Dark Cosmology Centre

DYNAMICS OF

STARS,

DARK MATTER AND THE UNIVERSE

JOHAN SAMSING

Dissertation
Submitted for the Degree

PHILOSOPHIÆ DOCTOR

Dark Cosmology Centre
The PhD School of Science
Faculty of Science
University of Copenhagen

Submission:	*10/07/2014*
Defence:	*-/08/2014*
Supervisor:	*Prof. Steen Hansen*
Opponents:	*Prof. -* *Prof. -* *Prof. -*

ACKNOWLEDGMENT

First I will thank my three main scientific advisors, Steen from DARK, Eric from Berkeley and Enrico from Santa Cruz. They have each greatly inspired me and helped to become the scientist I am today. I also thank the DARK Institute for its lively atmosphere and for always supporting me especially on my longer trips abroad. A further thank all my friends at DARK, my family and my wife Akiko.

Johan

ABSTRACT

For my PhD I have explored our Universe from its smallest scales with stars and galaxies to its largest to understand its cosmological evolution.

On the largest scales, the universe seems to be speeding up faster and faster. Our current cosmological model indicates this acceleration could be due to a component known as dark energy. From observations we find that the Universe consist of about 70% dark energy, but its fundamental properties are completely unknown. To address this question, I did a study on how well we can measure its possible evolution as a function of cosmic time. I especially focused on dark energy which could be present shortly after the big bang as this could strongly challenge our current picture of the Universe. My collaborators and I used temperature measurements of the cosmic microwave background (CMB) observed by the Planck satellite, to look for any variations in this dark component. This work was followed up by how we can measure dark energy at later times. For this we instead used redshift and angular position observations of millions of galaxies. The central problem was here how well we need to know how galaxies move and clump together during cosmic history. We explored a new model independent way of doing this which also seems promising for measuring modifications to the theory of gravity itself.

On slightly smaller scales I looked into what happens when two dark matter structures merge. Numerical simulations show that a smaller fraction of the dark matter particles are kicked out during the merger process. In my work I discovered a dynamical mechanism explaining this. The mechanism is a double scattering process where a particle gains energy by undergoing two gravitational defelctions during the merger. From this model I can explain recent observations of high velocity stellar systems. I further did a study on how well one can measure the 3D shape of a single dark matter structure. Dark matter structures attract a huge amount of gas during their formations which heats up and emit X-rays. I showed that one can estimate the dark matter structure shape from observation of these X-rays alone. This has implication for mass measurements which can be used for constraining the amount of matter and dark energy we have in our universe.

On even smaller scales I did an interesting study on the interaction between stars and black holes. I especially looked into the interaction where a binary stellar system is hit by a third object. In my calculations I included the possibity for the system to emit gravitational waves (GWs), these waves carry both energy and momentum. By performing millions of scatterings I discovered a new outcome where a binary stellar system quickly merge by sending out GWs. The two stars merge with very high eccentricty which makes this outcome very unique. My collaborators and I estimated this new outcome to be the most likely way of generating high eccentric binary mergers. This have huge implications for future GW measurements.

CONTENTS

1

INTRODUCTION

This thesis is a compilation of my recent work and includes very different topics from stellar interactions to the expansion of the universe. Instead of writing one big introduction I have instead chosen to write short individual resumes for each project I present. These resumes are found after the abstract of each paper presented in the chapters below.

The first part of the thesis will mainly be on dark energy and how to constrain its possible evolution in a model independent way. From my publications I present some recent work I did at Berkeley, where I explored the possibility of measuring dark energy at early times. As presented in chapter 2, my collaborators and I first did a principal component study on how well one is able to measure time evolving dark energy from CMB observations alone. Since the CMB originates from only 300.000 years after the big bang, we are here constraining what is known as an early dark energy component. This work was followed up by an analysis using real Planck data as presented in chapter 3. Further studies on varying dark energy and modified gravity is presented in chapter 4. In this work, my collaborator and I looked into the possibility of using galaxy redshift and angular position measurements to constrain dark energy and a possible deviation to general relativity. For this we used a framework known as redshift space distortions as described in the resume of the paper.

In the second part of the thesis I turn to more astrophysical topics including stellar dynamics, halo mergers and X-ray galaxy clusters. The first work presented in this part is a study I did on the role of gravitational wave (GW) emission in binary-single interactions between objects such as white dwarfs, neutron stars and black holes. This project was done with people at UCSC and is presented in chapter 5. The main result from this work was the discovery of a new channel for generating GW inspiraling binaries with extremely high eccentricity. My collaborators and I found this to be the leading channel for high eccentric inspirals which has major implications especially for GW detections using LIGO. In chapter 6 is presented work I have done on dark matter halo mergers. It is well known that halo particles are energetically kicked away during a merger between their own halo and a target halo, in my work I was able to explain this by introducing a new dynamical mechanism. It is a double scattering process where a particle gains energy by undergoing two deflections in two different velocity frames. From this model I am able to explain the velocity of the first observed high velocity globular cluster. In chapter 7 is presented earlier work I did with collaborators at DARK. We examined how to measure the 3D shape of X-ray galaxy clusters using the X-ray signal only. This have implications for halo mass measurements from which our cosmological model can be constrained.

Enjoy!

2

EARLY DARK ENERGY - FORECASTS

Model Independent Early Expansion History and Dark Energy

Johan Samsing[1,2], **Eric V. Linder**[1,3], **Tristan L. Smith**[1]

[1] Berkeley Center for Cosmological Physics & Berkeley Lab, University of California, Berkeley, CA 94720, USA
[2] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark
[3] Institute for the Early Universe WCU, Ewha Womans University, Seoul, Korea

Abstract

We examine model independent constraints on the high redshift and prerecombination expansion history from cosmic microwave background observations, using a combination of principal component analysis and other techniques. This can be translated to model independent limits on early dark energy and the number of relativistic species $N_{\rm eff}$. Models such as scaling (Doran-Robbers), dark radiation ($\Delta N_{\rm eff}$), and barotropic aether fall into distinct regions of eigenspace and can be easily distinguished from each other. Incoming CMB data will map the expansion history from $z = 0$–10^5, achieving subpercent precision around recombination, and enable determination of the amount of early dark energy and valuable guidance to its nature.

2.1 SUMMARY

Current observations are in agreement with a dark energy (DE) component behaving as a vacuum energy having negative pressure and constant energy density. In this picture the DE density is greatly suppressed at earlier times relative to the background. However, several physical mechanisms could give rise to an early DE excess, e.g. dilaton models, k-essence or dark radiation. Allowing for a non-standard DE in the analysis, could therefore lead to fundamental breakthroughs in our understanding of the acceleration of the universe.

To investigate this further, I first studied how we optimally extract information about a possible existence and evolution of an early DE component. I used the CMB power spectrum as observable which is currently the best probe of early time physics. An evolving DE affects this spectrum both through its possible clustering and its varying energy density which impacts the expansion. The expansion changes, e.g., the evolution of the horizon, the time of matter-radiation equality and the distance to last scattering, where the clustering affects the growth of perturbations by its inhomogenoues density distribution.

I modeled the DE as a fluid that only interacts with matter through gravity, but in contrast to the standard LCDM model, I allowed a completely free functional form for its density, $\rho_{DE}(t)$. I properly modified CAMB and used it to solve for the coupled background evolution and perturbations equations including our new DE model specified by $\rho_{DE}(t)$. The resultant CMB spectra including temperature and polarization, was then used in a Fisher code I wrote specifically for this project. In a first paper (first below), my collaborators and I investigated which functional forms for the DE density $\rho_{DE}(t)$ the new Planck data would allow us to constrain the best - before the data was released. From this we estimated an optimal way of binning the function $\rho_{DE}(t)$ in scale factor a into an MCMC tractable 5 bins. This was done by a joint Fisher information matrix and principle component analysis. In a following paper (second below), this optimal binning was then used in an analysis where we measured ρ_{DE} using the new Planck data. This gave model independent constraints equivalent to $\Delta N_{\nu,\mathrm{eft}} \sim 0.3$.

2.2 INTRODUCTION

The expansion history of the universe is a fundamental property of cosmology, reflecting the energy density constituents and their evolution. Yet remarkably little is known in detail about it, other than in a coarse grained average. For redshifts between 3000 and 10^9, the universe was mostly radiation dominated, for redshifts between 3000 and ~ 1 it was mostly matter dominated, but excursions are possible – in the effective number of relativistic species $N_{\rm eff}$ say, or even temporary breakdown of such domination – and the level of subdominant components is not well constrained. Only around the epoch of primordial nucleosynthesis and of recombination is the expansion rate (Hubble parameter) better constrained, but even there at the $\sim 5\%$ level averaged over the epoch (1; 2).

Given the importance of the expansion history, and the improvement in cosmic microwave background (CMB) data, we investigate what constraints can be placed on it in a model independent way, i.e. other than fitting for a deviation of a particular functional form such as extra $N_{\rm eff}$ or a specific dark energy model. This would fill in a vast range of cosmic expansion where almost no precise constraints have been placed. That is, an error band for the Hubble parameter $H(z)$ at $z > 1000$ should be a staple of cosmology textbooks, and yet does not exist.

The early expansion history has an important bearing on understanding the nature of dark energy as well, the question of *persistence* of dark energy. For a cosmological constant Λ, the dark energy density contributed at recombination is $\Omega_\Lambda \approx 10^{-9}$, while the current upper limit from data is above 10^{-2}. This gives substantial unexplored territory. Moreover, the current constraints use a specific functional form for the dark energy evolution (usually the Doran-Robbers form (12)), but other models could lead to significantly different limits (4). Thus, model independent limits on early dark energy are needed. Physics origins for early dark energy can be quite diverse, e.g. from dilaton models (as in some string theories) to k-essence (noncanonical kinetic field theories) to dark radiation (as in some higher dimension theories) (5). Establishing whether CMB observations could distinguish these classes is another important question.

Improvement of CMB data recently by higher resolution observations extending the temperature power spectrum to multipoles $\ell \approx 3000$ by the Atacama Cosmology Telescope (ACT (8)) and South Pole Telescope (SPT (9)) gives valuable leverage since higher multipoles are sensitive to modes crossing the cosmological horizon at earlier times. This advance was used in (4) to rule out in a model independent manner the presence of any epoch of cosmic acceleration between $z \approx 2$ and 10^5 (supplementing the limits from growth of structure post-recombination in (5)). Upcoming Planck and ground based polarization experiment data will also map out the polarization power spectra, giving additional constraints.

To carry out a model independent analysis of the early expansion history, we use a combination of redshift binning and principal component analysis. In Sec. 2.3 we lay out the methodology for describing arbitrary $H(z)$. Analyzing the results in Sec. 2.4, we identify the redshifts ranges where the CMB observations are most sensitive to expansion variations. We project three classes of models representing different physical origins onto the eigenmodes to explore the discriminating power of the data in Sec. 2.5. In Sec. 4.6 we discuss the results and future prospects.

2.3 EXPANSION HISTORY

The expansion Friedmann equation directly relates the expansion rate of the universe, or Hubble parameter, to the energy density constituents,

$$H^2(a) = \frac{8\pi G}{3} \sum \rho_i(a) \; , \tag{2.1}$$

where we neglect curvature (from a theoretical prior for flatness and because we mostly treat high redshift where it would be negligible). At high redshift the canonical expectation is that the universe is matter or radiation dominated, so we write

$$\begin{aligned} H^2(a) &= \frac{8\pi G}{3}\left[\rho_m(a) + \rho_r(a) + \rho_\Lambda\right] + \delta H^2(a) && (2.2)\\ &= H^2_{\rm fid} + \delta H^2(a) && (2.3)\\ &= H^2_{\rm fid}\left[1 + \delta(a)\right] \; . && (2.4) \end{aligned}$$

Deviations $\delta(a)$ to the fiducial expansion rate can also be interpreted as an effective dark energy density differing from that of the cosmological constant, with

$$\rho_{de}(a) = \rho_\Lambda + \frac{3H^2_{\rm fid}}{8\pi G}\,\delta \; . \tag{2.5}$$

We can write the dark energy density evolution as

$$\begin{aligned} \rho_{de}(a) &= \rho_{de,0}\, f(a) = \rho_\Lambda\, f(a) && (2.6)\\ f(a) &= 1 + \left(1 + \frac{\rho_{bg}}{\rho_\Lambda}\right)\delta \; , && (2.7) \end{aligned}$$

where ρ_{bg} is the background energy density excepting dark energy, i.e. usually the dominant component, matter or radiation. More simply, the fraction of critical density contributed by the effective dark energy is

$$\Omega_{de}(a) = \frac{8\pi G\rho_\Lambda}{3H^2_{\rm fid}\,(1+\delta)} + \frac{\delta}{1+\delta} = \frac{\Omega_\Lambda(a) + \delta}{1+\delta} \; . \tag{2.8}$$

We can readily see that at high redshift we obtain a fractional early dark energy density contribution of approximately $\delta(a)$, for $\Omega_\Lambda(a) \ll \delta \ll 1$. During epochs when δ is constant, this is a constant fractional contribution.

Our goal is to analyze constraints on the variations $\delta(a)$ from the canonical model with $\delta = 0$. We begin by writing $\delta(a)$ as a linear combination in an orthogonal bin basis,

$$\delta(a) = \sum \beta_i b_i(a) \; , \tag{2.9}$$

where $b_i(a)$ is a tophat of amplitude 1 over a given range of scale factor a, and 0 otherwise. That is, the Hubble parameter deviations $\delta(a)$ are given as a linear combination of piecewise constant values. We can then constrain $H(a)$ in bins of a, a model independent description. The bin basis is also the standard first step in principal component analysis (see, e.g., (10)), as we will pursue in the next section. We choose N bins per decade of scale factor over the range of $\log a = [-5, 0]$, beginning with $N = 20$.

Since we are interested in the expansion history we deal directly with the Hubble parameter (or effective dark energy density). Treating bins of the dark energy equation of state, or pressure to density, ratio $w(z)$ would have some drawbacks here. Most severe is that to obtain $H(z)$ one must integrate $w(z')$ over all redshifts from zero to z. This makes it difficult to explore the early expansion history in a model independent manner. Moreover, the instantaneous $w(z)$ is not fully informative: during matter domination, for example, any level of dark energy density from $\Omega_{de} = 10^{-9}$ to 10^{-2} or whatever that scales as the matter has $w = 0$. Thus we aim to derive constraints directly on variations in $H(z)$, and consider the interpretation of these as a further step.

The expansion history directly feeds into the CMB power spectra, through changing the distance scales, e.g. of the sound horizon or damping scale, and the relation of multipole ℓ (or angular scale θ) to wavenumber $k = \ell/\eta(z)$, where η is the conformal distance. It also affects the decoupling of photons from baryons and the growth of perturbations in both.

The treatment of perturbations requires some attention. The description of the cosmic expansion gives the evolution of the homogeneous background, but consistency of the field equations requires consideration of perturbations in all components of energy density. Unless the deviation in $H(z)$ is interpreted purely in terms of a cosmological constant (which indeed is purely homogeneous), spatial perturbations have to be accounted for, at least formally and generally in actual practice. The perturbation evolution equations for the additional energy density involve the quantities $w(z)$, $w'(z) = dw/d\ln a$, the initial conditions on the density perturbation, and the sound speed of the effective fluid $c_s(z)$. (One could also add a viscous sound speed or anisotropic stresses, see (11).)

The first three of these are fairly straightforward. For any deviation $\delta(a)$ one can define an effective equation of state

$$
\begin{aligned}
w &= -1 - \frac{1}{3}\frac{d\ln\rho_{de}}{d\ln a} && (2.10)\\
&= -1 - \frac{a}{3[1+\delta(1+\rho_{bg}/\rho_\Lambda)]}\left[\left(1+\frac{\rho_{bg}}{\rho_\Lambda}\right)\frac{d\delta}{da} + \frac{\delta}{\rho_\Lambda}\frac{d\rho_{bg}}{da}\right] && (2.11)\\
&\approx w_{bg} - \frac{1}{3}\frac{d\ln\delta}{d\ln a}\,, && (2.12)
\end{aligned}
$$

where the last line holds when ρ_Λ gives negligible contribution to the effective dark energy density. One can take a further derivative to obtain w'. Initial conditions are usually taken as adiabatic and stresses are taken to vanish. However, one does have to specify the sound speed. If one interprets the extra energy density as arising from quintessence, i.e. a minimally coupled, canonical scalar field, then $c_s = 1$. In general, the necessary inclusion of perturbations in whatever is the origin of the deviations in the expansion history prevents a purely model independent treatment – one has to make some assumptions about the physics. Here we fix c_s to that for the particular cases we consider, but in future work we will fit for it.

We modified CAMB (13) to allow a general $H(z)$, with the $w(z)$ that goes along with this. We then solve the coupled background evolution, and photon, matter, and effective dark energy perturbations equations to obtain the CMB power spectra. For evaluating binned $H(z)$ models, using the orthogonal bin basis introduced in Eq. (2.9), we slightly smooth the bin edges, using a Gaussian smoothing of width 0.075 times the bin width, to prevent infinite derivatives. We extensively test convergence and numerical stability of the results (also see (4) where this procedure was found to be robust).

Figure 2.1 shows the bins in log scale factor (20 bins per decade) and the conversion to multipole space (overlaid with the CMB temperature power spectrum) by $\ell \approx \eta_0/\eta(a)$, which approximately relates a given wavenumber $k \approx \ell/\eta_0$ to the time it entered the horizon. Note that a uniform binning in $\log a$, which is the expected characteristic scale for physical variations in the expansion, is not uniform in multipole space.

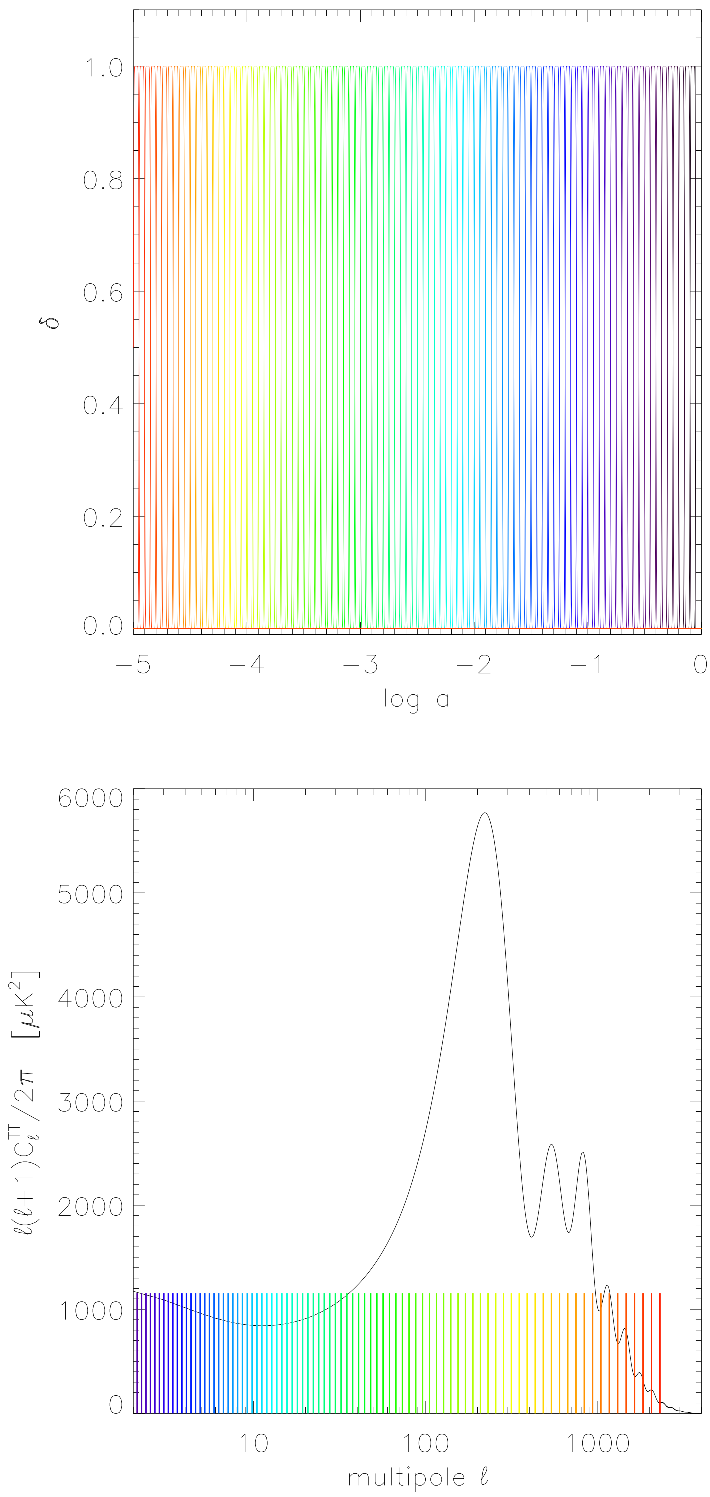


Figure 2.1 [Top] The bin basis for expansion variations $\delta(a_i)$ is plotted vs $\log a$. [Bottom] The scale factors a_i of the center of each bin in the top panel (with matching colors) are approximately mapped into multipoles by $\ell \approx \eta_0/\eta(a_i)$, with the CMB temperature power spectrum overplotted. (Amplitudes of the colored lines are arbitrary.)

To place constraints on allowed deviations $\delta(a)$ we carry out a Fisher matrix calculation. The Fisher matrix elements are given by

$$F_{ij} = \sum_l \sum_{X,Y} \frac{\partial C_{Xl}}{\partial p_i} (COV_l)^{-1}_{XY} \frac{\partial C_{Yl}}{\partial p_j} \tag{2.13}$$

where X, Y is any combination of the CMB temperature power spectrum (T), E-mode polarization power spectrum (E), and temperature-polarization cross power spectrum (TE). The covariance matrix COV is given by the measurement uncertainties of the CMB observations; we adopt the characteristics of the Planck satellite experiment (13). The parameter set $\{p_i\}$ includes the usual cosmological parameters – the physical baryon density $\Omega_b h^2$, physical cold dark matter density $\Omega_c h^2$, total present matter density Ω_m (the present Hubble constant h is a derived quantity), primordial scalar perturbation power law index n_s, optical depth τ, and present amplitude of mass fluctuations σ_8 – and the $N_{\rm bin}$ expansion variation parameters $\delta(a_i)$. The uncertainties on each $\delta(a_i)$ are given by the (square root of the) respective diagonal element of the inverse of the Fisher matrix.

Figure 2.2 shows the sensitivity $\sqrt{F_{ii,\ell}}$ of the weighted combination of CMB power spectra (T, E, TE) to the expansion deviations in each redshift bin for each multipole. Sensitivity peaks around the acoustic peaks and is reduced at low multipoles due to cosmic variance and at high multipoles due to the finite resolution from the instrument beam size. Since the polarization power spectrum is out of phase with the temperature power spectrum, dips in the temperature sensitivity are filled in by polarization information.

Figure 2.3 shows the actual Fisher and covariance submatrices corresponding to the expansion bin parameters (marginalized over other parameters in the case of the covariance matrix). First, we notice the maximum of the information content is near decoupling ($\log a \approx -3$), as expected. Earlier times, $\log a < -4$, map to multipoles on the damping tail and so have less leverage, while recent times, $\log a > -3$, are on the Sachs-Wolfe plateau and again have limited information. The Fisher matrix is not diagonal because expansion deviations affect all later times, e.g. perturbation evolution once the wavemode is within horizon and integral quantities such as the sound horizon. This will be one of the motivating factors for carrying out principal component analysis (PCA) in Sec. 2.4. The covariance matrix (inverse of the Fisher matrix), however, has a more diagonal structure, and so bins can be a useful parameter set if carefully chosen (see Sec. 4.6).

Most importantly, Figure 2.4 shows the constraints on the expansion history

$$\frac{\sigma(H)}{H_{\rm fid}} = \frac{\sigma(\delta)}{2\sqrt{1+\delta}} \,, \tag{2.14}$$

i.e. the fractional uncertainty on $H(a)$ due to deviations δ, marginalized over the other cosmological parameters. This is the "textbook" plot, showing the state of our knowledge of the early expansion history of the universe when given CMB data of Planck sensitivity. The constraints depend on what bandwidth we wish to constrain the expansion history: the top curve shows 10 bins per decade, the bottom curve 2 bins per decade. One can trade off sensitivity to fine features vs overall level of constraints. With 10 bins per decade one can achieve percent level constraints on $H^2(a)$ near decoupling, while with 2 bins per decade one obtains subpercent constraints over more than two decades in scale factor. The relation between 10 bins and 2 bins is not simply a $\sqrt{5}$ scaling due to correlations between bins (the offdiagonal elements of the covariance matrix), and the lowest redshift bin is particularly affected by covariance with the other cosmological parameters.

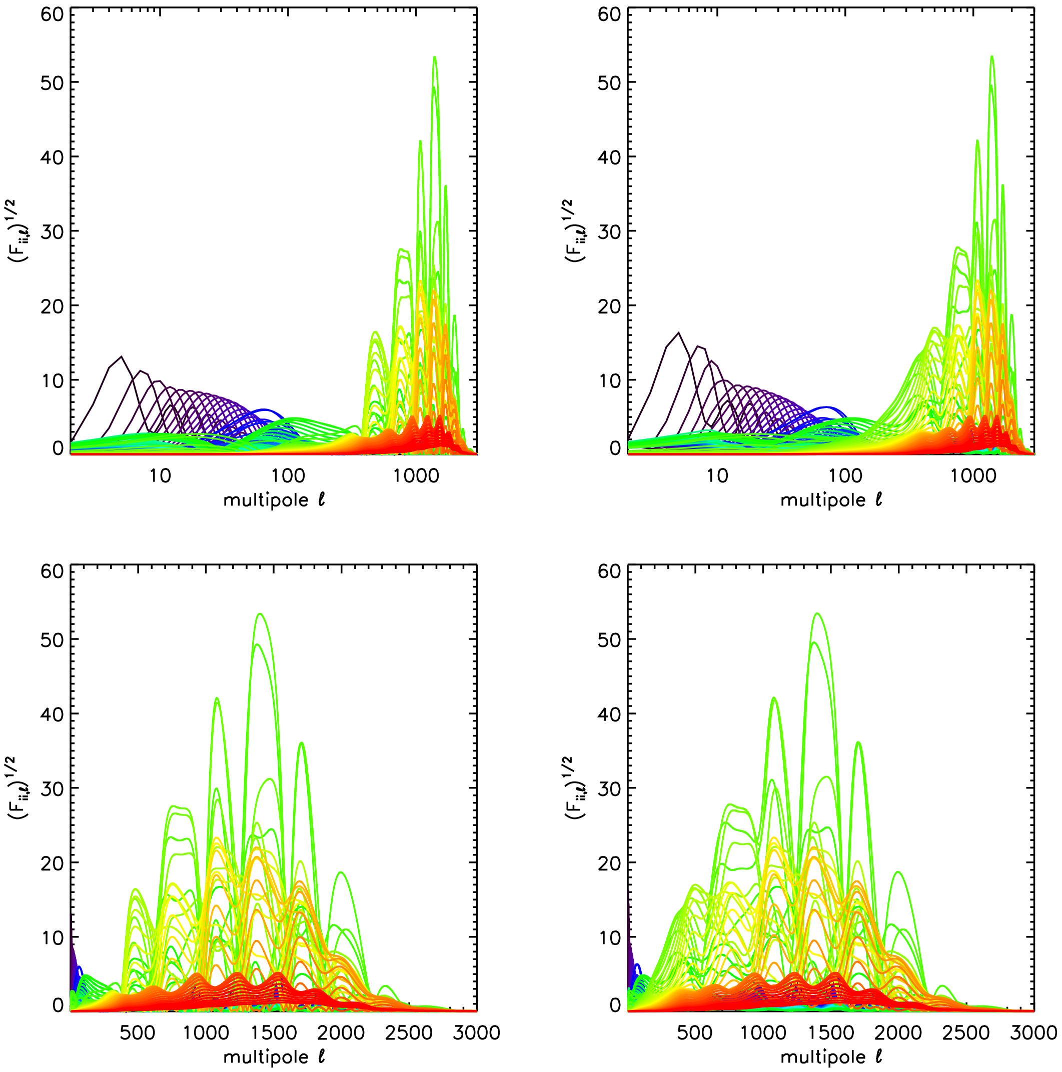


Figure 2.2 The fractional Fisher sensitivities, $\sqrt{F_{ii,\ell}}$, to the expansion variations $\delta(a_i)$ in each bin, color coded as in Fig. 2.1, are plotted vs multipole. The top panels use log scale in multipole, the bottom panels a linear scale to highlight different regions. The left panels are for the temperature power spectrum only, while the right panels use the variance weighted sum of the T, E, and TE power spectra entering the Fisher information matrix. Polarization information fills in the sensitivity gaps due to the acoustic troughs.

