

Astrostatistics: Wed 20 Feb 2019

<https://github.com/CambridgeAstroStat/PartIII-Astrostatistics-2019>

- Example Sheet online, Ex Class Tue Feb 19, 1pm MR5
- Today: continue Bayesian computation / Monte Carlo Methods / MCMC
- MacKay: Ch 29-30; Bishop: Ch 11; Gelman
- Givens & Hoeting “Computational Statistics”
(Free download through Cambridge Library iDiscover)
- Hogg & DFM, 2017 “Data analysis recipes: Using Markov Chain Monte Carlo.” <https://arxiv.org/abs/1710.06068>

Markov Chain Monte Carlo (MCMC)

Mapping the Posterior $P(\theta | D)$

- Markov Chain Monte Carlo (MCMC)
- Now:
 - Drawing Multivariate Gaussian random variables
 - N-D Metropolis Algorithm
 - Rules of thumb for proposal scale
 - assessing convergence (G-R Ratio)
 - Metropolis-Hastings algorithm
 - Gibbs sampling

d-dim Metropolis Algorithm:

Posterior $P(\theta | D)$,

Symmetric Proposal/Jump dist'n $J(\theta^* | \theta) = J(\theta | \theta^*)$

1. Choose a random starting point θ_0
2. At step $i = 1 \dots N$, propose a new parameter value $\theta^* \sim N(\theta_{i-1}, \Sigma_p)$.
The proposal distr'n is $J(\theta^* | \theta_{i-1}) = N(\theta^* | \theta_{i-1}, \Sigma_p)$
3. Evaluate ratio of posteriors at proposed vs current values. $r = P(\theta^* | \mathbf{y}) / P(\theta_{i-1} | \mathbf{y})$.
4. Accept θ^* with probability $\min(r, 1)$: $\theta_i = \theta^*$. If not accept, stay at same value $\theta_i = \theta_{i-1}$ for the next step & include in chain.
5. Repeat steps 2-4 until reach some measure of convergence (G-R) and gather enough independent samples to compute your inference (reduce Monte Carlo error)

Tuning d-dim Metropolis

- $\theta^* \sim N(\theta_i, \Sigma_p)$: if proposal scale Σ_p is too large, will get too many rejections and not go anywhere. If proposal scale too small, you will accept very many small moves: inefficient random walk
- Laplace Approximation:
$$P(\theta|D) \approx N(\theta|\hat{\theta}, \Sigma)$$

$$\hat{\theta} = \text{posterior mode}$$
$$(\Sigma^{-1})_{ij} = \frac{\partial^2 \log P(\theta|D)}{\partial \theta_i \partial \theta_j} \Big|_{\hat{\theta}}$$
- Choose $\Sigma_p = c^2 \Sigma$:
$$c \approx 2.4/\sqrt{d}$$
- Scale Proposal to aim for an acceptance ratio of 44% in 1D, 23% in $d > 5$

Astrostatistics Case Study: Supernova Cosmology

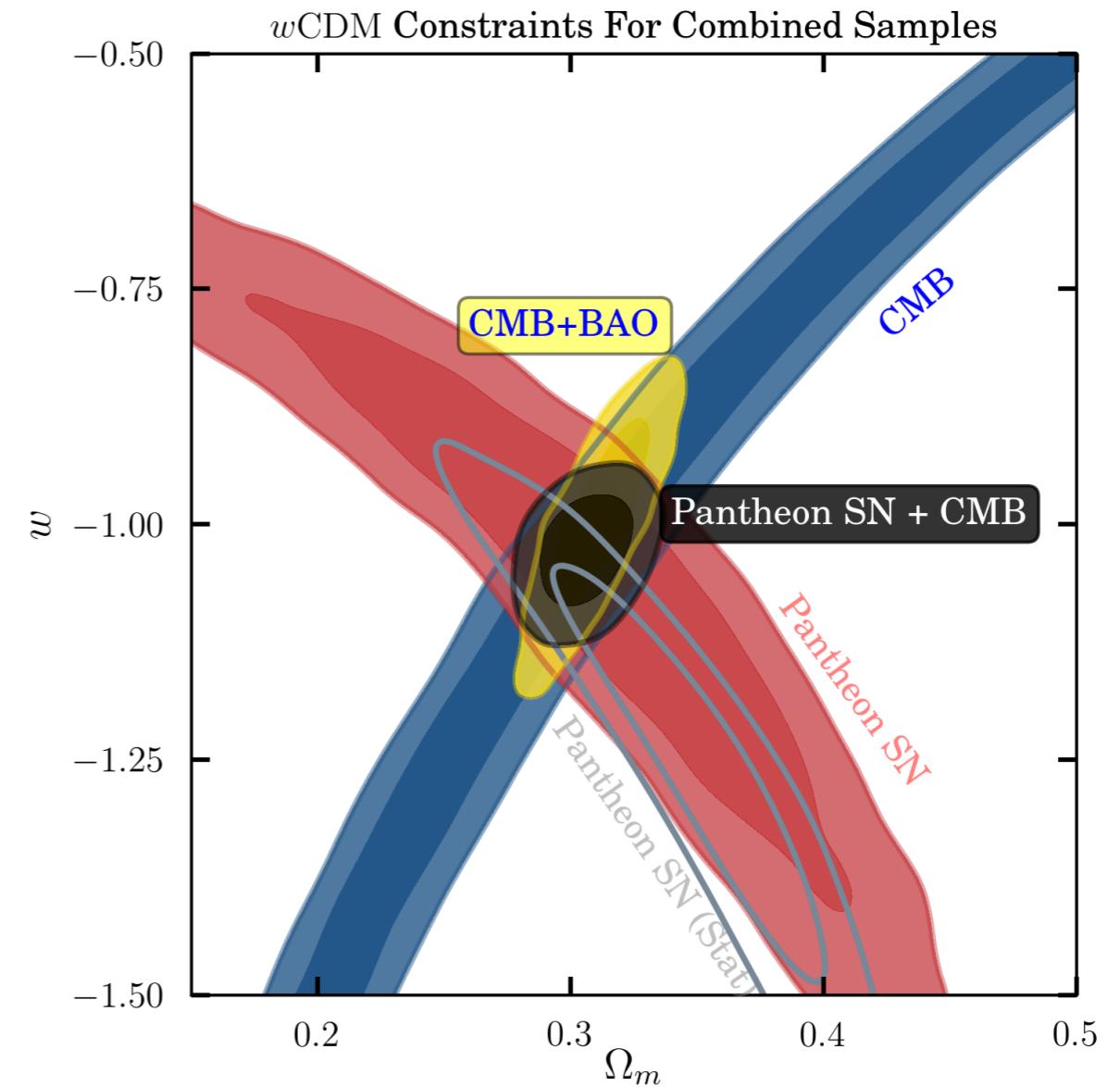
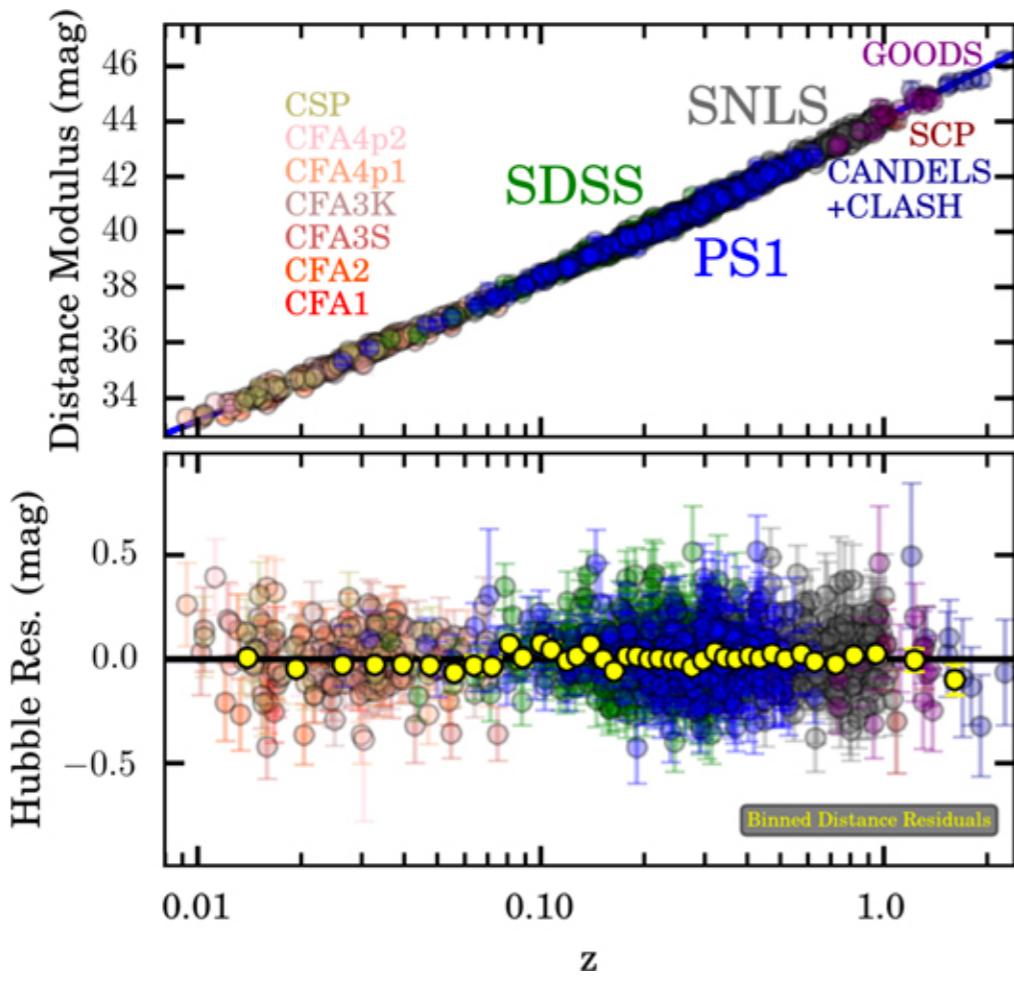
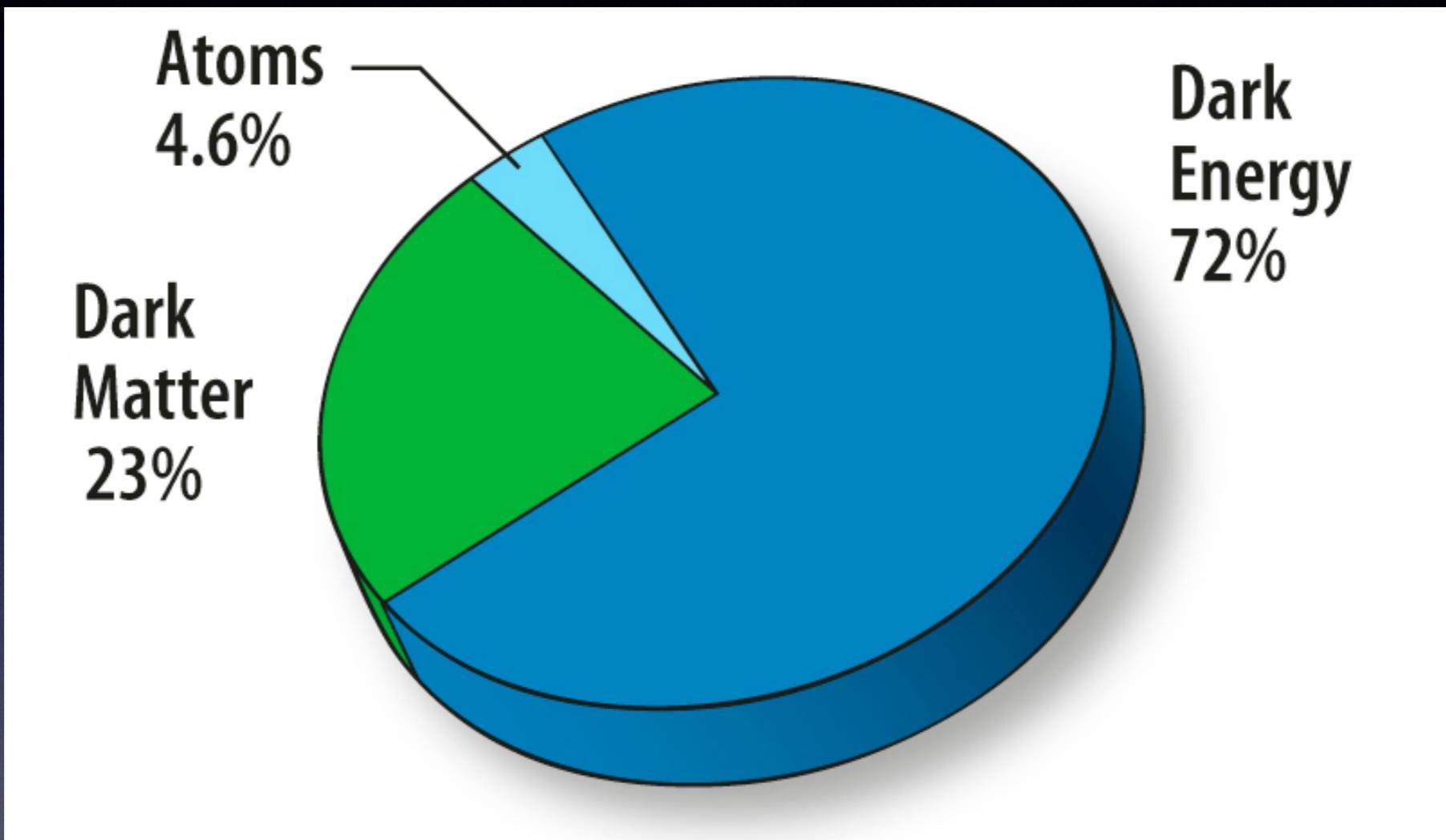


