PRESENTER: JAMES GUILLOCHON, HARVARD CFA

BUILDING BETTER MODELS FOR INFERENCE

HIGH-LEVEL GOALS

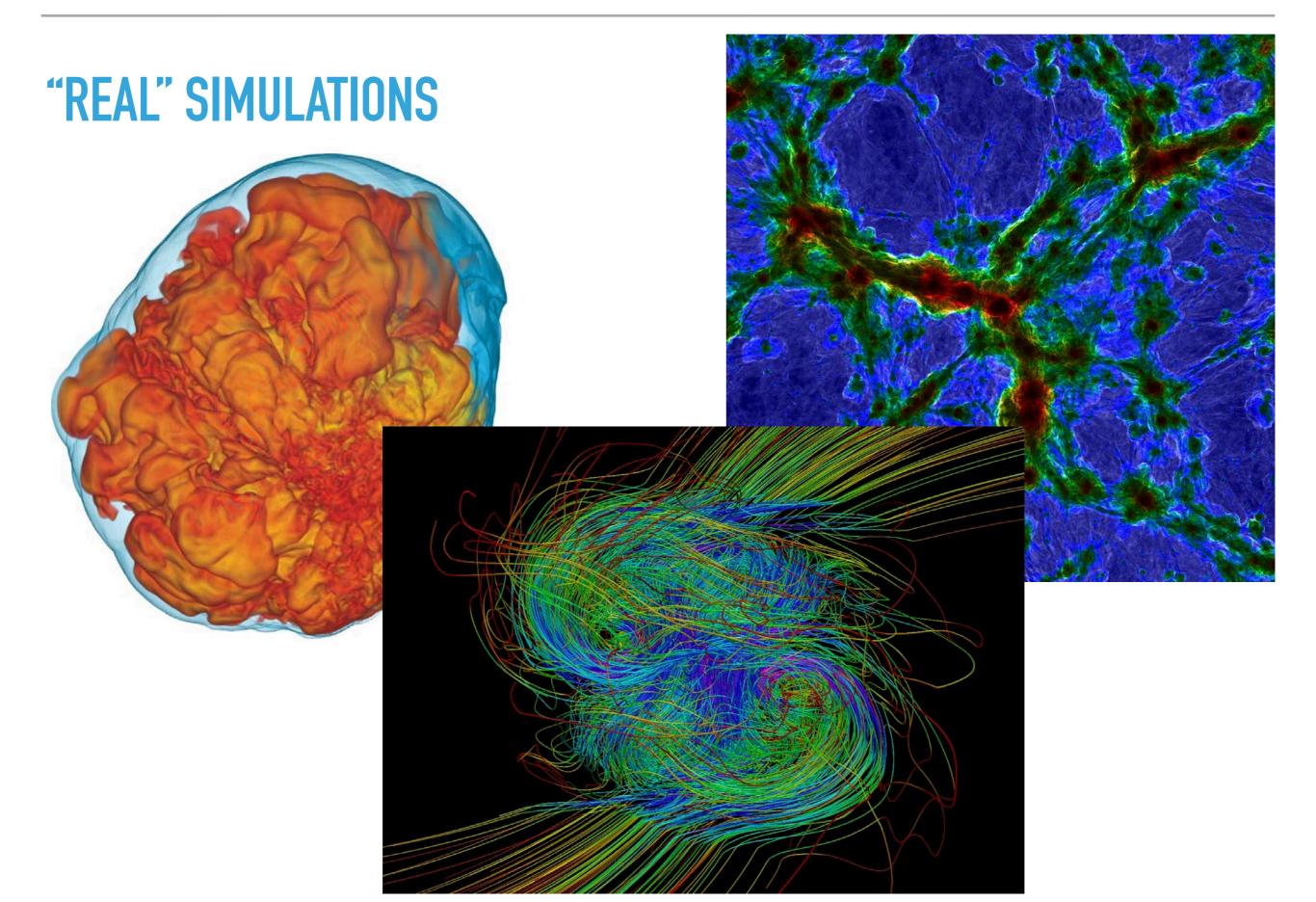
- We would like to be able to construct descriptions of reality in order to accurately make predictions.
- The behavior of a complicated system is, in detail, nearly impossible to model one-for-one on a computer.
- Thus, most models in astrophysics are *heuristics* for the real systems: inaccurate at some level, but cheaper to compute.
- When we use such heuristics, we must construct them carefully to make sure we're learning about the physics of the system, not the particulars of our heuristic.

