Data Analysis for LISA Pathfinder

Ę



One man's gap is another man's experiment

M Hewitson for the LPF Team Missing Data Workshop, Nice 11th May 2015

Contents



- 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



- Obtain the best geodesic motion possible
 - quietest differential acceleration of the two TMs
 - $3 \times 10^{-14} \text{ m s}^{-2} / \sqrt{\text{Hz}}$ 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



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		
			1		Disch		Duyo	Discharge	
		H2	2	Noise Run		at in an	Neles		
		H3	3		Work Poi		Noise Run	Stray Potentials	
		H4		Sys ID					
		HS	>						



LP Data Analysis Activities



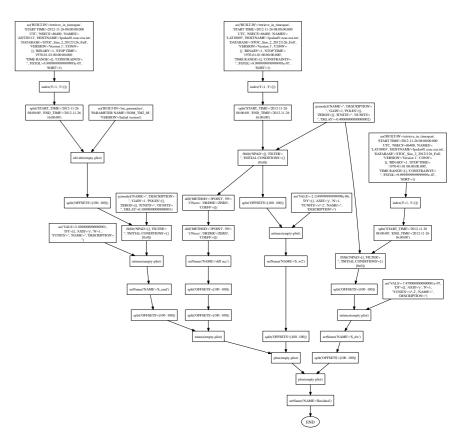
- Analysis software and infrastructure
- Defining, simulating and testing experiments
- Analysis pipelines
- End-to-end tests (including hardware wherever possible)
- Interfacing with ESA's STOC and MOC



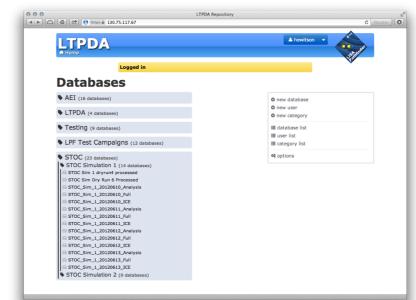
Analysis software and infrastructure

- 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/



lisa pathfinder





Science Operations



- 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

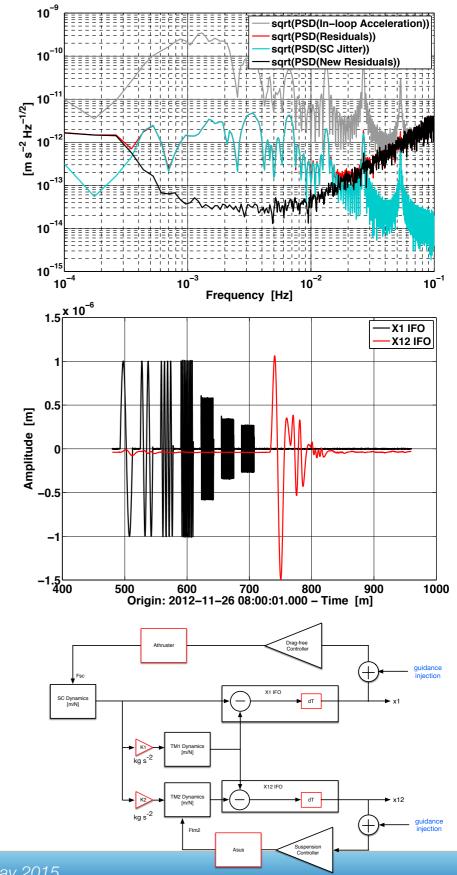


Spectral estimation

- 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?



- 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



- 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





From either missing data, or unusable data:

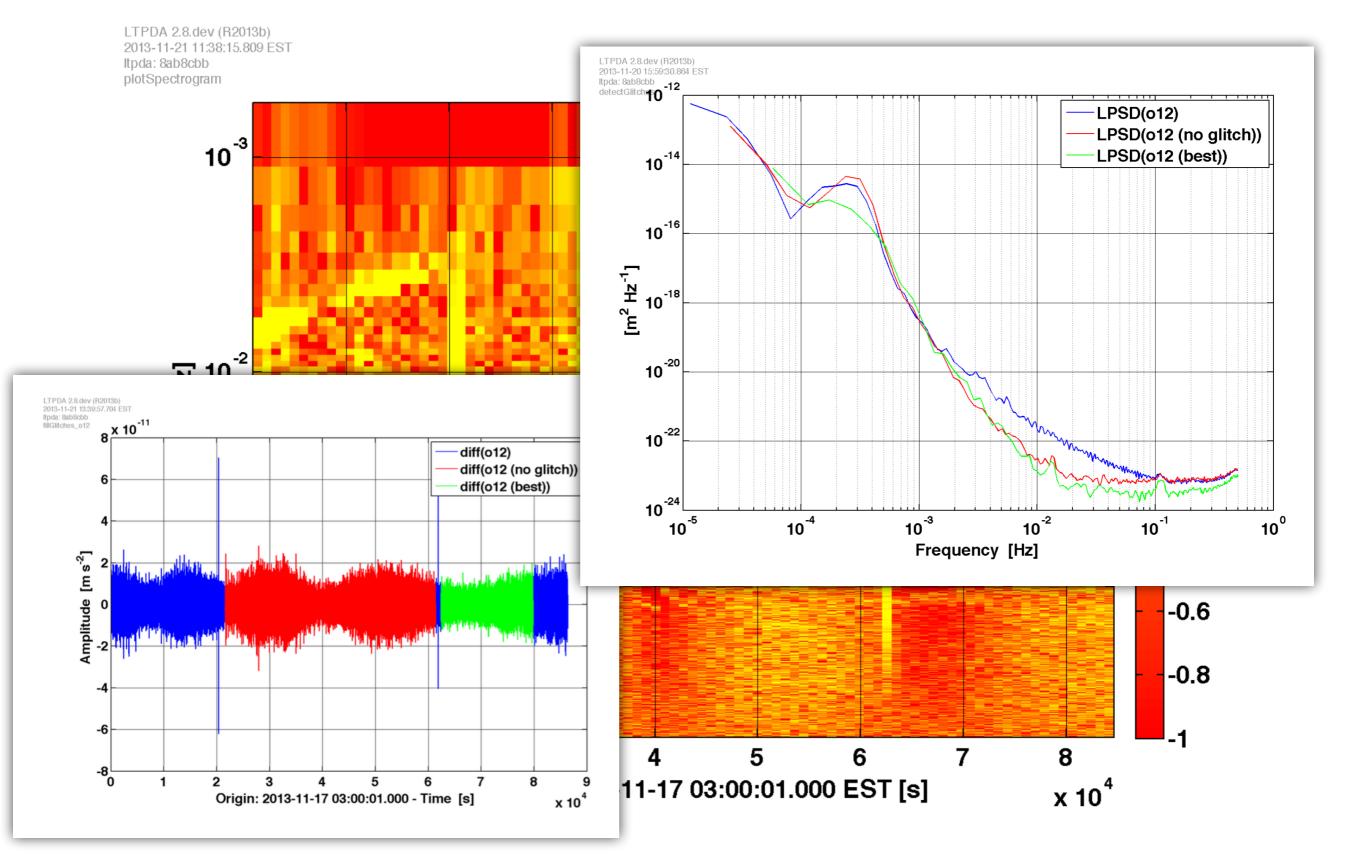
- Ignore
 - analyse segments either side
- Interpolate
 - for simple, large and/or 'slow' signals, this might be ok
 - for tiny gaps (a few samples), maybe even ok for noise only data
- Fill with representative simulated noise
 - with the goal to allow standard tools to be used
- Other techniques we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



OMS Test-campaign data

Ę





Gap filling



- 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

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

• Treat data inside and outside the gaps

$$\ln P(\vec{y}) = const - \frac{1}{2} \left[\sum_{j=1}^{\bar{\mathcal{G}},\bar{\mathcal{G}}'} C_{jk}y^jy^k + \left\{ \begin{array}{c} \text{Inside gaps to} \\ 0 \text{utside gaps} \\ 2\sum_{j=1}^{\bar{\mathcal{G}},\mathcal{G}} C_{jk}y^jy^k \\ 2\sum_{j=1}^{\bar{\mathcal{G}},\mathcal{G}} C_{jk}y^jy^jy^k \\ 2\sum_{j=1}^{\bar{\mathcal{G}},\mathcal{G}} C_{jk}y^jy^jy^k \\ 2\sum_{j=1$$





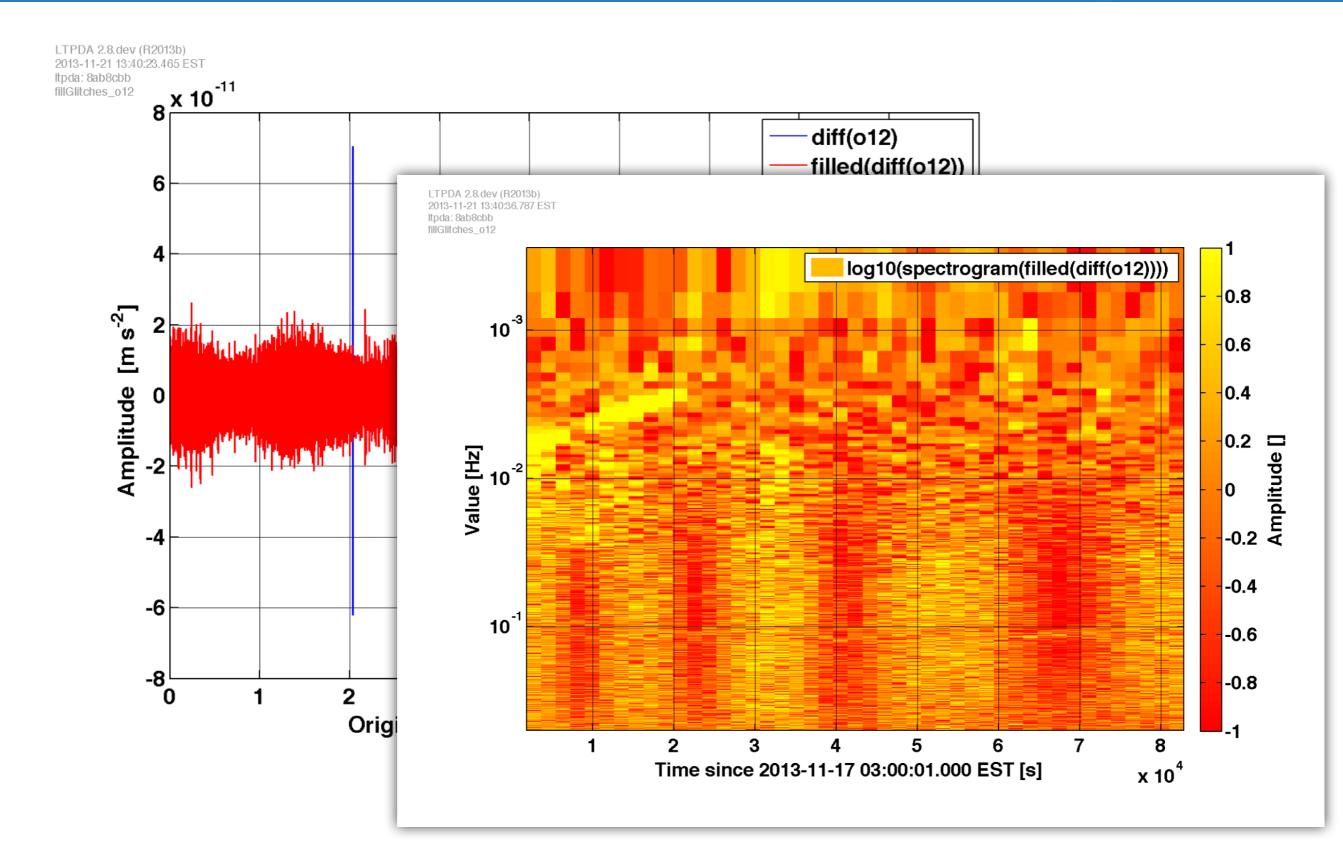
We want to generate replacement data for the gaps which has the correct spectral properties and correlates 'properly' with the data we have.

