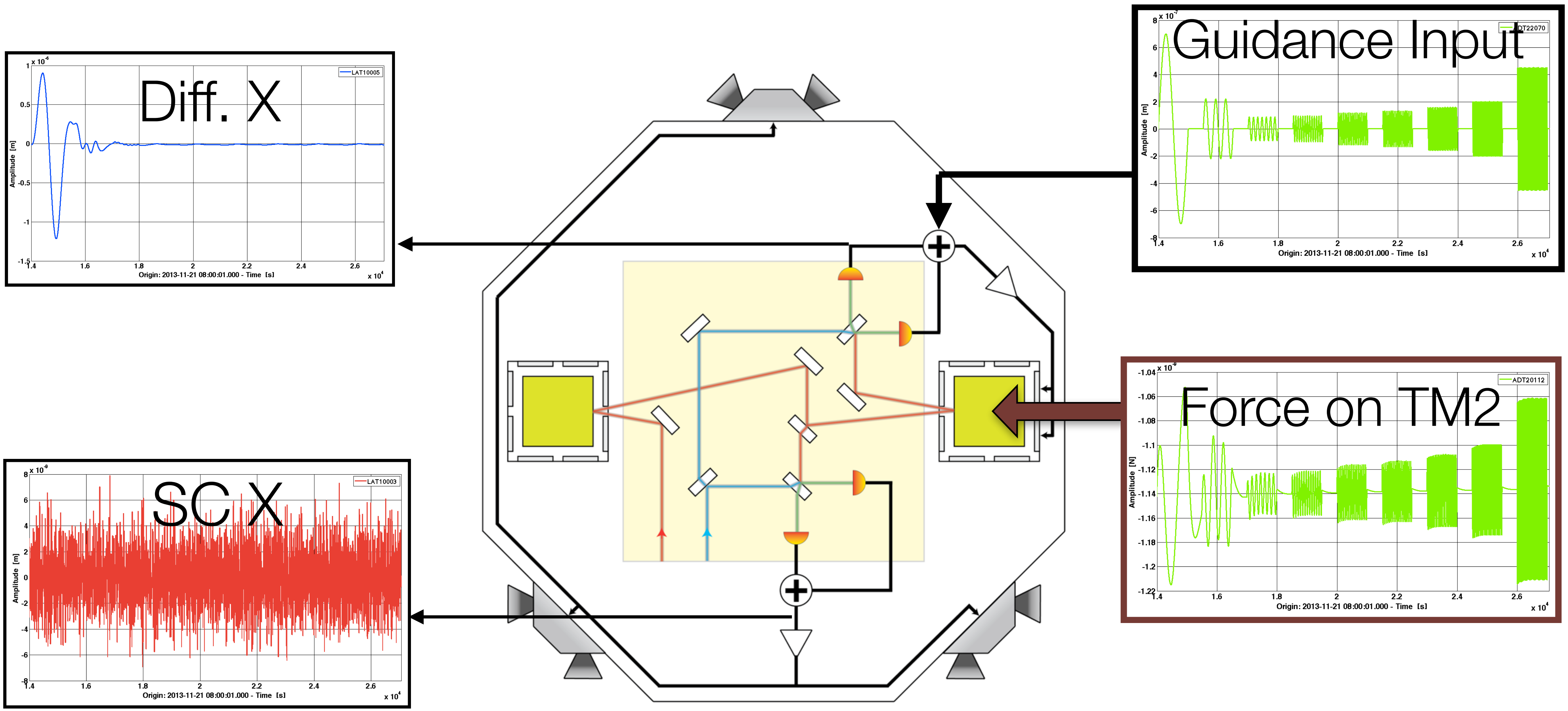
- Estimating our residual acceleration requires knowledge of certain system parameters
  - How do we gain that knowledge?
- At the beginning of operations, this comes from
  - ground measurements
  - system modelling
  - results of industrial commissioning campaign
- How do we improve and update that knowledge?
  - through dedicated investigations

