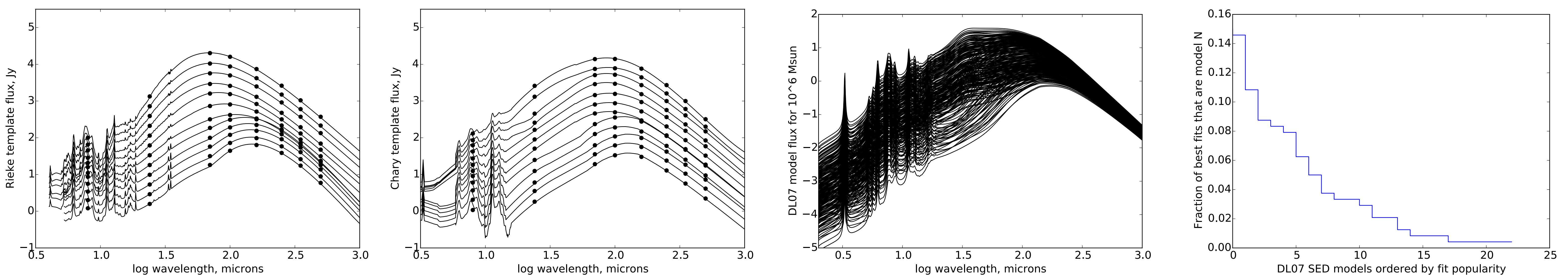


Fitting Galaxy Dust SEDs with Single Models is Biased - and how to fit composite models

Benjamin Weiner, MMT/Steward Observatory

Abstract: A common approach to estimate physical parameters of IR-emitting regions or galaxies is to fit the infrared SED with a family of models of dust emission, one at a time, and either take the best model or marginalize over the fits to estimate dust mass, temperature, and so on. We show that this method can yield biased estimates, because real galaxies are composites, and demonstrate a method that fits a linear combination of a few templates, allowing the method to select the most useful templates from a larger set.



Data - Galaxy Infrared SED families

from Rieke et al (2009) and Chary & Elbaz (2004), widely used templates that are representative of real galaxy SED shapes. We use the templates from $\log L_{IR} = 9.75$ to $12.5 L_{sun}$.

Models - 240 composite dust models

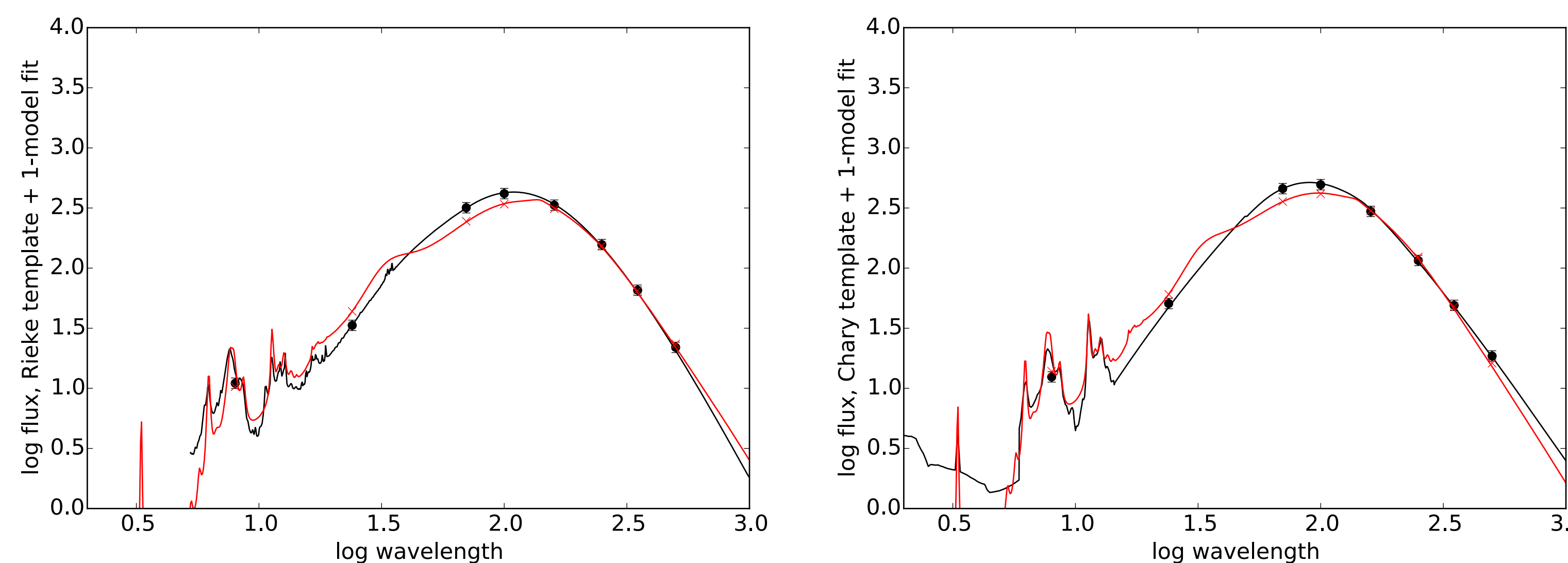
from Draine & Li (2007). A typical literature approach is to fit each of the models to the SED data one at a time, and either take the best model or marginalize over the fit probabilities to find the expectation values of dust mass, temperature, and other physical parameters.