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



Ę

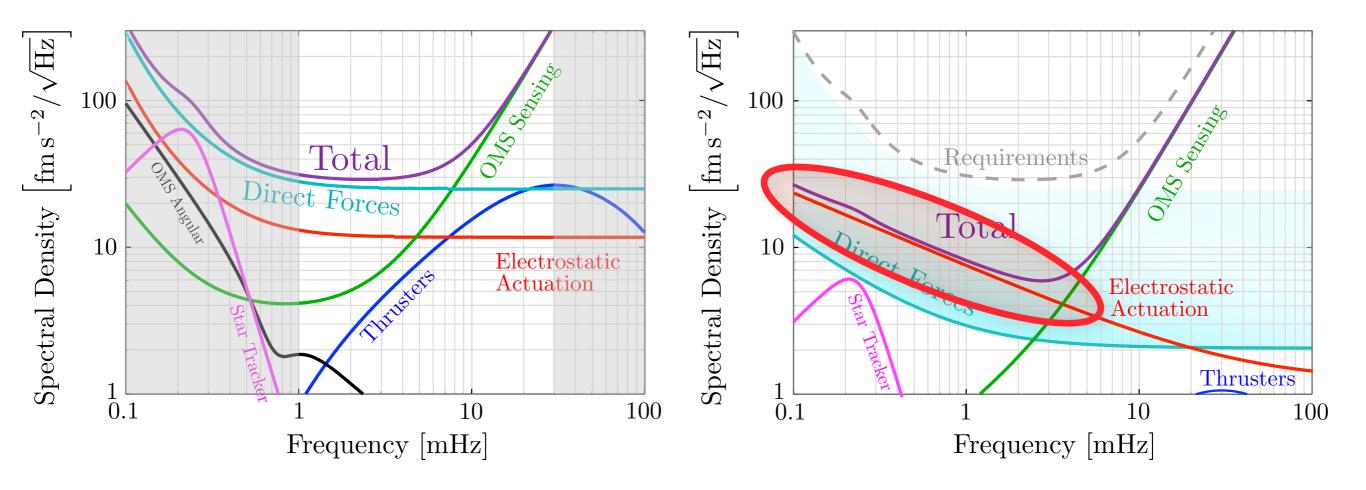
lisa pathfinder

10

Another source of 'gaps'



 Ones we create ourselves as a result of an experiment





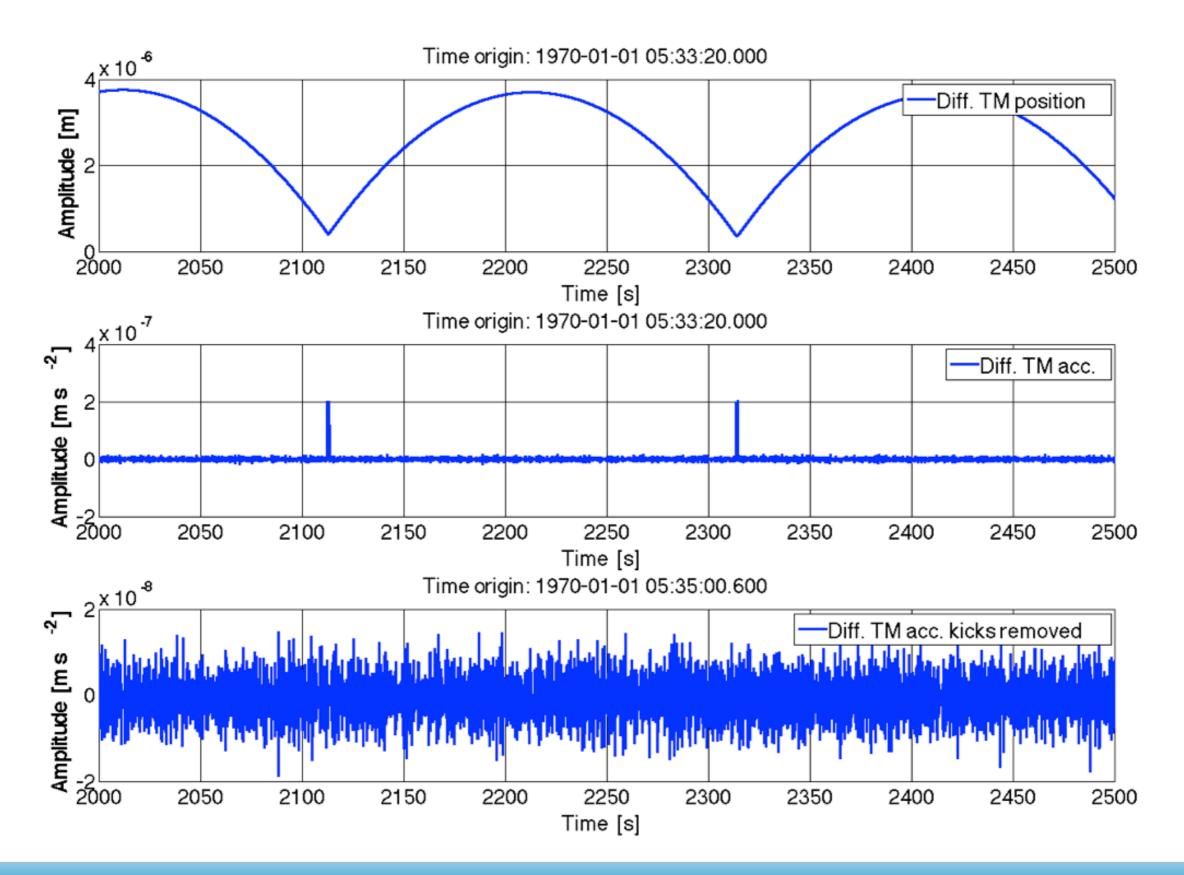
- Capacitive actuation may be limiting in our measurement band
- Do an experiment with the actuation off
 - must be short otherwise the TM will drift too far
 - repeat many times
- •Experiment:
 - kick test-mass away
 - turn off actuation
 - let the test-mass drift (in parabola)
 - repeat



Simulation

Ę

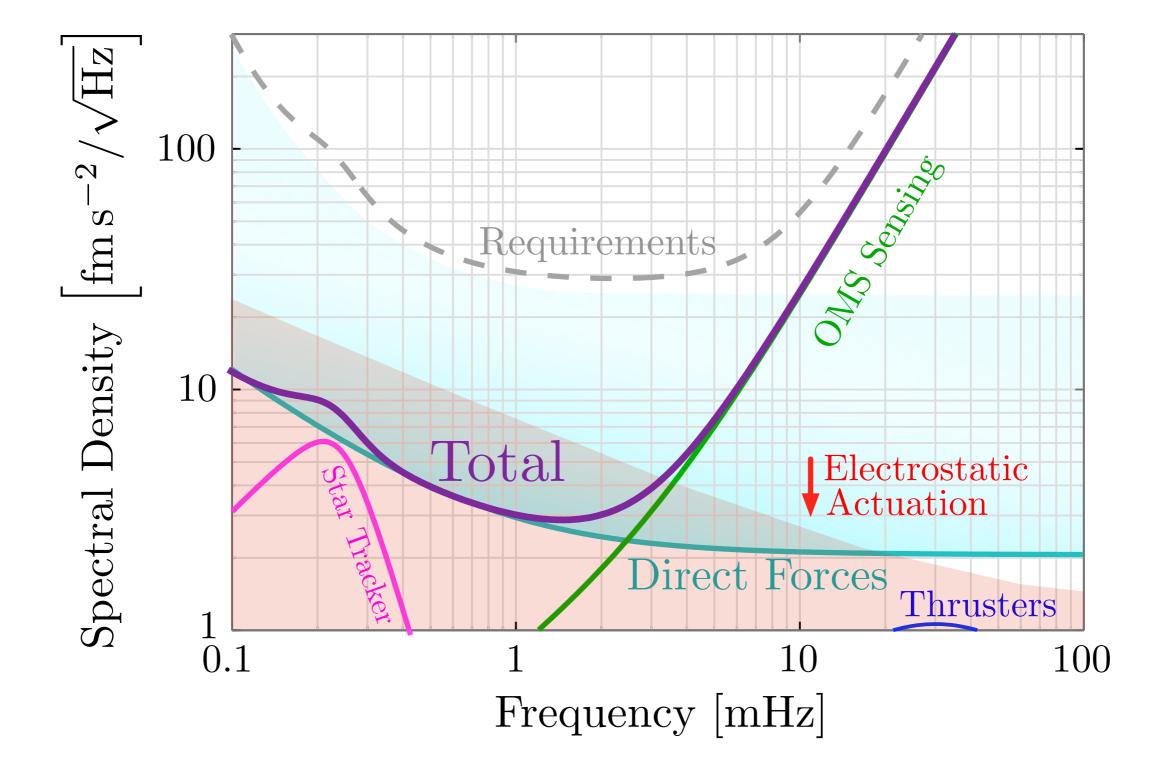






- Need to analyse the multiple short (200s) drift segments to estimate the spectrum at 1mHz
- One approach: window the data, and proceed with normal PSD estimate
- Alternative methods include replacing the noise data during the kicks with noise of same spectral content as drift phases to allow standard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

