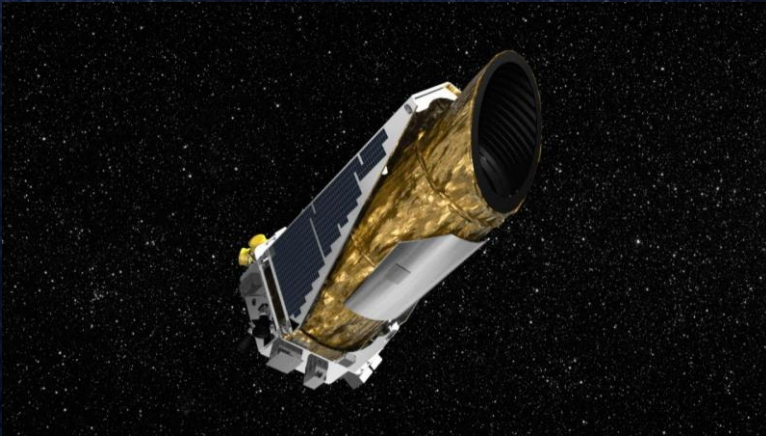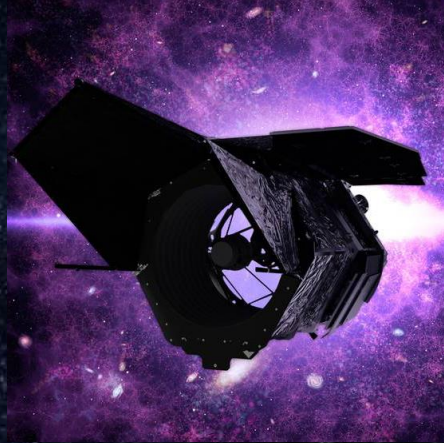Data mining astronomical time series using subMIT

Kevin Burdge

Massachusetts Institute of Technology

Department of Physics
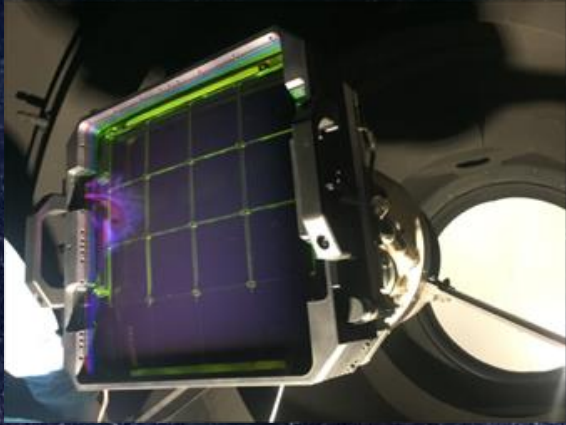
Astronomy is currently undergoing a revolution in gathering temporal information
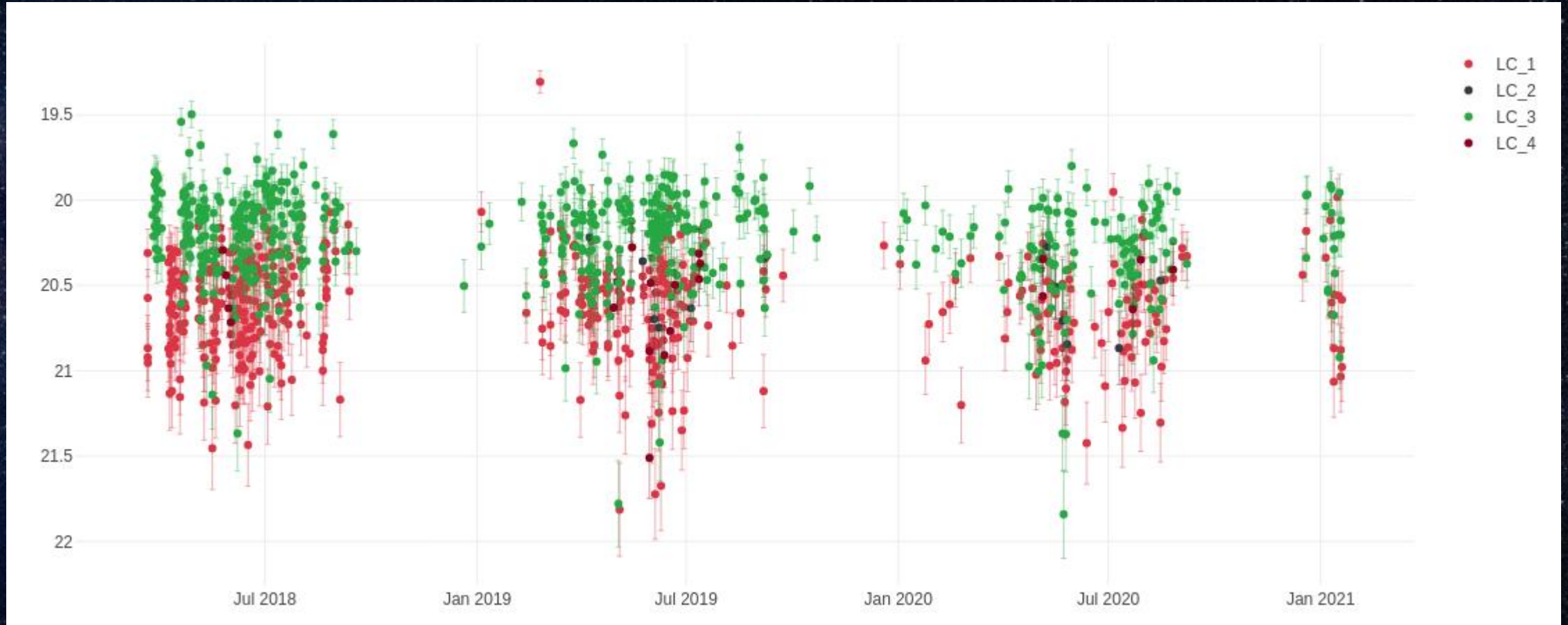
48 inch Samuel Oschin Schmidt telescope

