Near-infrared SN Ia as standard candles

Arturo Avelino

CfA, Harvard

CfA, April 24, 2018
Accelerated expansion of the Universe

Type Ia Supernova (SN Ia)
The problem

Optical samples of SN Ia for cosmology have reached their limit to constrain the nature of the dark energy (DE) because of the systematic uncertainties.

- More optical data *doesn’t* mean better DE constraints.

- **Optical** light is *dimmed* and *rereddened* by *dust* in the host galaxy, the Milky Way, and the extragalactic medium.
A solution: NIR observations!

- Near infrared (NIR) light is much less sensitive to dust than the optical wavelengths.
Low-z NIR sample

Compiled by Andrew Friedman (UCSD):

- CfA, CSP, Literature

- 190 SN Ia with optical + NIR (YJHK) light curves
sn2005el (la)

Observed Phase: MJD - T(B_{max})=53646.3 +/- 0.17 [days]

Magnitude + constant
Goal

Infer the distance modulus (luminosity distance) of each SNIa from their near-infrared time-series data (aka, light curves)

Method

★ Construct NIR light-curve templates

• Gaussian-Processes regression

• Hierarchical Bayesian model

★ Fit the NIR light-curve template to the time series data
Gaussian Processes

Interpolating the time series using Gaussian Processes regression

\[ M_{r,s}(\hat{t}_s) = \text{datum at a given time, for a given supernova “s” and band.} \]
Gaussian Processes

Interpolating the time series using Gaussian Processes regression

\[ M_{r,s}(\hat{t}_s) = \text{datum at a given time, for a given supernova "s" and band.} \]
Gaussian Processes

Interpolating the time series using Gaussian Processes regression

\[ M_{r,s}(\hat{t}_s) = \text{datum at a given time, for a given supernova "s" and band.} \]

\[ M_{r,s}(t^*_s) \mid M_{r,s}(\hat{t}_s), \hat{t}_s, t^*_s \sim N[\bar{M}_{r,s}(t^*_s), \text{cov}(M_{r,s}(t^*_s))] \]
Gaussian Processes

Interpolating the time series using Gaussian Processes regression

\[ M_{r,s}(\hat{t}_s) = \text{datum at a given time, for a given supernova "s" and band.} \]

\[ M_{r,s}(t^*_s) | M_{r,s}(\hat{t}_s), \hat{t}_s, t^*_s \sim N[\bar{M}_{r,s}(t^*_s), \text{cov}(M_{r,s}(t^*_s))] \]
Gaussian Processes

Interpolating the time series using Gaussian Processes regression

\[ M_{r,s}(t_s) = \text{datum at a given time, for a given supernova "s" and band.} \]

\[ M_{r,s}(t^*)|M_{r,s}({\hat{t}_s}), \hat{t}_s, t^*_s \sim N[\bar{M}_{r,s}(t^*_s), \text{cov}(M_{r,s}(t^*_s))] \]

Mean function is computed as:

\[ \bar{M}_{r,s}(t^*_s) = K(t^*_s, \hat{t}_s) \cdot [K(\hat{t}_s, \hat{t}_s) + W(\hat{t}_s, \hat{t}_s)]^{-1} \cdot M_{r,s}(\hat{t}_s) \]

Covariance matrix is computed as:

\[ \text{cov}(M_{r,s}(t^*_s)) = K(t^*_s, t^*_s) - K(t^*_s, \hat{t}_s) \cdot [K(\hat{t}_s, \hat{t}_s) + W(\hat{t}_s, \hat{t}_s) + \sigma_{\mu_{\text{pec},s}}^2 I_s \cdot I_s^\top]^{-1} \cdot K(\hat{t}_s, t^*_s) \]
Gaussian Processes

Kernel: \[ K(t, t') = \sigma_K^2 \exp \left[ -\frac{(t-t')^2}{2l^2} \right] \]

Notation: \[ W(\hat{t}, \hat{t}') = \sigma_M^2 \delta_{tt'} \]
\[ \bar{M}_{r,s}(t_s^*) = \{ \bar{M}_{r,s}(t_s^*) \} \]
\[ M_{r,s}(\hat{t}_s) = \{ M_{r,s}(\hat{t}_s) \} \]

