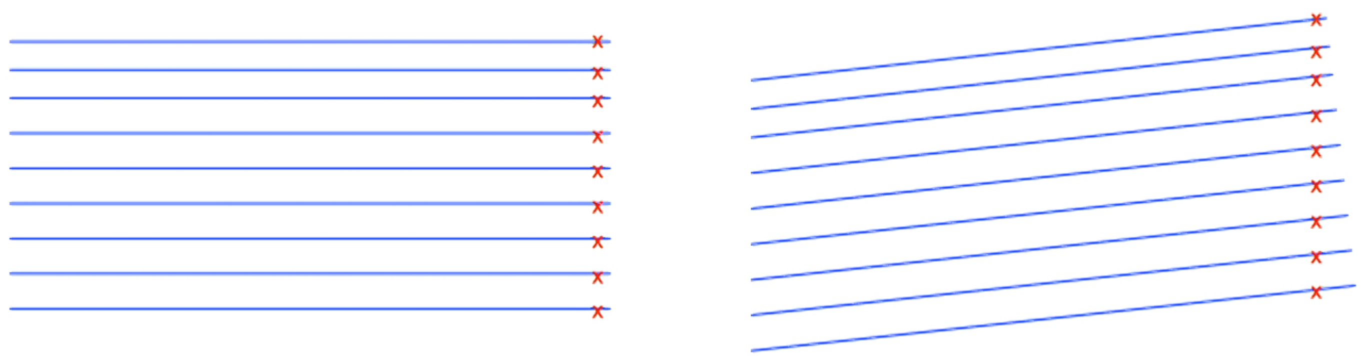
Title: CHIME: The Canadian Hydrogen Intensity Mapping Experiment

Date: Jan 17, 2017 11:00 AM

URL: http://pirsa.org/17010076

Abstract: <p>CHIME is a new interferometric telescope at radio frequencies 400-800 MHz. The mapping speed (or total statistical power) of CHIME is among the largest of any radio telescope in the world, and the technology powering CHIME could be used to build telescopes which are orders of magnitude more powerful. This breakthrough sensitivity has the power to revolutionize radio astronomy, but meeting the computational challenges will require breakthroughs on the algorithmic side. I'll give a status update on CHIME, with an emphasis on new algorithms being developed to search for fast radio bursts and pulsars.</p>

Phased-array interferometer

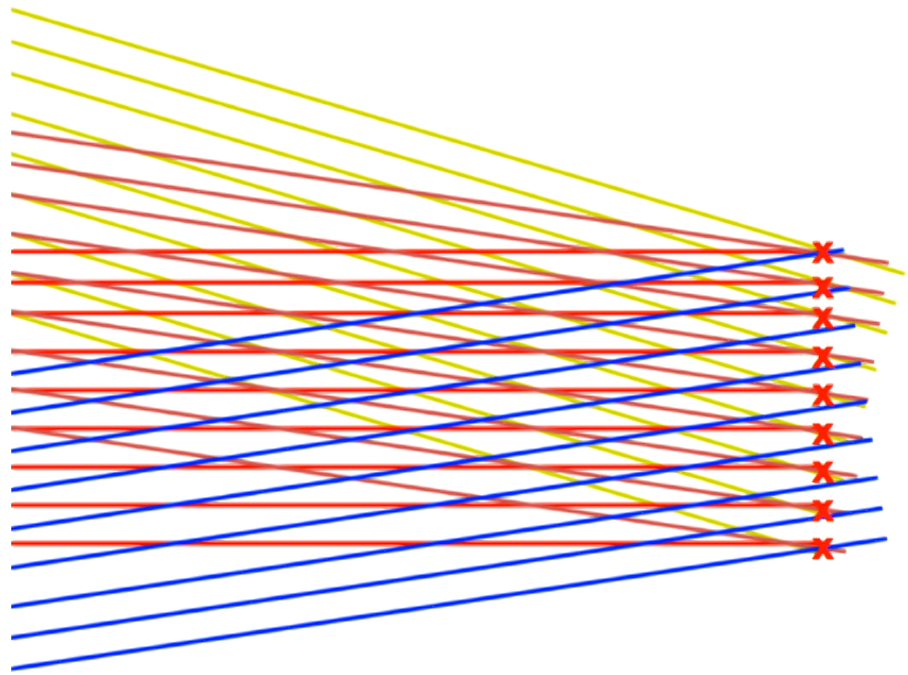


Dish is replaced by an array of antennas whose signals will be combined in software!

By summing signals with appropriate delays, can mimic the dish and focus on part of the sky.

Can "repoint" telescope by changing delays.

Beamforming interferometer



Given enough computational power, can form a complete set of N beams on the sky. Equivalent to N traditional telescopes with the same collecting area as the interferometer!

CHIME

Under construction now in Penticton B.C., first light expected this summer.

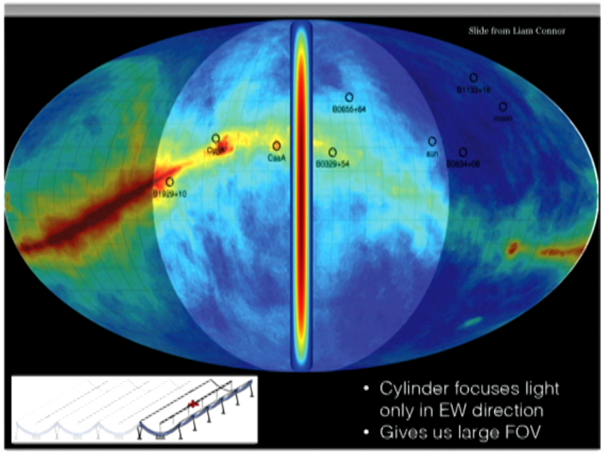
We have a 4 x 256 array of antennas and enough computing power to form all 1024 independent beams in real time.

Raw sensitivity is the same as 1024 world-class radio telescopes observing 24/7 !!

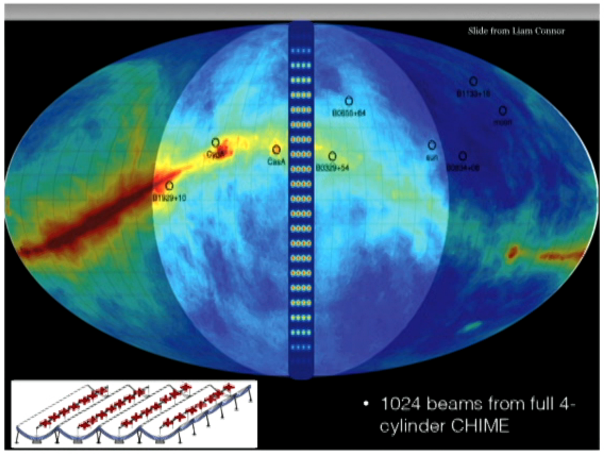
CHIME

At any instantaneous observing time, each antenna sees a narrow strip on the sky ("primary beam").

By beamforming in software as previously described, we can make 1024 "formed" beams with angular resolution ~20 arcmin.



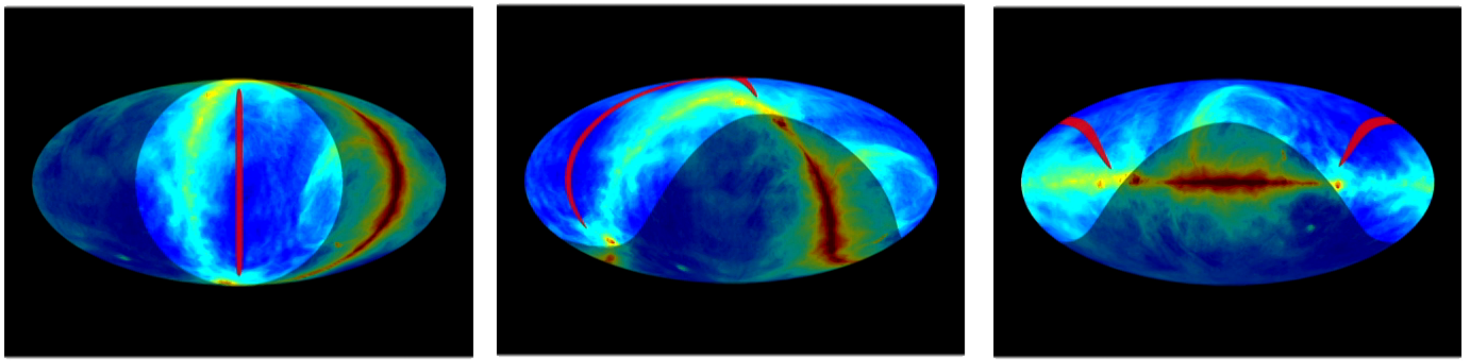
primary beam



formed beams

CHIME

The beam is fixed in telescope coordinates, but as the Earth rotates, the beam sweeps over the full sky.



Every 24 hours, we get a sky map with:

- Angular resolution: 20 arcmin
- Sky coverage: half the sky
- Frequency range: 400-800 MHz.
  (We see neutral hydrogen at z = 0.8-2.5 via the 21-cm line)

Mapping speeds (very approximate)

For many purposes, the statistical power of a radio telescope can be quantified by its mapping speed:

$$M \approx (\text{Collecting area } A) \times (\text{Number of beams}) \times (\text{order-one factors})$$

	A	$N_{\rm beams}$	$M/(10^5\ \mathrm{m}^2)$
Parkes 64m	3200 m^2	13	0.41
Green Bank 100m	7850 m^2	7	0.55
Aricebo 300m	70000 m^2	7	4.9
FAST 500m	200000 m^2	19	38
CHIME	6400 m^2	1024	66

=  CHIME ?!