Figure 20. Confidence contours at 68% and 95% for the Ω_m and w cosmological parameters for the w CDM model. Constraints from CMB (blue), SN - with systematic uncertainties (red), SN - with only statistical uncertainties (gray-line), and SN+CMB (purple) are shown.

Cosmological Energy Content



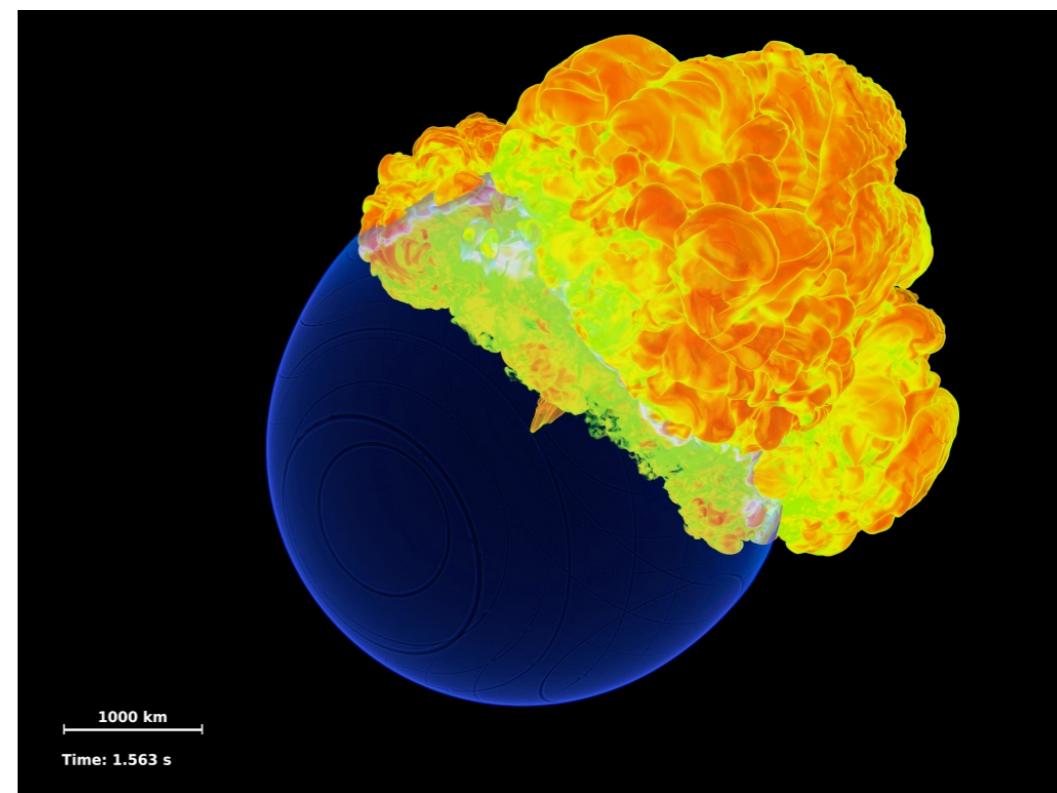
What is Dark Energy?

Dark Energy Equation of state $P = w\rho$

Is $w + l = 0$? (Cosmological Constant: $w = -l$)

Type Ia Supernovae (SN Ia) are Almost Standard Candles

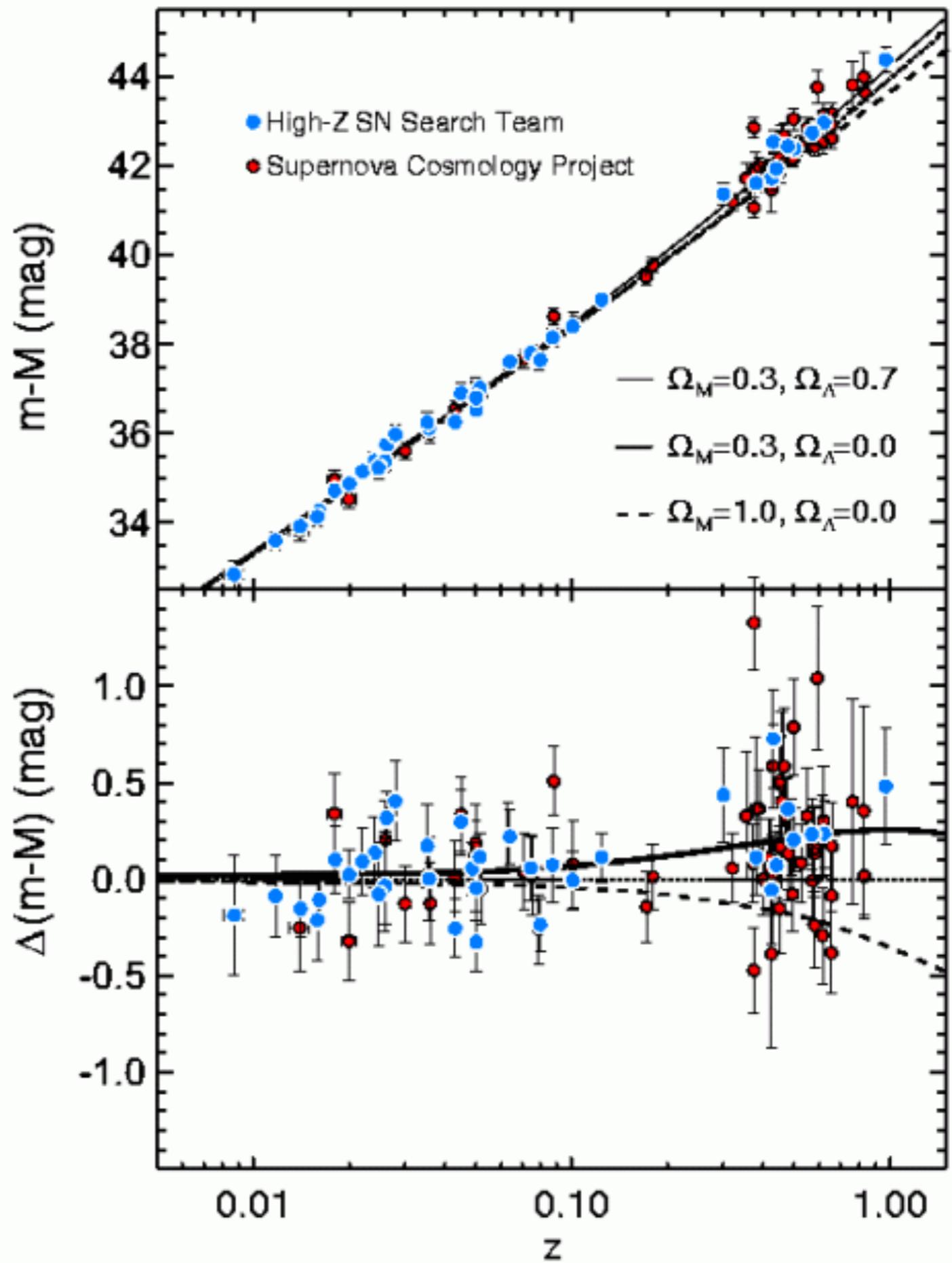
- Progenitor: C/O White Dwarf Star accreting mass leads to instability
- Thermonuclear Explosion: Deflagration/Detonation
- Nickel to Cobalt to Iron Decay + radiative transfer powers the light curve
- General Idea, but Theoretical Astrophysics simulations cannot quantitatively reproduce realistic observations (use empirical models)