Mean function is computed as:
\[ \bar{M}_{r,s}(t_s^*) = K(t_s^*, \hat{t}_s) \cdot [K(\hat{t}_s, \hat{t}_s) + W(\hat{t}_s, \hat{t}_s)]^{-1} \cdot M_{r,s}(\hat{t}_s) \]

Covariance matrix is computed as:
\[ \text{cov}(M_{r,s}(t_s^*)) = K(t_s^*, t_s^*) - K(t_s^*, \hat{t}_s) \cdot [K(\hat{t}_s, \hat{t}_s) + W(\hat{t}_s, \hat{t}_s) + \sigma_{\mu_{pec,s}}^2 I_s \cdot I_s^\top]^{-1} \cdot K(\hat{t}_s, t_s^*) \]
Gaussian Processes

Kernel:

\[ K(t, t') = \sigma_K^2 \exp \left[ -\frac{(t-t')^2}{2l^2} \right] \]

Hyperparameters computed as:

\[
\ln p(\{M_{r,s}\}|\{\hat{t}_s\}, \sigma_K, l) = \\
- \frac{1}{2} \sum_{s=1}^{N_{SN}} \left\{ M_{r,s}^T(\hat{t}_s) \cdot \left[ K_s(\hat{t}_s, \hat{t}_s) + W_s(\hat{t}_s, \hat{t}_s) + \sigma_{\text{pec},s}^2 I_s \cdot I_s^T \right]^{-1} \cdot M_{r,s}(\hat{t}_s) + \ln \left( \det \left[ K_s(\hat{t}_s, \hat{t}_s) + W_s(\hat{t}_s, \hat{t}_s) + \sigma_{\mu_{\text{pec},s}}^2 I_s \cdot I_s^T \right] \right) + N_{LC,s} \ln 2\pi \right\}, \quad (A6)
\]
Gaussian-Process fit

**Y band**

- $c_{15} = 1.348 \pm 0.005$
- $z_{\text{CMD}} = 0.0145$
- $t = 7.935$
- $\text{sigma}_\text{Y} = 1.00$

**J band**

- $c_{15} = 1.593 \pm 0.007$
- $z_{\text{CMD}} = 0.0146$
- $t = 7.935$
- $\text{sigma}_\text{J} = 1.00$

**H band**

- $c_{15} = 1.393 \pm 0.007$
- $z_{\text{CMD}} = 0.0146$
- $t = 8.335$
- $\text{sigma}_\text{H} = 0.89$

**K band**

- $c_{15} = 1.05 \pm 0.012$
- $z_{\text{CMD}} = 0.096$
- $t = 0.116$
- $\text{sigma}_\text{K} = 0.68$

MJD - $T_{\text{Bmax}}$ / $(1 + z_{\text{rel}})$
Templates
Hierarchical Bayesian model
Y band template ($z > 0$)

Phase = $\frac{(MJD - T_{Bmax})}{(1 + z_{rel})}$
Y band template ($z > 0$) (GP std dev)

Phase = (MJD - $T_{Brace}$)/($1+z_{rel}$)

Absolute magnitude
Y band template ($z > 0$)

Phase = ($MJD - T_{Bmax})/(1+z_{rel})$

- Absolute magnitude

- Gaussian Process mean (GP)
- Standard deviation
- 68.3% confidence interval GP
Bayesian Hierarchical model

Constructing the NIR light-curve templates

1st level of the hierarchy:

\( \tilde{M}_s \sim N(\tilde{M}_s, \sigma^2_{\tilde{M},s}) \)
Bayesian Hierarchical model

Constructing the NIR light-curve templates

1st level of the hierarchy:

\[ \tilde{M}_s \sim N(\tilde{M}_s, \sigma^2_{\tilde{M},s}) \]

2nd level of the hierarchy:

