Computing Model for ET

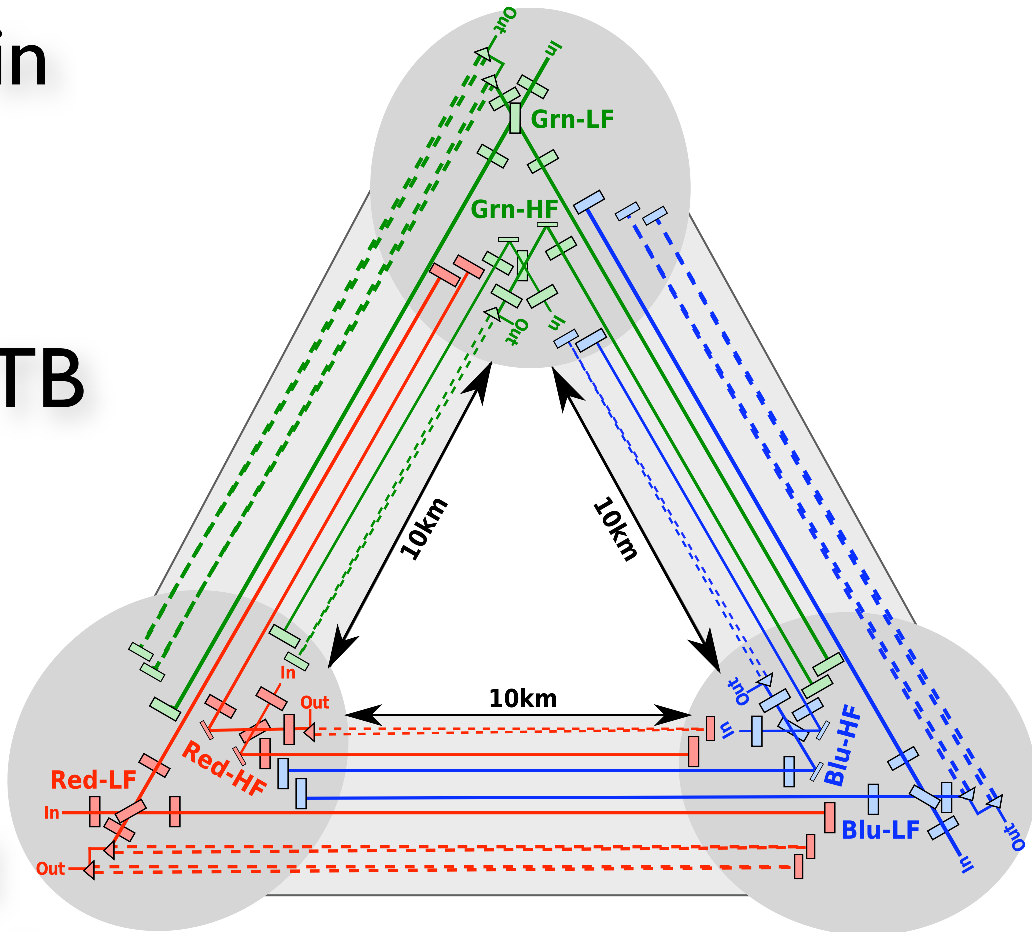
ASPERA meeting on Computing in Astroparticle Physics, Barcelona May 30-31 2011

B.S. Sathyaprakash
School of Physics and Astronomy, Cardiff University, UK
on behalf of the Einstein Telescope Design Study Team