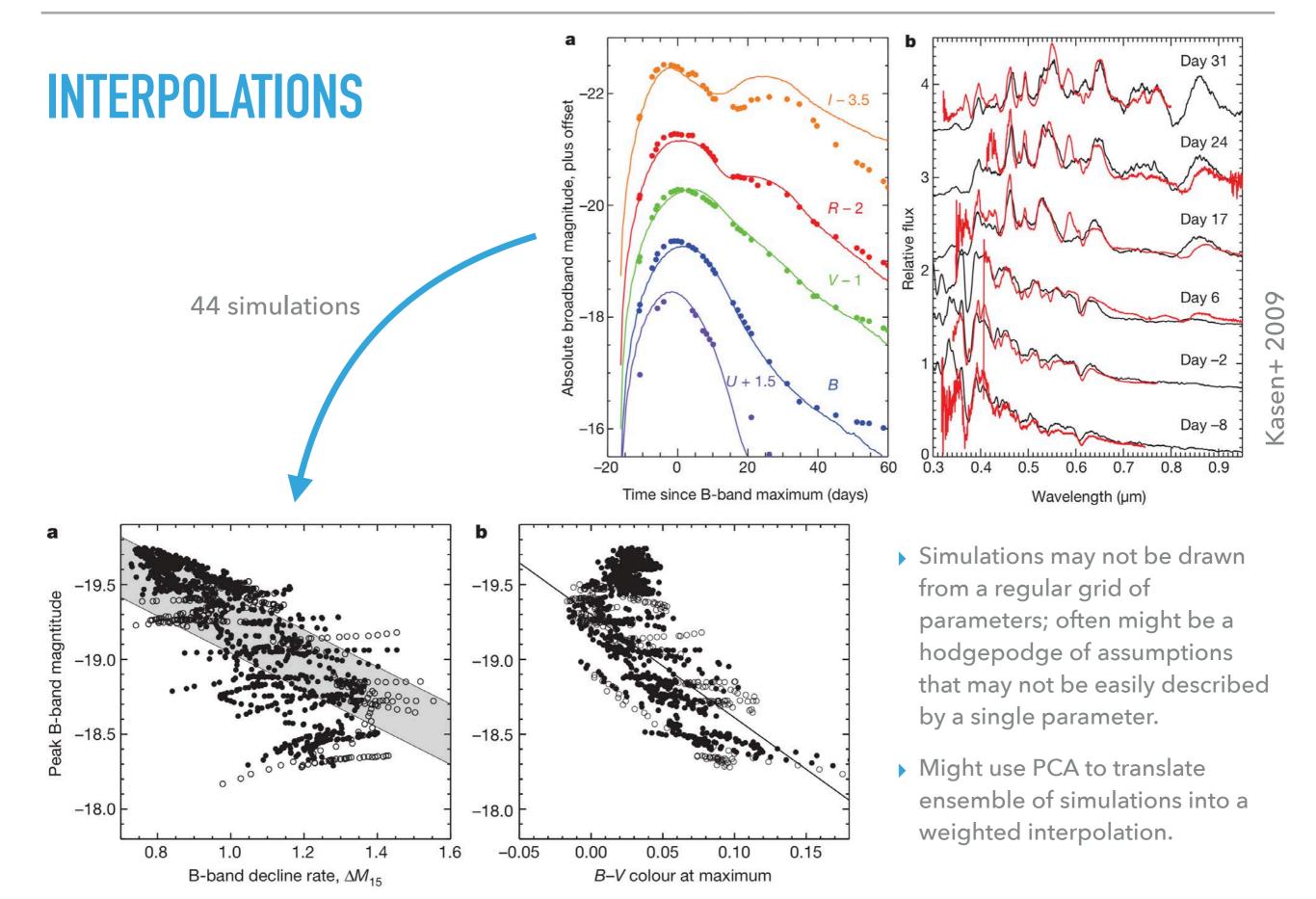
WHAT MAKES A GOOD HEURISTIC MODEL?

- Accurate as possible.
- Fast to compute.
- Simple to understand.
- Relatable to the underlying system.
- Friendly to optimizers/samplers.

ALL MODELS ARE HEURISTIC!

- No astrophysical simulations compute every particle's state at all times with perfect accuracy.
- We tend to study the same problem at a variety of different approximations.
 - ▶ 1/2/3D simulations with modules for all relevant physics, solving explicit/implicit transport equations at a set of discretized positions over a set of discretized times. Very expensive, some groups might only run a couple of these per year!
 - Interpolations of the above: construct a set of simulations that were expensive to run individually, build N-dimensional tables of various observables, interpolate between simulations to approximate. Accuracy depends on number of simulations available; the less expensive the simulation the more it's likely to be missing in terms of physics.
 - Simple semi-analytical models involving some integrals. Cheap to compute but not analytically differentiable, limited accuracy.
 - Purely analytical models; linear combinations of differentiable functions. Extremely cheap to compute, analytical derivatives available. Potentially very poor accuracy, except for self-similar classes of problems.





SEMI-ANALYTICAL MODELS

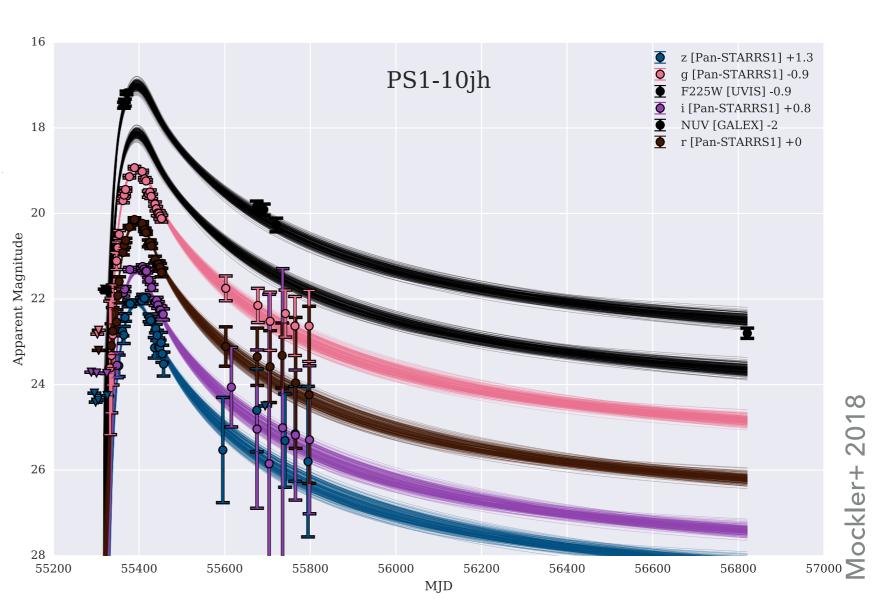
$$\dot{M}_{\rm d}(t) = \dot{M}_{\rm fb}(t) - M_{\rm d}(t)/T_{\rm viscous},$$

$$\dot{M}_{\rm d}(t) = \frac{1}{T_{\rm viscous}} \left(e^{-t/T_{\rm viscous}} \int_0^t e^{t'/T_{\rm viscous}} \dot{M}_{\rm fb}(t') dt'\right),$$