Credit: FLASH Center

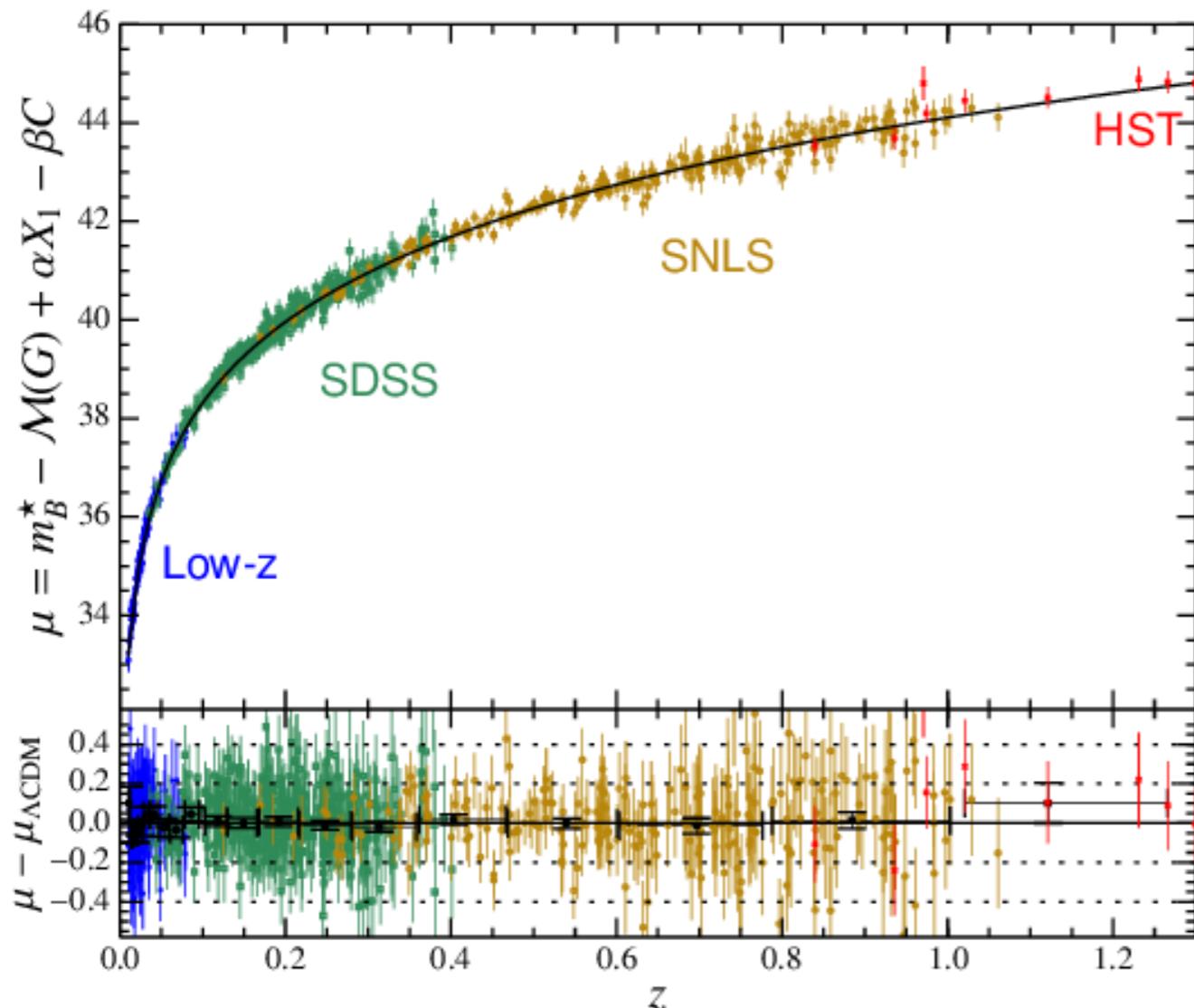
SN Ia Hubble Diagram (Distance Moduli vs. z):

The Universe is
accelerating
($\Omega_\Lambda > 0$)!



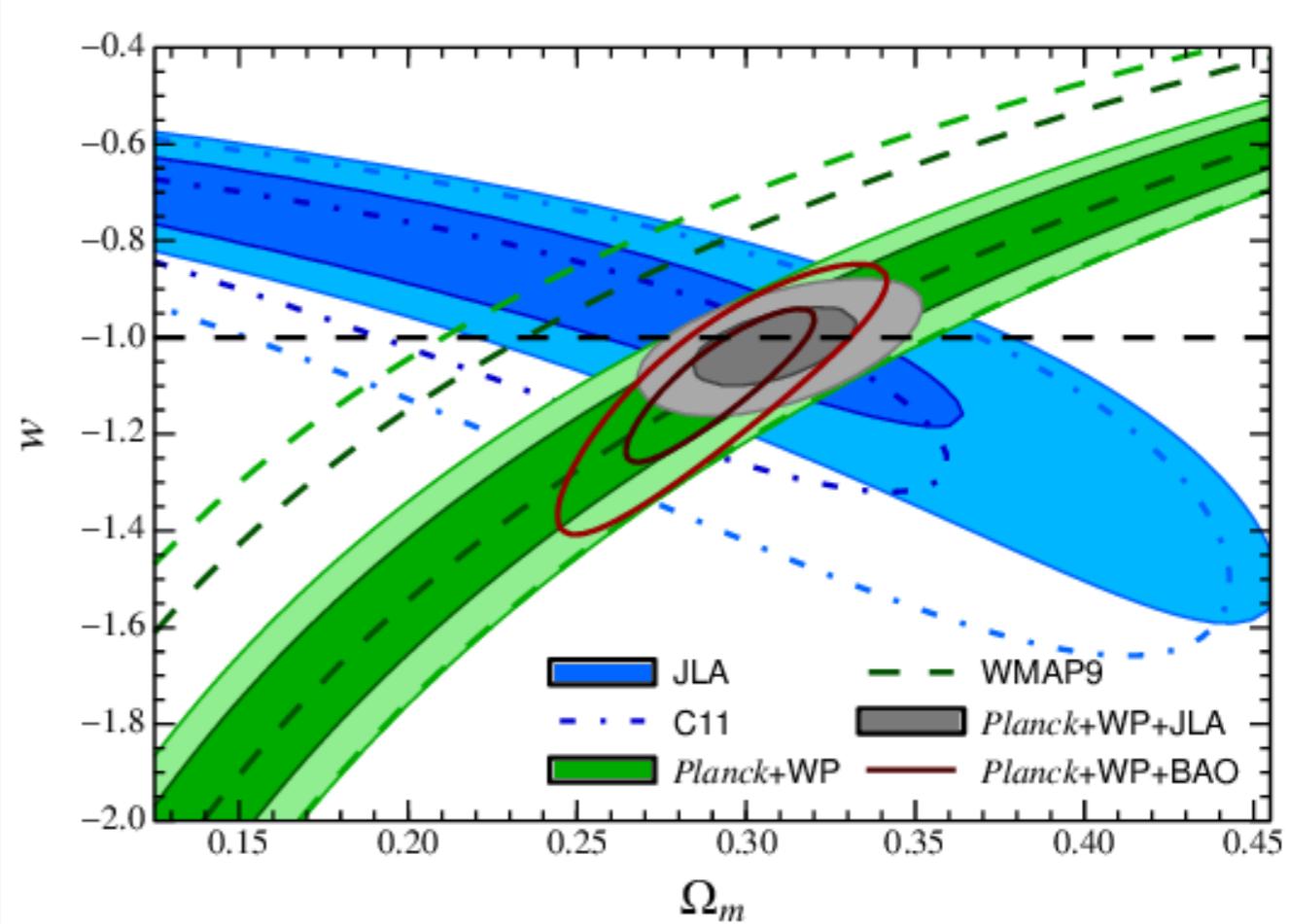
Type Ia SN Cosmology

Hubble Diagram Modern SN Ia Surveys



Joint Lightcurve Analysis
(JLA, Betoule et al. 2014)

Cosmological Constraints



$$w = -1.027 \pm 0.055 \text{ (stat+sys)}$$

Cosmology with standard candles

For us: Simplify:

Let's say Supernovae are *standard* candles

$$M_s \sim N(M_0, \sigma_{\text{int}}^2) \quad \text{Population Distribution}$$

$$m_s = M_s + \mu(z_s; \theta) \quad (\text{Log}) \text{ Inverse Square Law}$$

Assume apparent magnitudes and redshifts
 $\{m_s, z_s\}$ measured perfectly

$$\theta = (H_0, \Omega_M, \Omega_L, w) \quad \text{Cosmological Parameters}$$

Theoretical Distances

1 Comoving Distance

The dimensionless comoving distance \tilde{d} to an object with redshift z is:

$$\tilde{d}(z; \Omega_M, \Omega_\Lambda, w) = \int_0^z \frac{dz'}{\sqrt{\Omega_M(1+z')^3 + \Omega_k(1+z')^2 + \Omega_\Lambda(1+z')^{3(w+1)}}} \quad (1)$$

where Ω_M is the matter density, Ω_Λ is the dark energy density, the curvature is $\Omega_k \equiv 1 - \Omega_M - \Omega_\Lambda$, and w is the equation-of-state parameter of dark energy.

In terms of the scale factor $a \equiv (1+z)^{-1}$, the integral can be expressed as:

$$\tilde{d}(z; \Omega_M, \Omega_\Lambda, w) = \int_{(1+z)^{-1}}^1 \frac{da'/a'}{\sqrt{\Omega_M/a' + \Omega_k + \Omega_\Lambda/a'^{(1+3w)}}} \quad (2)$$

(derived from standard FLRW metric)

Theoretical Luminosity Distances (Arbitrary Curvature)

2 Luminosity Distance

The theoretical cosmological distance to an object with redshift z inferred from the standard candle method are given by the luminosity distance d_L . The dimensionless distance is:

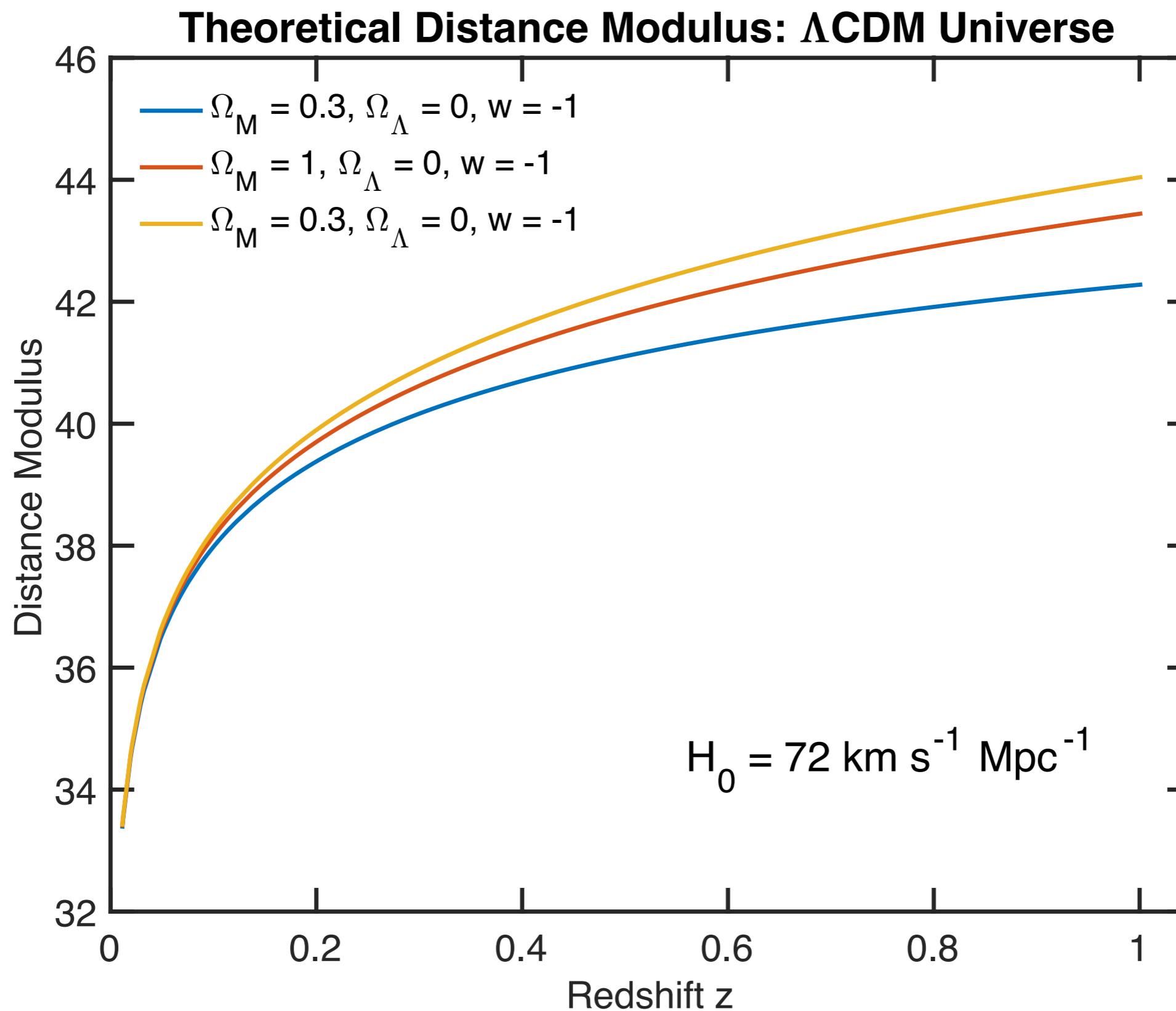
$$\tilde{d}_L(z; \Omega_M, \Omega_\Lambda, w) = (1+z) \begin{cases} |\Omega_k|^{-\frac{1}{2}} \sinh[\sqrt{|\Omega_k|} \tilde{d}(z; \Omega_M, \Omega_\Lambda, w)], & \Omega_k > 0 \\ \tilde{d}(z; \Omega_M, \Omega_\Lambda, w), & \Omega_k = 0 \\ |\Omega_k|^{-\frac{1}{2}} \sin[\sqrt{|\Omega_k|} \tilde{d}(z; \Omega_M, \Omega_\Lambda, w)], & \Omega_k < 0 \end{cases} \quad (3)$$