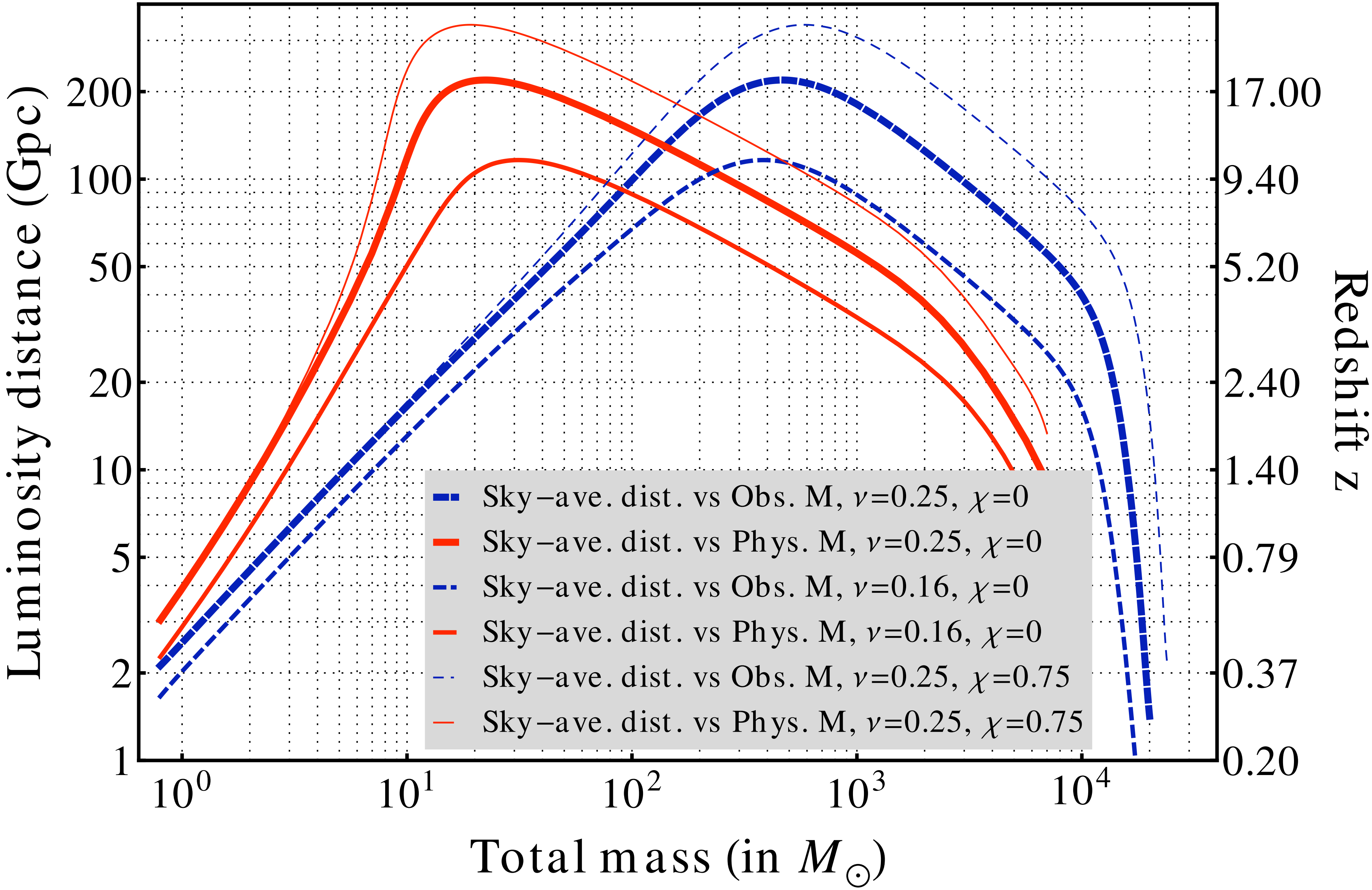

Einstein gravitational wave Telescope
Conceptual Design Study
ET
EINSTEIN
TELESCOPE

Einstein Telescope: An Observatory

- Six interferometers in a triangular array
- Timeline: 2025+
- Data rates of ~ 100 TB per year from the observatory
  - data handling is not a great challenge
- Data processing and science exploitation is the challenge



