The isocurvature fraction after WMAP 3–year data

Roberto Trotta
Oxford University, Astrophysics, Denys Wilkinson Building, Keble Road, OX1 3RH, United Kingdom

ABSTRACT
I revisit the question of the adiabaticity of initial conditions for cosmological perturbations in view of the 3–year WMAP data. I focus on the simplest alternative to purely adiabatic conditions, namely a superposition of the adiabatic mode and one of the 3 possible isocurvature modes, with the same spectral index as the adiabatic component.

I discuss findings in terms of posterior bounds on the isocurvature fraction and Bayesian model selection. The Bayes factor (models likelihood ratio) and the effective Bayesian complexity are computed for several prior ranges for the isocurvature content. I find that the CDM isocurvature fraction is now constrained to be less than about 10%, while the fraction in either the neutrino entropy or velocity mode is below about 20%. Model comparison strongly disfavours mixed models that allow for isocurvature fractions larger than unity, while current data do not allow to distinguish between a purely adiabatic model and models with a moderate (i.e., below about 10%) isocurvature contribution.

The conclusion is that purely adiabatic conditions are strongly favoured from a model selection perspective. This is expected to apply in even stronger terms to more complicated superpositions of isocurvature contributions.

Key words: Cosmology – Initial conditions – Model selection – Bayesian methods

1 INTRODUCTION
The detailed nature of the initial conditions for cosmological perturbations is one of the open questions in cosmology. The exquisite precision of the WMAP measurement of the first acoustic peak location in the cosmic microwave background (CMB) temperature power spectrum ($\ell = 220.7 \pm 0.7$, see Hinshaw et al. (2006)) is a strong indication in favour of adiabatic initial conditions, which predict for the first peak $\ell \approx 220$. The alternative possibility of cold dark matter (CDM) isocurvature initial conditions exhibits a sine wave (rather than the cosine excited by adiabatic conditions) in the photon–baryon plasma, resulting in a first acoustic peak displaced by half a period to $\ell \approx 330$, see e.g. Trotta (2004); Durrer (2004). Furthermore, the ratio of the Sachs–Wolfe plateau for $\ell \lesssim 50$ to the height of the peak is very different for the two modes.

A few years ago, Bucher and collaborators introduced two new isocurvature modes, called “neutrino density” (or, more appropriately, “neutrino entropy”) and “neutrino velocity” modes (Bucher et al. 2000). They are characterized by a non–vanishing initial entropy perturbation in the neutrino sector and by a non–vanishing difference in the neutrino to photon velocity, respectively. A superposition of the adiabatic and the three isocurvature modes (cold dark matter, neutrino entropy and neutrino velocity) constitutes the most general initial conditions for the perturbations, at least if the Universe is radiation dominated in its early phase (Trotta 2004). A baryon isocurvature mode is observationally indistinguishable from a CDM isocurvature one (Bucher et al. 2001; Gordon & Lewis 2003) and can thus be neglected without loss of generality.

Allowing for the most general type of initial conditions has two effects on cosmological parameter extraction from CMB measurements. First, the extra parameters associated with the initial conditions introduce severe degeneracies which limit our ability to reconstruct the cosmology (Trotta et al. 2001, 2003), even though this can fortunately be remedied by using polarization information (Bucher et al. 2001; Trotta & Durrer 2004). Secondly, it becomes difficult to constrain the type of initial conditions, i.e. the amount of isocurvature contributions allowed on top of the predominantly adiabatic mode (Moodley et al. 2004).

Recent works have investigated general isocurvature admixtures in the initial conditions (Beltran et al. 2005; Moodley et al. 2004; Bean et al. 2006). In this work I focus on the simplest alternative to a purely adiabatic power spectrum, namely a superposition with one totally (anti–)correlated isocurvature mode at the time with the same spectral index as the adiabatic one. This is partly motivated...
by models for the generation of initial conditions such as the curvaton (see e.g. Gordon & Lewis (2003); Lyth & Wands (2002) and references therein), where this kind of scenario arises as a generic prediction. A second justification comes from the model selection approach used in the second part of this work. In comparing the simplest (i.e., purely adiabatic) scenario with a more complex one, it makes sense to start by adding a minimal number of extra parameters, and see whether the extended model is justified by the data. This model selection perspective has been recently advocated by Beltran et al. (2005); Trotta (2005).