The catch

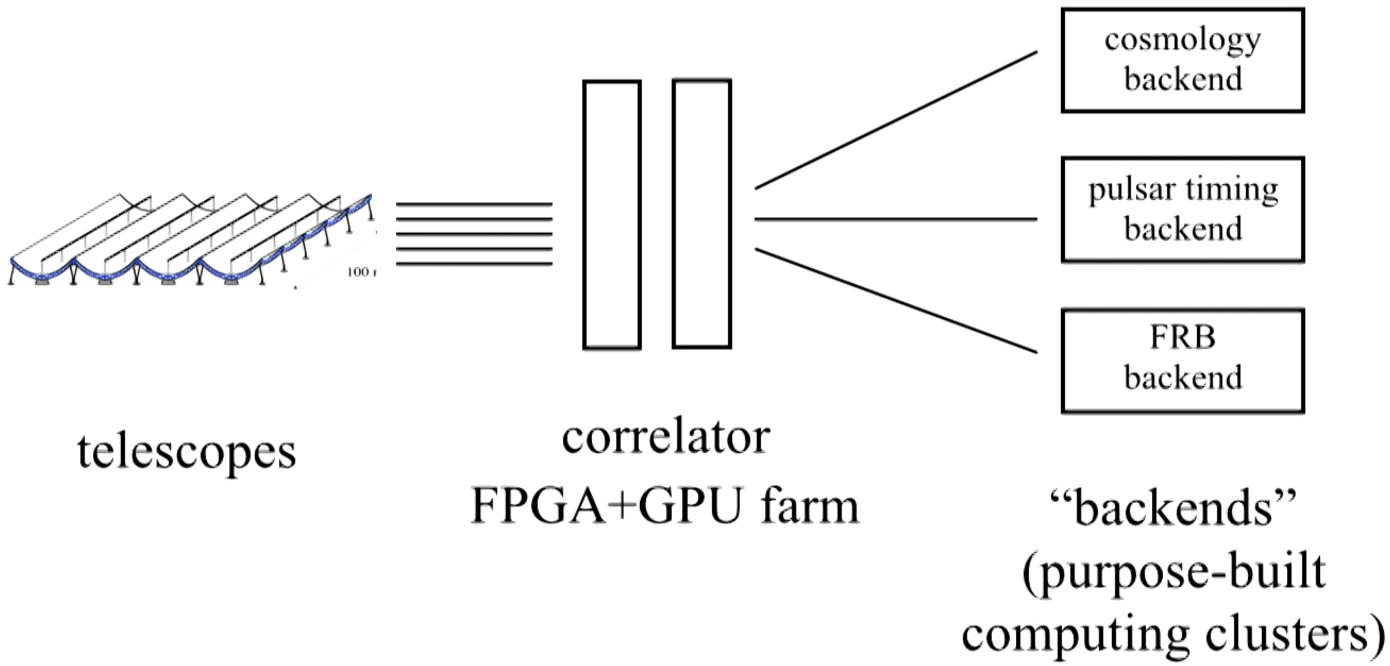
	A	N_{beams}	$M/(10^5\ \mathrm{m}^2)$
FAST 500m	200000 m^2	19	38
CHIME	6400 m^2	1024	66

CHIME gets its high mapping speed from many beams with low instantaneous sensitivity (compared to FAST)

In principle, sensitivity is roughly proportional to M, but computational cost is proportional to N_{beams} (or worse).

What we have really done is replace the problem of building a huge telescope by hard algorithmic/computational problems.

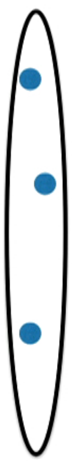
CHIME hardware, cartoon form



CHIME hardware, cartoon form

Each backend asks the correlator for a different data stream over the network. For example:

"Pulsar timing backend"

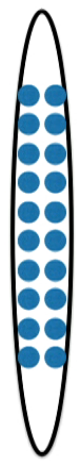


receives digitized electric field

with nanosecond sampling

at 10 specified sky locations

"FRB backend"



receives intensity, obtained from electric field by time-downsampling and polarization-averaging

with millisecond sampling

at a regular array of 1024 sky locations

=  CHIME ?

=  CHIME + a lot of math, for some problems?

1. Introduction
2. Fast radio bursts
3. Searching for new pulsars
4. Things I didn’t have time to talk about
5. Thoughts for the future

Can CHIME's high mapping speed be used to find FRB's?

In principle, FRB event rate is proportional to mapping speed, so CHIME should be incredibly powerful. The forecasted event rate is ~10 events per day (!!)

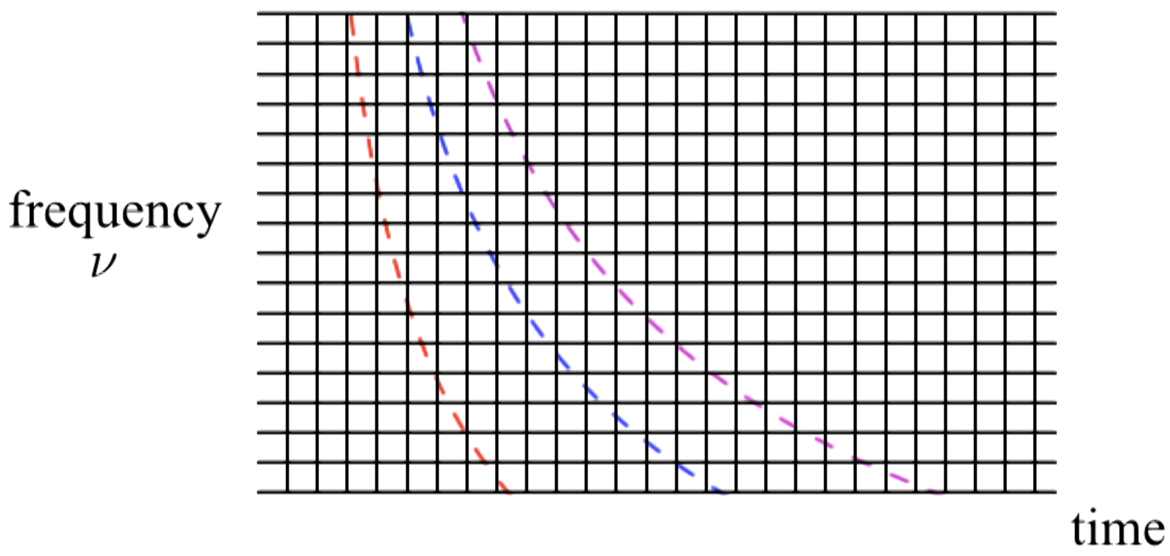
The search problem ("incoherent dedispersion") is very computationally challenging.

Very well-studied: there are many papers and some public software, but not nearly fast enough for CHIME!

Over the last year, we've developed a search algorithm which is much faster, and can search the CHIME data to its full statistical power.

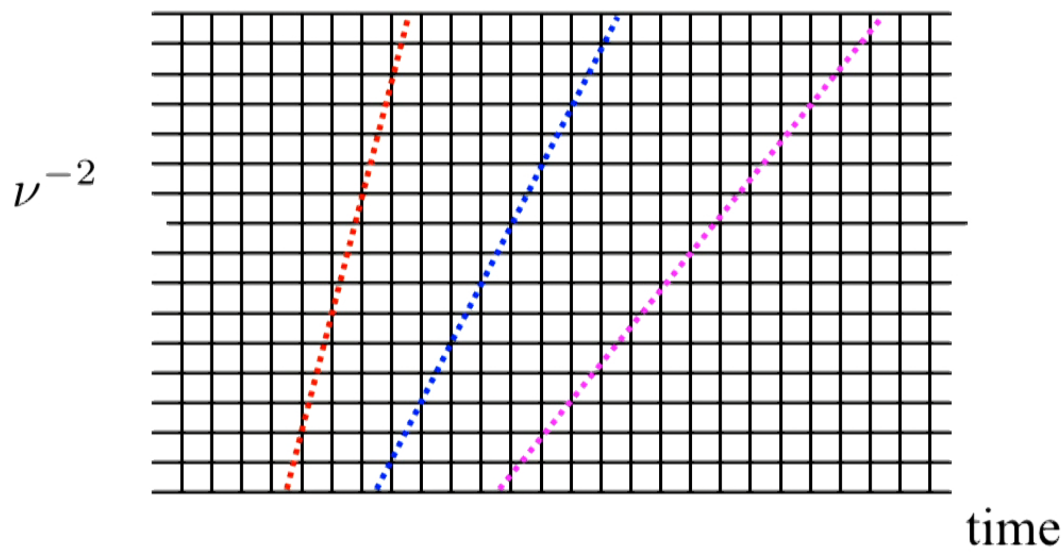
Incoherent dedispersion

Setting up the problem. The FRB backend incrementally receives a 2D array with (time, frequency) axes. We want to sum over all "tracks" with the shape shown.



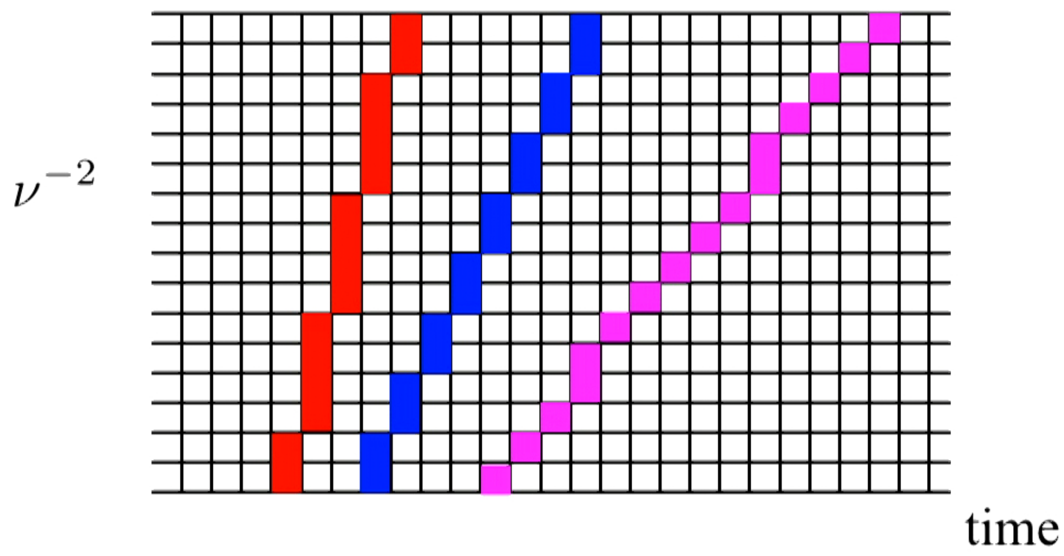
Incoherent dedispersion

The first step is to change variables $\nu \to \nu^{-2}$, in order to transform the dispersion tracks to straight lines.