ET Distance Reach for Compact Binary Mergers



ET Challenge for Data Analysts

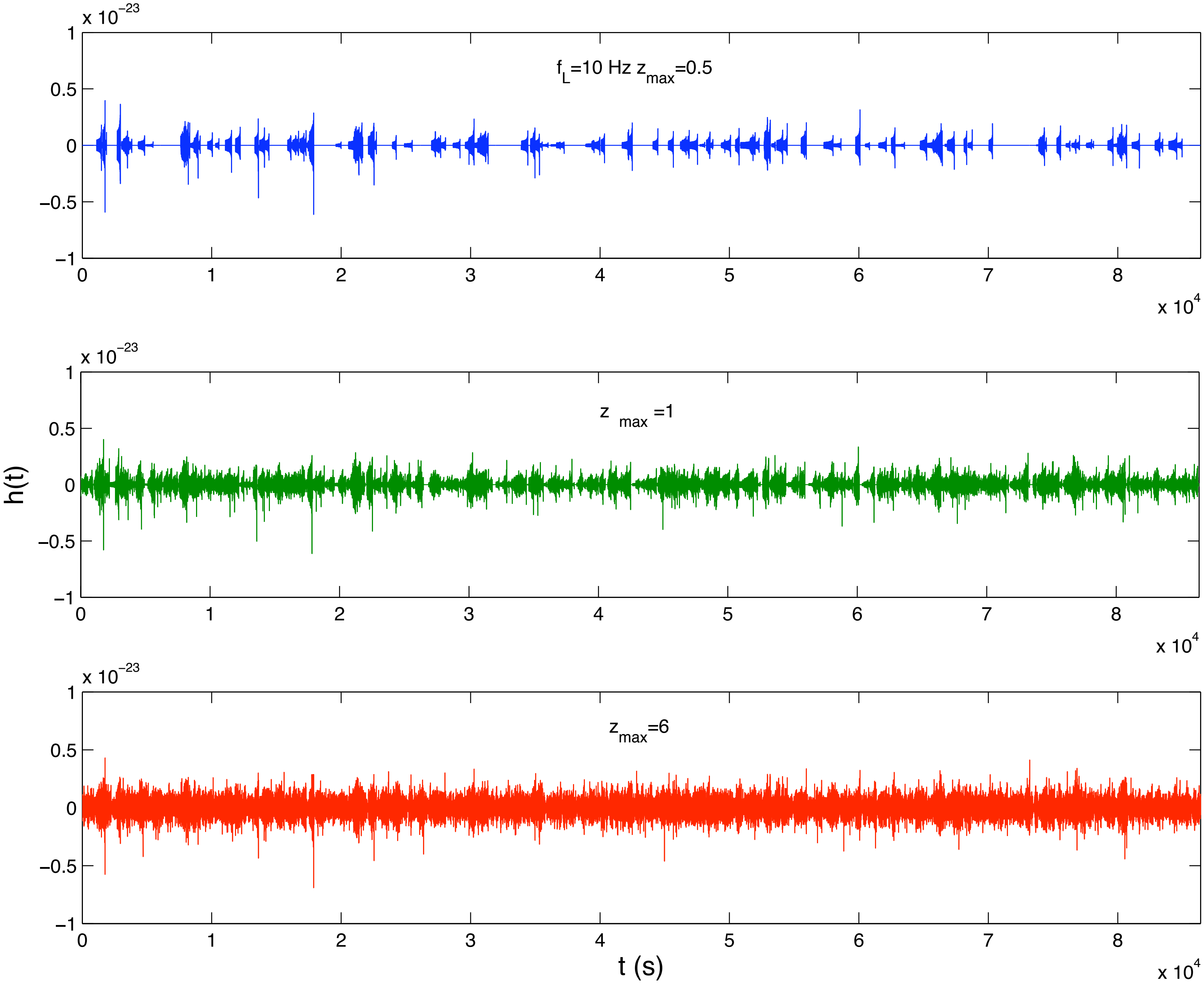
- Signal durations ranging over seven orders of magnitude
  - 10's of milliseconds for intermediate-mass black hole binaries, supernovae, neutron star glitches, etc.
  - Up to a week in the case of binary neutron stars
  - Continuous wave sources that last for weeks to years
- There could be many overlapping sources
  - Possible stochastic background but likely below detector noise
- Detecting rare transients from SN, pulsar glitches, etc.
  - Hidden amidst millions of sources observed each year might be a transient that lasts for 10's of milliseconds to a second

