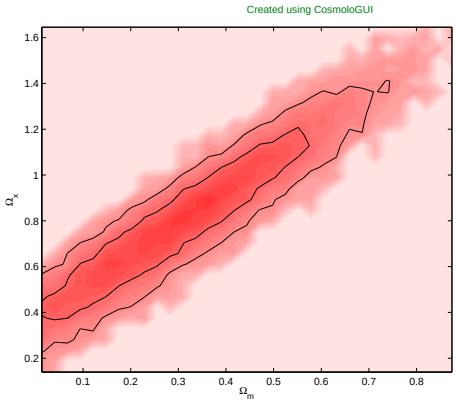
Master Thesis
Stockholm Observatory




# Investigating the dark energy using a Markov Chain Monte Carlo

Kristoffer Hniopek

Department of Astronomy, Stockholm University
May, 2006

**Abstract**

I have determined the mass density, $\Omega_m$, the energy density of the dark energy, $\Omega_x$, and the equation of state parameters $w_0$ and $w_1$ based on the measurement of 115 type Ia supernovae from the first year of the SuperNova Legacy Survey, ranging from $z = 0.015$ to $z = 1.01$. I am also using two additional datasets; the shift parameter, i.e. the reduced distance to the surface of last scattering, determined by the Wilkinson Microwave Anisotropy Probe and the baryonic acoustic peak determined from the Sloan Digital Sky Survey. The results favour a model corresponding to a flat universe with a cosmological constant, but are consistent with a large range of other models. For the best model I yield the results $0.25 < \Omega_m < 0.29$ and $0.71 < \Omega_x < 0.75$ (68 % confidence level) with the best fit $\Omega_m = 0.27$ and $\Omega_x = 0.73$.

# Contents

# 1 Introduction

Until the early 20th century, man believed that the whole Universe consisted only of the Milky Way. However, soon it became clear that some objects were in fact other galaxies. A breakthrough came in 1929 when Hubble published his famous paper, finding a relationship between the redshift and distance of galaxies. After that it became clear that the Universe is not static, which was before that considered obvious, but expanding. By backtracking this expansion, the theory of Big Bang evolved. The fate of the Universe, if it was to expand forever or at some point collapse, was now thought to be absolutely determined by its total density. If it contains enough matter, the expansion is halted and then the Universe would collapse into a Big Crunch. On the other hand, if the Universe contains too little matter, the expansion would go on forever.

This was changed when another giant breakthrough came in 1998, when it was discovered that the Universe was not only expanding but also accelerating. This means that something counter-acts the gravity and pushes the galaxies away from each other. This sent a shockwave through the scientific world and much effort was made to explain the results. The most prominent explanation is that the Universe is filled with a so-called dark energy. Now the fate of the Universe is determined both by the matter, baryonic and non-baryonic, and by this dark energy. This dark energy has a negative pressure, which leads to acceleration. The determination of the energy density of matter and dark energy, and also the equation of state of the dark energy, have become the main issue in cosmology the last years. A popular candidate to the dark matter is Einstein's cosmological constant.

I am writing this thesis as a last part toward a master in astronomy at the Department of Astronomy at Stockholm University. My supervisor is Edvard Mörtsell. I will be trying to determine the cosmological parameters, which describe the content in the universe, from Supernova Ia (SN Ia) observations and also additional measurements from the Cosmic Microwave Background (CMB) and from Baryonic Acoustic Oscillations (BAO) to put some more constraints on the parameters. The use of a fast Markov Chain Monte Carlo (MCMC) algorithm allow me to drop some assumptions usually made and see what happens; that is allowing a non-flat model and/or parameterize the equation of state of the dark energy and allow it to change in time.

# 2 Theoretical background

## 2.1 The metric

Hubble showed that all galaxies (except the local ones, which are bound by gravity) are moving away from us with a velocity proportional to the distance. Actually, the proper way to see it is that it is the universe itself that is expanding, described by the so-called scale factor, $R(t)$. In an expanding universe the metric is not what we are used to in our everyday life where the geometry obeys Pythagoras law, i.e. an infinitesimal distance $ds$ is described by

$$ds^2 = dx^2 + dy^2 + dz^2, \tag{1}$$

where $x$, $y$ and $z$ are the Euclidian coordinates. On large scales the metric of the universe is described by the four-dimensional space-time geometry, the so-called Friedmann, Lemaitre, Robertson-Walker (FLRW) metric (Raine & Thomas, 2001), named after its originators:

$$ds^2 = c^2dt^2 - R(t)^2\left(\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2sin^2(\theta)d\phi^2\right). \tag{2}$$

Here $c$ is the light speed in vacuum, $\theta$ and $\phi$ are the spherical angles and $k$ is the curvature factor. Since both $R(t)$ and $k$ is arbitrarily chosen, one can choose $R(t)$, so that $k$ takes three distinct, different cases. $k = 1$ means that the geometry of the universe is spherical (closed); two parallel rays eventually cross each other, given large enough scales. For $k = -1$ the geometry of the universe is hyperbolic (open); two parallel rays will eventually diverge from each other. The third case is $k = 0$. In this case the universe is perfectly flat and our intuitive view of the geometry is correct. It may look like the universe has a centre at $r = 0$, but that is not the case. As on the surface of the Earth the curvature is the same everywhere, so the origin on Earth (the meridian or the equator) is arbitrarily chosen. One can also note that for small $r$ the effect of $k$ is negligible, hence we must look at a large volume to see to effect of curvature. Now consider a light ray moving towards an observer. We can without loss of generality put the observer in the origin and letting $\theta = \phi = 0$. For a light ray $ds^2 = 0$ and we have

$$\frac{cdt}{R} = -\frac{dr}{\sqrt{1-kr^2}}. \tag{3}$$

The minus sign is because the ray is approaching the observer in the origin. Integrating gives

$$c\int \frac{\mathrm{d}t}{R} = -\int \frac{\mathrm{d}r}{\sqrt{1-kr^2}}. \tag{4}$$

The right hand, call it $\chi$, gives different results when evaluating, depending on $k$. for a closed universe we have $\chi = \sqrt{k}\arcsin(\sqrt{k}R)$. The equation yields $\chi = R$ for a flat universe and $\chi = \sqrt{k}\mathrm{arcsinh}(\sqrt{k}R)$ for an open universe. Thus, the metric looks a bit different for the different kinds of universes and one can write many of the equations in the more compact form by introducing the function $S(x)$.

$$S(x) = \begin{cases} \sin(x) & \text{if } k = +1 \\ x & \text{if } k = 0 \\ \sinh(x) & \text{if } k = -1 \end{cases} \tag{5}$$

## 2.2 Field equation and the equation of state

In order to combine the FLRW metric with matter, the field equations of Albert Einstein's general relativity is needed. The solutions, which are the ones of interest, generate two equations (Raine & Thomas, 2001). The first is the Friedmann equation

$$\left(\frac{dR}{dt}\right)^2 = \frac{8\pi G}{3}R^2\rho - kc^2. \tag{6}$$

Here $G$ is the gravitational constant and $\rho$ is the sum of all kinds of densities in the universe i.e. $\rho = \rho_r + \rho_m + \rho_x$ where index $r$, $m$ and $x$ represent radiation, matter and dark energy, respectively. The dark energy is here treated as an ordinary source of gravity; the difference from matter is that the dark energy has negative pressure (see below). The second equation is the equation of local energy conservation

$$\frac{d(\rho R^3)}{dt} + \frac{p}{c^2}\frac{dR^3}{dt} = 0. \tag{7}$$

Of great interest is also a combination of the two, the so-called acceleration equation

$$\frac{d^2R}{dt^2} = -\frac{4\pi G}{3}R\left(\rho + \frac{3p}{c^2}\right) \tag{8}$$

where $p$ is the pressure. The density and the pressure is related as what is known as the equation of state $p = p(\rho)$. Assuming that the dark energy

is a cosmological constant with a constant density $\rho_\Lambda$, the equation of local energy conservation gives

$$\rho_\Lambda \frac{dR^3}{dt} = -\frac{p}{c^2}\frac{dR^3}{dt} \tag{9}$$

which gives

$$\rho_\Lambda = -\frac{p}{c^2} \tag{10}$$

i.e. the cosmological constant has negative pressure. If we, in the usual astronomer way, put $c = 1$ we have $\rho_\Lambda = w_0 p$, where $w_0$ now equals $-1$. This is the usual way of parameterize the equation of state. One can also let it vary in redshift by putting

$$w(z) = w_0 + \frac{z}{1+z} w_1. \tag{11}$$

## 2.3 Critical density

Rearranging the Friedmann equations, equation (6), and substitute $H \equiv \frac{\dot{R}}{R}$ gives

$$-kc^2 = R^2 H^2 \left(1 - \frac{8\pi G}{3H^2}\rho\right). \tag{12}$$

$H$ is the Hubble parameter explained below. We now see that $\rho = \frac{3H^2}{8\pi G}$ gives $-kc^2 = 0$ which gives a flat universe ($k = 0$). We therefore call this the critical density

$$\rho_c \equiv \frac{3H^2}{8\pi G}. \tag{13}$$

$\rho < \rho_c$ gives $k = -1$ and we have a negatively curved universe, $\rho > \rho_c$ gives a positively curved universe. Note also that the critical density is a function of time through $H = H(t)$. We can now introduce the energy density parameter, $\Omega$, which is the density in units of the critical density. We have

$$\Omega_r = \frac{\rho_r}{\rho_c} = \frac{8\pi G \rho_r}{3H^2}, \tag{14}$$

$$\Omega_m = \frac{\rho_m}{\rho_c} = \frac{8\pi G \rho_m}{3H^2} \tag{15}$$

$$\Omega_x = \frac{\rho_x}{\rho_c} = \frac{8\pi G \rho_x}{3H^2}. \tag{16}$$

Since the universe today is not radiation dominated (Raine & Thomas, 2001), and has not been for a long time, I do not consider the effect of the radiation from now on. I now also define the energy density of the curvature from the remaining two densities: $\Omega_k = 1 - \Omega_m - \Omega_x$. $\Omega_k$ is now a continuous variable, $\Omega_k = 0$ represent a flat universe and $\Omega_k < 0$ and $\Omega_k > 0$ represent a closed and an open universe, respectively.

## 2.4 Expansion

The expansion of the universe means that every process, even the wavelength of light itself, is stretched out, as you look deeper into space. In this way the redshift is defined as the fractional change of the stretched wavelength i.e.

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} \tag{17}$$

which is equivalent with

$$1 + z = \frac{\lambda_o}{\lambda_e} \tag{18}$$

where index $e$ and $o$ is the emitted and observed wavelength, respectively. A theoretical distance of length $l_i$ at time $t_i$ is in the same way stretched according to $l(t) = l_i \frac{R(t)}{R(t_i)}$ where $R(t)$ is the scale factor at time $t$. In order to obtain the corresponding relation for the velocity we differentiate with respect to time $v(t) = \frac{l_i}{R(t_i)}\dot{R} = \frac{\dot{R}}{R(t)} l(t) = H(t)l(t)$ where we have introduced the Hubble parameter $H(t)$ which is describing the expansion rate of the universe. For small velocities $v(t) = cz$ and we have $cz = H(t)l(t)$ which is the famous Hubble law. He found that the redshift of galaxies was proportional to their distance and later on it was concluded that the redshifts actually originated from the velocity and that the galaxies were moving away from us.

When the universe is expanding all relevant properties change in the same way as the wavelength, hence we have

$$1 + z = \frac{\lambda_0}{\lambda_e} = \frac{\nu_e}{\nu_0} = \frac{l_0}{l_e} = \frac{R(t_0)}{R(t_e)} = \frac{dt_0}{dt_e}. \tag{19}$$

Index 0 is from now on assigned to the present day values. As seen the expansion is also affecting the time. Imagine a source sending out one photon per second. Since the galaxy has moved away a little bit from us between photon number one and photon two, the second one will have a longer distance to travel. It will therefore arrive at the detector on Earth more than

one second after the first one; the process will be slowed down by a factor $1 + z$.

## 2.5 Magnitude and luminosity distance

Astronomers measure the brightness of objects in the rather awkward unit of apparent or absolute magnitude, due to traditional reasons. The original definition is that the brightest star on the sky is of magnitude zero and the faintest visible star under perfect conditions is of magnitude six. Due to the sensitivity of the eye, this makes the scale logarithmic. The modern definition is (in the most common Vega–magnitude system)

$$m_1 - m_2 = -2.5 \log \frac{F_1}{F_2} \tag{20}$$

and the hot A0 star Vega sets the zero point (Sparke & Gallagher, 2000). Here $m_1$ and $m_2$ are the magnitude of two sources and $F_1$ and $F_2$ their fluxes, respectively. The absolute magnitude, $M$, is the apparent magnitude if the source would be at a standard distance of 10 parsec (32.6 light years). Now, the flux measured from a galaxy with observed redshift $z$ and comoving distance $r_e$ is

$$F = \frac{L}{4\pi S(R(t_0)r_e)^2(1+z)^2} \equiv \frac{L}{4\pi d_l^2}. \tag{21}$$

$L$ is the luminosity of the galaxy and $4\pi S(R(t_0)r_e)^2$ is an expanding sphere around the galaxy, where its light is spread. The two factors of $(1 + z)$ appear because the individual photons are redshifted and are also delayed in time. The introduced $d_l \equiv (1 + z)S(R(t_0)r_e)$ is called the luminosity distance. Note that $d_l$ depends in the curvature trough the function $S(x)$. The apparent magnitude of a supernova can now be written as $m = \mathcal{M} + 5 \log d_l'$, where $\mathcal{M} = 25 + M + 5 \log \frac{c}{H_0}$ contains the unknown (and for my analysis uninteresting) parameters $M$ and $H_0$ (Goliath et al, 2001; Ichikawa & Takahashi, 2005). $\mathcal{M}$ is a constant factor added to all the supernovae, it will not change the physics in a relevant way and is therefore of no importance. $d_l' \equiv \frac{H_0 d_l}{c}$ is the Hubble-parameter independent luminosity distance which can now also be written as (Goliath et al, 2001; Ichikawa & Takahashi, 2005)

$$d_l' = \frac{1+z}{\sqrt{\mid \Omega_k \mid}} S\left(\sqrt{\mid \Omega_k \mid} \int \frac{\mathrm{d}z}{H'(z)}\right) \tag{22}$$

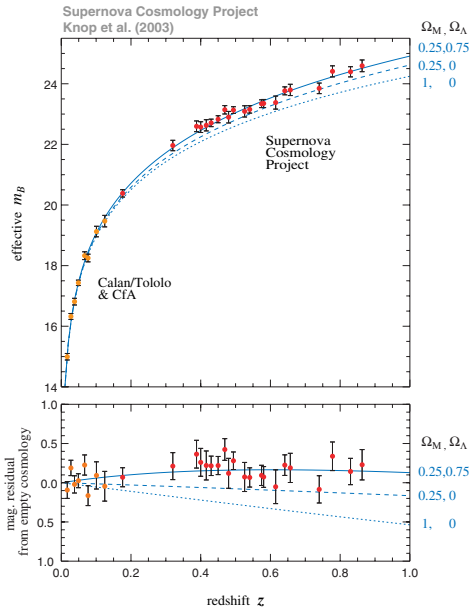



Figure 1: Fitting theoretical models to observational data, implying the existence of the dark energy. Picture taken from the Supernova Cosmology Project (Perlmutter et al, 1998).

where $S(x)$ is given by equation (5) and the factors $\sqrt{\mid \Omega_k \mid}$ are removed for the flat universe. Finally

$$H'(z) \equiv \frac{H(z)}{H_0} = \sqrt{(1+z)^3\Omega_m + (1+z)^2\Omega_k + (1+z)^{3(1+w_0+w_1)} e^{\frac{-3w_1 z}{1+z}} \Omega_x} \tag{23}$$

(see Appendix A). By fitting the theoretical value of the apparent magnitude to the observations, one is able to constrain the cosmological parameters. An example of such a fit is shown in figure (1).