Discovery: using the Zwicky Transient Facility
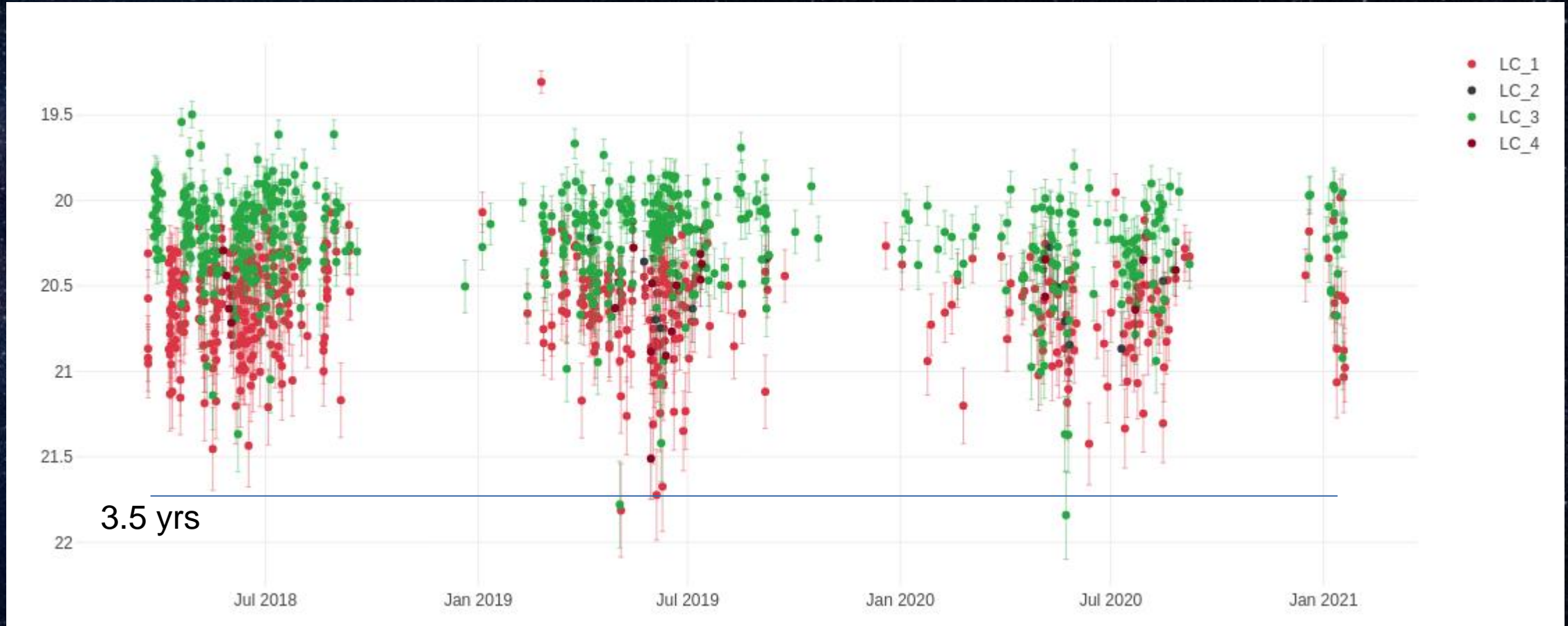
The crucial element: ZTF has a large field of view, and accumulates many images quickly