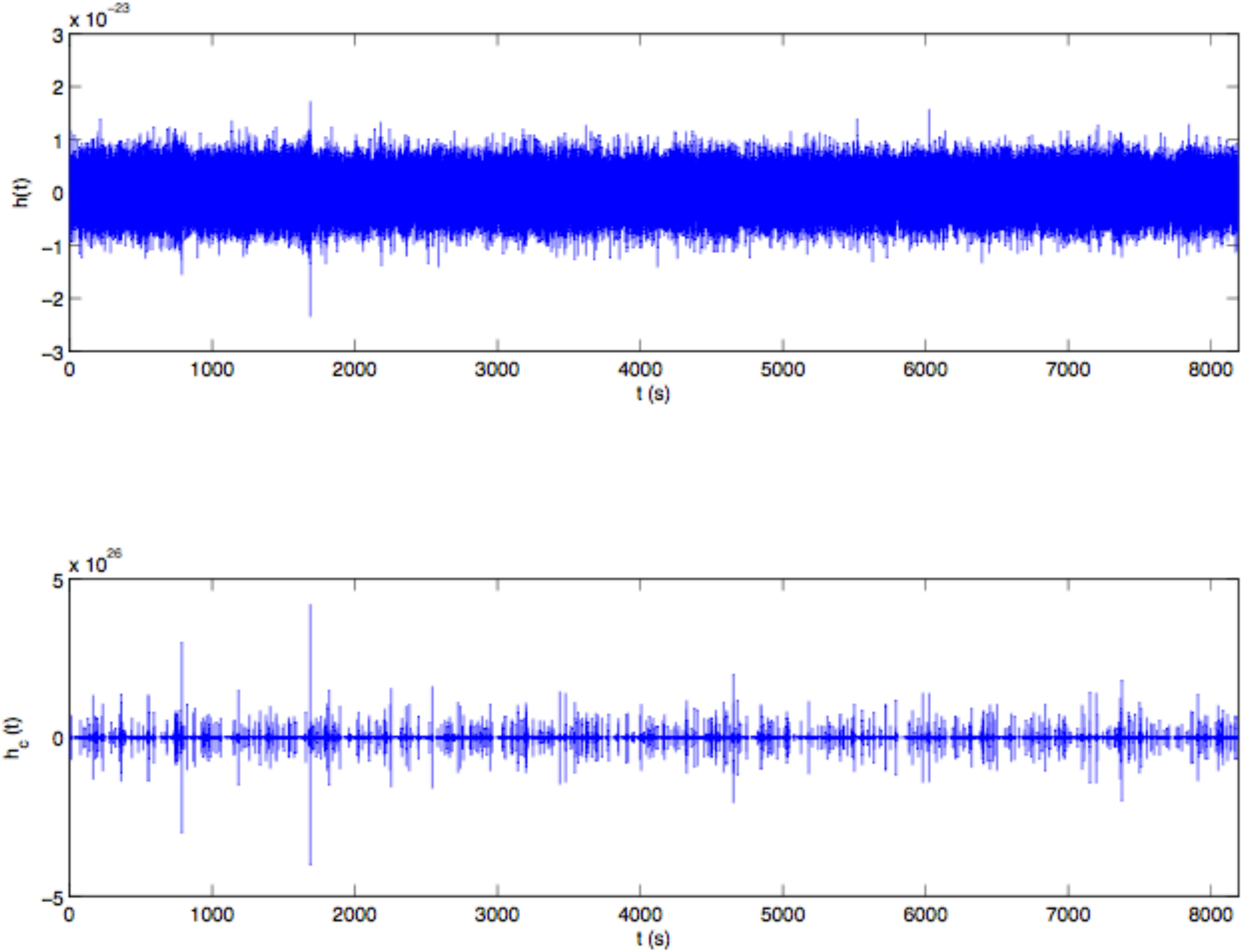
Duration of Inspiral Signals in ET

f_L (Hz)	NS-NS (1.4+1.4)	NS-BH (1.4+10)
40 (initial)	25 s	5.8 s
10 (Ad LIGO)	16.7 m	3.9 m
5 (Ad Virgo)	1.8 h	24.6 m
3 (possible ET)	6.9 h	1.6 h
1 (planned ET)	5.4 d	1.2 d