Tree dedispersion

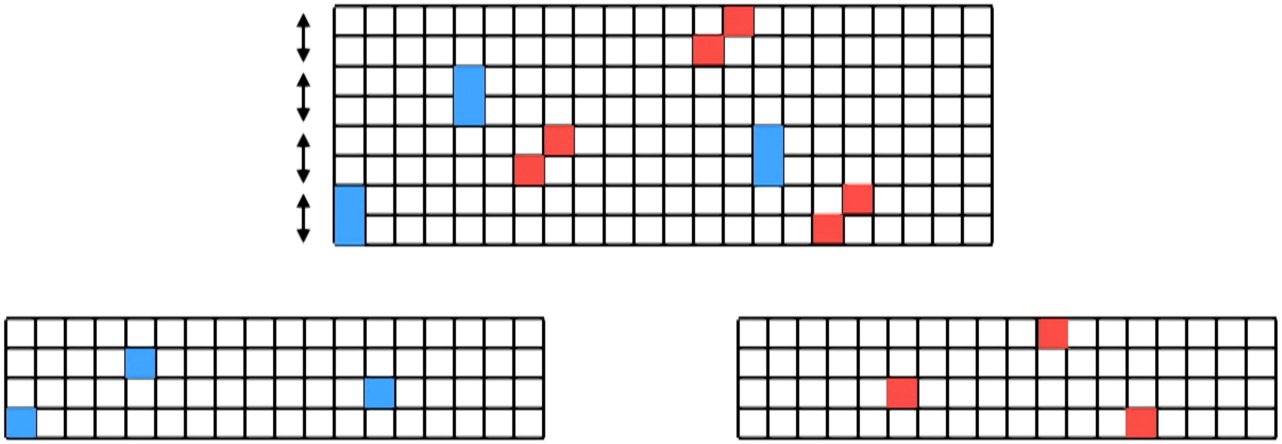
The algorithm I'll describe is a "tree" algorithm which ends up approximating each straight-line track by a jagged sum of samples.



Tree dedispersion

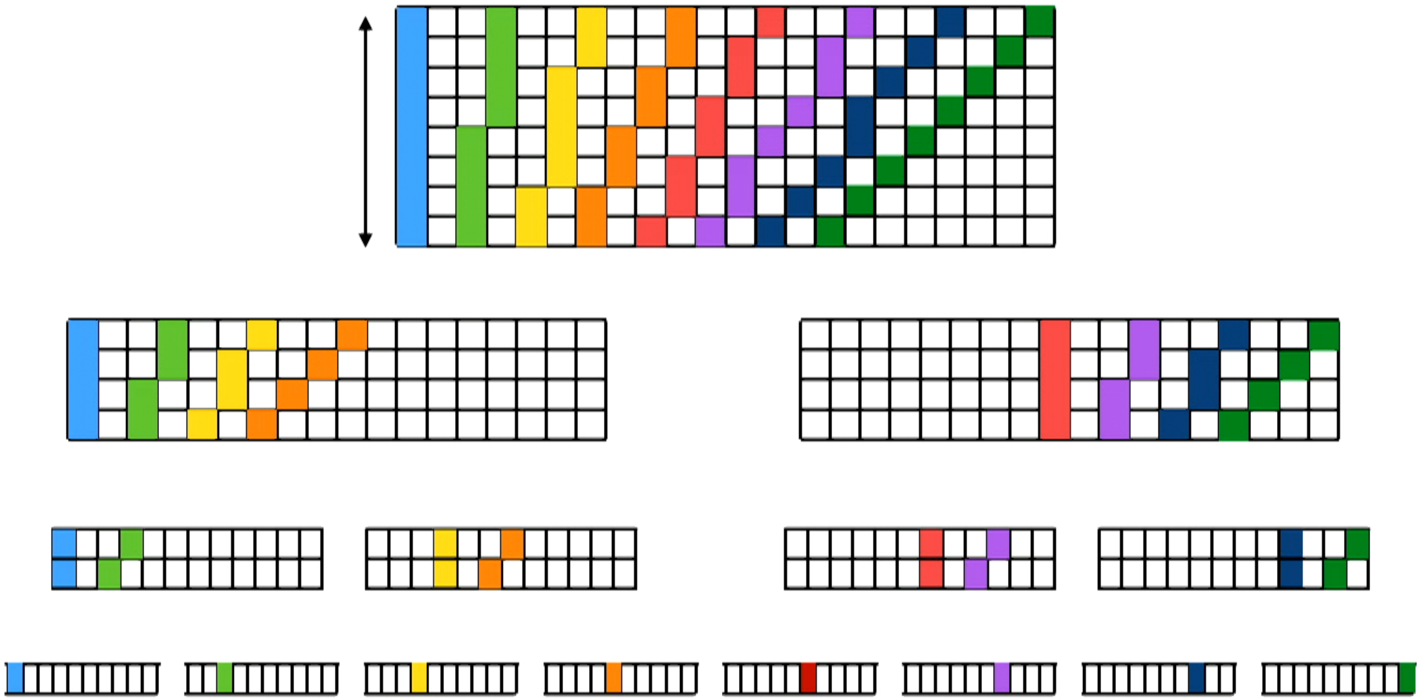
First iteration: group channels in pairs. Within each pair, we form all “vertical” sums (blue) and “diagonal” sums (red).

Output is two arrays, each half the size of the input array.



Tree dedispersion

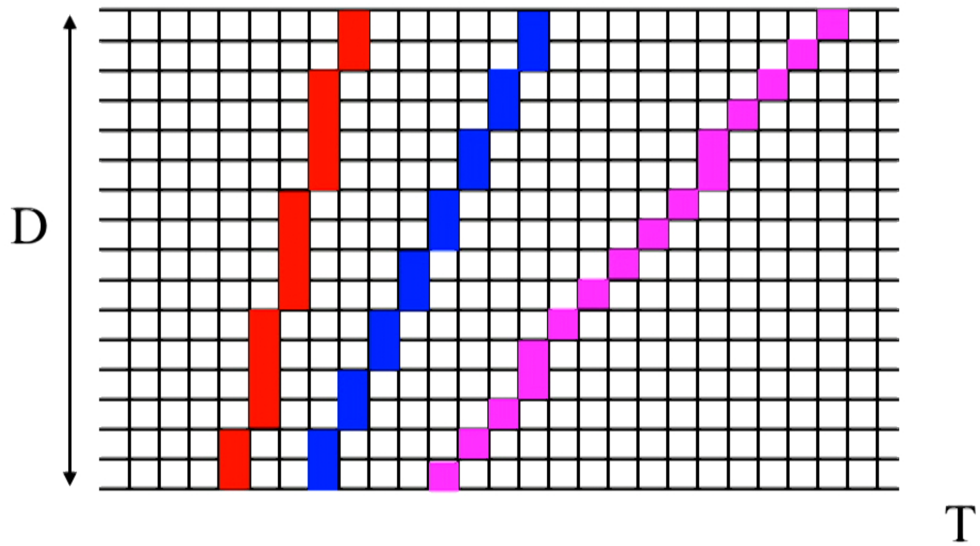
Last iteration: all channels summed.



Tree dedispersion

Total cost of tree summation is O(TD log D), where T=number of times, D=number of DM's.

Cost of direct summation would be $O(TD^2)$, so this is a big speedup!



Tree dedispersion

Total cost of tree summation is O(TD log D), where T=number of times, D=number of DM's.

Cost of direct summation would be $O(TD^2)$, so this is a big speedup!

An important result: the tree search can be made arbitrarily close to optimal if certain hidden choices are made appropriately.

Tree dedispersion

Total cost of tree summation is O(TD log D), where T=number of times, D=number of DM's.

Cost of direct summation would be $O(TD^2)$, so this is a big speedup!

An important result: the tree search can be made arbitrarily close to optimal if certain hidden choices are made appropriately.

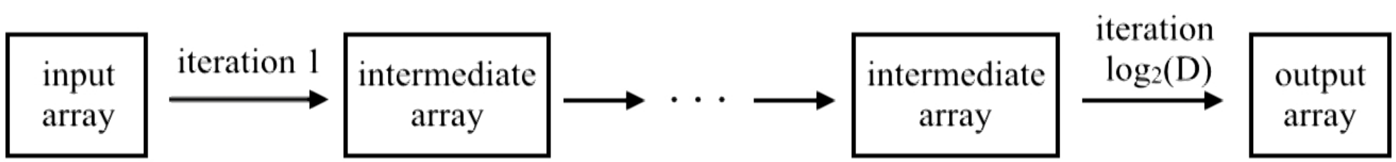
Tree dedispersion is very fast, but still not quite fast enough for CHIME!

Further speedups are possible. To describe them, let's take a little computer science detour…

"Blocking" the tree

We found a "blocked" formulation of the tree algorithm.

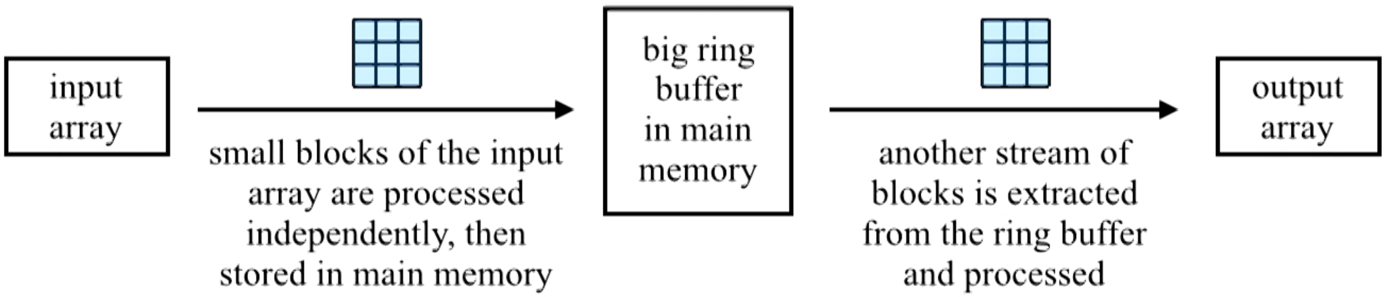
In a straightforward implementation, the entire array is exchanged between main memory and L3 cache in each of $\log_2(D)$ passes, leading to a memory bandwidth bottleneck.