For a given value of the Hubble constant, the dimensionful luminosity distance is

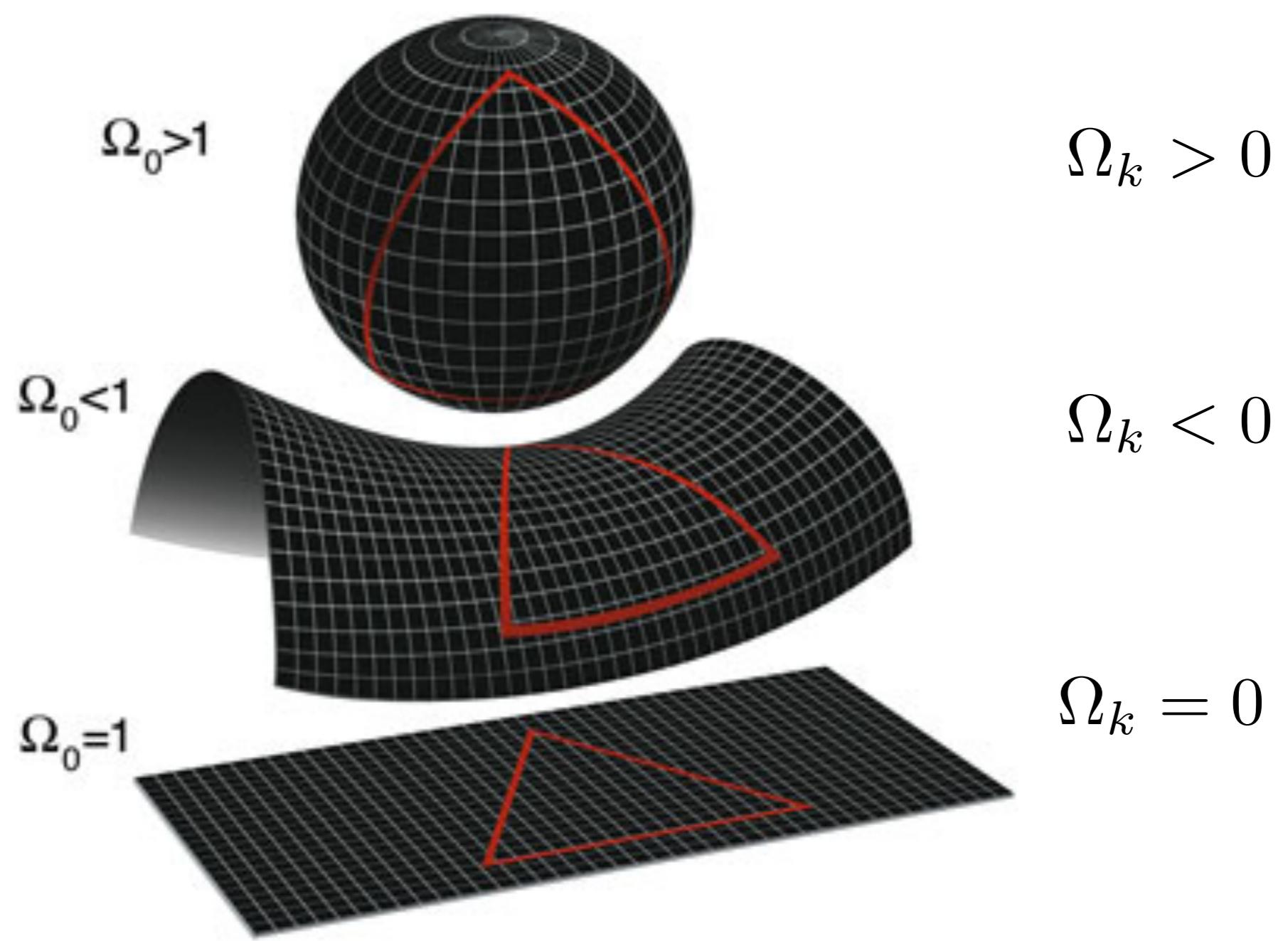
$$d_L(z; \Omega_M, \Omega_\Lambda, w, H_0) = \frac{c}{H_0} \tilde{d}_L(z; \Omega_M, \Omega_\Lambda, w) \quad (4)$$

(derived from standard FLRW metric)

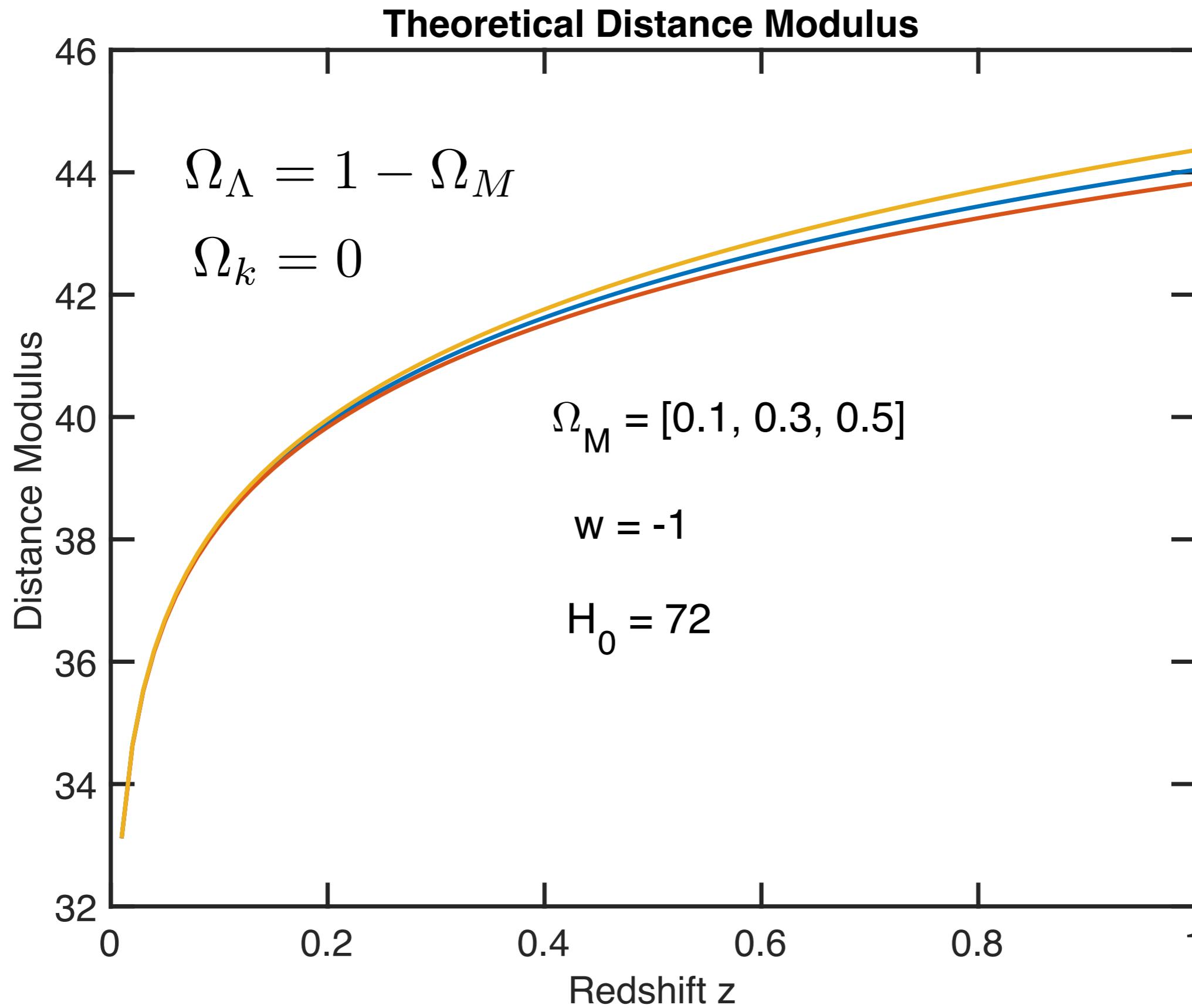
Theoretical Distance Modulus (Arbitrary Curvature)