Low Diversity in Model Fits!

When we generate Monte Carlo realizations of the Chary & Elbaz SEDs with 10% errors and fit models, the vast majority of the best fits are using only the ~10 most common models. So although there are 240 models, only ~10 are actually relevant. This is a little scary!

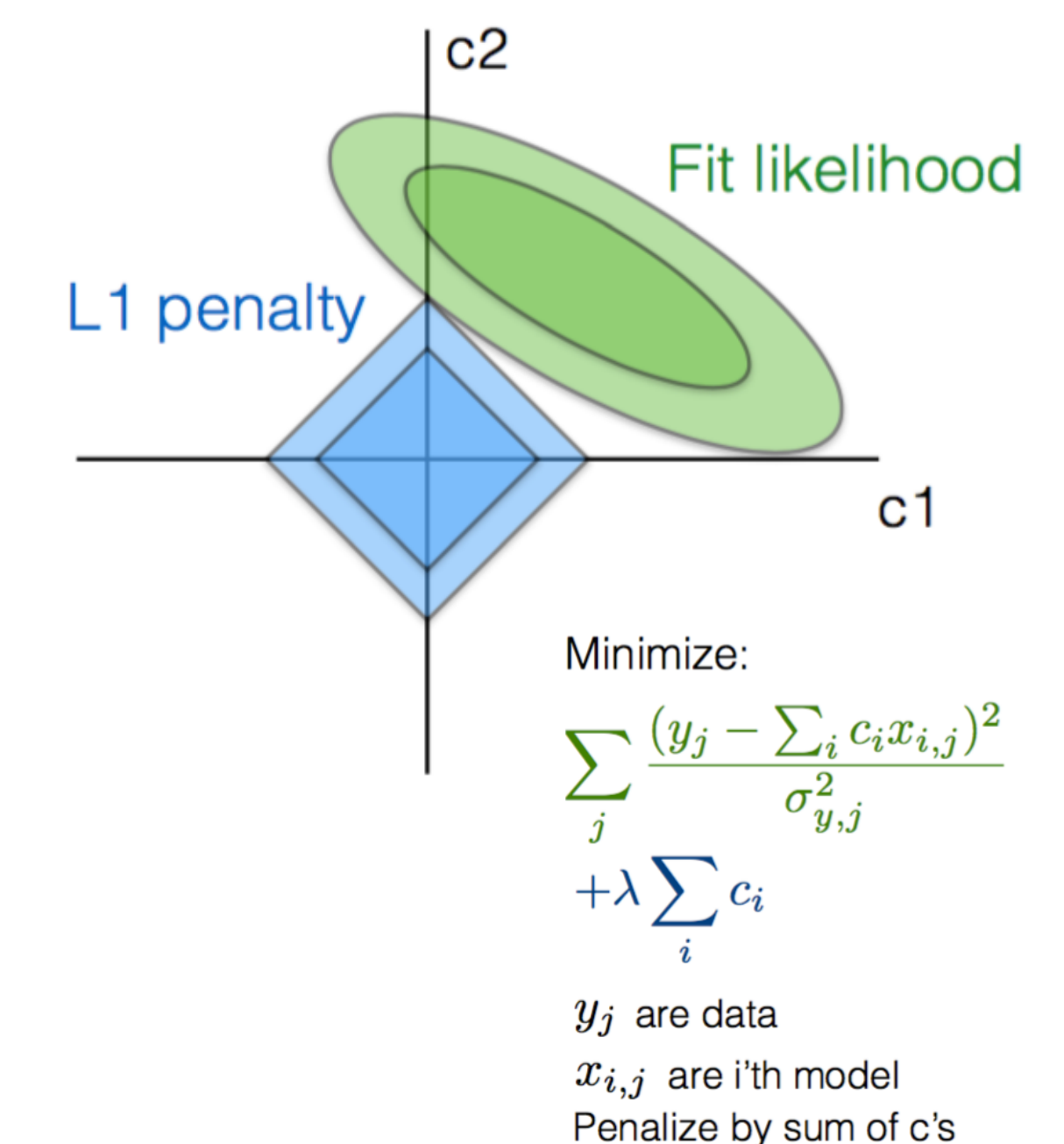
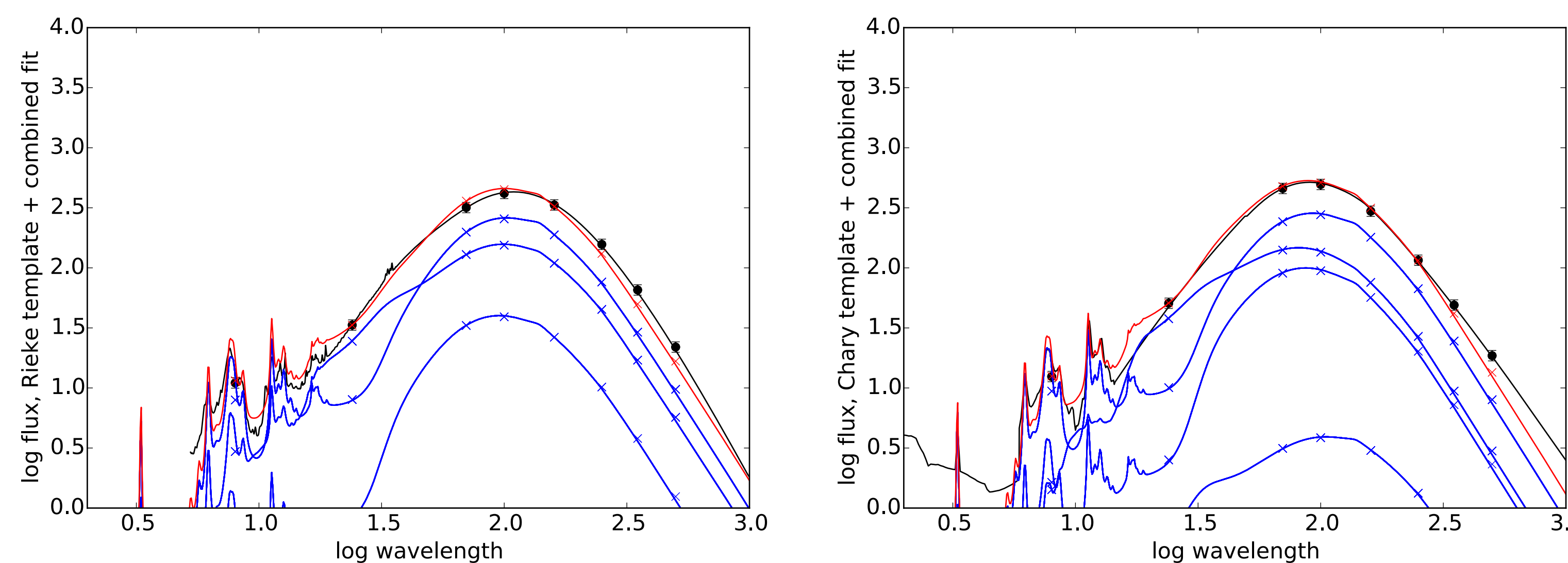
Best Fit Single Models Aren't So Good:

fits to the $L_{IR} = 10^{11}$ SEDs from Rieke et al (left) and Chary & Elbaz (right), with 10% errors. The model fits are tolerable, but statistically rejected (reduced chi-sq $\gg 1$), meaning that errors on the fit and probability of the model fits are not quantitatively correct. Marginalizing over the model probabilities is thus not reliable.