## 2.6 Standard candles

A very powerful tool in cosmology is the magnitude versus redshift test. The idea is to compare the magnitude of an object at different redshifts and hence calculating the distance to the object. The problem is that the object must

be equally luminous at all redshifts. Such a source is called a standard candle. In order to determine the energy densities of the universe from the standard candles, they must also be bright enough so that they can be observed at large redshifts, say up to $z = 1$. Until recently, finding a good standard candle has not been easy, at least not extra-galactic ones. For some reason the brightest galaxy in a rich cluster always have about the same luminosity and hence used to be a popular standard candle. The problem is that the principle of galaxy evolution is not well understood. A galaxy at $z = 1$ is observed as it looked like when the universe was very young, about one third of its current age. Since galaxies seem to evolve or merge they are not a suitable choice. Lately however, the improved knowledge of type Ia supernovae, SN Ia, has replaced them as the favorite standard candle.

### 2.6.1 Supernova Ia as standard candle

The source of a SN Ia is a carbon-oxygen white dwarf. A white dwarf is the end stage of a medium-mass star. When the star has burned all the helium in the core into carbon, it will not have high enough temperature to fuse the carbon any further. At this point the star ejects most of its material into the interstellar medium and is left only with a degenerate carbon-oxygen core (for details, see Christensen–Dalsgaard (2003)). In a degenerate gas the gravitational collapse is prevented only from the electrons repelling each other, forcing them into higher energy states. The core is now very luminous and dense; it has a mass of about 0.7–1.4 solar masses and is approximately the size of Earth. Therefore it is named a white dwarf. Chandrasekhar (Chandrasekhar, 1931) showed that all white dwarfs have a maximum mass limit, the Chandrasekhar mass, which is approximately 1.4 solar masses. If the white dwarf is part of a binary system and the companion swells into a red giant, it may lose mass to the nearby white dwarf. It is also possible for two white dwarfs to collide, they will lose angular momentum through gravitational wave radiation. In either case, the white dwarf will obtain a higher density and temperature. If the white dwarf accretes material enough to exceed the Chandrasekhar mass, it becomes hot enough to burn the central part into $^{56}$Ni. However, due to the degenerate state, the extended energy production does not change the pressure and causing the star to expand, as in the usual case. Instead, the energy will raise the central temperature and hurry the reaction rates even further. This is, of course, an unstable condition called a thermonuclear runaway. The white dwarf will explode

and completely annihilate, leaving nothing behind. The fact that the white dwarfs always explode at a certain mass makes the processes just before the explosion always the same, leading to very similar light curves. They are therefore easy to recognize and will have about the same absolute magnitude, making them very good as standard candles. There is one major difference in the light–curves, though (Phillips, 1993; Hamuy et al, 1993). Phillips showed that there is a correlation between the peak magnitude and the magnitude measured 15 days after the peak ($\Delta m_{15}$). The longer the decay time (i.e. smaller $\Delta m_{15}$) the brighter the supernova is. Phillips made this observation based on a small set of supernovae, but the correlation was confirmed later on (Hamuy et al, 1995; Riess, Press & Kirshner, 1994). The origin of this difference in the lightcurves is not fully understood and is under investigation. The good part however, is that after the different decays have been taken into account, the dispersion of these modified magnitudes can be very small, making SN Ia a very good candidate for a standard candle. There is one problem left to deal with though. The supernovae is seen at a high redshift, up to $z \sim 1.5$. A rule of thumb is that an event at $z = 1$ is seen when the universe is one third of its present age. The SN Ia must therefore not be a subject of evolution or at least evolve in an easily fashion which can be predicted and accounted for. Luckily, the SN Ia do not seem to evolve (Goobar & Perlmutter, 1995; Branch et al, 2001). Since SN Ia is a single event it does not care about the time since Big Bang, it only cares about the state of its progenitor. SN Ia is also seen in all types of galaxies, with different amounts of dust. Both dust in the host galaxy and the existance of a dark energy dimmer the supernovae, leading to similar results. However, dust also reddens the colour of the supernovae and it is therefore possible to compare nearby and distant supernovae. The answer to the question if the host galaxy dust extinction, after correction, still does contribute to a systematic error is inconclusive. Several authors claim that the effect is negligible, while others believe that some supernovae are over- or undercorrected. A positive thing is that since SN Ia is single event it is also possible to discard a supernova if it is somehow peculiar. In summary, SN Ia seems to be an excellent choice as standard candle, hence the great interest in them recently.

For my calculations I have used the supernova measurements from the first year of five from the SuperNova Legacy Survey (SNLS) program taken with the MegaPrime/MegaCam wide-field imager located at Canada-France-Hawaii Telescope (CFHT). For a review on how the measurements and data reductions were done see (Astier et al, 2005). I have 115 clear type Ia super-

novae, which are well analysed and, very important, they are analysed in the same way. The fact that the supernovae are taken by they same telescope make the SNLS a very robust and reliable dataset. The supernova are well distributed in distance and cover the range $0.015 < z < 1.01$, with a median redshift of $z = 0.43$.

## 2.7 Additional datasets

In addition to the supernova measurements I am also including measurements from the Cosmic Microwave Background (CMB) and Baryonic Acoustic Oscillations (BAO). These measurements are sensitive to different combinations of the parameters than the supernovae and will therefore obtain different confidence contour, hopefully intersecting the confidence contour from the supernovae. The CMB and BAO are good as complementary measurements, but as such I will not get into much detail about these measurements, I will only give a quick review. Both these measurements provide an excellent addition to the supernova results, when combined they constrain the cosmological parameters much tighter than the use of supernovae alone ever could.

### 2.7.1 Baryonic acoustic oscillations

The early universe consisted of a homogeneous plasma of electrons, protons, neutrons, dark matter, photons and neutrinos. However, due to the nature of quantum mechanics, very small perturbations arise; a few particles come together, forming a slightly denser region. These perturbations will propagate in the plasma, forming a sound wave where the velocity is dependent on the medium (Amanullah, 2005). The sound speed for baryons is higher than that of dark matter, a perturbation in baryonic matter will quickly travel away. At decoupling, the sound speed of the baryonic medium drop and the perturbation will freeze in at a characteristic distance. The dark matter will determine the mass profile, but it will have a small bump at this characteristic distance. The effect of this in the present-day universe is that it exists a characteristic length scale, approximately 150 Mpc, of the large-scale structure seen today (Eisenstein et al, 2005). The Sloan Digital Sky Survey (SDSS) (Eisenstein et al, 2005) observe the large structures in the universe and is from that able to determine a Hubble-constant-independent distance

to $z_1 = 0.35$, a typical redshift of the SDSS sample, defined as

$$A = \frac{\sqrt{\Omega_m}}{H'(z_1)^{\frac{1}{3}}} \left[ \frac{1}{z_1 \sqrt{\mid \Omega_k \mid}} S\left( \sqrt{\mid \Omega_k \mid} \int_0^{z_1} \frac{\mathrm{d}z}{H'(z)} \right) \right]^{\frac{2}{3}} \tag{24}$$

where $H'(z)$ is given by equation (23), $S(x)$ is, as before, given by equation( 5) and the factors $\sqrt{\mid \Omega_k \mid}$ are removed for a flat universe. Eisenstein finds $A$ to be $A = 0.469 \pm 0.017$. This distance measure acts as a good standard ruler, which can then be used to calibrate other distances.

### 2.7.2 Cosmic microwave background

The cosmic microwave background (CMB) is one of the most studied phenomena in astronomy. When the early universe recombined, forming neutral atoms, it became totally transparent to photons. This occurred at $z_2 \approx 1089$ and since the last scattering, the photons have been travelling in all directions, including right towards us, where we measure them today. This is thermal radiation producing a perfect blackbody spectrum. Because of the expansion of the universe the wavelength of the photons is stretched; the associated temperature of this blackbody spectrum is today $\sim 2.7$ K. The CMB is isotropic i.e. it looks similar in all directions. This isotropic, perfect blackbody spectrum is a very strong argument for the Big Bang theory, since no other known procedure can produce such a spectrum. The CMB measurement (Spergel et al, 2003; Bennet et al, 2003) I use is the reduced distance to the surface of last scattering, the so-called shift parameter, determined by the Wilkinson Microwave Anisotropy Probe (WMAP) and defined as (Ichikawa & Takahashi, 2005; Knop et al, 2003)

$$I = \frac{\sqrt{\Omega_m}}{\sqrt{\mid \Omega_k \mid}} S\left( \sqrt{\mid \Omega_k \mid} \int_0^{z_2} \frac{\mathrm{d}z}{H'(z)} \right). \tag{25}$$

However, Spergel et al (2003) does not provide with a value for $I$. Several authors (Knop et al, 2003; Ichikawa & Takahashi, 2005; Wang & Mukherjee, 2004) do that and then refer to Spergel et al (2003), with no obvious way of how the value is obtained. After some research it became clear that the value is constructed from equation (25) using Spergel's best fit for the so-called WMAPext dataset i.e. they assume a flat universe with a cosmological constant and $\Omega_m = 0.24$. However, when I calculated the shift parameter I

came up with a different value than the other authors. I found that Knop et al (2003) took the value for $\Omega_m$ from the wrong table, hence using other datasets than they claim to do. The value Ichikawa & Takahashi (2005) and Wang & Mukherjee (2004) presents is only slightly different than my number; it is possible they have access to an additional decimal value, not presented in the article. Anyhow I use the number I calculated myself, $I = 1.714 \pm 0.062$. The shift parameter is sensitive to the deviation from a flat universe; by comparing the results of a flat universe with a non-flat model including CMB, we can get a good picture of the flatness of the universe.

# 3 Markov Chain Monte Carlo

## 3.1 Principle of Markov Chain Monte Carlo

The simplest and most straightforward way of determining the best fit of a set of parameters $\theta = (\theta_1, \theta_2, \ldots)$ is to make a grid of points over the entire parameter space. The $\chi^2$–value is then computed at each point and the best fit is easily obtained as the point with the lowest $\chi^2$–value. This, however, requires a lot of computer time, since the number of points grows exponentially with the number of parameters. An alternative is the Markov Chain Monte Carlo (MCMC) method, which is a very powerful tool and is becoming a standard technique in cosmology nowadays. It is also an old and well-known method. There are different kinds of MCMC:s, such as Gibbs sampling and the Metropolis-Hastings (MH) algorithm (Metropolis et al, 1953; Hastings, 1970). I have used, and will focus on, the latter one. The advantage of an MCMC compared to a regular grid is that the time required to compute the MCMC grows approximately linear with the number of parameters (Neal, 1993) instead of exponentially, as for the grid method. For a large number of parameters, which is often required in cosmology, the grid method is therefore not viable.

A MCMC is a statistical method of obtaining samples from a posterior distribution. A series of point is drawn that is only dependent on the previous point, it will accept steps in the "right" way with high probability and accept steps in the "wrong" way with lower probability. So the chain will do a walk through the $\chi^2$–surface, accept and reject steps and hopefully find the global minimum of all parameters. The basic idea of a MCMC is hence quite simple and can be outlined as follows (see also figure (2))

1. Generate and save the starting point; a set of random parameters $\theta^0 = (\theta^{10}, \theta^{20}, \ldots)$.

2. Calculate $\chi^2_0$, the $\chi^2$–value corresponding to the starting set of parameters.

3. Generate a new set of points, $\theta_{new}$, obtained by moving a small step $\Delta\theta$ from the previous point. The size of this step can be chosen in many different ways and is crucial for the efficiency of the chain. I will discuss the step size later.

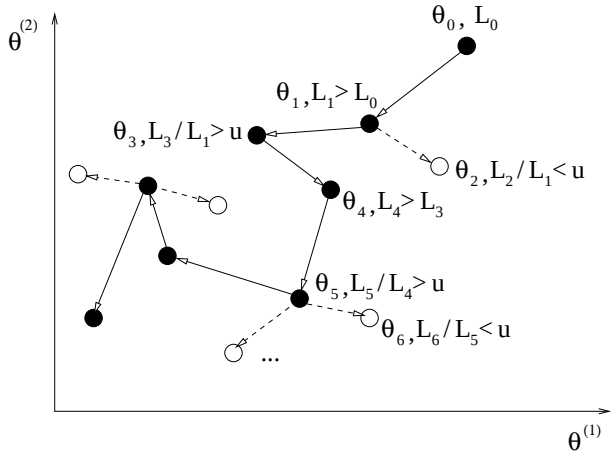



Figure 2: Illustrating the Metropolis algorithm for two parameters. Filled circles represent points belonging to the chain, empty circles are proposed but rejected points not belonging to the chain. In this example, the chain would be $[\theta_0, \theta_1, \theta_1, \theta_3, \theta_4, \ldots]$. The picture is taken from Doran & Müller (2004). Note here that the likelihood, $L$, is computed instead of $\chi^2$.

4. Calculate $\chi^2_{new}$, the $\chi^2$–value corresponding to the new set of parameters.

5. Calculate $\chi^2_0/\chi^2_{new}$.

   5a If $\chi^2_0/\chi^2_{new} \geq 1$ take the step. This means; take the new step to be the starting point (substitute $\theta_{new} \rightarrow \theta_0$), save the new point and go to 3.

   5b If $\chi^2_0/\chi^2_{new} < 1$ you still might take the step, with a certain probability. Draw a number $u$ in the region $[0, 1]$. If $u < \chi^2_0/\chi^2_{new}$ take the step according to 5a.

   5c If $u \geq \chi^2_0/\chi^2_{new}$ do not take the step. This means; discard $\theta_{new}$, save $\theta_0$ as an additional point at the same spot and go to 3.

6. Repeat step 3–5 until some condition is satisfied, i.e. until you believe you have points enough to reproduce the confidence contours.

The chain will approach regions of lower $\chi^2$–values, hopefully finding the global minimum. If it does it will stay in that region for most of the time

and not travel too far away from the region of interest. We say that the chain draws from the stationary distribution, the chain has converged. Since it is possible to take a step even if it goes in the "wrong" direction, the chain is prevented from getting stuck in a local minimum, at least if the $\chi^2$–well is not to steep. This is also making the chain exploring the surface around the whole region of interest, rather than staying in a small part of it. In fact, one property of a MCMC is that it is ergodic (Neal, 1993); it will visit every point on the $\chi^2$–surface, given an infinite time.

In practice there are some difficulties to overcome, which I will discuss in some detail below.

First, the initial steps will not be drawn from the stationary distribution. Since the starting point is randomly chosen it can take quite some time before the chain reaches the region of interest. An even bigger problem is that you do not know weather it has or not. The first few steps that are not drawn from the stationary distribution are in the literature called the burn-in period.

Second, and most important, is the choice of the step size. It is vital for the chain to be properly "mixed". Since we do not know the size or shape of the $\chi^2$–surface a priori, the chain must move quickly from one side of the distribution to the other. Otherwise a finite chain will miss a part of the region of interest, and there is no way to tell if it has or not. If the step size is too small, $\chi_0^2/\chi_{new}^2$ will be very close to 1 regardless of the direction and almost every step will be accepted. This will lead to a perfect random walk; the chain can go in every direction, probably not finding the global minimum. Even if it does, the chain will not stay there. It will also take a very long time for the chain to move from one side of the distribution to the other and we will experience poor mixing. On the other hand, if the step size is too large $\chi_0^2/\chi_{new}^2$ will be far from 1. Very few steps, only the ones in the "right" direction, will be accepted which will also lead to slow mixing. It is also possible to get stuck in a local minimum for a very long time. It is hence clear that there exist an optimum step size, which is not easily found.

The third problem is that of convergence. In order to trust the output results, one must know when the chain is drawing from the stationary distribution and also if it has visited the whole region if interest. This is not straightforward and is a very hard problem to solve. One must in some way come up with a numerical test, or more, to check for convergence. Fortunately, there exists a whole battery of such tests in the literature (Gelman & Rubin, 1992; Raftery & Lewis, 1992). A review of many of them is found in

Cowles & Carlin (1996).

The fourth and probably the easiest problem to solve is the length of the chain. If you stop it too soon it will not explore the whole target distribution. The chains must also consist of points enough to reproduce the confidence contours. The obvious solution is to run for as long as you can, weather it is a deadline or your patience that sets the limit.