This paper is organized as follows: in section 2 we introduce the parameterization of the initial condition parameters space we are considering, while in section 3 we review some concepts of Bayesian statistics and in particular the model selection approach. We present our results in terms of parameters constraint and model comparison outcome in section 4 and offer our conclusions in 5.

2 THE ISOCURVATURE FRACTION

The most general initial conditions for cosmological perturbations are described by a symmetric $4 \times 4$ matrix, $M$, with 10 free parameters representing the amplitudes of the pure modes (along the diagonal) and their correlations (off-diagonal elements). From a phenomenological point of view, there are also 10 more parameters describing the spectral tilt of each mode and correlator. If one is willing to consider running of the spectral index, then this would introduce another 10 free parameters in the problem. As motivated in the introduction, we consider here a minimal extension of the simplest adiabatic model, namely a diagonal matrix

$$M = \text{diag}(\zeta, S_c, S_v, V_o)$$

(1)

$$\zeta = \text{diag}(1, f_o, f_{ne}, f_{nv})$$

(2)

where $\zeta$ is the amplitude of the curvature perturbation (adiabatic mode), $S_c, S_v$ are the (gauge invariant) entropy perturbations in the CDM and neutrino component defining non-vanishing CDM isocurvature and neutrino entropy modes, respectively. The neutrino entropy mode is often referred to as “neutrino density”. The quantity $V_o$ corresponds to a non-zero neutrino-photon velocity giving rise to a neutrino velocity mode (see e.g. Trotta (2004) for precise definitions). The quantities $f_o, f_{ne}, f_{nv}$ give the isocurvature fractions with respect to the curvature perturbation, where the notation employed is $ci = CDM$ isocurvature, $ne = neutrino entropy$ and $nv = neutrino velocity$. The sign of $f_o$ (with $x = ci, ne, nv$) determines the nature of the correlation: a positive correlation ($f_x > 0$) results in extra power to the Sachs–Wolfe plateau, a negative correlation ($f_x < 0$) subtrahs power in the region $\ell \leq 50$. As already mentioned, we take a common spectral index for the adiabatic and the isocurvature mode, $n_x$, and we analyse separately the three scenarios where only one of the isocurvature modes is non-zero, in addition to the adiabatic mode.

An alternative parameterization for the isocurvature contribution that is common in the literature is given in terms of the variable $\alpha_x$, or $\alpha_x^2$ (used e.g. by Beltran et al. (2004) and Bean et al. (2008)). This is related to $f_x$ by

$$\alpha_x^2 = \text{sign}(f_x) \frac{f_x^2}{1 + f_x^2}$$

(3)

From a phenomenological perspective, there is little reason to prefer one parameterization over the other. However, from a model selection point of view the choice of the variable one puts flat priors on is of great importance, since the available parameter space under the prior enters in the calculation of the Occam’s factor for the model, see the discussion in Trotta (2005). We must consider the choice of priors as inherent to the specification of the extended model and different choices will lead to different conclusions since the Occam’s razor effect is not invariant under non-linear transformations of variables.

Once a fundamental model for the generation of the initial condition is specified, one can select the appropriate physical variable over which to impose a prior reflecting our state of knowledge before we see the data. For instance, it can be argued that the $f_x$ parameterization is closer to the curvaton setup, while the $\alpha_x$ choice of variable compresses the parameter space in the compact interval $-1 < \alpha_x < 1$. A flat prior of $\alpha_x$ gives equal a priori accessible volume to adiabatic–dominated ($|\alpha_x| \leq 0.5$) and to isocurvature–dominated ($|\alpha_x| > 0.5$) models. The prior on $f_x$ is very much dependent on what we think the available parameter space is under our extended model. Therefore we discuss below the results of model selection as a function of the prior width $\Delta f$, taking a flat prior in the range $-\Delta f \leq f_x \leq \Delta f$. This allows an easy comparison once a prior range under a specific model is given. We postpone to a future work a detailed analysis of prior selection based on first principles.

