The Mock LISA Data Challenges

Michele Vallisneri
(Jet Propulsion Laboratory, California Institute of Technology)

for the MLDC Task Force:
Stas Babak, John Baker, Matt Benacquista,
Neil Cornish, Jeff Crowder, Curt Cutler, Shane Larson,
Tyson Littenberg, Edward Porter, M.V., Alberto Vecchio

MLDC what?

- Coordinated, voluntary effort in GW community
- Periodically issue datasets with synthetic noise and GW signals from sources of undisclosed parameters; increasing difficulty
- Challenge participants return parameter estimates and descriptions of search methods

MLDC why?

- For LISA, data analysis is integral to the measurement concept
  - We must demonstrate that we can meet the LISA science requirements (which promise astrophysical significance)
  - We need to understand data analysis quantitatively to translate science requirements into design decisions
- Kickstart the development of a LISA data-analysis computational infrastructure (will be needed later!)
- Encourage, track, and compare progress in LISA data-analysis development in the open community

Genesis (fall 2005)

AMIGOS
LISA Data-Analysis
Planning Meeting

LIGO experience with LAL, etc…

LISA Computational Infrastructure

Open development

Sensitivity Generator
LISA Calculator

LISA Simulator
Synthetic LISA

MLDCs

Astrogravs

Prove that LISA data analysis is feasible

Establish common playground for experiments

Encourage data-analysis development

Mock LISA Data Challenge Task Force

Members (past and present)

- Alberto Vecchio (co-chair)
- Michele Vallisneri (co-chair)
- Keith Arnaud
- Stas Babak
- John Baker
- Matt Benacquista
- Neil Cornish (WG-1B co-chair)
- Jeff Crowder
- Curt Cutler
- Sam Finn
- Steffen Grunewald
- Shane Larson
- Tyson Littenberg
- Eric Plagnol
- Ed Porter
- Sathyaprakash
- Jean–Yves Vinet

Charter

- Specify pseudo-LISA model
- Identify standard source models
- Specify data format (lisaXML)
- Plan challenge progression
- Prepare and distribute training and challenge datasets
- Develop software infrastructure
- Compile challenge submissions

MLDC timeline (so far)

Jan 2006
work begins!

Jun 2006
Challenge 1 datasets released at 6th LISA Symposium [proceeds, gr-qc/0609105-6]

Dec 2006
Challenge 1 results due; presented at GWDAW-11 [CQG 24, S529 (2007), plus several by participants]

Jan 2007
Challenge 2 datasets released [CQG 24, S551 (2007)]

Jun/Jul 2007
Challenge 2 results due; presented at Amaldi (arxiv.org/0711.2667). Challenge 1B released

Dec/Jan 2007
Challenge 1B results due, presented at GWDAW-12. Challenge 3 released

	MLDC 1	MLDC 2	MLCD 1B	MLDC 3
Galactic binaries	• Verification ✓ • Unknown, isolated ✓ • Unknown, interfering ✓	• Galaxy of $3{\times}10^6$ ✓ ✓ ✓	• Verification ✓ • Unknown, isolated ✓ • Unknown, confused ✓	• Galaxy of $6{\times}10^7$ chirping
MBH binaries	• Isolated ✓	• 4–6×, over Galaxy with EMRIs ✓	• Isolated ✓ pre-subtracted	• Over Galaxy spinning, precessing
EMRIs		• Isolated ✓ • 4–6×, over Galaxy with SMBHs	• Isolated ✓	• 4–6× together, weaker
more…			raw observables, randomized noises	• Cosmic string cusp bursts • Cosmological background
	10 collaborations	**13** collaborations	**10** collaborations	we'll see…