“Blocking” the tree

We found a “blocked” formulation of the tree algorithm.

Non-obvious fact: tree dedispersion can be organized as two passes, each of which operates on a small (< 1 MB) block of memory.



"Blocking" the tree

We found a "blocked" formulation of the tree algorithm.

Non-obvious fact: tree dedispersion can be organized as two passes, each of which operates on a small (< 1 MB) block of memory.

This turns out to be faster, and leads to better parallelism (because all cores share bandwidth to L3 cache).

Another interesting property of the blocked algorithm: low latency real-time alerts. Dispersed pulses can be detected within a few seconds of entering the telescope.

Lots of low level optimizations

For example:

- Handwritten assembly language kernels for speed
- Special techniques for memory bandwidth bottlenecks (non-temporal/streaming writes)
- Hardware conversion between 32-bit and 16-bit floating-point when exchanging data between main memory and L3 cache
- and many more!

The bottom line: Our current implementation is a factor 7 faster than a C implementation of tree dedispersion which we originally believed to be highly optimized.

Public code to be released very soon!

This has brought a near-optimal CHIME search within reach. We are building an FRB backend now. This summer, we expect to start finding FRB's at the full event rate: O(10) per day!

There are currently many proposed explanations for FRB's, but with only 20 events, there is not enough data to test hypotheses. We hope CHIME will solve the mystery of what FRB's are!

=  CHIME for FRB's? Yes!

1. Introduction
2. Fast radio bursts
3. Searching for new pulsars
4. Things I didn't have time to talk about
5. Thoughts for the future

Over the last 50 years, pulsar science has been an exceptionally fertile area of astronomy.

- first observational evidence of gravitational waves
- first detection of extrasolar planets
- exceptionally precise tests of GR
- new astrophysics (e.g. magnetars)
- many astronomically interesting “oddball” systems

Over the last 50 years, pulsar science has been an exceptionally fertile area of astronomy.

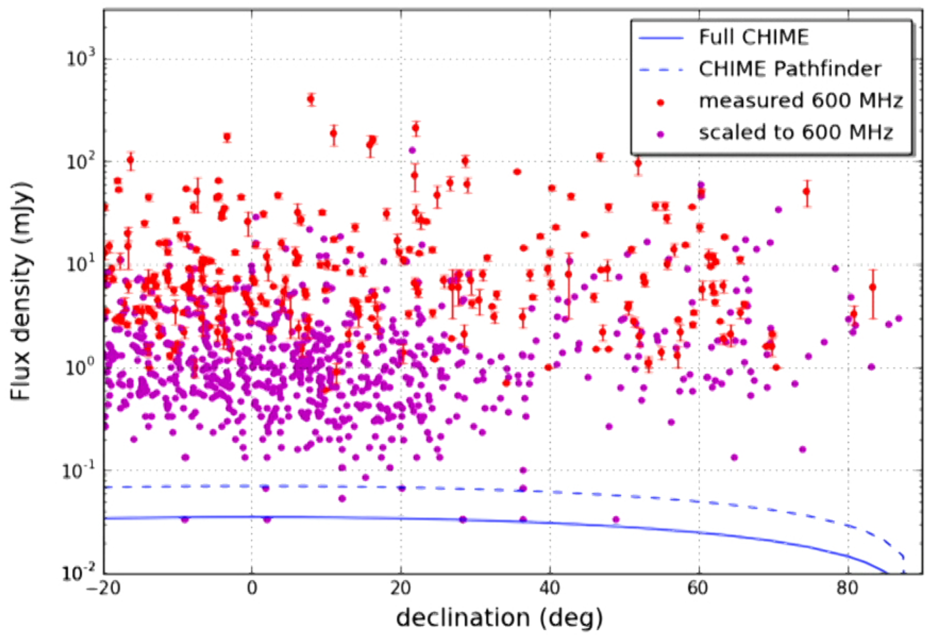
- first observational evidence of gravitational waves
- first detection of extrasolar planets
- exceptionally precise tests of GR
- new astrophysics (e.g. magnetars)
- many astronomically interesting "oddball" systems

Approx 2000 pulsars have been found, but the total population is predicted to be $\sim 10^5$. Many new discoveries to be made! E.g.

- detection of nanohertz gravity waves from pulsar networks
- pulsar + black hole binary system(s)

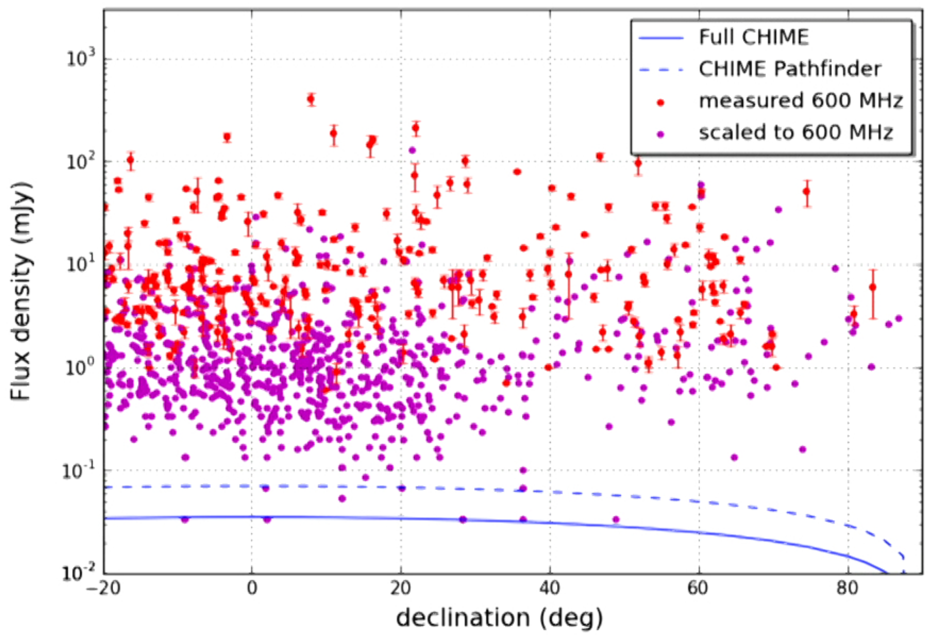
In principle, CHIME has the sensitivity to find many new pulsars!

However, CHIME builds up this sensitivity from a huge number of discontiguous timestreams with low instantaneous sensitivity. Searching this dataset near-optimally for pulsars is an unsolved problem.



In principle, CHIME has the sensitivity to find many new pulsars!

However, CHIME builds up this sensitivity from a huge number of discontiguous timestreams with low instantaneous sensitivity. Searching this dataset near-optimally for pulsars is an unsolved problem.



In principle, CHIME has the sensitivity to find many new pulsars!

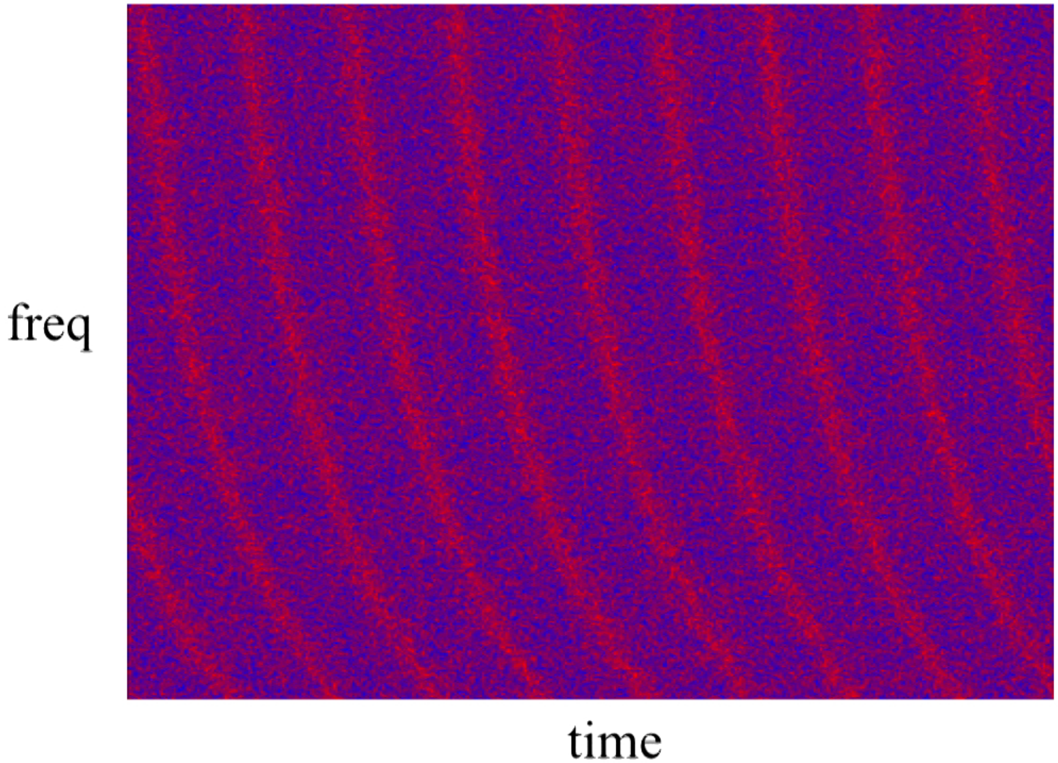
However, CHIME builds up this sensitivity from a huge number of discontiguous timestreams with low instantaneous sensitivity. Searching this dataset near-optimally for pulsars is an unsolved problem.