2.4 PRINCIPAL COMPONENTS OF EXPANSION

Expansion deviations at some redshifts may have substantially the same consequence for the observables as deviations at some other redshift, or deviations may be correlated in such a way that only the difference between them is important. This leads to the idea of compressing the 100 bins between $\log a = [-5, 0]$, or at least the information contained in them, into fewer

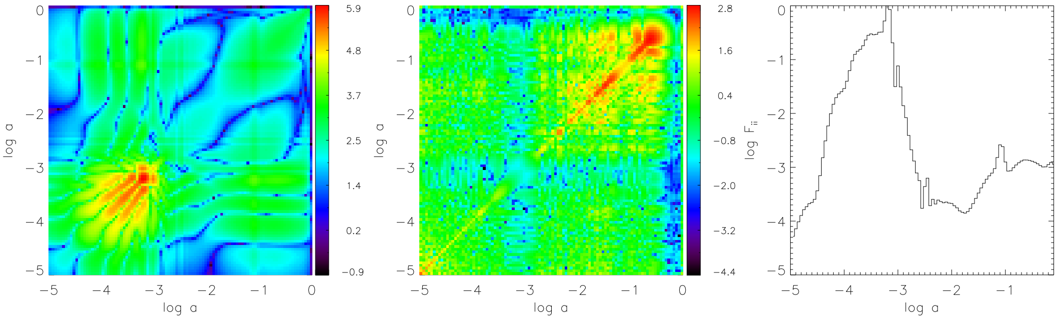


Figure 2.3 The Fisher information submatrix (left) corresponding to the expansion variation parameters $\delta(a_i)$ is plotted with redder shades representing larger absolute values (more information). The color bar gives the log of the absolute value of elements. The covariance matrix (middle), marginalized over other cosmological parameters, follows the same color scheme, so the best determined parameters (smallest errors) are bluest. Diagonal elements of the Fisher matrix (right), as a fraction of the largest diagonal element, quantify for which redshifts the CMB data is most sensitive to the expansion history. The bump at $\log a \approx -1.1$ reflects reionization.

variables. One might for example speculate that the major effect of expansion deviations for $a > 10^{-3}$ comes from shifting the distance to CMB last scattering, and whether the variation occurs at $a = 0.01$ or $a = 0.1$ is less crucial.

Principal component analysis (PCA) can provide an efficient way to compress the influence on the observables. For some uses probing dark energy and the CMB, see for example (14; 10; 15; 16; 17; 18) (and dark matter and the CMB in (19)). By diagonalizing the Fisher matrix we can find its eigenvectors that best summarize the sensitivity of the observations to the expansion deviations. We can then transform the bin basis to an orthogonal eigenmode basis of principal components (PCs), writing

$$\delta(a) = \sum m_i \, e_i(a) \ , \qquad (2.15)$$

where m_i is the amplitude of mode i and $e_i(a)$ is the eigenvector. Since the modes are orthogonal, the errors $\sigma_i \equiv \sigma(m_i)$ on the amplitudes are uncorrelated. Using the entire set of bins or the entire set of PCs is equivalent, but using only a few PCs with the highest eigenvalues (smallest σ_i) in general allows one to approximate the full set more accurately than the same number of bins. That is, the information can be compressed.

Figure 2.5 illustrates some of the PCs for the CMB Fisher matrix, ordered from highest to lowest eigenvalues (best to worst determined). Note that as expected most of the activity in the first PCs is prerecombination, associated with the acoustic peaks. In modes 7-10, there is some low redshift action coming primarily from the integrated distance to last scattering and the reionization epoch. Higher PC modes tend to be more oscillatory (essentially high derivatives of the expansion behavior) and localized.

By taking the cumulative sum of the eigenvalues, we find that the first 7 PCs contain 99% of the variance. That is, $\sum_1^7 \sigma_i^{-2} / \sum_1^{100} \sigma_i^{-2} = 0.99$. This means that the great majority of expansion behaviors, as far as their observational detectability is concerned, can be described with just 7 parameters $m_1, \ldots m_7$.

It is convenient to normalize the PCs such that $\sum_n e_i(a_n) e_j(a_n) = \delta_{ij}$, where δ_{ij} here is the

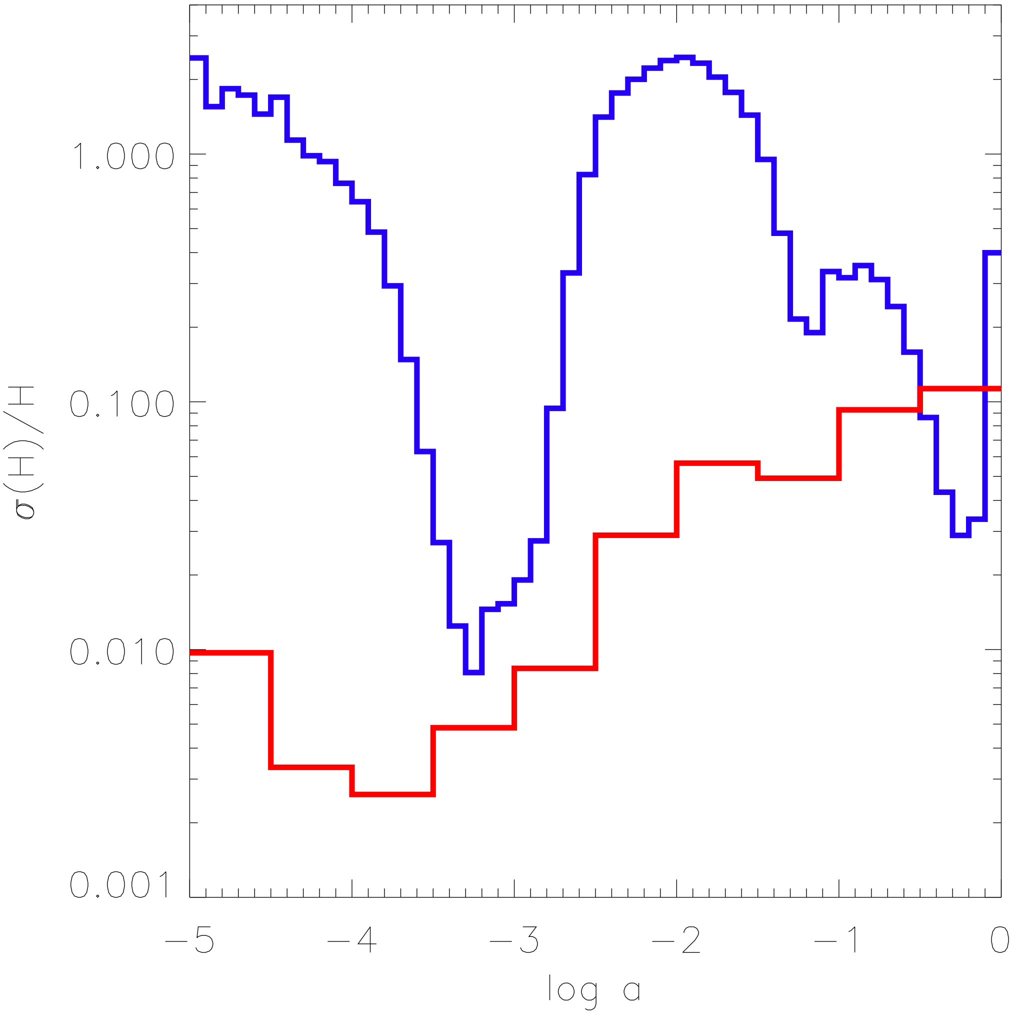


Figure 2.4 The fractional precision with which the expansion history can be determined by projected Planck CMB data is plotted vs scale factor, for two different bandwidths. The top (bottom) curve is for 10 (2) bins per decade in scale factor. Subpercent precision can be achieved around decoupling but large swaths of the cosmic history will remain unknown.

Kronecker delta, and then the mode coefficients m_i give the amplitudes for a given model of expansion history deviation,

$$m_i = \sum_n \delta(a_n) e_i(a_n) \,, \tag{2.16}$$

where a_n denotes the bin centers. For many narrow bins the sum can be converted to an integral. The amplitudes m_i only have meaning when discussing a specific model; note the canonical model ΛCDM has all $m_i = 0$. Since the mode amplitude m_i has no a priori magnitude, we do not know whether its uncertainty $\sigma_i = 0.01$, say, is a small or large number. We can compare

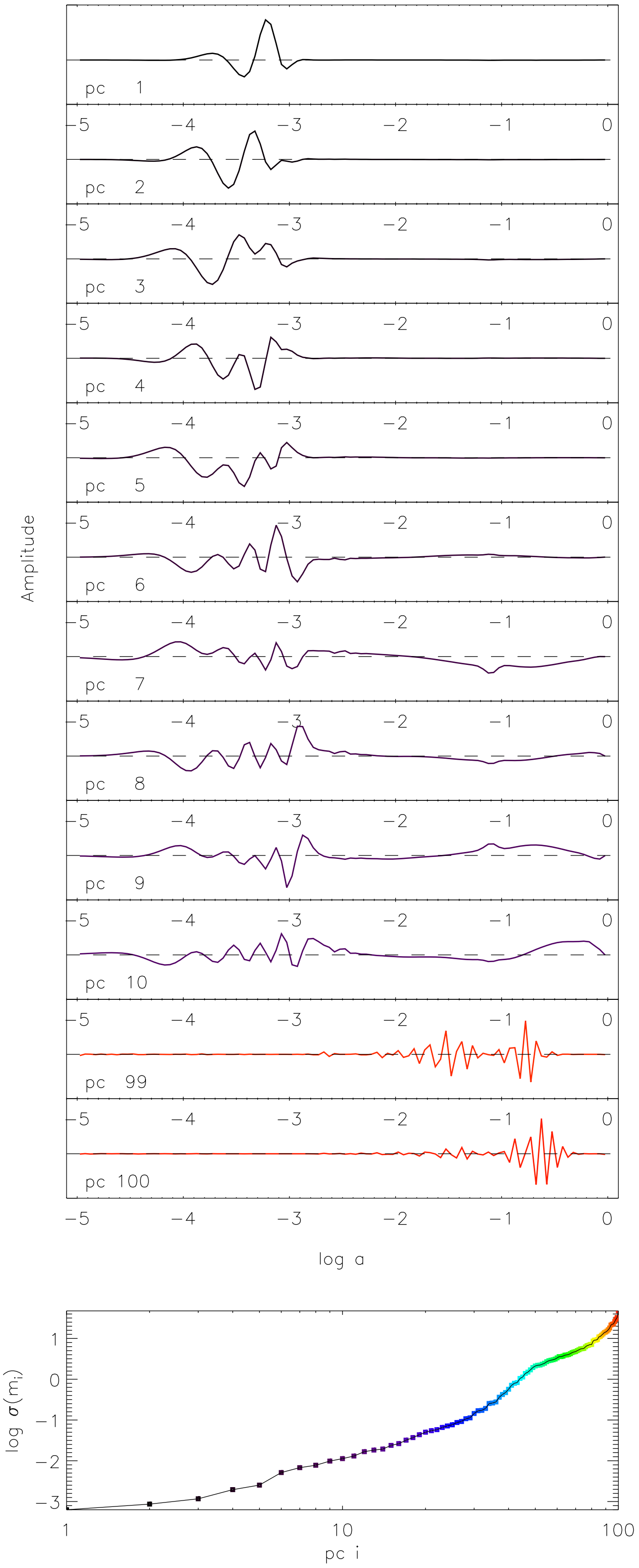


Figure 2.5 Principal components of the CMB observational sensitivity to expansion history are plotted vs $\log a$, for the first 10 and last 2 modes. The bottom plot shows the σ_i for all modes.

σ_i's to each other (but if the m_i are 0 this is irrelevant), or for a specific model we can form a signal to noise ratio for a mode,

$$(S/N)_i = m_i/\sigma_i \ . \tag{2.17}$$

If $(S/N)_i \ll 1$, then it does not matter if σ_i looks small, the mode cannot be well measured. Thus caution must be used in interpreting PCA (for more details see (17; 20)).

Although PCA certainly compresses the information content into fewer parameters, the number of PCs required to describe arbitrary expansion histories is still large for the purposes of, say, Monte Carlo simulations. We can therefore use PCA instead as a guide in two ways: to examine the ability to discriminate among different classes of models for the expansion deviation, and to indicate which regions in a show the most sensitivity to deviations and hence which of the original bins are most useful. These are treated respectively in Sec. 2.5 and Sec. 4.6.

2.5 COMPARING EARLY DARK ENERGY MODELS

While we wish to concentrate, as much as possible, on a model independent approach to constraining the expansion history, it is useful to make contact with various classes of models to make sure that important behaviors are captured. In broad strokes, one can consider cases where the expansion deviations decline at times earlier than recombination, increase at earlier times, or remain constant. This can be translated into early dark energy models that contribute a lower fraction of the total energy density in radiation vs matter domination, a higher fraction, or a constant fraction. This tilt of the expansion rate can be an important discriminant, somewhat analogous to the tilt of the primordial power spectrum for inflation.

Early dark energy models have been proposed with each of these behaviors. Examples of the three classes respectively are 1) Barotropic dark energy with sound speed $c_s^2 = 0$ (21), sometimes called aether models, 2) Barotropic dark energy with sound speed $c_s^2 = 1/3$ (21), sometimes called dark radiation, and 3) Scaling dark energy such as the commonly used Doran-Robbers model with $c_s^2 = 1$ (12).

Barotropic models have $w(a)$, $\rho_{de}(a)$, and hence $\delta(a)$, all determined by $c_s^2(a)$. The dynamics is given by (21)

$$w' = -3(1+w)(c_s^2 - w) \ , \tag{2.18}$$

with solution for constant c_s of

$$w(a) = \frac{c_s^2 B a^{-3(1+c_s^2)} - 1}{B a^{-3(1+c_s^2)} + 1} \ , \tag{2.19}$$

where $B = (1+w_0)/(c_s^2 - w_0)$. For the $c_s^2 = 0$ aether model,

$$\begin{aligned}
\rho_{ae}(a) &= \rho_\infty + \Omega_e \rho_{m,0} a^{-3} && (2.20)\\
\delta_{ae}(a) &= \Omega_e \frac{\Omega_m(a^{-3}-1)}{\Omega_m(a^{-3}-1) + \Omega_r a^{-4} + 1} && (2.21)\\
\rho_\infty/\rho_{crit,0} &= 1 - \Omega_m(1+\Omega_e) = (-w_0)(1-\Omega_m) && (2.22)\\
w_0 &= -1 + \frac{\Omega_e \Omega_m}{1-\Omega_m} \ , && (2.23)
\end{aligned}$$

where during matter domination the dark energy contributed a constant fractional density Ω_e, but this declines at earlier times as radiation becomes important.

For the $c_s^2 = 1/3$ dark radiation model,

$$\begin{aligned} \rho_{dr}(a) &= \rho_\infty + \Omega_e \rho_{r,0} a^{-4} && (2.24) \\ \delta_{dr}(a) &= \Omega_e \, \frac{\Omega_r(a^{-4}-1)}{\Omega_r(a^{-4}-1)+\Omega_m(a^{-3}-1)+1} && (2.25) \\ \rho_\infty/\rho_{crit,0} &\approx 1-\Omega_m && (2.26) \\ w_0 &\approx -1 \, , && (2.27) \end{aligned}$$

where during radiation domination the dark energy contributed a constant fractional density Ω_e, but this declines at later times as matter becomes important.

The most commonly used early dark energy is the Doran-Robbers form (12),

$$\Omega_{de}(a) = \frac{1-\Omega_m-\Omega_e(1-a^{-3w_0})}{1-\Omega_m(1-a^{3w_0})} + \Omega_e(1-a^{-3w_0}) \, , \qquad (2.28)$$

where during both matter and radiation domination the dark energy contributed a constant fractional density Ω_e. The sound speed is conventionally taken to be $c_s^2 = 1$.

Thus the three models we consider have expansion history deviations with complementary behaviors: rising, falling, and constant. This range should give a good indication of how PCA can characterize the expansion, while also having physical motivations. Of the physics origins mentioned in the Introduction, some string theories give Doran-Robbers behavior, some non-canonical kinetic fields give aether behavior, and some higher dimension theories give dark radiation.

Once we calculate the m_i for a model through Eq. (2.16), we have a better indication of the importance of a PC mode through the signal to noise criterion of Eq. (2.17) – recall that the σ_i alone say little about whether the mode is relevant. To choose the number of modes to keep in describing a model, we can simply impose a S/N cutoff, or ask that the cumulative S/N of the modes kept be above some fraction, e.g. 95% of the total S/N from all modes.

Another indicator is the statistical risk, or mean squared error. This takes into account that more modes decrease the bias from the true model, but increase the accumulated variance in δ. The risk R_N is

$$\begin{aligned} R_N &= \sqrt{b_N^2+\sigma_N^2} && (2.29) \\ b_N^2 &= \sum_n [\delta^{\rm model}(a_n) - \delta^{N\,{\rm PCs}}(a_n)]^2 && (2.30) \\ \sigma_N^2 &= \sum_n \sum_{i=1}^{N} \sigma_i^2 e_i^2(a_n) \, . && (2.31) \end{aligned}$$

One could choose to keep that number N PCs where R_N is minimized.

Figure 2.6 shows for each of the three models the approximation to $\delta(a)$ as more PCs are added, the values m_i, and the cumulative S/N and risk as a function of number of PCs used. In general we find the risk requires more PCs than the S/N criterion; this makes sense since S/N concentrates on those PCs fitting the observable (CMB power spectra) while risk attempts to fit the unobservable expansion deviation. Another drawback to risk is that it does not scale with the amplitude of the modes, i.e. while S/N increases linearly with Ω_e, the bias term in the risk scales but the variance does not, so the risk is weighted unevenly depending on deviation amplitude.

For the dark radiation, aether, and Doran-Robbers models, respectively, we should keep the first 6, 7, and 5 PCs according to S/N, and 10, 14, and 15 PCs according to risk. However, we note that this is if we keep the PCs in order according to σ_i. If we choose the highest S/N modes individually, we only require 4, 3, and 3 modes to attain 95% of the full S/N (but this requires knowing the correct model, or assuming a given set of models).

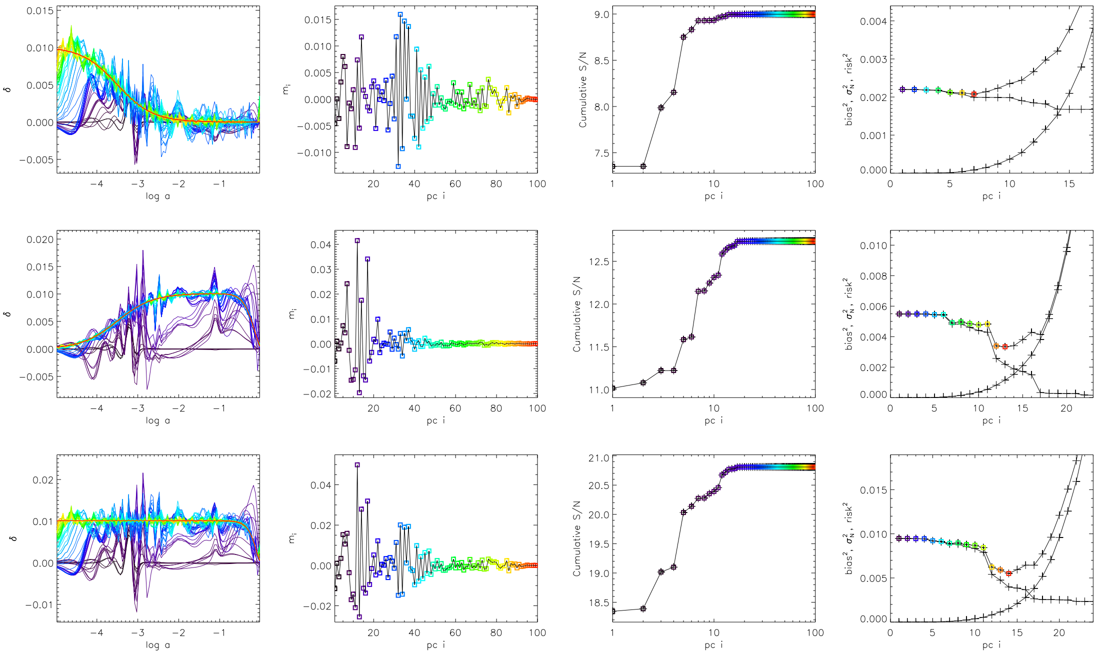
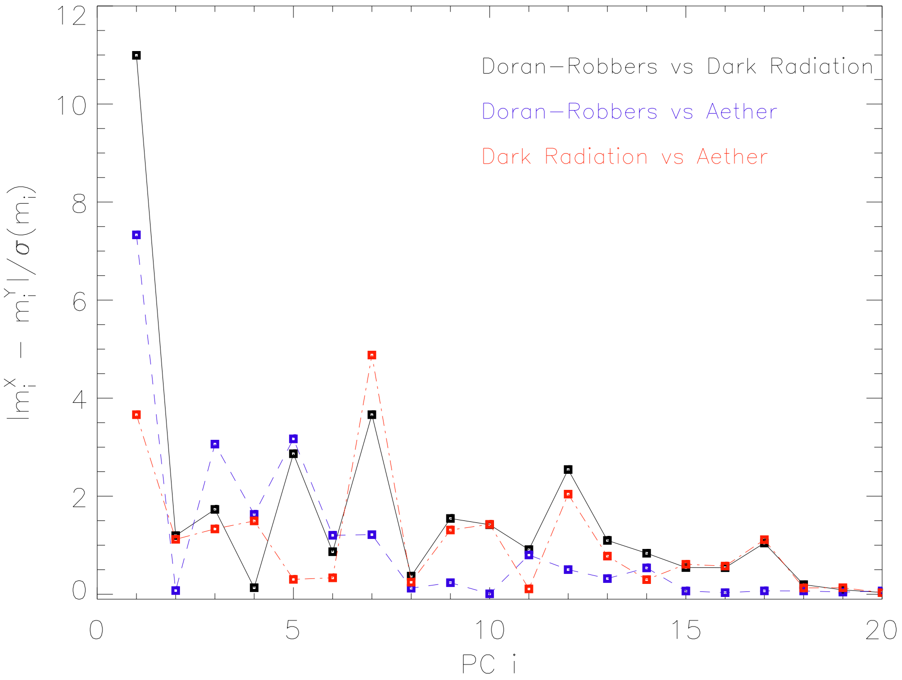


Figure 2.6 Three models of early dark energy are analyzed – dark radiation (top row), aether (middle row), and Doran-Robbers (bottom row) – with different recombination era behaviors. The leftmost column shows $\delta(a)$ built up out of PCs, with the thick red line giving the exact model. The second column gives the PC amplitudes m_i and the third column shows the cumulative S/N, summing Eq. (2.17) over the first i PCs in quadrature. The rightmost column shows the bias squared (falling curve), variance (rising curve), and risk squared (top curve) when including the first i PCs. One might choose to keep those PCs that either give the largest jumps in S/N, or all those up to the minimum in the risk curve.

Let us examine these models in more detail. Figure 2.7 shows how these models are well separated in eigenmode coefficient space. Considerable discriminating power occurs using modes 1 and 7, for example, with the separations between the three models many times the uncertainties σ_i. From the shape of the modes in Fig. 2.5, we can see that mode 1 is roughly measuring the amplitude of the expansion deviation at recombination, and whether it is increasing or decreasing (thus distinguishing all three models), and mode 7 is sensitive to more recent deviations such as occurring in Doran-Robber and aether, but not dark radiation, cases. Thus these two modes together are adept at distinguishing between the rising, falling, and constant deviation classes of expansion history, and early vs late deviations.

We emphasize that the PCs are really telling us about fits to the observable power spectra and not reconstruction of the expansion deviation in a fine grained sense. While Fig. 2.6 shows that ~ 50 PCs are needed to model $\delta(a)$ well for these example, Fig. 2.8 demonstrates that only ~ 10 PCs are needed to give accurate agreement in C_ℓ. Figure 2.8 shows the residual of the CMB temperature power spectrum from the sum of the first N PCs (in S/N ordering, and also



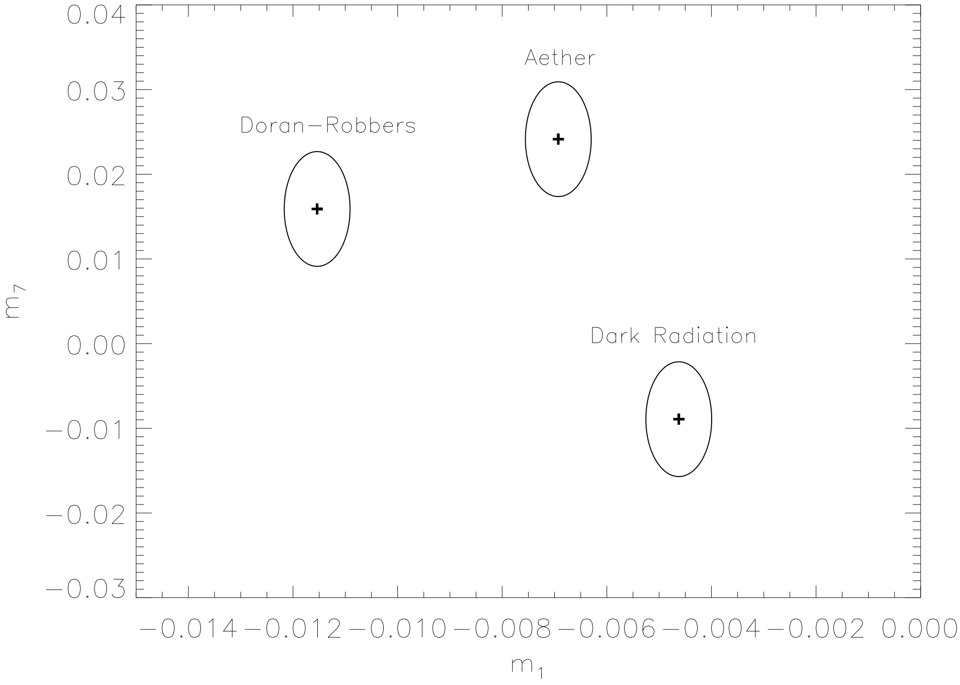


Figure 2.7 (Top) Differences between each PC amplitude m_i are shown for pairs (X, Y) of early dark energy models, with solid curves comparing Doran-Robbers to dark radiation, dashed Doran-Robbers to aether, and dot-dashed dark radiation to aether. The highest peaks indicate the modes with strong discriminating power. (Bottom) Amplitudes of modes 1 and 7 are plotted for the three models, with the +'s indicating the values m_i and the ellipses showing the uncertainties σ_i. These two modes can clearly distinguish between each of the three models.

shown in σ_i ordering for the aether model) relative to the true model for the three models, as well as the χ^2 summed over multipoles, and the reconstructed expansion history. Although the reconstructed expansion history may only agree over certain ranges of a, this can still give excellent agreement for the observables as the power spectra are not equally sensitive to all scale factors.

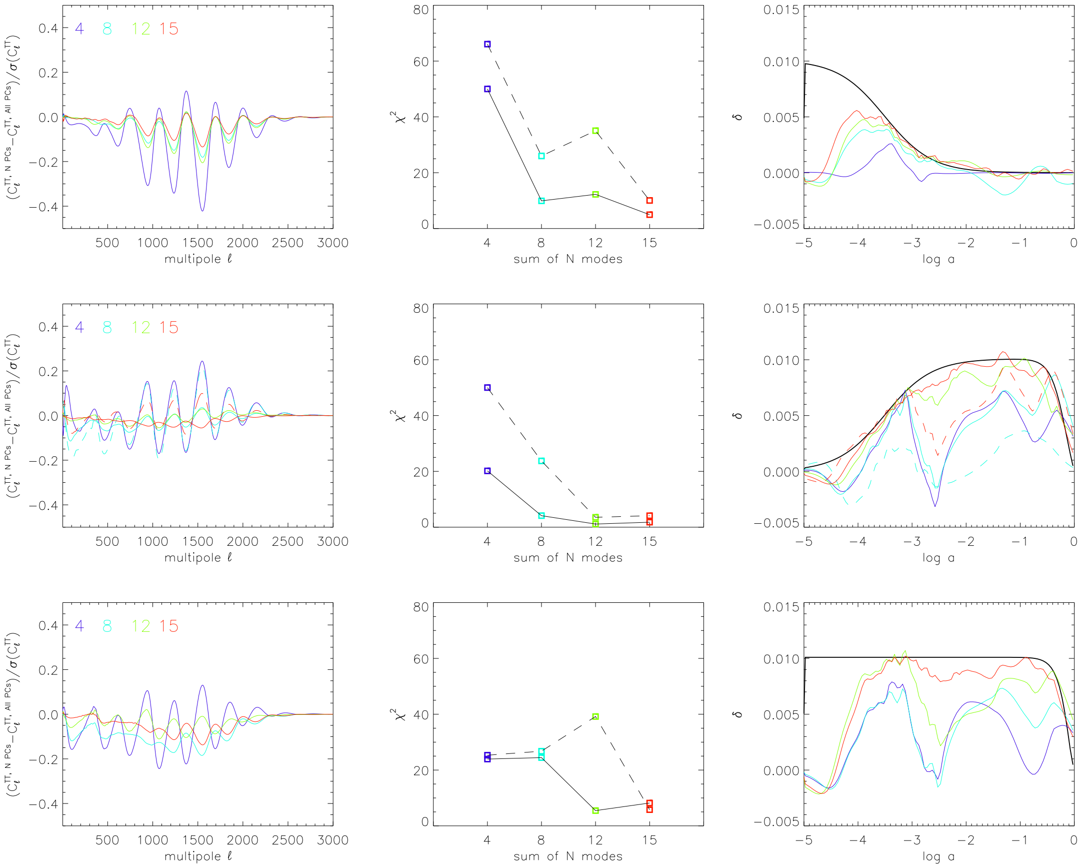


Figure 2.8 PCA of three models of early dark energy – dark radiation (top row), aether (middle row), and Doran-Robbers (bottom row) – is compared to the exact models in terms of the observable CMB power spectrum. The left column shows the deviation in the temperature power spectrum for the sum of the first 4, 8, 12, 15 modes using S/N ordering of the modes (see Fig. 2.6). The χ^2 of the deviations summing over multipoles is in the middle panels, with the dashed curve using ordering by σ_i instead. Reconstruction of the theoretical $\delta(a)$ appears in the right panels; note how only certain epochs need be fit well to reproduce the observable CMB. For the aether model (middle row) dashed curves show also the σ_i ordering results for the $N = 8, 15$ cases, which do much worse than S/N ordering.