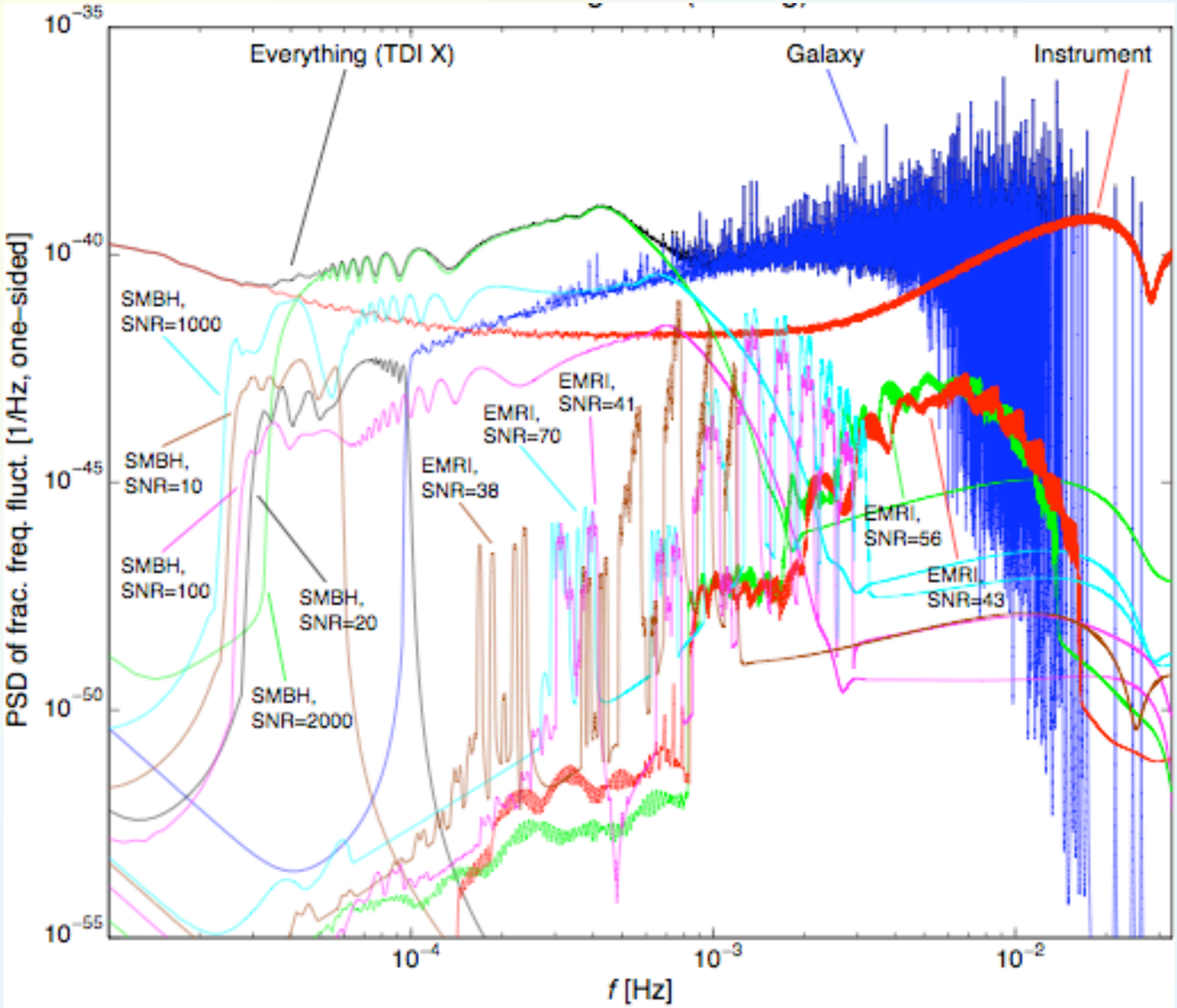
Contestants…

- NASA Ames
- U. of Auckland
- Chinese Academy of Sci., Beijing
- U. of Birmingham
- U. of Texas Brownsville
- Caltech/NASA JPL
- U. of Cambridge
- Cardiff U.
- Carleton College
- U. of Glasgow
- NASA Goddard
- Albert Einstein Institut Golm
- Albert Einstein Institut Hannover
- U. Illes Balears
- Indian Inst. of Tech., Kharagpur
- IMPAN Warszaw
- Montana State U.
- Nanjing U.
- CNRS Nice
- Northwestern U.
- CNRS APC Paris
- U. of Southampton
- U. of Wroclaw