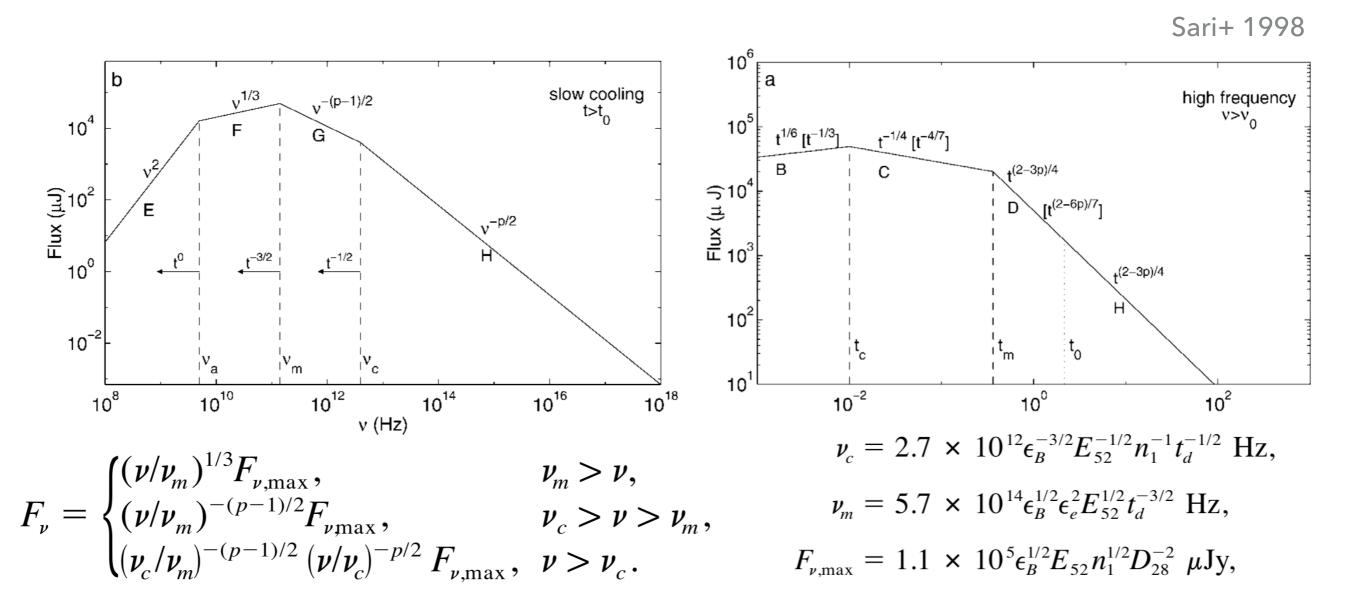
$$F_{\nu} = \frac{2\pi h \nu^3}{c^2} \frac{1}{\exp(h\nu/kT_{\rm eff}) - 1} \frac{R_{\rm phot}^2}{D^2},$$

$$T_{\rm eff} = \left(\frac{L}{4\pi\sigma_{\rm SB}R_{\rm phot}^2}\right)^{1/4}.$$

- Key challenge: Includes multiple integrals (both from the model and convolving filter with SED).
- The accretion rate above comes from an interpolation table of hydrodynamical simulations.



