1 Objectives

The objectives for this document are to

- Provide a detailed description of the basic supernova models and assumptions used, and give the probability theory that is represented by the probabilistic graphical model of our first hierarchical SALT2 inference.
- Outline the methods used to carry out this inference.
- Explore similar, but slightly different models, and any variations in the inference methods they require.

2 Basic SALT2 Inference

This is the model we started discussing at "Compute the Universe 2015."

2.1 Observed Values

There are three different quantities that are assumed to be observed values or 'data' in the Bayesian sense.

- The cosmological redshift z_i of the i^{th} supernova.
- The observed flux of the i^{th} supernova in the j^{th} epoch, denoted by $f_{i,j}$. (The bandpass and epoch of each observation are assumed known.)
- The noise in each of the measurements of flux, which is due to noise in the number of photons from both the supernova-source and the sky background, and in the reference image used in the image differencing. In this model, we will assume that this is entirely due to the sky background, which is measured independently, and perfectly. We denote the RMS value of this single, Gaussian noise component by $\sigma^{sky}_{i,j}$ where, as above, i indexes the supernova and j indexes the epoch of observation of the supernova.

2.2 Model Parameters

The SALT2 model can be summarized as follows:

- We have a light curve model for the i^{th} supernova with parameters $\{t_0, x_0, x_1, c, z\}_i$, that provides the predicted flux values that enter into the sampling distribution (likelihood) for the flux values at different times: $P(f_{i,j}|\{t_0, x_0, x_1, c, z\}_i, \{\sigma_{i,j}\})$.
- It is common to parameterize supernovae in terms of magnitudes rather than fluxes, writing,

$$m^\star_B = -2.5\log_{10}(f^\star_B), \quad f^\star_B = \int d\lambda T(\lambda)\frac{dS}{d\lambda}(\lambda, p=0). \tag{1}$$

  $m^\star_B$ is related mostly to x_0, since the contribution from terms involving c and x_1 to this quantity is constrained to be exactly 0 by defining the SALT2 model to give $p = 0$ at the central wavelength of the Bessell B band. In this very good approximation, $m^\star_B = -2.5\log_{10}(x_0) - 2.5\log_{10}(F_0)$, where F_0 is a constant which can be evaluated. Thus, we will treat x_0 and $m^\star_B$ as interchangeable quantities, and redefine for convenience:

$$m^\star_B = -2.5\log_{10}(x_0), \qquad M_i \to M_i + 2.5\log_{10}(F_0) \tag{2}$$

- An individual model supernova is completely specified by the set of parameters $\theta^{SN_{\text{true}}} = \{t_0, m^\star_B, x_1, c, z\}$: these are the parameters that can be inferred from the lightcurve data for a single supernova. Equivalently, to predict the observed fluxes for one supernova, we only need to specify the parameters in $\theta^{SN_{\text{true}}}$. Some may prefer to work entirely in flux units, with $\theta^{SN_{\text{true}}} = \{t_0, x_0, x_1, c, z\}$.

- When considering an ensemble, we make a Tripp ansatz that relates the peak absolute magnitude M_B in the rest frame Bessell B band to the stretch and color parameters x_1, c and a standard candle absolute magnitude M_i:
$$M_{B,i} = M_i + \beta c_i - \alpha x_{1,i}, \tag{3}$$
so that the peak magnitude $m^\star_B$ in Bessell B band calculated from the model is related to the distance modulus $\mu(z, \Omega)$ by
$$\begin{aligned} m^\star_{B,i} &= \mu_i + M_{B,i} &(4)\\ &= \mu_i + M_i + \beta c_i - \alpha x_{1,i}. &(5) \end{aligned}$$
In terms of probability distributions, we can write $P(m^\star_B|z, \Omega, M, \beta, c, \alpha, x_1) = \delta(m^\star_B - \mu - M - \beta c + \alpha x_1)$.

2.3 Conditional PDFs

We model the intrinsic dispersion of supernova properties as follows:

- It is well known that standardization as described above is not complete, and that the distance moduli derived using the above equations have residual scatter. There is some evidence that this scatter is different in different bands, and therefore cannot be completely modelled by a scatter in M_i above, but we shall ignore this for now and model this by assuming that
$$M_i \sim \mathcal{N}(\bar{M}, \sigma_{\rm int}), \tag{6}$$
where $\sigma_{\rm int}$ is the intrinsic dispersion hyper-parameter which we aim to infer.

- Likewise, the stretch and color parameters x_1 and c are taken to be similarly simply distributed, with
$$\begin{aligned} x_{1,i} &\sim \mathcal{N}(\bar{x_1}, \sigma_{x_1}), &(7)\\ c_i &\sim \mathcal{N}(\bar{c}, \sigma_c). &(8) \end{aligned}$$

- The distribution of peak times t_0 is taken to be uniform over the survey window, with no hyper-parameters to be inferred.

2.4 Hyper-parameters

At the top level of inference we have the following hyper-parameters, on top of the ones that parametrize the conditional PDF for the SN properties:

- α and β. These are gradient-like parameters, for which a prior PDF something like $1/(1 + \theta^2)$ might be appropriate.

- Cosmological parameters Ω, with the usual priors.

2.5 Probabilistic Graphical Model

Figure 1 shows probabilistic graphical models for the inference outlined above, for an individual supernova, and then an ensemble.

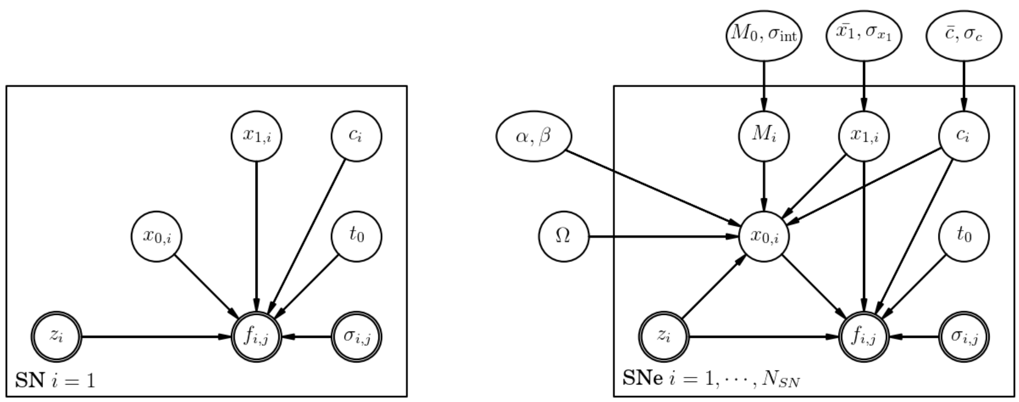

Figure 1: The basic hierarchical SALT2 model for type Ia supernovae. Left: individual objects; right: ensemble.