3 PARAMETER ESTIMATION, MODEL SELECTION AND MODEL COMPLEXITY

Bounds on the isocurvature fraction are derived in terms of high probability regions in the posterior probability density function (pdf) for the parameters $\theta$ given the data $d$, $p(\theta|d)$. This is obtained through Bayes theorem,

$$p(\theta|d) = \frac{p(d|\theta)\pi(\theta)}{p(d)}$$

(4)

where $p(d|\theta)$ is the likelihood function, $\pi(\theta)$ is the prior pdf and $p(d)$ is the model likelihood (sometimes called “the evidence”) of the data under the model. The model under consideration is defined by the parameter set $\theta$ and the choice of the prior $\pi(\theta)$ (we shall return to this point below).

The model likelihood is a normalization constant independent on the parameters of the model, and it can be ignored as far as the parameter estimation step is concerned. It becomes the key quantity for model selection, and in particular we are interested in the relative odds between the simplest, purely adiabatic model $M_0$ and a model augmented by an extra isocurvature contribution, $I_x = (M_0, f_x)$, with $x = ci, ne, nv$ as above. The change in our degree of belief in the two models after we have seen the data is described by the Bayes factor

$$B = \frac{p(d|M_0)}{p(d|I_x)}$$

(5)

which is the ratio of the normalization constants for the two models in Bayes theorem, Eq. (4). Since the two models are nested (i.e., we obtain $M_0$ from $I_x$ by setting the isocurvature fraction to zero, $f_x = 0$), the Bayes factor can
be conveniently computed using the Savage–Dickey density ratio (SDDR) (see Trott (2005) and references therein)
\[
B = \frac{\pi(f_x|d, I_x)}{\pi(f_x|I_x)} |_{f_x=0}.
\]

This is easy to compute from a Monte Carlo Markov Chain (MCMC), requiring only knowledge of the properly normalized posterior over the extra variable \(f_x\) of the extended model. Furthermore, using the SDDR has the advantage that the impact of a change of prior can usually be evaluated by simply post-processing a chain including the new prior. This is the approach used below in section 4. If the posterior pdf is well approximated by a Gaussian distribution with mean \(\mu\) and standard deviation \(\sigma\), and for a flat prior in the range \(-\Delta f \leq f_x \leq \Delta f\), the Bayes factor becomes
\[
B = \frac{\sqrt{8}}{\pi \sigma} \left[ \text{erfc} \left( \frac{|\mu| - \Delta f}{\sqrt{2} \sigma} \right) - \text{erfc} \left( \frac{|\mu| + \Delta f}{\sqrt{2} \sigma} \right) \right]^{-1}.
\]

For \(\ln B > 0\) model \(M_0\) is favoured over \(I_x\) because the extra complexity (in terms of wasted parameter space) of \(I_x\) is not warranted by the data, while for \(\ln B < 0\) \(I_x\) is favoured since the data require the extra parameter. A useful rule of thumb (Kass & Raftery 1995) is that a positive (strong) preference requires \(|\ln B| \gtrsim 1\) (\(\gtrsim 3\)). A model likelihood ratio \(|\ln B| > 5\) (corresponding to odds > 150 : 1) is deemed to constitute “decisive” evidence.

Finally, the last relevant quantity for our analysis is the Bayesian model complexity, which measures the number of parameters the data can support, regardless of whether the parameters in question are actually detected or not (for more details, see Kunz et al. 2004). The Bayesian complexity is defined as
\[
C_b = \chi^2(\hat{\theta}) - \chi^2(\bar{\theta}),
\]
where the effective \(\chi^2(\bar{\theta})\) is derived from the likelihood as \(\chi^2(\bar{\theta}) = -2 \ln p(d|\bar{\theta})\). The bar denotes an average over the posterior pdf, while the hat denotes a point–estimator which in our case we take to be the mean under the posterior, i.e. \(\hat{\theta} = \bar{\theta}\). We will use \(C_b\) to quantify the number of supported parameters in our extended models \(I_x\), in order to verify whether the isocurvature fraction is a variable that could have been detected using current data. A detailed discussion of the meaning and interpretation of the Bayesian complexity can be found in Kunz et al. 2004.