We assume that the \( \tilde{M}_s(t_*) \) are drawn from a Gaussian distribution with mean \( \mathcal{M}(t_*) \) and variance \( \sigma^2_{\mathcal{M}} \):

\[
p \left( \{\tilde{M}_s\} | \mathcal{M}, \sigma_{\mathcal{M}} \right) = \prod_{s=1}^{N_{\text{SN}}} N \left( \tilde{M}_s | \mathcal{M}, \sigma^2_{\mathcal{M}} \right)
\]
Hierarchical Bayesian model

\[ p \left( \{ \tilde{M}_s \} | \mathcal{M}, \sigma_\mathcal{M} \right) = \prod_{s=1}^{N_{SN}} N \left( \tilde{M}_s | \mathcal{M}, \sigma^2_\mathcal{M} \right) \]
Arturo Avelino, "Near-infrared SN Ia as standard candles"
Arturo Avelino, "Near-infrared SN Ia as standard candles"
Distance modulus

\[ \Delta m_s(t) \equiv m_s(t) - M(t) - m_{0,s} \]  \hspace{1cm} (11)

where \( m_s(t) \) and \( M(t) \) are the apparent magnitude and the magnitude of the normalized template at phase \( t \), respectively. We can express this difference for all the \( N_{LC,s} \) phases in a given LC as the vector,

\[ \Delta m_s \equiv \begin{pmatrix} \Delta m_s(t_1) \\ \Delta m_s(t_2) \\ \vdots \\ \Delta m_s(t_{N_{LC,s}}) \end{pmatrix} \]  \hspace{1cm} (12)

Then, to determine \( m_{0,s} \) we minimize the negative of the log likelihood function \( L(m_{0,s}) \) defined as

\[ -2 \ln L(m_{0,s}) = \Delta m_s^T \cdot C^{-1} \cdot \Delta m_s \]  \hspace{1cm} (13)
Distance modulus

$$\Delta m_s(\hat{t}) \equiv m_s(\hat{t}) - M(\hat{t}) - m_{0,s}$$ \hspace{1cm} (11)

where $m_s(\hat{t})$ and $M(\hat{t})$ are the apparent magnitude and the magnitude of the normalized template at phase $\hat{t}$, respectively. We can express this difference for all the $N_{\text{LC},s}$ phases in a given LC as the vector,

$$\Delta m_s \equiv \begin{pmatrix} \Delta m_s(\hat{t}_1) \\ \Delta m_s(\hat{t}_2) \\ \vdots \\ \Delta m_s(\hat{t}_{N_{\text{LC},s}}) \end{pmatrix}.$$ \hspace{1cm} (12)

Then, to determine $m_{0,s}$ we minimize the negative of the log likelihood function $L(m_{0,s})$ defined as

$$-2 \ln L(m_{0,s}) = \Delta m_s^\top \cdot C^{-1} \cdot \Delta m_s$$ \hspace{1cm} (13)

where $C$ is the $N_{\text{LC},s}$-dimensional covariance matrix where the $(\hat{t}_i, \hat{t}_j)$ component is given by:

$$C_{ij} \equiv \text{Cov}(\Delta m_s(\hat{t}_i), \Delta m_s(\hat{t}_j))$$ \hspace{1cm} (14)

$$= \sigma_{M}(\hat{t}_i) \sigma_{M}(\hat{t}_j) \exp \left[ - \frac{(\hat{t}_i - \hat{t}_j)^2}{2l^2} \right] + \hat{\sigma}_{m,s}^2(\hat{t}_i) \delta_{ij} \hspace{1cm} (15)$$

where $\sigma_{M}(\hat{t})$ is the population standard deviation of the sample distribution of magnitudes at time $\hat{t}$, determined from Eq. (B2) during the training process used to construct the mean LC template, with the hyperparameter $l$ computed via Eq. (A6), while $\hat{\sigma}_{m,s}^2(\hat{t}_i)$ is the photometric error of the datum $m_s(\hat{t}_i)$. 

Distance modulus

\[ \Delta m_s(\hat{t}) \equiv m_s(\hat{t}) - M(\hat{t}) - m_{0,s} \]  

(11)

where \( m_s(\hat{t}) \) and \( M(\hat{t}) \) are the apparent magnitude and the magnitude of the normalized template at phase \( \hat{t} \), respectively. We can express this difference for all the \( N_{LC,s} \) phases in a given LC as the vector,

\[ \Delta m_s \equiv \left( \begin{array}{c} \Delta m_s(\hat{t}_1) \\ \Delta m_s(\hat{t}_2) \\ \vdots \\ \Delta m_s(\hat{t}_{N_{LC,s}}) \end{array} \right). \]  