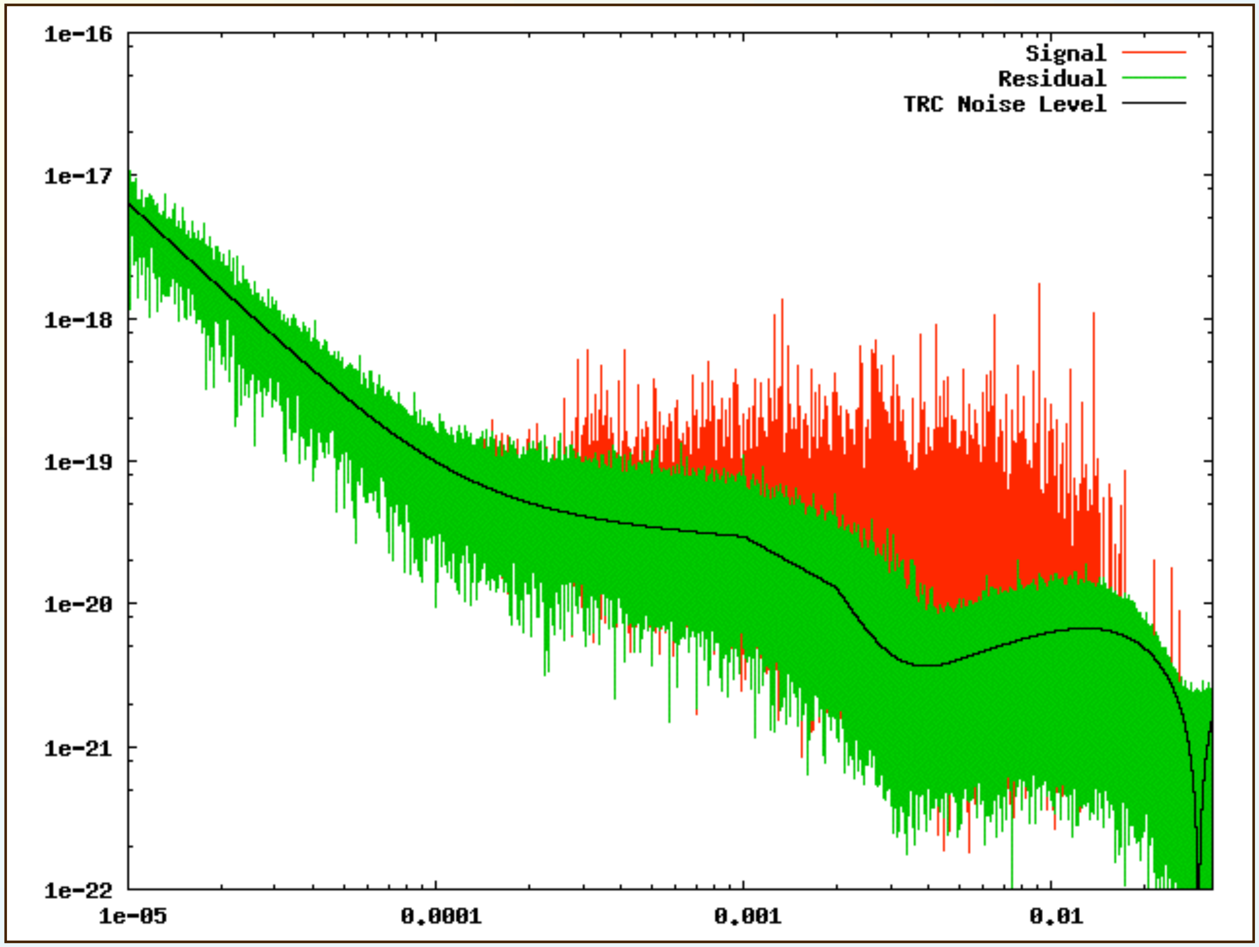
…and techniques

- Template-bank matched filtering
- Markov-Chain Monte Carlo matched filtering
- Genetic optimization
- Time-frequency track scans
- Tomographic reconstruction
- Hilbert transform
- F-statistic, hierarchical schemes
- …

Challenge 2 highlights: the “Whole Enchilada”