And what do we do to this data?

And what do we do to this data?

3.5 yrs

Find a signal in it!

Find a signal in it!

6.91 minutes!
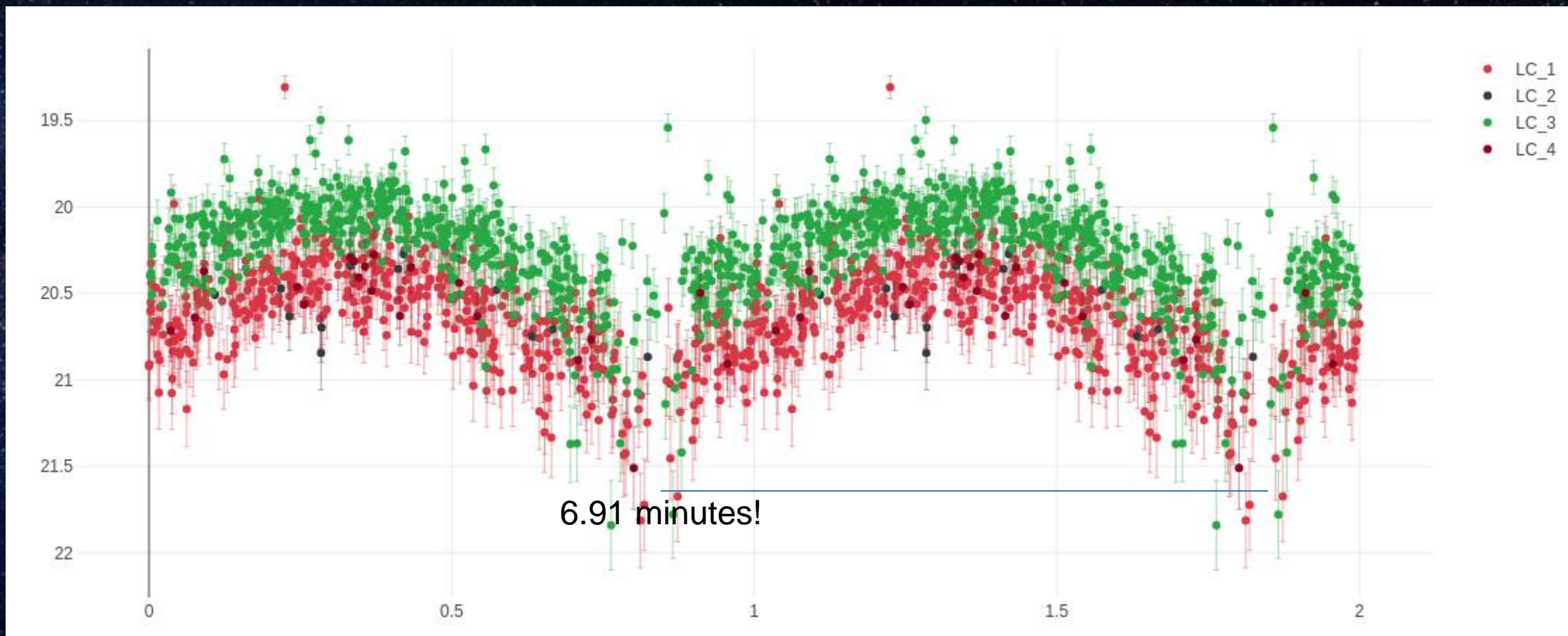
But it's not trivial to find these objects…

Searching for minute periods in data sampled over years->enormous frequency grids

Graphics processing units help a lot with this

But it's not trivial to find these objects...