ANALYTIC



BUILDING BETTER MODELS

GETTING A LITTLE CARRIED AWAY

TABLE 1 NORMALIZATION OF THE DIFFERENT POWER-LAW SEGMENTS

PLS	β	$F_{\nu}(k=0)$ (mJy)	$F_{\nu}(k=2)$ (mJy)	
A 5/2		$1.18(4.59 - p)10^{8}(1 + z)^{9/4}\epsilon_{B}^{-1/4}n_{0}^{-1/2}E_{52}^{1/4}t_{days}^{5/4}d_{L28}^{-2}\nu_{14}^{5/2}$	$2.96(4.59 - p)10^{7}(1 + z)^{7/4}\epsilon_{B}^{-1/4}A_{*}^{-1}E_{52}^{3/4}t_{days}^{7/4}d_{L28}^{-2}\nu_{14}^{5/2}$	
В	2	$4.20 \frac{3p+2}{3p-1} 10^9 (1+z)^{5/2} \bar{\epsilon}_e n_0^{-1/2} E_{52}^{1/2} t_{\text{days}}^{1/2} d_{L28}^{-2} \nu_{14}^2$	$1.33 \frac{3p+2}{3p-1} 10^9 (1+z)^2 \bar{\epsilon}_e A_*^{-1} E_{52} t_{\text{days}} d_{L28}^{-2} \nu_{14}^2$	
C	11/8	$8.01\times 10^5(1+z)^{27/16}\epsilon_B^{-1/4}n_0^{-5/16}E_{52}^{7/16}t_{\rm days}^{11/16}d_{L28}^{-2}\nu_{14}^{11/18}$	$3.28\times 10^{5}(1+z)^{11/8}\epsilon_{B}^{-1/4}A_{*}^{-5/8}E_{52}^{3/4}t_{\rm days}d_{L28}^{-2}\nu_{14}^{11/8}$	
D	1/3	$27.9 \frac{p-1}{3p-1} (1+z)^{5/6} \overline{\epsilon_c}^{-2/3} \epsilon_B^{1/3} n_0^{1/2} E_{52}^{5/6} t_{\text{days}}^{1/2} d_{L28}^{-2} \nu_{14}^{1/3}$	$211\frac{p-1}{3p-1}(1+z)^{4/3}\overline{\epsilon_e}^{-2/3}\overline{\epsilon_B}^{1/3}A_*E_{52}^{1/3}d_{L28}^{-2}\nu_{14}^{1/3}$	
E	1/3	$73.0(1+z)^{7/6}\epsilon_B n_0^{5/6} E_{52}^{7/6} t_{\rm days}^{1/6} d_{\rm L28}^{-2} \nu_{14}^{1/3}$		
F	-1/2	$6.87(1+z)^{3/4}\epsilon_{B}^{-1/4}E_{52}^{3/4}t_{\rm days}^{-1/4}d_{\rm L28}^{-2}\nu_{14}^{-1/2}$	$6.68(1+z)^{3/4}\epsilon_B^{-1/4}E_{52}^{3/4}t_{days}^{-1/4}d_{L28}^{-2}\nu_{14}^{-1/2}$	
G	(1 - p)/2	$0.461(p-0.04)e^{2.53p}(1+z)^{(3+p)/4}\bar{e}_c^{p-1}\epsilon_B^{(1+p)/4}n_0^{1/2}E_{52}^{(3+p)/4}r_{\mathrm{days}}^{3(1-p)/4}d_{L28}^{-2}\nu_{14}^{(1-p)/2}$	$3.82(p-0.18)e^{2.54p}(1+z)^{(5+p)/4}e_e^{p-1}\epsilon_B^{(1+p)/4}A*E_{\mathfrak{Q}}^{(1+p)/4}t_{\mathrm{days}}^{(1-3p)/4}d_{228}^{-2}\nu_{14}^{(1-p)/2}$	
Н	-p/2	$0.855(p-0.98)e^{1.95p}(1+z)^{(2+p)/4}\epsilon_e^{p-1}\epsilon_B^{(p-2)/4}E_{52}^{(2+p)/4}t_{\rm days}^{(2-3p)/4}d_{L28}^{-2}\nu_{14}^{-p/2}$	$0.0381(7.11-p)e^{2.76p}(1+z)^{(2+p)/4}\epsilon_r^{p-1}\epsilon_B^{(p-2)/4}E_{52}^{(2+p)/4}t_{\rm days}^{(2-3p)/4}d_{L28}^{-2}\nu_{14}^{-p/2}$	

Note.—First two columns give the labels and the spectral slope, β , of the different PLSs (see Fig. 1), while the last two columns give the asymptotic flux density within each PLS for k=0 and k=2. The reader is reminded that $\ell_c=\epsilon_c(p-2)/(p-1)$ depends on p. The notation Q_s stands for the quantity Q in units of 10^s times the (cgs) units of Q_s while t_{days} is the observed time in days, and A_s is A in units of 5×10^{11} g cm⁻¹ (Chevalier & Li 2000).

* For PLS E, the emission becomes dominated by the contribution from small radii for k > 23/13. This new regime is described in a separate work (J. Granot & R. Sari 2002. in