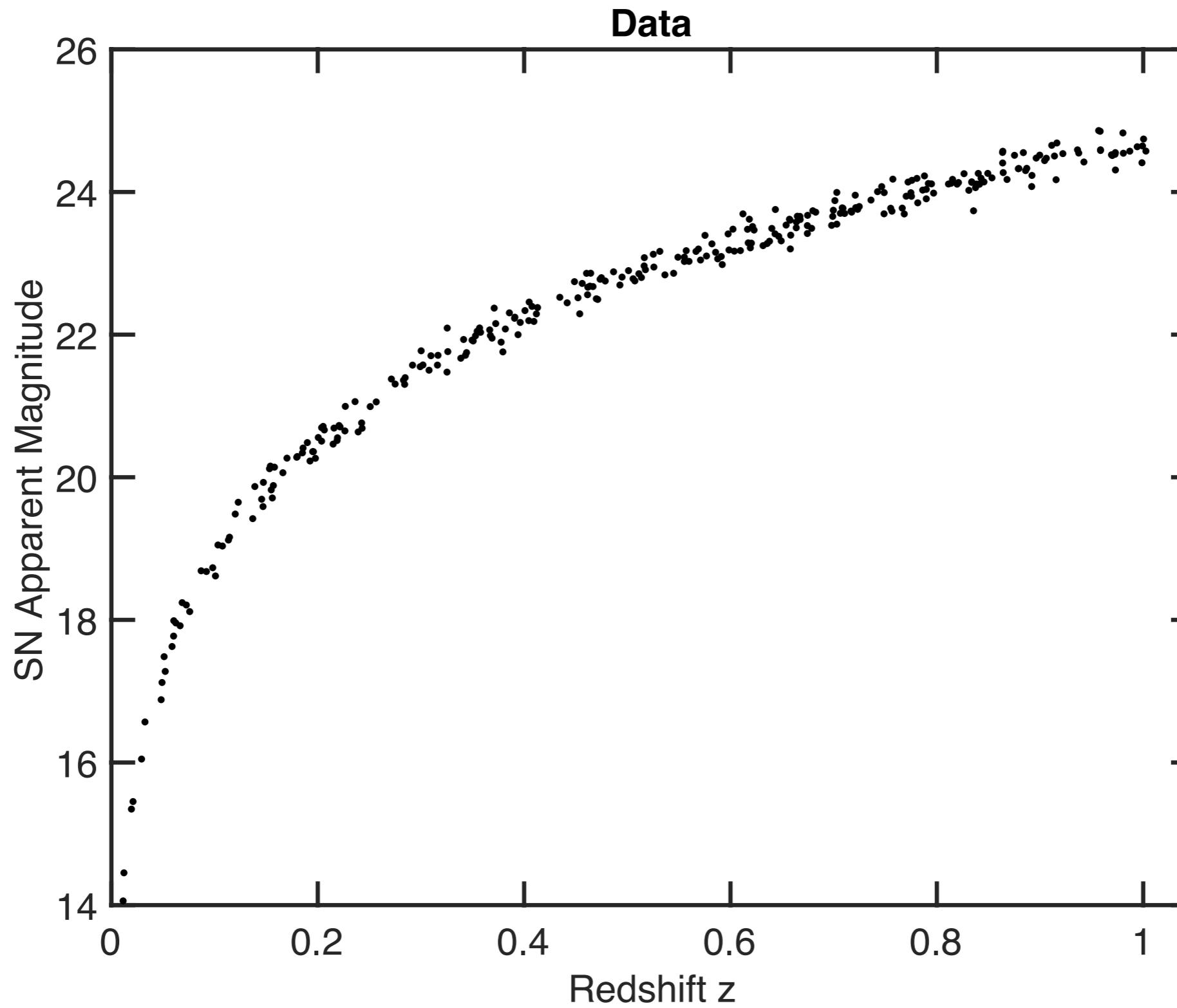
Curvature of the Universe



Theoretical Distance Modulus



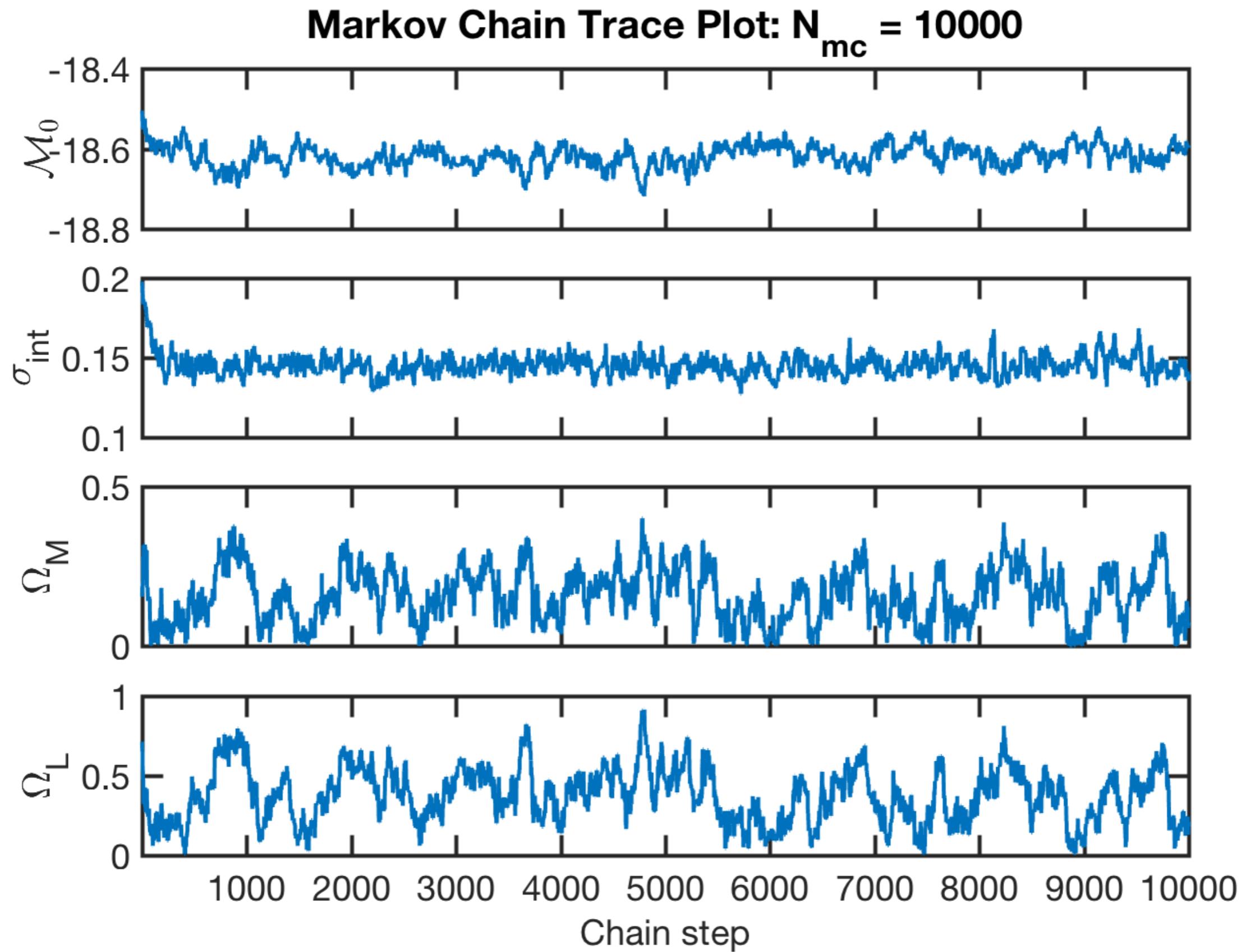
Idealised supernova dataset



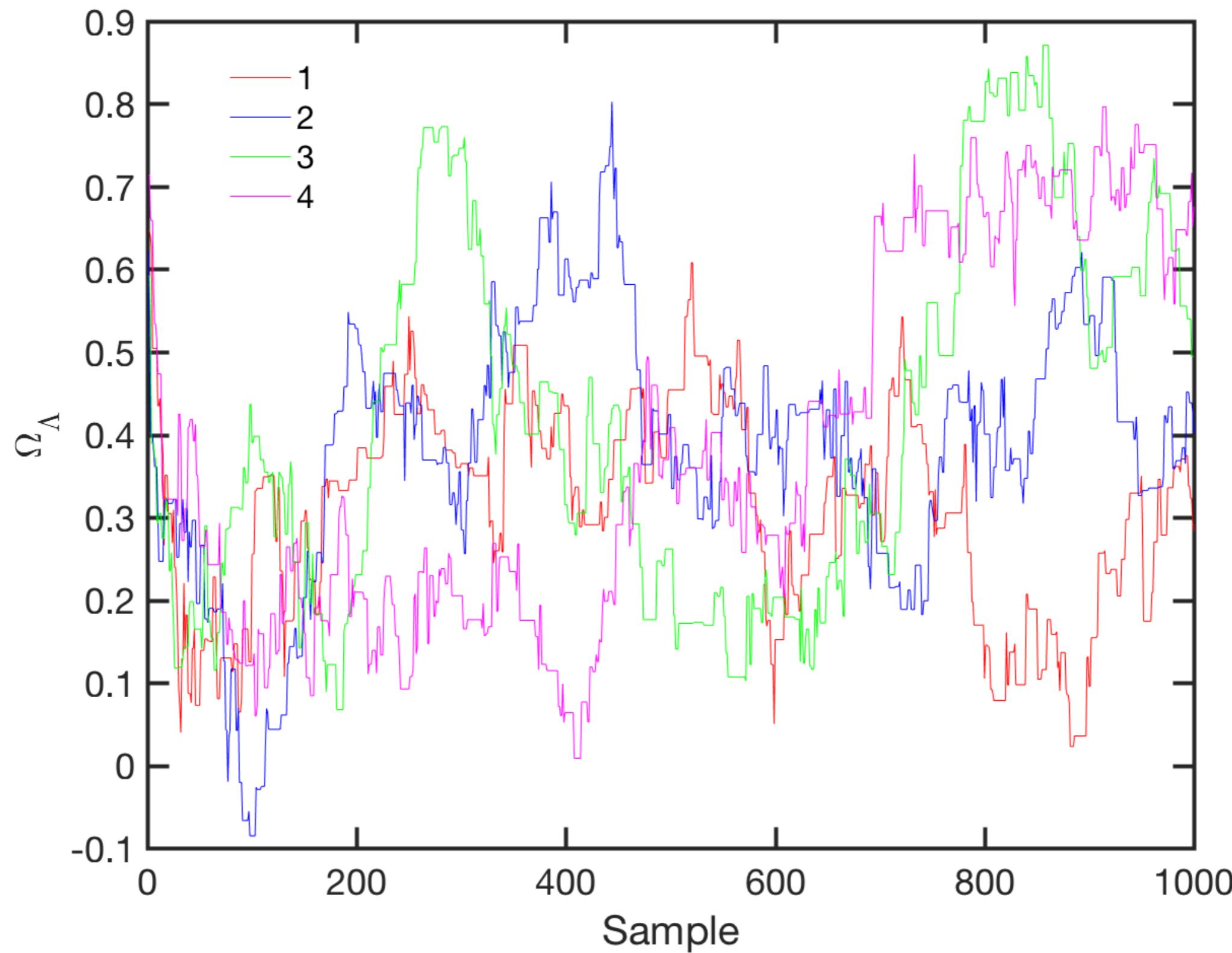
First assume $w = -1$

Write down model & posterior

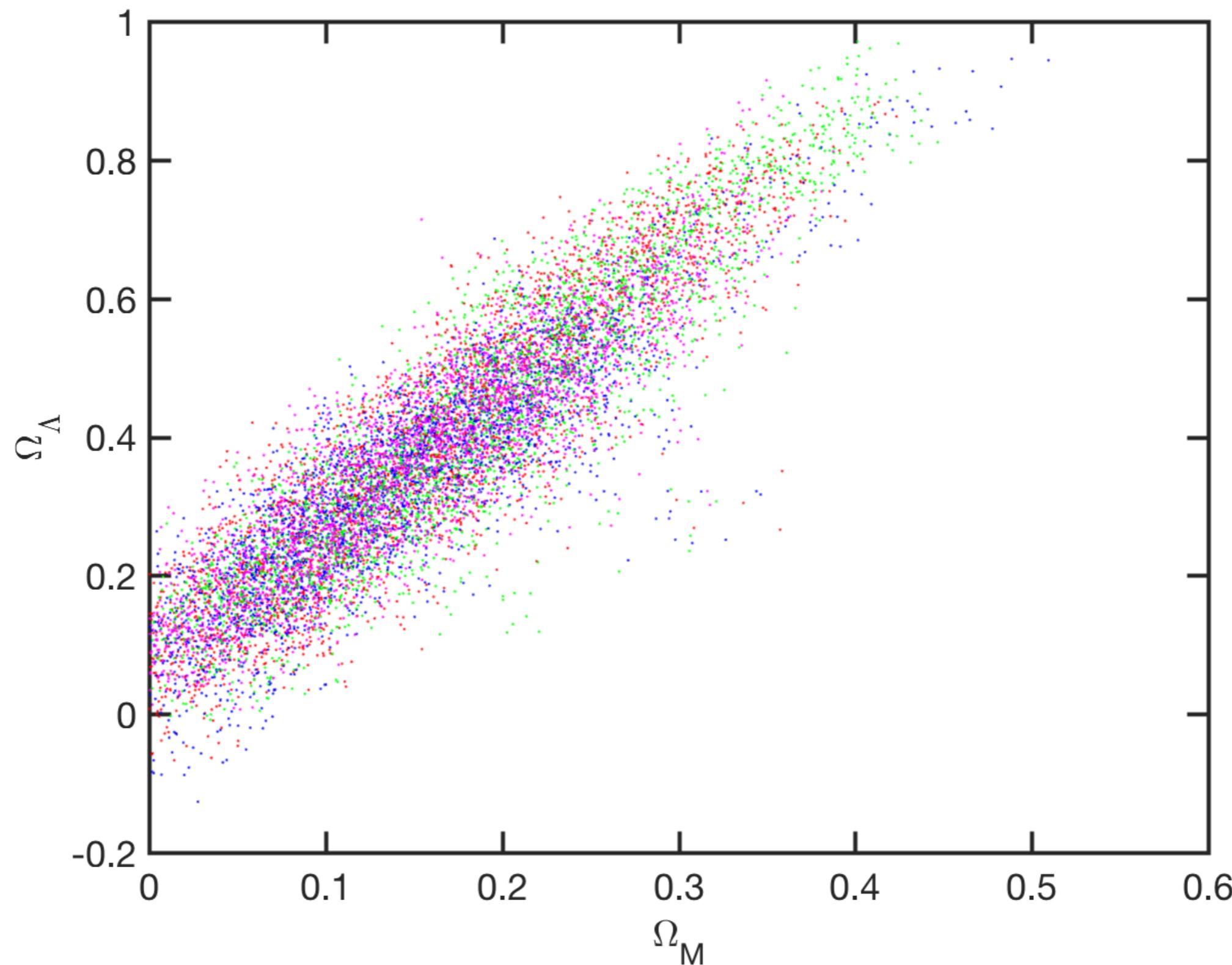
4D Metropolis: one chain



Multiple Independent Chains



Multiple Independent Chains



Assessing Convergence with multiple chains: Gelman-Rubin (G-R) ratio