## 3.2 My Markov Chain Monte Carlo

In this section I will describe how I made my MCMC and how I dealt with the above-mentioned problems.

I have tried to build my chain somewhat general. The user can change which parameters that shall be fixed at some value and the ones that will vary, this is defined through an input file. The user can also choose which datasets to use; SN, CMB or BAO or any combination of them. In the input file the user can also choose the number of chains to run simultaneously and the length of each individual chain. It is also possible to put constraints on the parameters; that is restricting the parameters from an unphysical or for any reason unwanted region. This is useful when you have a prior knowledge, for example that $\Omega_m > 0$. It is rather easy to add additional datasets and measurements in my MCMC as long as it not introduces additional parameters. If this is the case one has to do some extensive changes to the source code of my program.

To solve the first of the above mentioned problems, the burn–in, I have followed Tegmark et al (2004). I am running several chains, which is good for several reasons, and computes the overall median $\chi^2$–value combined for all chains. The length of the burn-in period for each chain is from point number one up to the point that is less than this median $\chi^2$–value. The burn-in period is removed before any other calculation is made. By doing this I probably discard a small number of interesting points, but that is unavoidable. By running a very long chain, this has a negligible effect.

The most important problem probably, the step size, is of course the hardest one to solve. The most common step function to draw from is the Gauss function i.e. $G(\Delta\theta) \propto e^{\Delta\theta^2/(2\sigma^2)}$. This is sufficiently good for many problems, but it is far from optimal. First of all, one has to determine the standard deviation of the Gauss function (i.e. the step size) which can be very different from parameter to parameter. There is also difficulties if the parameters are degenerate or if the $\chi^2$–surface takes peculiar, banana-shaped

forms (see figure (3)). To solve this I use an adaptive Gaussian sampling function (Doran & Müller, 2004; Tegmark et al, 2004; Haario et al, 2003; Sahlén, Liddle & Parkinson, 2005). This method determines the step size based on the previous points in the chain and it also take account for the degeneracy of the parameters. This method is as follows; I first calculate the covariance matrix

$$\mathbf{S} = \begin{pmatrix} \sigma_1^2 & \rho_{12} & \dots & \rho_{1N_{para}} \\ \rho_{21} & \sigma_2^2 & \dots & \rho_{2N_{para}} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N_{para}1} & \rho_{N_{para}2} & \dots & \sigma_{N_{para}}^2 \end{pmatrix} \tag{26}$$

from the points in the chain. $\sigma_i^2$ is the variance of parameter $i$, $\rho_{ij}$ is the covariance of parameters $i$ and $j$ and $N_{para}$ is the number of parameters. I then diagonalize this matrix to obtain the eigenvalues, $\lambda_j$, and the eigenvectors which I put together to form the transformation matrix $T$. Then I draw Gaussian samples, $\Delta\theta'$, with the computed eigenvalues $\lambda_j$ as standard deviations. Those are now the step size in the rotated system, i.e. the sampling distribution is now aligned with the $\chi^2$ surface. I then rotate back to the original coordinate system to obtain the wanted step size by the transformation

$$\Delta\theta = T\Delta\theta'. \tag{27}$$

For many occasions it is sufficient to run a test chain, calculate the covariance matrix of all the points in the test chain and then use that as input to the real chain. This is an effective method, but even this method may have problems where we have a banana-shaped target distribution. It is also not working in my case, where I am trying to build a general program. The covariance matrix is of course completely different when you change dataset or the parameters to vary. I therefore start with a random covariance matrix and then update it every 100 steps (Haario et al, 2003; Sahlén, Liddle & Parkinson, 2005; Tegmark et al, 2004). The starting covariance matrix $S_0$ is usually some best guess. Since my best guess would be different from run to run I start with a very loosely random covariance matrix, as mentioned. This is of course not optimal, but I have found empirically that it is working quite well. In order to improve the efficiency of the chain even further, especially in the case where the target distribution takes on banana or irregular shapes, I also scale the covariance matrix with a dynamic factor $\alpha$ (Doran & Müller, 2004). If the $\chi^2$–surface is irregular in some way, I want the chain to take

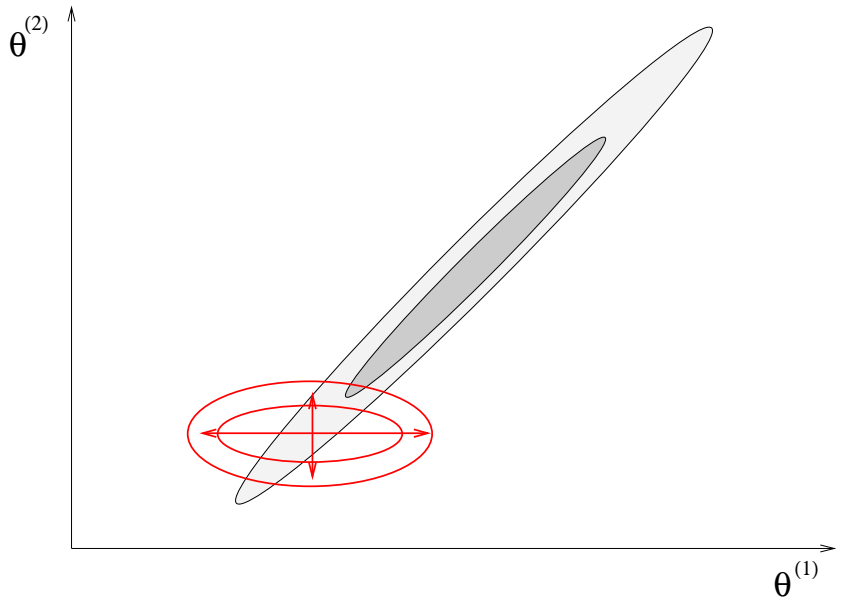



Figure 3: An ordinary Gaussian sampler does not take into account the degeneracy between the parameters $\theta^{(1)}$ and $\theta^{(2)}$ leading to slow mixing. The picture is taken from Doran & Müller (2004).

small steps in some areas and large in some. I initially set $\alpha = 1$ and before calculating the covariance matrix (every 100 steps) I check how many of the last 100 steps that has been accepted. If the acceptance rate of the last 100 steps in the chains is less than 20 % I scale $\alpha$ according to $\alpha = \alpha/1.2$. The covariance matrix will get smaller, smaller steps will be proposed, leading to increased acceptance rate. Similarly, if the chain accepts more than 70 % I increase $\alpha$ with $\alpha = \alpha \cdot 1.2$. Doran & Müller (2004) did not specify when or how much they scaled $\alpha$, but after some trial and error I how found my algorithm to be satisfactory. A problem is that the scaling of the covariance matrix changes the step size of all the parameters, which may of course not be optimal.

Now to the third problem, convergence. As previously mentioned there are a lot of available convergence tests. My primary weapon has been the most famous one, namely the Gelman–Rubin test (Gelman & Rubin, 1992). This test both look for converge and good mixing. Assume I am running $M$ chains and have ran all chains for $2N$ points. I consider only the last $N$ points of each chain, which are nere labelled $y_i^j$ where $i = 1, 2; \ldots, N$ and $j = 1, 2, \ldots, M$. I calculate the mean of each chain as

$$\overline{y}^j = \frac{1}{N} \sum_{i=1}^{N} y_i^j \tag{28}$$

and the mean of the whole set of chains as

$$\overline{y} = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} y_i^j. \tag{29}$$

Next, I calculate variances, both between the $M$ chains and within each chain as

$$B_n = \frac{1}{M-1} \sum_{j=1}^{M} (\overline{y}^j - \overline{y})^2 \tag{30}$$

and

$$W = \frac{1}{M(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{M} (y_i^j - \overline{y}^j)^2 \tag{31}$$

respectively. The actual test consists of monitoring the quantity $\hat{R}$, defined as

$$\hat{R} = \frac{\frac{N-1}{N} W + B_n(1 + \frac{1}{M})}{W}. \tag{32}$$

$\hat{R}$ goes asymptotically towards 1 and convergence is obtained if $\hat{R}$ is always below a number somewhat larger than 1, for all parameters. Gelman and Rubin themselves set the convergence criteria to be 1.2, but I choose the limit to be 1.1 as several other authors (Tegmark et al, 2004). Since $\hat{R}$ is approximately $1 + B_n/W$, a value near 1 simply states that the variances between the chains is negligible compared to the variances within the chain. That is, if the chains are indistinguishable it is hopefully because they all draw samples from the stationary distribution. By running several chains one is also able to look at them separately to see if something has gone wrong. If one or more chains is drawing from another part of the $\chi^2$–surface, then the chain has obviously not converged. It is not advisable to rely your result on only one chain (Neal, 1993).

To further check for convergence I am also calculating the autocorrelation of one, say $\theta^1$, of the parameters. Since each point in the chain is just a small step away from the previous one, they are highly correlated. So, a chain of $N$ points is not equivalently to $N$ different measurements. I therefore calculate the effective length, $N_{eff}$, which is approximately the number of unrelated points, that is; we have a chain that is equivalent to $N_{eff}$ different measurements (Tegmark et al, 2004). Of course, to obtain some meaningful results, $N_{eff}$ must be much larger than 1. I monitor $N_{eff}$ as an additional convergence test; a small $N_{eff}$ means the chain is not properly mixed (Gilks, Richardson & Spiegelhalter, 1996). The autocorrelation, $c_j$, is now defined as

$$c_j = \frac{\sum_{k=0}^{N-j-1}(\theta_k^1 - \overline{\theta}^1)(\theta_{k+j}^1 - \overline{\theta}^1)}{\sum_{k=0}^{N-1}(\theta_k^1 - \overline{\theta}^1)^2} \tag{33}$$

where $j$ is the lag i.e. the lag is specifying the signed distances between indexed elements of $\theta^1$. $c_0$ is by definition 1 and $c_j$ is then decreasing with increasing $j$. I define the correlation length, $c_l$, as the lag where the autocorrelation drops below 0.5. The correlation length is the number of steps required so that point $\theta_{i+c_l}$ is independent of point $\theta_i$. Then the effective length is

$$N_{eff} = \frac{N}{c_l}. \tag{34}$$

All of my chain, except maybe two, has a very long effective length, thus indicating that they are properly mixed.

To solve the fourth problem, which is not really a problem, I run quite long chains; usually 4 chains with 10 000 points in each chain. This is sufficient

to accurately reproduce the confidence contours.

Another problem I had was the construction of the confidence contours. To obtain the contour for two parameters you marginalize over all the others. The $\chi^2$–value of two points just next to each other may therefore be very different; it is dependent on the value of the other parameters. It is therefore not possible to plot the contours using their $\chi^2$–value; one must instead obtain them from the density of points. Finally I found and used a program by Sarah Bridle called cosmoloGUI available at:
http://www.sarahbridle.net/cosmologui/. The program is made in MATLAB and is very good at obtaining the confidence contours and displaying them in various ways.

## 3.3 Additional options and enhancements

There are a few more tricks available to enhance the efficiency of the MCMC. I have not incorporated them in my chain but I will discuss them here briefly, for the sake of the argument. Green & Mira (2001) suggest a method called delayed rejection. When rejecting a step, instead of staying at the original point they immediately propose a second candidate, which depends both on the original point and the first candidate. This can then be iterated either a fixed or a random number of times, before finally rejecting. Tierney & Mira (1999) proved that the delayed rejection method outperform the regular MH.

A second improvement, reparameterization, is efficient in cases of severe degeneracy or banana shaped distributions. By changing variables one may be able to completely remove or at least reduce the degeneracy. However, this may be hard or impossible to do, completely dependent on the science you are doing. In theory the adaptive covariance with scaling $\alpha$ method explained earlier also reduces degeneracy and normally only one of these two methods should be needed.

# 4 Results

I have made several simulations, each corresponding to a certain combination of models and datasets. I am considering the case of a cosmological constant as well as a time-dependent equation of state of the dark energy. For each of the models I have determined the best fit of each of the parameters, marginalizing over the others, as well as the confidence contours for two of the parameters at a time. As for the equation of state of the dark energy, I have parameterized it as given by equation (11). $w_0$ is the value of the equation of state today and $w_1$ describes its history. A non-zero $w_1$ imply that the equation of state is a subject of evolution and is hence more complex than a cosmological constant. This particular parameterization gives reasonable values in both cases $z \to 0$ and $z \to \infty$. Another common parameterization $w = w_0 + w_1 z$ does not give a reasonable value as $z \to \infty$. If I do not fit one of the parameters in a run it is fixed at a default value, which is specified in the input file mentioned earlier. For all my simulations the default values for the parameters are

$$\theta = (\Omega_m, \Omega_x, w_0, w_1) = (0.3, 0.7, -1, 0). \tag{35}$$

In some cases I force the universe to be flat, in that case $\Omega_x$ is always fixed at $\Omega_x = 1 - \Omega_m$. Each of my runs consists of 4 independent chains with an individual length of 10 000 steps. An advantage of the MCMC is that all the necessary information can be obtained from the chain itself. When I determine the best value for one parameter I again use the cosmoloGUI program. I plot the parameter against probability, essentially counting the number of points in a large number of bins. The program gives the approximate 68 % and 95 % percent interval for that parameter and the best value is at the top of this distribution, which for some reason is normalized to one. An example of such a probability plot is given in figure (4) below. I obtain the best value by reading from the figure; hence the last digit in this value should not be taken too seriously. In many cases the distribution had a wide and flat top and the best value was not clearly defined. In addition of knowing which parameter value is the best, one also want to know how good the fit of that value is. I quantify such a number known as goodness of fit (GoF), which is defined as (Hannestad & Mörtsell, 2004)

$$GoF = \frac{\Gamma(\nu/2,\ \chi^2/2)}{\Gamma(\nu/2)} \tag{36}$$

where $\Gamma(x)$ is the Gamma function, $\Gamma(x, y)$ is the incomplete gamma function and $\nu$ is the degrees of freedom defined as the number of measurements minus the fitted parameters. A GoF of, say 0.2, means that 20 % of some randomly generated measurements based on the same underlying model would give a worse fit. One is also interested to know if the addition of a varying parameter is making the fit better or worse. Generally, the addition of a parameter lowers the $\chi^2$–value, since there are more combinations of parameters to fit to the data. Therefore, when determining if a paramterer improves the model or not it is more useful to calculate another quantity; the Bayesian Information Criterion (BIC) which is defined as (Liddle, 2004)

$$BIC = \chi^2_{min} + N_{para} \ln N_{data} \tag{37}$$

where $N_{para}$ is the number of fitted parameters and $N_{data}$ is the number of measured data points. A lower value is a better fit, but the absolute value is not interesting, it is the difference in BIC between different models that is important. A model which has a BIC lower than 2 compared to another model is usually considered an improvement whereas a BIC lower than 6 is considered a strong improvement.

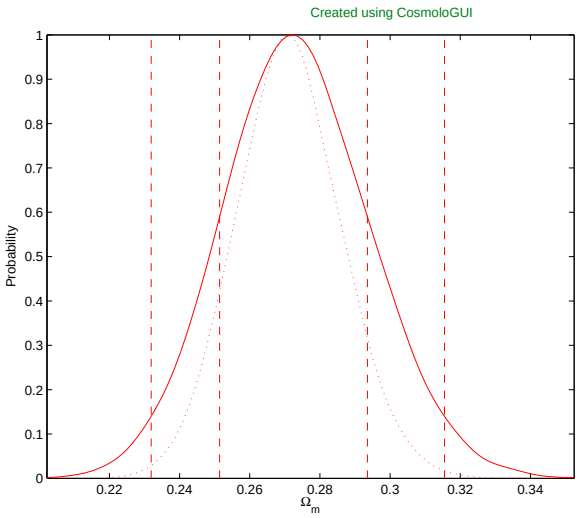



Figure 4: The solid line is the marginalized probability, which is the one I use for my result. The dotted line is the mean likelihood of the sample, which I am not using. See (Lewis & Bridle, 2002) for details.