x-axis system identification: part 1

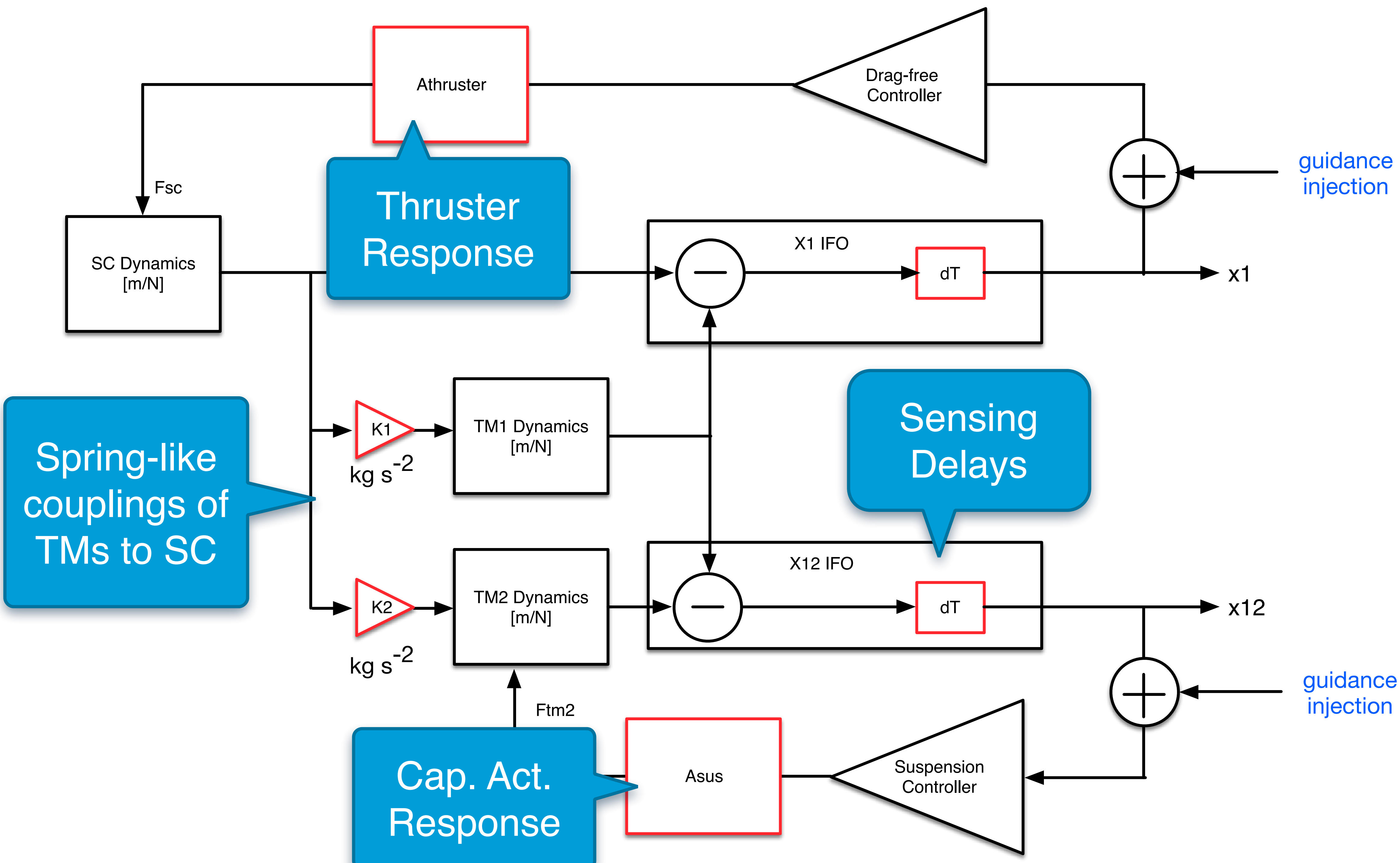
- Goal is to measure the key parameters needed for estimating the residual differential acceleration can be done by