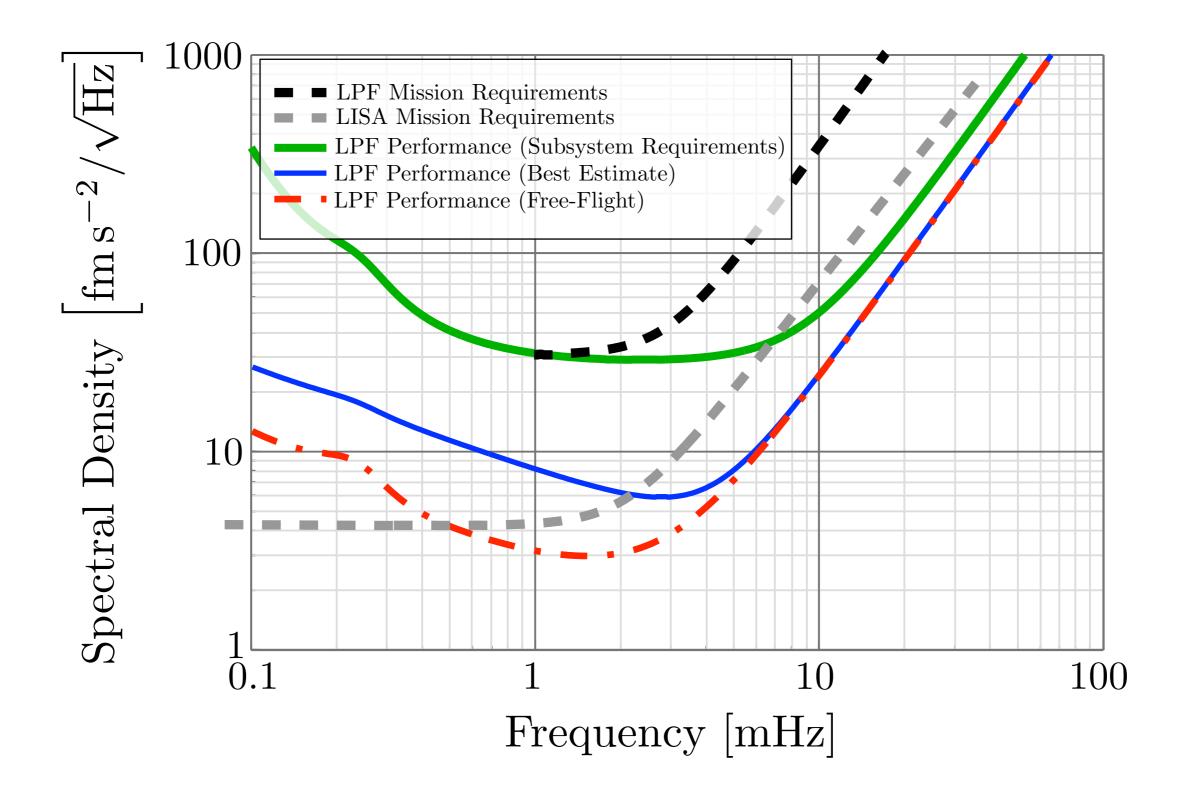




lisa pathfinder

ß



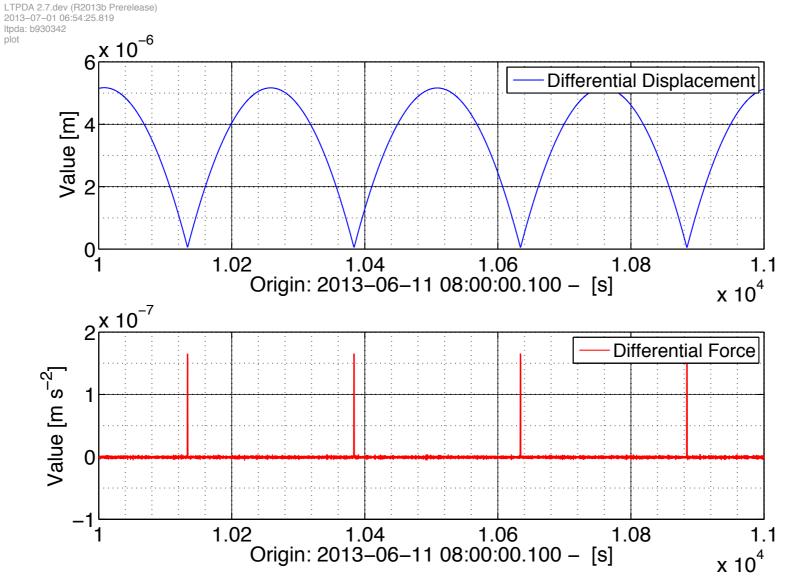


Suppress gaps: "windowing"



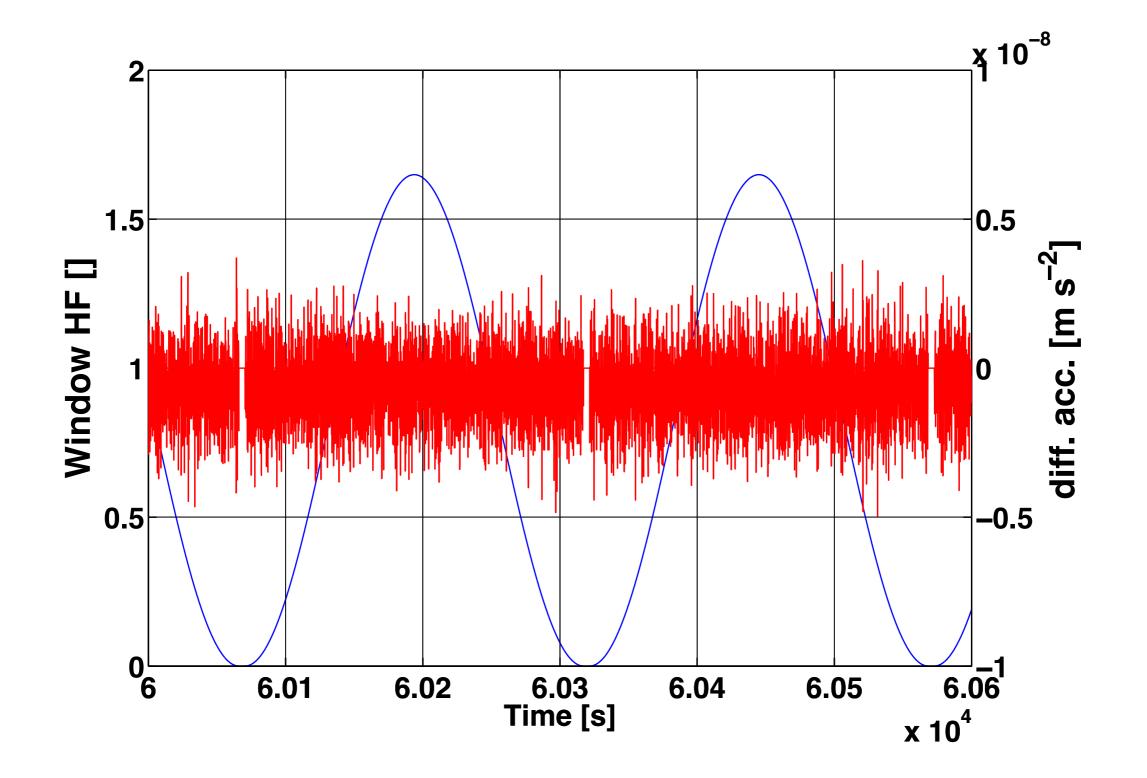
- For regular gaps, we can suppress the contribution by windowing
- Take second derivative of position:

Ę



Apply windows





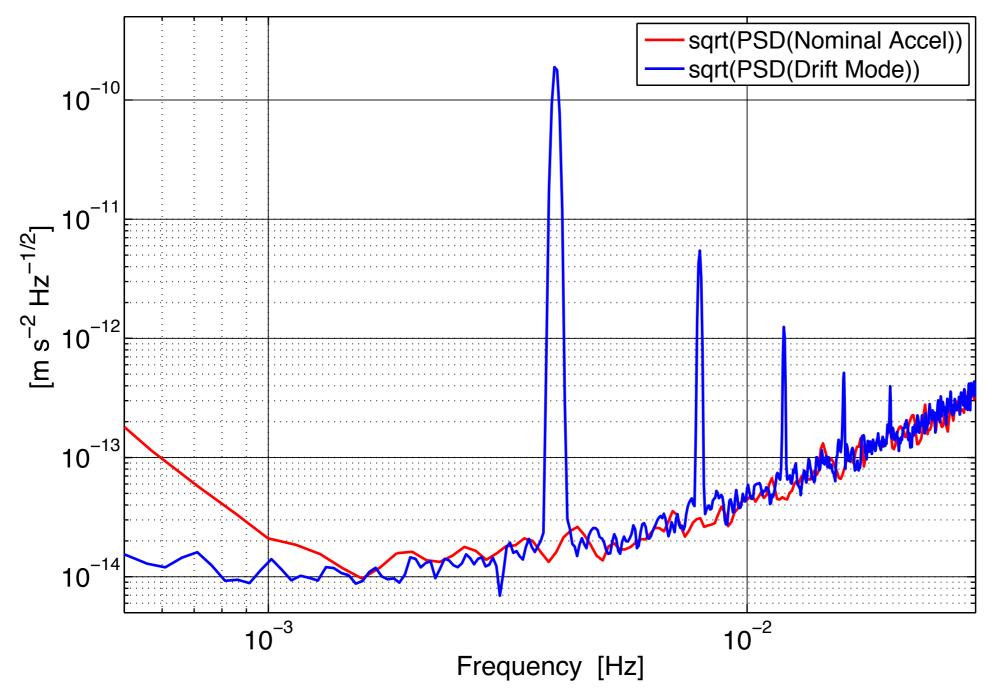


23

Take spectrum



Residual acceleration







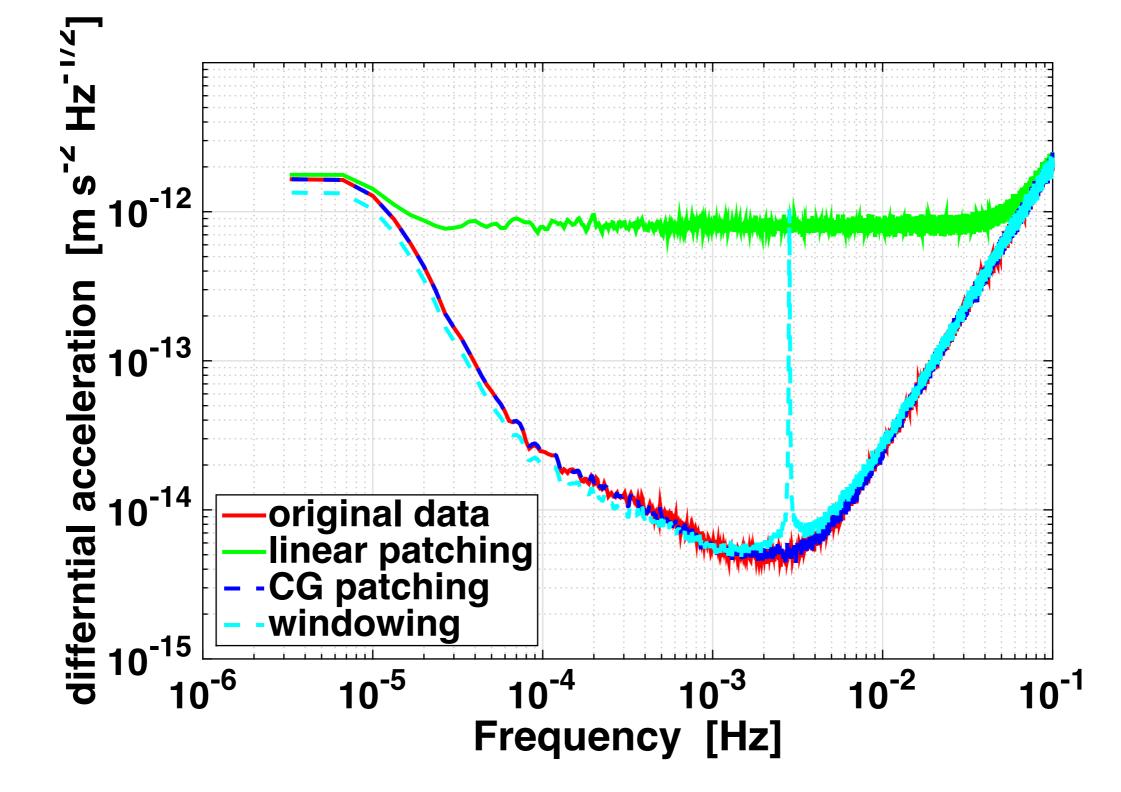
- Again, start from a model spectrum
 - fit to data from before free-flight experiment
- Identify each 'gap' corresponding to the noisy force kicks
- Generate data which has the correct spectral characteristics
 - correct gaussian noise within the gaps
 - correct correlations on longer timescales



Simulated LPF noise with gaps

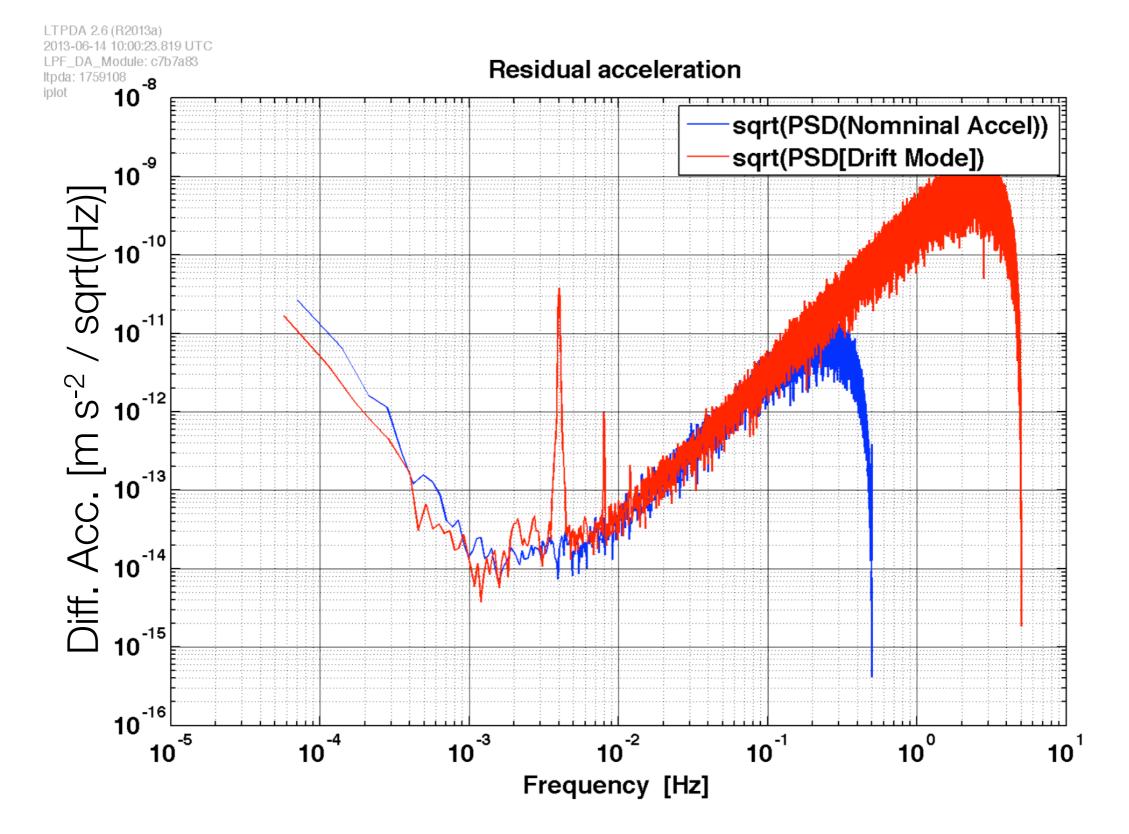
G





Simulated LPF free-flight data









- 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?







- 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:

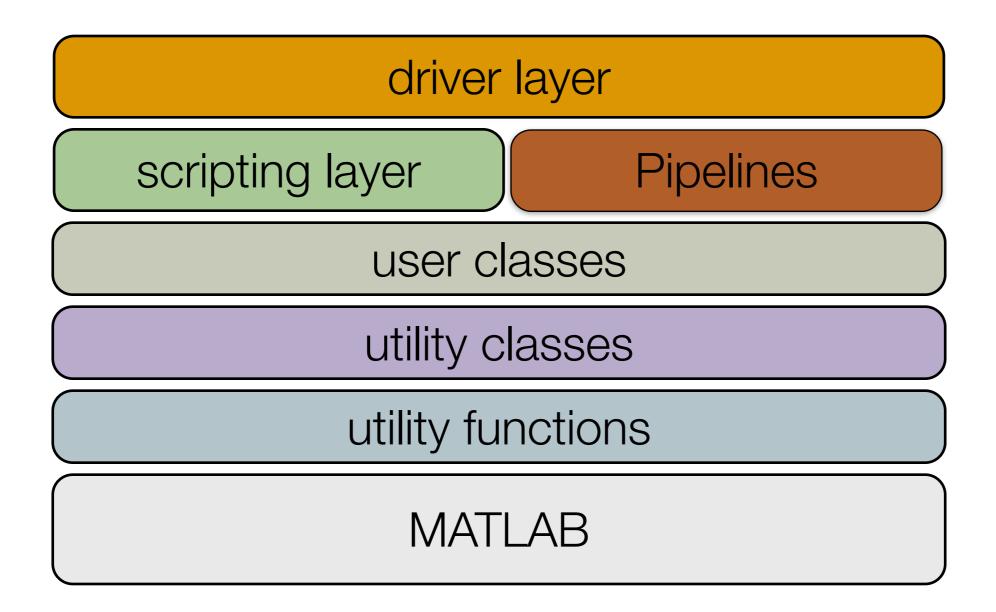
 provide a data analysis environment suitable for analysing the experiment data in 'real-time'

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





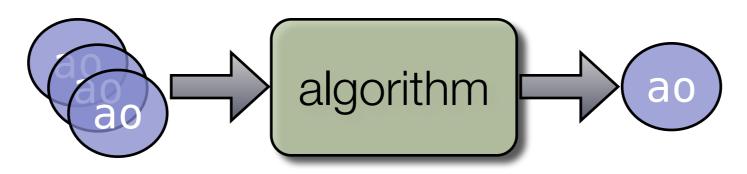




Objects, objects everywhere...



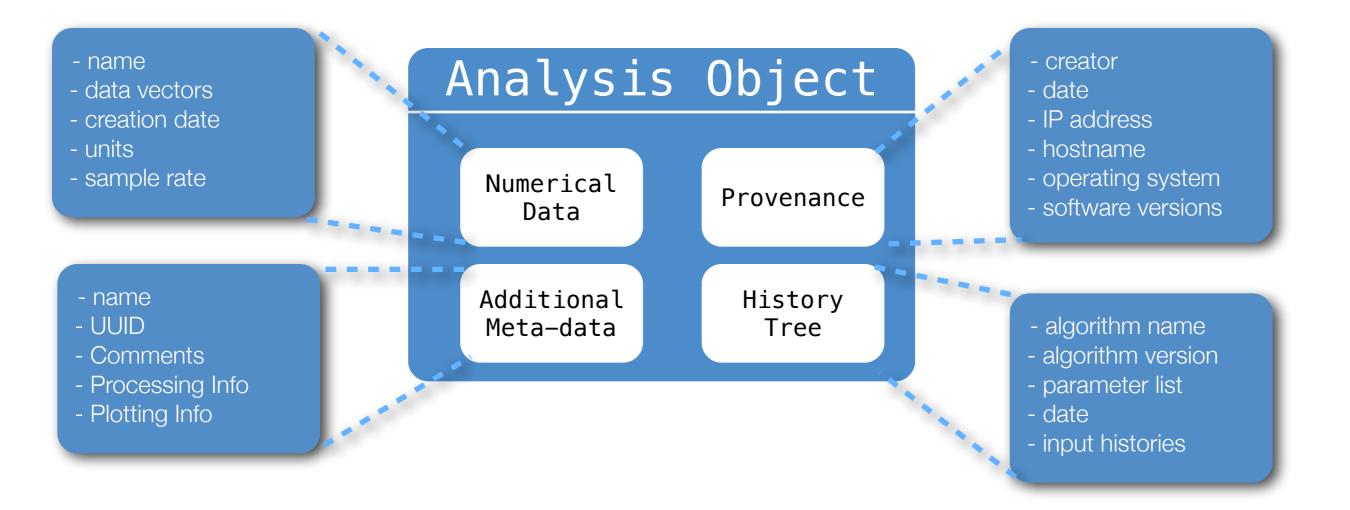
- LTPDA offers an object-oriented data analysis framework
 - we encapsulate (describe) different data analysis concepts with classes
 - users instantiate (build) these classes to get ltpda objects
 - DA algorithms are methods (functions) of these different classes
 - users act on the objects using the 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 abs acos angle ao asin atan atan2 bilinfit bin_data bsubmit buildWhitener1D cat char cohere complex compute confint conj consolidate conv convert copy corr COS COV cpsd crbound created creator csvexport ctranspose curvefit delay delayEstimate demux det