The numbers of PCs required to ensure an accurate estimation of the observables, say $\chi^2 < 10$ (summed over 3000 multipoles), is generally greater than 10, rendering cumbersome a straight application of mode amplitudes as parameters in a Monte Carlo simulation. (And recall that for a model independent analysis we do not know a priori the S/N ordering, so one would need to include many more PCs using σ_i ordering.) Moreover, PCs per se are not always easy to interpret in terms of physical effects. Therefore it is both clearer and more efficient to use the PCA instead to guide an effective, low order binning basis.

2.6 CONCLUSIONS

Our knowledge of the expansion history of our universe, even at the level of degree of matter domination or radiation domination at early epochs, is remarkably imprecise. Cosmic microwave background radiation measurements from ACT, Planck, and SPT (and later ACTpol and SPTpol) will shed light on the times around recombination and reionization. We quantify the model independent state of our knowledge through a combination of redshift bin and principal component analysis, finding that subpercent level constraints will be placed by Planck over $\log a = [-2.5, -5]$ for a bandwidth of $\Delta \log a = 0.5$.

CMB data will address one of the key aspects of dark energy – its persistence, a characteristic of many high energy physics origins – and we find that several different classes of early dark energy are well separated in PCA space. The limits can also be interpreted in terms of the number of effective relativistic species, $N_{\rm eff}$, such as an extra neutrino type, with current data mildly preferring further contributions. A thermal relativistic neutrino species adds 23% to the photon energy density, so $\delta = 0.13\,\Delta N_{\rm eff}$, giving tight limits on extra relativistic degrees of freedom from the forthcoming data.

We explore three specific models, representing different classes for early dark energy, possibly corresponding to different physical origins. The commonly used Doran-Robbers form has a dark energy fraction Ω_e that is constant through the recombination epoch. We also investigate a dark radiation model with $\Omega_{de}(a)$ rising to the past and a barotropic aether model with $\Omega_{de}(a)$ falling to the past, and find that the dominant PC mode is well able to distinguish between these behaviors. Since the amplitude of that mode is greatest for the Doran-Robbers model, we expect that data constraints on Ω_e in the other classes will be weaker than in this model (such as from (22) using current CMB data), allowing for nonnegligible persistence of dark energy (see (4) for further demonstrations of this). Our general approach, however, does not rely on assuming the form for the new component or expansion deviation.

Figure 2.3 is in a sense the textbook picture of what Planck CMB data will say in a model independent manner about early universe expansion. For postrecombination epochs this will improve with further ground based polarization measurements (especially of CMB lensing) and inclusion of growth of structure data. Understanding early expansion is in fact crucial for accurate interpretation of large scale structure, and feeds directly into the early time gravitational growth calibration parameter $g_\star$ (23); ignorance of this can bias cosmological parameter estimation and tests of gravity.

Expansion history is not the whole story as the effective fluid behind the expansion deviations has perturbations and can have internal degrees of freedom. We treat the perturbations consistently – the dark radiation and barotropic aether models for example have sound speeds different from the speed of light. We do not include viscosity, however, as the data has poor leverage on this (24; 4; 25). Another difficulty for model independent analysis is having $\delta < 0$, since perturbations are difficult to treat when the effective density deviation passes through zero; models such as nonthermal neutrinos, with energy densities below the standard, could realize such a condition. We will consider such cases in future work.

Principal component analysis provides a valuable guide to the key epochs of sensitivity and the amount of information contributed from different times. However, we emphasize and demonstrate that the raw uncertainty σ_i on an eigenmode has very limited meaning; the first 15 modes ordered by σ_i can give a highly inaccurate reconstruction relative to a smaller number of modes ordered by signal to noise. Redshift bins can be more clearly interpreted. Employing the best aspects of each can result in physically clear, well characterized expansion history constraints.

BIBLIOGRAPHY

[1] S.M. Carroll & M. Kaplinghat, Phys. Rev. D 65, 063507 (2002) [arXiv:astro-ph/0108002]

[2] S. Galli, A. Melchiorri, G.F. Smoot, O. Zahn, Phys. Rev. D 80, 023508 (2009) [arXiv:0905.1808]

[12] M. Doran and G. Robbers, JCAP 0606, 026 (2006) [arXiv:astro-ph/0601544]

[4] E. Calabrese, D. Huterer, E.V. Linder, A. Melchiorri, L. Pagano, Phys. Rev. D 83, 123504 (2011) [arXiv:1103.4132]

[5] E.J. Copeland, M. Sami, S. Tsujikawa, Int. J. Mod. Phys. D 15, 1753 (2006) [arXiv:hep-th/0603057]

[8] S. Das, T.A. Marriage, P.A.R. Ade et al., ApJ 729, 62 (2011); J. Dunkley, R. Hlozek, J. Sievers et al., ApJ 739, 52 (2011)

[9] E. Shirokoff, C.L. Reichardt, L. Shaw et al., ApJ 736, 61 (2011); R. Keisler, C.L. Reichardt, K.A. Aird et al., ApJ 743, 28 (2011)

[4] E.V. Linder & T.L. Smith, JCAP 1106, 001 (2011) [arXiv:1009.3500]

[5] E.V. Linder, Phys. Rev. D 82, 063514 (2010) [arXiv:1006.4632]

[10] D. Huterer & G. Starkman, Phys. Rev. Lett. 90, 031301 (2003) [arXiv:astro-ph/0207517]

[11] W. Hu, ApJ 506, 485 (1998) [arXiv:astro-ph/9801234]

[13] A. Lewis, A. Challinor, & A. Lasenby, ApJ 538, 473 (2000) [arXiv:astro-ph/9911177]; `http://camb.info`

[13] Planck Collaboration, P.A.R. Ade, N. Aghanim et al., A&A, 536, A1 (2011); `http://planck.esa.int`

[14] W. Hu & T. Okamoto, Phys. Rev. D 69, 043004 (2004) [arXiv:astro-ph/0308049]

[15] S. Leach, MNRAS 372, 646 (2006) [arXiv:astro-ph/0506390]

[16] M.J. Mortonson & W. Hu, ApJ 672, 737 (2008) [arXiv:0705.1132]

[17] R. de Putter & E.V. Linder, Astropart. Phys. 29, 424 (2008) [arXiv:0710.0373]

[18] C. Dvorkin & W. Hu, Phys. Rev. D 82, 043513 (2010) [arXiv:1007.0215]

[19] D.P. Finkbeiner, S. Galli, T. Lin, T.R. Slatyer, Phys. Rev. D 85, 043522 (2012) [arXiv:1109.6322]

[20] R. de Putter & E.V. Linder, arXiv:0812.1794

[21] E.V. Linder & R.J. Scherrer, Phys. Rev. D 80, 023008 (2009) [arXiv:0811.2797]

[22] C.L. Reichardt, R. de Putter, O. Zahn, Z. Hou, ApJL 749, L9 (2012) [arXiv:1110.5328]

[23] E.V. Linder, Phys. Rev. D 79, 063519 (2009) [arXiv:0901.0918]

[24] E. Calabrese, R. de Putter, D. Huterer, E.V. Linder, A. Melchiorri, Phys. Rev. D 83, 023011 (2011) [arXiv:1010.5612]

[25] T.L. Smith, S. Das, O. Zahn, Phys. Rev. D 85, 023001 (2012) [arXiv:1105.3246]

3

EARLY DARK ENERGY - MEASUREMENTS

New Constraints on the Early Expansion History

Alireza Hojjati[1], **Eric V. Linder**[1,2], **Johan Samsing**[3]

[1] Institute for the Early Universe WCU, Ewha Womans University, Seoul 120-750, Korea
[2] Berkeley Center for Cosmological Physics & Berkeley Lab, University of California, Berkeley, CA 94720, USA
[3] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark

Abstract

Cosmic microwave background measurements have pushed to higher resolution, lower noise, and more sky coverage. These data enable a unique test of the early universe's expansion rate and constituents such as effective number of relativistic degrees of freedom and dark energy. Using the most recent data from Planck and WMAP9, we constrain the expansion history in a model independent manner from today back to redshift $z = 10^5$. The Hubble parameter is mapped to a few percent precision, limiting early dark energy and extra relativistic degrees of freedom within a model independent approach to 2–16% and 0.71 equivalent neutrino species respectively (95% CL). Within dark radiation, barotropic aether, and Doran-Robbers models, the early dark energy constraints are 3.3%, 1.9%, 1.2% respectively.

Except for the last e-fold of cosmic expansion, our knowledge of the state of the universe arises directly only through measurements of the cosmic microwave background (CMB) radiation or indirectly (as in models of its influence on growth of large scale structure). Recent CMB data (1; 2) provides a clear window on an additional 10 e-folds of history (back to redshift $z = 10^5$), a vast improvement in mapping the universe.

The expansion rate, or Hubble parameter, is a fundamental characterization of our universe, and includes information on its matter and energy components, their evolution, and the overall curvature of spacetime. Moreover, the CMB encodes linear perturbations in the photons and the gravitational potentials they experience, providing sensitivity to the microphysics of components, e.g. their sound speed.

These observations lead to constraints on quantities such as early dark energy and extra neutrino species or other relativistic degrees of freedom. However most analyses assume a specific model for these deviations, enabling stringent but model dependent constraints. In this Letter our approach is to map the cosmic state and history in as model independent fashion as practical, guided by the data. We utilize the results of the principal component analysis of (3) to define localized bins of Hubble parameter in log scale factor that are most sensitive to the data, and then carry out a Markov Chain Monte Carlo (MCMC) analysis to constrain them. Finally we discuss the implications for dark energy, relativistic degrees of freedom, and spacetime curvature. For data we use the most recent CMB results from the Planck satellite (1) and WMAP satellite (2).

Cosmic History Mapping – For robust, model independent results we adopt a combination of principal component analysis (PCA) and binning. This avoids assuming a specific functional form for the Hubble parameter or dark components and allows the data itself to inform where the greatest sensitivity lies. Such PCA on the Hubble parameter for projected mock CMB data was used in (3) to predict the strength of constraints at various epochs of cosmic history.

This identification of the key epochs where physical conditions most affect the observations enables informed choice of bins in log scale factor to use in a MCMC fit. Bins have several advantages over the raw PCA: 1) they are localized and can be clearly interpreted physically – the Hubble parameter during a specific epoch, 2) they avoid negative oscillations that can cause unphysical results (while the sum of all PCs will give a positive, physical Hubble parameter squared, this is not guaranteed for a subset), and 3) they are well defined, not changing when new data is added.

The Hubble parameter, or logarithmic derivative of the scale factor, $H = d\ln a/dt$, is then written as

$$H^2(a) = \frac{8\pi G}{3}\left[\rho_m(a) + \rho_r(a) + \rho_\Lambda\right]\left[1 + \delta(a)\right], \tag{3.1}$$

where δ accounts for any variation from ΛCDM (cosmological constant plus cold dark matter plus standard radiation) expansion history, and ρ_m is the matter density, ρ_r the radiation density, and ρ_Λ the cosmological constant density. The bins in the deviation $\delta(a)$ are slightly smoothed for numerical tractability, with

$$\delta = \sum_i \delta_i \left[\frac{1}{1 + e^{(\ln a - \ln a_{i+1})/\tau}} - \frac{1}{1 + e^{(\ln a - \ln a_i)/\tau}}\right]. \tag{3.2}$$

Within bin i, $\delta = \delta_i$ and far from any bin $\delta = 0$. A smoothing length $\tau = 0.08$ was adopted after numerical convergence tests. (A similar binned approach was used in (4; 5) to bound early cosmic acceleration.)

We modify CAMB (13) to solve the Boltzmann equations for the photon perturbations in this cosmology. The dark energy density contributed by the deviations δ and the cosmological constant term (which becomes negligible at high redshift) has an effective equation of state

$$1+w=\frac{Q\delta}{1+\delta(1+Q)}\,(1+w_{bg})-\frac{1}{3}\frac{1+Q}{1+\delta(1+Q)}\frac{d\delta}{d\ln a}\,, \tag{3.3}$$

where $Q=(\rho_m+\rho_r)/\rho_\Lambda$ and w_{bg} is the background equation of state of the combined matter and radiation (e.g. $1/3$ during radiation domination, transitioning to 0 during matter domination). Thus w and $w'=dw/d\ln a$, entering into the Boltzmann equations, are defined fully by Eq. (3.2) for δ. We choose the associated sound speed to be the speed of light, as in quintessence dark energy, but explore variations of this later.

Guided by the PCA of (3), where the first few PCs show greatest sensitivity in $\log a\in[-4,-2.8]$, we choose bins δ_{1-5} in the logarithmic scale factor $\log a=[-5,-4]$, $[-4,-3.6]$, $[-3.6,-3.2]$, $[-3.2,-2.8]$, $[-2.8,0]$ so the finest binning is near CMB recombination at $a\approx 10^{-3}$. (Future CMB data could change the PCs, but we could keep the same bins, or not.) The cosmological parameters we fit for are the six standard ones: physical baryon density $\Omega_b h^2$, physical cold dark matter density $\Omega_c h^2$, acoustic peak angular scale θ, primordial scalar perturbation index n_s, primordial scalar amplitude $\ln(10^{10}A_s)$, and optical depth τ, plus the five new deviation parameters δ_{1-5}. Additional astrophysical parameters enter from the data, as discussed next.

Constraints – To constrain the cosmology with the data we use MCMC analysis, modifying CosmoMC (7). The likelihood involves the temperature power spectrum from the two satellite experiments, and the E-mode polarization and TE cross spectrum from WMAP (the first Planck likelihood release does not include polarization, or the high multipole likelihoods from Atacama Cosmology Telescope (8) or South Pole Telescope (9); in the future such data should become available). Astrophysical nuisance parameters characterizing foregrounds (see (1)) are marginalized over.

Figure 3.1 shows the constraints on the standard cosmological parameters, in the ΛCDM case (fixing $\delta_i=0$) and when allowing variations in the expansion history (fitting for the δ_i). Here the Hubble constant H_0 replaces the θ parameter and we omit showing τ. Including the fitting for expansion history deviations induces roughly a factor of two larger marginalized estimation uncertainties for most of the standard cosmology parameters, and significantly shifts the cold dark matter density value. This is due to the deviations in the Hubble parameter having similar effects on the expansion near recombination to those in matter, so δ takes the place of some of ρ_m. We discuss this degeneracy further later. The best fit for the ΛCDM case remains within the 68% confidence contour when allowing expansion deviations.

Figure 3.2 shows the constraints on the expansion history deviations. Note that to ensure positive energy density (and Hubble parameter squared) we restrict $\delta\geq 0$, i.e. equal or more early energy density than in the ΛCDM case (which has $\Omega_\Lambda\approx 10^{-9}$ at $a=10^{-3}$; allowing the limiting non-negative energy density $\delta\approx -10^{-9}$ would have negligible impact on the distributions). Note that these binned deviations do not have appreciable covariances with each other, with the correlation coefficients under 0.26 except for δ_2–δ_3 at 0.49. This is a useful feature adding near independence to localization, making their interpretation transparent, and is a result of the careful choice of bins based on the PCA of (3).

Table 3.1 gives the 95% confidence upper limits on each expansion deviation parameter, showing that recent CMB data provides 2–16% constraints on the expansion history back to $z=10^5$. The earliest bin, of δ_1, is reasonably constrained despite being well before recombination, and should improve further when adding high resolution (high multipole l) measurements. The second bin has equivalent constraints when taking into account its narrowness. Around recombination, however, δ_3 and δ_4 have looser bounds because all the standard cosmological parameters also enter strongly at this epoch, and so the increased covariance dilutes

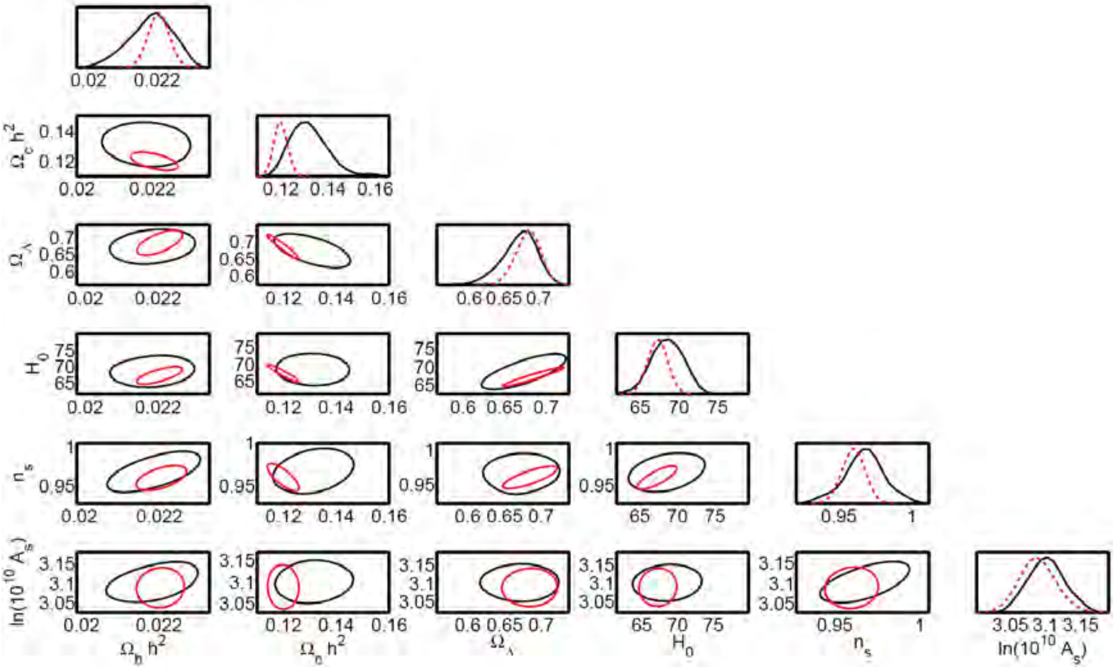


Figure 3.1 Joint 68% confidence contours on the standard cosmological parameters are shown when allowing for expansion history deviations from ΛCDM (black), and fixing to ΛCDM (smaller contours or dashed curves). Plots on the diagonal give the 1D marginalized probability distributions.

their estimation. They have the two highest correlation coefficients, of 0.89 between δ_3 and $\Omega_c h^2$ and -0.76 between δ_4 and $\Omega_b h^2$. Finally, the late, broad bin of δ_5 has strong constraints. These behaviors are all consistent with the pre-Planck, Fisher matrix predictions of (3) (see their Fig. 4). Adding late time data or priors (which we avoid; see concluding section) can shrink some uncertainties by up to 60%.

The expansion history does not completely define the system of Boltzmann equations: the effective dark component can have internal degrees of freedom such as sound speed c_s that determine the behavior of its perturbations and hence the gravitational clustering of the photons (10). Therefore we also show in Table 3.1 the constraints when this sound speed is equal to that of a relativistic species ($c_s^2 = 1/3$), or is much smaller than the speed of light, cold dark energy with $c_s = 0$. The $c_s = 0$ case has looser bounds, due to the additional influence on the photon clustering with the strengthened gravitational potentials, and covariance with matter parameters during matter domination. For the $c_s^2 = 1/3$ case, where the extra expansion rate corresponds to extra relativistic degrees of freedom, the constraints are weaker during radiation domination. This is a combination of the expansion deviation acting just like the photons, and a slight preference of the data for additional radiation energy density, in accord with previous hints that the number of effective neutrino species, $N_{\rm eff}$, might be greater than the standard model value of 3.046. Indeed, the mean value of $\delta_2 = 0.026$ in this case corresponds to $\Delta N_{\rm eff} = 0.31$, in good agreement with the Planck values of $N_{\rm eff} = 3.39$. Recall that $\Delta N_{\rm eff}$ denotes the equivalent number of relativistic neutrino species corresponding to the extra energy density. Since δ_2 is not in the fully radiation dominated era, we must translate it to the constant early

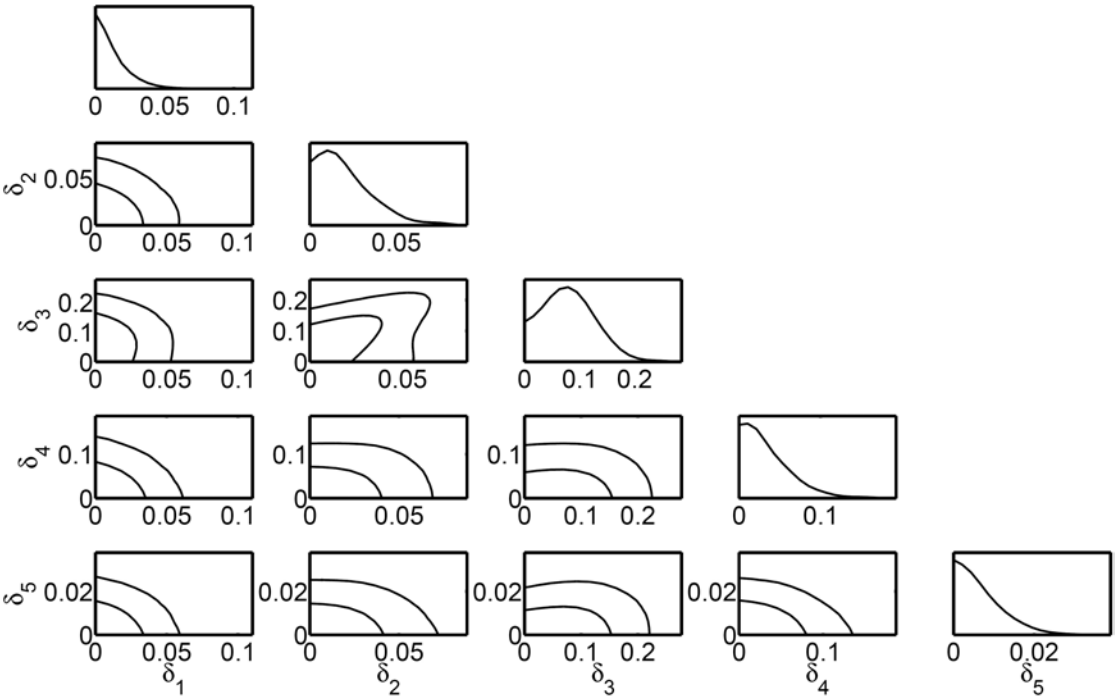


Figure 3.2 Joint 68% and 95% confidence contours on the expansion deviation parameters are shown. Plots on the diagonal give the 1D marginalized probability distributions.

dark energy density using Eq. (25) of (3) and then to the asymptotic relativistic $\Delta N_{\rm eff}$ using Eq. (6) of (11).

In all other parts of the article we keep $c_s = 1$.

Case	$\delta_1(10^{-4.5})$	$\delta_2(10^{-3.8})$	$\delta_3(10^{-3.4})$	$\delta_4(10^{-3.0})$	$\delta_5(10^{-1.4})$
$c_s^2 = 1$	0.036	0.050	0.160	0.095	0.018
$c_s^2 = 1/3$	0.053	0.054	0.067	0.038	0.013
$c_s^2 = 0$	0.060	0.069	0.109	0.184	0.223

Table 3.1 95% confidence upper bounds are given for the expansion history deviations δ, listed by the bin number and midpoint of the $\log a$ bins, for cases with different sound speeds.

Figure 3.3 shows the mean value and 68% uncertainty band of the expansion deviations $\delta(a)$ given by the Monte Carlo reconstruction using the recent CMB data. This figure represents the best current model-independent knowledge of the early expansion history of our Universe. Setting all $\delta_i = 0$, i.e. ΛCDM, is consistent with these results at the 95% confidence level. The mean value does show a very slight preference for a faster expansion rate, as in early dark energy or extra relativistic degrees of freedom, before recombination.

Physical Implications – This analysis has been model independent, allowing individual epochs to float freely without assuming a functional form. If we do assume a specific model, then constraints will in general be tighter, with each epoch having leverage on others through the restricted form.

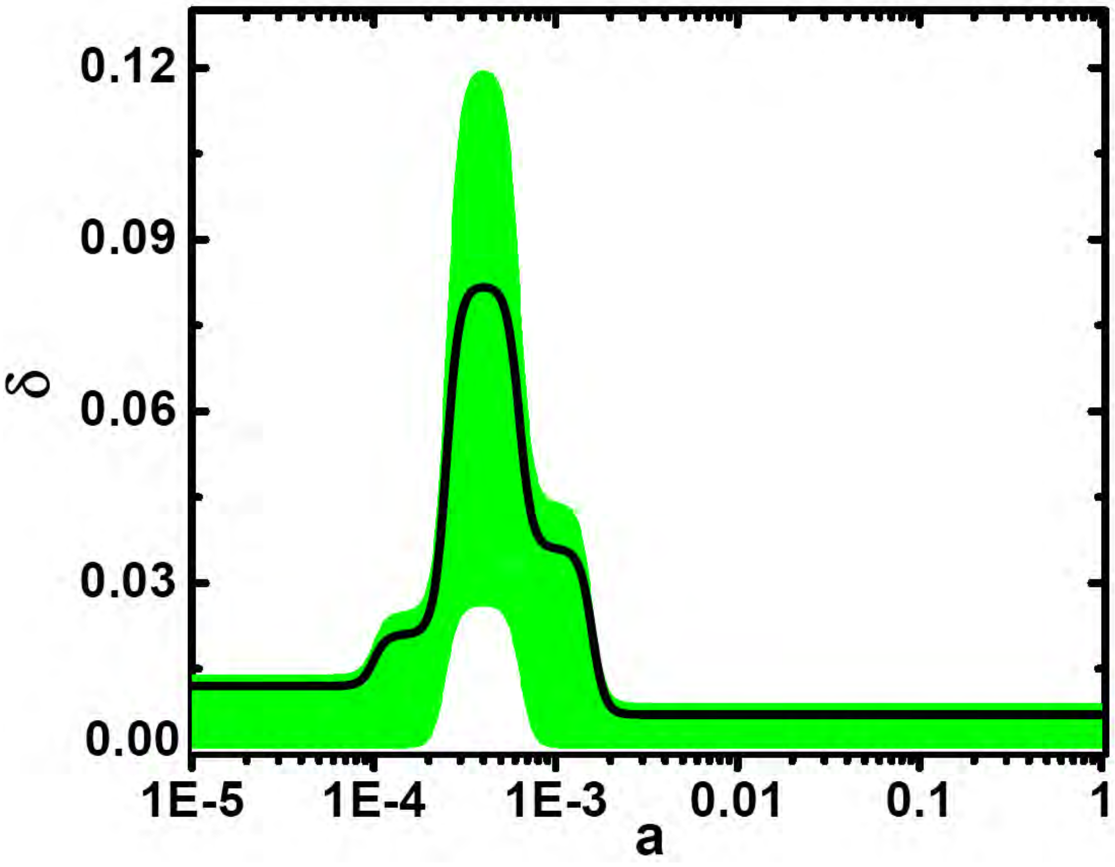


Figure 3.3 Reconstruction of the expansion history deviations $\delta(a)$ from ΛCDM is shown, with the mean value (solid line) and 68% uncertainty band (shaded area).

Three distinct families of early dark energy might be considered: where the early dark energy density rises, falls, or stays constant across CMB recombination. These were investigated in (3) in terms of the (somewhat motivated) models of barotropic aether, dark radiation, and Doran-Robbers (12) forms, respectively (see (3) for more detailed discussion). We compute the constraints on the fraction Ω_e of critical density contributed by early dark energy (approximately equivalent to δ) within each of these models (not using the δ_i bins), giving the results in Table 3.2. (Note that Planck finds $\Omega_e < 0.009$ at 95% CL for the Doran-Robbers model when also including high multipole data (13).)

	Aether	Dark Radiation	Doran-Robbers
Ω_e	0.019	0.033	0.012

Table 3.2 The 95% confidence level uncertainties are presented for three early dark energy models. For small values, $\Omega_e \approx \delta$. The Doran-Robbers model has an additional parameter w_0; we find $w_0 = -1.49^{+0.69}_{-0.57}$ (95% CL).