TESS poses different challenges:

- Equispaced sampled data

- Baseline is short, but large number of samples

- Limited depth, but high precision

What we could do with subMIT?

- Equispaced sampled data

- Baseline is short, but large number of samples

- Limited depth, but high precision

What we could do with subMIT?

- A single RTX 2080 can process an entire sector of 2 min cadence data in 5 mins
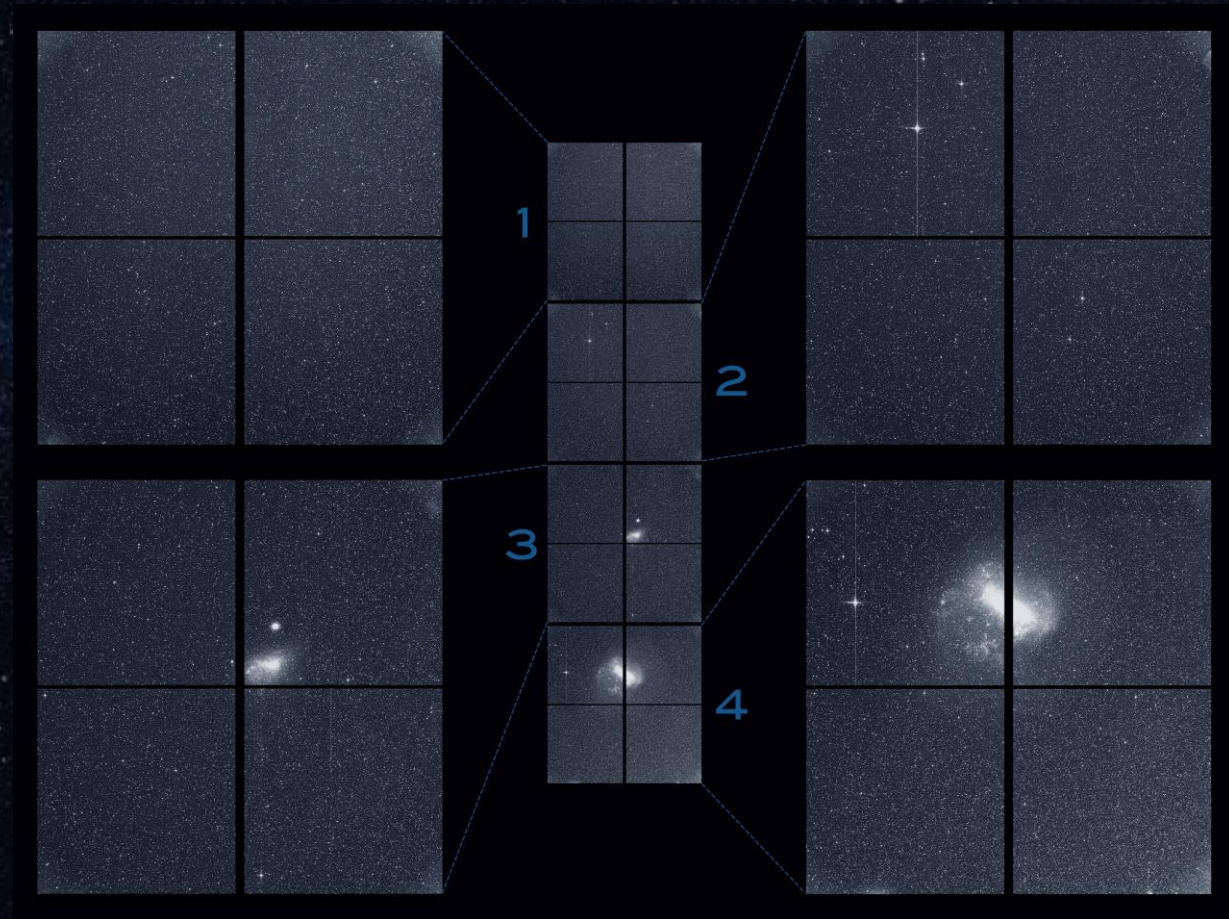
- There are only ~45 sectors of data

What we could do with subMIT?