It is important to stress that both the model likelihood and the Bayesian complexity depend not only on the data but also on the model description one chooses to adopt, i.e. on the prior choices one makes for \(\pi(f_x)\) (see Trott 2006 for an example applied to the case of dark energy models). This is an irreducible feature of the Bayesian model selection approach. It seems to us that there cannot be an absolute notion of “a best model”, but only relative statements about the support the data give to different models when compared to each other. Furthermore, the very concept of Bayesian complexity is only meaningful when the constraining power of the data is compared to the scale of the problem at hand, which again must be defined by specifying the prior.

Our simplest model \(M_0\) is a flat ΛCDM Universe with purely adiabatic conditions, described by following set of 6 parameters
\[
\theta = (\zeta, \omega_b, \omega_c, \Theta, n_s, \tau_s)
\]
where \(\zeta\) is the curvature perturbation, \(\omega_b, \omega_c\) are the physical densities of baryons and CDM, respectively, \(\Theta\) is the ratio of the sound horizon to the angular diameter distance to last scattering, \(n_s\) is the spectral tilt and \(\tau_s\) the optical depth to reionization. An extra bias parameter \(b\) for the matter power spectrum is treated as a nuisance parameter and marginalized over, hence we do not count it as an additional parameter. We do not consider tensor modes nor extra neutrino species nor running of the spectral index. We take our 3 neutrino families to be massless and we fix the dark energy equation of state to \(w = -1\) at all redshifts. All of those choices are motivated by the fact that inclusion of any of the above extra parameters is presently not required by the data. This means that a comparison between a model including both the isocurvature fraction and one of the above extra parameters against the simple adiabatic model would favour even more strongly the latter, as a consequence of the extra Occam’s factor effect coming from the extra parameter. In this sense, our model selection is actually conservative.

The situation is different for parameter constraints, since in this case strong degeneracies between the isocurvature fraction and other extra parameters might change the posterior bounds of \(f_x\). In particular, one can expect a strong degeneracy between the CDM isocurvature mode and the presence of a tensor mode from gravity waves, when considering temperature power spectrum information alone. The extra power contributed by the CDM isocurvature to the Sachs–Wolfe plateau for small \(\ell\)’s is strongly anti–correlated with the tensor mode amplitude. However, inclusion of polarization data would help in breaking this degeneracy, at least partially. The impact of allowing a tensor mode contribution is very mild for the neutrino modes, since the Sachs–Wolfe plateau is lower than the first acoustic peak for these modes, and as a consequence constraints on their amplitudes are dominated by the height of the peak, not by the height of the plateau. Similar considerations also apply for a possible running of the spectral index. Other parameters that mainly affect the angular diameter distance to last scattering and therefore the position of the acoustic peaks in the spectrum (such as the dark energy equation of state, the curvature of spatial sections or an extra background of relativistic particles) present only weak degeneracies with the isocurvature fractions, since the peaks’ position is strongly constrained by the data.

In the following, we therefore limit our analysis to the 6 parameters model \(M_0\) described above, complemented in the extended models by the isocurvature fractions as follows. The extended models \(I_x\) contain a non–vanishing isocurvature fraction
\[
I_x = (\theta, f_x)
\]
where \(f_x\) is defined in Eq. (2) and \(x = c, i, n, d\). The spectral index of the isocurvature mode is the same as the adiabatic one, \(n_s\). The correlation coefficient between the adiabatic and the isocurvature mode is ±1, depending on the sign of \(f_x\).
For a given value of \( \sigma \), the CDM isocurvature is the mode \( f_{\text{CDM}} \). We find that the 1-dimensional marginalized regions encompassing 95% of probability. We give the posterior mean \( \mu \) and standard deviation \( \sigma \), as well 1-dimensional marginalized regions encompassing 95% of posterior probability.

| Model          | \( \mu \)     | \( \sigma \)     | 95% interval on \( f_x \) |
|----------------|--------------|------------------|----------------------------|
| CDM iso        | \(-2.5 \times 10^{-2}\) | \(4.0 \times 10^{-2}\) | \(-0.10 \ldots 0.06\)    |
| \( \nu \) entropy | \(-4.4 \times 10^{-2}\) | \(8.0 \times 10^{-2}\) | \(-0.20 \ldots 0.12\)    |
| \( \nu \) velocity | \(-1.2 \times 10^{-2}\) | \(1.0 \times 10^{-1}\) | \(-0.18 \ldots 0.22\)    |

with the largest contribution to the CMB power spectrum. Also, all of the 1-dimensional pdf’s for \( f_x \) are very close to Gaussian. Hence we expect that Eq. (7) is a good approximation to the Bayes factor, Eq. (6), as we now show.