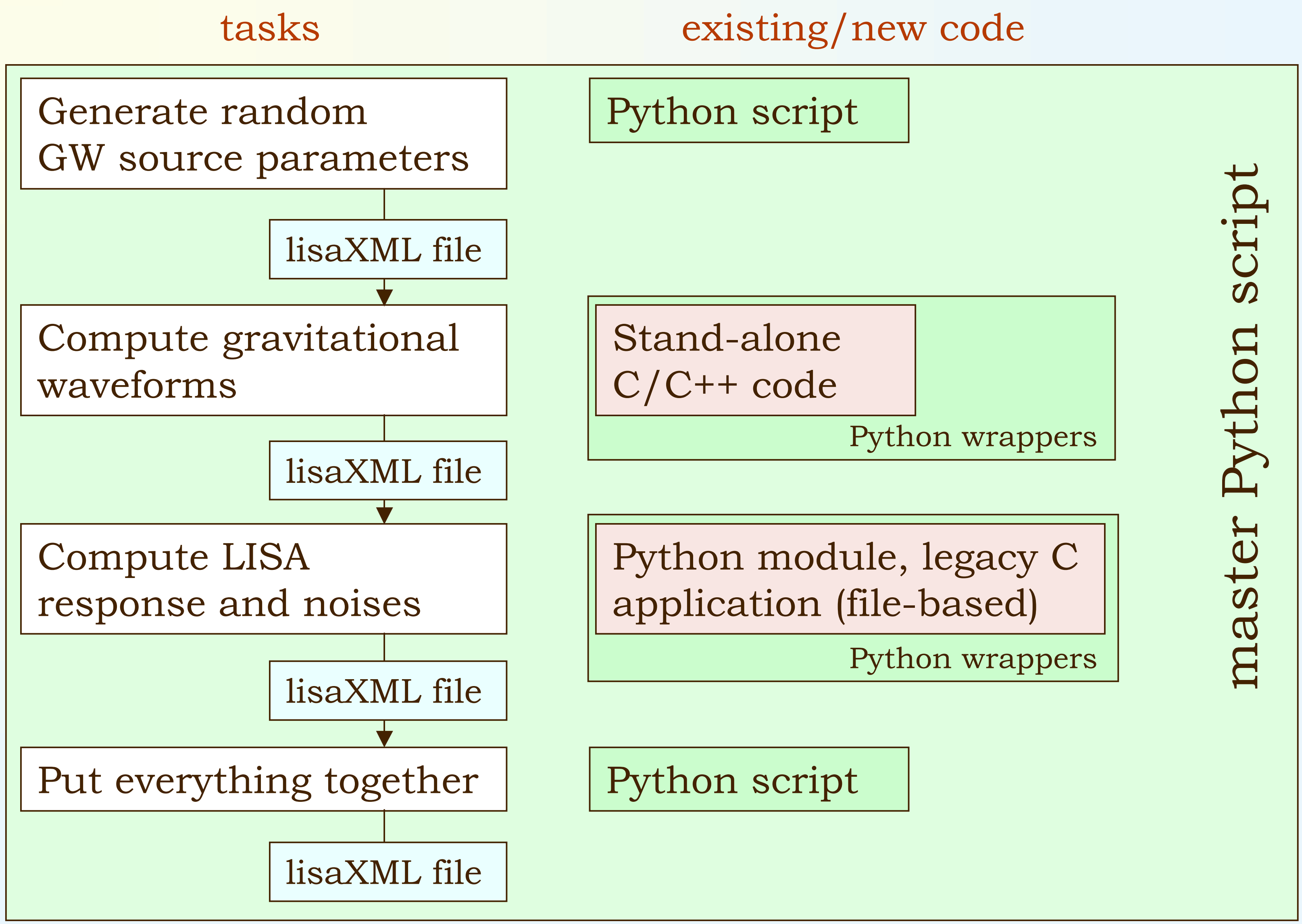


Challenge 2 highlights: Galaxy subtraction



- Using the MT/JPL catalog (thanks to Jeff Crowder) of 19324 sources, minus 1712 rejected by Bayesian Information Criterion

The MLDC workflow



The MLDC workflow

```
makesource-GalacticBinary.py --seed=1234
                 --requestSN=10 mybinary-parameters.xml

makebarycentric.py mybinary-parameters.xml
                               mybinary-barycentric.xml

makeTDIsignal-synthlisa.py mybinary-barycentric.xml
                                       mybinary-TDI.xml
```

```
<?xml version="1

<XSIL>
    <Param Name=

    <XSIL Type="
        [...]
    </XSIL>
        <XS
    <XSI
```

simulated
data strea

bi
de

```
</XS
        </XS
    </XSIL>
```

```
</Param>

>0</Param>
</Param>
0</Param>

>

nary,BigEndian">
```