## 4.1 A model with a cosmological constant

### 4.1.1 SN

I first consider a universe with a cosmological constant, i.e. I vary $\Omega_m$ and $\Omega_x$. $w_0$ and $w_1$ are fixed at their default values. My first run includes only the SN dataset and the confidence contours are shown in figure (5). The marginalized one-dimensional uncertainties and some other useful information about the chain itself are given in table (1). We see that the confidence contour is quite large when I use SN only.

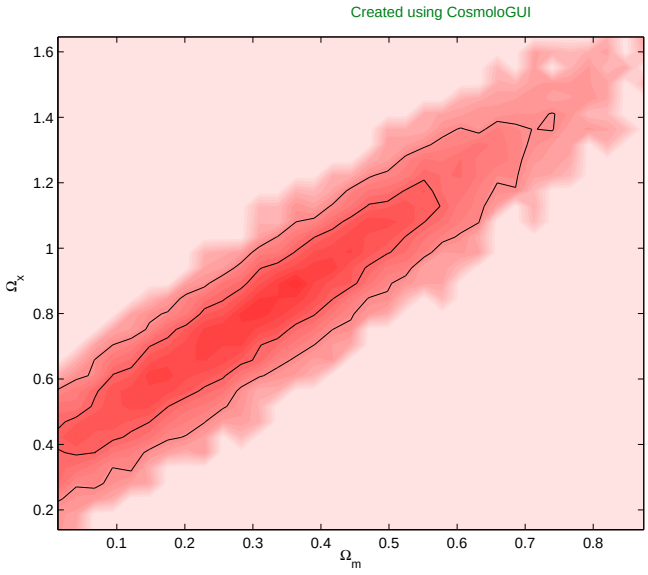



Figure 5: Confidence contours for a model with a cosmological constant. Only SN is used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.10 < \Omega_m < 0.48$ | 68 % | Convergence | 455 |
| | $0.005 < \Omega_m < 0.63$ | 95 % | Burn-in of chain 1 | 24 |
| | 0.30 | Best fit | Burn-in of chain 2 | 11 |
| $\Omega_x$ | $0.49 < \Omega_x < 1.09$ | 68 % | Burn-in of chain 3 | 0 |
| | $0.25 < \Omega_x < 1.37$ | 95 % | Burn-in of chain 4 | 24 |
| | 0.79 | Best fit | $c_l$ of chain 1 | 5 |
| $\chi^2_{min}$ | 112.18 | | $c_l$ of chain 2 | 5 |
| GoF | 0.48 | | $c_l$ of chain 3 | 4 |
| BIC | 126.41 | | $c_l$ of chain 4 | 5 |
| | | | Acc. rate of chain 1 | 0.44 |
| | | | Acc. rate of chain 2 | 0.44 |
| | | | Acc. rate of chain 3 | 0.44 |
| | | | Acc. rate of chain 4 | 0.44 |

Table 1: Results for a model with a cosmological constant. Only SN is used. Convergence is the iteration–number where the Gelman–Rubin test $\hat{R}$–value is below 1.1 for all parameters. The Burn-in is the number of discarded points not belonging to the stationary distribution. The correlation length, $c_l$, is the number of steps correlated to eachother. A small number implies good mixing. Acc. rate is the acceptance rate; the fraction of proposed steps that were accepted.

### 4.1.2 SN + BAO

I then combine the SN dataset with that of BAO. Since these two datasets has different kind of degeneracies, as mentioned, the confidence contour for this simulation is much smaller. The results are shown in figure (6) and table (2).

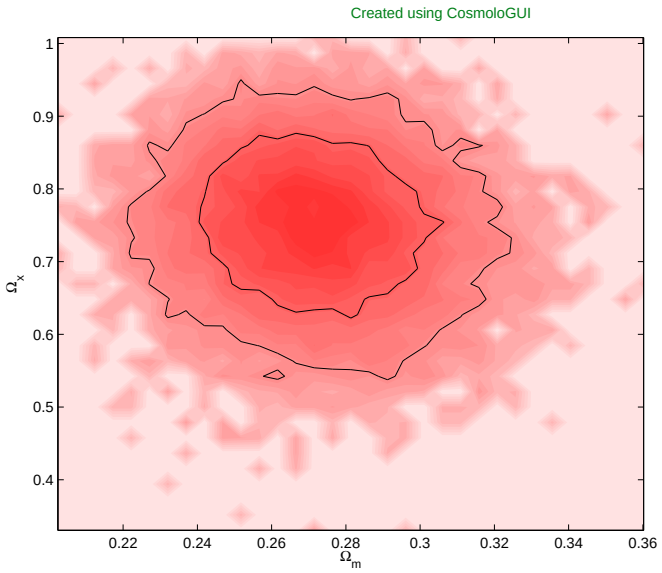



Figure 6: Confidence contours for a model with a cosmological constant. SN and BAO are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.30$ | 68 % | Convergence | 352 |
| | $0.22 < \Omega_m < 0.32$ | 95 % | Burn-in of chain 1 | 55 |
| | 0.27 | Best fit | Burn-in if chain 2 | 33 |
| $\Omega_x$ | $0.65 < \Omega_x < 0.86$ | 68 % | Burn-in of chain 3 | 19 |
| | $0.54 < \Omega_x < 0.96$ | 95 % | Burn-in of chain 4 | 49 |
| | 0.75 | Best fit | $c_l$ of chain 1 | 5 |
| $\chi^2_{min}$ | 112.22 | | $c_l$ of chain 2 | 5 |
| GoF | 0.50 | | $c_l$ chain of 3 | 5 |
| BIC | 126.48 | | $c_l$ chain of 4 | 5 |
| | | | Acc. rate of chain 1 | 0.45 |
| | | | Acc. rate of chain 2 | 0.45 |
| | | | Acc. rate of chain 3 | 0.44 |
| | | | Acc. rate of chain 4 | 0.44 |

Table 2: Results for a model with a cosmological constant. SN and BAO are used.

### 4.1.3 SN + CMB + BAO

I then present my best result for a non–flat universe with a cosmological constant; I combine all three datasets. The results are shown in figure (7) and table (3), respectively. The addition of the CMB dataset does not affect $\Omega_m$, but $\Omega_x$ is constrained more tightly.

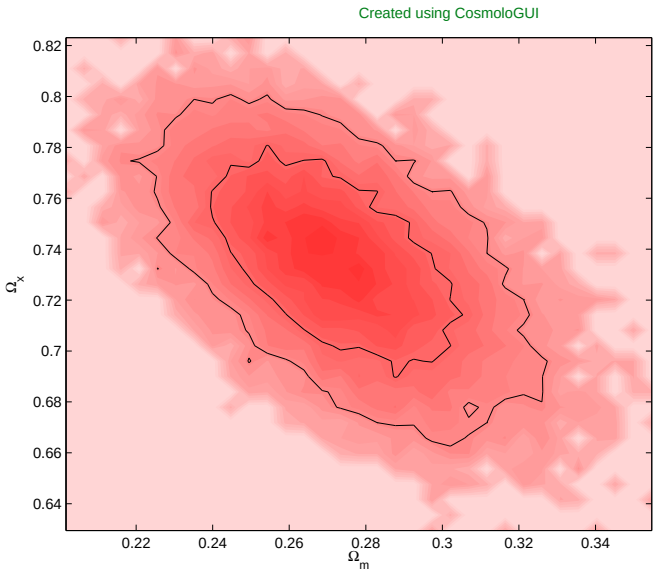



Figure 7: Confidence contours for a model with a cosmological constant. All datasets are used are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.30$ | 68 % | Convergence | 158 |
| | $0.22 < \Omega_m < 0.32$ | 95 % | Burn-in of chain 1 | 25 |
| | 0.27 | Best fit | Burn-in of chain 2 | 22 |
| $\Omega_x$ | $0.70 < \Omega_x < 0.77$ | 68 % | Burn-in chain of 3 | 80 |
| | $0.67 < \Omega_x < 0.80$ | 95 % | Burn-in of chain 4 | 61 |
| | 0.74 | Best fit | $c_l$ of chain 1 | 5 |
| $\chi^2_{min}$ | 112.28 | | $c_l$ of chain 2 | 5 |
| GoF | 0.53 | | $c_l$ of chain 3 | 5 |
| BIC | 126.57 | | $c_l$ of chain 4 | 5 |
| | | | Acc. rate of chain 1 | 0.48 |
| | | | Acc. rate of chain 2 | 0.48 |
| | | | Acc. rate of chain 3 | 0.50 |
| | | | Acc. rate of chain 4 | 0.49 |

Table 3: Results for a model with a cosmological constant. All datasets are used.

## 4.2 A flat model with a cosmological constant

### 4.2.1 SN + CMB + BAO

Here I now present the result from the simplest model of them all; flat universe with a cosmological constant. I will vary only $\Omega_m$. Since this simulation obtains a low $\chi^2$–value, it gets the lowest BIC–value of them all, implying that this is the best model for my data. Since the plot is forced to a straight line I will not present a figure showing the confidence contour here, I will only give the one–dimensional probability plot for $\Omega_m$ in figure (8) and give the additional results in table (4).

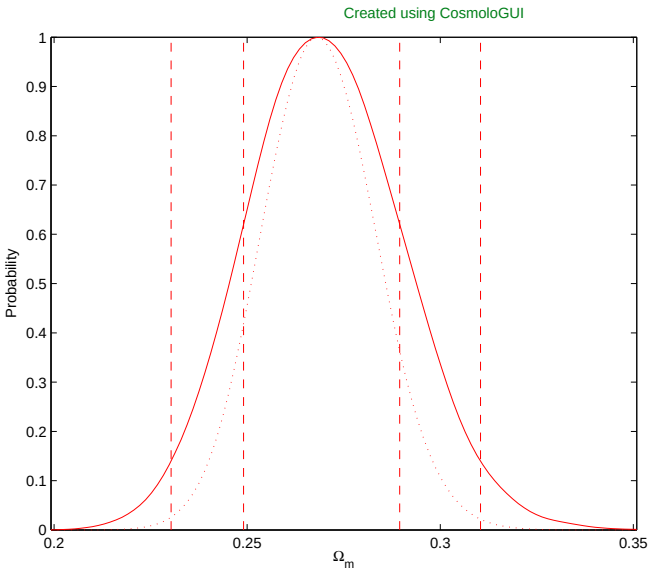



Figure 8: Probability plot for a flat model with a cosmological constant. All datasets are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.29$ | 68 % | Convergence | 138 |
| | $0.23 < \Omega_m < 0.31$ | 95 % | Burn-in of chain 1 | 56 |
| | 0.27 | Best fit | Burn-in of chain 2 | 24 |
| $\Omega_x$ | $0.71 < \Omega_x < 0.75$ | 68 % | Burn-in chain of 3 | 24 |
| | $0.69 < \Omega_x < 0.77$ | 95 % | Burn-in of chain 4 | 53 |
| | 0.73 | Best fit | $c_l$ of chain 1 | 4 |
| $\chi^2_{min}$ | 112.40 | | $c_l$ of chain 2 | 4 |
| GoF | 0.55 | | $c_l$ of chain 3 | 4 |
| BIC | 121.92 | | $c_l$ of chain 4 | 4 |
| | | | Acc. rate of chain 1 | 0.55 |
| | | | Acc. rate of chain 2 | 0.55 |
| | | | Acc. rate of chain 3 | 0.55 |
| | | | Acc. rate of chain 4 | 0.54 |

Table 4: Results for a flat model with a cosmological constant. All datasets are used.

## 4.3 A model with a constant equation of state

In the next case I the drop the assumption that the dark energy is a cosmological constant and could instead consist of something else. I am now varying $\Omega_m$ and $w_0$ whereas $w_1$ is still fixed, i.e. I do not let the equation of state evolve in time.

### 4.3.1 SN

In this scenario I am only including the SN dataset. Unfortunately, the SN dataset only is not able to constrain these three parameters; my chain has not converged at all. Hence the results of this simulation are not reliable. I therefore choose not to include the results of this simulation.

### 4.3.2 SN + BAO

By combining the dataset with BAO I am able to constrain $\Omega_m$ quite well. However, the addition of BAO does not constrain $\Omega_x$ or $w_0$ much and additional datasets is required if smaller confidence contours is wanted. The result of this simulation as shown below anyway. Since this simulation has a quite irregular distribution I ran 4 chain with 20 000 points each, i.e. twice as much as for a normal chain, in order to be sure that the whole region of interest was explored. Unfortunately, the improvement of this effort was minor. All four chains also have a large correlation length, $c_l$, indicating that the chains are not properly mixed.

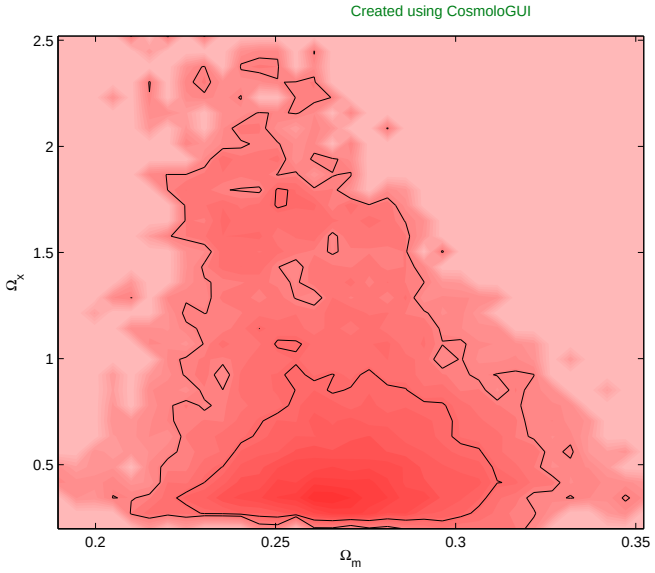



Figure 9: Confidence contours for a model with a constant equation of state. SN and BAO are used.

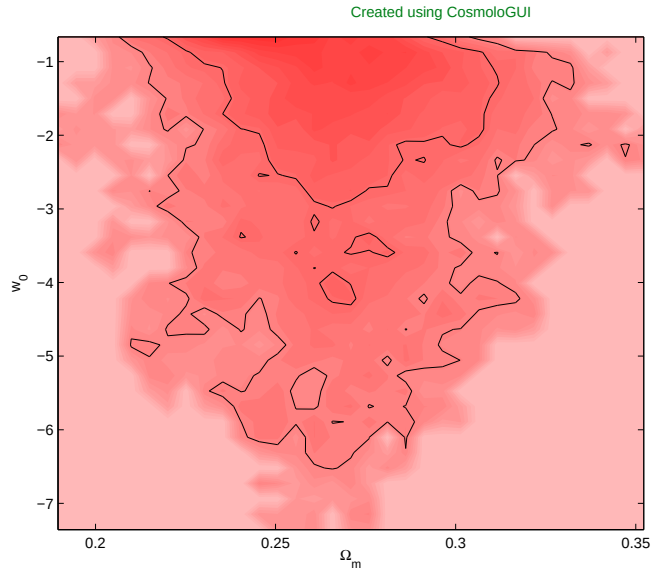



Figure 10: Confidence contours for a model with a constant equation of state. SN and BAO are used.