Two aspects of the models impact their detectability: the presence of the expansion history deviation at a sensitive epoch and its persistence over time, and its clustering behavior. The common Doran-Robbers form has the tightest bounds (despite the extra parameter), due to its persistence pre- and post-recombination and its distinction from matter clustering since it has $c_s^2 = 1$. The aether model only begins to deviate around recombination, and has $c_s^2 = 0$ so there is more covariance with the dark matter component. Dark radiation has influence only before recombination and its $c_s^2 = 1/3$ makes it more covariant with the photons (and neutrinos). A key conclusion is that early dark energy could in fact be more prevalent than apparent from bounds in the literature on the Doran-Robbers model.

Since dark radiation density at early times scales like radiation, it acts like the addition of relativistic degrees of freedom. Taking into account the definition of extra degrees in terms of the number of effective neutrino species $N_{\rm eff}$, the constraint on Ω_e within the dark radiation model translates to (11)

$$\Delta N_{\rm eff}(a \ll a_{\rm eq}) = 7.44\,\Omega_e/(1-\Omega_e)\ . \tag{3.4}$$

Thus $\Omega_e < 0.033$ for the dark radiation model becomes $\Delta N_{\rm eff} < 0.25$ at 95% CL. This puts a tighter global bound on $\Delta N_{\rm eff}$ compared to our model independent value from δ_2 before recombination ($\Delta N_{\rm eff} < 0.71$ at 95% CL, where again we have to account for δ_2 not being in the fully radiation dominated era).

Another implication of the expansion history is its relation to the spacetime itself. The Ricci scalar curvature is the central quantity in the Einstein-Hilbert action for general relativity, and plays a key role as well in extensions to gravity such $f(R)$ theories. The curvature history of the Universe has been explored from a theoretical perspective recently by (14). Since

$$\begin{aligned} R &= 3H^2\left[1 - 3w_{bg}\frac{H_{\rm fid}^2}{H^2} - 3w\frac{\delta H^2}{H^2}\right] && (3.5)\\ &= 3H_{\rm fid}^2\left[1 - 3w_{bg} + \delta(1-3w)\right]\ , && (3.6)\end{aligned}$$

observational constraints on δ (and hence w through Eq. 3.3) can be used to cast light on the curvature history.

Conclusions – We have used the recent advances in CMB data to constrain the fundamental quantity of the expansion history of our Universe. The results from the model independent analysis bound deviations from ΛCDM at 2–16% (95% CL), depending on the epoch. This constrains any deviations, whether due to, e.g., some form of dark energy or a nonstandard number of relativistic degrees of freedom. It also relates directly to the Ricci spacetime curvature.

Adding late time data that helps to constrain H_0 or Ω_m, say, would help break the degeneracy around recombination that led to the loosest, 16% upper bound on deviations. However, proper treatment of this would require many low redshift bins to reflect the density of the data, while our focus here is on the early expansion history.

We regard the model independence of the analysis as a signal virtue; however we can also compare the bounds for specific early dark energy models. For the barotropic aether, dark radiation, and Doran-Robbers models we derive 95% CL limits of less than 0.019, 0.033, 0.012 in early dark energy density Ω_e, respectively. We emphasize that bounds appear tightest when assuming the conventional Doran-Robbers form, and so early dark energy should be not be thought ruled out based purely on constraining this model. In terms of extra effective neutrino species the model independent results imply $\Delta N_{\rm eff} < 0.71$ at 95% CL.

Future CMB data, such as the release of polarization data from Planck, ACTpol (15), PolarBear (16), SPTpol (17) experiments will enhance our knowledge of the history back to $z \approx 10^5$. Exploring the even earlier universe will require neutrino, dark matter, or gravitational wave astronomy. Late time probes will continue to map the last e-fold of cosmic expansion in greater detail. Over just a few years cosmological observations have taken us from order unity uncertainty (with $\gtrsim 10\%$ in narrow epochs around recombination and today) to a few percent level knowledge over more than 10 e-folds of cosmic history.

Acknowledgments – We thank Stephen Appleby, Scott Daniel, and Tristan Smith for helpful discussions. AH acknowledges the Berkeley Center for Cosmological Physics for hospitality. This work has been supported by WCU Korea grant R32-2009-000-10130-0 and the Director, Office of Science, Office of High Energy Physics, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The Dark Cosmology Centre is funded by the Danish National Research Foundation.

BIBLIOGRAPHY

[1] Planck Collaboration XV, arXiv:1303.5075

[2] C.L. Bennett et al, arXiv:1212.5225

[3] J. Samsing, E.V. Linder, T.L. Smith, Phys. Rev. D 86, 123504 (2012) [arXiv:1208.4845]

[4] E.V. Linder & T.L. Smith, JCAP 1104, 001 (2011) [arXiv:1009.3500]

[5] E.V. Linder, Phys. Rev. D 82, 063514 (2010) [arXiv:1006.4632]

[13] A. Lewis, A. Challinor, & A. Lasenby, ApJ 538, 473 (2000) [arXiv:astro-ph/9911177]; `http://camb.info`

[7] A. Lewis & S. Bridle, Phys. Rev. D 66, 103511 (2002) [arXiv:astro-ph/0205436] `http://cosmologist.info/cosmomc`

[8] S. Das et al, arXiv:1301.1037

[9] K.T. Story et al, arXiv:1210.7231; `http://pole.uchicago.edu/public/data/story12/index.html`

[10] W. Hu, ApJ 506, 485 (1998) [arXiv:astro-ph/9801234]

[11] E. Calabrese, D. Huterer, E.V. Linder, A. Melchiorri, L. Pagano, Phys. Rev. D 83, 123504 (2011) [arXiv:1103.4132]

[12] M. Doran and G. Robbers, JCAP 0606, 026 (2006) [arXiv:astro-ph/0601544]

[13] Planck Collaboration XVI, arXiv:1303.5076

[14] R.R. Caldwell, S.S. Gubser, Phys. Rev. D 87, 063523 (2013) [arXiv:1302.1201]

[15] M.D. Niemack et al., SPIE 7741, 77411S (2010) [arXiv:1006.5049]

[16] Z. Kermish et al., SPIE 8452, 84521C (2012) [arXiv:1210.7768]

[17] J.E. Austermann et al., SPIE 8452, 84520E (2012) [arXiv:1210.4970]

4

DARK ENERGY EVOLUTION AND MODIFIED GRAVITY

Power Spectrum Precision for Redshift Space Distortions

Eric V. Linder[1,2] **& Johan Samsing**[1,3]

[1] Berkeley Center for Cosmological Physics & Berkeley Lab, University of California, Berkeley, CA 94720, USA
[2] Institute for the Early Universe WCU, Ewha Womans University, Seoul, Korea
[3] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark

Abstract

Redshift space distortions in galaxy clustering offer a promising technique for probing the growth rate of structure and testing dark energy properties and gravity. We consider the issue of to what accuracy they need to be modeled in order not to unduly bias cosmological conclusions. Fitting for nonlinear and redshift space corrections to the linear theory real space density power spectrum in bins in wavemode, we analyze both the effect of marginalizing over these corrections and of the bias due to not correcting them fully. While naively subpercent accuracy is required to avoid bias in the fixed case, in the fitting approach the Kwan-Lewis-Linder reconstruction function for redshift space distortions is found to be accurately selfcalibrated with little degradation in dark energy and gravity parameter estimation for a next generation galaxy redshift survey such as BigBOSS.

4.1 SUMMARY

The growth of structures is a promising way of testing gravity and the evolution of the energy density of DE. The basic idea is to observationally estimate how the force from the matter field itself and the friction force term from the expanding background each impact the dynamical evolution of structures. Both modified gravity and evolving DE models can change these terms which directly impacts the growth amplitude δ and the growth rate f_G of matter at different redshifts. However, the parameters δ and f_G are not easily measured because we observe angles and redshifts rather than real positions and velocities. To extract full cosmological information from observations we therefore need to model how we go from the observed redshift space to position and velocity space. Current analytical models fail at this reconstruction on small scales where the density field stops being linear, $k \gtrsim 0.1\,h/\mathrm{Mpc}$, which greatly limit us in extracting the full information.

To address this reconstruction problem, I did a study with a collaborator on how well we need to do this modeling in order not to unduly bias cosmological conclusions. For this I wrote a Fisher code to study uncertainties and biases in redshift space reconstructions for a galaxy redshift survey experiment similar to BigBOSS. We included modified gravity and DE using the parametrizations $w(a) = w_0 + w_a(1-a)$ and $f_G = \Omega_m(a)^\gamma$ for the DE equation of state and growth rate, respectively. Our approach was to first model the full non-linear redshift space matter power spectrum using the linear Kaiser formula corrected at non-linear scales using the numerically estimated KLL reconstruction function $F(k, \mu, z)$. We then allowed F to vary, dividing it into bins in wavenumber k to study how the bin values, or reconstruction parameters, impact the cosmological fit including gravity and DE.

First we studied the bias due to systematic errors in these reconstruction parameters. We found that the cosmological fit depends mostly on parameters associated with the non-linear part of the real space matter power spectrum and secondarily on the peculiar velocity effects beyond the KLL reconstruction. This serves as a guidance to where future models should improve. We then investigated the possibilities for including the reconstruction parameters in the fit to remove biases. Surprisingly, we found the parameters to be highly 'self-calibrating' leading to only weak degradations in the cosmological parameters.

4.2 INTRODUCTION

The pattern of galaxy clustering in three dimensions, and its evolution, encodes abundant information on the cosmological parameters affecting matter growth. Ongoing and next generation spectroscopic galaxy surveys will vastly increase our measurements of this clustering, and our knowledge of cosmology if we can accurately interpret the results in terms of theory. Measurements accrue an extra contribution to the redshift, and hence apparent position along the sight, from the galaxy peculiar velocities induced by the inhomogeneous density field; this gives rise to an anisotropy in the observed clustering known as redshift space distortions (RSD).

These distortions carry information on the growth rate, as opposed to just the growth amplitude, and so are valuable for probing cosmology, as well as the gravitational strength driving the growth. However, linear theory is insufficient for accurate relation of the redshift space galaxy power spectrum to the true (real space) matter density power spectrum, even on quite large scales, or wavenumbers $k > 0.05\,h/\mathrm{Mpc}$, where the vast majority of the statistical leverage lies (1; 2; 3; 4; 5; 6). Numerous corrections involving higher order perturbation theory have been employed (7; 8; 9; 10) that extend the validity but the region $k > 0.1\,h/\mathrm{Mpc}$ is still problematic, especially for quantities involving the growth rate and the gravitational growth characterization. For example, (6) demonstrates that these first principles approaches deliver results for the growth rate that are biased by several standard deviations when using modes out to $k = 0.2\,h/\mathrm{Mpc}$.

Here we investigate a basic question: how accurately does one actually need to know the redshift space distortion mapping in order to extract the cosmological and gravitational parameter information without substantial bias or degradation? Similar questions have been considered for weak gravitational lensing, for example, where one asks how well the nonlinear matter power spectrum needs to be known to estimate cosmology from the lensing shear power spectrum (11; 12).

In Section 4.3 we introduce the correction, or reconstruction, function for the redshift space power spectrum and review the KLL (6) form for it. Section 4.4 uses the Fisher bias method to compute both the individual parameter bias and joint confidence contour bias due to misestimated RSD, thus giving criteria for the accuracy to which the RSD effects must be known. Adding fit parameters for uncertainties in the reconstruction function in Sec. 4.5, we assess the impact of marginalizing over them on the cosmological parameters, in particular for tests of dark energy and gravity. Section 4.6 summarizes the results and conclusions.

4.3 GALAXY POWER SPECTRUM

4.3.1 REDSHIFT SPACE POWER SPECTRUM

In real space the matter density power spectrum is expected to be isotropic, and the linear power spectrum grows in a scale independent manner through the growth factor $D(z)$, where z is the redshift. The observed, redshift space galaxy power spectrum involves a transformation to redshift space due to the velocity effects, and a bias relation $b(z)$, usually taken to be scale independent, converting the dark matter overdensity to galaxy overdensity, and the effects of nonlinear structure formation. Each of these is modeled in various ways, with attendant uncertainties.

We write the anisotropic redshift space galaxy power spectrum as

$$P(k,\mu,z) = P^r(k,z)\,M(k,\mu,z)\,F(k,\mu,z) \tag{4.1}$$

where P^r is the isotropic real space matter power spectrum, M is an approximate model for redshift space distortions (including galaxy bias), and F is the reconstruction function accounting for nonlinearities and more exact velocity effects.

The linear mass power spectrum P^r is given by a Boltzmann numerical code such as CAMB (13). It depends on the cosmological parameters through its shape as a function of wavenumber k and through the growth factor $D(z)$ giving its amplitude evolution. Since we will correct the RSD modeling by the reconstruction function, we choose M to be simply given by the linear theory prediction, the Kaiser approximation (14),

$$M(k, \mu, z) = [b(z) + f(z)\,\mu^2]^2\ , \tag{4.2}$$

where b is the galaxy bias, $f = d\ln D/d\ln a$ the growth rate of density perturbations that in the linear regime grow as $\delta \sim D(a)$, where the scale factor $a = 1/(1+z)$, and μ is the cosine of the angle made by the perturbation wavevector $\vec{k}$ with respect to the line of sight. Beyond the linear regime, b could be scale dependent, i.e. $b(k)$, but we will absorb that possibility into the reconstruction function. The reconstruction function is fitted to N-body simulations by the analytic form of Kwan, Lewis, & Linder (KLL; (6)),

$$F(k, \mu, z) = \frac{A(k, z)}{1 + B(k, z)k^2\mu^2} + C(k, z)k^2\mu^2\ . \tag{4.3}$$

This form has been found to reproduce accurately results of N-body simulations over a wide range of redshifts, and for halos of various masses as well as dark matter; see (6; 15) for details. Note that A, B, C may be cosmology dependent, just as f and P^r are, and their universality is a subject of ongoing research, but here we treat them as independent parameters (as an analogy, recall how people treat coefficients within Halofit also as universal, though here we let the values of A, B, C float; also see Sec. 4.5.4). The factor A characterizes nonlinearity of the real space power spectrum, B describes velocity effects such as damping from a Lorentzian velocity dispersion but also higher order multipole terms, while C describes nonlinear enhancement for large $k\mu$ and possibly breaks the degeneracy in the two roles of B.

4.3.2 GALAXY CLUSTERING INFORMATION

The cosmological information inherent in the galaxy power spectrum can be estimated through the Fisher information matrix. The full set of parameters $\{p_i\}$ includes the cosmological parameters, astrophysical parameters such as galaxy bias, and parameters for the reconstruction function. Sensitivity to cosmology enters through the derivatives $\partial P/\partial p_i$ and the error covariance matrix for the redshift space galaxy power spectrum P.

We follow the usual prescription (16; 17; 18) where the covariance matrix comes from sample variance (finite volume) and shot noise (finite resolution of the density field by sparse galaxies). Taken together, the error can be thought of as depending on the number of modes that the galaxy redshift survey samples. Treated as Poisson sampling of the density field, the statistical error is

$$\sigma_P = P + n^{-1} \tag{4.4}$$

from these two effects. The number of Fourier modes is the volume of a Fourier cell times the number of cells,

$$N_{\rm modes} = 2\pi k^2 dk\, d\mu \times V_{\rm survey}/(2\pi)^3\ . \tag{4.5}$$

Therefore the error covariance matrix C is

$$C = P^2 \left(\frac{1+nP}{nP}\right)^2 \frac{8\pi^2}{V_{\rm survey}k^2 dk d\mu}\ . \tag{4.6}$$

Since the Fisher information matrix is constructed from C^{-1} multiplied by the sensitivity derivatives $\partial P/\partial p_i$, we can use the P^2 factor to convert the derivatives to involve $\ln P$, which will be useful in treating the multiplicative factors in Eq. (4.1). In summary, the Fisher matrix is

$$F_{ij} = \sum_z \sum_\mu \sum_k \frac{\partial \ln P}{\partial p_i} \frac{\partial \ln P}{\partial p_j} V_{\rm eff}(k,\mu,z) \frac{k^2 \Delta k\, \Delta\mu}{8\pi^2} \,, \tag{4.7}$$

where the survey volume is reduced by the shot noise to a z, k, and μ dependent effective volume

$$V_{\rm eff}(k,\mu,z) = V_{\rm survey}(z) \left[\frac{n(z)P(k,\mu,z)}{n(z)P(k,\mu,z)+1}\right]^2 \,. \tag{4.8}$$

When the galaxies densely sample the underlying field, the effective volume approaches the survey volume, otherwise modes are lost, diluting the effective volume due to increased noise.

Note that the logarithmic derivatives can be written as

$$\begin{aligned} \frac{\partial \ln P}{\partial p_i} \frac{\partial \ln P}{\partial p_j} &= \frac{(\partial \ln P^r + \partial \ln M + \partial \ln F)}{\partial p_i} \\ &\quad \times \frac{(\partial \ln P^r + \partial \ln M + \partial \ln F)}{\partial p_j} \end{aligned} \tag{4.9}$$

so only the $\partial \ln F$ term depends on A, B, and C.

The reconstruction function derivatives are

$$\frac{\partial F}{\partial A} = \frac{1}{1+Bk^2\mu^2} \tag{4.10}$$

$$\frac{\partial F}{\partial B} = \frac{-Ak^2\mu^2}{(1+Bk^2\mu^2)^2} \tag{4.11}$$

$$\frac{\partial F}{\partial C} = k^2\mu^2 \,, \tag{4.12}$$

and are otherwise taken not to depend on cosmology. This is because we use A, B, C purely as fiducial values, and investigate how their variation (from astrophysics or cosmology) impacts the cosmological parameter estimation.

Our fiducial case attempts to match F to the simulation results in (6), with estimated

$$A(k) = 1 + \left(\frac{k}{0.39\, h/\mathrm{Mpc}}\right)^{1.58} \tag{4.13}$$

$$B(k) = 20\, (\mathrm{Mpc}/h)^2 \tag{4.14}$$

$$C(k) = 8\, e^{-k/(0.176 h/\mathrm{Mpc})}\, (\mathrm{Mpc}/h)^2 \,. \tag{4.15}$$

The resulting redshift space distortion reconstruction function of Eq. (4.3) is shown in Fig. 4.1. We emphasize that these are merely the fiducials; we allow the values of A, B, C to float freely in bins of wavenumber. This provides a model independent variation of the power spectrum (within the reconstruction form) and we can then investigate the influence of such variations on the cosmological parameter estimation. Conversely, the question can be phrased as "what is the accuracy required on knowledge of the galaxy power spectrum in order to deduce the cosmology with confidence?"

We later contrast this fiducial with fiducial $(A,B,C) = (1,0,0)$, i.e. assuming that perturbation theory (for example linear theory in the Kaiser case, although F also works with higher order perturbation theory (6)) fully captures RSD effects in the model M.

Figure 4.1 The redshift space distortion reconstruction function $F(k, \mu)$ is plotted for the fiducial expressions for A, B, C for three values of angle μ.

The analysis is carried out in the next sections in two ways: in Sec. 4.4 we compute the effect that a given level of unrecognized power spectrum deviation in some k bin, i.e. a systematic error in modeling, has in biasing the cosmological conclusions, and in Sec. 4.5 we recognize the existence of systematic uncertainties and treat them by marginalizing over the A, B, C values for each k bin.

4.3.3 SURVEY CHARACTERISTICS AND PARAMETERS

For the galaxy redshift survey data we consider a next generation spectroscopic survey of the quality proposed for BigBOSS (19), covering 14000 deg^2 from $z = 0.1 - 1.8$, with a galaxy number density n of approximately $3 \times 10^{-4}\, h^3\,\mathrm{Mpc}^{-3}$. For the exact distribution adopted see Table 4.1. There are actually two populations of galaxies: luminous red galaxies (LRG) and emission line galaxies (ELG), each with their own galaxy bias value. These galaxy biases are taken as free parameters to be marginalized over, for each redshift bin of width 0.1. Their fiducials are $b(z) = b_0 D(z = 0)/D(z)$ with $b_0^{\mathrm{ELG}} = 0.8$ and $b_0^{\mathrm{LRG}} = 1.6$, which provide good fits to observations. Galaxy populations with different biases can help reduce sample variance (20), with the Fisher matrix involving a sum over populations, i.e.

$$\sum_{XY} \frac{\partial \ln P_X}{\partial p_i} \frac{\partial \ln P_Y}{\partial p_j} \left[\frac{n_X P_X}{n_X P_X + 1}\right] \left[\frac{n_Y P_Y}{n_Y P_Y + 1}\right] . \tag{4.16}$$

Note that for multiple populations the factor V_{eff} in Eq. (4.8) involves the shot noise, i.e. nP, of each population.

For the cosmological parameters we use the physical baryon density $\Omega_b h^2$ and physical cold dark matter density $\Omega_c h^2$, reduced Hubble constant h, scalar perturbation tilt n_s and amplitude A_s, dark energy equation of state parameters w_0 and w_a, and gravitational growth index γ. The gravitational growth index gives an accurate description of the growth rate for both general relativity and a range of modified gravity models, and looking for deviations from its general relativistic value of $\gamma = 0.55$ acts as a test of gravity (21; 22). The growth index is treated as an independent parameter (not a function of w_0, w_a) and determines the growth factor at scale factor $a = 1/(1 + z)$,

$$D(a) = e^{\int_0^a (da'/a')\, \Omega_m(a')^\gamma} , \tag{4.17}$$

that in this ansatz is used to convert the linear power spectrum delivered by CAMB at $z = 0$ to another redshift z, to account for the effects of modified gravitational growth. Note that the growth rate $f = \Omega_m(a)^\gamma$, and redshift space distortions were highlighted as a test of gravity in (23).

For the central question of RSD uncertainties we employ up to 12 free parameters, taking A, B, C with independent values in each bin of width 0.1 in wavenumber above $k = 0.1\, h/\mathrm{Mpc}$ out to some $k_{\max}$. This corresponds to uncertainties ΔP_k. For the current work we follow (11; 12) and consider the uncertainties only as a function of wavenumber, not redshift, except in Sec. 4.5.3; we also take the KLL form to be accurate while allowing freedom in the parameters A, B, C. In summary we fit for 8 cosmological parameters and up to 39 systematics parameters.

z	$n_{\rm ELG}$	$n_{\rm LRG}$
0.15	22.6	30.1
0.25	8.45	3.04
0.35	4.02	3.07
0.45	2.65	3.09
0.55	2.99	3.10
0.65	3.99	3.11
0.75	5.15	3.12
0.85	5.36	1.89
0.95	5.02	0.33
1.05	4.80	0.04
1.15	4.49	–
1.25	4.04	–
1.35	3.02	–
1.45	2.00	–
1.55	1.15	–
1.65	0.43	–
1.75	0.12	–

Table 4.1 Spectroscopic survey number densities adopted for emission line galaxies and luminous red galaxies, in units of $10^{-4}\,h^3/{\rm Mpc}^3$, for each redshift shell.

4.4 FISHER BIAS

The first question we are interested in answering is what is the sensitivity of the cosmological parameter determination to errors in modeling RSD. One might have M or F wrong, but if this does not mimic a change in cosmology then no harm is done. The Fisher bias formalism (see, e.g., (24; 25)) propagates misestimation of the observable or theoretical prediction, in this case the redshift space power spectrum, into biases on the fit parameters. Specifically, we consider the effect of errors in the k bin values of A, B, C.

The Fisher bias on a parameter p_i from misestimating parameter p_a is

$$\delta p_i = \delta p_a \sum_j (F^{\rm sub})^{-1}{}_{ij}\,(F^{\rm full})_{ja}\ , \tag{4.18}$$

where $\delta p_a = p_a({\rm true}) - p_a({\rm fiducial})$. The superscript "sub" denotes the Fisher submatrix without entries for the parameters whose misestimation we are studying. (For the specific case here, the submatrix will be 35×35 for the cosmology and galaxy bias parameters, and the full matrix adds the reconstruction parameters one at a time. We later consider all the reconstruction parameters at once.) By evaluating the ratio dp_i/dp_a for $a = A, B, C$ we can assess the sensitivity of the parameter estimation to the RSD modeling. To take a weak lensing example, (26) found that a particular form of matter power spectrum distortion with amplitude A_{NL} at high k distorted estimation of w_a derived from shear power spectrum measurement by a leverage factor of 18: a misestimation of 10% in A_{NL} yielded a 1.8σ bias in w_a.

The bias δp_i can be compared to the statistical uncertainty $\sigma(p_i)$ on the parameter, either directly or through the risk statistic

$$R_i \equiv \sqrt{\sigma^2(p_i) + \delta p_i^2}\ . \tag{4.19}$$

Treating the bias as a systematic error in this way, to restrict the degradation in the statistical error to less than 20%, say, requires $\delta p_i/\sigma(p_i) < 0.66$, thus putting a constraint on the allowable modeling error δp_a etc. We examine two, converse statistics: the cosmological degradation caused by a certain fractional misestimation of the reconstruction parameters $\delta p_a/p_a$, and the requirement on the reconstruction parameter to bound the cosmological parameter bias to less than a given factor of the statistical uncertainty, $\delta p_i/\sigma_i$. These are respectively

$$\frac{R_i}{\sigma_i} = \sqrt{1+\left(\frac{\delta p_i}{\delta p_a}\frac{\delta p_a}{p_a}\frac{p_a}{\sigma_i}\right)^2} \tag{4.20}$$

$$\frac{\delta p_a}{p_a} = \left(\frac{\delta p_i}{\delta p_a}\right)^{-1}\frac{\delta p_i}{\sigma_i}\frac{\sigma_i}{p_a}\;. \tag{4.21}$$

Figure 4.2, left panel, shows the matrix of degradations R_i/σ_i for fixed $\delta p_a/p_a = 0.01$ (i.e. 1% uncertainty on the reconstruction parameters), where the columns are the dark cosmological parameters and the rows are the reconstruction parameters. The right panel gives a similar matrix of the reconstruction requirements $\delta p_a/p_a$ for fixed $\delta p_i/\sigma_i = 1$ (which corresponds to $R_i/\sigma_i = 1.41$). One can scale the results for different fixed values according to the above equations. The stripe structure arises because the bias from $A_{0.45}$, where the subscript indicates the center of the k bin, is of opposite sign from that of $A_{0.25}$, and $A_{0.35}$ lies in between near null effect, and similar for B and C.

The degradations in determination of the dark parameters w_0, w_a, γ are less than 22% for 1% shifts in the reconstruction parameters in all cases except $A_{0.25}$, $A_{0.45}$, and $B_{0.45}$. For $A_{0.25}$ and $A_{0.45}$ the risk error on w_a can exceed the statistical uncertainty by a factor 3. For the B parameters the worst case is degradation by 1.5. These results offer indications of what physics must be most accurately understood, i.e. the nonlinearity from A and, somewhat less critically, the velocity effects from B.

In the converse analysis of what accuracy is required on the reconstruction parameters to ensure that a bias is restricted to below 1σ, we find that 5% accuracy is sufficient for all parameters except for the A_k, plus $B_{0.25}$ and $B_{0.45}$. Knowledge of $A_{0.25}$ and $A_{0.45}$ are needed to 0.3%, $B_{0.45}$ to 0.9%, $B_{0.25}$ to 1.4%, $A_{0.35}$ to 1.5%, and $A_{0.15}$ to 3.1%.

While these approaches give indications of sensitivity, they treat the cosmological parameters one by one while a power spectrum misestimation will generally impact several at once. This can either tighten or loosen overall requirements, depending on the covariances. To take this into account we use the $\Delta\chi^2$ method (27; 28). This describes the fuller impact of biasing cosmology through quantifying how far from the fiducial the best fit cosmology is shifted relative to the confidence contour, taking into account degeneracies between the reconstruction and cosmological parameters. This measure is given by

$$\Delta\chi^2 = \sum_{ij} \delta p_i\, F^{(\rm red)}_{ij}\, \delta p_j\;, \tag{4.22}$$

where the sum runs only over the reduced parameter set whose bias we are interested in, e.g. w_0 and w_a for a 2D w_0–w_a joint likelihood contour plot. The reduced Fisher matrix $F^{(\rm red)}$ is marginalized over all other cosmological and galaxy bias parameters (the reconstruction parameters have already been taken into account in obtaining δw_0 etc.). The bias $\Delta\chi^2$ accounts for the property that biases in, say, the direction of the narrow axis of the Fisher ellipse are more detrimental than those along the degeneracy direction.