We now evaluate the Bayes factor between the models including an isocurvature contribution and the simplest, purely adiabatic model. As we have seen above in the parameter extraction step, there is no indication that the data require and isocurvature component, since the isocurvature fraction is compatible with 0. This is consistent with the findings of Bean et al. (2006). We therefore expect the Bayes factor to favour the purely adiabatic model on the ground of the Occam’s razor argument. The strength of evidence in favour of the adiabatic model depend on the amount of wasted parameter space for the isocurvature fraction, i.e. by the prior range \( \Delta f \). In the top panel of Figure 2, we plot the Bayes factor as a function of the prior width \( \Delta f \), while in the bottom panel we plot the Bayesian complexity, i.e. the number of parameters effectively constrained by the data. We can see that for models with poor predictivity, i.e. a large prior accessible range \( \Delta f \gg 1 \), one finds strong \((> 20:1)\) to decisive \((> 150:1)\) posterior odds against the extended model for all of the three isocurvature modes. We also plot the Gaussian approximation to the SDDR for the Bayes factor, Eq. (7), for the CDM isocurvature mode, and find a very good match with the value computed numerically from the Monte Carlo chain.

For a prior choice \( \Delta f \leq 1 \), the Bayesian complexity is close to 7, indicating that all of the 7 parameters of the extended model have been measured. We therefore conclude that models predicting up to the same amount of isocurvature to adiabatic power (the case \( \Delta f = 1 \)) are strongly disfavoured for the CDM mode, and mildly disfavoured in the case of the two neutrino modes. However, if the prior range is reduced below \( \Delta f = 1 \), i.e. for models predicting predominantly adiabatic initial conditions with subdominant isocurvature contribution, the Bayes factor gives an inconclusive result, with about equal odds for the purely adiabatic and the mixed models. At the same time, the Bayesian complexity decreases, indicating that \( f_x \) is only poorly constrained with respect to the scale of the prior, especially for the neutrino density and velocity modes. This reinforces the conclusion that current data are not strong enough to select among a purely adiabatic model and one which predicts up to 10% isocurvature contribution and we need to acquire better data in order to obtain a higher-odds result.

### 5 CONCLUSIONS

We have submitted the question of the type of initial conditions for cosmological perturbations to renewed scrutiny...
Bottom panel: the Bayesian complexity gives the effective number to the Bayes factor, Eq. (7), for the CDM isocurvature mode. Out (\( \ln B \) < about 10% in any of the three modes cannot presently be ruled out. For \( \Delta f \) strongly disfavours models with \( \Delta f \gtrsim 1 \) because of the Occam’s razor effect, while models predicting an isocurvature fraction below about 10% in any of the three modes cannot presently be ruled out (\( \ln B < 1 \)). The dotted line gives the Gaussian approximation to the Bayes factor, Eq. (7), for the CDM isocurvature mode.

Figure 2. Result of model selection between a purely adiabatic model and an extended model featuring a totally (anti-)correlated isocurvature component, as a function of the prior available range for the isocurvature fraction, \( \Delta f \). Top panel: the Bayes factor strongly disfavours models with \( \Delta f \gg 1 \) because of the Occam’s razor effect, while models predicting an isocurvature fraction below about 10% in any of the three modes cannot presently be ruled out (\( \ln B < 1 \)). The dotted line gives the Gaussian approximation to the Bayes factor, Eq. (7), for the CDM isocurvature mode.

We have derived posterior bounds on the isocurvature fraction from WMAP 3-year data combined with other CMB measurements and SDSS. We have constrained the isocurvature fraction in the CDM mode to be less than about 10%, while the maximum allowed neutrino isocurvature contribution (either density or velocity) is about 20%.