We're making progress on developing new search algorithms:

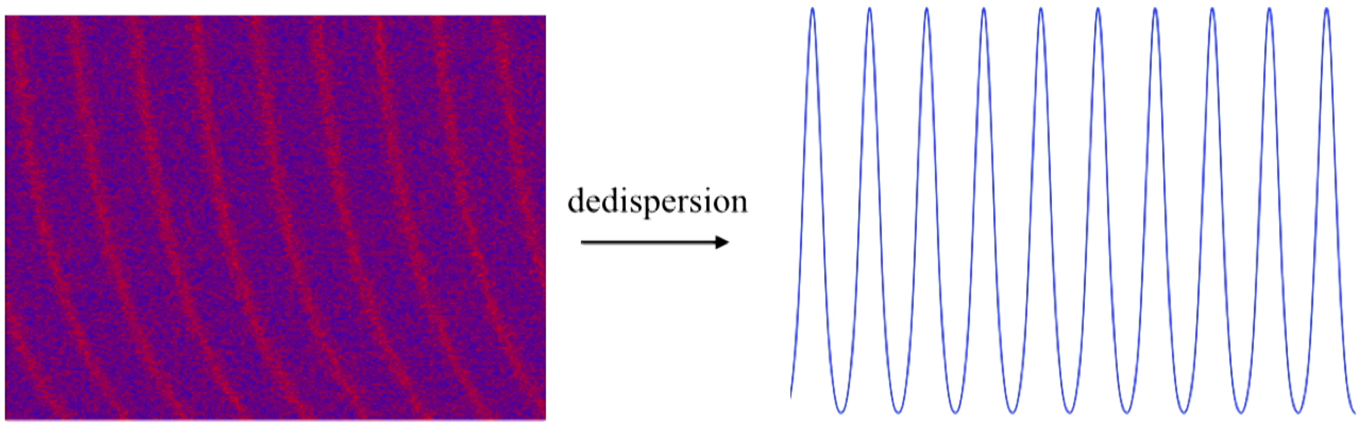
- For the case of a long contiguous timestream (more common than the CHIME case!) can do much better than current state-of-the-art (KMS, arxiv:1610.06381).
- The CHIME case (many discontiguous timestreams) is an extension we plan to study next!

Setting up the pulsar search problem

A pulsar appears as a quasiperiodic train of dispersed pulses.
(Plotted here with artificially high signal-to-noise.)



If the DM of the pulsar were known in advance, then applying dedispersion would collapse this to a 1D quasiperiodic timestream.



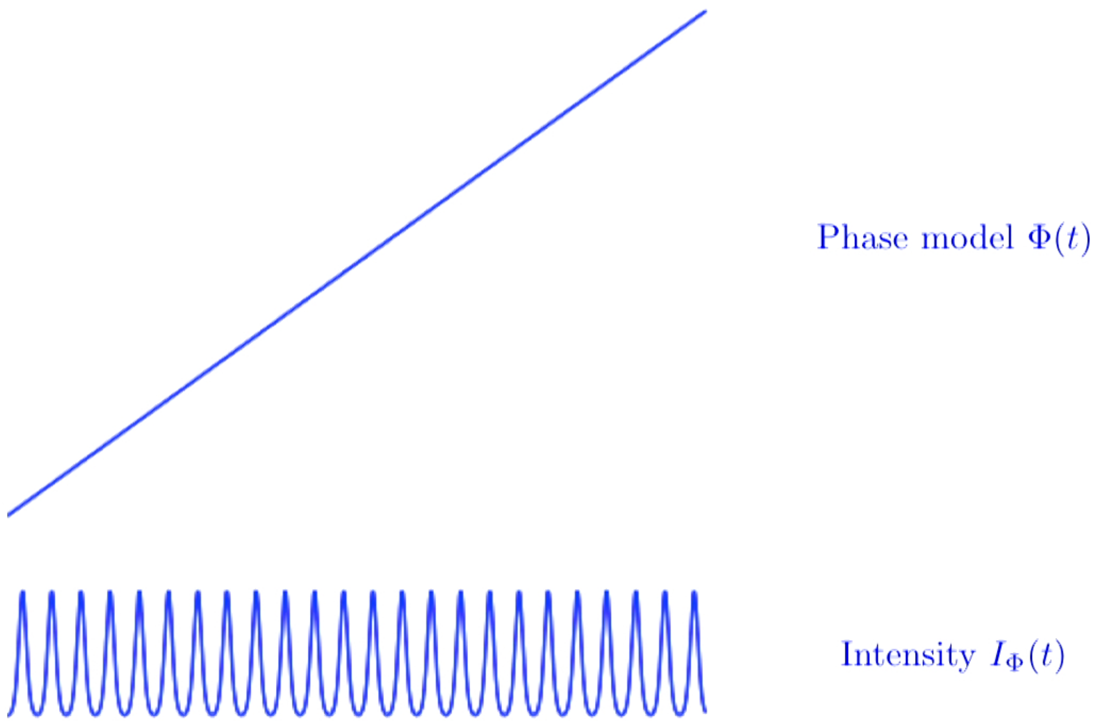
All pulsar search algorithms start with a big outer loop over:

- trial sky locations (~ 10^5 in CHIME)
- trial dispersion measures (10^3-10^4 in CHIME)

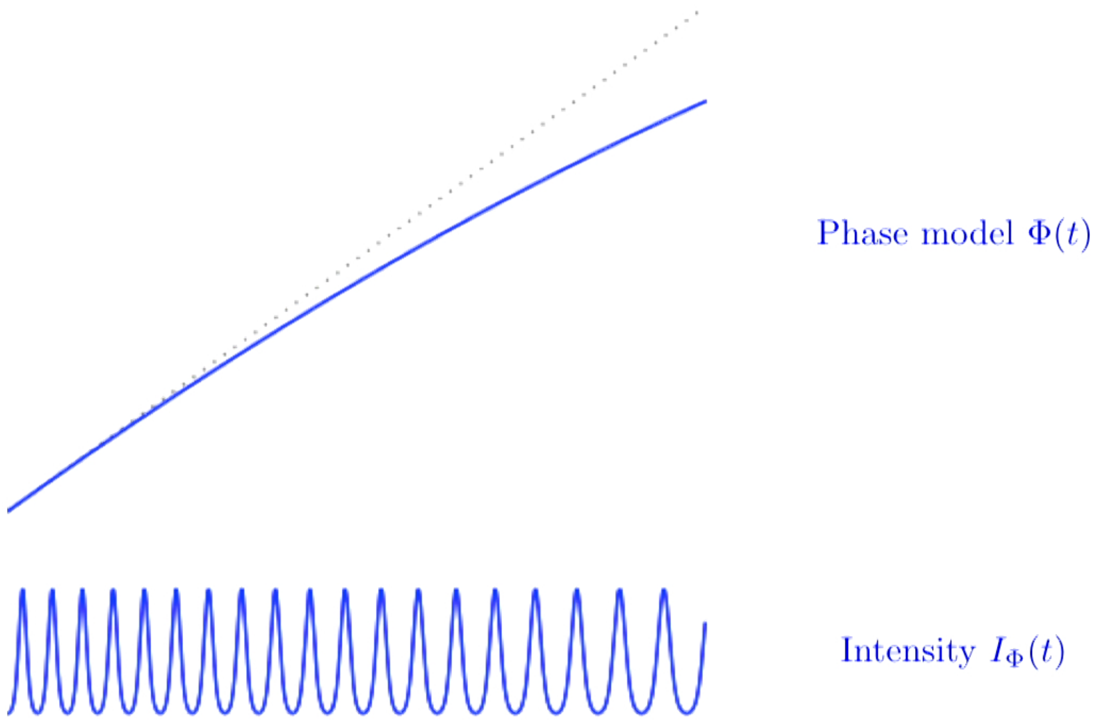
to reduce to the case of searching a 1D timestream.

Next we want to inner-loop over all phase models. The phase model defined as a function $\Phi(t)$ such that pulses appear when Φ is a multiple of 2π .

Example: regular pulsar, linearly evolving phase $\Phi(t) = \omega t$



Decelerating pulsar $\Phi(t) = \omega t - \alpha t^2$
(Plotted with size of spin-down parameter α exaggerated!)



Coherent search

Here is a manifestly optimal but slow algorithm. Loop over all possible phase models Φ and compute

$$\hat{\mathcal{E}}[\Phi] = \int d(t) I_\Phi(t)\, dt \quad \text{where} \left\{ \begin{array}{rl} d(t) = & \text{data timestream} \\ I_\Phi(t) = & \text{model timestream} \end{array} \right.$$

Computational cost is naively O(ST), where

S = size of the search space (# of independent phase models)
T = size of timestream

Coherent search

Here is a manifestly optimal but slow algorithm. Loop over all possible phase models Φ and compute

$$\hat{\mathcal{E}}[\Phi] = \int d(t) I_\Phi(t)\, dt \quad \text{where} \quad \begin{cases} d(t) = \text{data timestream} \\ I_\Phi(t) = \text{model timestream} \end{cases}$$

Simplest case: linearly evolving phase $\Phi(t) = \omega t + \Phi_0$
Interpret $\hat{\mathcal{E}}$-evaluation as a "transform" $d(t) \to \hat{\mathcal{E}}(\omega, \Phi_0)$

Result #1: this transform factors as a sequence of Fourier transforms (and other fast operations, e.g. interpolation).

Therefore, in this search space, the optimal search statistic can be evaluated very quickly! (O(S log T) vs O(ST))

Next case: "constant-acceleration" search $\Phi(t) = \Phi_0 + \omega_0 t - \alpha t^2$

Result #2: the transform $d(t) \to \hat{\mathcal{E}}(\Phi_0, \omega_0, \alpha)$ can be computed quickly with a tree algorithm.

Computational cost is O(S) [versus O(ST) for brute-force], where S is size of the search space.

This is best possible for a coherent search, where (by definition) we compute $\hat{\mathcal{E}}$ for every point in search space.