TABLE 2 BREAK FREQUENCIES AND CORRESPONDING FLUX DENSITIES

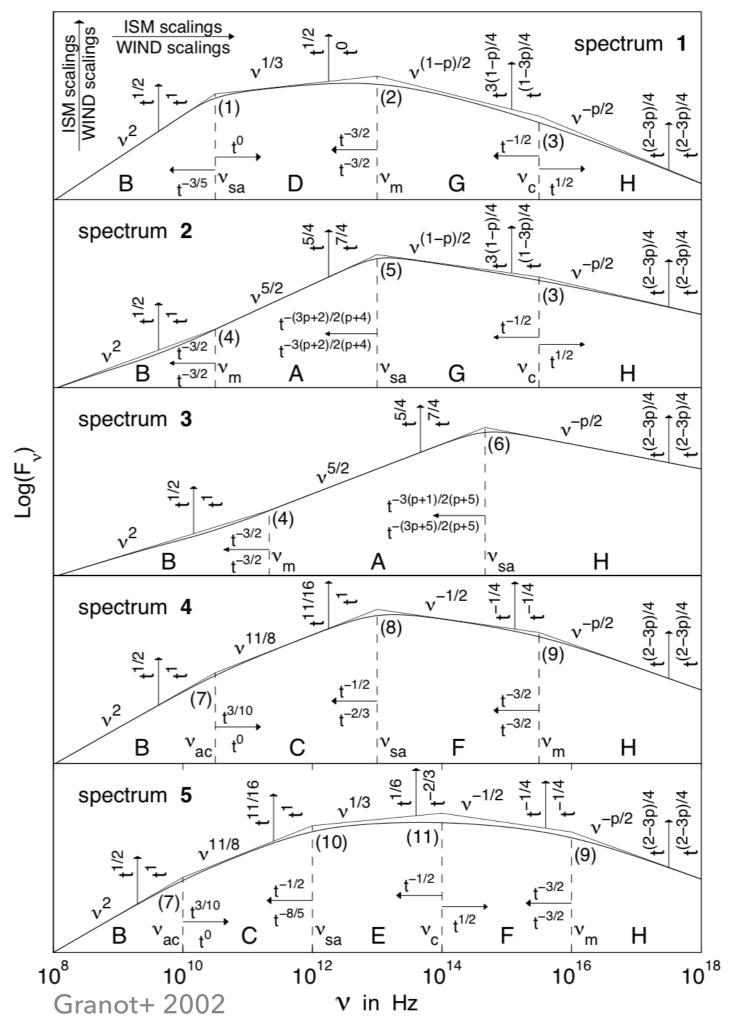
b	β_1	β_2	ν_b	$ u_b(p) $ (Hz)	$F_{v_0, ext}(p)$ (mJy)	s(p)	MRD (%)
	2	$\frac{1}{3}$	$\nu_{\rm sa}$	$1.24 \frac{(p-1)^{3/5}}{(3p+2)^{3/5}} 10^9 (1+z)^{-1} \bar{\epsilon}_e^{-1} \epsilon_B^{1/5} n_0^{3/5} E_{52}^{1/5}$	$0.647 \frac{(p-1)^{6/5}}{(3p-1)(3p+2)^{1/5}} (1+z)^{1/2} \bar{\epsilon_c}^{-1} \epsilon_e^{2/5} n_0^{7/10} E_{52}^{9/10} t_{\rm days}^{1/2} d_{L28}^{-2}$	1.64	6.68
				$8.31 \frac{(p-1)^{3/5}}{(3p+2)^{3/5}} 10^9 (1+z)^{-2/5} \bar{\epsilon}_e^{-1} \epsilon_B^{1/5} A_*^{6/5} E_{52}^{-2/5} r_{\rm days}^{-3/5}$	$9.19 \frac{(p-1)^{6/5}}{(3p-1)(3p+2)^{1/5}} (1+z)^{6/5} \bar{\epsilon}_e^{-1} \epsilon_B^{2/5} A_*^{7/5} E_{52}^{1/5} t_{\rm days}^{-1/5} d_{L28}^{-2}$	1.06	1.02
	$\frac{1}{3}$	$\frac{1-p}{2}$	ν_m	$3.73(p-0.67)10^{15}(1+z)^{1/2}E_{52}^{1/2}\hat{\epsilon}_{6}^{2}\hat{\epsilon}_{8}^{1/2}I_{\rm days}^{-3/2}$	$9.93(p+0.14)(1+z)\epsilon_B^{1/2}n_0^{1/2}E_{52}d_{L23}^{-2}$	1.84-0.40p	5.9
				$4.02(p-0.69)10^{15}(1+z)^{1/2}E_{52}^{1/2}\dot{\epsilon}_{c}^{2}\epsilon_{B}^{1/2}i_{\rm days}^{-3/2}$	$76.9(p+0.12)(1+z)^{3/2}\epsilon_B^{1/2}A_*E_{52}^{1/2}I_{\rm days}^{-1/2}d_{L28}^{-2}$	1.76-0.38p	7.2
	$\frac{1-p}{2}$	$-\frac{p}{2}$	ν_{ϵ}	$6.37(p-0.46)10^{13}e^{-1.16p}(1+z)^{-1/2}\epsilon_B^{-3/2}n_0^{-1}E_{52}^{-1/2}t_{\rm days}^{-1/2}$	$4.68e^{4.82(p-2.5)}10^3(1+z)^{(p+1)/2}\tilde{e}_c^{p-1}e_B^{p-1/2}n_0^{p/2}E_{52}^{(p+1)/2}i_{\rm days}^{(1-p)/2}d_{228}^{-2}$	1.15-0.06p	1.9
				$4.40(3.45-p)10^{10}\epsilon^{0.45p}(1+z)^{-3/2}\epsilon_B^{-3/2}A_*^{-2}E_{52}^{1/2}t_{\rm days}^{1/2}$	$8.02e^{2.02(p-2.5)}10^5(1+z)^{p+1/2}e_r^{p-1}e_B^{p-1/2}\mathcal{A}_s^pE_{52}^{1/2}I_{\rm days}^{1/2-p}d_{L28}^{-2}$	0.80-0.03p	4.4
4	2	$\frac{5}{2}$	ν_m	$5.04(p-1.22)10^{16}(1+z)^{1/2}\tilde{\epsilon}_{e}^{2}\epsilon_{B}^{1/2}E_{52}^{1/2}\epsilon_{\rm days}^{-3/2}$	$3.72(p-1.79)10^{15}(1+z)^{7/2}\bar{\epsilon}_{5}^{4}\epsilon_{B}n_{0}^{-1/2}E_{52}^{3/2}I_{\rm days}^{-5/2}d_{L28}^{-2}$	$3.44p - 1.41^{a}$	0.7a
				$8.08(p-1.22)10^{16}(1+z)^{1/2}\tilde{e}_{e}^{2}\epsilon_{B}^{1/2}E_{52}^{1/2}t_{\rm days}^{-3/2}$	$3.04(p-1.79)10^{15}(1+z)^3\bar{\epsilon}_{c}^{5}\epsilon_{B}A_{a}^{-1}E_{52}^{2}t_{\mathrm{days}}^{-2}d_{128}^{-2}$	$3.63p - 1.60^a$	1.8 ^a
5	. 5/2	$\frac{1-p}{2}$	$\nu_{\rm sz}$	$3.59(4.03 - p)10^9e^{2.34p}\left[\frac{\overline{c}_e^{4(p-1)}e_\beta^{p+2}n_0^4E_{52}^{p+2}}{(1+z)^{6-p}r_{days}^{3p+2}}\right]^{1/2(p+4)}$	$20.8(p-1.53)e^{2.56p}d_{L28}^{-2}\left[\frac{(1+z)^{7p+3}e_B^{2p+3}E_{32}^{1p+7}}{e_e^{10(1-p)}r_{days}^{3(p-1)}}\right]^{1/2(p+4)}$	1.47 - 0.21p	5.9
				$1.58(4.10 - p)10^{10}e^{2.16p} \left[\frac{\tilde{\epsilon}_r^{4(p-1)}\tilde{\epsilon}_B^{p+2}A_a^8}{(1 + z)^{2-p}E_{52}^{2-p}\tilde{\epsilon}_{332}^{4(p+2)}} \right]^{1/2(p+4)}$	$158(p-1.48)e^{2.24p}d_{-28}^{-2}\left[\frac{(1+z)^{6p+9}\epsilon_B^{2p+3}E_{52}^{4p+1}}{\epsilon_e^{10(1-p)}A_s^{2(p-\theta)}f_{days}^{4p+1}}\right]^{1/2(p+4)}$	1.25 - 0.18p	7.2