Ę

detrend dft diag diff display dopplercorr downsample dropduplicates dsmean dx dy eiq eq eamotion evaluateModel exp export fft fftfilt filtSubtract filter filtfilt find firwhiten fixfs fngen fromProcinfo fs gapfilling gapfillingoptim qe get aetdof gnuplot at heterodyne hist

hist_gauss hypot ifft imaa index integrate interp interpmissing inv iplot iplotyy isprop isvalid ioin 1cohere lcpsd le len linSubtract lincom linedetect linfit lisovfit ln log log10 lpsd lscov lt ltfe ltp_ifo2acc max mcmc md5 mdc1_cont2act_utn mdc1 ifo2acc fd mdc1_ifo2acc_fd_utn

mdc1_ifo2acc_inloop mdc1 ifo2cont utn mdc1_ifo2control mdc1 x2acc mean median min minus mode mpower mrdivide mtimes ne noisegen1D noisegen2D norm normdist nsecs offset optSubtraction phase plot plus polyfit polynomfit power psd psdconf pwelch quasiSweptSine rdivide real rebuild removeVal report resample rms

rotate round sDomainFit save scale scatterData search select setDescription setDx setDy setFs setMdlfile setName setPlotinfo setProcinfo setT0 setUUID setX setXY setXunits setY setYunits setZ sign simplifyYunits sin sineParams smallvector_lincom smallvectorfit smoother sort spectrogram spikecleaning split spsd sqrt

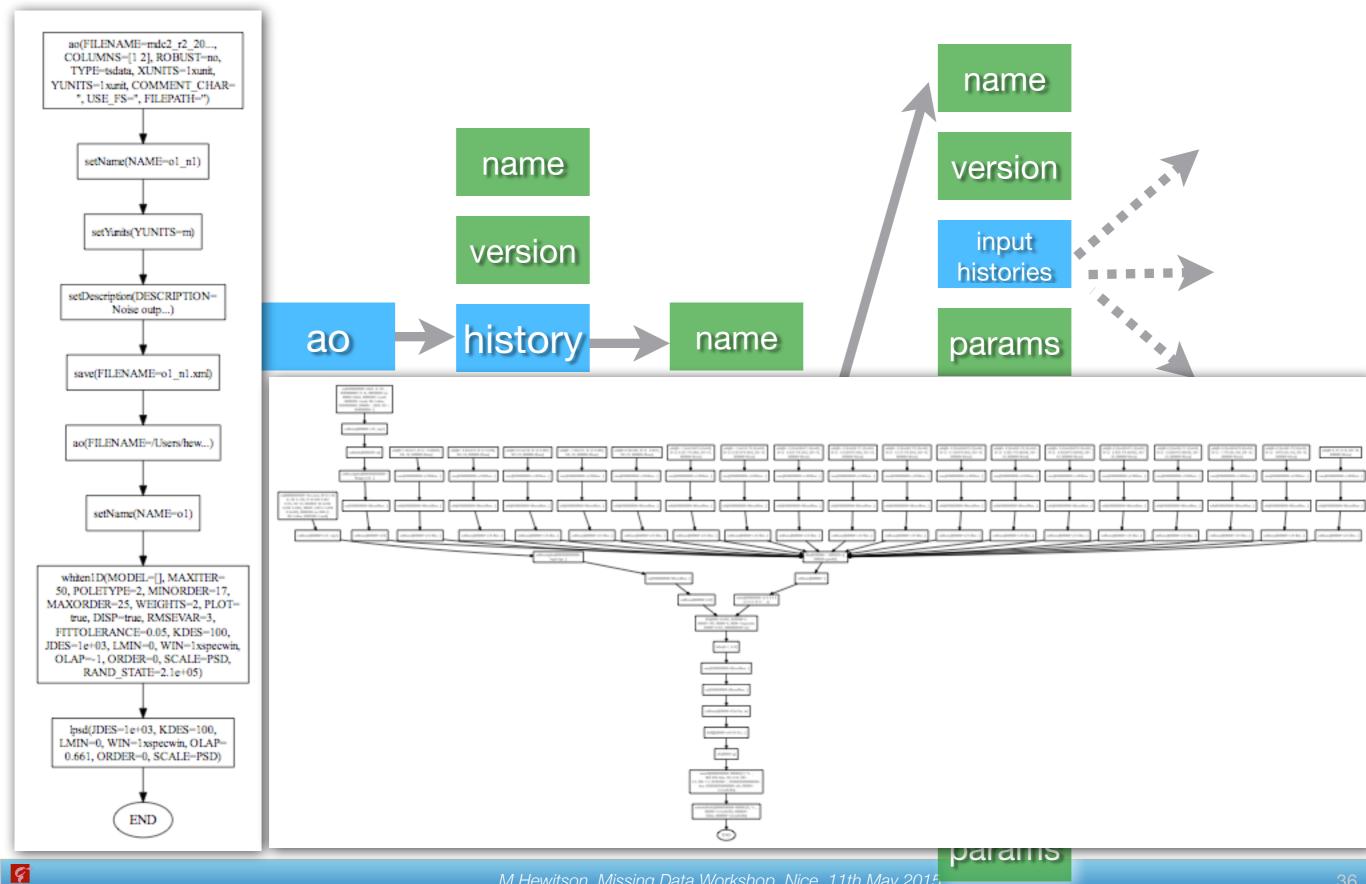
std straightLineFit string submit sum sumjoin svd svd fit t0 table tan tdfit tfe timeaverage timedomainfit times timeshift transpose type uminus unwrap update upsample validate var viewHistory whiten1D whiten2D х xcorr xfit xunits У vunits

zDomainFit

zeropad

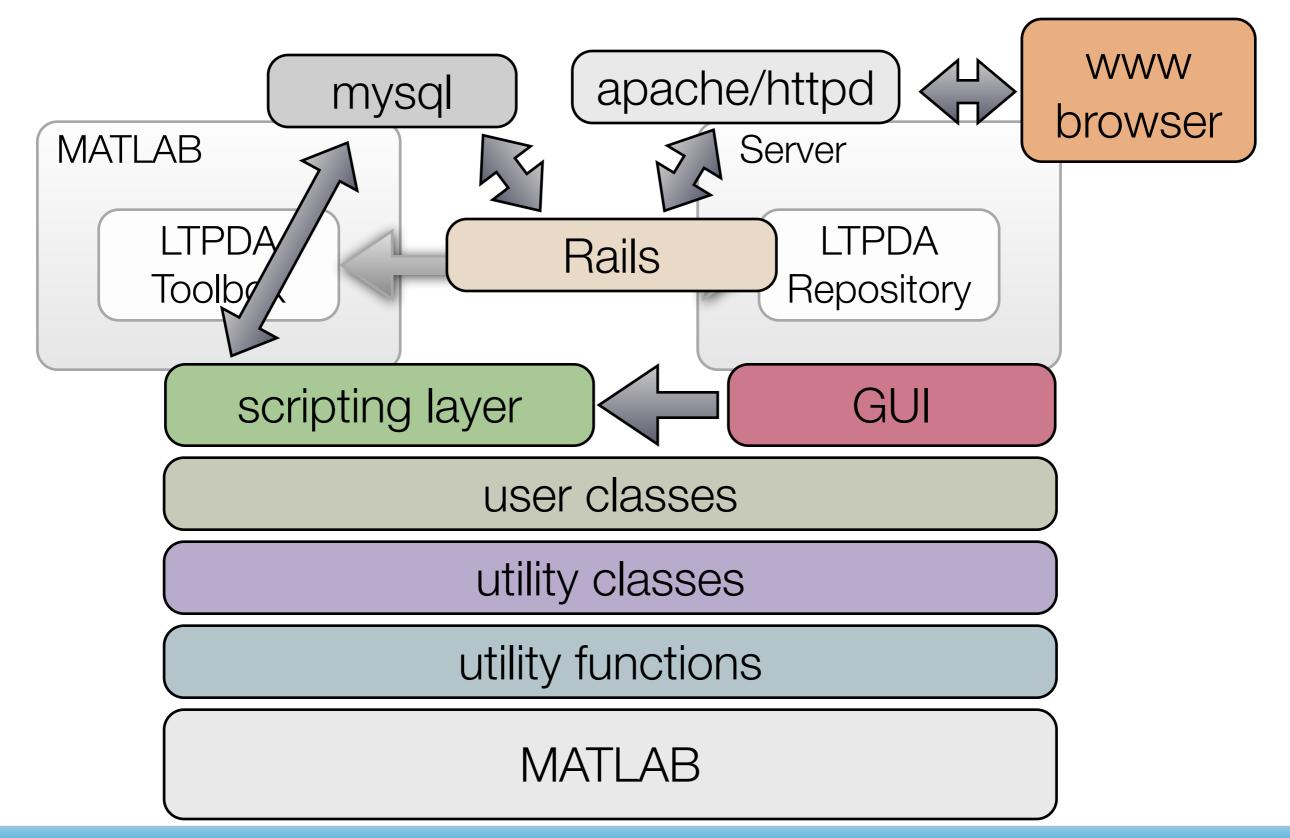
Don't touch me with that AO, I don't know where it's been!





Where can I put my AO?