Figure 4.3 illustrates the 2D bias induced in the dark energy and growth parameters, here for a 1% misestimation in the reconstruction parameters one by one. Most such reconstruction

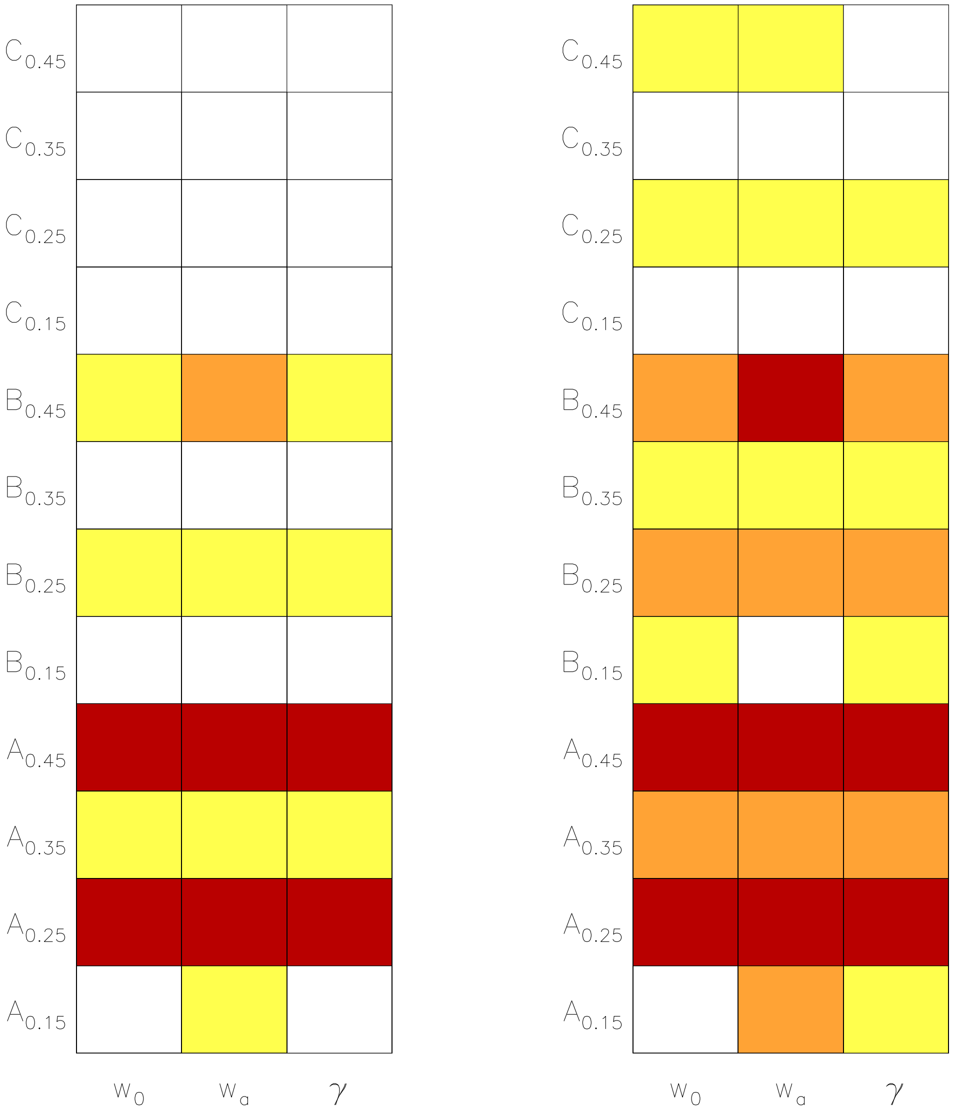


Figure 4.2 [Left panel] The ratio of the root mean squared error, or risk, to the statistical uncertainty, R_i/σ_i, is plotted for each dark cosmological parameter in the case of a 1% deviation in a reconstruction parameter. [Right panel] The fractional requirement on each reconstruction parameter $\delta p_a/p_a$ needed to ensure bias less than 1σ, i.e. $\delta p_i/\sigma_i < 1$ is plotted. Dark red indicates danger (high risk or tight requirement), with lighter colors showing reduced impact. For the left panel the color scale is $R_i/\sigma_i > 2$ (dark red), 1.4–2 (medium orange), 1.05–1.4 (light yellow), and 1–1.05 (white). The right panel has $|\delta p_a/p_a| < 0.01$ (dark red), 0.01–0.05 (medium orange), 0.05–0.2 (light yellow), > 0.2 (white). Here $k_{\rm max} = 0.5\,h/$Mpc.

parameter errors do not significantly affect the joint parameter likelihood. In the w_0–w_a plane, none of the C parameters and three of the B parameters do not bias the best fit outside the 1σ contour, and $B_{0.45}$ remains within the 2σ contour. Only errors on the $A_{0.25}$ and $A_{0.45}$ parameters are particularly damaging, causing a bias of up to $\Delta\chi^2 = 39$ (approximately at the 6σ level). The covariance between the shifts in w_0 and w_a is crucial; if the same bias in w_a and an even larger bias in w_0 occurred such that the joint shift lay along the degeneracy axis, then the 2D bias would be scarcely outside the 2σ uncertainty contour. For the w_a–γ plane, the biases are less severe, with only $A_{0.25}$ and $A_{0.45}$ causing more than a 2σ joint bias, at $\Delta\chi^2 \approx 16$.

Treating the reconstruction errors one by one effectively takes a localized bump in the reconstruction function. A smooth variation would instead affect several of the reconstruction parameters at once; this has a different effect on the cosmological parameter bias. As an example, we simultaneously vary all four A parameters by 1%. Since $A_{0.25}$ and $A_{0.45}$ have nearly opposite effects this reduces substantially the $\Delta\chi^2$ due to varying just one of them, e.g. from 39 to 7.7. The 2D bias due to such smooth variation is shown in the figures by the magenta squares.

Figure 4.3 The biases in the w_0–w_a and w_a–γ planes due to 1% misestimation in the 12 reconstruction parameters, one by one, are shown by x's (along with the $\Delta\chi^2$ if larger than 2.8). The contours give the joint 2D 1σ and 2σ confidence levels on the cosmological parameters when the reconstruction parameters take their fiducial ("true") values. Magenta squares show the biases when varying all bins of A simultaneously; such smooth variations are much less damaging, e.g. reducing the individual $\Delta\chi^2 = 39$ and 34 biases to a joint $\Delta\chi^2 = 7.7$ offset (or the 16 and 15 in the w_a–γ panel to 2.8).

While we have thus far been model independent in taking A, B, C to be independent from one k bin to the next, we can now consider the difference between two fiducial models for the overall reconstruction function F. This then includes the effects of variations at all k's simultaneously, and allows a study of bias as a function of $k_{\rm max}$. As we consider each successive bin at higher k, we increase the number of modes, reducing the statistical uncertainty, but also often increase the deviation in the power spectra, increasing the bias in the cosmological parameters if we assume the wrong fiducial as the truth. The truth is taken to be F as given by Eqs. (4.13)–(4.15) in Eq. (4.3), while the incorrect assumption is pure linear theory, i.e. simply the Kaiser form for redshift space distortions, equivalent to $A = 1$, $B = C = 0$.

This misestimation of the redshift space galaxy power spectrum causes a bias in cosmological parameters of

$$\begin{aligned}\delta p_i &= \left(F^{\rm sub}\right)^{-1}_{ij} \sum_z \sum_\mu \sum_k \frac{P(A,B,C) - P(1,0,0)}{P(A,B,C)} \\ &\quad \times \frac{\partial \ln P}{\partial p_j}\, V_{\rm eff}(k,\mu,z) \frac{k^2 \Delta k\, \Delta\mu}{8\pi^2}\ . \end{aligned} \tag{4.23}$$

The systematic biases tend to worsen with increasing $k_{\rm max}$, reaching 1.4 in w_0, -8 in w_a, and 0.2 in γ for $k_{\rm max} = 0.5\, h/$Mpc, and are much larger than the statistical uncertainties for all $k_{\rm max}$. This is not surprising since $F_{\rm Kaiser}$ can deviate by a factor 2 from the KLL form. Thus, neglecting the uncertainties in the reconstruction parameters is not a viable option: we must take them into account.

4.5 MARGINALIZATION AND SELFCALIBRATION

As an alternative to requiring the power spectrum to subpercent accuracy and computing the bias from misestimated reconstruction parameters, we can fit for those parameters and calculate the increased uncertainty in cosmological parameters due to marginalization over the extra inputs. The basic question is how well the model needs to be known for precision determination of cosmology with RSD. This is similar to what (11; 12) did for matter power spectrum uncertainties applied to weak lensing cosmology. They used fractional power spectrum uncertainties in wavenumber bins, assumed constant with redshift, and applied some level of priors.

4.5.1 GLOBAL FIT

Now our quantities A_k, B_k, C_k in each wavenumber bin become fit parameters. Again, we can study the effects as we extend $k_{\rm max}$, using more bins and hence more parameters. Including these parameters means that we will not be biased any more with respect to the fiducial, but the enlarged parameter space will lead to some level of degradation of the statistical uncertainties, relative to fixing the reconstruction parameters, at the same $k_{\rm max}$.

Table 4.2 shows the effect of extending the data to higher $k_{\rm max}$, while simultaneously allowing for the additional reconstruction parameters in each bin. Despite the additional degrees of freedom in the fit, the cosmological parameter estimation sharpens with increasing $k_{\rm max}$. As long as the *form* of the reconstruction function holds, we obtain an accurate and unbiased cosmology even allowing for fitting variation in the amplitudes of A, B, C in each k bin. This is an extremely promising initial result for use of the reconstruction.

	$\Omega_b h^2$	$\Omega_c h^2$	h	n_s	$10^9 A_s$	w_0	w_a	γ	Ω_m
Fiducial	0.0226	0.112	0.7	0.96	2.47	-0.99	0	0.55	0.275
$\sigma(k_{\rm max} = 0.1)$	0.00524	0.0189	0.0542	0.0524	0.538	0.599	2.23	0.177	0.0302
$\sigma(k_{\rm max} = 0.2)$	0.00284	0.0102	0.0284	0.0288	0.325	0.197	0.779	0.0519	0.0159
$\sigma(k_{\rm max} = 0.3)$	0.00219	0.00760	0.0219	0.0198	0.248	0.112	0.478	0.0272	0.0122
$\sigma(k_{\rm max} = 0.4)$	0.00148	0.00508	0.0150	0.0130	0.170	0.0824	0.347	0.0193	0.00834
$\sigma(k_{\rm max} = 0.5)$	0.00141	0.00478	0.0142	0.0119	0.158	0.0713	0.306	0.0163	0.00794

Table 4.2 1σ constraints from future galaxy power spectrum data on cosmological parameters, marginalized over galaxy bias and redshift space distortion reconstruction. Note Ω_m is a derived parameter; $k_{\rm max}$ is in units of $h/{\rm Mpc}$. Despite the addition of more fit parameters when increasing $k_{\rm max}$, the cosmological parameters can be better determined.

To better understand why the added fit parameters do not cause an overall degradation, we look at the correlation matrix of the 47 parameters in Fig. 4.4. The block of parameters 36–47, representing the reconstruction parameters, is not highly correlated with other parameters: correlation coefficients are below 0.58 (0.38 for parameters other than n_s). (Even within the block, only $B_{0.15}$ and $C_{0.15}$, other than between the A_k, have a correlation coefficient exceeding 0.8, reaching 0.90; this is expected since for a low k expansion both B and C contribute as μ^2.) This means that the change in power spectrum shape due to adjusting the amplitudes of these parameters in F is not degenerate with a change due to w_0 or other such parameters. That is, the influence of these parameters have different k and μ dependences than those of cosmological parameters and so we find they can be separately fit.

Moreover, the reconstruction parameters are selfcalibrated by the data to good precision.

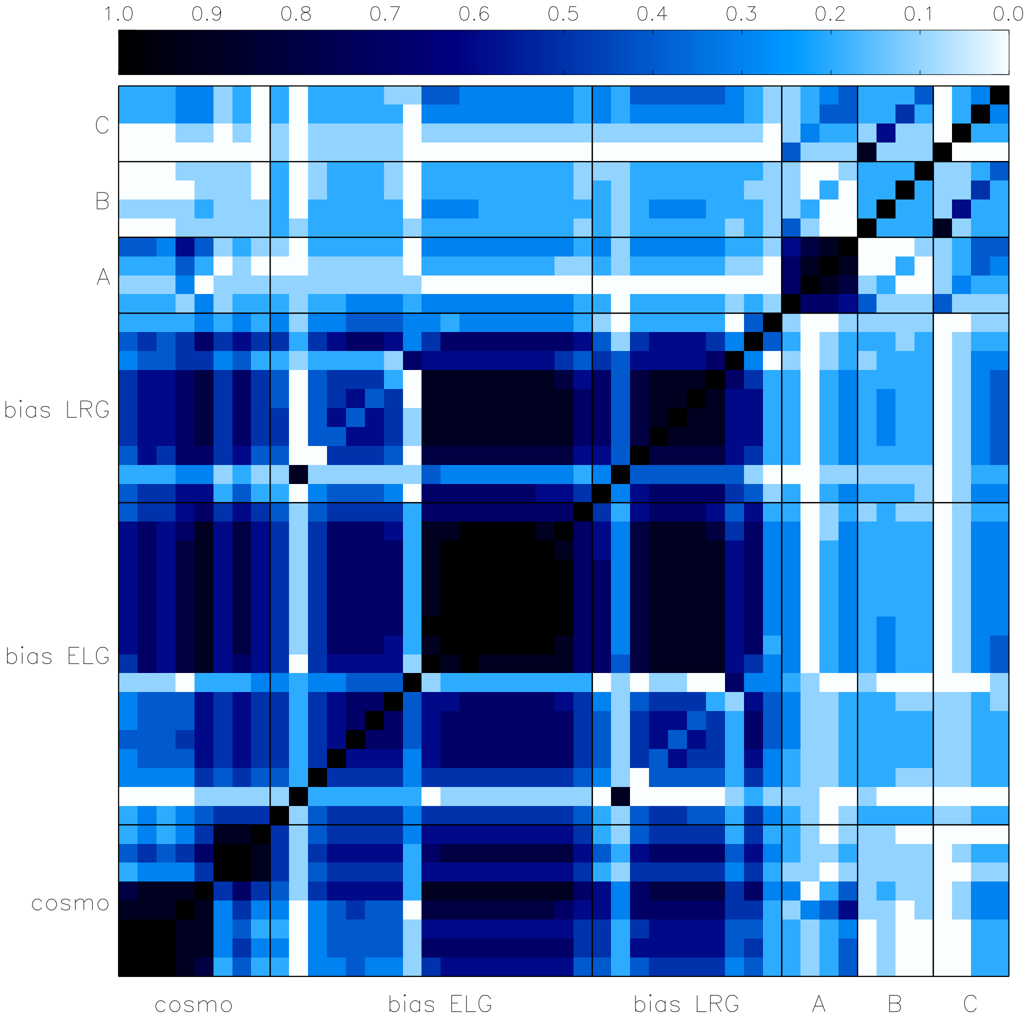


Figure 4.4 Correlation matrix of the 47 parameters for $k_{\rm max} = 0.5\,h/{\rm Mpc}$ is shown with color shading reflecting the absolute value of the correlation coefficient $r_{ij} = C_{ij}/\sqrt{C_{ii}C_{jj}}$. The correlation matrix is mostly block diagonal and the cosmological parameters are not strongly correlated with the reconstruction (or galaxy bias) parameters, so marginalization does not badly degrade cosmological parameter estimation.

All are determined to better than 3% (except $C_{0.15}$, to 8%) and most to subpercent level. These propagate into determination of the power spectrum to the subpercent level for variation of each one individually by 1σ, except for the extreme cases of $\mu = 1$ and $B_{0.15}$ ($C_{0.15}$) which gives 1.1% (1.2%) uncertainty. Most combinations, however, give subpercent precision. Adding all their uncertainties in the most unfavorable way generates an extreme of 2.6% power spectrum uncertainty. Thus unlike the weak lensing probe analyzed by (11; 12), redshift space distortions do not require any priors to be placed on the power spectrum parameters (assuming that the

KLL reconstruction form is valid).

Remarkably, in addition to selfcalibration, the additional fit parameters have little impact on the cosmological parameter estimation, enlarging the uncertainties by only 9%, 22%, and 7% on w_0, w_a, and γ. And of course including the extra parameters removes any cosmology bias as suffered in the previous section (modulo model validity). Regarding the 12 extra reconstruction parameters, from Fig. 4.1 we see that F is smooth in k so taking bins of width 0.1 in k is reasonable. For completeness, for bins of width 0.02 (and hence 60 extra parameters) we find that cosmological parameter estimation is mildly degraded, with uncertainties on w_0, w_a, γ increasing relative to 0.1 width by 16%, 33%, 17%.

4.5.2 MAXIMUM WAVENUMBER

Let us examine the dependence of the results on the maximum wavenumber $k_{\rm max}$ used. Note that for the $k_{\rm max} = 0.1$ case, the cosmology parameters are not determined particularly well even though no reconstruction parameters are used for $k \leq 0.1\,h/{\rm Mpc}$. This is due to strong covariance with the 27 galaxy bias parameters. Once beyond the linear regime, this degeneracy is broken and the correlation coefficients drop, greatly improving the cosmological parameter determination (e.g. by factors of 2.9–3.4 on the dark parameters, for $k_{\rm max} = 0.2$ relative to $k_{\rm max} = 0.1$). This continues for higher $k_{\rm max}$, despite the addition of further reconstruction parameters, but gradually saturates. For example, while relative to the $k_{\rm max} = 0.5$ case the uncertainties on w_0, w_a, or γ at $k_{\rm max} = 0.2$ are greater by a factor $\sim$ 3, at $k_{\rm max} = 0.3$ the factor is 1.6, and at $k_{\rm max} = 0.4$ is 1.15, as seen in Fig. 4.5. Thus, having an accurate reconstruction form out to $k_{\rm max} \approx 0.4 - 0.5$ is sufficient for robust cosmological parameter estimation, while selfcalibration obviates the need for any prior knowledge of the values of the reconstruction parameters.

Figure 4.5 Extending $k_{\rm max}$ to values above $0.1\,h/{\rm Mpc}$ breaks degeneracies, leading to improvements in cosmological parameter estimation as shown here, even given the addition of reconstruction parameters to marginalize over. Reconstruction to $k_{\rm max} = 0.4 - 0.5\,h/{\rm Mpc}$ is sufficient to plateau the cosmology estimation precision.

Binning such as we use is model independent and closest to the weak lensing work. This model independence is important since in the absence of a large suite of simulations we may have no particular confidence in parametrizations such as those in Eqs. (4.13-4.15). Recall that those equations merely give the fiducial values in each k bin, and then we allow the bin values to float freely and marginalize over them. Given simulations we might be able to adopt specific forms and fit for a reduced set of parameters, e.g. the coefficients in those equations.

4.5.3 REDSHIFT DEPENDENCE

To give a first indication of whether adding redshift dependence to A, B, C affects the conclusions we include a variation of the characteristic wavenumber scale entering in the nonlinearity amplitude A in Eq. (4.13), i.e. the $0.39\,h/{\rm Mpc}$, writing this as

$$k_\star(z) = 0.39\,(1+z)^{\alpha/1.58}\,h/{\rm Mpc}\,, \tag{4.24}$$

and adding this evolution parameter α to the fit. The simulation results in (15) indicate that relative to A, the parameters B and C have negligible additional redshift dependence. Therefore we scale B and C by the same factor as A, i.e.

$$\begin{aligned} B(k,z) &= B(k,0)\,A(k,z)/A(k,0) && (4.25)\\ &= B(k,0)\,\frac{1+[A(k,0)-1](1+z)^{-\alpha}}{A(k,0)}\,, && (4.26)\end{aligned}$$

and the same for C.

The introduction of redshift dependence through α has little impact on the cosmological parameter estimation; the largest correlation coefficient of α cosmologically is 0.31, with γ, and overall 0.83 with $B_{0.45}$, while α itself is determined to within 0.025. Figure 4.6 shows the influence on dark cosmology parameter estimation of marginalization over the reconstruction parameters with and without redshift dependence, and fixing the reconstruction parameters (i.e. with a total of 48, 47, or 35 parameters). Uncertainties on w_0, w_a, and γ increase by only 0.8%, 0.2%, 5% respectively upon including α. Other forms of redshift dependence may have different cosmological impact, and this deserves further analysis through simulations, but the scaling of the characteristic wavenumber as used here should give a reasonable first indication.

Figure 4.6 Joint 2D 1σ confidence contours on the dark cosmology parameters are shown for the cases of all reconstruction parameters being fixed (dotted red), marginalized over without redshift dependence as in most of the article (solid black), and additionally marginalizing over a redshift evolution parameter α (dashed blue). Here $k_{\rm max} = 0.5\,h/{\rm Mpc}$. Note that the fixed F case, shown here centered on the true cosmology, could be strongly biased if F was misestimated (see Sec. 4.4).

4.5.4 NONLINEAR POWER SPECTRUM

The greatest effect of uncertainty in the reconstruction function comes from the parameters A_k, as seen in Fig. 4.1 and in Sec. 4.4 regarding the parameter bias. Recall that $A(k)$ arises from the nonlinear effects in the density field, and even exists when $\mu = 0$. In this limit $A(k)$ acts to map the linear real space density power spectrum to the nonlinear regime. Therefore, if we had a robust nonlinear (or quasilinear) real space power spectrum then we would have no need of a separate parameter as then $A(k) = 1$ (this has been tested and found accurate to subpercent level by (15)). Substantial effort is going into developing cosmic emulators (29) that could provide accurate nonlinear power spectra, eventually including the full set of cosmological parameters and redshifts considered here. Since that is still in the future, we consider the nonlinear prescription of Halofit (30) to get an indication of the potential impact on our conclusions.

The linear power spectrum at a given redshift is fed into Halofit to give the approximate nonlinear form. This removes the need for $A(k)$, setting this equal to one for all k and z. As indicated earlier in this section, the simulation results from (15) imply that for such a normalized A then the quantities B and C become substantially redshift independent. Therefore we do not need any hypothetical model such as the α parametrization, making the entire analysis more robust. Furthermore, Halofit includes cosmology dependence and so no assumption about universality of A is needed.

We show the results for cosmological parameter estimation in Table 4.3, for $k_{\rm max} = 0.5\,h/{\rm Mpc}$, for the three cases of using the model independent approach of fitting for $A(k)$ in bins, using Halofit and $A = 1$, and using the revised version of Halofit from (31) and $A = 1$. In all cases we still fit for the binned values of B and C. The use of functional forms for the nonlinear real space power spectrum allows better determination of the cosmological parameters, by 12%, 28%, 12% for w_0, w_a, γ respectively (15%, 33%, 15% for revised Halofit, which has slightly more quasilinear power). This offers some promise for the future use of cosmic emulators, but in this paper we prefer to be conservative in the estimations and use the model independent approach.

	w_0	w_a	γ	Ω_m
Fit $A(k)$	0.0713	0.306	0.0163	0.00794
Halofit	0.0624	0.220	0.0143	0.00678
New Halofit	0.0603	0.206	0.0139	0.00608

Table 4.3 1σ constraints as in Table 4.2, using $k_{\rm max} = 0.5\,h/{\rm Mpc}$, but for three different methods of treating nonlinearities.

4.6 CONCLUSIONS

With the ability to map galaxy clustering in three dimensions over large volumes of the universe comes the necessity for accurate theoretical interpretation. This entails linking the isotropic, linear theory real space density power spectrum to the observed anisotropic, nonlinear redshift space galaxy power spectrum. We have investigated here some of the relevant issues involving nonlinearities in the density field and velocity effects, using the Kwan-Lewis-Linder analytic redshift space reconstruction function calibrated from numerical simulations.

The main question addressed is to what accuracy the anisotropic redshift space power spectrum must be known in order to achieve robust cosmological conclusions. We propagate uncertainties in the power spectrum through a model independent binning of reconstruction amplitudes with wavenumber and assess the effects of deviations from fiducial values. To avoid biasing cosmological parameters such as the dark energy equation of state and gravitational growth index requires down to 0.3% accuracy on the reconstruction parameters in the most stringent cases, while smoother deviations give more tractable requirements. Note that it is only those deviations that mimic cosmological variations that are most important.

A more flexible and robust approach is to carry out a global fit for the binned reconstruction parameters simultaneously with the cosmological parameters, which avoids biasing the results so long as the *form* of the KLL function is a good approximation. With 8 cosmology parameters and 39 systematics parameters we find that a next generation galaxy redshift survey such as BigBOSS can tightly and accurately constrain cosmology, for example determining the equation of state time variation w_a to 0.3 and testing gravity through the growth index γ to 3%. No external priors on the reconstruction parameters are necessary as they are selfcalibrated by the survey, most at the subpercent level. This also corresponds to subpercent calibration of the redshift space power spectrum.

We have tested the robustness of the conclusions by adding redshift evolution, which has little effect, varying the number of wavenumber bins, and exploring the leverage from increasing the maximum wavenumber used, $k_{\rm max}$. Cosmological leverage plateaus by $k_{\rm max} = 0.4 - 0.5\,h/{\rm Mpc}$ so the KLL form need only apply up to this scale. We made no assumptions about the departure from linearity, allowing the nonlinearity amplitude to float in a model independent manner in bins of k, but also then analyzed the impact of adopting a nonlinear prescription such as Halofit (or its revision). This improved the cosmology estimation and offers a promising sign to motivate continued development of cosmic emulators for the nonlinear power spectrum.

Several areas exist for further development. The KLL form has been tested for dark matter, but (15) indicates it is successful for halos as well. Eventually this must be extended to galaxies, a major undertaking. On the positive side, we achieved excellent results using reconstruction starting from simple linear theory (Kaiser approximation); higher order perturbation theory approaches extend the range where reconstruction is milder. Universality, i.e. cosmology dependence, of the reconstruction is a major topic for future investigation, requiring large suites of cosmological simulations, again suited for emulators. We have taken a first step toward addressing this effect by exploring the influence of using Halofit and new Halofit cosmological dependences for the nonlinearity. Simulations may also enable us to compress the information in bins down to a smaller set of parameters.

Redshift space distortions provide a powerful tool for measuring the growth rate of cosmic structure, and delivering insights on the competition between the gravitational laws driving clustering and accelerated expansion suppressing it. The results here give encouraging indications, and quantitative measures, that theoretical analysis can take into account robustly the nonlinear and velocity effects to extract accurate cosmology from the forthcoming large volume redshift surveys.

We thank Sudeep Das, Juliana Kwan, and Alberto Vallinotto for helpful discussions. This work has been supported by DOE grant DE-SC-0007867 and the Director, Office of Science, Office of High Energy Physics, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. EL acknowledges World Class University grant R32-2009-000-10130-0 through the National Research Foundation, Ministry of Education, Science and Technology of Korea; JS is supported by the Dark Cosmology Centre, funded by the Danish National Research Foundation, and thanks LBNL for additional support.