(12)

Then, to determine \( m_{0,s} \) we minimize the negative of the log likelihood function \( L(m_{0,s}) \) defined as

\[ -2 \ln L(m_{0,s}) = \Delta m_s^\top \cdot C^{-1} \cdot \Delta m_s \]  

(13)

where \( C \) is the \( N_{LC,s} \)-dimensional covariance matrix where the \( (i, j) \) component is given by:

\[ C_{ij} \equiv \text{Cov}(\Delta m_s(\hat{t}_i), \Delta m_s(\hat{t}_j)) \]  

\[ = \sigma_M(\hat{t}_i) \sigma_M(\hat{t}_j) \exp \left[ -\frac{(\hat{t}_i - \hat{t}_j)^2}{2\ell^2} \right] + \delta_{m,s}(\hat{t}_i) \delta_{ij} \]  

\[ \hat{\sigma}^2_{m,s}(\hat{t}_i) \delta_{ij} \]  

(14)

(15)

From Eq. (13), we can calculate an analytic expression for the maximum likelihood estimator (MLE) of the apparent magnitude at \( B \)-band maximum light, \( \hat{m}_{0,s} \), given by:

\[ \hat{m}_{0,s} = \left[ \sum_{i}^{N_{LC,s}} \left( C^{-1} \right)_{ij} \right]^{-1} \times \left[ \sum_{i}^{N_{LC,s}} \left( m_s(\hat{t}_i) - M(\hat{t}_i) \right) \sum_{j}^{N_{LC,s}} \left( C^{-1} \right)_{ij} \right], \]  

(16)

with the MLE of the uncertainty of \( \hat{m}_{0,s} \) given as

\[ \sigma_{0,s} = \left[ \sum_{i}^{N_{LC,s}} \left( C^{-1} \right)_{ij} \right]^{-1/2} \]  

(17)
Distance modulus

\[ \Delta m_s(t) \equiv m_s(t) - M(t) - m_{0,s} \quad (11) \]

where \( m_s(t) \) and \( M(t) \) are the apparent magnitude and the magnitude of the normalized template at phase \( t \), respectively. We can express this difference for all the \( N_{LC,s} \) phases in a given LC as the vector,

\[ \Delta m_s \equiv \begin{pmatrix} \Delta m_s(t_1) \\ \Delta m_s(t_2) \\ \vdots \\ \Delta m_s(t_{N_{LC,s}}) \end{pmatrix}. \quad (12) \]

Then, to determine \( m_{0,s} \) we minimize the negative of the log likelihood function \( L(m_{0,s}) \) defined as

\[ -2 \ln L(m_{0,s}) = \Delta m_s^\top \cdot C^{-1} \cdot \Delta m_s \]

\[ \mu_s = \hat{m}_{0,s} - \langle M_0 \rangle \]

with uncertainty given as

\[ \sigma_{\mu,s} = \sqrt{\sigma^2_{0,s} + \sigma^2_{\text{int}}}. \quad (20) \]

where \( C \) is the \( N_{LC,s} \)-dimensional covariance matrix where the \((i, j)\) component is given by:

\[ C_{ij} \equiv \text{Cov} \left( \Delta m_s(i), \Delta m_s(j) \right) \]

\[ = \sigma_M(t_i) \sigma_M(t_j) \exp \left[ -\frac{(t_i - t_j)^2}{2l^2} \right] + \]

From Eq. (13), we can calculate an analytic expression for the maximum likelihood estimator (MLE) of the apparent magnitude at B-band maximum light, \( \hat{m}_{0,s} \), given by:

\[ \hat{m}_{0,s} = \left[ \sum_{i,j} (C^{-1})_{ij} \right]^{-1} \times \sum_{i} \left( m_s(i) - M(t_i) \right) \sum_{j} (C^{-1})_{ij} \]

\[ \text{with the MLE of the uncertainty of} \hat{m}_{0,s} \text{given as} \]

\[ \sigma_{0,s} = \left[ \sum_{i,j} (C^{-1})_{ij} \right]^{-1/2}. \quad (17) \]
Distance modulus

\[ \Delta m_s(\hat{t}) \equiv m_s(\hat{t}) - M(\hat{t}) - m_{0,s} \]  

where \( m_s(\hat{t}) \) and \( M(\hat{t}) \) are the apparent magnitude and the magnitude of the normalized template at phase \( \hat{t} \), respectively. We can express this difference for all the \( N_{LC,s} \) phases in a given LC as the vector,

\[ \Delta m_s \equiv \begin{pmatrix} \Delta m_s(\hat{t}_1) \\ \Delta m_s(\hat{t}_2) \\ \vdots \\ \Delta m_s(\hat{t}_{N_{LC,s}}) \end{pmatrix}, \]