- A single RTX 2080 can process an entire sector of 2 min cadence data in 5 mins

- There are only ~45 sectors of data
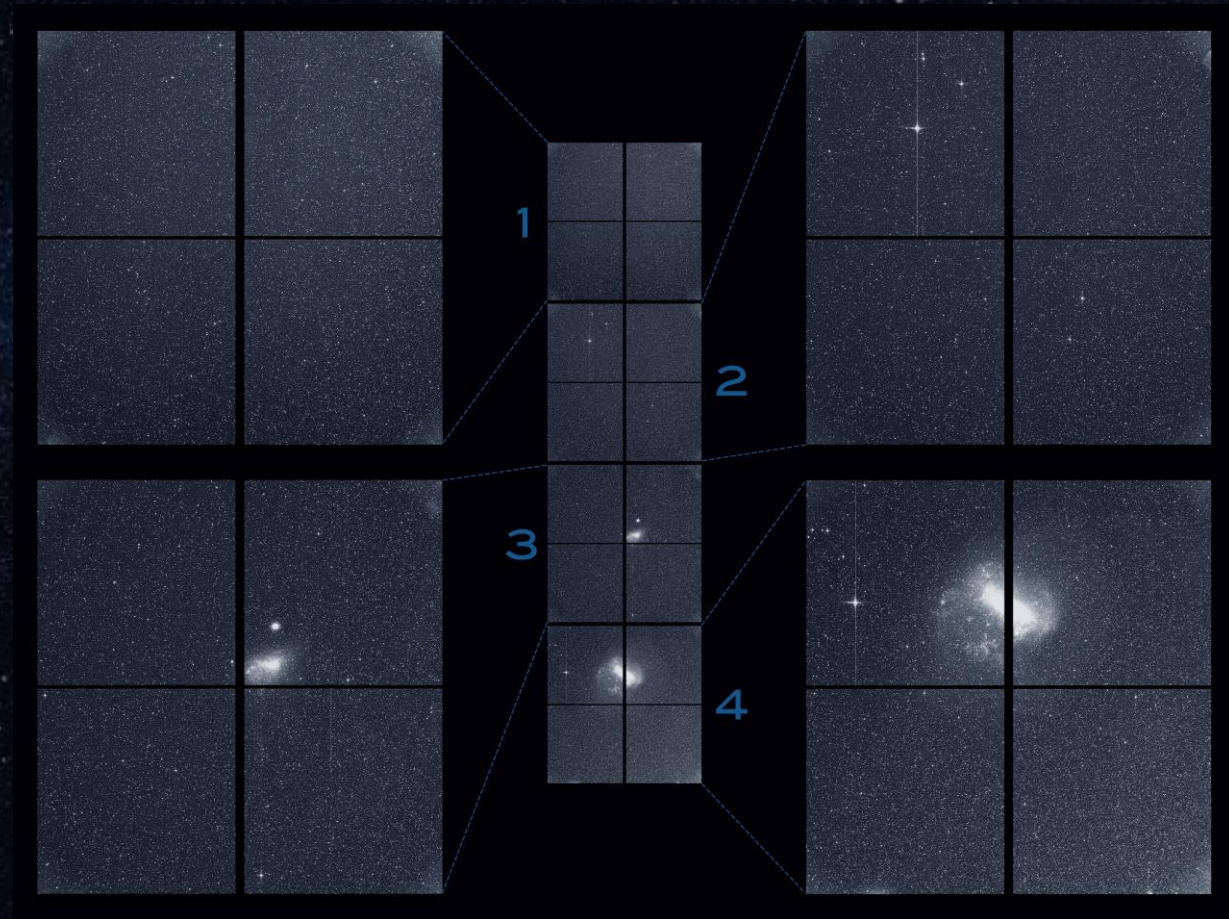
- We need more ambitious tasks

What if we took an FFT of every pixel?

- A single RTX 2080 can process an entire sector of 2 min cadence data in 5 mins

- There are only ~45 sectors of data

- We need more ambitious tasks

What if we took an FFT of every pixel?