BIBLIOGRAPHY

[1] W.J. Percival & M. White, MNRAS 393, 297 (2009) [arXiv:0808.0003]

[2] T. Okumura & Y.P. Jing, ApJ 726, 5 (2011) [arXiv:1004.3548]

[3] E. Jennings, C.M. Baugh, S. Pascoli, MNRAS 410, 2081 (2011) [arXiv:1003.4282]

[4] E. Jennings, C.M. Baugh, S. Pascoli, ApJL 727, 9 (2011) [arXiv:1011.2842]

[5] J. Tang, I. Kayo, M. Takada, MNRAS 416, 2291 (2011) [arXiv:1103.3614]

[6] J. Kwan, G.F. Lewis, E.V. Linder, ApJ 748, 78 (2012) [arXiv:1105.1194]

[7] R. Scoccimarro, Phys. Rev. D 70, 083007 (2004) [arXiv:astro-ph/0407214]

[8] A. Taruya, T. Nishimichi, S. Saito, Phys. Rev. D 82, 063522 (2010) [arXiv:1006.0699]

[9] B.A. Reid & M. White, MNRAS 417, 1913 (2011) [arXiv:1105.4165]

[10] T. Okumura, U. Seljak, P. McDonald, V. Desjacques, JCAP 1202, 010 (2012) [arXiv:1109.1609]

[11] D. Huterer & M. Takada, Astropart. Phys. 23, 369 (2005) [arXiv:astro-ph/0412142]

[12] A.P. Hearin, A.R. Zentner, Z. Ma, JCAP 1204, 034 (2012) [arXiv:1111.0052]

[13] A. Lewis, A. Challinor, A. Lasenby, ApJ 538, 473 (2000) [arXiv:astro-ph/9911177]; `http://camb.info`

[14] N. Kaiser, MNRAS 227, 1 (1987)

[15] J. Kwan, Ph.D. thesis, University of Sydney (2011)

[16] H.A. Feldman, N. Kaiser, J.A. Peacock, ApJ 426, 23 (1994) [arXiv:astro-ph/9304022]

[17] H-J. Seo & D.J. Eisenstein, ApJ 598, 720 (2003) [arXiv:astro-ph/0307460]

[18] A. Stril, R.N. Cahn, E.V. Linder, MNRAS 404, 239 (2010) [arXiv:0910.1833]

[19] D. Schlegel et al, arXiv:1106.1706 ; `http://bigboss.lbl.gov`

[20] P. McDonald & U. Seljak, JCAP 0910, 007 (2009) [arXiv:0810.0323]

[21] E.V. Linder, Phys. Rev. D 72, 043529 (2005) [arXiv:astro-ph/0507263]

[22] E.V. Linder & R.N. Cahn, Astropart. Phys. 28, 481 (2007) [arXiv:astro-ph/0701317]

[23] E.V. Linder, Astropart. Phys. 29, 336 (2008) [arXiv:0709.1113]

[24] L. Knox, R. Scoccimarro, S. Dodelson, Phys. Rev. Lett. 81, 2004 (1998) [arXiv:astro-ph/9805012]

[25] E.V. Linder, Astropart. Phys. 26, 102 (2006) [arXiv:astro-ph/0604280]

[26] S. Das, R. de Putter, E.V. Linder, R. Nakajima, JCAP 1211, 011 (2012) [arXiv:1102.5090]

[27] S. Dodelson, C. Shapiro, M. White, Phys. Rev. D 73, 023009 (2006) [arXiv:astro-ph/0508296]

[28] C. Shapiro, ApJ 696, 775 (2009) [arXiv:0812.0769]

[29] E. Lawrence, K. Heitmann, M. White, D. Higdon, C. Wagner, S. Habib, B. Williams, ApJ 713, 1322 (2010) [arXiv:0912.4490] ; `http://www.lanl.gov/projects/cosmology/CosmicEmu`

[30] R.E. Smith et al, MNRAS 341, 1311 (2003) [arXiv:astro-ph/0207664]

[31] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, M. Oguri, arXiv:1208.2701

5

STELLAR INTERACTIONS AND GRAVITATIONAL WAVE ASTRONOMY

The Formation of Eccentric Compact Binary Inspirals and the Role of Gravitational Wave Emission in Binary-Single Stellar Encounters

Johan Samsing[1], Morgan MacLeod[2], Enrico Ramirez-Ruiz[2]

[1] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark
[2] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064

Abstract

The inspiral and merger of eccentric binaries leads to gravitational waveforms distinct from those generated by circularly merging binaries. Dynamical environments can assemble binaries with high eccentricity and peak frequencies within the *LIGO* band. In this paper, we study binary-single stellar scatterings occurring in dense stellar systems as a source of eccentrically-inspiraling binaries. Many interactions between compact binaries and single objects are characterized by chaotic resonances in which the binary-single system undergoes many exchanges before reaching a final state. During these chaotic resonances, a pair of objects has a non-negligible probability of experiencing a very close passage. Significant orbital energy and angular momentum are carried away from the system by gravitational wave (GW) radiation in these close passages and in some cases this implies an inspiral time shorter than the orbital period of the bound third body. We derive the cross section for such dynamical inspiral outcomes through analytical arguments and through numerical scattering experiments including GW losses. We show that the cross section for dynamical inspirals grows with increasing target binary semi-major axis, a, and that for equal-mass binaries it scales as $a^{2/7}$. Thus, we expect wide target

binaries to predominantly contribute to the production of these relativistic outcomes. We estimate that eccentric inspirals account for approximately one percent of dynamically assembled non-eccentric merging binaries. While these events are rare, we show that binary-single scatterings are a more effective formation channel than single-single captures for the production of eccentrically-inspiraling binaries, even given modest binary fractions.

5.1 SUMMARY

In globular clusters and galactic nuclei a large fraction of the stars are believed to be in binaries with other stars or compact objects. The density in these environments can be up to a million times higher than that in our solar neighborhood, making it highly possible that a binary will experience at least one single encounter within its lifetime. This leads to a three-body, or binary-single, interaction. These interactions not only give rise to a large variety of observable phenomena, but also play an important role in the evolution of e.g. globular clusters, by storing or releasing binding energy into the surrounding field stars.

I investigated the role of gravitational wave (GW) radiation in these binary-single interactions. For this study I wrote an N-body code including GWs in the post-Newtonian framework. By performing millions of binary-single scattering experiments using this code, I discovered a completely new outcome never characterized before: during a chaotic three-body interaction a temporarily formed binary can spiral in by emitting GWs while the third object is still bound. This inspiraling binary has an extremely high eccentricity when it passes through the observable LIGO band, which is in contrast to ordinary field binaries that circularize long before they become observable. This new outcome therefore gives a unique possibility to observe how GR impacts high eccentricity systems.

To estimate the probability for this new GR outcome, I calculated the associated cross section $\sigma(a)$ as a function of the semi-major axis a of the target binary by numerical scattering experiments. I found surprisingly that $\sigma(a)$ *increases* with a - a very counterintuitive result. I managed to also develop a theoretical description and calculated analytically the relation $\sigma(a) \propto a^{2/7}$, which is in perfect agreement with the numerical simulations. I performed a broad range of simulations, from equal mass scatterings to situations with white dwarf - neutron star target binaries. Using the resultant cross sections folded with a typical globular cluster model, my collaborators and I estimated this new outcome to be the dominant GW source of high eccentricity mergers visible by LIGO. This work has therefore great impact on not only stellar dynamics, but also future searches for GWs.

5.2 INTRODUCTION

The density of stars in galactic nuclei and in the centers of some globular clusters can be more than a million times higher than that in our solar neighborhood (76). In such cases, a primordial binary will undergo a close encounter with at least one other star with high probability within its lifetime (e.g. 100). It is in these environments, called dense stellar systems, that binary populations will no longer be truly primordial as their stellar composition, eccentricity, and period distributions will be largely determined by past interactions with other stars (e.g. 81; 56; 57; 60; 51; 62; 28; 61; 59). This transformation of binary systems was envisioned by Hills (49), who suggested that exchanging neutron stars into preexisting binaries might be a natural way to form X-ray binaries as byproducts.

Dynamical friction causes the heaviest stars and primordial binaries to concentrate towards the cluster's core (85; 32; 31). Since the heaviest stars tend to be left in the binary following such three-body encounters (this can be understood as consequence of the tendency toward energy equipartition, in which the lighter star would have the highest velocity in the final state), binaries are quite effective at soaking up heavy stars such as neutron stars and heavy white dwarfs (50; 116; 117; 45), even if none of them originally had a companion.

After such an exchange, the binary will not only be slightly wider but also heavier, which will result in gravitational focusing being more effective. The binary's cross section for encounters will thus be larger than before the exchange. For this reason, a binary likely to undergo one exchange over some time period is likely to have several more encounters coming rapidly after the first exchange (116). The tendency to exchange the heaviest compact stars also has the consequence that the rates of ejection of binaries involved in three-body exchanges are less than those predicted by models in which all stars have equal masses. The recoil speeds of the light, single stars are consequently larger.

A large fraction of the encounters where the field star approaches within approximately a binary semi-major axis (SMA), a_0, of the binary center of mass result in resonant interactions, in which the three stars wander for a long time on chaotic orbits and approach each other repeatedly (43; 54). During these chaotic encounters, the stars have many opportunities for close encounters. If the stars are compact, angular momentum loss due to gravitational radiation may become a noticeable effect during close passages (97), and could cause the two stars to be driven together. It is the interplay between binaries and compact objects in such dense environments and their ability to manufacture eccentric merging binaries in three-body exchanges that forms the main topic of this work.

Our main goal in this paper is to study how the inclusion of gravitational wave (GW) losses modifies the compact binary outcomes that originate from three-body scatterings, in particular during resonant interactions. The inclusion of GW losses into the binary-single dynamical system, we argue, introduces a new potential outcome in which a pair of objects may dynamically inspiral and merge while the three-body system is still in resonance. These outcomes are rare, and they are typically only realized during resonant interactions. Chaotic, resonant orbits augment the probability of very close passages when compared to direct interactions, and they can produce systems with correspondingly short GW inspiral time. Gültekin et al. (40) first explored the cross section for these inspiral outcomes in the context of IMBH formation and growth. A surprising result of Gültekin et al.'s simulations is that the cross section for inspiral outcomes *increases* with increasing binary SMA. This is perhaps counterintuitive because one might expect that the cross section for relativistic outcomes would be largest in very tight binaries. However, we will show that this is a natural consequence of resonant binary-single interactions, and that the scaling with binary SMA can be analytically derived.

In this paper, we explore the cross section for dynamical inspiral outcomes during binary-single interactions through numerical experiments and analytic calculations. In Section 5.3, we review some of the dynamical properties and outcomes of binary-single interactions. In order to build intuition for how the inclusion of GW losses modifies binary-single interaction dynamics, in Section 5.4 we summarize the results binary-single scatterings with point masses in Newtonian gravity. Readers familiar with previous work in binary-single dynamics may wish to skip to Section 5.5, in which we describe the inclusion of post-Newtonian (PN) corrections to the binary-single system equation of motion. Section 5.6 describes the formation of dynamical inspirals from resonant interactions between hard binaries and single objects. We explain the origin of these inspirals through numerical scattering experiments, and use our results to motivate an analytic derivation of the scaling of the inspiral cross section with binary SMA. In Section 5.7, we show that dynamical inspirals give rise to inspirals that pass with high eccentricity through the *LIGO*[1] band. We compare this process to eccentric inspirals arising from single-single interactions and show that the cross section is greatly enhanced in binary-single interactions. In Section 5.8, we extend our calculations to consider binaries containing white dwarfs, we discuss binary lifetimes and the role of GW emission, and we estimate whether the products of binary-single interactions are ejected or retained in their host stellar system. Finally, we estimate the rates of eccentric inspirals given typical globular cluster core properties.

5.3 BINARY-SINGLE ENCOUNTERS

Binary-single stellar encounters in dense stellar systems may be broadly divided into a few well-defined categories. In the majority of encounters, the incoming object passes on a hyperbolic trajectory relative to the binary at a distance large compared to the binary separation (43). The passage time is greater than the binary's orbital period and the binary is subjected to a weak perturbation (WP). A strong perturbation (SP) is possible (43) when the incoming object approaches the binary on a hyperbolic trajectory that happens to pass at a distance comparable to the binary SMA. In this case, the interaction time is less than or similar to the binary's orbital period.

The accumulation of WPs and SPs across the lifetime of a binary in a dense stellar system modifies the expected eccentricity and SMA distributions as compared to more isolated binaries. To quantify this effect, one must rely on integrations of the coevolution of binaries and their parent clusters over the cluster's relaxation time (e.g. 1; 47; 48; 43; 76; 80; 6; 30; 58; 33; 31).

5.3.1 CLOSE INTERACTIONS AND THEIR CROSS SECTION

A close interaction (CI), by contrast, occurs when the incoming object passes within a sphere of influence marked by the binary's separation. In these cases, the gravitational interaction between all three bodies may be of similar strength, and the outcomes are chaotic. In this work, we will focus on CIs and the dramatic role they play in reshaping binaries. Figure 5.1 shows a schematic overview of the different interactions and their expected outcomes.

We define a CI to have occurred when the third body passes within a distance $r_{\rm CI}$ from the binary center of mass. We choose $r_{\rm CI}$ as the distance from the center of mass to the lighter object in the binary,

$$r_{\rm CI} = \frac{m_2}{m_1 + m_2} a_0, \tag{5.1}$$

---

[1] http://www.ligo.caltech.edu/

where $1, 2$ are the binary members in order of ascending mass ($m_2 > m_1$), 3 is the incoming object, and $m_1 + m_2$ is the mass of the target binary. This value is always between $a_0/2$ (if $m_1 = m_2$) and a_0 (if $m_2 \gg m_1$).

Whether a CI will occur is analytically predictable given the impact parameter, b, and velocity, v_∞, of the third body relative to the target binary. At large separations between the binary and the incoming object, the fact that the binary is composed of two objects is unimportant and thus the encounter can realistically be treated as the interaction between two point masses: the binary with total mass $m_{\rm bin} = m_1 + m_2$ and the incoming object with mass m_3. In this case, a given distance of closest approach between the incoming single and the center-of-mass of the binary, $r_{\rm min}$, corresponds directly to an impact parameter, b, defined at infinity (116),

$$b = r_{\rm min}\sqrt{1 + \frac{2Gm_{\rm tot}}{r_{\rm min}v_\infty^2}}, \tag{5.2}$$

where v_∞ is the initial relative velocity at infinity of the binary center of mass and the single object, and $m_{\rm tot} = m_{\rm bin} + m_3$. The second term in this expression corresponds to the gravitational focusing of trajectories from an initially large impact parameter to a closer pericenter distance. Because the argument of the square root is always larger than unity, b is always greater than $r_{\rm min}$.

If we now consider the interactions with a closest approach less than the sphere of the binary, $r_{\rm CI}$, then we see that all encounters with impact parameter less than the corresponding $b_{\rm CI} = b(r_{\rm CI})$ will have $r_{\rm min} < r_{\rm CI}$. Therefore, all encounters coming from within the area $\sigma_{\rm CI} = \pi b_{\rm CI}^2$ will lead to an interaction with $r_{\rm min} \leq r_{\rm CI}$. This area $\sigma_{\rm CI}$ is defined as the *cross section* for a close interaction. Given the definition of b above, this may be written

$$\sigma_{\rm CI} = \pi b_{\rm CI}^2 = \pi r_{\rm CI}^2\left(1 + \frac{2Gm_{\rm tot}}{r_{\rm CI}v_\infty^2}\right). \tag{5.3}$$

Whether the first (geometric) or second (gravitational focus) term in parenthesis dominates depends on the relative binding energy of the binary and the kinetic energy of the incoming object.

Given a distribution of single stars, the CI cross section, $\sigma_{\rm CI}$, gives an estimate of how often such interactions can occur. As $\sigma_{\rm CI}$ increases, the more encounters will be focused into the binary system. In a stellar system with an isotropic stellar density, n, and typical relative velocity, v_∞, this rate of CIs per binary may be approximated as

$$\Gamma_{\rm CI} \simeq n\sigma_{\rm CI}v_\infty. \tag{5.4}$$

Thus, given a stellar distribution, the cross section is the only factor that determines the relative rates of different processes. For this reason, significant effort will be invested in deriving the cross sections of the various outcomes of CIs as fractions of the total CI cross section. In the following section, we explore the role of the relative energy of the binary and the single object in shaping binary-single interactions.

5.3.2 HARD AND SOFT TARGET BINARIES

The relative velocity of the binary and the single object, v_∞, as compared to the characteristic velocity of a binary, $v_{\rm c}$, determines the outcomes that are possible in a binary-single interaction. A binary's characteristic velocity is defined as (55)

$$v_{\rm c}^2 = G\frac{m_1m_2(m_1 + m_2 + m_3)}{m_3(m_1 + m_2)}\frac{1}{a_0}. \tag{5.5}$$

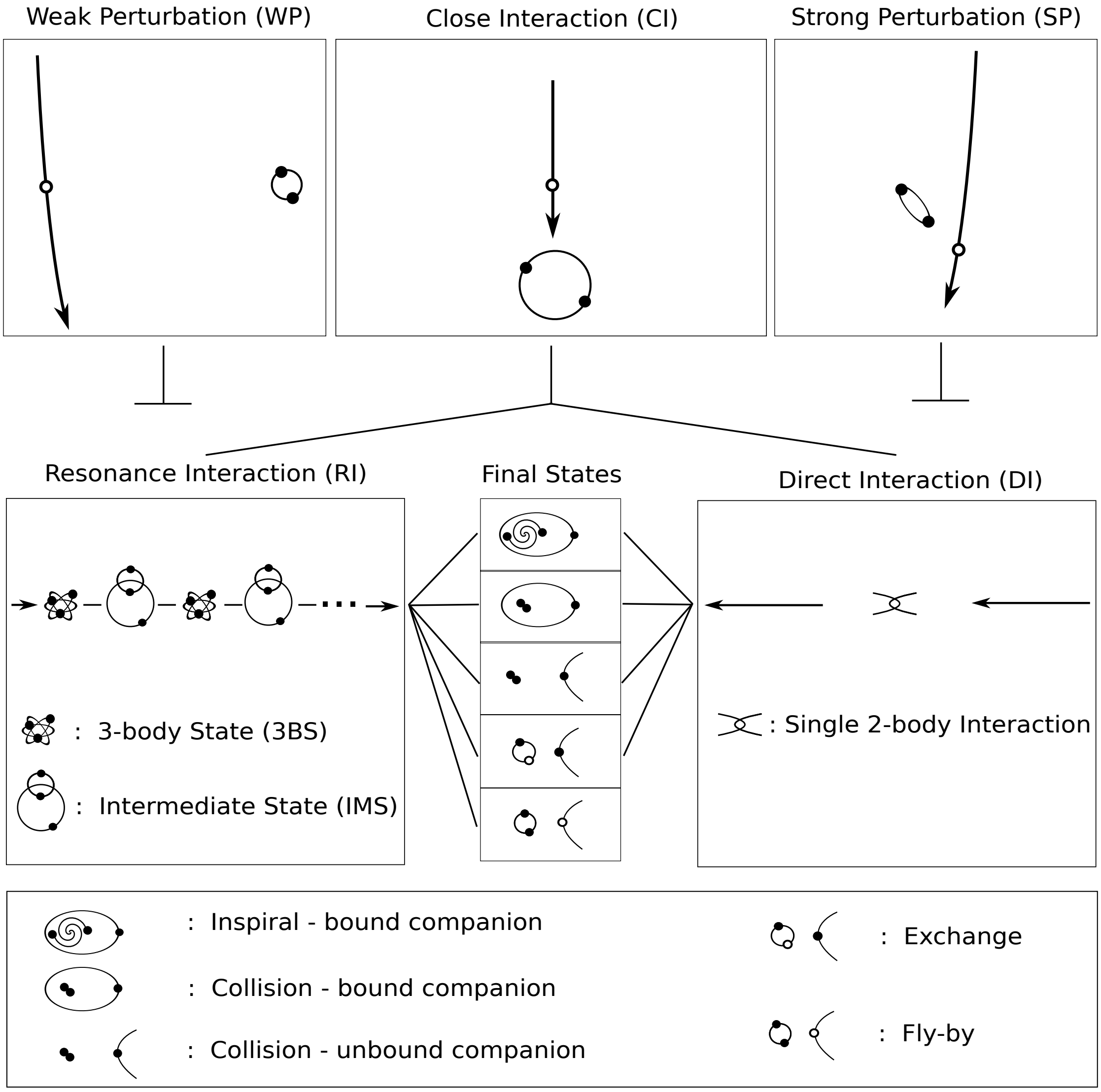


Figure 5.1 Schematic illustration of binary-single interactions and their final states. The *top panel* shows three different types of interactions. The *top left panel* shows a weak perturbation (WP) where the single encounter is only weakly perturbing the binary, but over several orbital periods. The *top right panel* shows a short but strongly perturbing encounter (SP). A close interaction (CI) is shown in the *middle panel*. The evolution of the system from this CI channel can further be divided into the two interaction channels: direct interaction (DI) and resonant interaction (RI). These are illustrated in the *middle panel*. The RI channel can be decomposed into intermediate binary-single states (IMS), where an intermediate binary is formed with a bound companion. Several IMS are created and destroyed in the chaotic RI before a final state is reached. The RI erases any information of initial conditions. The DI channel is on the other hand very fast and, as a result, the endstate depends sensitively on the initial state. Which channel dominates depends particularly on the mass ratio between the objects and the velocity of the incoming object. The set of endstates from both the RI and the DI interactions are listed in the *middle panel* where the individual interaction diagrams are defined in the *bottom panel*. There is, in general, a similar final state scheme for each permutation of the objects.

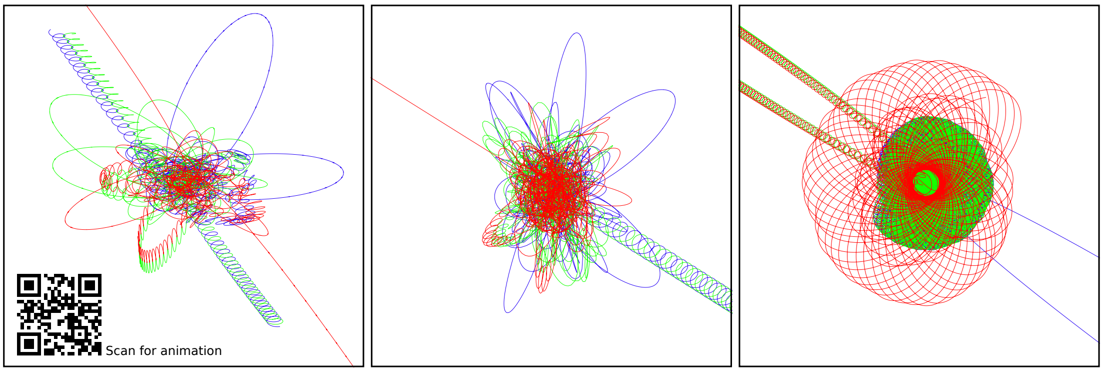


Figure 5.2 Examples of binary-single resonance interactions between equal mass black holes. *Left:* The close interaction forms a few intermediate binary states (red-green) with a bound companion (blue). We denote these intermediate binary-single states (IMS). Scan the `QR code`[2] to see an animation of this interaction. The last frame in the animation is shown in the plot. *Middle*: This interaction have the same total energy as in the *left* panel but the energy is distributed differently in the system. *Right*: A relative rare interaction class is displayed in this panel, with a system composed of a binary (blue and green) that remains bound to a companion (red) for many orbits. The final state is a collision. The interactions shown in the *left* and *middle* panels are generally refer to as *democratic resonances* where the right panel shows an example of an *hierarchical resonance*. The many IMS created during these resonant interactions aid the formation of eccentric binaries with short inspiral times. Binaries that inspiral and merge due to GW radiation during a resonance interaction are called *inspirals*. An example is shown in Figure 5.3.

This velocity is written such that if the relative velocity at infinity is larger than v_c ($v_\infty > v_\mathrm{c}$), then the total energy of the three-body system is positive (43).

A binary with $v_\infty > v_\mathrm{c}$ is described as a soft binary (SB) relative to its environment. The cross section for close interaction, equation (5.3), is well approximated by the binary's geometrical cross section, πr_CI^2. Because the velocity at infinity is greater than the binary's orbital velocity, the binary appears nearly static during the interaction. The resultant encounters can thus be viewed mainly as two-body interactions that are well described by *impulsive* approximations (43; 53). Additionally, with $v_\infty > v_\mathrm{c}$ the incoming body carries a large amount of energy when compared to the binary's binding energy. That excess of energy can effectively be utilized to split the binary (43).

Hard binaries (HB) are characterized by $v_\infty < v_\mathrm{c}$. In this case, the cross section for CI is dominated by the gravitational focus term, and

$$\sigma_\mathrm{CI} \simeq \frac{2\pi G m_\mathrm{tot} r_\mathrm{CI}}{v_\infty^2} \tag{5.6}$$

Thus, in this limit, $\sigma_\mathrm{CI} \propto a_0/v_\infty^2$. Further, the energy carried from the encounter into the system is relatively small and a temporary bound triple state can be formed (53).

In dense stellar systems, the HB limit is typically the relevant limit for the steady-state binary population. Equation (5.5) can be re-written for equal mass encounters as

$$v_\mathrm{c} \approx 36.5\,(m/M_\odot)^{1/2}\,(a_0/\mathrm{AU})^{-1/2}\,\mathrm{km\ s}^{-1}. \tag{5.7}$$

Values for v_∞ are in the 10-50 km s^{-1} range for galactic GCs (76). Thus any binaries with SMA smaller than ≈ 1 AU will be in the HB limit. In clusters, HBs tend to be the ones that survive as encounters tend to split soft binaries (43). Further, based on a statistical trend toward energy equipartition (43; 48), hard binaries tend to become harder (as energy is transferred from the binary to put the single on an unbound orbit) while soft binaries get softened or disrupted (as the incoming single star pumps energy into the system before leaving). This natural selection makes a hard binary population even harder and causes a soft binary population to evaporate.

Binary-single CIs involving HBs may be decomposed into direct interactions (DIs) and resonant interactions (RIs). DIs are brief, two-body interactions which occur when the incoming body passes very close to only one of the binary members. In these cases, the interaction is brief and the initial conditions with which the single object entered the binary are key in determining the outcome. By contrast, a RI is comprised of many intermediate exchanges of binary and single star hierarchy. We denote these temporary triple-object states, comprised of a binary and a bound single, as intermediate states (IMSs). The IMS decomposition is illustrated in Figure 5.1.

The number of resonances a system undergoes during a RI depends on the mass ratio of the interacting objects and is maximized for equal mass objects (116). In the equal mass case, these RIs can have lifetimes extending from one to several hundred times the orbital period of the initial target binary. If one of the objects is lighter compared to the others, this object is likely to be dynamically kicked out, leaving the heavier objects behind as a binary (116). An illustration of the possible orbital morphologies of RIs is shown in Figure 5.2. Examples of both democratic (similar pairwise binding energy) and hierarchical (disparate pairwise binding energies) resonances can be clearly seen in Figure 5.2.

Compared to the entire duration of a RI, the lifetimes of individual IMSs are relatively short. This implies that a single RI encompasses many IMS exchanges in which close encounters occur and the binary-single system is transformed. The IMSs themselves are unstable because they are disrupted every time the current bound single object makes a close passage. Over the course of several such IMS changes (through three-body interaction *knots*), the triple system evolves chaotically, loosing memory of the initial conditions with which the single object first entered the binary (43). Rare outcomes may be achieved with higher likelihood in RIs for the simple reason that the single object makes many randomized close passages through the binary system. This is particularly significant when GW radiation is included into the three-body equation of motion because there is a non-negligible probability that a very close (and thus highly dissipative) passage will take place.

5.3.3 OUTCOMES OF CLOSE INTERACTIONS

In the previous section we have described how CIs arise in binary-single star encounters and how their likelihood can be quantified by their cross section, σ_{CI}. During a CI, the system is in a three-body state, but no three-body state is stable (54) and the system will thus invariably evolve (through the DI or the RI channel) into one out of the several possible final-states (or *outcomes*) as illustrated in Figure 5.1. In general, there is a given cross-section for each of these possible outcomes to occur. In Section 5.3.4, we describe how we compute these outcome cross sections statistically based on the fraction of binary-single scatterings that can generate a given outcome. In the two sections below we describe in detail the particular final outcomes expected from CI interactions.

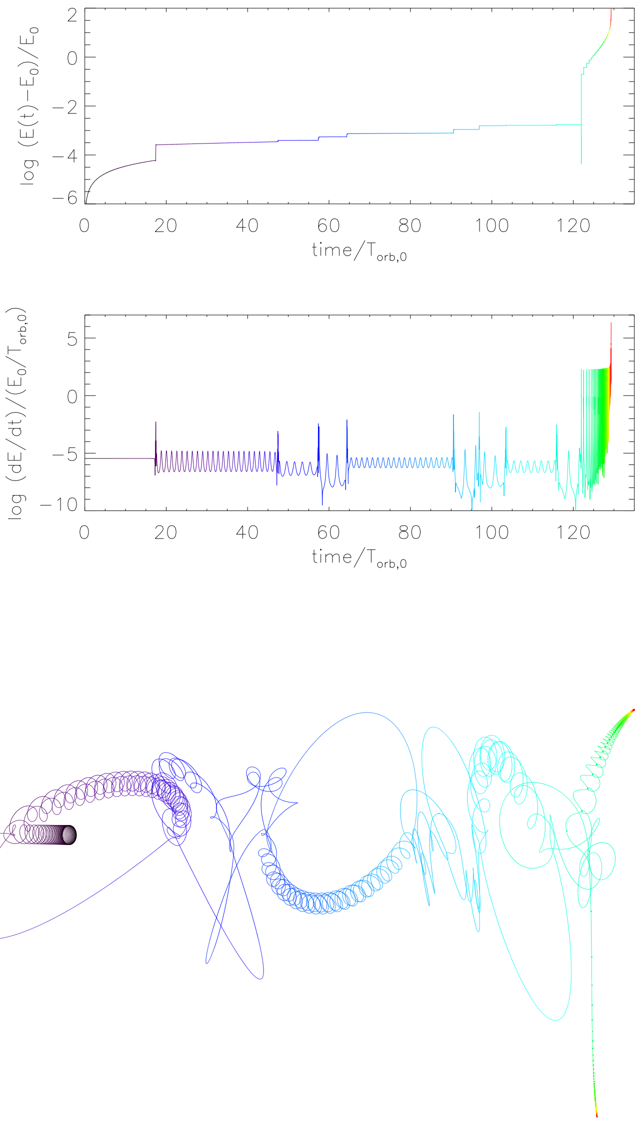


Figure 5.3 Example of a binary-single interaction which ends as an *inspiral*. These are the new general relativistic (GR) endstates that are the main focus of our work. *Inspirals* are IMS-formed binaries that merge due to GW radiation during the resonance interaction, i.e. while the single object is still bound. The color across all plots denotes time. *Top*: Fraction between the energy loss of the system after a time t and the initial energy of the system E_0. Any deviations from zero are due to energy radiated away by GWs. *Middle*: Energy loss dE/dt as a function of time. The oscillating form of dE/dt arises because the system evolves between multiple IMS. *Bottom*: The ensuing binary-single trajectories. The interaction starts at the left where the binary interacts with the incoming object. The final state seen at the far right is an IMS binary that inspirals due to GW radiation while the third object is still bound. The final inspiral is as expected characterized by a large and rapid increase in GW losses. These inspirals can by observed with *LIGO* and as we will show later are likely to be highly eccentric at the time of observation, which makes them particularly interesting.