Then, to determine \( m_{0,s} \) we minimize the negative of the log likelihood function \( L(m_{0,s}) \) defined as

\[ -2 \ln L(m_{0,s}) = \Delta m_s^\top \cdot C^{-1} \cdot \Delta m_s \]

where \( C \) is the \( N_{LC,s} \)-dimensional covariance matrix where the \((\hat{t}_i, \hat{t}_j)\) component is given by:

\[ C_{ij} \equiv \text{Cov} \left( \Delta m_s(\hat{t}_i), \Delta m_s(\hat{t}_j) \right) = \sigma_M(\hat{t}_i) \sigma_M(\hat{t}_j) \exp \left[ -\frac{(\hat{t}_i - \hat{t}_j)^2}{2l^2} \right] \]

From Eq. (13), we can calculate an analytic expression for the maximum likelihood estimator (MLE) of the apparent magnitude at B-band maximum light, \( \hat{m}_{0,s} \), given by:

\[ \hat{m}_{0,s} = \left[ \sum_{i,j} (C^{-1})_{ij} \right]^{-1} \times \sum_i N_{LC,s} \left( m_s(\hat{t}_i) - M(\hat{t}_i) \right) \sum_j (C^{-1})_{ij}, \]

with the MLE of the uncertainty of \( \hat{m}_{0,s} \) given as

\[ \sigma_{0,s} = \left[ \sum_{i,j} (C^{-1})_{ij} \right]^{-1/2}. \]
Intrinsic dispersion

Scatter in the Hubble residuals after accounting for peculiar-velocity and photometric uncertainties.

Intrinsic dispersion $\sigma_{\text{int}}$:

$$-2 \ln L(\sigma_{\text{int}}^2) = \sum_{s}^{N_{\text{SN}}} \left[ \ln \left( \sigma_{0,s}^2 + \sigma_{\text{int}}^2 + \sigma_{\mu_{\text{pec}},s}^2 \right) + \frac{\delta \mu_s^2}{\sigma_{0,s}^2 + \sigma_{\text{int}}^2 + \sigma_{\mu_{\text{pec}},s}^2} \right]$$

Blondin, Mandel, Kirshner, 2011
Arturo Avelino, "Near-infrared SN Ia as standard candles"
Gaussian-Process method

Arturo Avelino, "Near-infrared SN Ia as standard candles"
Gaussian-Process Method

Y band

\[ c_{15} = 1.346 \pm 0.005 \]
\[ z_{\text{CMB}} = 0.0145 \]
\[ i = 7.535 \]
\[ \text{sigma}_{\text{kern}} = 1.003 \]

J band

\[ c_{15} = 1.33 \pm 0.007 \]
\[ z_{\text{CMB}} = 0.0146 \]
\[ i = 8.338 \]
\[ \text{sigma}_{\text{kern}} = 0.969 \]

H band

\[ c_{15} = 1.32 \pm 0.012 \]
\[ z_{\text{CMB}} = 0.0146 \]
\[ i = 8.116 \]
\[ \text{sigma}_{\text{kern}} = 0.882 \]

K band

\[ c_{15} = 1.33 \pm 0.007 \]
\[ z_{\text{CMB}} = 0.0145 \]
\[ i = 7.535 \]
\[ \text{sigma}_{\text{kern}} = 1.003 \]
Gaussian-Process Method

sn2005el

Y band

J band

K band

H band
Arturo Avelino, "Near-infrared SN Ia as standard candles"
Combining multiple NIR bands
4.3. Distance modulus from the combined NIR bands