Bayesian model selection tends to favour purely adiabatic initial conditions with strong odds (\( > 20 : 1 \)) when compared to models predicting isocurvature fractions larger than unity. For such models – having a large prior range on the isocurvature fraction – we have shown that the data can support 7 parameters, but that only 6 of them are required, with no need to include isocurvature modes from a model selection point of view. These findings confirm the conclusions of Kunz et al. (2006). However, mixed models that limit the isocurvature contribution to less than about 10% cannot presently be ruled out. We have shown that the constraining power of the data for this class of models is insufficient, and therefore we must hold our judgement until better data becomes available. These findings are however dependent on the parameterization chosen for the isocurvature fraction, that in this work is motivated by the curvaton scenario. The question of prior selection will be further addressed in a future publication.

It is reasonable to expect that the same conclusion would apply in even stronger terms to the case of more complicated models, e.g. those involving a superposition of different isocurvature modes at the same time, or with arbitrary correlations among them. In fact, more complicated models (such as the class considered by Bean et al. (2006)) ought to be even more disfavoured because of their larger volume of wasted parameter space. At present, Occam’s razor is perfectly compatible with the simplest possibility, namely purely adiabatic initial conditions.

We think that this model comparison approach can be a useful complement to parameter constraints analysis, and that it can offer valuable guidance in building models for the generation of primordial perturbations.

References R.T. is supported by the Royal Astronomical Society through the Sir Norman Lockyer Fellowship. I am grateful to Andrew Liddle and David Parkinson for comments. I acknowledge the use of the package cosmomc, available from cosmologist.info and the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science.

References
Bean R., Dunkley J., Pierpaoli E., 2006, astro-ph/0606685
Beltran M., Garcia-Bellido J., Lesgourgues J., Slosar A., 2005, Phys. Rev., D71, 063532
Beltran M., Garcia-Bellido J., Lesgourgues J., Riazuelo A., 2004, Phys. Rev., D70, 103530
Beltran M., Garcia-Bellido J., Lesgourgues J., Viel M., 2005, Phys. Rev., D72, 103515
Bucher M., Moodley K., Turok N., 2000, Phys. Rev., D62, 083508
Bucher M., Moodley K., Turok N., 2001, Phys. Rev. Lett., 87, 191301
Durrer R., 2004, Lect. Notes Phys., 653, 31
Freedman W. L., et al., 2001, Astrophys. J., 553, 47
Gordon C., Lewis A., 2003, Phys. Rev., D67, 123513
Hinshaw G., et al., 2006, astro-ph/0603451
Kass R., Raftery A., 1995, J. Amer. Stat. Assoc., 90, 773
Kunz M., Trotta R., Parkinson D., 2006, Phys. Rev. D 74, 023503
Kuo C.-I., et al., 2004, Astrophys. J., 600, 32
Lyth D. H., Wands D., 2002, Phys. Lett., B524, 5
Moodley K., Bucher M., Dunkley J., Ferreira P. G., Skordis C., 2004, Phys. Rev., D70, 103520
Page L., et al., 2006, astro-ph/0603450
Readhead A. C. S., et al., 2004, Astrophys. J., 609, 498
Tegmark M., et al., 2004, Astrophys. J., 606, 702
Trotta R., 2004, PhD thesis, University of Geneva, Thesis N. 3534, astro-ph/0410115
Trotta R., 2005, astro-ph/0504022
Trotta R., Durrer R., 2004, Proceedings of the MG10 Meeting, Rio de Janeiro, Brazil 20–26 July 2003 (M.
Roberto Trotta

Novello and S. Perez Bergliaffa, Eds) World Scientific, astro-ph/0402032
Trotta R., Riazuelo A., Durrer R., 2001, Phys. Rev. Lett., 87, 231301
Trotta R., Riazuelo A., Durrer R., 2003, Phys. Rev., D67, 063520
Trotta R., To appear in Proceedings of the Conference on Cosmology, Galaxy Formation and Astro-Particle Physics on the Pathway to the SKA, Oxford, England, 10–12 Apr 2006. arXiv:astro-ph/0607496