Mock LISA Data Challenge XML F

Mock LISA Data Challenge XML File Format, v. 1.0

File Info

Authors	MLDC Task Force	
GenerationDate	2007-08-10T18:12:06CEST	ISO-8601

```
Full dataset for challenge1B.1.1c (synthlisa version), source seed = 733424,
noise seed = 733424, LISAtools SVN revision 491
lisaXML 1.0 [M. Vallisneri, June 2006]
```

LISA data

Standard MLDC PseudoLISA (PseudoLISA)

TimeOffset	0	Second
InitialPosition	0	Radian
InitialRotation	0	Radian
Armlength	16.6782	Second

Source data

GB-1.1.1c (PlaneWave)

SourceType	GalacticBinary	
EclipticLatitude	-0.575706071762	Radian
EclipticLongitude	3.68595734709	Radian
Polarization	3.2062766975	Radian
Frequency	0.00974356389768	Hertz
InitialPhase	0.523693531091	Radian
Inclination	1.69786387662	Radian
Amplitude	1.98421310881e-23	1

TDI data

t,Xf,Yf,Zf (TDIObservable)

DataType	FractionalFrequency	

TimeSeries: t,Xf,Yf,Zf

TimeOffset	0.0	Second
Cadence	15.0	Second
Duration	31457280.0	Second
Array Stream: t,Xf,Yf,Zf	Filename	challenge1B.1.1c-training-frequency-0.bin
	Encoding	Binary,LittleEndian
	Type	double
	Unit	

lisaXML's natural Python interface

```
<?xml version="1.0"?>

<XSIL>
 <Param Name="Author">
  Michele Vallisneri
 </Param>

<XSIL Type="SourceData">
  <XSIL Name="Galactic binary 1.1"
                              Type="PlaneWave">
   <Param Name="SourceType">
    GalacticBinary
   </Param>
   <Param Name="EclipticLatitude"
                                 Unit="Radian">
    0.9806443268
   </Param>

   [...more Params...]
  </XSIL>

  [...more PlaneWave sources...]
 </XSIL>
```

```
>>> fileobj = lisaXML('test.xml','r')
>>> fileobj
<lisaXML file 'test.xml'>

>>> fileobj.Author
'Michele Vallisneri'

>>> fileobj.SourceData
<XSIL SourceData (2 ch.)>

>>> gb = fileobj.SourceData[0]
>>> gb
<XSIL PlaneWave 'Galactic binary 1.1'>

>>> gb.Name
'Galactic binary 1.1.1a'
>>> gb.EclipticLatitude
0.9806443268
>>> gb.EclipticLatitude_Unit
'Radian'
>>> gb.parameters
['EclipticLatitude',
'EclipticLongitude', 'Polarization',
'Frequency', 'InitialPhase',
'Inclination', 'Amplitude']
```

load lisaXML file

access metadata

select XSIL container

access attributes and Params

In conclusion

- It's been a lot of work, but we're showing that LISA data analysis is possible, we're developing new techniques, publishing many papers
- Cross-pollination with ground-based efforts is crucial
- The MLDC infrastructure (LISAtools) can be used to generate data for many other experiments outside the mainline challenges
- The LISA standard model (pseudo-LISA, source models) can be used to compare data-analysis results (see beginning investigations of LISA science performance)
- In the future: more realistic noise, sources (e.g., inspiral + merger + ringdown); use MLDCs as testbed for prototypes of LISA core analysis tools

Discussion

- What can mock challenges do for numerical-relativity vs. data analysis?
  - (The MLDCs have been assuming perfectly known waveform models)
  - Test a parameter estimation protocol?
- What can you do in 15 years? [with funding, of course]
  - Produce a perfect interpolated inspiral + merger + ringdown template family covering parameter space
  - (related:) Can numerical-relativity waveforms ever be produced as a commodity?
  - To match statistical errors, reduce systematic errors to 1 in 10^{-3} (for SNR = 10^3, FF must be $1 - 10^{-6}$, although errors in some directions are benign)
  - Solve modified Einstein equations
- What else?

See for yourself

- MLDC official site:
  astrogravs.nasa.gov/docs/mldc
- MLDC taskforce wiki:
  www.tapir.caltech.edu/dokuwiki/listwg1b:home
- Mailing lists:
  lisatools-mldc@gravity.psu.edu (formulation)
  lisatools-challenge@gravity.psu.edu (participants)
- LISAtools software (including full MLDC pipeline):
  lisatools.googlecode.com