LTPDA Repository

G



● ● ● LTPDA Reposit ● ● ● ● ● ● LTPDA Reposit ● ● ● ● ● ● LTPDA Reposit ● ● ● ● ● ● □ LTPDA Reposit ● ● ● ● ● ● □ LTPDA Reposit ● ● ● ● ● 130.75.117.67 □ □ LTPDAA ● ● ● ● ● ● ● ● Home ● <th>ory</th> <th>Manage User Accounts</th>	ory	Manage User Accounts
Logged in Databases AEI (25 databases)	• new database	Manage Databases
 LTPDA (19 databases) Testing (9 databases) LPF Test Campaigns (14 databases) 	 new user new category iii database list iii user list iii category list 	Browse/Query Databases
STOC (92 databases) LTPDA Interface Version 2.6 LTPDA Database Version: 2.2 Repository Connection	*8 options	Submit Objects
hostname 130.75.117.67 130.75.117.67 username hewitson	O Query: 130.75.117.67/ltpda_test From table objmeta select id obj_id	Retrieve Objects
database Itpda_test Cet list	obj_type name created version ip hostname os submitted experiment title	
	order by id DESC : SELECT obj_type,name,created,ip FROM objmeta ORDER BY id DESC;	
	Done Execute	

Documentation



Documentation

Ţ

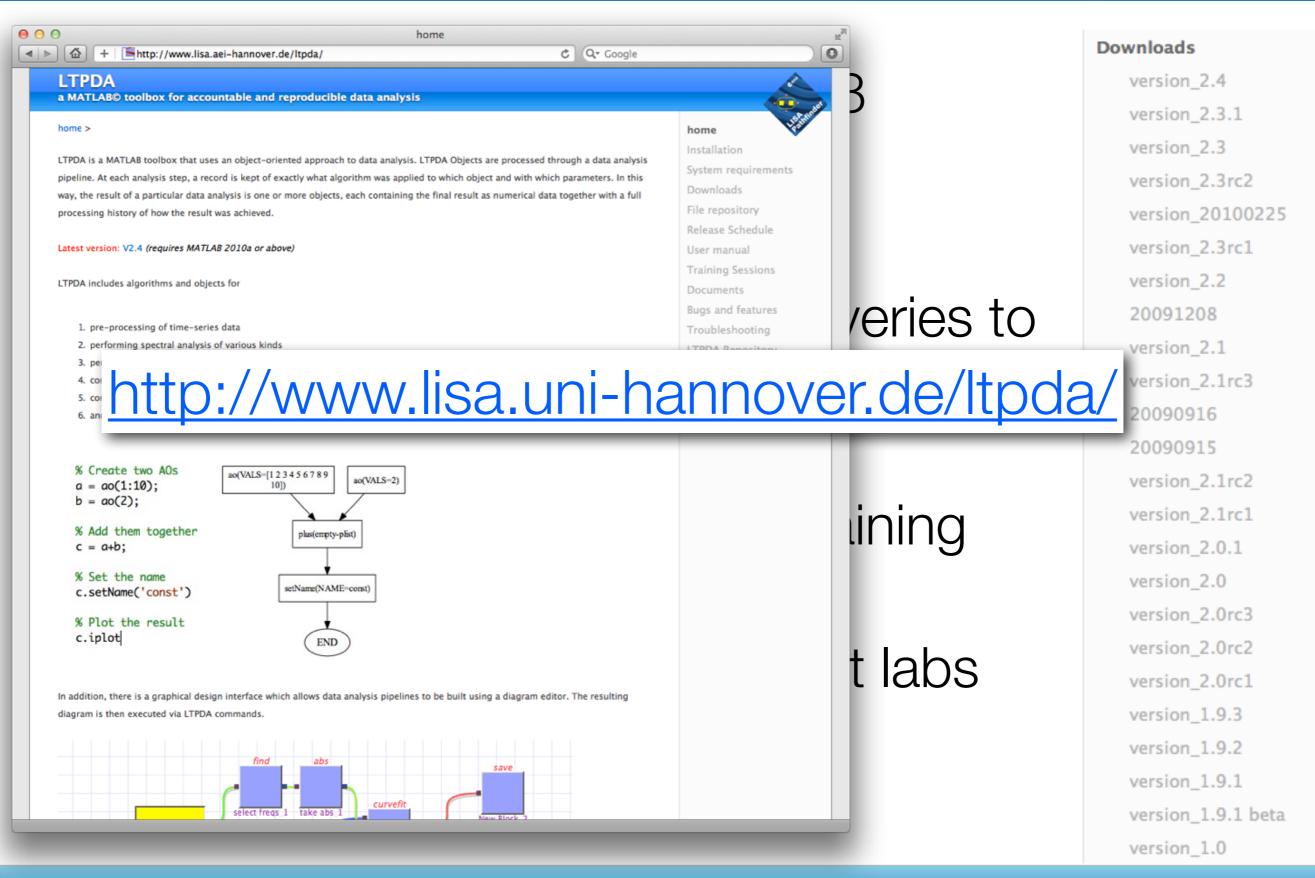
• user-manual ~700 pages

~4500 documented m-files (methods/functions)

	000	Help	😋 analysis object class constructor.
	Qr Search *	★ → ★· ♀ → LTPDA Toolbox → LTPDA Training Session 1 →	
	Contents Search Results	LTPDA Toolbox	ESCRIPTION: ao analysis object class constructor.
Model Report for matrix/m	► 🤣 Release Notes ► 🤣 Installation ► 🥔 MATLAB	LTPDA Training Session 1	Create an analysis object.
A built-in model that constructs a model of the MDC3 LT	 MATLAB Code Generation from MATLAB 		Possible constructors:
The control matrix is defined as:	 Control System Toolbox Fixed-Point Toolbox 	This series of help pages consitute the first training session of LTPDA. The various data-packs used throughout the tutorials are available for download on the LTPDA web-site.	<pre>a = ao() - creates an empty analysis obje a = ao('a1.xml') - creates a new ao by loading a</pre>
/ (1+dHdf)*Hdf 0 \ H = \ 0 (1+dHsus)*Hsus /	 	1. <u>Topic 1 – The basics of LTPDA</u> 2. <u>Topic 2 – Pre-processing of data</u>	<pre>a = ao('a1.mat') a = ao('a1.mat') - creates a new ao by loading t</pre>
Model Description	 Generative Introducing LTPDA Objects Parameter Lists 	3. <u>Topic 3 - Spectral Analysis</u> 4. <u>Topic 4 - Transfer function models and digital filtering</u> 5. Topic 5 - Model fitting	a = ao('file.txt') - creates a new ao by loading t
LSS v2.0.0 This version is designed to mimic the c The following fit characteristics were ac • drag free accurate to 4e-9 MSE or	Built-in models of LTPDA	In addition, throughout the course of this training session, we will perform a full analysis of some lab data. The inputs to the analysis are two time-series data streams, the first is the	<pre>a = ao('file.dat') a = ao('file',pl) (Set: From ASCII File)</pre>
 ency new accurate to the share of a share	 Parameteric models 	recorded output of an interferometer, the second is a recording of the room temperature in the vicinity of the interferometer. Both are recorded with different sample rates and on different sampling grids. The temperature data is unevenly sampled, and may evem have	<pre>a = ao(data) - creates an ao with a data objo a = ao(constant) - creates an ao from a constant</pre>
STOC Ex 6 This version was used in STOC Exercise The following fit characteristics were ac	 Transfer Function Modellin Signal Pre-processing in L1 	missing samples. During each topic of the training session, the data will be manipulated using the tools	a = ao(specwin) - creates an ao from a specwin
 drag free accurate to 1e-8 MSE or s^-2, input is in m electrostatic suspension accurate 	Graphical User Interfaces ir	introduced in that topic of the framing session, the data will be manipulated using the tools introduced in that topic (and previous topics). The aim of the data analysis is to determine the influence of temperature on the interferometer output. In particular the steps will be:	a = ao(pzm, nsecs, fs)
output is in kg m s^-2, input is MDC3 This version was used in MDC3. Implemented transfer functions are designed.	LTPDA Extension Modules	 Topic 1 Loading and calibrating the raw data. Read in the raw data files and convert them to AOs 	<pre>a = ao(smodel) - creates an ao from a symbolic a = ao(pest) - creates an ao from a paramete</pre>
Some information of the method matrix/matrix_mode	fx Functions - By Category fx Functions - By Category Y Category Image: Category Category	 Plot the two data streams Calibrate the interferometer output to meters (from radians) Calibrate the temperature data to degrees Kelvin from degrees Celcius 	a = ao(x,y) - creates an ao with xy data a = ao(y, fs) - creates an ao with time-series
Class name matrix Method name matrix_model_H	 Topic 1 - The basics of LTI Topic 2 - Pre-processing c 	 Save the calibrated data series to XML files, ready for the input to the next topic Topic 2 Pre-processing and data conditioning. 	a = ao(y, 1s) = creates an ab with time-series a = ao(x,y,fs) = creates an ab with time-series
CVS Version \$Id: matrix_model_H.m,v 1.		 Read in the calibrated AOs from XML files Trim the data streams to the same time segments 	a = ao(x,y,pl) - creates an ao depending from
LSS v2.0.0 This version is designed to mimic the controllers in LSS V	Topic 5 - Model fitting Examples	 Resample the temperature on to an even sampling grid with no missing samples Resample to the two data streams to a common 1Hz sample rate Interpolate the two data streams on to the same time grid 	a = ao(plist) - creates an ao from a paramete
The following fit characteristics were achieved: • drag free accurate to 4e-9 MSE on the frequency ra	P P Demos P P Release Notes	6. Save the cleaned data to AO XML files 3. <u>Topic 3</u> Spectral analysis.	kamples
 electrostatic suspension accurate to 7e-11 MSE on s^-2, input is in m 	LTPDA Web Site A Ontimization Toolbox	 Load the time-series data from Topics 1 and 2 Compare PSDs of the time-series data before and after pre-processing Check the coherence of temperature and IFO output before and after 	arameters Description
Parameter List for version: LSS v2.0.0		nro_prococcipa	EDCTON: 4Td. 20 m v 1 240 2011/05/16 07:15:27 haveters 5
Key Default Value	Options	Description	
	1.09174245874697e-05;1.140723270 none	a frequency vector to be used as xvals (Hz)	

LTPDA







- 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



Launch	IOCR										
Launch, LEOP, Transfer, Separation, De-spin	Commiss	ioning	LT	P Science	Ops	Con	DRS nmissioning	DRS Operations			
60 days	14 da	ys		3 months	6	1	0 days	3 months			
					Day	/2	Day 3	Day 4			
		H		Day 1		arge	,	Discharge			
		H2	?	Noise Run	Mori	din a	Neise				
		H3	3		Working Point		Noise Run	Stray Potentials			
		H4		Sys ID				rotentials			
		Hŝ	>								



Week 1: Gentle Probing

- 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

	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	1 Noise run in Sci 1.2																					
2	CE1	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			S	Station Keeping Transition Acc3 -> Sci 1.2									F	D1	FI	D2						







Charge estimate TM1

Charge estimate TM2

Fast Discharge TM1

Fast Discharge TM2

CE1

CE2

FD1

FD2



- 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

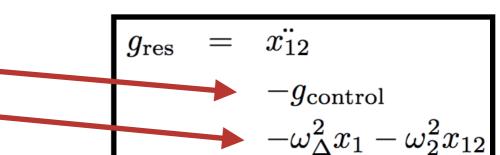


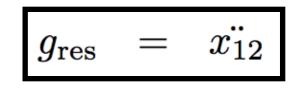
Estimating Residual Differential Acceleration

- 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
- Try to account for the contributions of g_res that we know
 - applied control forces