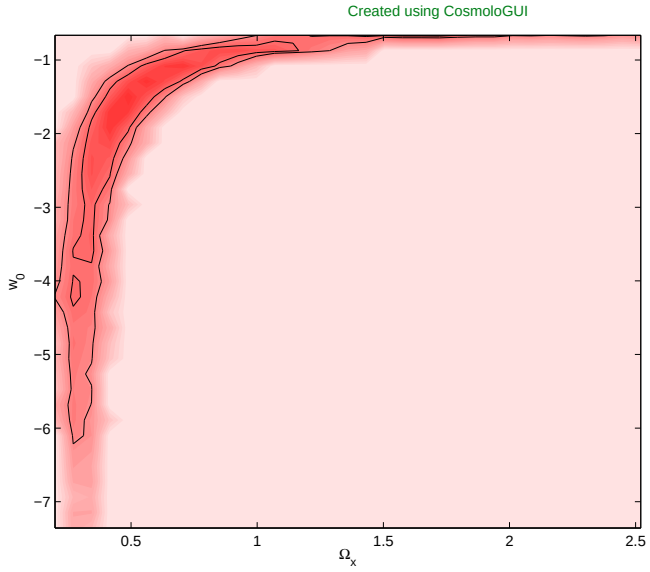



Figure 11: Confidence contours for a model with a constant equation of state. SN and BAO are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.24 < \Omega_m < 0.29$ | 68 % | Convergence | 708 |
| | $0.22 < \Omega_m < 0.31$ | 95 % | Burn-in of chain 1 | 69 |
| | 0.27 | Best fit | Burn-in of chain 2 | 41 |
| $\Omega_x$ | $0.21 < \Omega_x < 0.73$ | 68 % | Burn-in of chain 3 | 11 |
| | $0.17 < \Omega_x < 1.81$ | 95 % | Burn-in of chain 4 | 55 |
| | 0.37 | Best fit | $c_l$ of chain 1 | 49 |
| $w_0$ | $-2.16 < w_0 < -0.60$ | 68 % | $c_l$ of chain 2 | 49 |
| | $-4.96 < w_0 < -0.60$ | 95 % | $c_l$ of chain 3 | 49 |
| | -0.70 | Best fit | $c_l$ of chain 4 | 49 |
| $\chi^2_{min}$ | 112.15 | | Acc. rate of chain 1 | 0.27 |
| GoF | 0.48 | | Acc. rate of chain 2 | 0.33 |
| BIC | 131.16 | | Acc. rate of chain 3 | 0.27 |
| | | | Acc. rate of chain 4 | 0.33 |

Table 5: Confidence contours for a model with a constant equation of state. SN and BAO are used.

### 4.3.3 SN + CMB + BAO

When I now include all three datasets in order to obtain my best result for the time-independent dark energy universe, the confidence contours gets smaller. This time all of the three parameters are well constrained, as we can see in figures (12), (13) and (14) below. Other results are shown in table (6), as usual. Figure (12) is almost identical to figure (7), indicating that letting $w_0$ loose has no effect on $\Omega_m$ and $\Omega_x$.

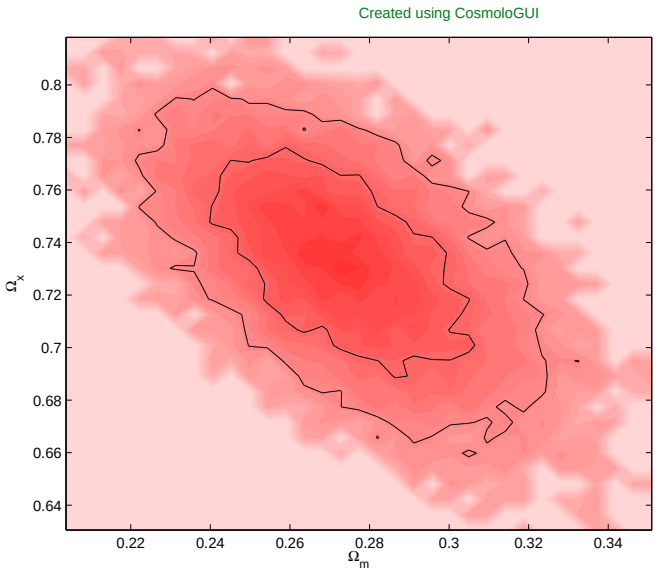



Figure 12: Confidence contours for a model with a constant equation of state. All datasets are used.

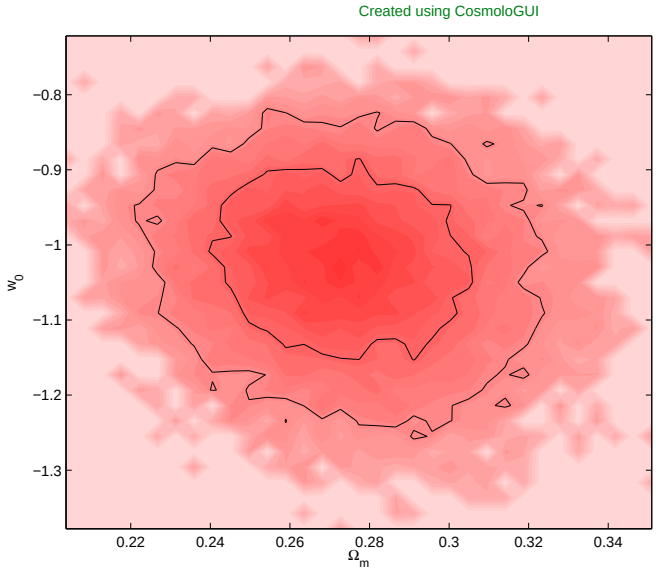



Figure 13: Confidence contours for a model with a constant equation of state. All datasets are used.

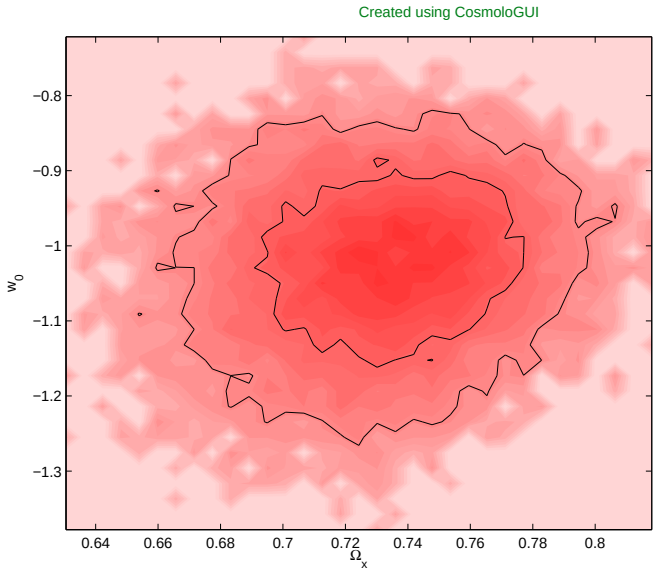



Figure 14: Confidence contours for a model with a constant equation of state. All datasets are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.29$ | 68 % | Convergence | 494 |
| | $0.23 < \Omega_m < 0.32$ | 95 % | Burn-in of chain 1 | 208 |
| | 0.27 | Best fit | Burn-in of chain 2 | 203 |
| $\Omega_x$ | $0.71 < \Omega_x < 0.76$ | 68 % | Burn-in of chain 3 | 235 |
| | $0.68 < \Omega_x < 0.79$ | 95 % | Burn-in of chain 4 | 16 |
| | 0.73 | Best fit | $c_l$ of chain 1 | 6 |
| $w_0$ | $-1.11 < w_0 < -0.94$ | 68 % | $c_l$ of chain 2 | 6 |
| | $-1.20 < w_0 < -0.86$ | 95 % | $c_l$ of chain 3 | 6 |
| | -1.0 | Best fit | $c_l$ of chain 4 | 6 |
| $\chi^2_{min}$ | 112.24 | | Acc. rate of chain 1 | 0.39 |
| GoF | 0.50 | | Acc. rate of chain 2 | 0.38 |
| BIC | 131.29 | | Acc. rate of chain 3 | 0.39 |
| | | | Acc. rate of chain 4 | 0.38 |

Table 6: Confidence contours for a model with a constant equation of state. All datasets are used.

## 4.4 A flat model with a constant equation of state

I also see how the parameters behave when I am forcing the universe to be flat. I am now only varying $\Omega_m$ and $w_0$. Since the CMB is sensitive to the curvature of the universe it should not be needed to include that dataset when looking at a flat universe. However, I will also see what happens if I assume a flat universe and include CMB anyway. Will the CMB data improve the results or is it completely equivalent of assuming a flat universe? For these simulations I will only show the plots containing $\Omega_m$ and $w_0$ here; the plot showing $\Omega_m$ and $\Omega_x$ is locked at a straight line and is therefore not so interesting. The plot of $\Omega_x$ and $w_0$ is just the reflection of the $\Omega_m - w_0$ plot and is therefore unnecessary too.

### 4.4.1 SN

Here I am only including SN. The results are shown below in figure (15) and table (7).

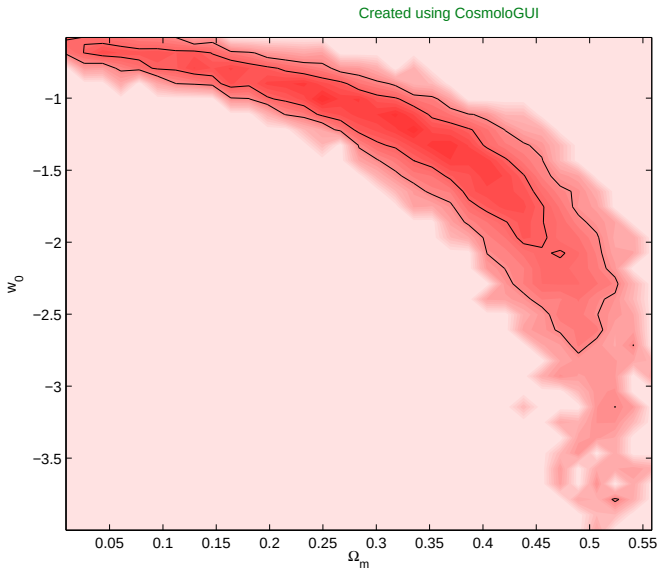



Figure 15: Confidence contours for a flat model with a constant equation of state. Only SN is used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.23 < \Omega_m < 0.48$ | 68 % | Convergence | 913 |
| | $0.05 < \Omega_m < 0.52$ | 95 % | Burn-in of chain 1 | 98 |
| | 0.40 | Best fit | Burn-in of chain 2 | 42 |
| $\Omega_x$ | $0.52 < \Omega_x < 0.77$ | 68 % | Burn-in of chain 3 | 76 |
| | $0.48 < \Omega_x < 0.95$ | 95 % | Burn-in of chain 4 | 44 |
| | 0.60 | Best fit | $c_l$ of chain 1 | 19 |
| $w_0$ | $-1.58 < w_0 < -0.65$ | 68 % | $c_l$ of chain 2 | 20 |
| | $-2.32 < w_0 < -0.54$ | 95 % | $c_l$ of chain 3 | 28 |
| | -1.0 | Best fit | $c_l$ of chain 4 | 21 |
| $\chi^2_{min}$ | 112.21 | | Acc. rate of chain 1 | 0.23 |
| GoF | 0.48 | | Acc. rate of chain 2 | 0.23 |
| BIC | 126.44 | | Acc. rate of chain 3 | 0.22 |
| | | | Acc. rate of chain 4 | 0.23 |

Table 7: Confidence contours for a flat model with a constant equation of state. Only SN is used.

### 4.4.2 SN + BAO

As usual I am now including BAO to see what happens. The results are shown below. This figure looks very much like figure (13), in this case implying a flat universe is equivalent with including the CMB dataset.

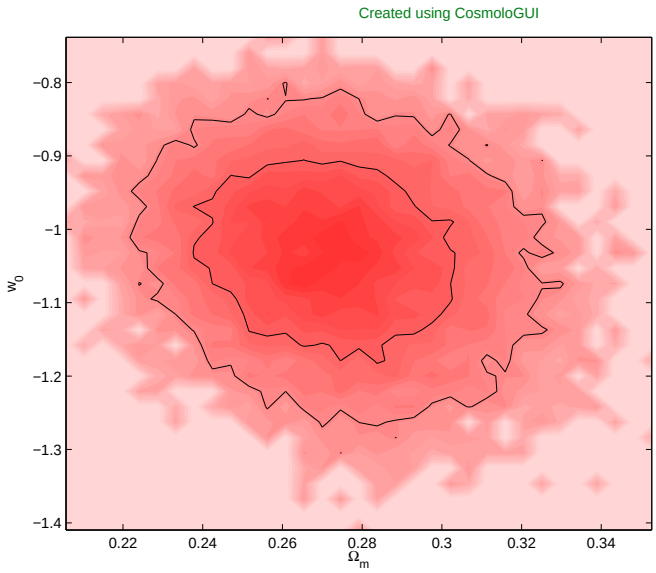



Figure 16: Confidence contours for a flat model with a constant equation of state. SN and BAO are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.29$ | 68 % | Convergence | 505 |
| | $0.23 < \Omega_m < 0.31$ | 95 % | Burn-in of chain 1 | 71 |
| | 0.27 | Best fit | Burn-in of chain 2 | 8 |
| $\Omega_x$ | $0.71 < \Omega_x < 0.75$ | 68 % | Burn-in of chain 3 | 84 |
| | $0.69 < \Omega_x < 0.77$ | 95 % | Burn-in of chain 4 | 114 |
| | 0.73 | Best fit | $c_l$ of chain 1 | 5 |
| $w_0$ | $-1.12 < w_0 < -0.94$ | 68 % | $c_l$ of chain 2 | 4 |
| | $-1.22 < w_0 < -0.86$ | 95 % | $c_l$ of chain 3 | 4 |
| | -1.0 | Best fit | $c_l$ of chain 4 | 5 |
| $\chi^2_{min}$ | 112.22 | | Acc. rate of chain 1 | 0.44 |
| GoF | 0.50 | | Acc. rate of chain 2 | 0.45 |
| BIC | 126.49 | | Acc. rate of chain 3 | 0.45 |
| | | | Acc. rate of chain 4 | 0.44 |

Table 8: Confidence contours for a flat model with a constant equation of state. SN and BAO are used.

### 4.4.3 SN + CMB + BAO

Now I include CMB even though I imply a flat universe to see if any improvement is obtained relative to the preceding figure. The plots including $\Omega_x$ are not shown. Since the difference between figure (17) and (16) is minimal we see that CMB is equivalent with implying a flat universe, at least when looking at the time-independent case.

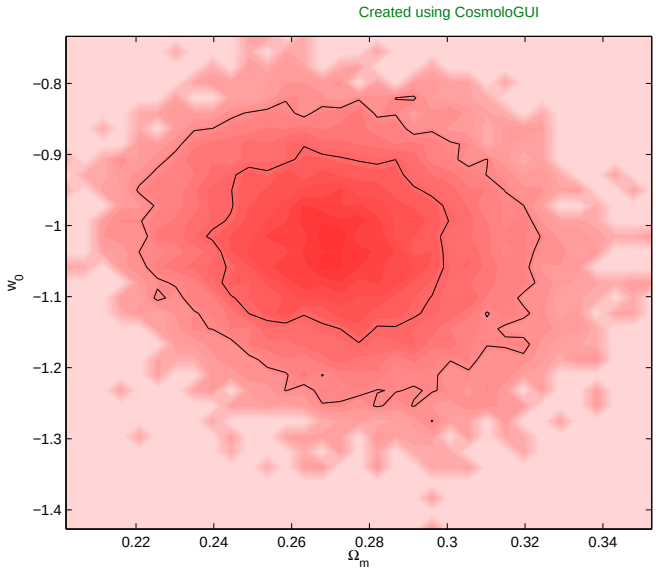



Figure 17: Confidence contours for a flat model with a constant equation of state. All datasets are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.29$ | 68 % | Convergence | 351 |
| | $0.23 < \Omega_m < 0.31$ | 95 % | Burn-in of chain 1 | 22 |
| | 0.27 | Best fit | Burn-in of chain 2 | 207 |
| $\Omega_x$ | $0.71 < \Omega_x < 0.75$ | 68 % | Burn-in of chain 3 | 36 |
| | $0.69 < \Omega_x < 0.77$ | 95 % | Burn-in of chain 4 | 65 |
| | 0.73 | Best fit | $c_l$ of chain 1 | 5 |
| $w_0$ | $-1.11 < w_0 < -0.94$ | 68 % | $c_l$ of chain 2 | 4 |
| | $-1.20 < w_0 < -0.86$ | 95 % | $c_l$ of chain 3 | 5 |
| | -1.0 | Best fit | $c_l$ of chain 4 | 5 |
| $\chi^2_{min}$ | 112.34 | | Acc. rate of chain 1 | 0.45 |
| GoF | 0.53 | | Acc. rate of chain 2 | 0.44 |
| BIC | 126.63 | | Acc. rate of chain 3 | 0.43 |
| | | | Acc. rate of chain 4 | 0.44 |

Table 9: Confidence contours for a flat model with a constant equation of state. All datasets are used.