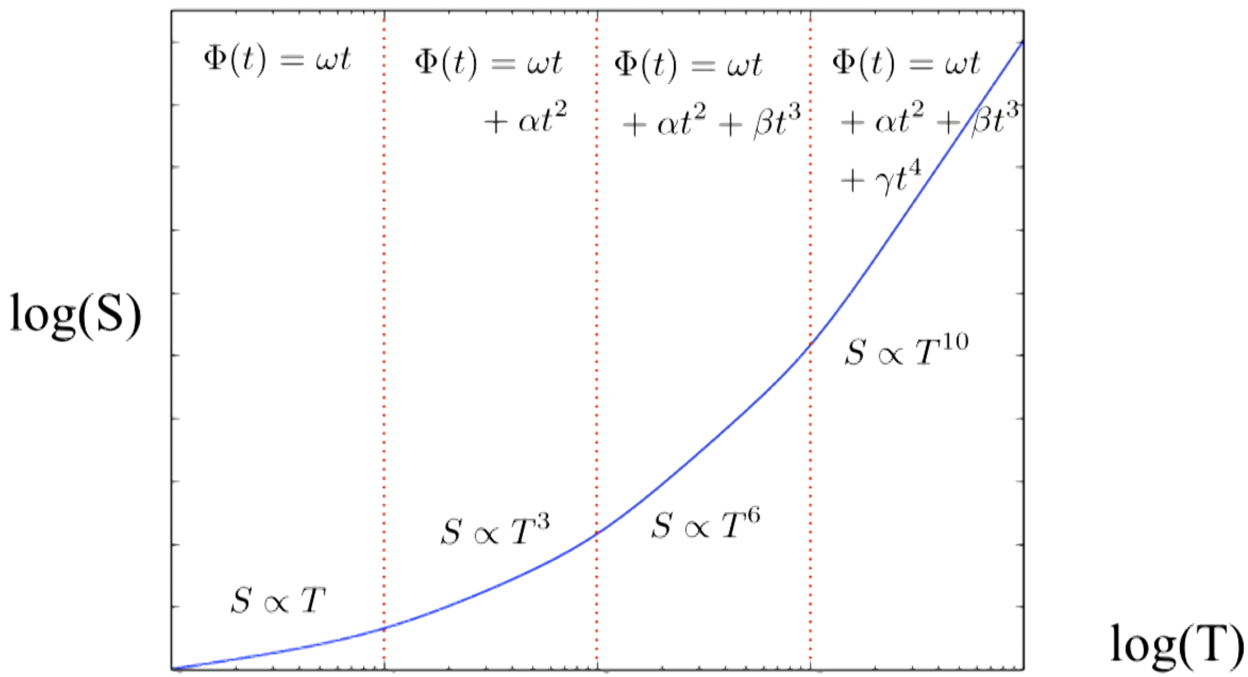
This is a big improvement, is it enough for CHIME?

Why pulsar search is so hard

The size S of the search space is a very rapidly increasing function of the timestream size T.

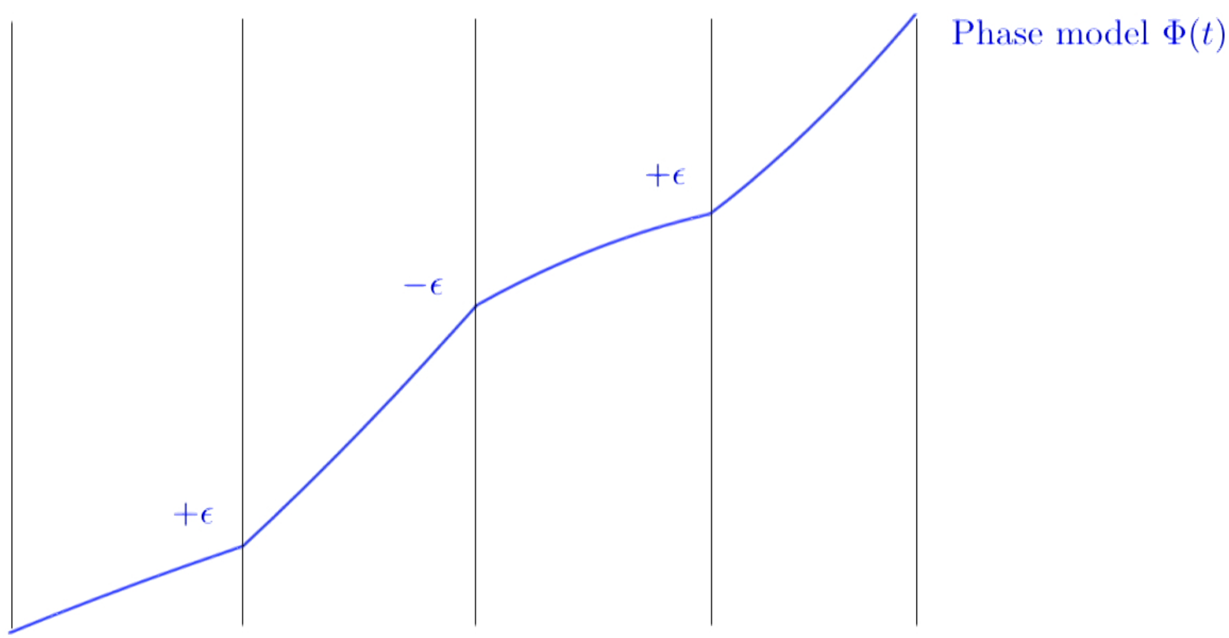
Example: consider modeling $\Phi(t)$ by a low-order polynomial.

In CHIME, S ~ 10^{30}! New ideas are needed.



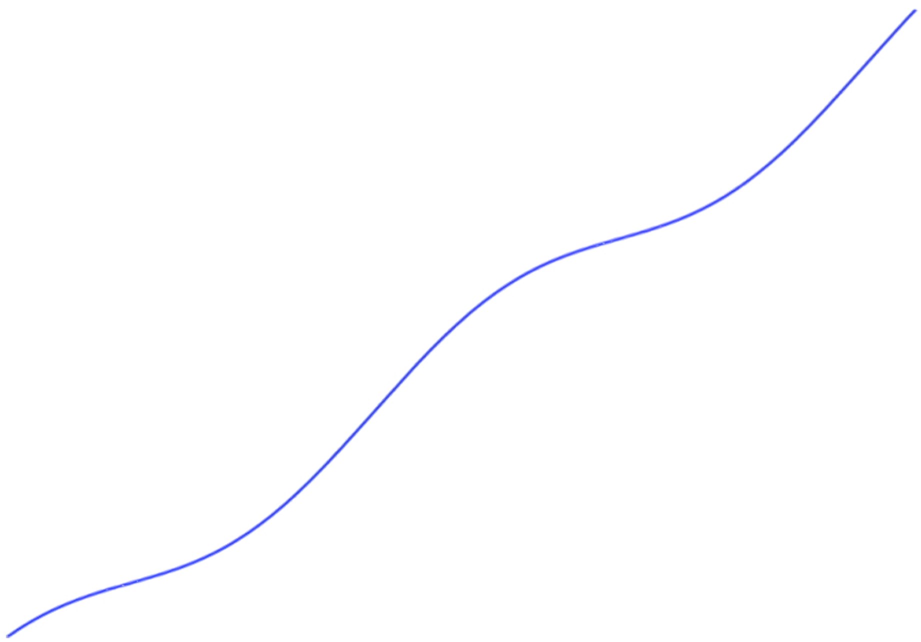
Semicoherent search

Divide the timestream into chunks of a fixed size. Consider "jagged" phase models whose acceleration $\ddot{\Phi}$ is constant in each chunk, but changes by $\pm\epsilon$ at chunk boundaries ($\Phi, \dot{\Phi}$ evolve continuously)



Semicoherent search

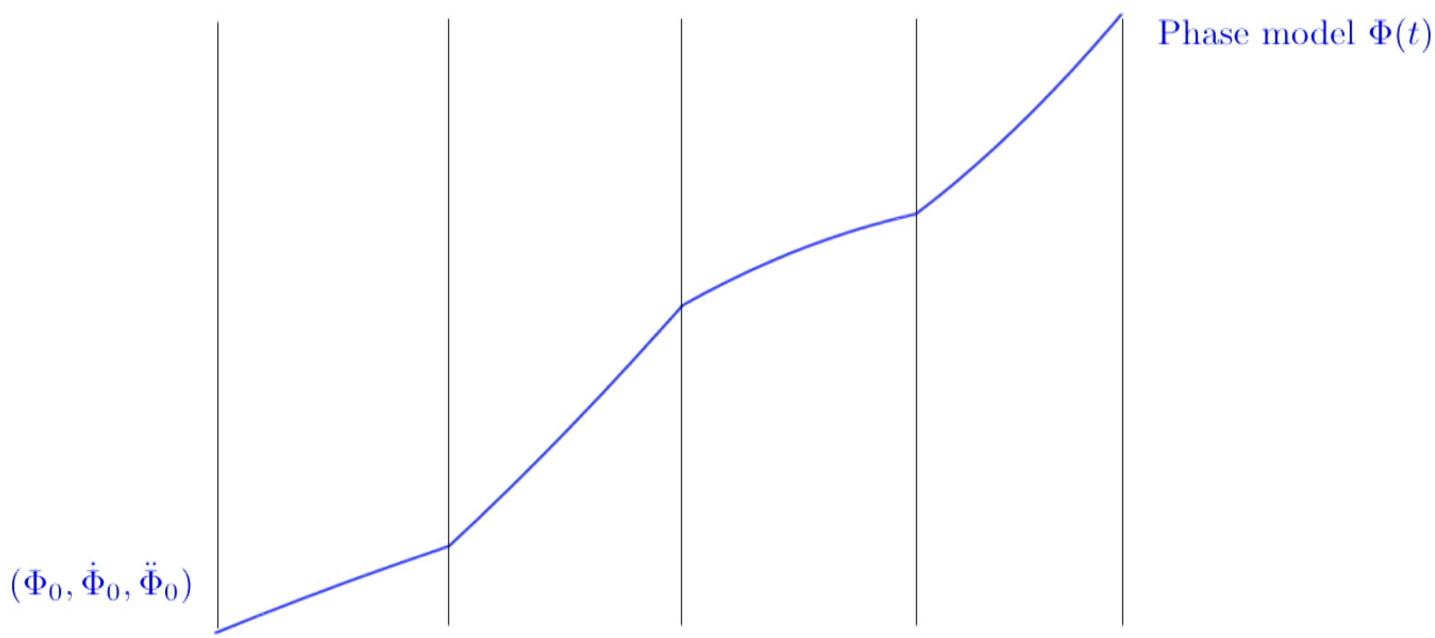
Note that complicated phase models (e.g. binary pulsar) can be approximated by jagged phase models.