Ę

• couplings due to force gradients







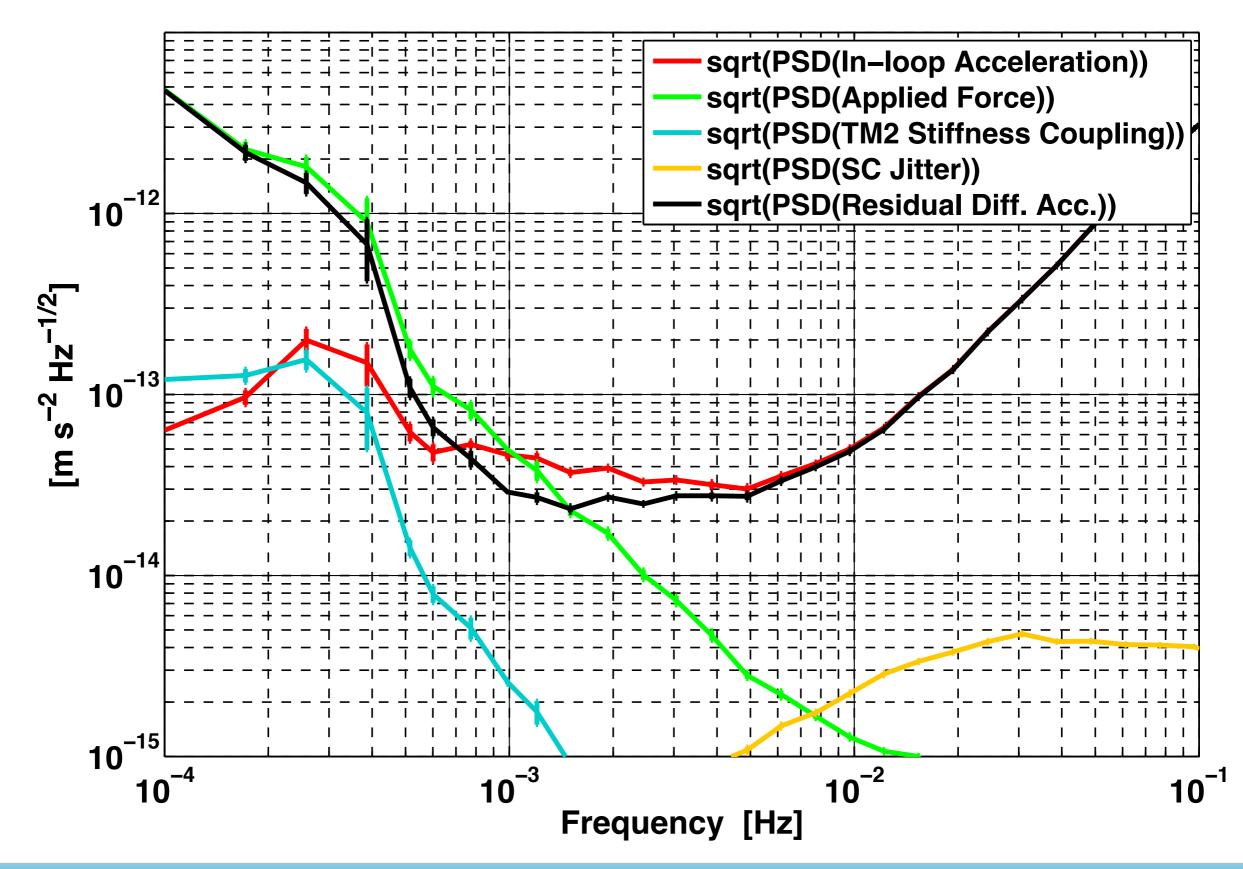


- 1. Download the time-series
- 2. Assemble the current best estimate of the required system parameters
 - actuator gains, delays, stiffnesses, ...
- 3. Form linear combination of the time-series
 - with delays, and filtering as necessary
- 4. Take spectrum of the residuals



The contributions





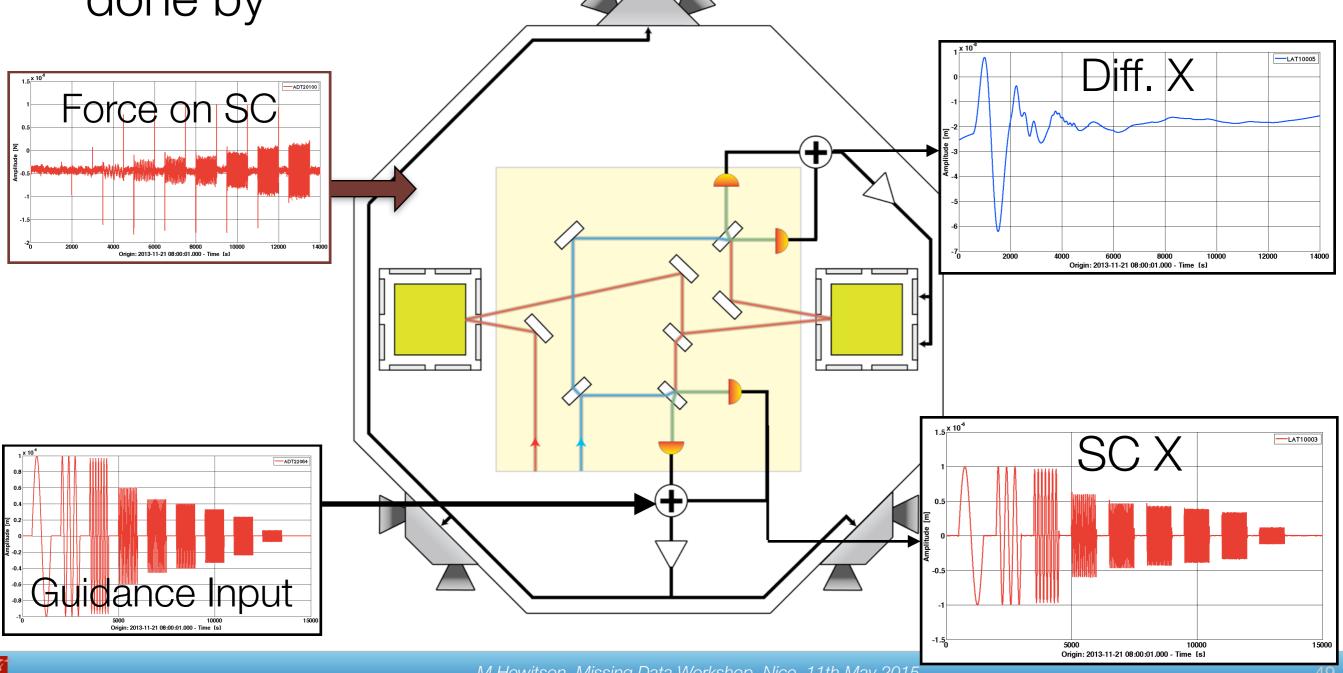


- Estimating our residual acceleration requires knowledge of certain system parameters
 - How do we gain that knowledge?
- At the beginning of operations, this comes from
 - ground measurements
 - system modelling
 - results of industrial commissioning campaign
- How do we improve and update that knowledge?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