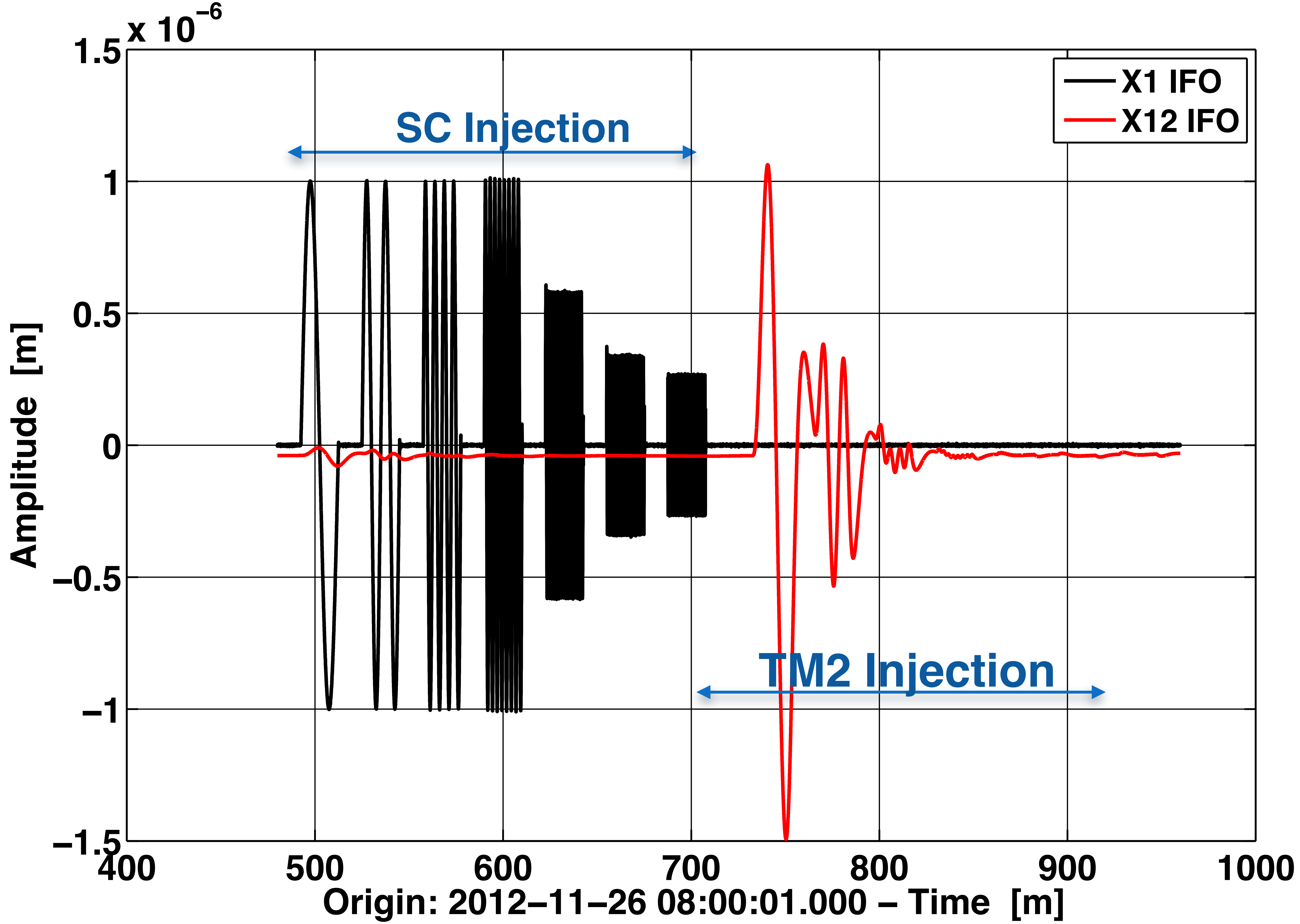
x-axis system identification: part 2



What do we learn from that?