Outcomes from Newtonian Gravity

In Newtonian gravity, a binary-single interaction can result in a binary with an unbound companion, a collision, or three unbound objects. The cases in which a binary is left behind may be further subdivided based on the properties of the surviving binary (43; 55). If the binary is composed of the original two objects (1,2) then we refer to the encounter as a *fly-by* even though the endstate binary may be the result of a more complex interaction than the fly-by label suggests. If instead the binary is composed of one of the original binary members and the third body, we denote the encounter an *exchange*. In this case, the binary may either be (1,3) or (2,3). An outcome in which all three members are mutually unbound is possible when the total system energy is positive, $v_{\infty} > v_{\rm c}$. This outcome is denoted as an *ionization*. Collisions are possible at all values of v_{∞}, but they are most likely to occur at negative total binding energies where the gravitational focus cross sections of the individual objects are larger.

Inspiraling Binaries due to GW Emission

With GW emission included in the three-body equation of motion, a new outcome is possible: *dynamical inspirals*. Inspirals are characterized by the gravitational radiation driven inspiral of an IMS binary, while the third object is bound to the binary. Inspirals are particularly likely to occur during RIs. The magnitude of GW emission depends strongly on the distance of closest approach between two objects (e.g. 97). In relatively widely separated binaries, inspirals do not result from tightly bound circular orbits, but rather they are the product of orbits of very high eccentricity in which the objects experience close pericenter passages that generate significant GW emission and thus substantially reduce their orbital energy and angular momentum. High eccentricity orbits are most readily achieved in the chaotic environment of RIs, where despite the $e = 0$ initial conditions we impose on the binaries, the angular momenta of the three bodies is randomized and approaches an isotropic distribution with increasing number of passages.

Figure 5.3 shows an inspiral from one of our simulations. The binary-single interaction happens at the left of the plot and then propagates towards the right, terminating with the inspiral. One important feature of this interaction is that the bulk of the energy losses occur in three body knots, where the relative orbital angular momenta of the bodies is randomized, and the objects undergo very close pericenter passages, which in turn give rise to the spikes seen in the energy loss rate. Inspirals are of particular interest, as we will show in this paper, because they occur more frequently in widely separated target binaries, and they give rise to eccentric compact object mergers.

5.3.4 NUMERICAL APPROACH

Here we study the outcomes of binary-single interactions and their associated cross sections by performing large sets of numerical scattering experiments. To this end, we have developed a new N-body code to integrate the equation of motion of the three bodies using a fourth order Hermite integration scheme. The equation of motion including the effect from GW emission is discussed in Section 5.5.1. For a full description of the code and the exact state classification criteria employed, the reader is referred to the Appendix.[3] For each scattering experiment the target binary was randomly orientated in phase and orbital plane orientation.

We estimate the cross section numerically for a given outcome type O_i by performing $N_{\rm tot}$

---

[3] In the Appendix we also directly test the code against the Peters (97) analytic solution for binaries inspiraling due to GW emission.

binary-single interactions with isotropic sampling across a disc at infinity with radius b. If the total number of outcomes of type O_i from that scattering set is denoted by N_i, then the corresponding cross section for outcome O_i can be estimated by

$$\sigma_i = \frac{N_i}{N_{\text{tot}}}\pi b^2 \tag{5.8}$$

with a corresponding error given by

$$\Delta\sigma_i = \frac{\sqrt{N_i}}{N_{\text{tot}}}\pi b^2. \tag{5.9}$$

This, in turn, implies a rate of a given outcome O_i,

$$\Gamma_i \simeq n\sigma_i v_\infty \tag{5.10}$$

expected from a distribution of single objects with number density n and typical relative velocity v_∞. Thus, the rate of outcomes of type O_i compared to the rate of CIs is defined by the ratio of their cross sections, $\Gamma_i/\Gamma_{\text{CI}} = \sigma_i/\sigma_{\text{CI}}$.

5.4 NEWTONIAN POINT-PARTICLE LIMIT

To build intuition and to provide a direct link to previous studies in Newtonian gravity, we will first describe the most salient features of binary-single encounters of point masses in Newtonian gravity. These interactions and their final states, or outcomes, are well studied numerically and theoretically, especially in the pioneering series of work by Hut & Bahcall (55); Hut (53, 54); Heggie & Hut (44); Goodman & Hut (36); McMillan & Hut (82); Heggie et al. (46). More recent work by Fregeau et al. (29) and Fregeau & Rasio (33) have extended such studies to calculate the probability for collisions, and the coevolution of binaries and their host clusters.

When the three objects are equal point-masses, the outcome of an interaction will always be either a *fly-by*, an *exchange* or an *ionization*. These outcomes were described in Section 5.3.3. In this Section, we calculate their associated cross section over a broad range of encounter velocities v_∞/v_c using a series of numerical scattering experiments. In our equal mass case,

$$\frac{v_\infty}{v_c} = v_\infty\sqrt{\frac{2a_0}{3m}}, \tag{5.11}$$

thus any defining characteristics of the system can be rescaled using this ratio. We perform a total of 8×10^5 binary-single scatterings divided into 40 sets each with 2×10^4 interactions. For each scattering experiment, the target binary is randomly orientated in phase and orbital plane. The velocities of the encounters for the 40 sets are equally spaced in $\log(v_\infty/v_c)$ from 0.01 to 8. The maximum impact parameter, b_{max}, is kept fixed for all scatterings at $5a_0$. In this setup, outcomes from all the three interaction channels WP, SP and CI will occur depending on v_∞/v_c. Our numerical approach is closely related to the one used in (55). We also refer the reader to the Appendix for further details on our numerical approach.

Figure 5.4 shows the results from our scattering experiments. Both panels show the cross sections for exchange, fly-by, and ionization outcomes as a function of v_∞/v_c. The upper panel includes outcomes from all interactions including DIs and RIs, while the lower panel shows the outcomes coming from the RIs only. In what follows, we detail the outcomes and their dependence with v_∞/v_c.

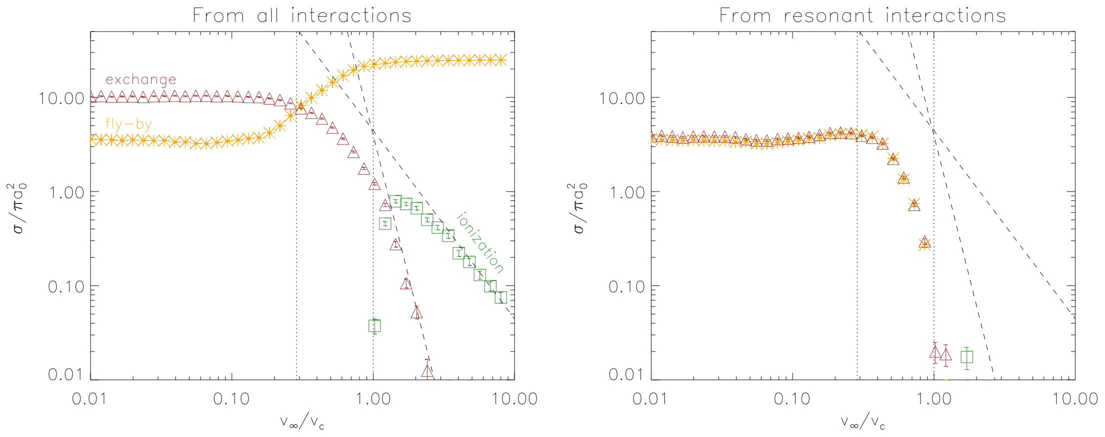


Figure 5.4 Integrated cross sections for the classical outcomes: *exchange* (brown triangles), *ionization* (green squares) and *fly-by* (orange stars) as a function of v_∞/v_c, where v_∞ is the relative velocity of the incoming object at infinity and v_c the characteristic velocity given by equation (5.5). The dashed lines show analytical approximations to the *exchange* (equation 5.13) and *ionization* (equation 5.14) cross sections. The vertical dotted lines indicate two characteristic velocities, the gravitational focusing velocity $v_\mathrm{foc}/v_\mathrm{c}(b_\mathrm{max} = 5a_0) \approx 0.28$ and the velocity that divides the system into having total positive or negative energy, $v_\infty = v_\mathrm{c}$. *Top*: Cross sections calculated from all interactions including RIs and DIs. *Bottom:* Cross sections only including endstates coming from RI encounters. This channel erases any information about initial conditions and all the three objects have thus equal probability to be kicked out. As a result, the *fly-by* and *exchange* cross sections are identical. Because a *fly-by* can not result from a DI, the *exchange* and *fly-by* cross sections are separated in the *top* panel. As can be clearly seen, the cross section for a RI is independent of v_∞ as long as $v_\infty < v_\mathrm{foc}$. Each plot is based on a total of 8×10^5 scatterings.

5.4.1 Low Velocity ($v_\infty/v_\mathrm{c} \ll 1$)

At low velocities, gravitational focus leads to all interactions happening via the CI channel. Therefore, all final state outcomes will be a result from either the DI or the RI channel. Since the total energy of the three-body system is initially negative and no bound triple state can form a stable final state (54), the only possible outcome is a binary (carrying the negative energy part in form of binding energy) and a single unbound object. Depending on which two objects that form the binary the outcome will either be labeled as an exchange or a fly-by.

Within the CI channel the probability for a given outcome depends on whether the binary has experienced a RI or a DI. If the outcome is a result of the RI channel, then any permutation of the three objects in the final state is equally likely since the RI erases any memory of the binary's initial configuration. As a result, the exchange and fly-by outcomes have the same cross section when the system has evolved through a RI. This can be seen in the lower panel of Figure 5.4.

For interactions passing through the DI channel, fly-bys have a negligible probability to occur. The reason is that a DI is characterized by having only a single interaction that in the majority of cases leads to an exchange between the incoming object and one of the binary members. A typical fly-by involves at least two closest IMS pairs leading these interactions to be classified as arising from the RI channel. This leads to the cross section difference between exchange and fly-by when all interactions are included as seen in the upper panel of Figure 5.4.

The critical velocity that defines the transition to all interactions happening through the CI channel, v_{foc}, is found from equation (5.3),

$$\frac{v_{\mathrm{foc}}}{v_{\mathrm{c}}} = \sqrt{2}\left(\frac{a_0}{b_{\mathrm{max}}}\right), \tag{5.12}$$

which, in our numerical setup with $b_{\mathrm{max}} = 5a_0$, gives $v_{\mathrm{foc}}/v_{\mathrm{c}} = 0.28$. This critical velocity transition is illustrated with a vertical dotted line in Figure 5.4. It is clear in the lower panel in Figure 5.4 that this line accurately separates the plot into two regimes. The cross sections are approximately flat to the left of this line, when $v_\infty < v_{\mathrm{c}}$. This tells us that the relative numbers of RIs and DIs are nearly constant, and as a result, independent of the exact impact parameter and encounter velocity as long as the interaction is a CI.

5.4.2 INTERMEDIATE VELOCITY ($v_\infty/v_{\mathrm{c}} \approx 1$)

At intermediate velocities, the resultant encounters are a mixture of CIs, SPs and WPs, and the velocity dependence shapes the resultant cross-sections. CIs can still occur at intermediate velocities, but their probability decreases as $\sigma_{\mathrm{CI}} \propto (v_\infty/v_{\mathrm{c}})^{-2}$, as given by equation (5.3). This scaling solely determines the shape of the exchange cross section in this regime, since exchanges only can happen via a CI. This is seen in Figure 5.4 where the exchange cross section is observed to clearly transition from being flat at low velocities to decreasing as $\propto (v_\infty/v_{\mathrm{c}})^{-2}$ at intermediate velocities.

WPs and SPs happen with increasing frequency as the velocity is increased since more encounters pass by the binary instead of making a CI. These perturbative encounters necessarily result in a fly-by classification since the encounter never comes close enough to make an exchange, thus leading to a velocity dependent increase in the associated cross section.

5.4.3 HIGH VELOCITY ($v_\infty/v_{\mathrm{c}} > 1$)

In high velocity interactions ($v_\infty > v_{\mathrm{c}}$), the total energy of the three-body system is positive and ionization becomes a possible outcome. Ionization occurs when all three objects are unbound with respect to each other. This outcome dominates over the exchange outcome in this high velocity regime as seen in the upper panel in Figure 5.4.

Because of the high velocity, CIs are rare. The CI cross section is determined by the geometrical term in equation (5.3). Since the geometrical term only depends on the size of the target binary, the occurrence of a CI is independent of velocity. By contrast to the intermediate velocity range, the observed steep decrease in both the exchange and ionization cross sections as the velocity increases is a result of properties of the interactions themselves, rather than a varying number of CIs.

As observed in the lower panel of Figure 5.4, RIs do not occur at high velocity. All outcomes from the CI channel are, therefore, only arising from the DI channel. The main reason for this is that the incoming object enters the binary with such a high velocity that the pair appears to be approximately stationary. The majority of interactions between the single and the binary will therefore be a DI between the incoming single and its nearest binary object. The problem therefore reduces to a two-body interaction between the encounter and one of the binary members. This setup has an analytical solution and cross sections for exchange and ionization can be analytically estimated in this so called impulsive regime. This was first done by (53) who calculated in this high velocity regime the exchange cross section

$$\sigma_{\mathrm{ex}} = \frac{320}{81}\frac{\pi a_0^2}{v_\infty^6}, \tag{5.13}$$

and the ionization cross section

$$\sigma_{\mathrm{ion}} = \frac{40}{9}\frac{\pi a_0^2}{v_\infty^2}. \tag{5.14}$$

These scalings are also shown in Figure 5.4. The similarity of this three-body scattering problem to atomic physics can be seen by comparing the exchange scenario, in the limit where one of the binary members are very light, with electron capture (or charge transfer) in heavy nucleus interactions (113).

5.5 GRAVITATION WAVE LOSSES AND THREE BODY DYNAMICS

In this section, we describe how general relativity (GR) corrections are included into the equation of motion in our three-body integration code, and highlight the dynamical consequences of these loss terms.

5.5.1 ADDING GENERAL RELATIVISTIC CORRECTIONS

In this work, we include the energy and angular momentum losses by GW radiation using the PN formalism (10). In this formalism, the acceleration experienced by an object of mass m_1 due to the gravitational force from a second object of mass m_2 is expanded in series as

$$\mathbf{a} = \mathbf{a}_0 + c^{-2}\mathbf{a}_2 + c^{-4}\mathbf{a}_4 + c^{-5}\mathbf{a}_5 + \mathcal{O}(c^{-6}). \tag{5.15}$$

The standard Newtonian force per unit mass, $\mathbf{a}_0$ is

$$\mathbf{a}_0 = -\frac{Gm_2}{r_{12}^2}\hat{\mathbf{r}}_{12}, \tag{5.16}$$

where the separation vector is $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$, its magnitude is $r_{12} = |\mathbf{r}_{12}|$, and its direction is $\hat{\mathbf{r}}_{12} = \mathbf{r}_{12}/r_{12}$. The terms $\mathbf{a}_2$ and $\mathbf{a}_4$ account for the periastron shift. The leading order term that represents the radiation of energy and momentum from the system, $\mathbf{a}_5$, is also known as the 2.5PN term. This term takes the following form

$$\begin{aligned}\mathbf{a}_5 =& \frac{4}{5}\frac{G^2 m_1 m_2}{r_{12}^3}\left[\left(\frac{2Gm_1}{r_{12}} - \frac{8Gm_2}{r_{12}} - v_{12}^2\right)\mathbf{v}_{12}\right.\\ &\left.+ (\hat{\mathbf{r}}_{12}\cdot\mathbf{v}_{12})\left(\frac{52Gm_2}{3r_{12}} - \frac{6Gm_1}{r_{12}} + 3v_{12}^2\right)\hat{\mathbf{r}}_{12}\right],\end{aligned} \tag{5.17}$$

where the relative velocity scalar, v_{12}, and vector, $\mathbf{v}_{12}$, are defined following the same conventions as in Blanchet (10). We use the modified acceleration $\mathbf{a} = \mathbf{a}_0 + c^{-5}\mathbf{a}_5$ in our numerical treatment instead of the Newtonian $\mathbf{a}_0$. A fundamental difference between the purely Newtonian acceleration and the 2.5PN acceleration is that $\mathbf{a}_5$ depends not only on the separation between the objects but also on their relative velocity.

The energy and angular momentum losses through the 2.5PN term should coincide with those calculated using the quadripolar formalism for two bodies. To this end, the orbit-averaged equations for the time dependent evolution of SMA, a, and eccentricity, e, of a two-body system emitting GWs derived by Peters (97) have provided a useful test framework to many authors,

$$\frac{da}{dt} = -\frac{64}{5}\frac{G^3 m_1 m_2 (m_1+m_2)}{c^5 a^3 (1-e^2)^{7/2}}\left(1+\frac{74}{24}e^2+\frac{37}{96}e^4\right), \tag{5.18}$$

and

$$\frac{da}{de} = \frac{12}{19}\frac{a}{e}\frac{\left[1+(73/24e^2)+(37/96)e^4\right]}{(1-e^2)\left[1+(121/304)e^2\right]}. \tag{5.19}$$

By including the comparable 2.5PN terms directly in our three-body integration of the equation of motion we can capture losses in three-body interaction knots as well as reproduce equations (5.18) and (5.19) in the case where the system develops strong hierarchy and two bodies evolve following the secular evolution described by Peters (97). In the Appendix, we show comparisons between the orbit-averaged equations (5.18) and (5.19) and a direct numerical integration in our code.

With the inclusion of losses to GW radiation, binaries have a finite lifetime. If, for example, we consider a binary with objects of equal mass, m, and a circular orbit with initial SMA a_0, equation (5.18) reduces to the form $da/dt \propto (m/a)^3$ with the solution

$$t_{\mathrm{life}}(a_0) = 1.6\times 10^{17}\left(\frac{a_0}{\mathrm{au}}\right)^4\left(\frac{m}{M_\odot}\right)^{-3}\ \mathrm{yr}. \tag{5.20}$$

Here t_{life} is the GW inspiral time, or the time it takes for the initial binary to evolve from $a = a_0$ to $a = 0$. The dependence on the SMA to the fourth power makes the lifetime very sensitive to small changes in a_0. In the other limit, where the initial eccentricity e_0 is not far from unity, the inspiral time is

$$t_{\mathrm{life}}(a_0, e_0) \simeq t_{\mathrm{life}}(a_0)\frac{768}{425}\left(1-e_0^2\right)^{7/2}. \tag{5.21}$$

The lifetime of a very eccentric binary is shorter than that of a binary in a circular orbit with similar SMA because as the eccentricity increases the pericenter distance, which is given by $r_{\mathrm{min}} = (1-e)a$, decreases. This results in a higher GW flux every pericenter passage, which in turn decreases the lifetime and gradually circularizes the orbit of the binary.

An analytical solution for the coupled evolution in a and e also exists (97)

$$a(e) = \frac{c_0 e^{12/19}}{1-e^2}\left(1+\frac{121}{304}e^2\right)^{870/2299}, \tag{5.22}$$

where c_0 is a constant with dimensions of length, set according to the initial conditions (a, e) of the binary system. From this expression we see that in the high eccentricity limit, where $e \approx 1$, the SMA scales as $a(e) \propto (1-e)^{-1}$. As a result, the orbital SMA (and thus also the orbital energy) must change by many orders of magnitude before the eccentricity becomes significantly less than unity. Inspiraling binaries thus only become approximately circular during the last phases of their inspiral.

5.5.2 SIGNIFICANCE OF PN CORRECTIONS

The binary's compactness determines many of the important dynamical properties of the system, especially the importance of PN corrections and collisions. A dimensionless compactness can be defined as (10)

$$\gamma = \frac{Gm}{rc^2}. \tag{5.23}$$

Using γ, we can write the acceleration, $\mathbf{a} = \mathbf{a}_0 + c^{-5}\mathbf{a}_5$, in terms of the dimensionless radius and mass, $\tilde{r} = r/r_\mathrm{u}$ and $\tilde{m} = m/m_\mathrm{u}$. In these units, the acceleration is $\tilde{\mathbf{a}} = \mathbf{a}/(Gm_\mathrm{u}/r_\mathrm{u}^2)$ and we have

$$\tilde{\mathbf{a}}_\mathrm{tot} = \tilde{\mathbf{a}}_0(\tilde{m}, \tilde{r}) + \gamma^{5/2}\tilde{\mathbf{a}}_5(\tilde{m}, \tilde{r}, \tilde{v}). \tag{5.24}$$

For systems that are strongly relativistic, the SMA $a_0 \approx Gm/c^2$ and, as a result, PN corrections become very important. For weakly PN systems, $a_0 \gg Gm/c^2$ and the compactness of the orbit provides an estimate for the importance of the PN corrections to the equation of motion of a circular, $e \approx 0$, orbit. However, a key point that we emphasize in this work is that measuring the strength of the PN corrections only in terms of the compactness of the initial binary orbit can be misleading. In chaotic three-body interactions, the eccentric orbits and close passages that arise make it possible for strong PN corrections to be realized even in systems with initially wide SMA. As we will discuss later, the initial compactness of the binary system still determines the probability that a very strong encounter will occur.

Close approaches in eccentric orbits lead to strong PN corrections to the equation of motion. They also may lead to direct collisions. The maximal strength of PN corrections to the acceleration is therefore set by the physical size and mass of the objects, rather than by the initial SMA of their orbits. This can be quantified by calculating the compactness γ for the interacting objects themselves using their mass and radius. For example, if the objects are black holes, their compactness $\gamma \sim 1$, and PN corrections can therefore reach their maximal strength. If the constituent objects are not black holes, then $\gamma < 1$, and the magnitude of the maximal PN corrections for that three-body system is reduced. Neutron stars have typical dimensionless compactness of $\gamma \approx 0.2$, while a $0.6\ M_\odot$ white dwarf is characterized by a $\gamma \approx 10^{-4}$. Interacting WDs will therefore in general collide before PN corrections become strong.

If a system of N interacting objects is only composed of BHs, then the dynamics of the system becomes scale free (e.g. 114; 40). The reason is that the equation of motion scales with the masses of the BHs, as do the BH gravitational radii. For example, for a binary-single interaction involving three equal mass BHs, the expected dynamics for a system with $a_0 = 10^{-3}$ AU and $m_\mathrm{BH} = 1\ M_\odot$ will be equivalent to that of a system with $a_0 = 10^{-1}$ AU and $m_\mathrm{BH} = 10^2\ M_\odot$. This allows us to identify dynamically similar systems that occur in different astrophysical contexts. If the N interacting objects are not BHs, then the system looses its scale-free behavior as the object radius no longer scales with mass. Neutron stars, for example, exhibit relatively constant radius across their observed mass range (119), while white dwarfs have an inverse mass radius relationship $R_\mathrm{WD} \propto m_\mathrm{WD}^{-1/3}$.

5.5.3 ENERGY LOSSES

The effects of GW energy loss can be most easily seen by examining equation (5.17) in the context of a circular binary of equal mass objects. In that case, $\hat{\mathbf{r}} \cdot \mathbf{v}_{12} = 0$, leaving only the first term in equation (5.17). For equal mass objects, the term in parenthesis in equation (5.17) evaluates to a negative number and the direction of $\mathbf{a}_5$ is determined by $-\mathbf{v}_{12}$, directly against the motion of the two bodies. As a result, the orbiting objects essentially experience a drag force

$$F_\mathrm{2.5PN} = \frac{32\sqrt{2}}{5}\frac{G^{7/2}}{c^5}\left(\frac{m}{r}\right)^{9/2} \tag{5.25}$$

This follows directly from equation (5.17) by substituting $v = \sqrt{2Gm/r}$. The energy leaving the system per unit time can be easily calculated by using ΔE_orb = force×distance = $F_\mathrm{2.5PN} 2\pi r$,

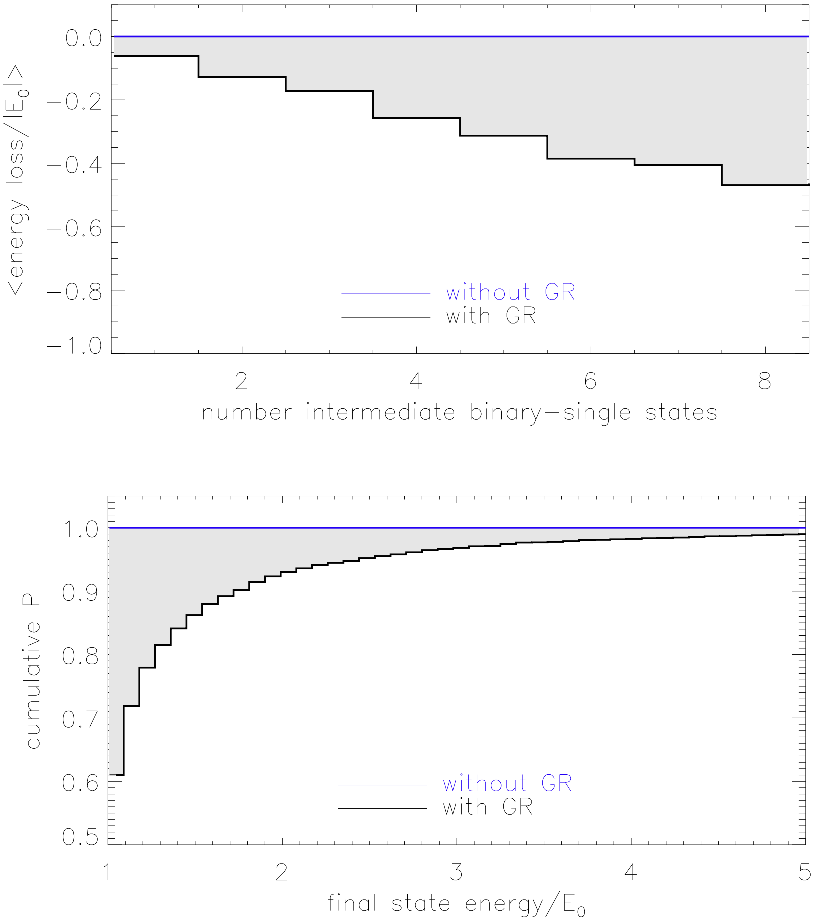


Figure 5.5 GW energy loss in binary-single interactions between equal mass BHs. The panels show an extreme HB case with $a_0 = 10^{-5}$ AU and $m_{\rm BH} = 1\ M_{\odot}$. The black lines indicate scatterings where GR is included in the simulation and the blue lines indicate those for which GR is not included. Softer binaries will have energy losses within the grey shaded region and quickly end up near the blue line, which indicates no energy loss. Both panels include only states from the RI channel with a finale state where the single object is unbound. *Top:* Average energy change scaled by the initial energy E_0 after a certain number of intermediate binary-single states. Fractional energy losses of the order of $\sim 10\%$ can be achieved just after the second instance a new binary-single state is produced. The average energy loss increases with the number of identified IMS, indicating that energy has being extracted from the system. *Bottom:* Cumulative distribution for the fractional energy difference between the total final state energy and the initial energy. The figures are based on 2×10^4 binary-single interactions with $v_{\infty} \ll v_{\rm c}$.

from which it follows that

$$\frac{dE}{dt} \simeq \frac{\Delta E_{\rm orb}}{T_{\rm orb}} = -\frac{64}{5}\frac{G^4}{c^5}\left(\frac{m}{r}\right)^5 \tag{5.26}$$

where $T_{\rm orb} = 2\pi(2Gm/r^3)^{-1/2}$ is the orbital period. One should notice that the distance r is changing as a function of time with a rate that can be calculated by using the Newtonian

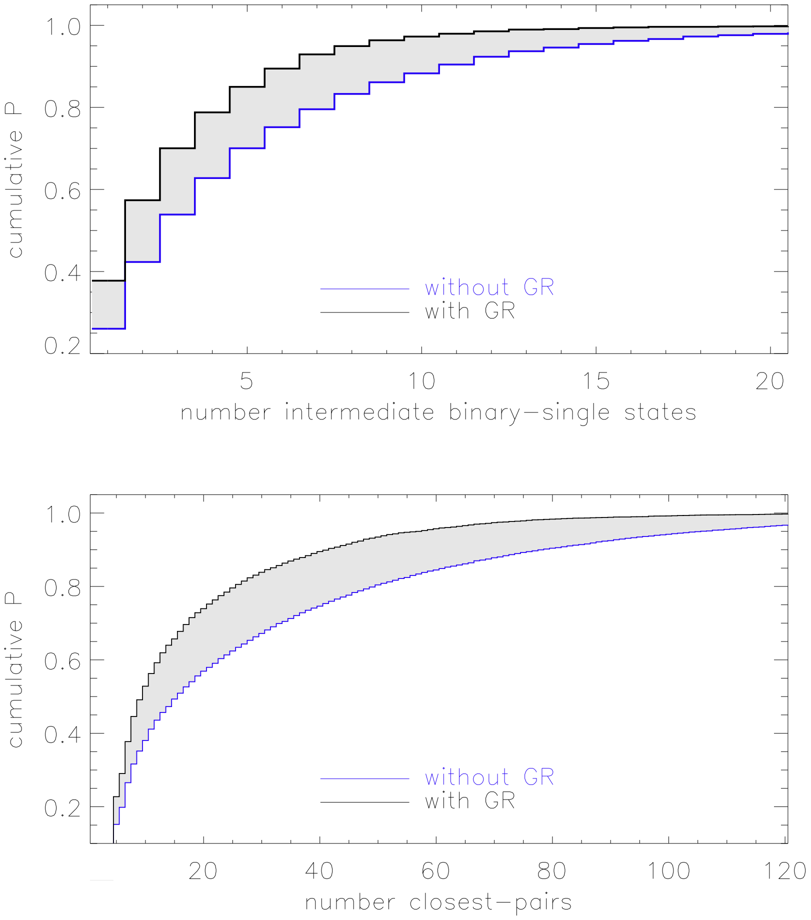


Figure 5.6 Number of three-body interactions between equal mass BHs arising from binary-single scatterings. Both panels include only states from the RI channel. The target binary is chosen to be initially very hard with $a_0 = 10^{-5}$ AU and $m_{\rm BH} = 1\ M_\odot$. The black lines indicate scatterings where GR corrections have been added while blue lines show experiments with no GR corrections included. The two plots differ in the way the number of interactions are counted. *Top:* Number of times an intermediate binary-single state (IMS) is observed to occur during a resonant interaction. *Bottom:* Number of times a new closest pair has been identified during the resonant interaction. A high number of close-pairs indicates highly chaotic motion during the encounter (see Figure 5.1) which occurs between each IMS.

relation $dE/dr = -Gm^2/2r^2$.