6	. 5/2	$-\frac{p}{2}$	$\nu_{\rm sz}$	$3.23(p-1.76)10^{12}\begin{bmatrix}\bar{c}_{e}^{A(p-1)}c_{B}^{p-1}n_{0}^{2}E_{52}^{p+1}\\(1+z)^{7-p}c_{033}^{3(p+1)}\end{bmatrix}^{1/2(p+5)}$	$76.9(p-1.08)e^{2.06p}dL_{23}^{-2}\left[\frac{(1+z)^{7p-5}\epsilon_{R}^{2p-5}E_{32}^{3p-5}}{\tilde{\epsilon}_{e}^{P(1-p)}\eta_{0}^{p}\epsilon_{dys}^{2p-1}}\right]^{1/2(p+5)}$	0.94 - 0.14p	12.4
				$4.51(p-1.73)10^{12} \left[\frac{e^{A(p-1)}e_B^{p-1}A_a^AE_{52}^{p-1}}{(1+z)^{5-p}r_{2p+3}^{2p+3}} \right]^{1/2(p+5)}$	$78.6(p-1.12)e^{1.50p}d_{L_{28}^{-2}}^{-2}\left[\frac{(1+z)^{6p+5}e_{B}^{2p-5}E_{32}^{4p+5}}{\overline{\epsilon}_{c}^{10(1-p)}A_{a}^{2p}f_{c}^{4p-5}}\right]^{1/2(p+5)}$	1.04 - 0.16p	11.0
7	2	$\frac{11}{8}$	$\nu_{\rm ac}$	$1.12 \frac{(3p-1)^{8/5}}{(3p+2)^{8/5}} 10^8 (1+z)^{-13/10} \epsilon_e^{-8/5} \epsilon_B^{-2/5} n_0^{3/10} E_{52}^{-1/10} t_{\text{days}}^{3/10}$	$5.27 \frac{(3p-1)^{11/5}}{(3p+2)^{11/5}} 10^{-3} (1+z)^{-1/10} \epsilon_e^{-11/5} \epsilon_B^{-4/5} \eta_0^{1/10} E_{52}^{3/10} t_{days}^{11/10} d_{L28}^{-2}$	1.99 - 0.04p	1.9
				$1.68 \frac{(3p-1)^{8/5}}{(3p+2)^{8/5}} 10^8 (1+z)^{-1} \bar{\epsilon}_e^{-8/5} \epsilon_B^{-2/5} A_u^{3/5} E_{52}^{-2/5}$	$\begin{split} &5.27\frac{(3p-1)^{11/5}}{(3p+2)^{11/5}}10^{-3}(1+z)^{-1/10}\epsilon_e^{-11/5}\epsilon_B^{-4/5}a_0^{1/10}\epsilon_{52}^{3/10}t_{\mathrm{days}}^{11/10}d_{L23}^{-2}\\ &3.76\frac{(3p-1)^{11/5}}{(3p+2)^{11/5}}10^{-3}\epsilon_e^{-11/5}\epsilon_B^{-4/5}A_0^{1/5}E_{52}^{1/5}t_{\mathrm{days}}d_{L23}^{-2} \end{split}$	1.97 - 0.04p	1.9
8	. 11	$-\frac{1}{2}$	$\nu_{\rm sa}$	$1.98\times 10^{11}(1+z)^{-1/2}n_0^{1/6}E_{\rm 52}^{1/6}t_{\rm days}^{-1/2}$	$154(1+z)\epsilon_B^{-1/4}n_0^{-1/12}E_{52}^{2/3}d_{L38}^{-2}$	0.907	1.71
				$3.15 \times 10^{11} (1+z)^{-1/3} A_*^{1/3} t_{\rm days}^{-2/3}$	$119(1+z)^{11/12}\epsilon_B^{-1/4}A_*^{-1/6}E_{52}^{3/4}\iota_{\rm days}^{1/12}d_{\rm C28}^{-2}$	0.893	2.29
9	$-\frac{1}{2}$	$-\frac{p}{2}$	ν_m	$3.94(p-0.74)10^{15}(1+z)^{1/2} \hat{\epsilon}_e^{2} \hat{\epsilon}_B^{1/2} E_{52}^{1/2} t_{\rm days}^{-3/2}$	$0.221(6.27-p)(1+z)^{1/2}\bar{\epsilon}_e^{-1}\epsilon_B^{-1/2}E_{52}^{1/2}t_{\rm days}^{1/2}d_{L28}^{-2}$	3.34 - 0.82p	4.5
				$3.52(p-0.31)10^{15}(1+z)^{1/2}\tilde{\epsilon}_e^2\tilde{\epsilon}_B^{1/2}E_{52}^{1/2}t_{\rm days}^{-3/2}$	$0.165(7.14-p)(1+z)^{1/2}\bar{\epsilon}_{r}^{-1}\epsilon_{B}^{-1/2}E_{52}^{1/2}t_{days}^{1/2}d_{L28}^{-2}$	3.68 - 0.89p	4.2
10.	$\frac{11}{8}$	$\frac{1}{3}$	$\nu_{\rm s_3}$	$1.32\times 10^{10}(1+z)^{-1/2}\epsilon_B^{6/5}n_0^{11/10}E_{52}^{7/10}t_{\rm days}^{-1/2}$	$3.72(1+z)\epsilon_B^{7/5}n_0^{6/5}E_{52}^{7/5}d_{L23}^{-2}$	1.213	5.22
				6	b	ь	b
11.	. 1/3	$-\frac{1}{2}$	ν_{ϵ}	$5.86\times 10^{12}(1+z)^{-1/2}\epsilon_B^{-3/2}n_0^{-1}E_{52}^{-1/2}t_{\rm days}^{-1/2}$	$28.4(1+z)\epsilon_B^{1/2}n_0^{1/2}E_{52}d_{228}^{-2}$	0.597	0.55
				, , b	b	ьь	ь