## 4.5 A model with a time-dependent equation of state

Now I am making the full simulation with a time-dependent equation of state. Both $w_0$ and $w_1$ are varying. For this set of parameters it is not possible to obtain reasonable constraints of SN data only, all my simulations here therefore include CMB, BAO or both.

### 4.5.1 SN + BAO

I am first considering the flat universe case; the use of the CMB dataset is therefore not needed. The result of this simulation is shown below; in figures (18), (19), (20) and table (10). As before I do not present plots including $\Omega_x$ since it is the reversion of $\Omega_m$. In this case (compared to the previous one) the intervall for $w_0$ gets significantly larger. It is usually very hard to put constraints on the parameters when I let $w_1$ free and the possible time-variation of $\Omega_x$ is hard to find.

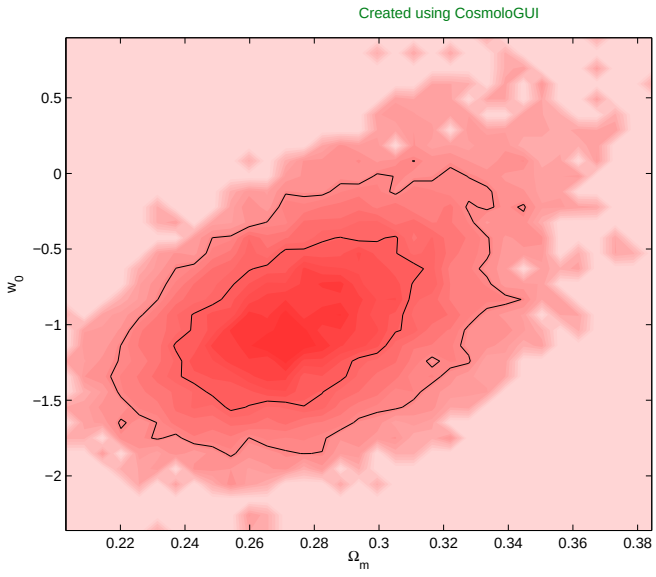



Figure 18: Confidence contours for a flat model with a time-dependent equation of state. SN and BAO are used.

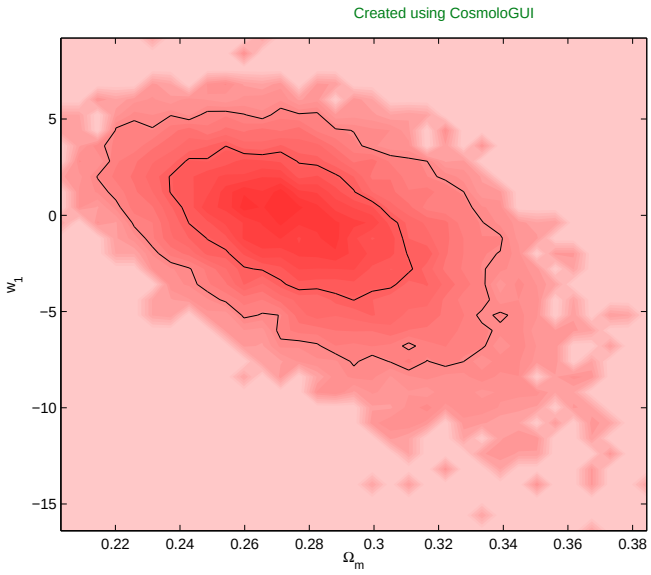



Figure 19: Confidence contours for a flat model with a time-dependent equation of state. SN and BAO are used.

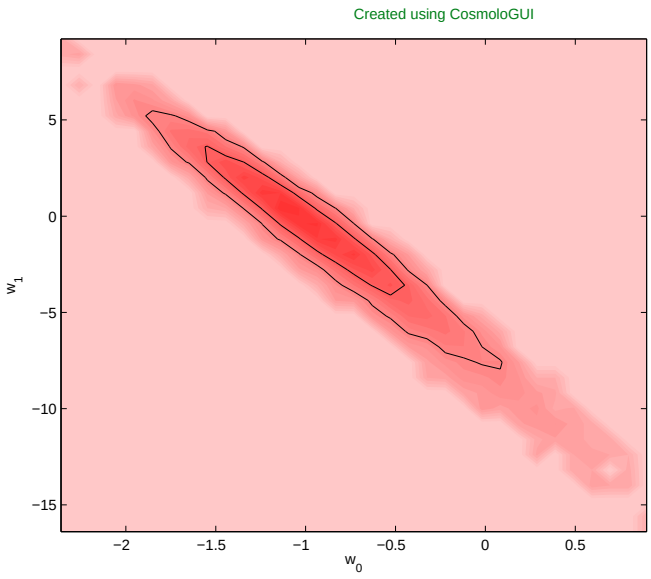



Figure 20: Confidence contours for a flat model with a time-dependent equation of state. SN and BAO are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.30$ | 68 % | Convergence | 440 |
| | $0.23 < \Omega_m < 0.33$ | 95 % | Burn-in of chain 1 | 26 |
| | 0.27 | Best fit | Burn-in of chain 2 | 47 |
| $\Omega_x$ | $0.70 < \Omega_x < 0.75$ | 68 % | Burn-in of chain 3 | 85 |
| | $0.67 < \Omega_x < 0.77$ | 95 % | Burn-in of chain 4 | 22 |
| | 0.73 | Best fit | $c_l$ of chain 1 | 5 |
| $w_0$ | $-1.37 < w_0 < -0.61$ | 68 % | $c_l$ of chain 2 | 6 |
| | $-1.73 < w_0 < -0.18$ | 95 % | $c_l$ of chain 3 | 6 |
| | -1.1 | Best fit | $c_l$ of chain 4 | 6 |
| $w_1$ | $-2.87 < w_1 < 2.38$ | 68 % | Acc. rate of chain 1 | 0.37 |
| | $-6.21 < w_1 < 4.71$ | 95 % | Acc. rate of chain 2 | 0.38 |
| | 0.0 | Best fit | Acc. rate of chain 3 | 0.37 |
| $\chi^2_{min}$ | 112.23 | | Acc. rate of chain 4 | 0.37 |
| GoF | 0.48 | | | |
| BIC | 131.24 | | | |

Table 10: Confidence contours for a flat model with a time-dependent equation of state. SN and BAO are used.

### 4.5.2 SN + CMB + BAO

Here I run a simulation for a flat universe but includes CMB anyway. As before, the reason for this is to see if CMB improve the results any further. The result is as always seen below; figures (21), (22), (23) and table (11). The plots including $\Omega_x$ are not shown, as always for a flat universe. This time we see a clear difference relative to figures (18)–(20); especially for $w_1$. Now the CMB dataset improves the results. When investigating the time-dependent case CMB is not equivalent to implying a flat universe.

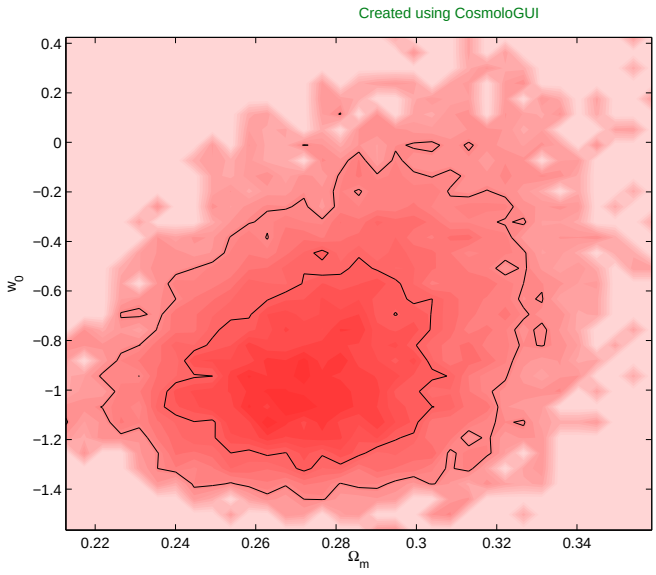



Figure 21: Confidence contours for a flat model with a time-dependent equation of state. All datasets are used.

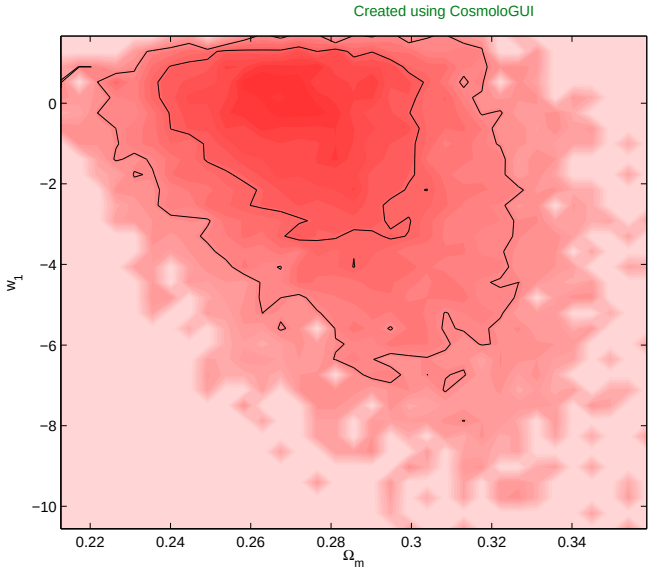



Figure 22: Confidence contours for a flat model with a time-dependent equation of state. All datasets are used.

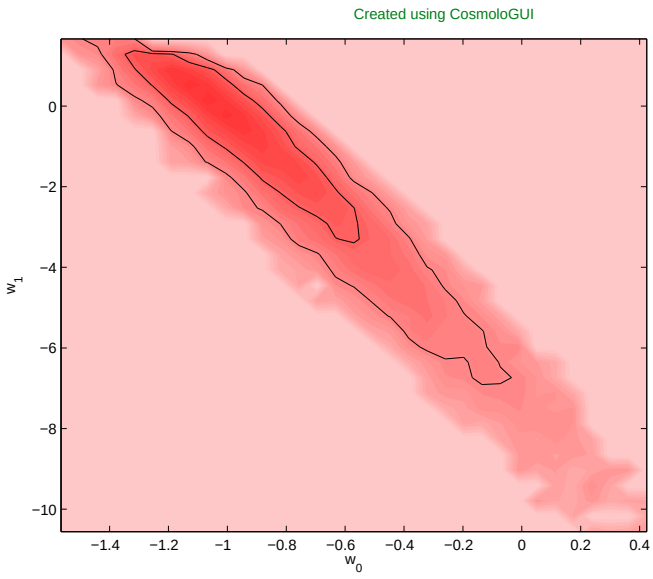



Figure 23: Confidence contours for a flat model with a time-dependent equation of state. All datasets are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.25 < \Omega_m < 0.30$ | 68 % | Convergence | 648 |
| | $0.23 < \Omega_m < 0.32$ | 95 % | Burn-in of chain 1 | 16 |
| | 0.27 | Best fit | Burn-in of chain 2 | 33 |
| $\Omega_x$ | $0.70 < \Omega_x < 0.75$ | 68 % | Burn-in of chain 3 | 100 |
| | $0.68 < \Omega_x < 0.77$ | 95 % | Burn-in of chain 4 | 59 |
| | 0.73 | Best fit | $c_l$ of chain 1 | 6 |
| $w_0$ | $-1.22 < w_0 < -0.67$ | 68 % | $c_l$ of chain 2 | 6 |
| | $-1.38 < w_0 < -0.28$ | 95 % | $c_l$ of chain 3 | 6 |
| | -1.0 | Best fit | $c_l$ of chain 4 | 6 |
| $w_1$ | $-2.07 < w_1 < 1.20$ | 68 % | Acc. rate of chain 1 | 0.33 |
| | $-5.00 < w_1 < 1.80$ | 95 % | Acc. rate of chain 2 | 0.34 |
| | 0.2 | Best fit | Acc. rate of chain 3 | 0.34 |
| $\chi^2_{min}$ | 112.29 | | Acc. rate of chain 4 | 0.33 |
| GoF | 0.50 | | | |
| BIC | 131.34 | | | |

Table 11: Confidence contours for a flat model with a time-dependent equation of state. All datasets are used.

### 4.5.3 SN + CMB + BAO

Here I am even dropping the flat universe prior and let all parameters vary, including all the datasets. I am doubling the lengths of the chains to 20 000 in this simulation. As seen in the pictures below some of the parameters has a "tail" which was probably not fully explored by all four chains when they were 10 000 points long. Since this is the "full" simulation were I include everything I got, I also believed this run was worth some extra time. The contours for $\Omega_m$ and $w_0$ looks quite the same as in the previous simulation.

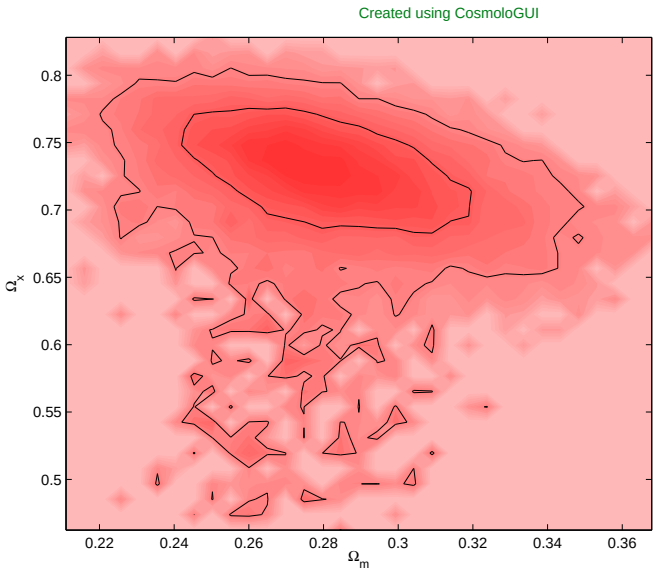



Figure 24: Confidence contours for a model with a time-dependent equation of state. All datasets are used.

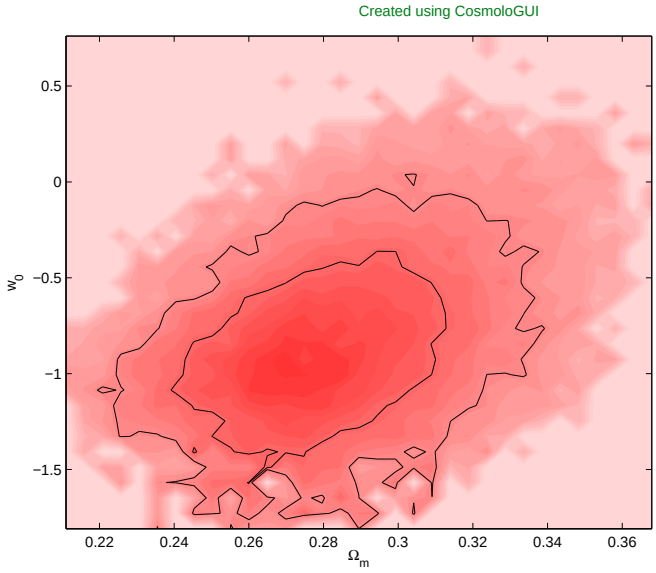



Figure 25: Confidence contours for a model with a time-dependent equation of state. All datasets are used.

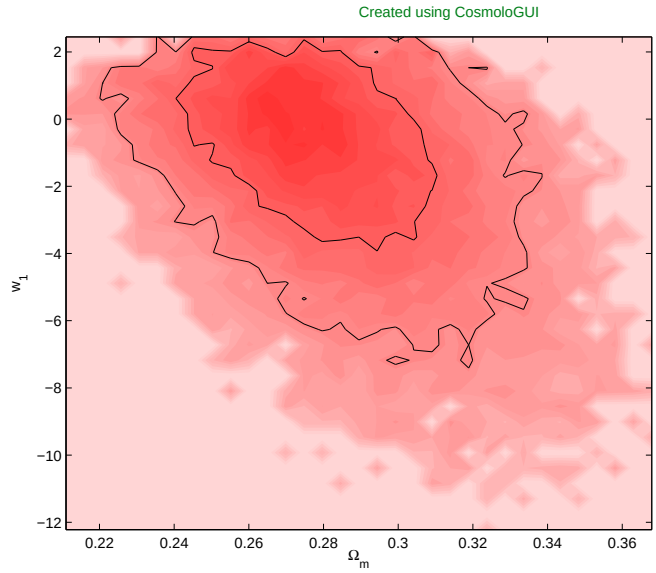



Figure 26: Confidence contours for a model with a time-dependent equation of state. All datasets are used.