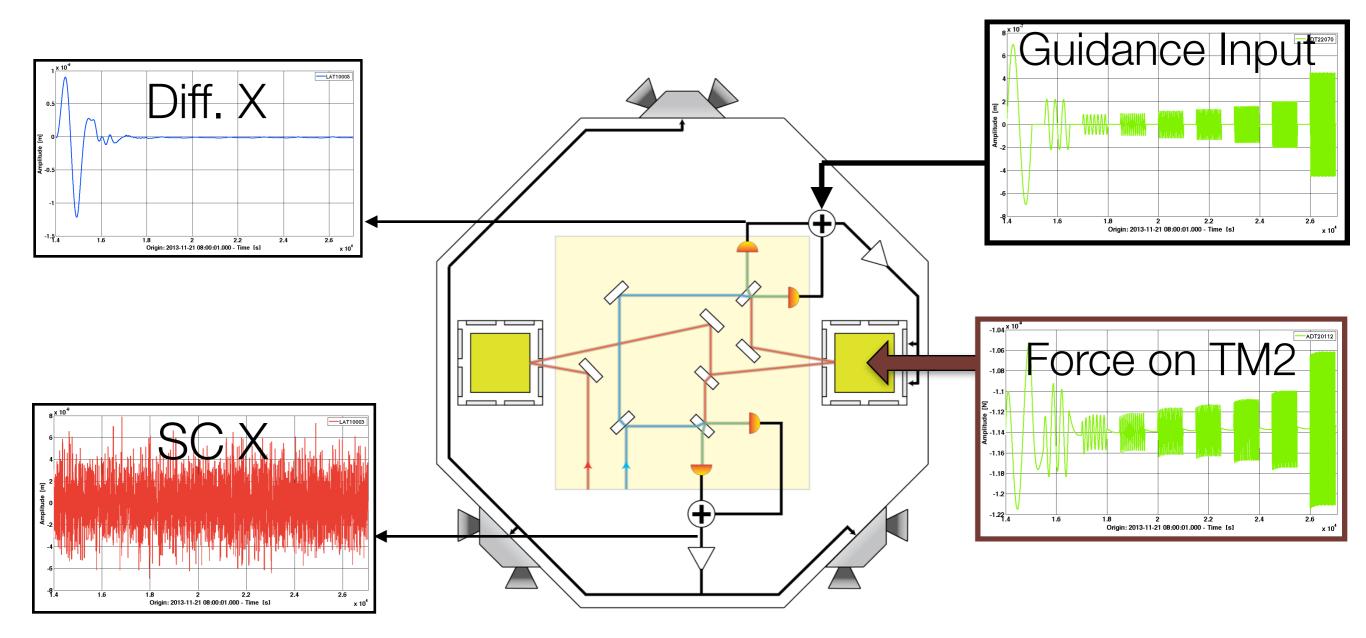
x-axis system identification: part 1

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



lisa pathfinder

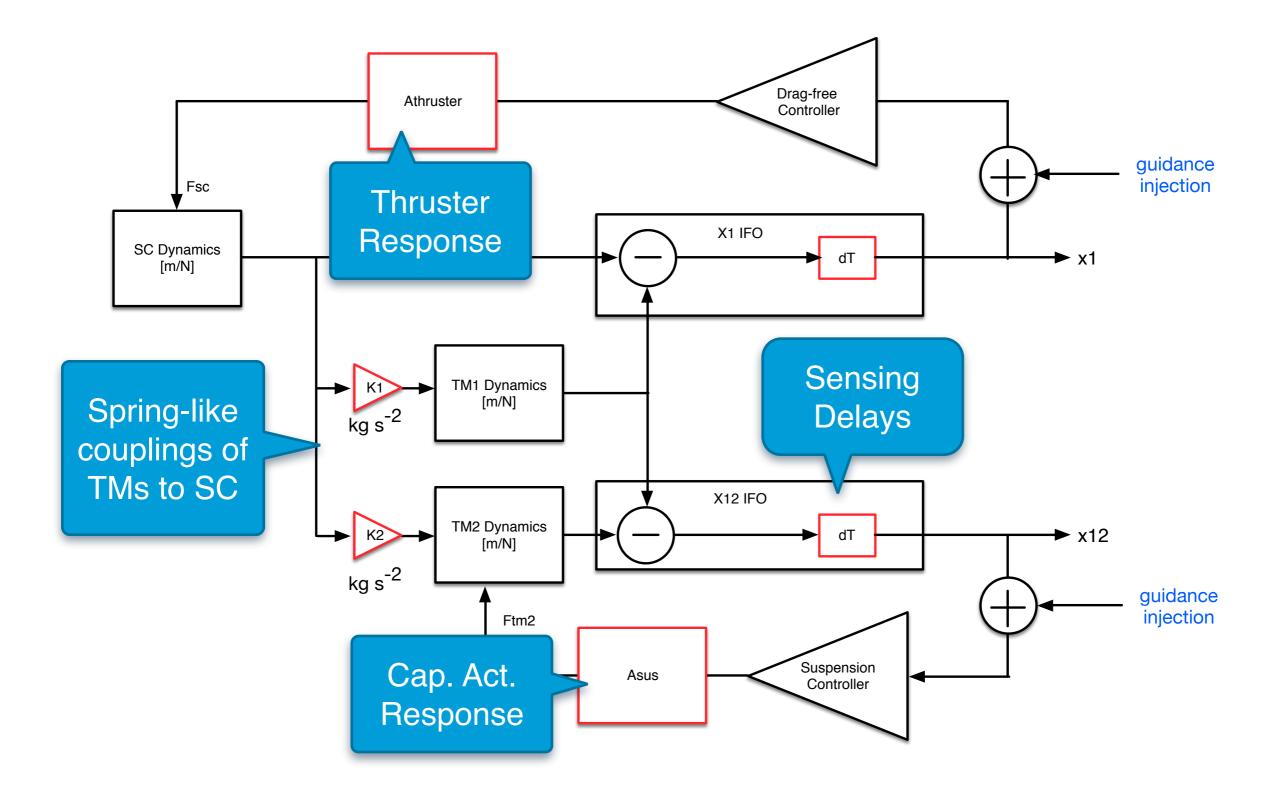
x-axis system identification: part 2



lisa pathfinder

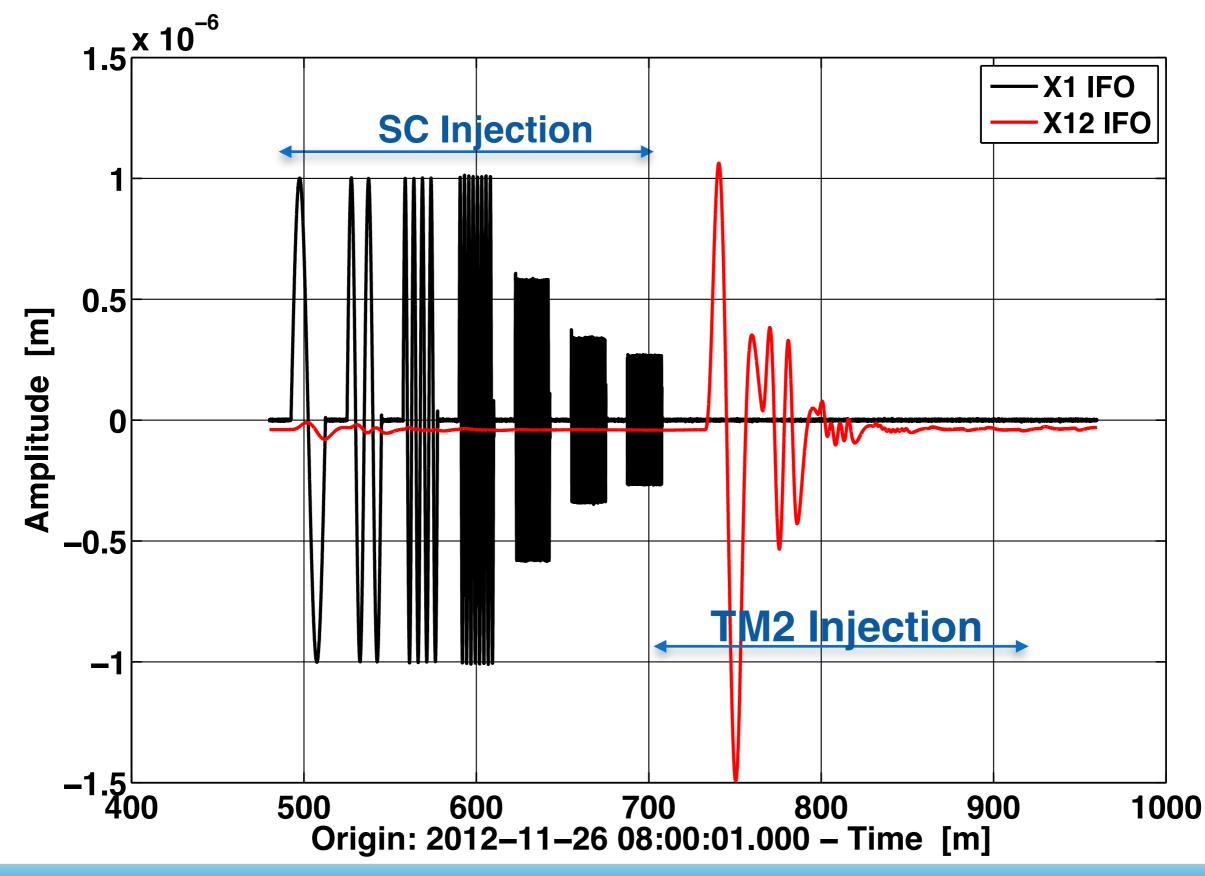
What do we learn from that?





The data





M Hewitson, Missing Data Workshop, Nice, 11th May 2015

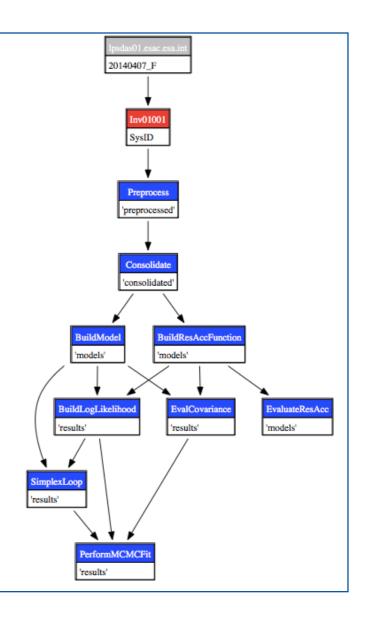
Analysis

- 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

G.

observation =
$$x_{12}[k]$$

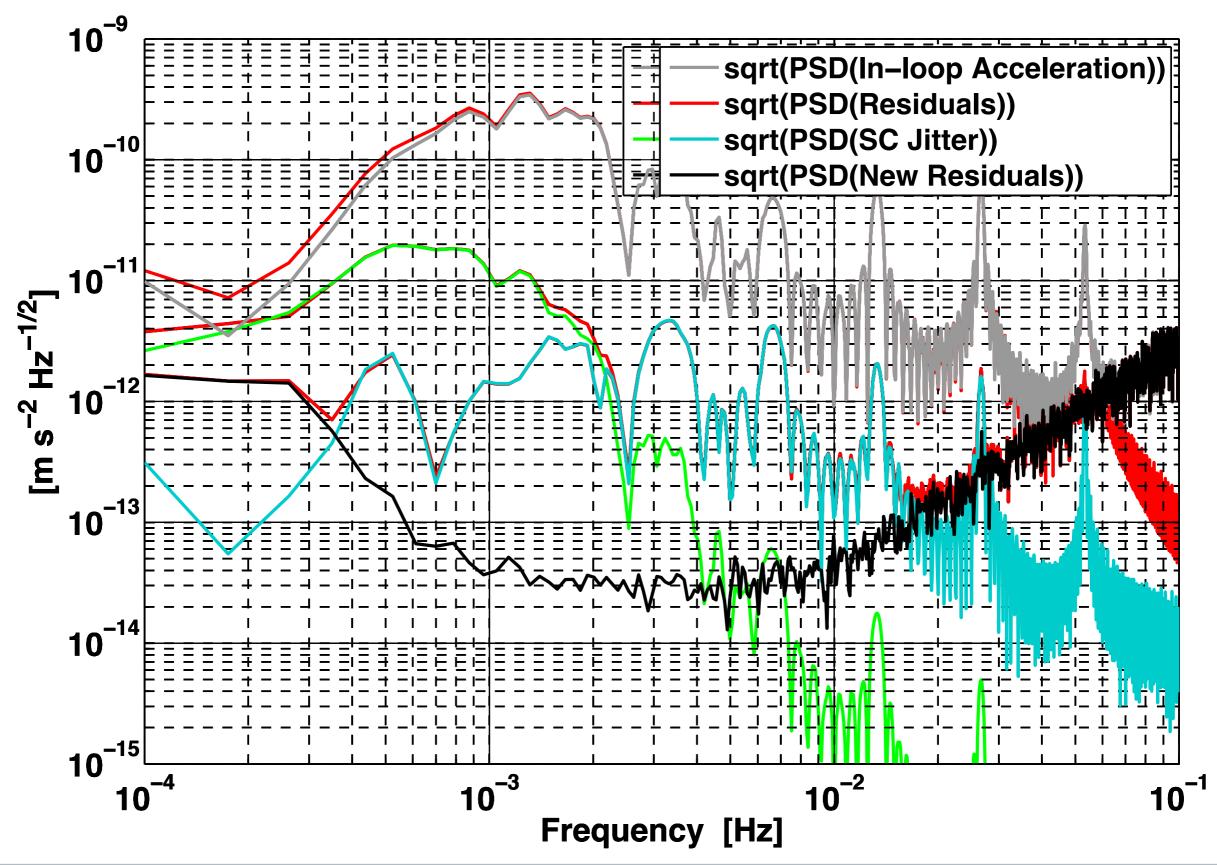
$$\begin{array}{lll} \mathrm{model} &=& -A_{\mathrm{sus}}F(g_{\mathrm{control}}[k],\Delta T) \\ && -(\omega_2^2 - \omega_1^2)x_1[k] \\ && -\omega_2^2 x_{12}[k] \end{array}$$





Residuals







• 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



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