For initial data $(\Phi_0, \dot{\Phi}_0, \ddot{\Phi}_0)$ define a statistic $\hat{\mathcal{H}}(\Phi_0, \dot{\Phi}_0, \ddot{\Phi}_0)$ which sums over 2^N jagged paths with initial conditions $(\Phi_0, \dot{\Phi}_0, \ddot{\Phi}_0)$

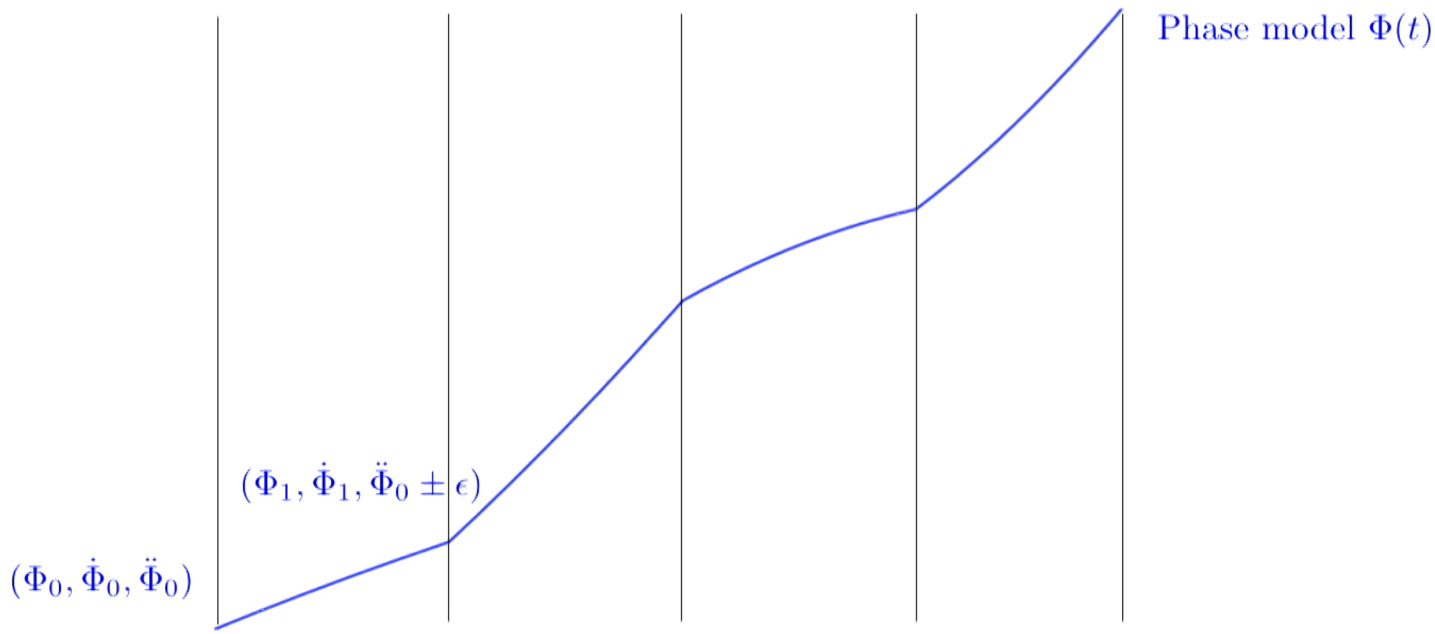
$$\hat{\mathcal{H}} = \sum_{\Phi(t)} \exp(r\hat{\mathcal{E}}[\Phi])$$

where $\hat{\mathcal{E}}[\Phi] = \int dt\, d(t) I_\Phi(t)$ is the coherent search statistic.



Key fact: $\hat{\mathcal{H}}$ is computable in O(N) time using recursion relations.

$$\hat{\mathcal{H}}(\Phi_0, \dot{\Phi}_0, \ddot{\Phi}_0) = \sum_{\Phi(t)} \exp(r\hat{\mathcal{E}}_0 + r\hat{\mathcal{E}}_1)$$
$$= e^{r\hat{\mathcal{E}}(\Phi_0, \dot{\Phi}_0, \ddot{\Phi}_0)}\big(\hat{\mathcal{H}}(\Phi_1, \dot{\Phi}_1, \ddot{\Phi}_0 + \epsilon)$$
$$+\hat{\mathcal{H}}(\Phi_1, \dot{\Phi}_1\, \ddot{\Phi}_0 - \epsilon)\big)$$



The $\hat{\mathcal{H}}$-search is fully coherent within each chunk. On longer timescales it sums over all ways of connecting coherent subsearches in a consistent way. ("jagged paths")

Computational cost no longer blows up!

cost = N_{chunks} x (cost of single-chunk coherent search)

Key question: how suboptimal is the search as N_{chunks} increases?
Answer: surprisingly optimal!
E.g. for $N_{chunks} = 64$, only ~20% suboptimal.

This is a big improvement over existing algorithms:

- For timestreams which are too long to search coherently, can search much deeper, should find lots of new pulsars!
- For timestreams which are short enough to search coherently, can "chunk" the search to make it much faster

Analogy with quantum-mechanical path integral

Just for fun! There is an analogy with the QM path integral:

classical action	$\hat{\mathcal{E}}$-statistic
$S = \int (\frac{1}{2}\dot{x}^2 - V(x))dt$	$\hat{\mathcal{E}} = \int d(t) I(\Phi(t))$
path integral	$\hat{\mathcal{H}}$-statistic
$Z = \int_{x(t)} \exp(i\hbar^{-1} S[x])$	$\hat{\mathcal{H}} = \sum_{\Phi(t)} \exp(r\hat{\mathcal{E}}[\Phi])$
Schrodinger equation	recursion relation
$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2} \frac{\partial^2 \psi}{\partial t^2} + V\psi$	$\hat{\mathcal{H}}(t_2) = e^{r\hat{\mathcal{E}}_2} \sum_{\ddot{\Phi}} \hat{\mathcal{H}}(t_1)$

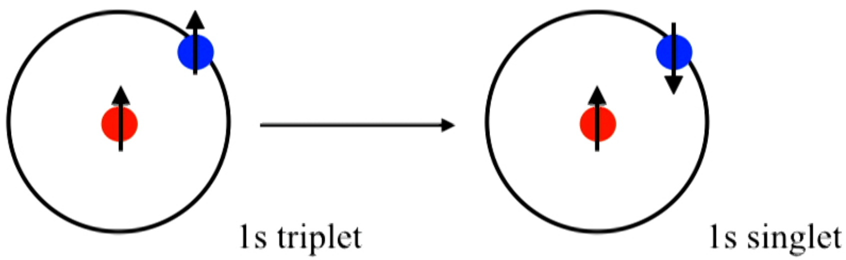
Pulsar search: summary

- We have developed some very interesting new algorithms!
- A very interesting near-term goal: use these algorithms to find new pulsars in existing datasets. (There is plenty of public data which can be searched more optimally with these algorithms.)
- Are these new algorithms useful for other experiments, e.g. SKA?
- Holy grail: develop these ideas further, find large numbers of new pulsars with CHIME. Will this be possible? I don't know yet!

1. Introduction
2. Fast radio bursts
3. Searching for new pulsars
4. Things I didn’t have time to talk about
5. Thoughts for the future

Cosmology via the 21-cm line

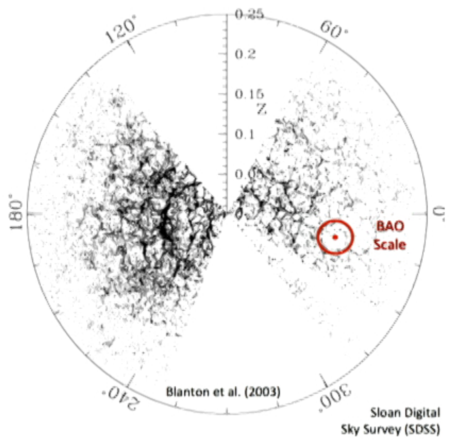
Neutral hydrogen (HI) has a long-lived emission line at λ_0=21cm



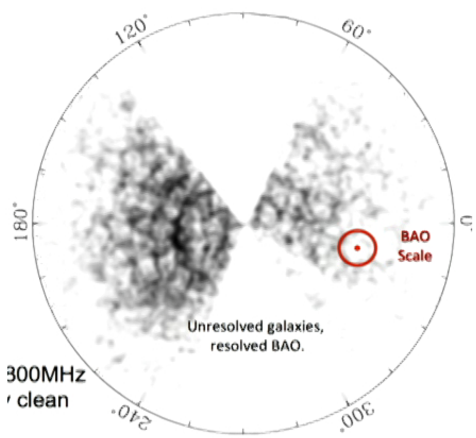
We observe the intensity of this emission as a function of sky angles θ,ϕ and wavelength λ_{obs} = (21 cm)(1+z).

The resulting 3D map traces cosmological structure.

At low redshifts ($z \lesssim 5$), hydrogen is mostly ionized, but some neutral hydrogen survives in "self-shielding" systems.



Spectroscopic galaxies: number density $n(\theta,\phi,z)$ traces large-scale structure.



21-cm intensity mapping: brightness temperature $T(\theta,\phi,z)$ traces LSS.

CHIME:

redshift range $0.8 \leq z \leq 2.5$

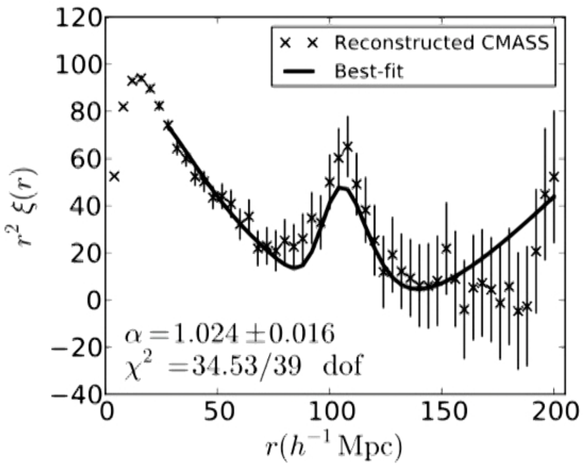
radial resolution $\Delta z = 0.002$ (~5 Mpc)

angular resolution 0.3 deg (~20 Mpc)

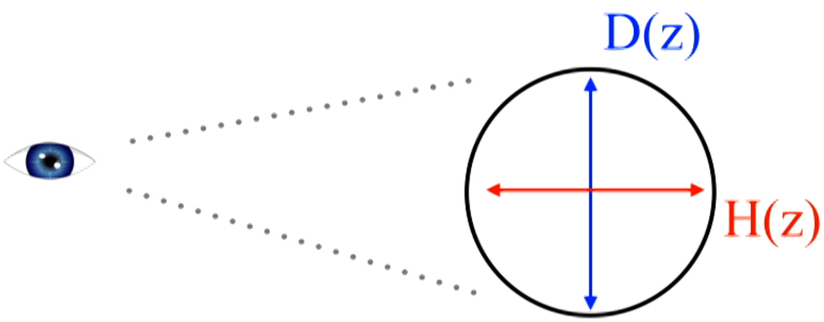
Cosmology via the 21-cm line

Can use this 3D map to do large-scale structure: baryon acoustic oscillations, lensing, redshift-space distortions, etc.