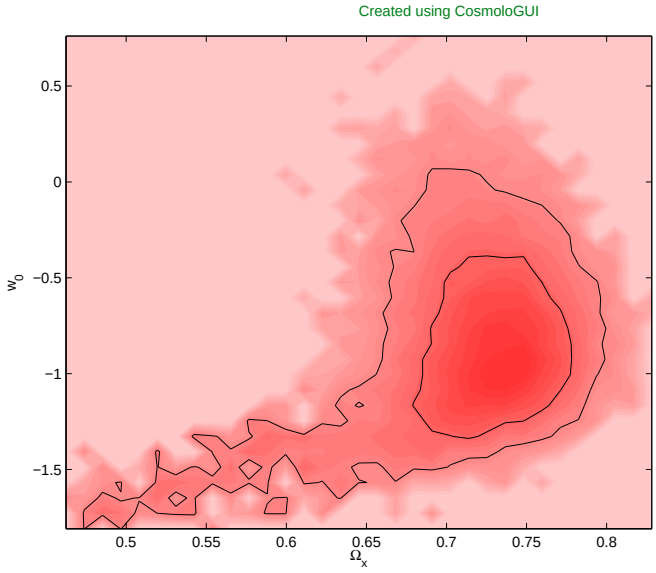



Figure 27: Confidence contours for a model with a time-dependent equation of state. All datasets are used.

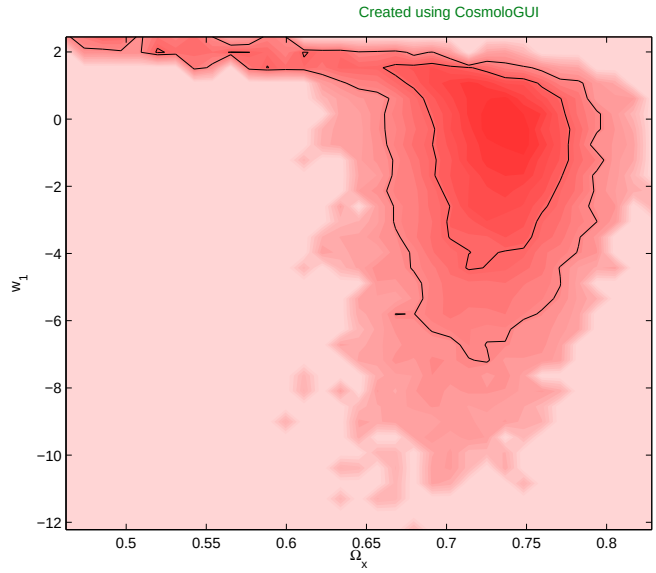



Figure 28: Confidence contours for a model with a time-dependent equation of state. All datasets are used.

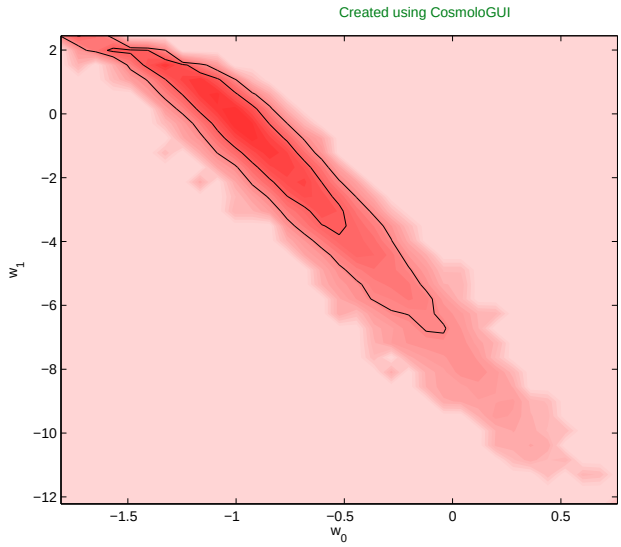



Figure 29: Confidence contours for a model with a time-dependent equation of state. All datasets are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.26 < \Omega_m < 0.30$ | 68 % | Convergence | 806 |
| | $0.23 < \Omega_m < 0.33$ | 95 % | Burn-in of chain 1 | 59 |
| | 0.27 | Best fit | Burn-in of chain 2 | 134 |
| $\Omega_x$ | $0.69 < \Omega_x < 0.76$ | 68 % | Burn-in of chain 3 | 39 |
| | $0.53 < \Omega_x < 0.80$ | 95 % | Burn-in of chain 4 | 154 |
| | 0.73 | Best fit | $c_l$ of chain 1 | 10 |
| $w_0$ | $-1.25 < w_0 < -0.58$ | 68 % | $c_l$ of chain 2 | 10 |
| | $-1.62 < w_0 < -0.20$ | 95 % | $c_l$ of chain 3 | 11 |
| | -1.0 | Best fit | $c_l$ of chain 4 | 10 |
| $w_1$ | $-2.38 < w_1 < 1.67$ | 68 % | Acc. rate of chain 1 | 0.29 |
| | $-5.06 < w_1 < 2.60$ | 95 % | Acc. rate of chain 2 | 0.31 |
| | -0.2 | Best fit | Acc. rate of chain 3 | 0.32 |
| $\chi^2_{min}$ | 112.27 | | Acc. rate of chain 4 | 0.31 |
| GoF | 0.48 | | | |
| BIC | 136.08 | | | |

Table 12: Confidence contours for a model with a constant equation of state. All datasets are used.

## 4.6 Future results

As an additional bonus I have made runs with simulated datasets. This is just to get a hunch of what the future holds. I have simulated 1200 supernovae; this is what we could expect from SLNS and other SN Ia search programs in five years. These simulated data are constructed to give the result $\Omega_m = 0.3, \Omega_x = 0.7, w_0 = -1$ and $w_1 = 0$. Therefore the results themselves are not important, it is the size of the confidence contours that is interesting. I have also simulated a result which the Planck satellite may measure; the shift parameter is $I = 1.74969 \pm 0.7\%$. BAO will give results for three different redshifts; 0.35, 1.0 and 3.0. For these the according distances are $0.486831 \pm 3.6\%$, $0.381531 \pm 0.67\%$ and $0.208112 \pm 1\%$. Since the magnitude of these simulated supernovae are normalized in a different way than the real ones, I do not get a $\chi^2$–value of the same magnitude. Hence the GoF value does not make sense and it is not meaningful to compare the BIC–value of these simulations with the real ones, it is only meaningful to compare the BIC–values of future results with each other.

### 4.6.1 SN

Just for comparison I am running a universe with a cosmological constant and the SN data only. This is just to compare the size of the confidence contours with the well-known result in figure (5), here they are much smaller.

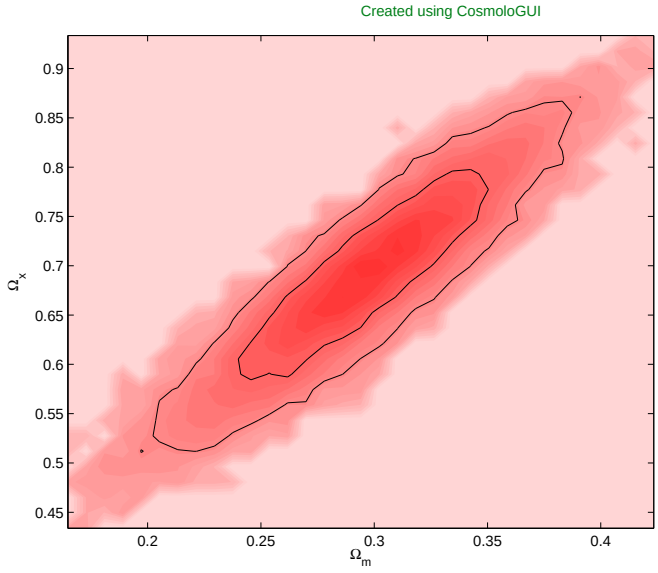



Figure 30: Confidence contours for a model with a cosmological constant. Only the simulated future SN is used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.26 < \Omega_m < 0.34$ | 68 % | Convergence | 458 |
| | $0.22 < \Omega_m < 0.37$ | 95 % | Burn-in of chain 1 | 39 |
| | 0.30 | Best fit | Burn-in of chain 2 | 7 |
| $\Omega_x$ | $0.62 < \Omega_x < 0.77$ | 68 % | Burn-in of chain 3 | 76 |
| | $0.54 < \Omega_x < 0.83$ | 95 % | Burn-in of chain 4 | 214 |
| | 0.70 | Best fit | $c_l$ of chain 1 | 7 |
| $\chi^2_{min}$ | 4.81 | | $c_l$ of chain 2 | 7 |
| GoF | — | | $c_l$ of chain 3 | 6 |
| BIC | 26.08 | | $c_l$ of chain 4 | 6 |
| | | | Acc. rate of chain 1 | 0.57 |
| | | | Acc. rate of chain 2 | 0.58 |
| | | | Acc. rate of chain 3 | 0.57 |
| | | | Acc. rate of chain 4 | 0.56 |

Table 13: Results for a model with a cosmological constant. Only the simulated future SN is used. I was not able the compute GoF for such a large number of datapoints.

### 4.6.2 SN + BAO + CMB

Now I am running the whole lot. All parameters with all datasets included. The results are shown below in figure (31)–(36) and table (14). This is the kind of results one is expecting in five years. By looking at this plot one can decide whether it is worth the money or not.

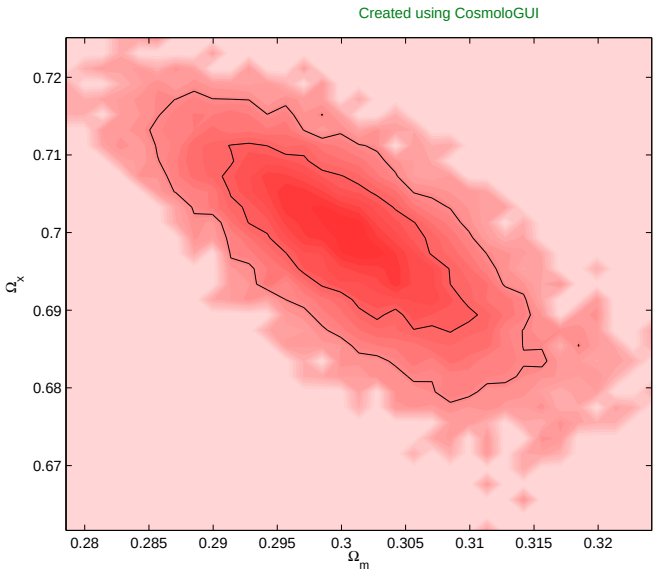



Figure 31: Confidence contours for a model with a time-dependent equation of state. All simulated future datasets are used.

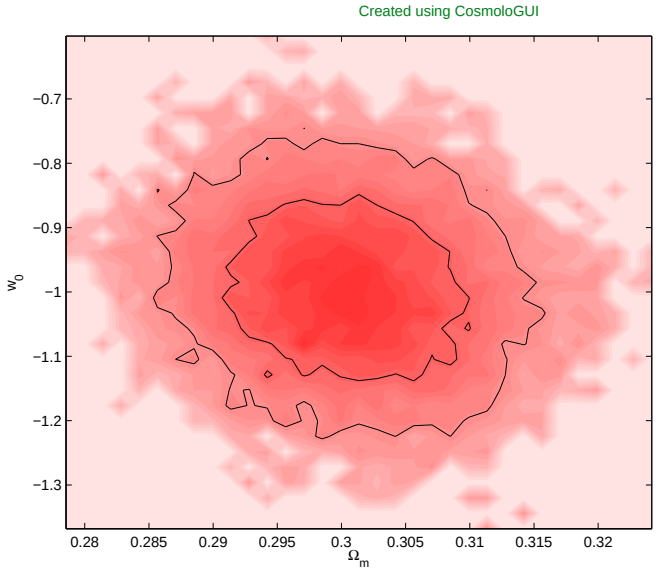



Figure 32: Confidence contours for a model with a time-dependent equation of state. All simulated future datasets are used.

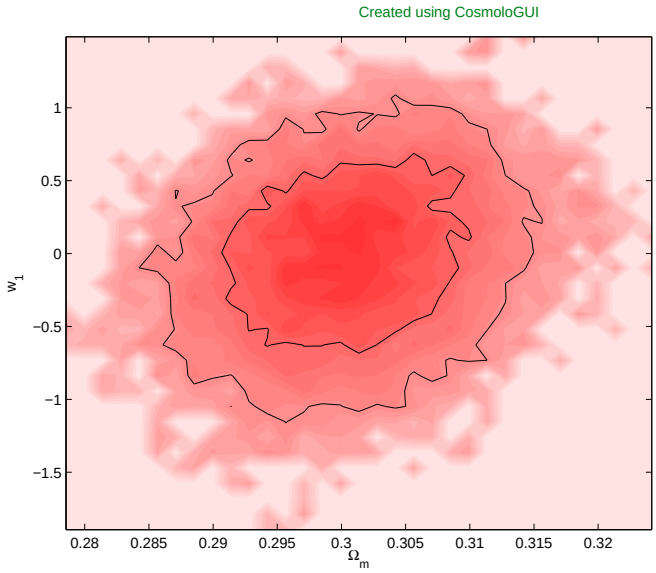



Figure 33: Confidence contours for a model with a time-dependent equation of state. All simulated future datasets are used.

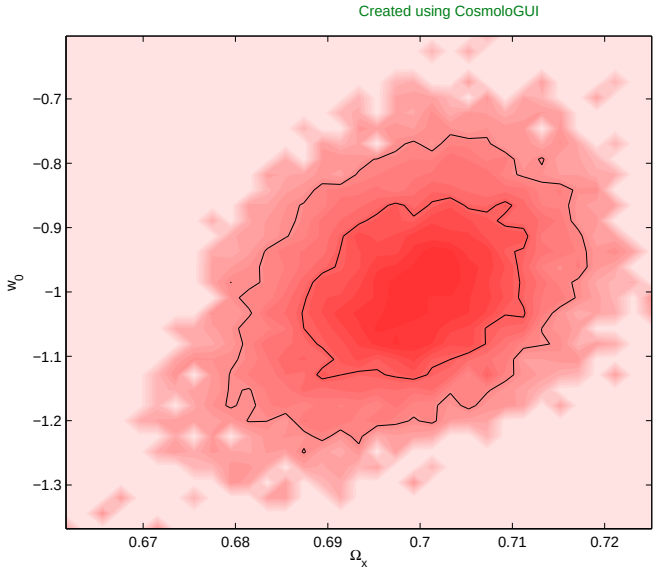



Figure 34: Confidence contours for a model with a time-dependent equation of state. All simulated future datasets are used.

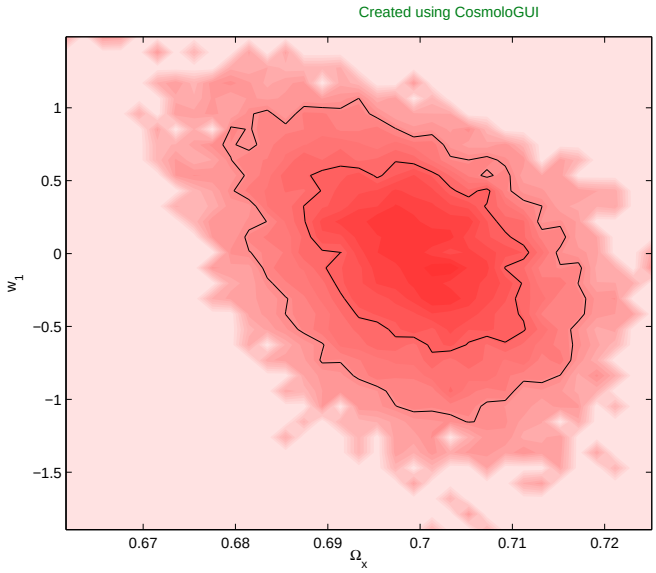



Figure 35: Confidence contours for a model with a time-dependent equation of state. All simulated future datasets are used.