From the distance moduli \( \mu_s^Y, \mu_s^J, \mu_s^H, \mu_s^K \) for a given supernova \( s \) determined from each NIR band following either of the two methods described above, we determine the "total" distance modulus \( \hat{\mu}_s \) in each method. First we define the vector of residuals

\[
\delta \mu_s \equiv \begin{pmatrix}
\mu_s^Y - \hat{\mu}_s \\
\mu_s^J - \hat{\mu}_s \\
\mu_s^H - \hat{\mu}_s \\
\mu_s^K - \hat{\mu}_s
\end{pmatrix}.
\]

(25)

where \( \mu_s^Z \) is given by either Eq. (19) or (23). Then, to determine \( \hat{\mu}_s \) we minimize the negative of the likelihood function \( L(\hat{\mu}_s) \) defined as

\[
-2 \ln L(\hat{\mu}_s) = \delta \mu_s^\top \cdot C_\mu^{-1} \cdot \delta \mu_s
\]

(26)
Arturo Avelino, "Near-infrared SN Ia as standard candles"
Arturo Avelino, "Near-infrared SN Ia as standard candles"
How good or bad are these results?
Optical Hubble diagram
Fitting the *optical* light curves only

Arturo Avelino, "Near-infrared SN Ia as standard candles"
Arturo Avelino, "Near-infrared SN Ia as standard candles"
### Intrinsic dispersion and wRMS summary

<table>
<thead>
<tr>
<th>Band</th>
<th>Method</th>
<th>$\sigma_{\text{int}}$</th>
<th>wRMS (mag)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y$</td>
<td>Template</td>
<td>0.095 ± 0.018</td>
<td>0.129</td>
</tr>
<tr>
<td>$Y$</td>
<td>GP</td>
<td>0.091 ± 0.020</td>
<td>0.125</td>
</tr>
<tr>
<td>$J$</td>
<td>Template</td>
<td>0.118 ± 0.015</td>
<td>0.156</td>
</tr>
<tr>
<td>$J$</td>
<td>GP</td>
<td>0.099 ± 0.017</td>
<td>0.137</td>
</tr>
<tr>
<td>$H$</td>
<td>Template</td>
<td>0.061 ± 0.015</td>
<td>0.113</td>
</tr>
<tr>
<td>$H$</td>
<td>GP</td>
<td>0.057 ± 0.022</td>
<td>0.117</td>
</tr>
<tr>
<td>$K_s$</td>
<td>Template</td>
<td>0.138 ± 0.028</td>
<td>0.180</td>
</tr>
<tr>
<td>$K_s$</td>
<td>GP</td>
<td>0.096 ± 0.056</td>
<td>0.170</td>
</tr>
<tr>
<td>any $YJHK_s$</td>
<td>Template</td>
<td>0.089 ± 0.012</td>
<td>0.123</td>
</tr>
<tr>
<td>any $YJHK_s$</td>
<td>GP</td>
<td>0.081 ± 0.015</td>
<td>0.118</td>
</tr>
<tr>
<td>$YJH$</td>
<td>Template</td>
<td>0.055 ± 0.015</td>
<td>0.097</td>
</tr>
<tr>
<td>$YJH$</td>
<td>GP</td>
<td>0.061 ± 0.025</td>
<td>0.105</td>
</tr>
<tr>
<td>$JHK_s$</td>
<td>Template</td>
<td>0.089 ± 0.026</td>
<td>0.134</td>
</tr>
<tr>
<td>$JHK_s$</td>
<td>GP</td>
<td>0.098 ± 0.047</td>
<td>0.149</td>
</tr>
<tr>
<td>Optical</td>
<td>SALT2</td>
<td>0.154 ± 0.023</td>
<td>0.216</td>
</tr>
<tr>
<td>Optical</td>
<td>SNooPy</td>
<td>0.140 ± 0.020</td>
<td>0.146</td>
</tr>
</tbody>
</table>
RAISIN = SN IA in the IR
Tracing cosmic expansion with SN Ia in the Near Infrared

RAISIN-1
• 23 SN Ia, redshift ~ 0.3

RAISIN-2
• 24 SN Ia, redshift ~ 0.5

Arturo Avelino, "Near-infrared SN Ia as standard candles"
Take away

• NIR SN Ia are very good standard candles compared with optical observations.

• Very promising for cosmology when combining optical+NIR observations: RAISIN program, WFIRST.