- Each camera has four 2Kx2K CCDs, and there are four cameras, so 64 million pixels

- One GPU could process an entire TESS sector's pixels in ~74 days

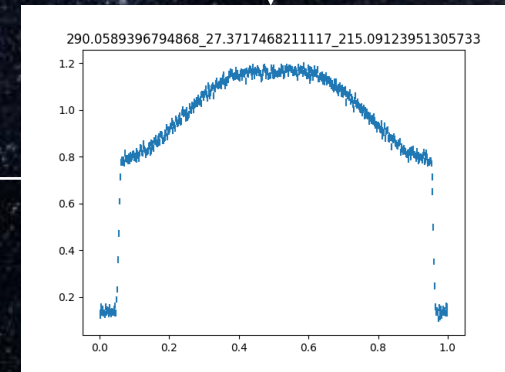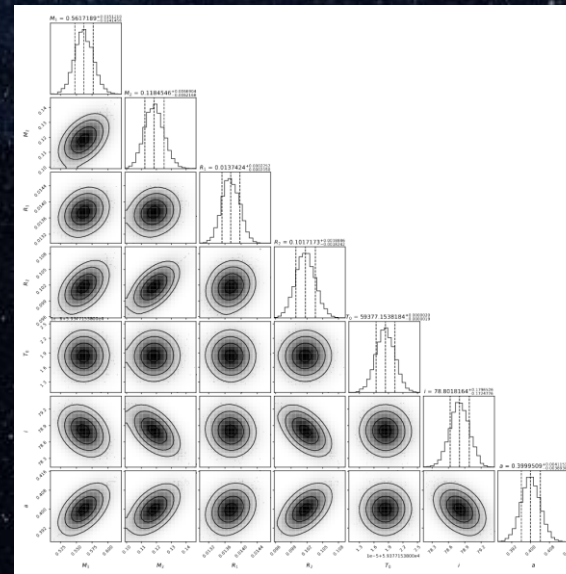- This is getting into interesting territory

Student projects

- The MIT undergraduate physics curriculum sorely lacks exercises in programming

- SubMIT could be a valuable asset in teaching undergraduates how to work in a cluster environment

- Python is the language of choice for young scientists, but few have experience with optimizing code
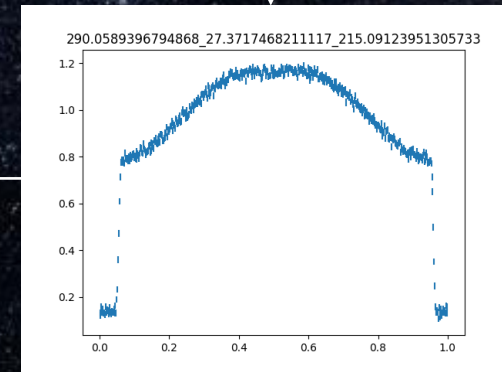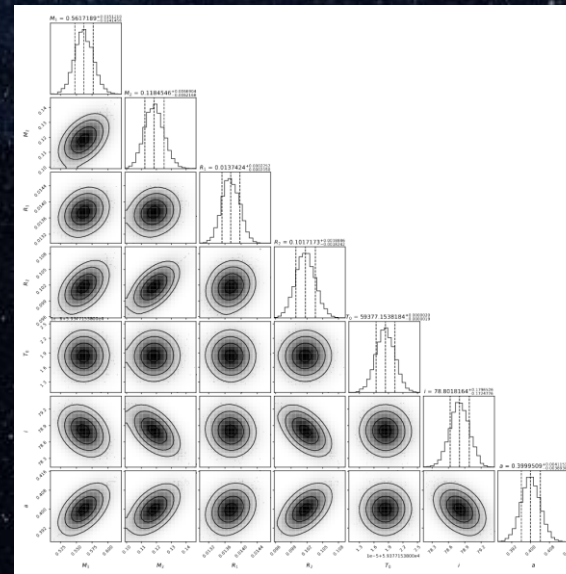
Student projects

- The MIT undergraduate physics curriculum sorely lacks exercises in programming

- SubMIT could be a valuable asset in teaching undergraduates how to work in a cluster environment