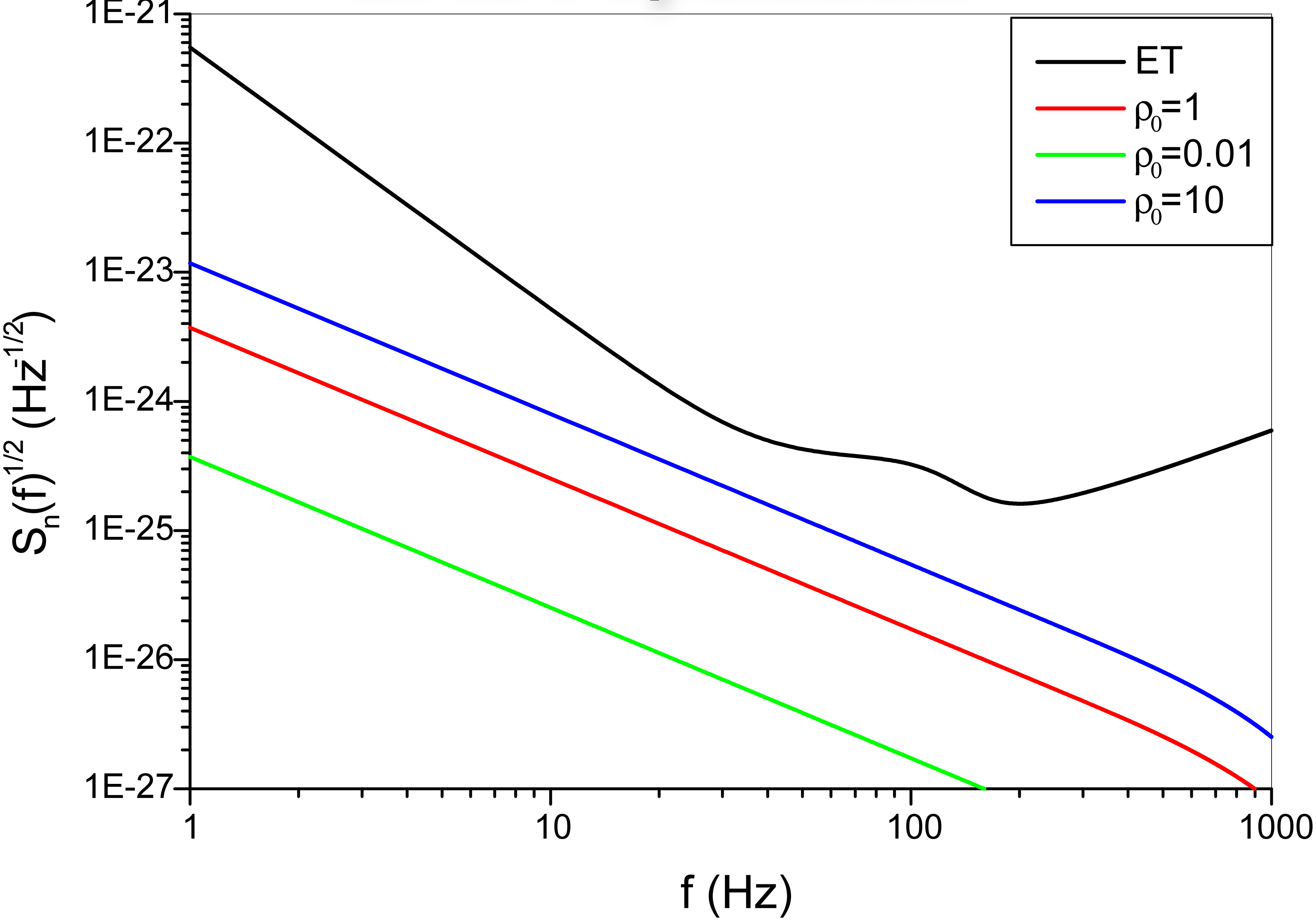
Searching for binary inspirals in ET

- Inspirals are traditionally searched using matched filtering
- Whitened data is filtered through a bank of templates that are copies of the expected signal
  - The search space and algorithm is already limited due computational requirements
  - The full parameter space is 17-dimensional
  - A coincidence search is carried out instead of a coherent search although the latter has a great potential to strengthen the searches


x 10^−23
f_L=10 Hz z_max=0.5
z max =1
z_max=6
h(t)
t (s)
x 10^4
1
0.5
0
−0.5
−1
0 1 2 3 4 5 6 7 8


x 10^-23
h(t)
3
2
1
0
−1
−2
−3
0
1000
2000
3000
4000
5000
6000
7000
8000
t (s)
x 10^26
h_c (t)
5
0
−5
0
1000
2000
3000
4000
5000
6000
7000
8000
t (s)