Note.—First column numbers the breaks. The following two columns are the asymptotic spectral slopes below (β_1) and above (β_2) the break. The fourth column gives the name of the break frequency. The following two columns are $v_b(p)$ and $F_{n_b,r_b}(p)$. The last two columns are the parameter s(p), which determines the shape of each break according to eq. (1) (except for b=4, where it applies to eq. [3]), and the maximal relative difference (MRD) between this analytic formula and our exact numerical results. For each break frequency there are two lines; the first is for an ISM surrounding (k=0) and the second for a stellar wind environment (k=2). The reader is reminded that $\ell_b=\ell_b(p-2)/(p-1)$ depends on p.

For b=4, the values of s(p) and the corresponding MRD refer to eq. (3), and not to eq. (1) as for the other breaks.

The breaks b=10, 11 involve PLS E, where the emission is dominated by the contribution from small radii for k>23/13. This new regime is described in a separate work (J. Granot & R.

Sari 2002, in preparation)



CHOICES

- OK, you have some data you'd like to make inferences about with a model! Which direction should you go?
 - Full simulations: Best you can do is qualitatively identify features in your sim vs. reality, but quantification/parameter determination very difficult.
 - Interpolations: Some parameter inference possible, but limited by the size of your library. May be quite shaky in regions of parameter space far from sim coverage.
 - **Semi-analytic:** Parameter inference very doable, models limited by your creativity. May however not be conducive to certain optimizers/samplers due to lack of accurate derivatives.
 - Analytic: Inference here straightforward, can use all sorts of software to help you optimize/sample it. Rarely directly related to underlying physics.

OTHER CONSIDERATIONS

- Pure Python or Python wrapper for a compiled language?
 - If Python: which libraries to use for the math?
- Support Python 2? Windows?
- Data format of inputs/outputs? (LSST will use AVRO)
- Metrics used to evaluate models? (Information criteria, evidence).
- Particular needs of the optimizer/sampler the model will be passed to?

My principles for making useful modeling software

- Adaptability
 Can I test out new models and model variants quickly?
- VoracityCan it work with a variety of inputs?
- Objectivity

 Are different models compared in a reasonable way?
- Efficiency

 Am I performing my calculations as fast as possible?
- Physicality
 Do my models directly relate to the laws of nature?
- Accessibility
 Can it be used by both "beginners" and "professionals"?

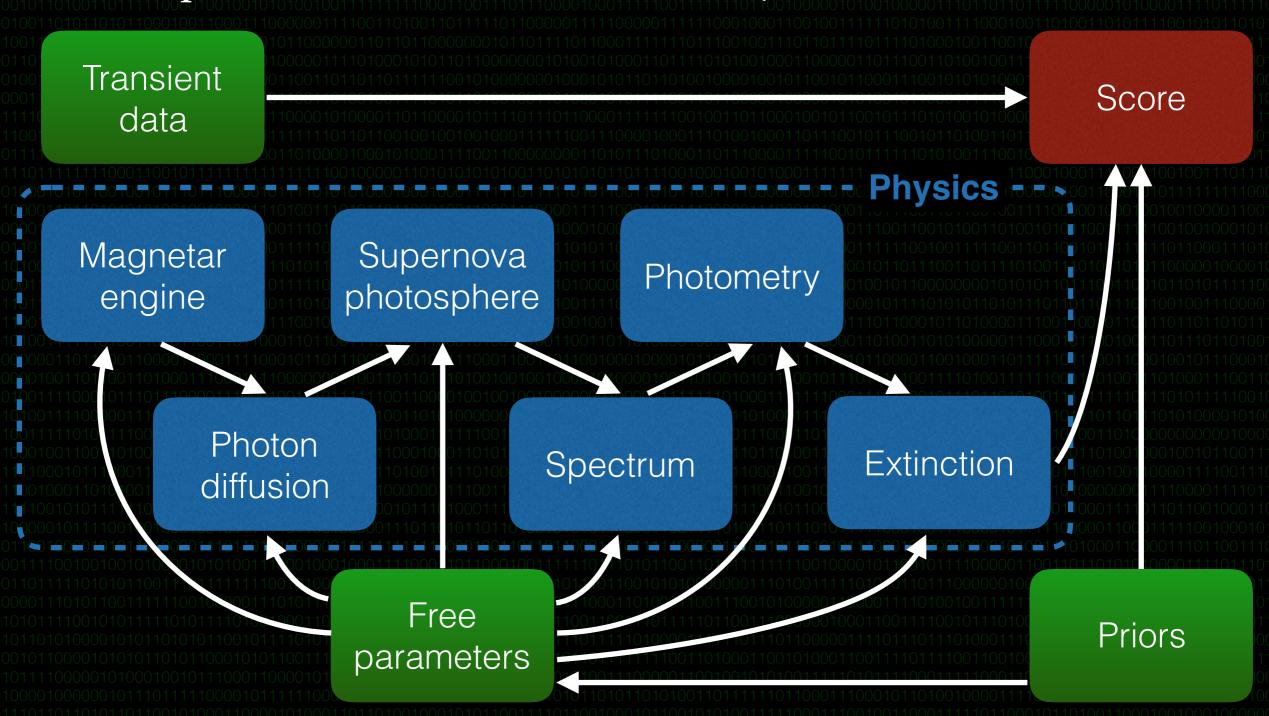
Modular Open-Source
Fitter for Transients
http://mosfit.readthedocs.io