Monitoring convergence of each scalar estimand

Suppose we have simulated m parallel sequences, each of length n (after discarding the first half of the simulations). For each scalar estimand ψ , we label the simulation draws as ψ_{ij} ($i = 1, \dots, n; j = 1, \dots, m$), and we compute B and W , the between- and within-sequence variances:

$$B = \frac{n}{m-1} \sum_{j=1}^m (\bar{\psi}_{\cdot j} - \bar{\psi}_{..})^2, \text{ where } \bar{\psi}_{\cdot j} = \frac{1}{n} \sum_{i=1}^n \psi_{ij}, \quad \bar{\psi}_{..} = \frac{1}{m} \sum_{j=1}^m \bar{\psi}_{\cdot j}$$
$$W = \frac{1}{m} \sum_{j=1}^m s_j^2, \text{ where } s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (\psi_{ij} - \bar{\psi}_{\cdot j})^2.$$

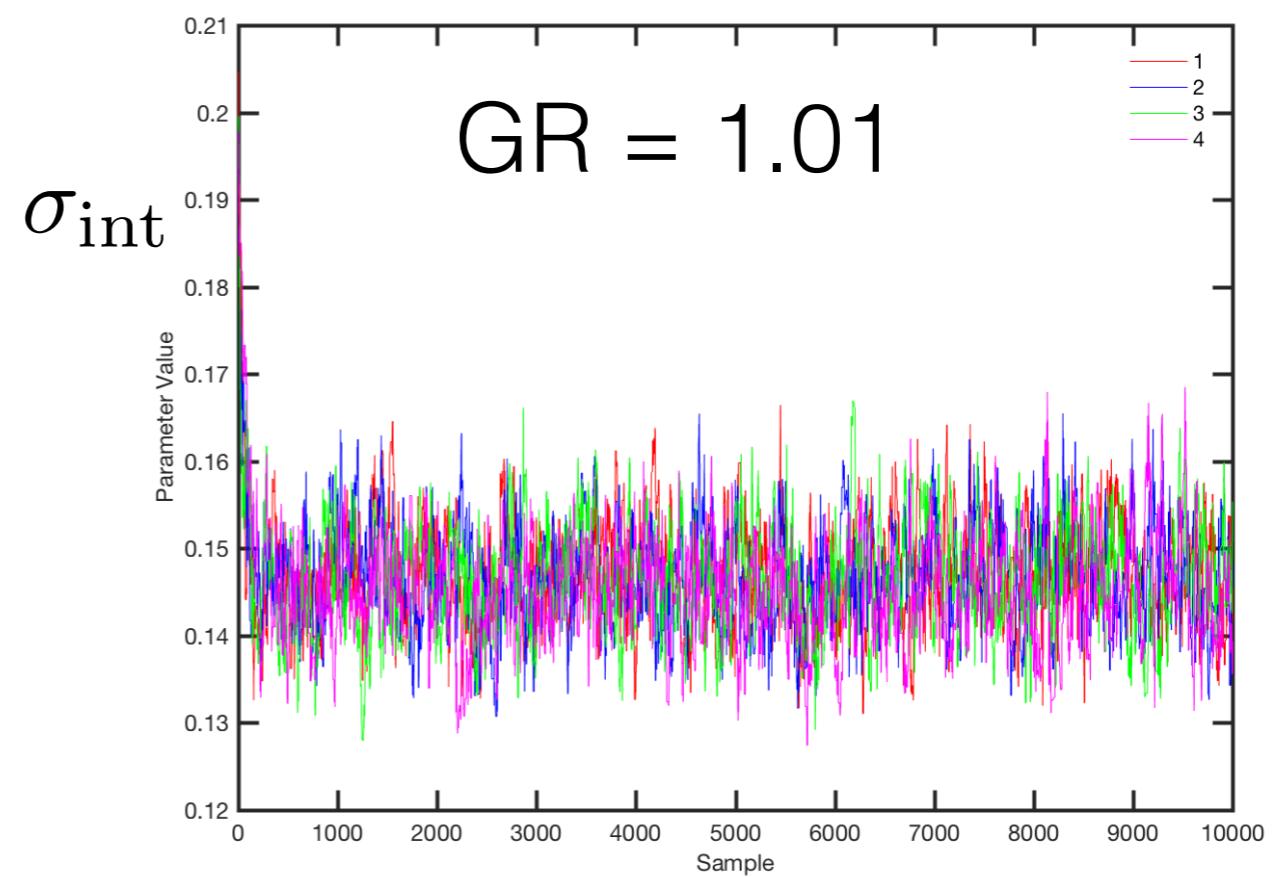
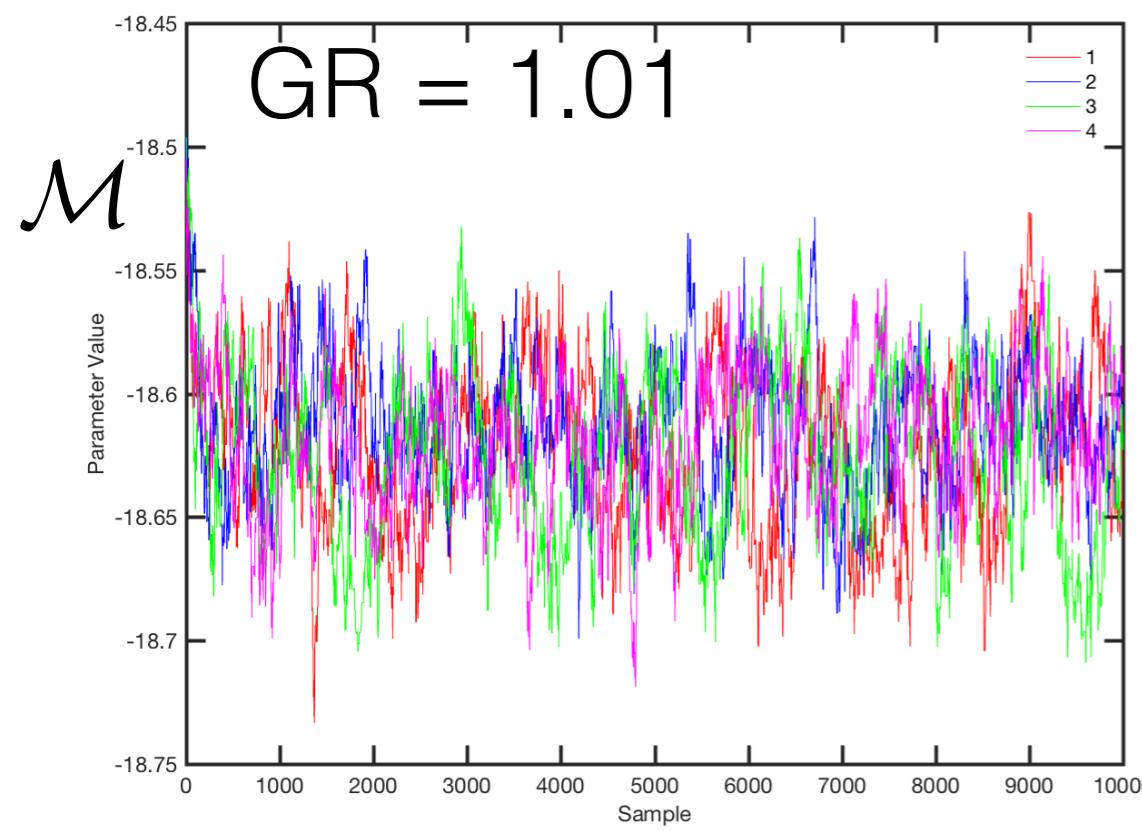
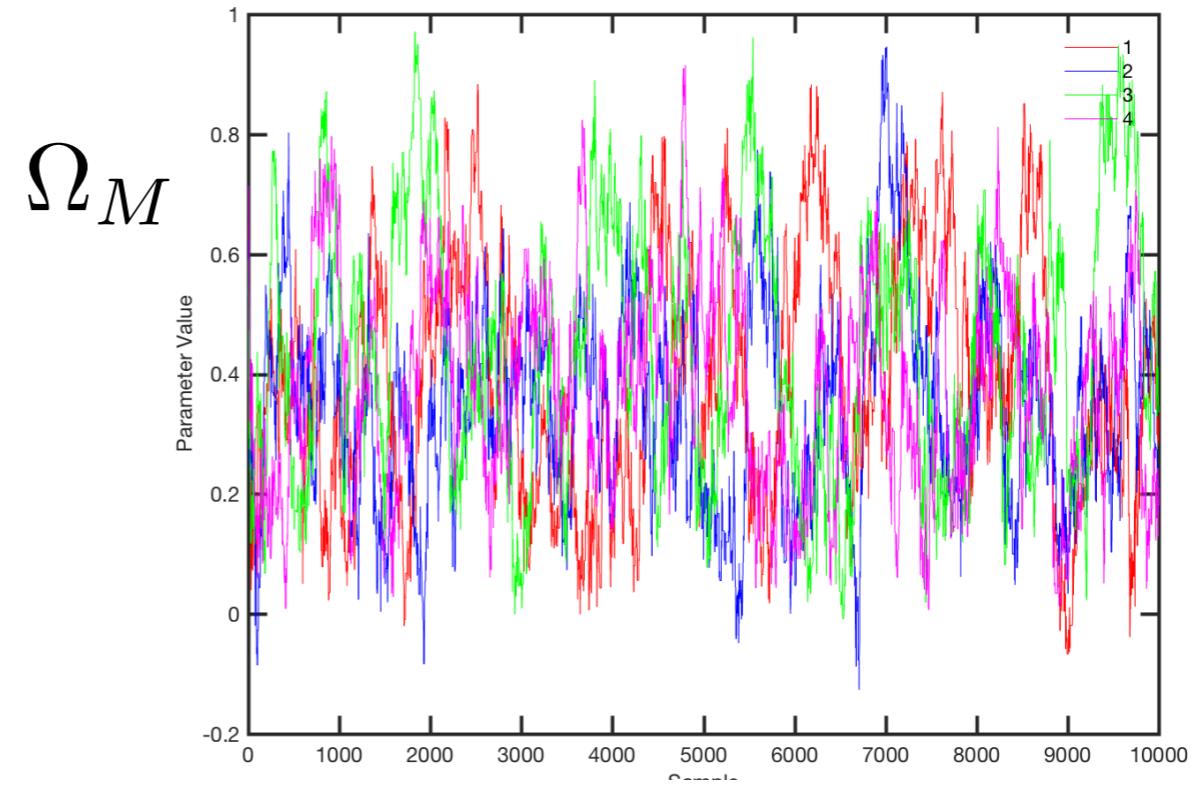
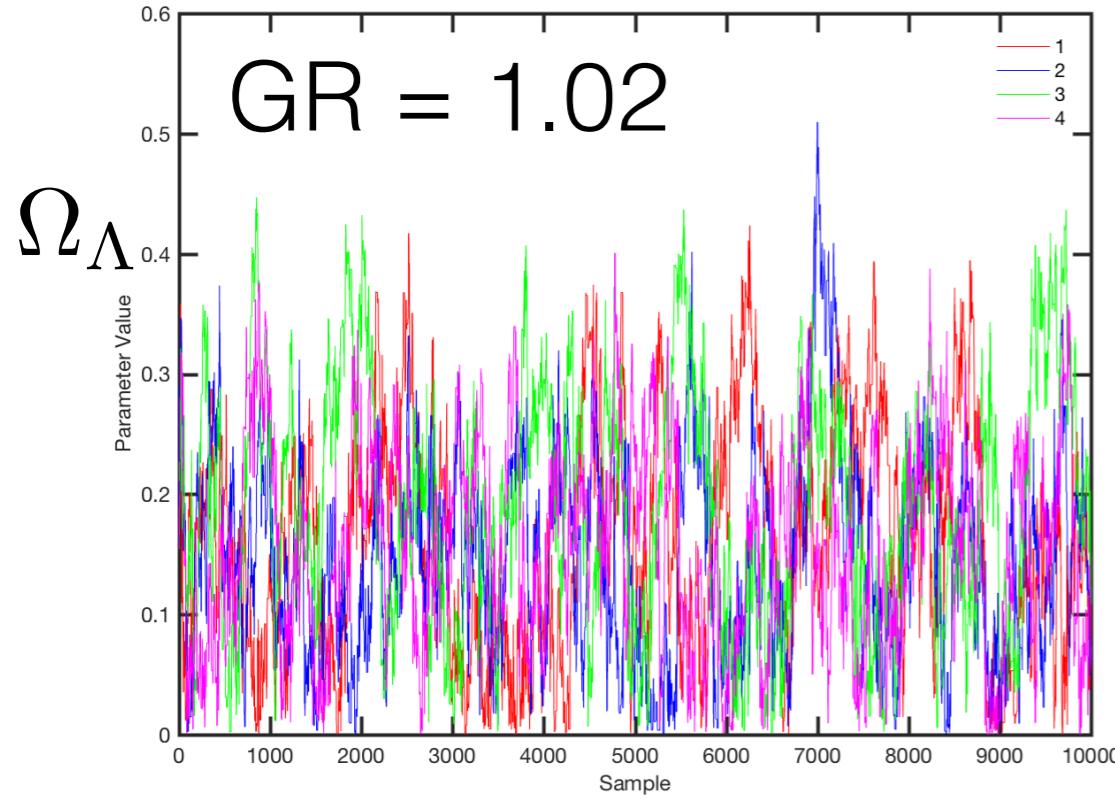
We can estimate $\text{var}(\psi|y)$, the marginal posterior variance of the estimand, by a weighted average of W and B , namely