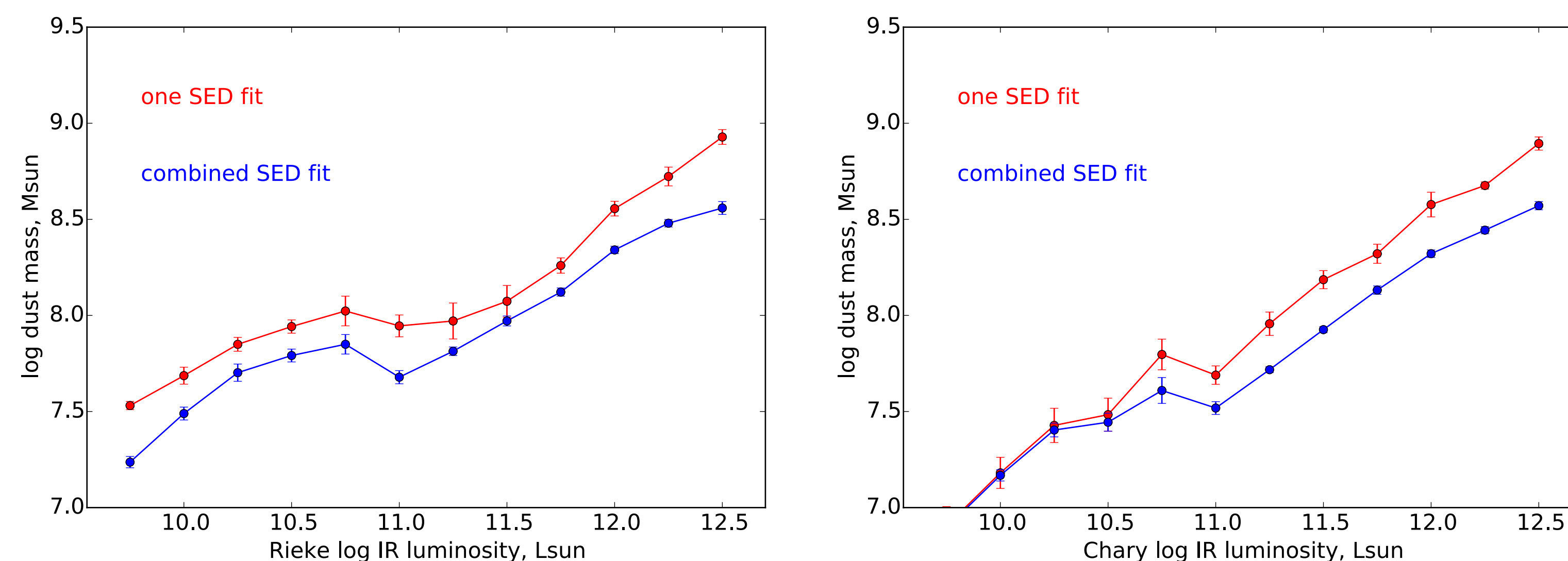
Best Fit "3 of 9" Composite

Models: for the $L_{IR} = 10^{11}$ SEDs from Rieke et al (left) and Chary & Elbaz (right). We feed in the 9 most popular models from the Draine & Li set, and apply a penalized fit to use a linear combination of as few models as possible, so there are only 1, 2 or 3 non-zero components (blue lines). The fits are better, of course, but the more important issue is that the total model for dust emission (red) is not biased by model mismatch.



Conclusion: Single Model Dust

Masses are Biased High. We generate Monte Carlo realizations of the SED families, with 10% errors on each flux, fit single models (red points) and best-3 composite models (blue line), and measure the mean and dispersion of inferred dust mass, for Rieke et al (left) and Chary & Elbaz (right). The main takeaway is that the single model fit masses are biased 0.15-0.2 dex high, 1.5x, perhaps because the composites use less dust mass to produce the emission of both warm and cool components. These biases may be worse with missing data or larger errors, common in real datasets.



How to fit composite

models: We use the "L1 norm" = sum of model coefficients, to regularize the fit by penalizing models with non-zero coefficients. This drives as many models to 0 as possible, and reduces degeneracy. See eg the LASSO method, R. Tibshirani (1996).

We thank George Rieke for discussions on IR SEDs, and David Hogg for suggesting the L1 norm penalty. The code is released at: <http://github.com/bjweiner/sedfitting>