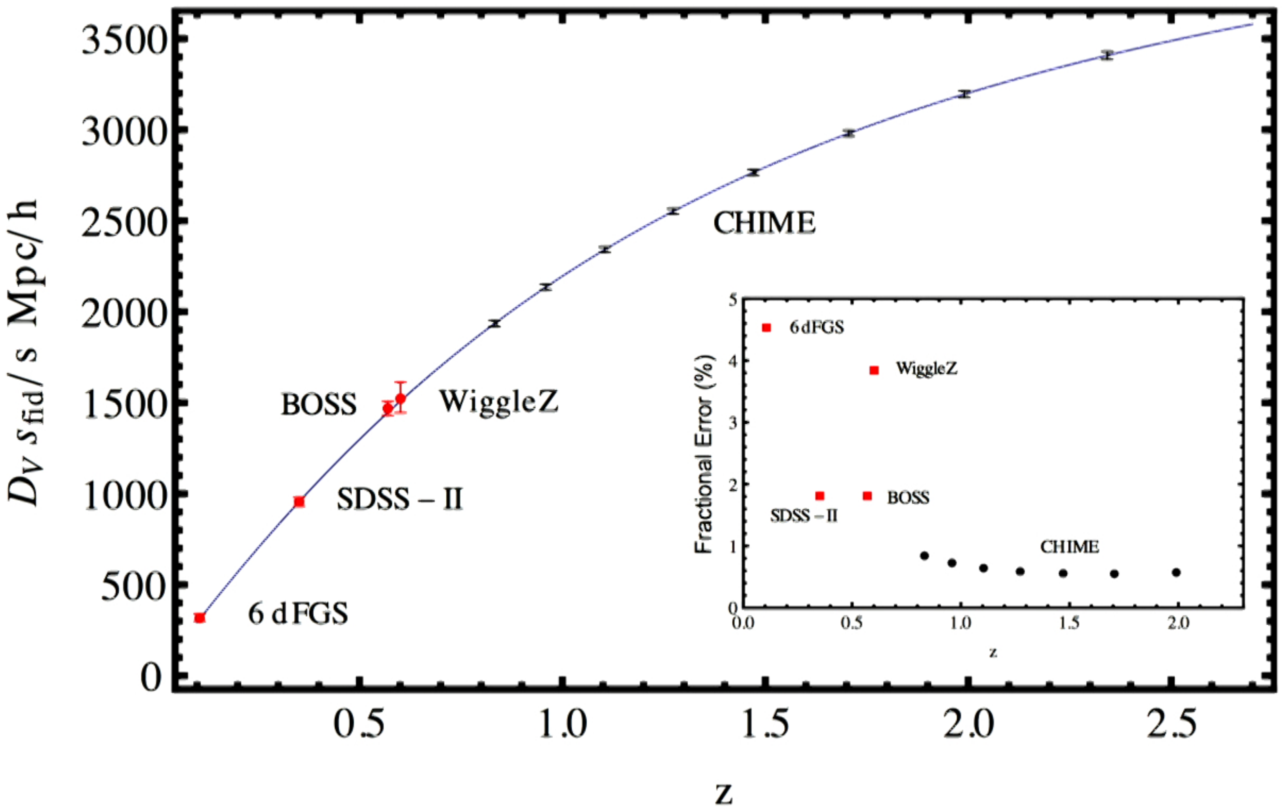
I emphasized transients in this talk, but the main goal of CHIME is to measure the BAO "standard ruler"!



SDSS (2012)

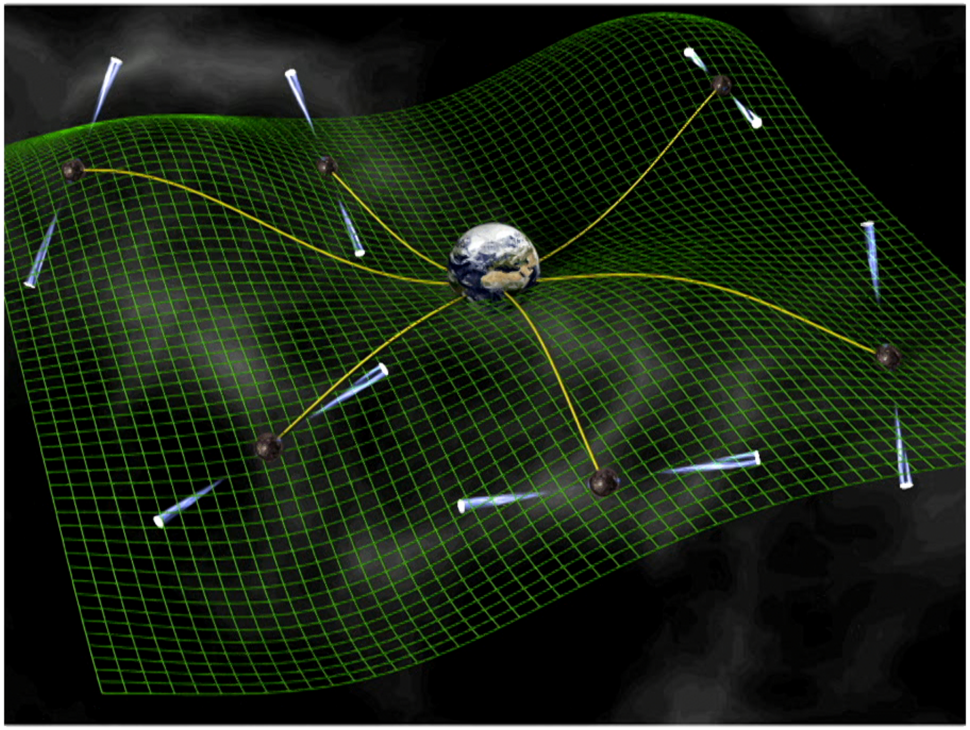


Baryon acoustic oscillations



Precision pulsar timing

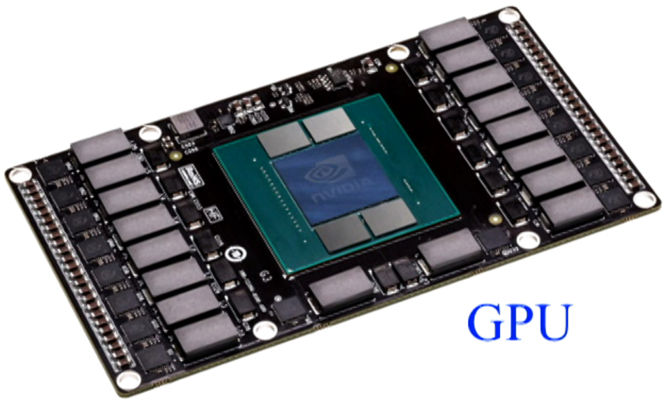
Daily timing of known pulsars at CHIME frequencies (400-800 MHz) can contribute to global efforts to detect gravity waves using pulsar networks.



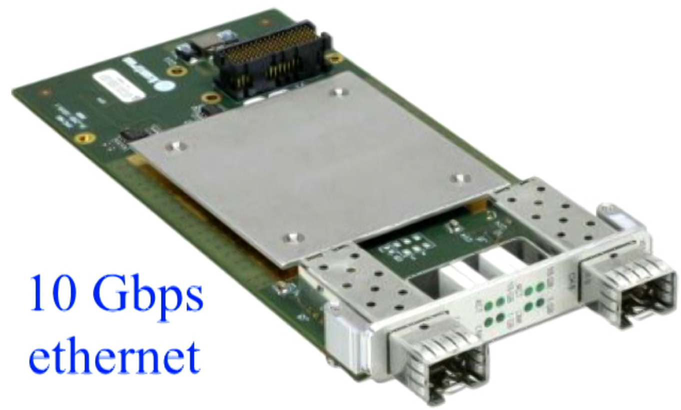
1. Introduction
2. Fast radio bursts
3. Searching for new pulsars
4. Things I didn't have time to talk about
5. Thoughts for the future

The CHIME design is a breakthrough if mapping speed per dollar is the metric.

Main advance on hardware side: using inexpensive commodity hardware in ways that haven't been done before.



GPU



10 Gbps ethernet

Leveraging Moore's law!

Consider the cost of scaling up CHIME.
Suppose we increase collecting area A at fixed antenna density.

Cost of computing = $\mathcal{O}(A^2 e^{-T/T_{\rm Moore}})$
Cost of everything else = $\mathcal{O}(A)$
Mapping speed $M = \mathcal{O}(A^2)$

So total cost depends on mapping speed as

$$\text{Cost} = \begin{cases} \mathcal{O}(Me^{-T/T_{\rm Moore}}) & \text{if computation-dominated} \\ \mathcal{O}(M^{1/2}) & \text{otherwise} \end{cases}$$

Whether computation-dominated or not, a much shallower dependence than other areas of astronomy.

CHIME costs are computation-dominated, but not by a large factor.
End of computation-dominated era in ~10 years?
An $M^{1/2}$ scaling would be really amazing!

A radio astronomy slogan: "radio telescopes are paperclips connected to supercomputers".

Nowadays the supercomputers aren't much more expensive than the paperclips, and might even be cheaper soon.

CHIME is demonstrating the technology to digitize the electric field over huge collecting areas very cheaply. This technology is scalable to much larger instruments.

The impact of this technology will partly depend on the answer to this question: which algorithms in radio astronomy can be scaled up to these new data volumes?


Thanks!