$$\widehat{\text{var}}^+(\psi|y) = \frac{n-1}{n} W + \frac{1}{n} B. \quad (11.3)$$

G-R ratio: $\hat{R} = \sqrt{\frac{\widehat{\text{var}}^+(\psi|y)}{W}}, \approx 1$

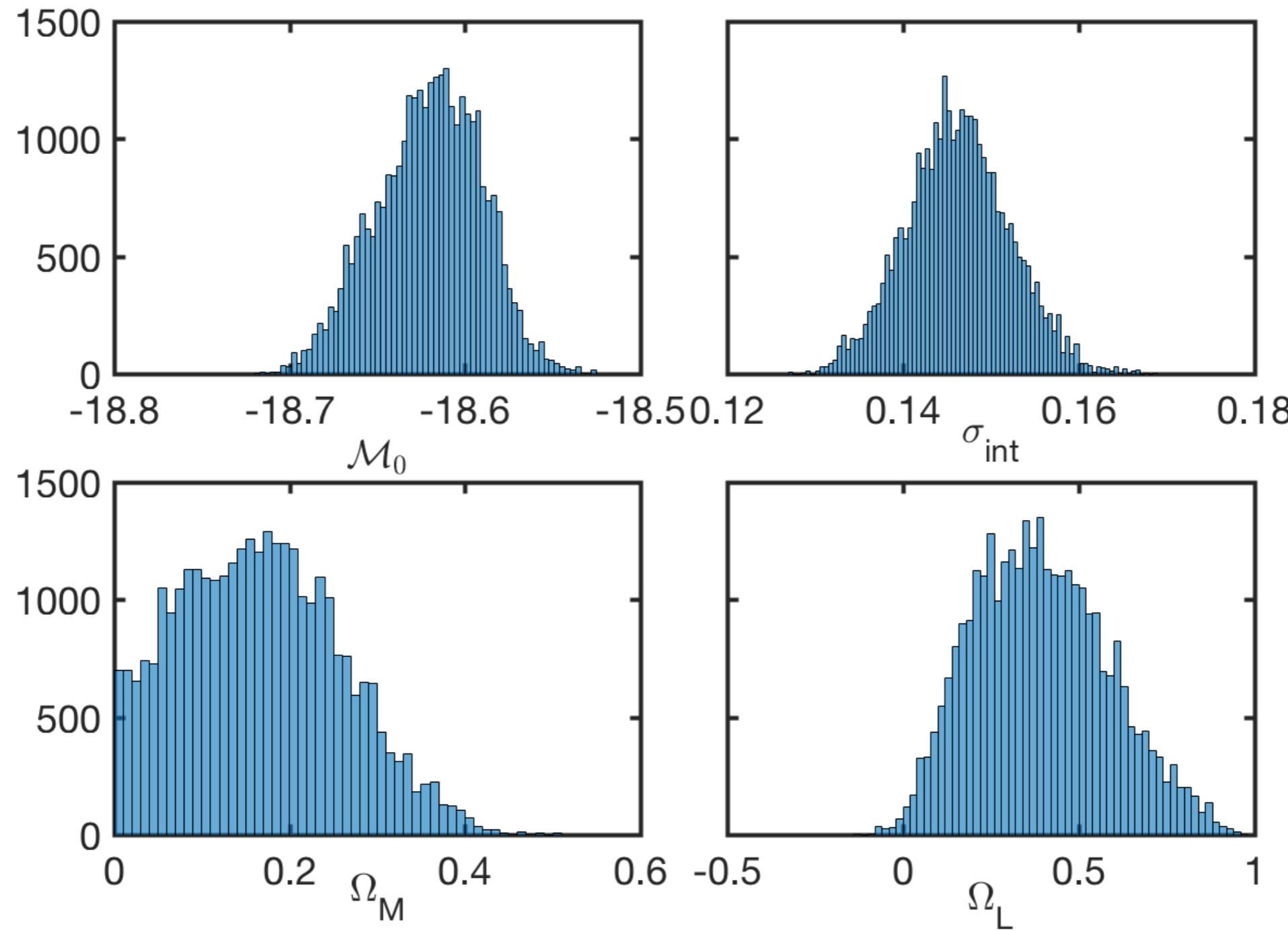
Multiple Independent Chains

GR = 1.02

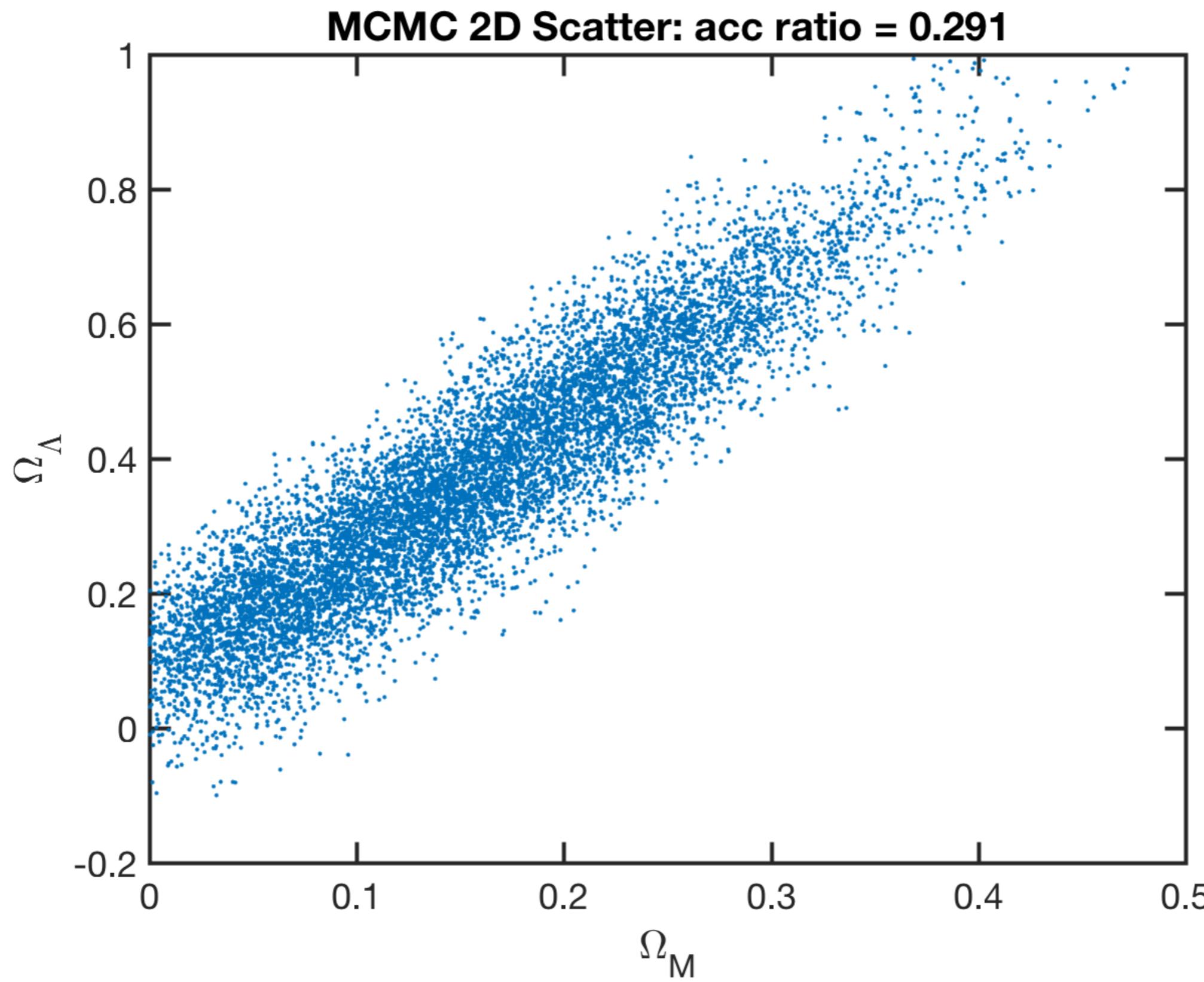


Cut burn-in (20%) and combine chains

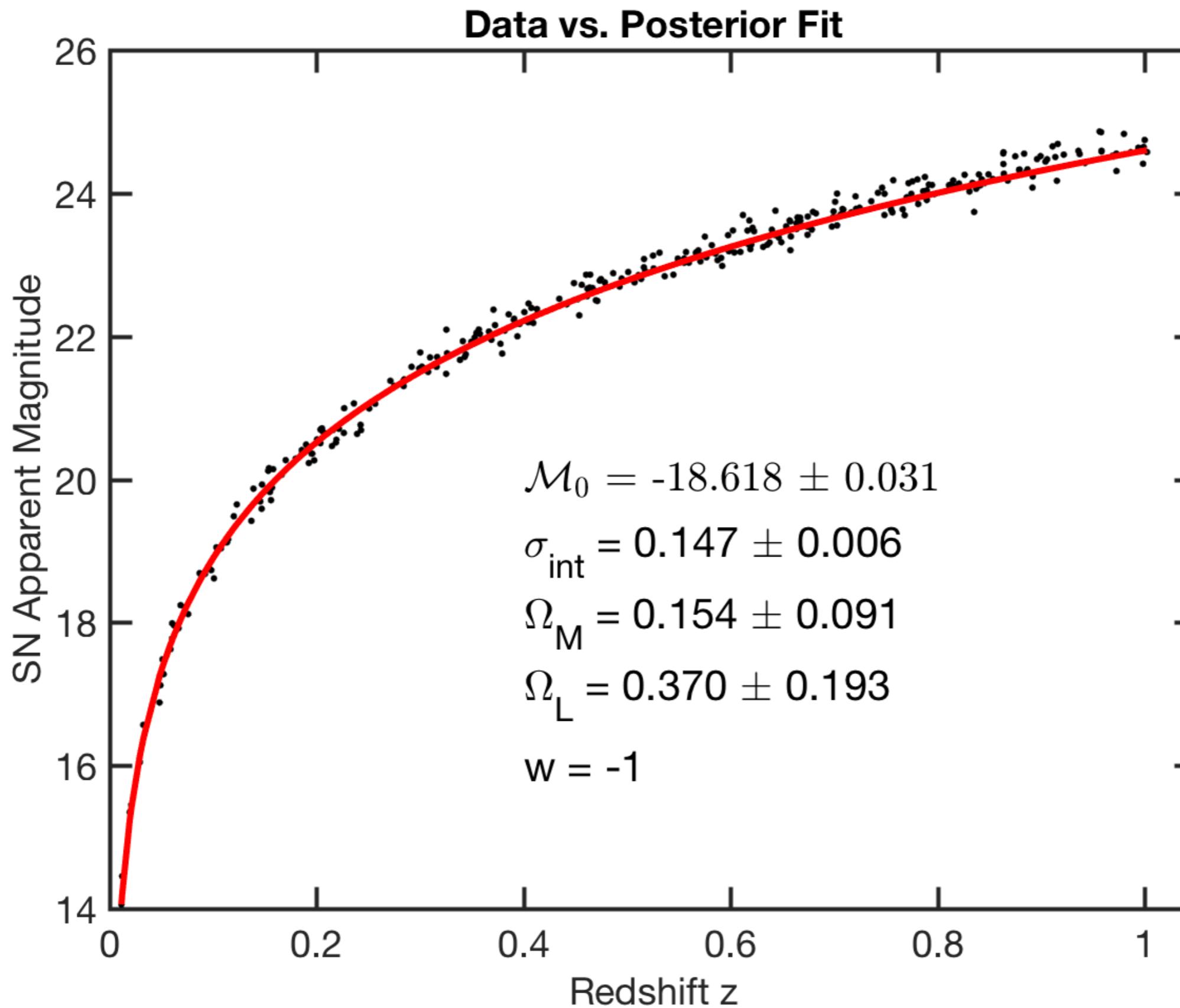
Posterior Histograms



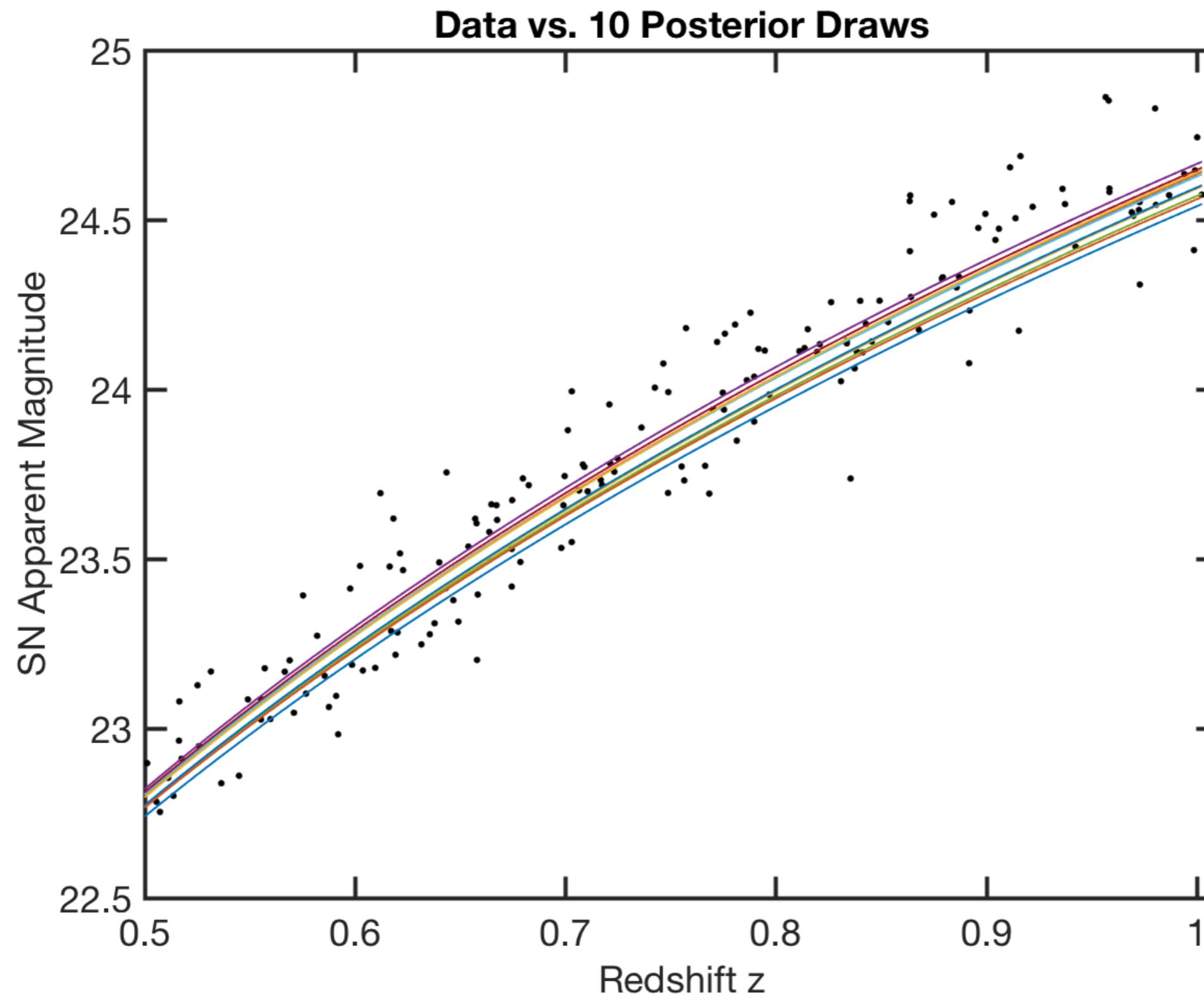
Posterior Scatter plot



Fit with posterior mean



Random Posterior Draws of parameters from chain



The deceleration parameter q_0

of the cosmological models. "It is a striking and slightly puzzling fact that almost all current cosmological observations can be summarized by the simple statement: The jerk of the Universe equals one" [25].

Let us make Taylor expansion of the scale factor in time using the above introduced parameters

$$a(t) = a_0 \left[1 + H_0(t - t_0) - \frac{1}{2}q_0 H_0^2(t - t_0)^2 + \frac{1}{3!}j_0 H_0^3(t - t_0)^3 + \frac{1}{4!}s_0 H_0^4(t - t_0)^4 + \frac{1}{5!}l_0 H_0^5(t - t_0)^5 + O((t - t_0)^6) \right]. \quad (2.49)$$

$$\begin{aligned} a(0) &= 0 \text{ (Big Bang)} \\ a(1) &= 1 \text{ (Today)} \end{aligned}$$

$$q_0 = \Omega_M/2 - \Omega_\Lambda$$

The posterior of derived quantities

Deceleration Parameter: $q_0 = \Omega_M/2 - \Omega_\Lambda$