Stochastic Background from BNS Population



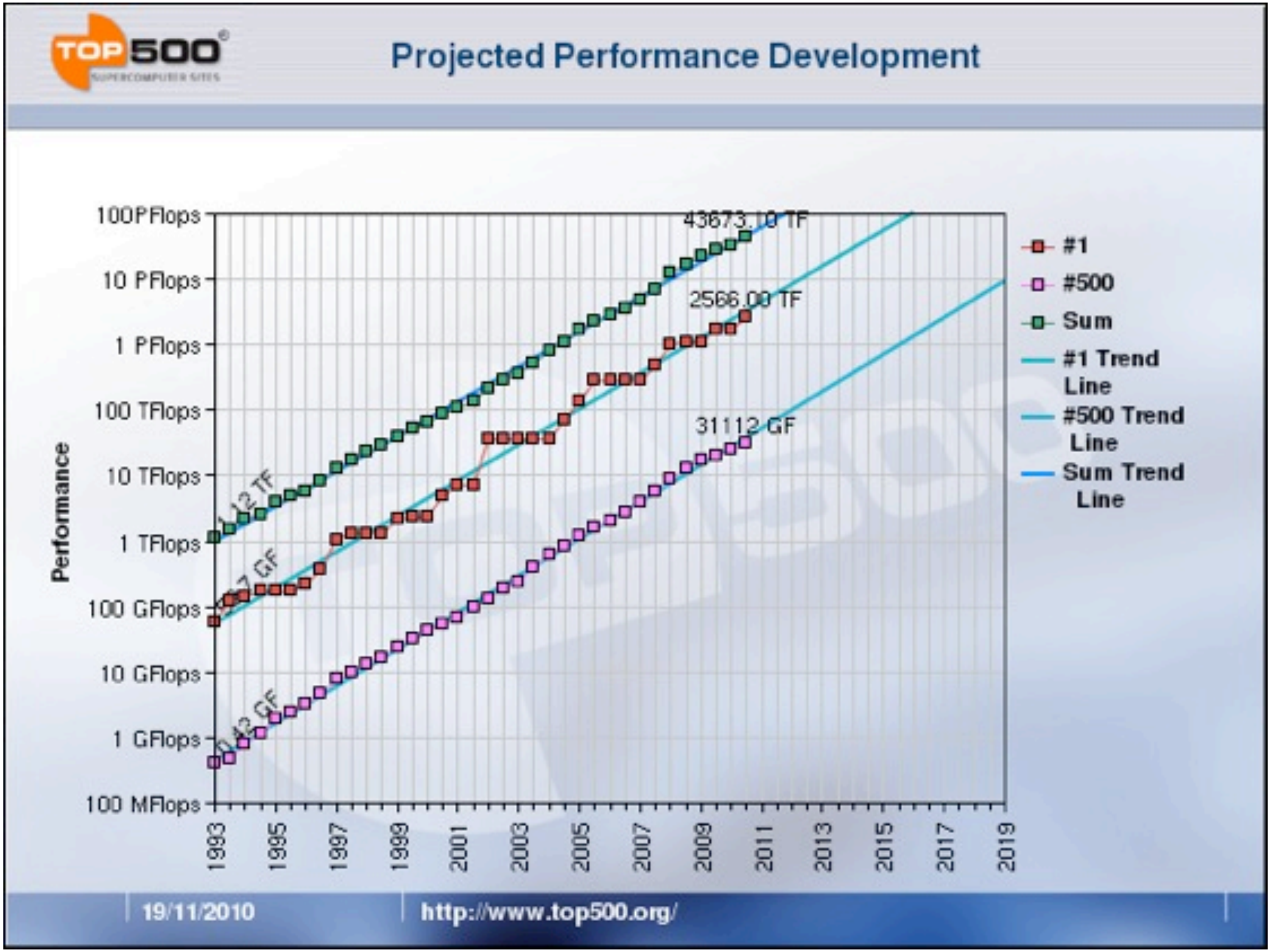
Cost of Matched Filtering

To search for one day's worth of data

		GPU C2050	WLCG	Tianhe-1A	Low-cost
Power (TFlops)		0.25	2000	2567	100
# of Params	# of Templates	Computational Time			
2	10^{9}	128 d	24 m	18 m	7 h
4	10^{13}	> 1 y	182 d	128 d	9 y
6	10^{18}	> 1 y	> 1 y	> 1 y	> 1 y

Projected Performance of Top 500

Computing power available in the ET era (even assuming Moore's law) is not sufficient, need other ways of analysis



New Computing Paradigms

- Conventional computing paradigms for detecting binary inspirals simply won't work
  - Signal lengths are far too long - need to break up the data set as in CW case
  - Number of overlapping signals means focus should be on those that are brightest
- At several events every minute the signal rate is far too high for post-processing pipelines
  - Current follow-up analyses take ~ days per event
  - E.g., posterior distribution of parameters is essential for delivering the science but takes too much time

Multi-messenger Astronomy

- A subset of events (50%) might be followed up in radio, X-ray, gamma-ray, etc
  - Processing power is required for not just GW data
  - Need to process EM observations
- Current set up and that envisaged for advanced detectors still work with semi-automated set ups for EM follow-ups
  - Must have fully automated data pipelines
  - Automated alerts, observation and analysis

A new model for data processing

- Current data analysis of most GW data is centrally controlled
  - A software library that is contributed and maintained by the collaboration as a whole
- This is not a good model in the ET era
  - It might be impossible to centrally process all the data and keep up also with progress in algorithms, theory, models, etc.
- Computation and analysis should perhaps be decentralized
  - End users apply for "observatory time" and they get a week's worth of data or a month's worth
  - They could have the ownership for a set period (say six months), after which the data will be made public
- The decentralized models might help distribute the load

Summary

- Raw data rates in ET are not expected to be too different from advanced detectors
  - Each site might host a number (~10) interferometers but even so the data rate might not be a big issue
- ET will be a signal-dominated detector
  - 1 event per week in advanced detectors translates to millions per year in ET
  - Data products could overwhelm the raw data rate
- A new model for data products might be needed
  - ET could see a variety of different signals and we need a way to store all the information associated with them
  - Need to develop a new formats/structures/databases, to store data products