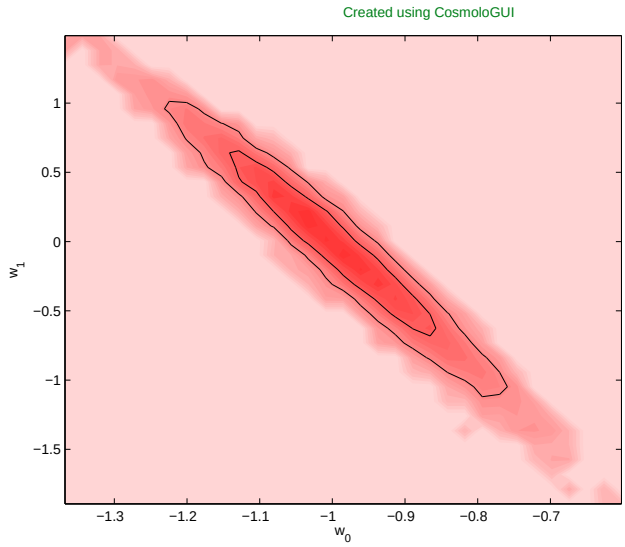



Figure 36: Confidence contours for a model with a time-dependent equation of state. All simulated future datasets are used.

| Results from chain | | | Performance of the 4 chains | |
|---|---|---|---|---|
| $\Omega_m$ | $0.294 < \Omega_m < 0.306$ | 68 % | Convergence | 1422 |
| | $0.288 < \Omega_m < 0.312$ | 95 % | Burn-in of chain 1 | 214 |
| | 0.30 | Best fit | Burn-in of chain 2 | 339 |
| $\Omega_x$ | $0.692 < \Omega_x < 0.708$ | 68 % | Burn-in of chain 3 | 301 |
| | $0.683 < \Omega_x < 0.715$ | 95 % | Burn-in of chain 4 | 325 |
| | 0.70 | Best fit | $c_l$ of chain 1 | 10 |
| $w_0$ | $-1.10 < w_0 < -0.90$ | 68 % | $c_l$ of chain 2 | 8 |
| | $-1.19 < w_0 < -0.80$ | 95 % | $c_l$ of chain 3 | 8 |
| | -1.0 | Best fit | $c_l$ of chain 4 | 10 |
| $w_1$ | $-0.45 < w_1 < 0.45$ | 68 % | Acc. rate of chain 1 | 0.36 |
| | $-0.91 < w_1 < 0.85$ | 95 % | Acc. rate of chain 2 | 0.37 |
| | 0.1 | Best fit | Acc. rate of chain 3 | 0.37 |
| $\chi^2_{min}$ | 4.84 | | Acc. rate of chain 4 | 0.37 |
| GoF | — | | | |
| BIC | 40.30 | | | |

Table 14: Confidence contours for a model with a constant equation of state. All datasets are used.

# 5 Conclusion

One can look at any of the plots, say figure (5), to confirm the results of Riess et al (1998) and Perlmutter et al (1998) that the universe is expanding. This is usually described by the deceleration parameter[1]

$$q_0 = \frac{\Omega_m}{2} - \Omega_x. \tag{38}$$

With very much more than 95 % certainty the deceleration parameter is negative i.e. the universe is accelerating.

By comparing figures (16) and (13), i.e. the flat model with a time-independent equation of state with the non-flat version where I instead include CMB, we see that the difference is very small. This means that instead of including the CMB one can just assume a flat universe. However, when comparing figures (18)–(20) with figures (24)–(29), where I now vary $w_1$, I get a slightly different result. Figure (27) and (28) show that $\Omega_x$ has a clear "tail" with very low values, which are associated with low values for $w_0$ and high values for $w_1$. This specific combination for the equation of state also fits the data. For the flat universe the low values of $\Omega_x$ gives a high value for $\Omega_m$, which are inconsistent for the model. Hence figure (18) lacks the tail of $\Omega_x$. When exploring the time-dependence of the dark energy it is therefore wise not to assume a flat universe, but include CMB instead.

In the simulations where I assume a flat universe but include CMB anyway, the contours are getting somewhat smaller relative to the figures where CMB are excluded, especially in the time-dependent case. The confidence contour showing $w_1$ (figure (22) and (23) shrinks quite remarkably. As before, the CMB measurement is not exactly equivalent with assuming a flat universe. One can improve the results further by including CMB in all cases.

When comparing figure (7), where I include all datasets in the cosmological constant model, with figure (12), where I include all datasets in the time-independent equation of state model, we see that they are almost identical. Hence, dropping the prior that the dark energy is a cosmological constant makes no effect on the confidence regions of the energy densities.

One can also note that by letting the equation of state vary in time, figures (18)–(29), there is little effect on $\Omega_m$ and $\Omega_x$ when comparing with the

[1]Until 1998 it was taken for granted that the universe was decelerating, the question was how much. This was the second time the universe fooled us completely. I wonder when it is going to happen the third time.

corresponding simulations for the time-independent model. The confidence contours of $w_0$ on the other hand are increasing dramatically when I let $w_1$ free. Hence the strength of the energy density of the dark energy can be determined quite well from the datasets I use, but the nature of it, described by the equation of state, is harder to determine.

My best model, according to BIC, is the simplest one where I vary only one parameter; $\Omega_m$ (plus the uninteresting $\mathcal{M}$), corresponding to a flat model with a cosmological constant. This means that the data do not support the introduction of more parameters, in most simulations they take values near this model even though they are free to take any value at all. However, saying that this model is the best does not mean that it is the true model. In many simulations the confidence contours are large, especially for $w_0$ and $w_1$, so there are many models which are consistent with data. The BIC only suggests the simplest one because that is what physics is all about; we want to describe the universe as simple as possible.

In figures (24)–(29), where I let all parameters free, we see that the data is still consistent with a cosmological constant. The best-fit values are $(w_0, w_1) = (-1.0, -0.2)$ and the values of the cosmological constant, $(w_0, w_1) = (-1, 0)$, is well within the 68 % contour. These contours, however, are very large, especially for $w_1$ and this model can therefore not rule out other theories for the dark energy. In figure (36) where the $w_0 - w_1$ plot is made for the simulated future results, we see that the confidence contours have shrunk considerably, but not enough as to rule out many models. It is very difficult to determine the possible time variation of dark energy, so in order to determine the nature of the dark energy, we must therefore continue to do more and better measurements, especially of distant supernovae, even more than those we will probably have in five years.

# References


Amanullah, R., *Baryon Acoustic Oscillations*, 2005

Astier, P., et. al. *The Supernova Legacy Survey: Measument of* $\Omega_M$, $\Omega_\Lambda$ *and w from the First Year Data Set*, arXiv:astro-ph/0510447, 2005

Bennet, C. L., et. al. *First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results*, arXiv:astro-ph/0302207, 2003

Branch, D., Perlmutter, S., Baron, E., Nugent, P., *Coping with Type Ia Supernova "Evolution" When Probing the Nature of the Dark Energy*, arXiv:astro-ph/0109070, 2001

Chandrasekhar, S., *The maximum mass of ideal white dwarfs*, Astrophysical Journal, 1931

Christensen–Dalsgaard, J., *Lecture Notes on Stellar Structure and Evolution*, Institut for Fysik og Astronomi, Aarhus Universitet, 2003

Cowles, M. K., Carlin, B. P., *Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review*, Journal of the American Statistical Association, 1996

Doran, M., Müller, C. M., *Analyze This! A cosmological constraint package for cmbeasy*, arXiv:astro-ph/0311311, 2004

Dunkley, J., Bucher, M., Ferreira, P. G., Moodley, K., Skordis, C., *Fast and reliable MCMC for cosmological parameter estimation*, arXiv:astro-ph/0405462, 2004

Efstathiou, G., Bridle, S. L., Lasenby, A. N., Hobson, M. P., Ellis, R. S., *Constraints on* $\Omega_\Lambda$ *and* $\Omega_M$ *from Distant Type Ia Supernovae and Cosmic Microwave Background Anisotropies*, arXiv:astro-ph/ 9812226, 1998

Eisenstein, J. D., et. al. *Detection of the Baryon Acoustic Peak in the Large-Scale Corelation Function of SDSS Luminous Red Galaxies*, arXiv:astro-ph/0501171, 2005

Folatelli, G., *Type Ia Supernova Cosmology: Quantitative Spectral Analysis*, Department of Physics, Stockholm University, 2004

Gelman, A. G., Roberts, G. O., Gilks, W. R., *Efficient Metropolis jumping rules*, Oxford University Press, 1996

Gelman, A. G., Rubin, D.B., *Interference from Iterative Simulation using Multiple Sequences*, Statistical Science, 1992

Gilks, W. R., Richardson, S., Spiegelhalter, D. J., *Markov Chain Monte Carlo in Practice*, Chapman and Hall, 1996

Goliath, M., Amanullah, R., Astier, P., Goobar, A., Pain, R., *Supernovae and the Nature of the Dark Energy*, arXiv:astro-ph/0104009, 2001

Goobar, A., Perlmutter, S., *Feasibility of Measuring the Cosmological Constant $\Lambda$ and Mass Density $\Omega$ using Type Ia Supernovae*, arXiv:astro-ph/9505022, 1995

Green, P. J., Mira, A., *Delayed rejection in reversible jump Metropolis-Hastings*, Biometrika, 2001

Haario, H., Laine, M., Mira, A., Saksman, E., *DRAM: Efficient adaptive MCMC*, AMS Classification, 2003

Hamuy, M., et. al., *The 1990 Calan/Tololo supernova search*, Astrophysical Journal, 1993

Hamuy, M., Phillips, M. M., Maza, J., Suntzeff, N. B., Aviles, R., *A Hubble diagram of distant Type Ia supernovae*, Astrophysical Journal, 1995

Hannestad, S., Mörtsell, E., *Cosmological constraints on the dark energy equation of state and its evolution*, arXiv:astro-ph/0407259, 2004

Hastings, W. K., *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, 1970

Ichikawa, K., Takahashi, T., *Dark Energy Evolution and the Curvature of the Universe from Recent Observations*, arXiv:astro-ph/0511821, 2005

Jha, S., and the High-Z Supernova Search Team, *Testing Cosmic Acceleration with Type Ia Supernovae*, arXiv:astro-ph/0101521, 2001

Kim, A., et. al., *K Corrections for type Ia Supernovae and a Test for Spatial Variation of the Hubble Constant*, arXiv:astro-ph/9602123, 1996

Knop, R. A., et. al. *New Constraints on $\Omega_M$, $\Omega_\Lambda$, and w from an Independent Set of Eleven High-Redshift Supernovae Observed with HST*, arXiv:astro-ph/0309368, 2003

Lewis, A., Bridle, S., *Cosmological parameters from CMB and other data: a Monte-Carlo approach*, arXiv:astro-ph/ 0205436, 2002

Liddle, A. R, *How many cosmological parameters?*, arXiv:astro-ph/0401198, 2004

Lindahl, J., *Probing the Positron Phase of a Thermonuclear Supernova*, Department of Astronomy, University of Stockholm, 2003

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E., *Equation of State Calculations by Fast Computing Machines*, Journal of Chemical Physics, 1953

Neal, R. M., *Probabilistic Inference Using Markov Chain Monte Carlo Methods*, Department of Computer Science, University of Toronto, 1993

Perlmutter, S., et. al. *Measurements of $\Omega$ and $\Lambda$ from 42 High-Redshift Supernovae*, arXiv:astro-ph/9812133, 1998

Phillips, M. M., *The Absolute Magnitudes of Type Ia Supernovae*, Astrophysical Journal , 1993

Raftery, A. E., Lewis, S., *How many iterations in the Gibbs Sampler?*, Oxford University Press, 1992

Raine, D.J, Thomas, E.G., *An Introduction to the Science of Cosmology*, Institute of Physics Publishing, Bristol and Philadelphia, 2001

Riess, A. G., et. al. *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, arXiv:astro-ph/9805201, 1998

Riess, A. G., et. al., *Type Ia Supernova Discoveries at $z > 1$ From the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution*, arXiv:astro-ph/ 0402512, 2004

Riess, A. G., Press, W. H., Kirshner, R. P., *Using SN Ia Light Curve Shapes to Measure the Hubble Constant*, astro-ph/ 9410054, 1994

Ryden, B., *Introduction to Cosmology*, Addison Wesley, 2003

Sahlén, M, Liddle, A.R., Parkinson, D., *Direct reconstruction of the quintessence potential*, arXiv:astro-ph/0506696, 2005

Sparke, L. S., Gallagher, J. S., *Galaxies in the Universe: An Introduction*, Cambridge University Press, 2000

Spergel, D. N., et. al., *First Year Wilkinson Microwave Anisotrophy Probe (WMAP) Observations: Determination of Cosmological Parameters*, arXiv:astro-ph/0302209, 2003

Sunesson, C., *Gravitational Lensing with Multiple sources*, Department of Astronomy, University of Stockholm, 2004

Tegmark, M, et. al., *Cosmological Parameters from SDSS and WMAP*, arXiv:astro-ph/0310723, 2004

Tierney, L., Mira, A., *Some adaptive Monte Carlo methods for Bayesian inference*, Statistics in Medicine, 1999

Tonry, J. L., and the High-Z Supernova Search Team, *Type Ia Supernovae, the Hubble Constant, the Cosmological Constant, and the Age of the Universe*, arXiv:astro-ph/0105413, 2001

Vacca, W. D., Leibundgut, B, *Modeling the Lightcurves of Type Ia Supernovae*, arXiv:astro-ph/ 9511114, 1995

Verde, L., et. al. *First Year Wilkinson Microwave Anisotrophy Probe (WMAP) Observations: Parameter Estimation Methodology*, arXiv:astro-ph/0302218, 2003

Wang, Y., Mukherjee, P., *Model-Independent Constraints on Dark Energy Density from Flux-averaging Analysis of Type Ia Supernova Data*, arXiv:astro-ph/0312192, 2004

# 6 Appendix A - The evolution of the Hubble parameter

The Friedmann equation, equation (6), gives

$$H^2 = \frac{8\pi G}{3}(\rho_m - \rho_\lambda) - \frac{k}{R^2} \tag{39}$$

where $c$ as usual is set to 1. Since matter is diluted over a sphere as the universe expand we have

$$\rho_m \propto R^{-3} \tag{40}$$

whereas the cosmological constant by definition goes as

$$\rho_\lambda \propto constant. \tag{41}$$

By using the critical density, equation (13), and re-write equation (39) using the present-day values, equation (40) and (41), we obtain

$$\frac{H^2}{H_0^2} = \Omega_\lambda + \Omega_m \left(\frac{R}{R_0}\right)^{-3} - (\Omega - 1)\left(\frac{R}{R_0}\right)^{-2}. \tag{42}$$

Since $\Omega_k = 1 - \Omega$ and $\frac{R_0}{R} = 1 + z$ we now have

$$\frac{H^2}{H_0^2} = (1+z)^3\Omega_m + (1+z)^2\Omega_k + \Omega_\lambda \tag{43}$$

If we assume that the dark energy is something different than a cosmological constant things get a little more complicated. We must also consider the evolution of the density of the dark energy, $\rho_x$, defined as, if the equation of state of the dark energy is parameterized as equation (11) (Ichikawa & Takahashi, 2005)

$$\rho_x(z) = \rho_{x0}(1+z)^{3(1+w_0+w_1)}e^{\frac{-3w_1 z}{1+z}}. \tag{44}$$

This gives, finally

$$\frac{H^2}{H_0^2} = (1+z)^3\Omega_m + (1+z)^2\Omega_k + (1+z)^{3(1+w_0+w_1)}e^{\frac{-3w_1 z}{1+z}}\Omega_\lambda \tag{45}$$