- **MOSFiT** downloads any dataset from one of the open catalogs (or private, user-supplied data) and fits against a user-specified model.
- Many semi-analytical models employ similar physics with slightly different assumptions. This redundancy motivates a **modular** design, which is what MOSFiT implements.
- Can utilize photometry, radio observations, X-ray observations, and spectra when model matching.
- Performs minimization and sampling of the maximum likelihood via a combination of MCMC and global optimization.
- Can be run in parallel and/or on a list of events and/or models.
- Written in Python, available now, paper: 1710.02145. Intended for use by both observers and theorists.

conda install mosfit

Modular model fitting

- Models defined by a tree structure defined in JSON file which specifies how model inputs are related to outputs.
- Example SLSN model (see Nicholl+ 2017):



CHOICES

- OK, you have some data you'd like to make inferences about with a model! Which direction should you go?
 - Full simulations: Best you can do is qualitatively identify features in your sim vs. reality, but quantification/parameter determination very difficult.
- Interpolations: Some parameter inference possible, but limited by the size of your library. May be quite shaky in regions of parameter space far from sim coverage.
- Semi-analytic: Parameter inference very doable, models limited by your creativity. May however not be conducive to certain optimizers/samplers due to lack of accurate derivatives.
- Analytic: Inference here straightforward, can use all sorts of software to help you optimize/sample it. Rarely directly related to underlying physics.

1. MOSFiT users fit data to transient using the built-in or their own custom models.

2. Data is uploaded to GitHub by MOSFiT, where it can be absorbed by the OACs.



MOSFiT User A

JSON Payload



MOSFiT User B

JSON Payload

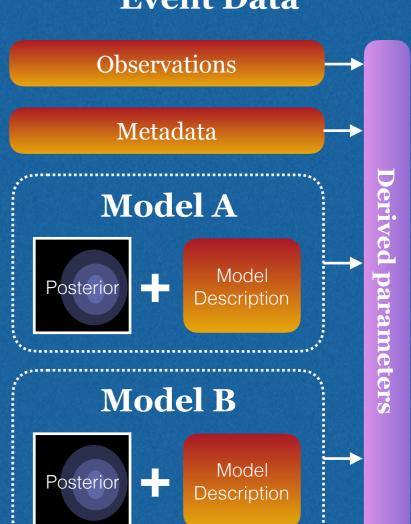


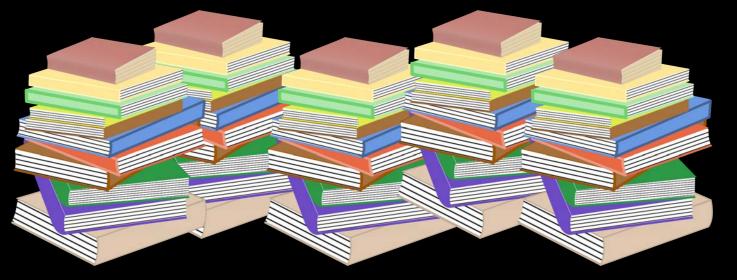
GitHub

3. Data is delivered to an interested user of the catalog. Full information about the models fit against the transient is encapsulated alongside the observed data.

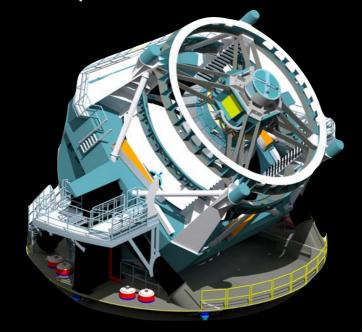
Interested party







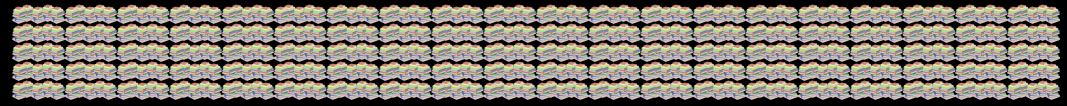
Current supernova data collection could fit in the above books (in 3pt font)



LSST will collect ~10,000 SNe a week



10 years of LSST:



Rare transients (TDEs, SLSN, kilonovae) occur at rates a thousandth (or less) frequently than supernovae. When LSST comes online, these "rare" events will be observed almost daily.

Transient modelers must prepare for the data deluge!