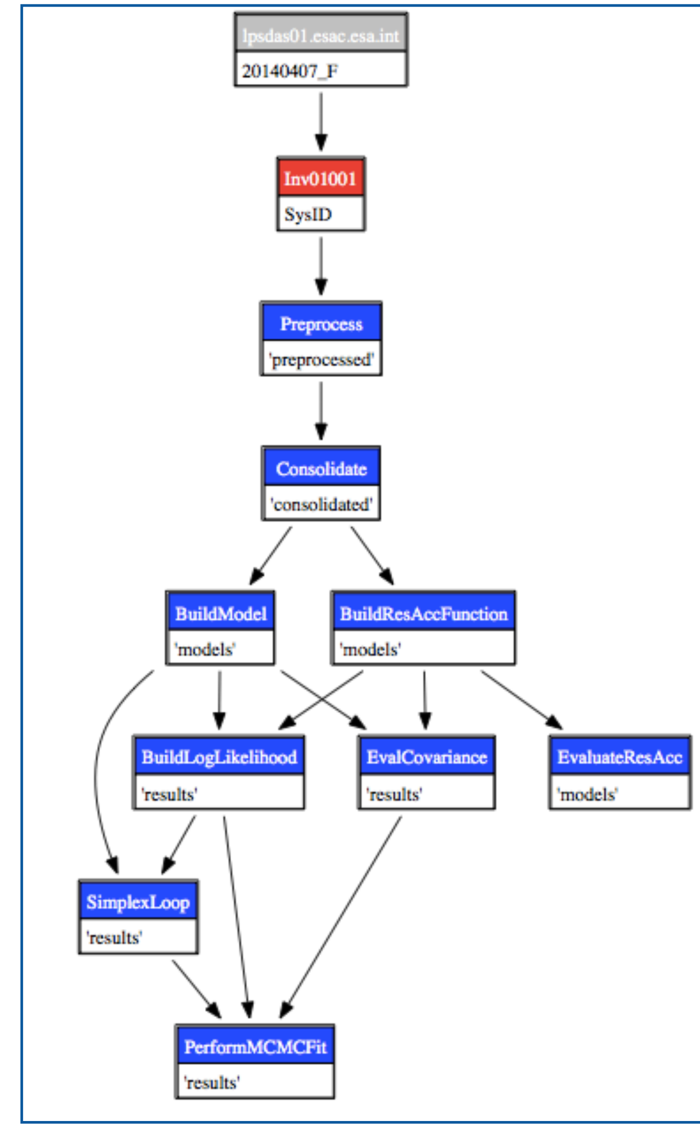


The data


x 10^−6
1.5
1
0.5
0
−0.5
−1
−1.5
Amplitude [m]
X1 IFO
X12 IFO
SC Injection
TM2 Injection
400
500
600
700
800
900
1000
Origin: 2012−11−26 08:00:01.000 − Time [m]

Analysis

- Follows the same form as for estimating residual differential acceleration
- But now the coefficients in the model are fit so that the linear combination of terms fit the observation
- When a good fit is found, the residuals contain no trace of the injected signals