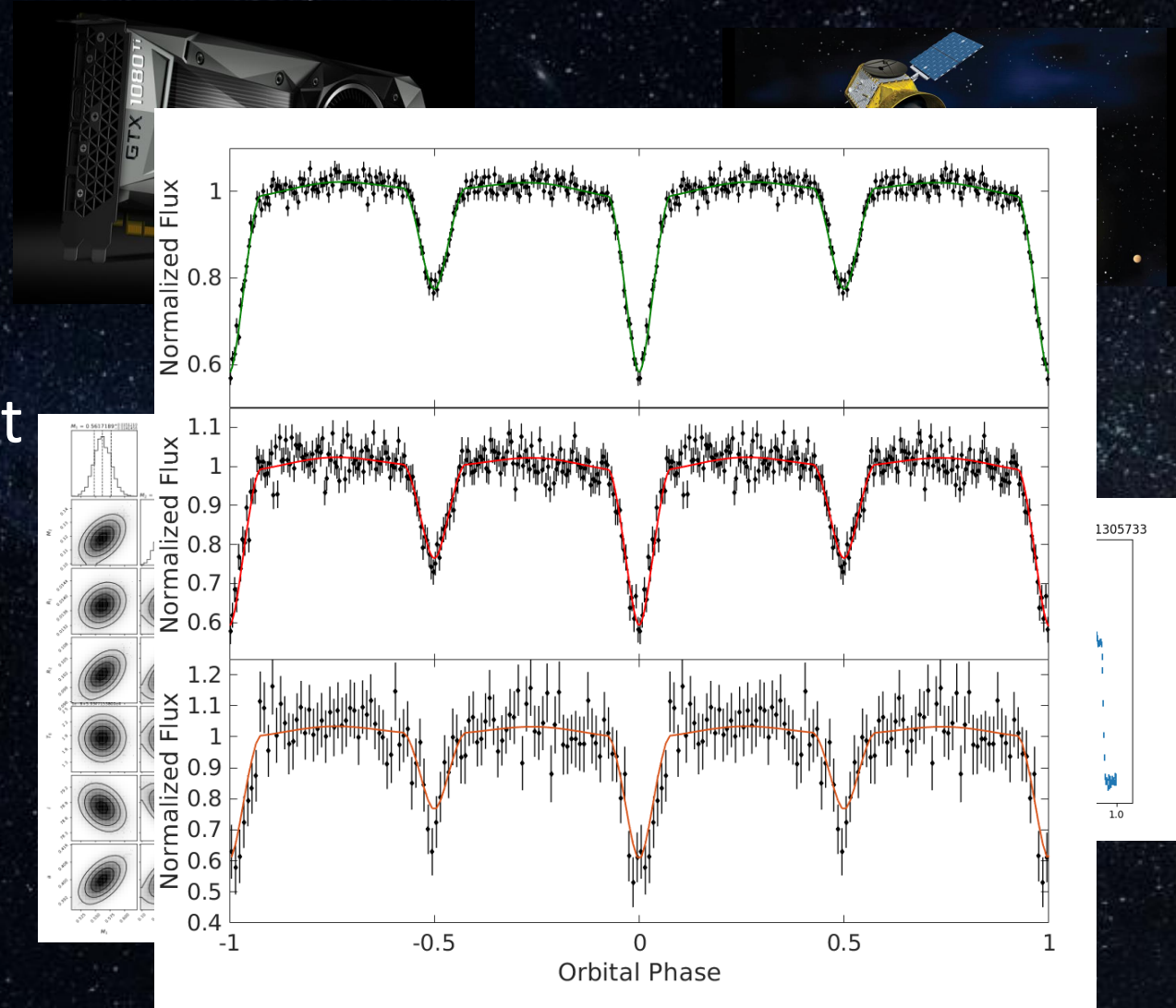- Python is the language of choice for young scientists, but few have experience with optimizing code

Student projects

- Basic python programming

- How to data mine a public dataset

- Modeling real data

- How to use an MCMC

Student projects

- Basic python programming

- How to data mine a public dataset

- Modeling real data

- How to use an MCMC

Conclusions

- The volume of large public datasets is rapidly increasing

- The computational resources availible to the average researcher are also rapidly increasing

- SubMIT is an opportunity to ensure that the next generation of researchers are ready to leverage the computational resources and datasets of the future

-