Fit

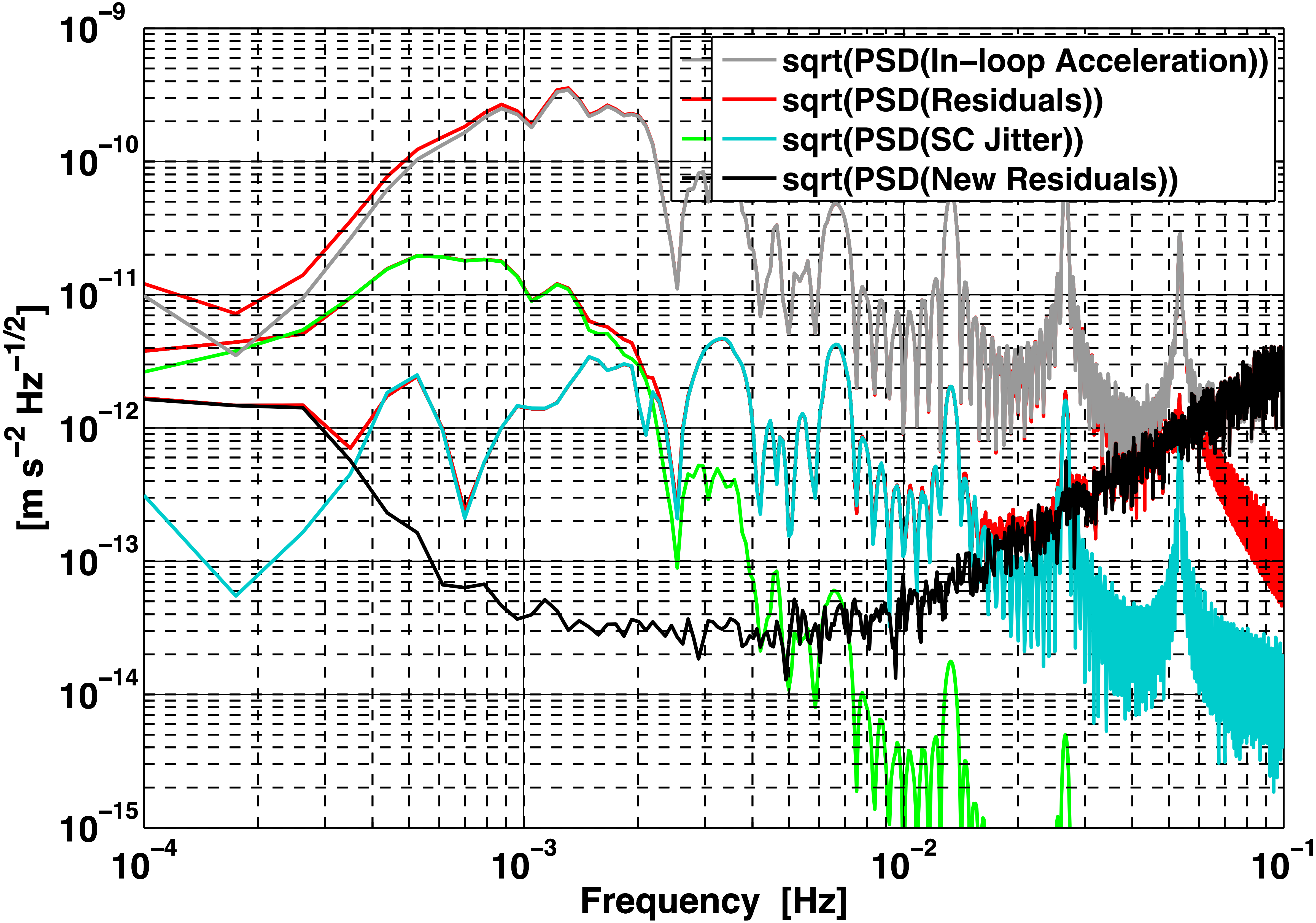
$$\text{observation} = \ddot{x}_{12}[k]$$

to

$$\begin{aligned}\text{model} \;=\;& -A_{\text{sus}}F(g_{\text{control}}[k], \Delta T)\\ & -(\omega_2^2 - \omega_1^2)x_1[k]\\ & -\omega_2^2 x_{12}[k]\end{aligned}$$

Residuals

lisa pathfinder



A general scheme

- **Balancing forces**:
  - improves physical modelling and interpretation
  - simplifies the analysis a great deal
- This ‘acceleration’ scheme can be used to account for other contributions
  - cross-talk
  - thermal
  - magnetic
  - free-flight experiments

Noise Budget

lisa pathfinder

- How does our observed residual differential acceleration differ from what we expect?
- Why does it differ?
  - this drives the next activities to be performed