The above formalism can be extended to a binary-single interaction. The distribution of GW energy radiated during a resonant encounter is shown in Figure 5.5. The upper panel shows how energy from the system is depleted as new intermediate binary-single state are created. The fractional energy loss is relatively small, especially for binaries with large SMA, but at each encounter the binaries are effectively hardened and the relative likelihood for the system to undergo a collision or a merger is increased. The lower panel shows the cumulative distribution of the fractional energy loss between the initial state and the final state for the same

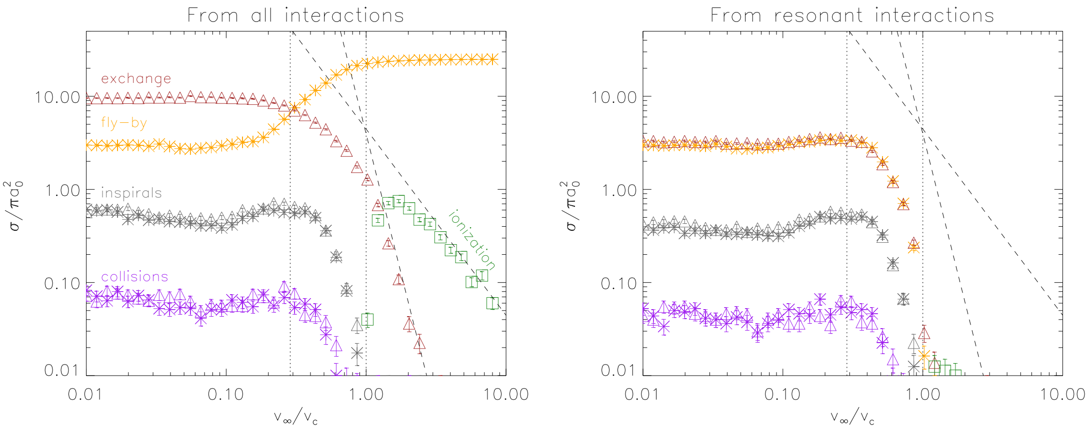


Figure 5.7 Integrated outcome cross sections from binary-single interactions between equal mass BHs including 2.5PN corrections. Similar to Figure 5.4 but now including collisions (purple) and *inspirals* (grey). The initial SMA of the target binary is $a_0 = 10^{-4}$ AU and $m_{\rm BH} = 1\ M_{\odot}$. *Top:* All interactions including RIs and DIs. *Bottom:* Outcomes arising from the RI channel only. The new outcome from including GR is a population of objects that gravitationally inspiral during the interaction. We denote such endstates as *inspirals*. The *inspiral* cross section flattens out below the gravitational focusing velocity, $v_{\rm foc}$, which implies that these end states are not sensitive to the exact value of the impact parameter, b, and velocity, v_∞, as long as the interaction is a CI.

set of interactions. Figure 5.6 shows the corresponding cumulative distributions of the number of IMS (*top panel*) and the number close-pairs (*bottom panel*) in a binary-single interaction. The number of close-pairs is greater than the number of IMS since it also includes all close passings that can occur within a single state (see Figure 5.1). For the set of scatterings ending with an unbound companion (exchange or fly-by) the number of three-body interactions are reduced when GR is included. For example, without GR 20% of all scatterings shown in Figure 5.6 have more than 50 close interactions, but only about 25 when GR is included. The reason is simply that the possibility of the system inspiraling when GR is included, truncates the chain of resonance interactions.

5.6 THE FORMATION OF DYNAMICAL INSPIRALS

With the inclusion of energy and angular momentum losses from GW emission a new class of dynamical outcomes appears, which we denote here as *inspirals*. These are interactions in which two of the objects inspiral and merge while all three objects are still in a bound three body state; that is, before one of the classical outcomes of exchange, flyby or ionization is achieved. An example of an inspiral end state is shown in Figure 5.3.

In order to understand how the inclusion of GR corrections changes the binary-single outcome landscape, we recompute the Newtonian scattering experiments shown in Figure 5.4 with the addition of the 2.5PN term in the equation of motion. Our results are illustrated in Figure 5.7. The revised cross sections include inspirals and collisions between solar mass black holes with an initial binary SMA of 10^{-4} AU. The top panel shows the resultant cross sections from all interaction channels including DIs and RIs while the bottom panel includes only endstates arising from the RI channel. By comparing the two panels one can conclude that inspirals (and collisions) are dominated by the RI channel, an observation that will become useful when we derive the analytical treatment for inspiral occurrence in Section 5.6.2.

Another important point is that the cross section for inspirals is approximately flat when $v_{\infty} < v_{\text{foc}}$. This implies that the probability for an inspiral to occur is not sensitive to the exact value of the impact parameter, b, or velocity, v_{∞}, as long as the single object experiences a CI with the binary. The lack of a dependence on the initial conditions arises because nearly all inspirals are generated from RIs (for which memory of the initial conditions is rapidly lost through ensuing resonances) and because the fraction of RIs and DIs is approximately constant for $v_{\infty} \ll v_{\text{c}}$ (see Section 5.4.1). This observation makes it possible to write the probability for an outcome to be an inspiral given the interaction is a CI as

$$P_{\text{insp}} \equiv N_{\text{insp}}/N_{\text{CI}}, \tag{5.27}$$

and the corresponding inspiral cross section as

$$\begin{aligned} \sigma_{\text{insp}} &= P_{\text{insp}}\sigma_{\text{CI}}, \\ &\simeq P_{\text{insp}}\frac{3\pi G m a_0}{v_{\infty}^2}, \end{aligned} \tag{5.28}$$

where the last equality holds for the equal mass case. This factorization is useful in the sense that it separates the contribution coming from the chaotic RIs from the standard focusing cross section that simply acts as a weight factor. It is important to notice that P_{insp} depends on the compactness of the initial binary, i.e. its SMA a_0 and mass m_{bin}, as we will show in Section 5.6.2.

5.6.1 PHASE SPACE DISTRIBUTION OF INSPIRALS

Figure 5.8 shows distributions of the orbital parameters (a, e) for all exchange and fly-by binaries (*orange*) and intermediate state binaries (*blue*) from 2×10^4 HB binary-single interactions. The division at a/a_0 indicates energy conservation between the newly formed binary with SMA a and initial binary with SMA a_0. The target binary must shrink if the single object becomes unbound, i.e. exchange and fly-by binaries have $a < a_0$ while IMS binaries have $a > a_0$.

Inspirals appear in grey in the *right-hand* panel in Figure 5.8 where the 2.5PN term is included in the equation of motion. These inspirals form from the subset of IMS binaries that merge while the three-body system is still bound and are therefore (mainly) initially created with $a > a_0$. Since GWs in general carry energy out of the system before an endstate is reached, then IMS can flow across the initial $a/a_0 = 1$ border line. This means that all outcome distributions are slightly changed when GR is included. Inspiral states are, however, those that experience the highest energy losses.

Immediately after an inspiralling binary has formed, it evolves according to equation (5.22). Several of these evolutionary trajectories are shown with thin black lines in Figure 5.8. GW emission circularizes the binary as its SMA is decreased. This migrates binaries from their initial formation region in the right hand side of the (a, e) phase space to the lower left. Therefore, the exact location of the inspiral event in Figure 5.8 depends on when the system was identified in the code (see Appendix A for a discussion of the selection criteria for states). It is therefore not necessarily representative of the binary's initially assembled position in the formation locus for inspirals.

The phase space accessible for inspirals depends on γ (equation 5.23). At particular (a, e) combinations with close pericenter approaches, direct collisions can also occur. A direct tradeoff

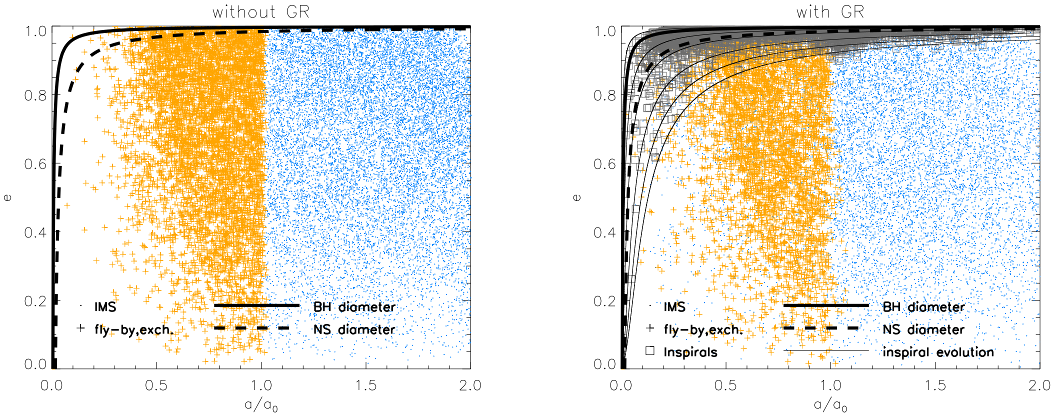


Figure 5.8 Distribution of orbital parameters (a, e) for all identified binaries from 2×10^4 binary-single interactions between equal mass BHs. The target binaries have a SMA of $a_0 = 10^{-5}$AU and $m_{\mathrm{BH}} = 1\ M_{\odot}$. The encounters occur with an incoming velocity of 10 km s^{-1} and as such are in the extreme HB limit where $v_\infty \ll v_{\mathrm{c}}$. *Left:* Without GR. *Right:* With GR. The blue points represent IMS binaries, which are the candidates for *inspiral* end states. The sampling of these IMS is nearly homogenous. The orange symbols show the endstate binaries from the classical outcomes: *exchanges* and *fly-bys*. *Inspirals* that arise when GR is included are seen in the right panel as *grey squares*. Each identified binary separation, a, is scaled with the initial a_0. This is because if there is no energy loss and $v_\infty \ll v_{\mathrm{c}}$, then all intermediate states must have $a/a_0 > 1$ and all final states with an unbound companion should satisfy $a/a_0 < 1$. This follows directly from conservation of energy ($a \propto 1/E$) and helps illustrate how binaries tend to harden after a HB interaction. If energy is leaking out of the system by GW emission, then the resultant states shown by the *blue-points* can flow across the $a/a_0 = 1$ boundary as seen on the *right* panel. The orbital parameters for *inspirals* are fast evolving and the grey region is therefore only showing a snapshot of the phase space distribution of these states. The thin black lines show a few examples of the evolution contours the *inspirals* follow in the (a, e) space. The black solid line shows the diameter of a $1\ M_{\odot}$ BH, and the dashed shows the diameter of a NS with 12 km radius. In this example, where the interacting objects are three stellar mass BHs, any formed binary above the BH limit would lead to a collision. Many of the BH *inspirals* would have been collisions instead, if the objects would have been NSs.

can the be found between the number of collisions and the number of inspirals. The rates for these particular end states cannot be independent because they originate from a similar phase space region. Not surprisingly, extended objects produce relatively fewer inspirals and more collisions than compact ones. The importance on th object's size is illustrated in Figure 5.8, in which we plot the boundaries defined by the BH and NS diameters, respectively.

5.6.2 ANALYTIC DERIVATION OF INSPIRAL CROSS SECTIONS

In this section, we develop an analytical understanding of what determines the occurrence rate of inspirals and collisions, including how the outcomes depend on the initial SMA and on the mass of the target binary. Each IMS is characterized by three parameters: the SMA (a) and eccentricity (e) of the IMS binary and the orbital period of the bound companion, which we denote here as the isolation time (t_{iso}). Since the single object is bound to the binary during

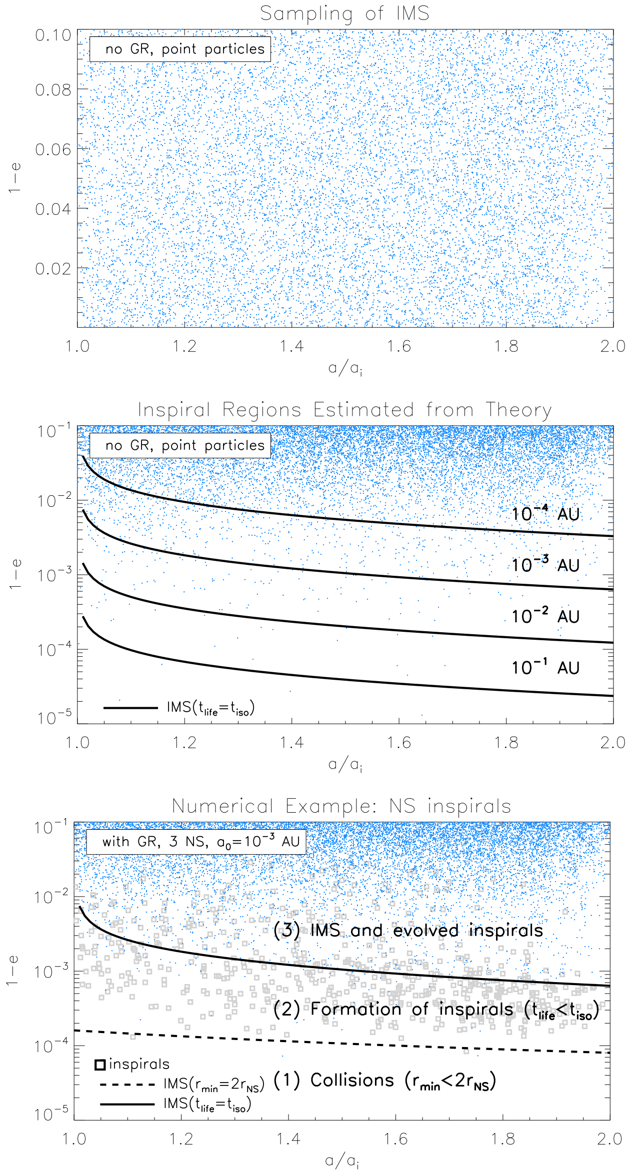


Figure 5.9 Inspirals in $(a, 1-e)$ space produced by IMS binaries. Blue points indicate IMSs whose endstate is not an inspiral. Grey squares show inspirals at the time of identification. The *x*-axis on all three panels show $a/a_i = E_{\mathrm{tot},i}/E_{\mathrm{bin},i}$ where $E_{\mathrm{tot},i}$ and $E_{\mathrm{bin},i}$ are respectively the total energy of the three-body system and the energy of the IMS binary at the time of identification: t_i. *Top:* Sampling of IMS near $e \approx 1$. This distribution is relatively uniform in $(a, 1-e)$ space, an observation that makes it possible to estimate how the number of inspirals produced scales with initial SMA a_0 (equation 5.36). *Middle:* Same distribution as in the top panel but with the y-axis in logarithmic scale. To illustrate which IMS can form inspirals when GR is included, we have plotted using equation (5.34) a few lines showing where $t_{\mathrm{life}} = t_{\mathrm{iso}}$. An inspiral can form when $t_{\mathrm{life}} < t_{\mathrm{iso}}$. *Bottom:* Results from numerical scattering experiments including GR and a finite radius, r_{NS}, for the interacting objects. The NS radius introduces a collision boundary given by $r_{\mathrm{min}} = 2r_{\mathrm{NS}}$ where r_{min} is the the pericenter distance of the IMS binary. As shown in the plot, the $(a, 1-e)$ IMS space divides into three distinct regions: (1) IMS with $r_{\mathrm{min}} < 2r_{\mathrm{NS}}$ will produce direct collisions, (2) IMS with $r_{\mathrm{min}} > 2r_{\mathrm{NS}}$ and $t_{\mathrm{life}} < t_{\mathrm{iso}}$ will form inspirals, and (3) IMS with $r_{\mathrm{min}} > 2r_{\mathrm{NS}}$ and $t_{\mathrm{life}} > t_{\mathrm{iso}}$ can be followed by further interactions. As inspirals formed in region (2) spiral in, they diffuse into region (3). All three panels are based on 2×10^4 scatterings between equal $1.4\ M_{\odot}$ objects with $r_{\mathrm{NS}} = 12$ km.

an IMS, $t_{\rm iso}$ is finite. It then follows that if an IMS binary is formed with $t_{\rm life} < t_{\rm iso}$, then the binary will inspiral before the return of the bound companion. The lifetime, $t_{\rm life}$, is determined by equations (5.18) and (5.19) but can be estimated by equations (5.20) and (5.21) in the circular and eccentric limits, respectively. In all of the following calculations, we assume the hard binary limit ($v_\infty \ll v_c$).

The probability for a particular outcome to be an inspiral can be estimated by considering the fraction of states during a RI that satisfies $t_{\rm life}(a,e) < t_{\rm iso}(a)$. The isolation time $t_{\rm iso}$ is described by Keplers law

$$t_{\rm iso} = 2\pi\sqrt{\frac{a_{\rm bs}^3}{Gm_{\rm tot}}}, \tag{5.29}$$

where $a_{\rm bs}$ is the SMA of the hierarchical triple. This SMA, $a_{\rm bs}$, can be expressed in terms of the initial binary SMA, a_0, and the SMA of the IMS binary, a, by making use of energy conservation

$$E_{\rm tot} \simeq -\frac{Gm_1m_2}{2a_0} = E_{\rm bin} + E_{\rm bs} = -\frac{Gm_im_j}{2a} - \frac{Gm_{\rm bin}m_{\rm sin}}{2a_{\rm bs}} \tag{5.30}$$

where 'bin' and 'sin' respectively refer to the binary and the single bound object in the hierarchical triple. In the equal mass case, equation (5.30) reduces to

$$a_{\rm bs} = \frac{2a_0}{1-1/a\prime} \tag{5.31}$$

such that

$$t_{\rm iso} = \left(\frac{2}{1-1/a\prime}\right)^{3/2} 2\pi\sqrt{\frac{a_0^3}{Gm_{\rm tot}}}, \tag{5.32}$$

where $a\prime = a/a_0$ and the last term in equation (5.32) is the orbital time of the initial binary system, $T_{\rm orb,0}$. Equation (5.32) relates the normalized SMA, $a\prime$, of a given IMS binary to the time it remains isolated from its bound companion. Since $a\prime > 1$ during a resonance, it follows that $t_{\rm iso} > T_{\rm orb,0}$.

We can now compare $t_{\rm iso}$ to $t_{\rm life}$, which, in the high eccentricity limit, is given by equation (5.21). The ratio $F_{\rm insp} = t_{\rm life}/t_{\rm iso}$ describes the lifetime relative to the binary isolation time and can be written as

$$F_{\rm insp} = \frac{C_{\rm F}c^5}{G^{5/2}}\left(\frac{a_0}{m}\right)^{5/2}(1-e^2)^{7/2}a\prime^{5/2}(a\prime-1)^{3/2} \tag{5.33}$$

where $C_{\rm F} = (3\sqrt{3})/(680\pi\sqrt{2}) \approx 1.7\times10^{-3}$. If $F_{\rm insp} < 1$, the binary will inspiral before the third body returns. If, on the other hand, $F_{\rm insp} > 1$ another three-body encounter will take place. The boundary defined by $F_{\rm insp} = 1$ produces a clear division in the $(a\prime, e)$ phase space plane, clearly separating IMSs that will inspiral to those that can be followed by further three-body interactions (Figure 5.9).

Defining the allowed phase space region for inspirals as $\Delta_{\rm insp} = 1-e$ and setting $F_{\rm insp} = 1$ in equation (5.33), we get

$$\Delta_{\rm insp} \approx \frac{1}{2}\frac{G^{5/7}}{C_{\rm F}^{2/7}c^{10/7}}\left(\frac{m}{a_0}\right)^{5/7} a\prime^{-5/7}(a\prime-1)^{-3/7}, \tag{5.34}$$

which implies $\Delta_{\rm insp} \propto (m/a_0)^{5/7}$. Assuming that the (a,e) sampling of IMSs is relatively uniform where $e \sim 1$, as observed in Figure 5.9, we conclude that the number of IMSs within the inspiral region is $\propto (m/a_0)^{5/7}$. This means that the probability for an outcome to be an inspiral given that the interaction is a CI (see 5.27) scales as

$$P_{\text{insp}} \propto \left(\frac{m}{a_0}\right)^{5/7} \propto \gamma^{5/7}, \tag{5.35}$$

such that

$$\sigma_{\text{insp}} = P_{\text{insp}}\sigma_{\text{CI}} \propto a_0^{2/7}\frac{m^{12/7}}{v_\infty^2}. \tag{5.36}$$

This illustrates that the cross section for inspirals is expected to *increase* with the SMA of the target binary. The dominant inspiral-producing targets in a cluster are thus not extremely compact binaries, but instead wide ones.

Collisions occupy a similar phase space region to that populated by inspirals, with the size of the interacting objects and the initial SMA of the target binary determining their relative cross sections. If an IMS binary is formed with a periapsis $r_{\text{min}} = a(1-e)$ that is smaller than twice the radius r_{obj} of the interacting objects, then a collision will occur. Using $\Delta_{\text{coll}} = 1 - e$, the collision boundary is simply given by

$$\Delta_{\text{coll}} = (2r_{\text{obj}}/a_0)(a\prime)^{-1}, \tag{5.37}$$

which leads to the result that the probability for a collision is $P_{\text{coll}} \propto a_0^{-1}$. The associated cross section, σ_{coll}, can be estimated using equation (5.28), and it is thus independent of a_0.

If we compare equations (5.34) and (5.37), we can see that the probability for a collision ($\propto a_0^{-1}$) decreases faster than the inspiral probability ($\propto a_0^{-5/7}$) as a_0 increases. This means that collisions will occupy a progressively smaller fraction of the available inspiral phase space as the SMA of the target binary increases. Inspirals arising from widely separated binaries are therefore less likely to be depleted by collisions, which in turn makes widely separated binaries even better targets for inspiral production.

5.6.3 NUMERICAL DETERMINATION OF THE CROSS SECTION

Figure 5.10 shows the formation probability and corresponding cross sections of inspirals and collisions as a function of initial SMA derived using numerical scattering experiments. The symbols show results from our numerical simulations while the dashed lines show the results from our analytical estimates giving by equation (5.36). As discussed in Section 5.6.2, the inspiral cross section increases with SMA. This is because the gravitational focusing cross section for a CI increases faster with SMA ($\propto a_0$) than the probability for an inspiral decreases ($\propto a_0^{-5/7}$).

As can be seen in Figure 5.10, the numerical and analytical scalings are in agreement in the asymptotic limit but show small differences in slope at low SMA. These differences are caused by having neglected a series of physical effects in the analytical scaling, such as collisions and GW energy losses before the interaction has reached its final endstate. However, these corrections are only important for target binaries in the high compactness limit. From an astrophysical perspective, these binaries are believed to be a negligible target population as these they are expected to merge before a CI can take place. The reader is refered to Section 5.8 for further discussion.

Since we have now shown that inspirals are a likely outcome even from widely separated binaries, it is important to compare them with mergers arising from the widely-discussed single-single GW capture scenario (41; 120; 67; 23; 22).

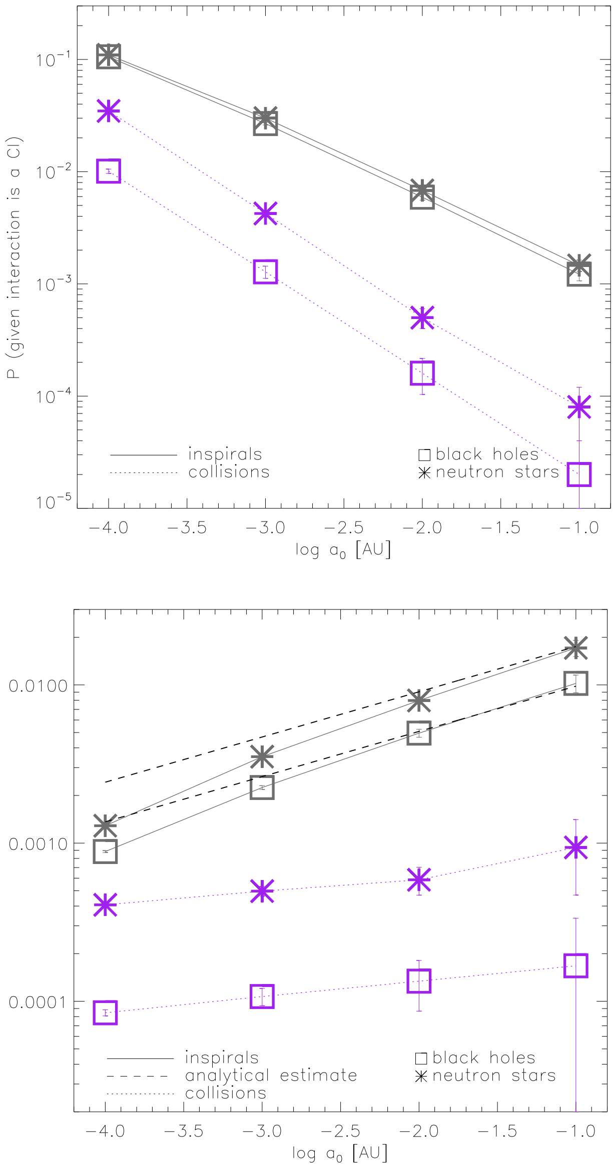


Figure 5.10 Formation of inspirals (*grey*) and collisions (*purple*) in equal mass binary-single interactions between either BHs (*squares*) or NSs (*stars*) as a function of the initial SMA of the target binary. All BHs have a $1\ M_{\odot}$ mass where the NSs have a $1.4\ M_{\odot}$ mass and a $12\ \mathrm{km}$ radius. The corresponding analytical estimates, given by equation (5.36), are shown as dashed lines. The general normalization is found by numerical experiments, but as can be seen our analytical model correctly separates the cross sections between NSs and BHs based solely on their mass difference. The reader is refered to the text for a discussion explaining the slight difference at low SMA between the simple analytical scaling and the simulations. *Top:* The probability for an outcome to result either in a collision or an inspiral given a CI. *Bottom:* The corresponding total integrated cross sections for each outcome. As expected, the probability for an inspiral decreases with SMA (equation 5.35) while the total cross section increases with SMA (equation 5.36). Widely separated binaries are thus expected to be the dominant target for producing inspirals. Our numerical results used 2×10^5 scatterings per SMA.

5.6.4 COMPARISON TO SINGLE-SINGLE CAPTURE

Inspirals resulting from binary-single interactions and mergers resulting from single-single GW capture can create binaries with extremely short merger times and, in some cases, with very high eccentricity. Comparing the formation probabilities for eccentric mergers arising from both mechanisms is thus of great interest.

A single-single capture occurs when two objects pass close enough to each other that the resulting GW energy losses are larger than their initial positive energy. To first order, the energy radiated away during the first passage can be obtained by integrating the GW energy losses along the initial, *unperturbed* unbound orbit (41):

$$\Delta E = -\frac{2}{15}\frac{G^{7/2}}{c^5}\frac{m_1^2 m_2^2 (m_1+m_2)^{1/2}}{r_{\mathrm{min}}(a,e)^{7/2}}h(e), \tag{5.38}$$

where $r_{\mathrm{min}} = a(1-e)$ is the minimum distance between the two objects in the unbound orbit and $h(e)$ is a dimensionless constant for which $h(e=1) = 425\pi/(8\sqrt{2})$. For capture to occur, we require $\Delta E > (1/2)\mu v_{\infty}^2$, where μ is the reduced mass. Combining this with equation (5.38), we find the maximum allowed r_{min} for a capture, which we denote r_{cap} (71),

$$r_{\mathrm{cap}} = \left(\frac{85\pi}{6\sqrt{2}}\right)^{2/7} \frac{G m_1^{2/7} m_2^{2/7} (m_1+m_2)^{3/7}}{c^{10/7} v_{\infty}^{4/7}}. \tag{5.39}$$

All single-single encounters with pericenter distance smaller than r_{cap} become bound.

In analogy with the CI interaction cross section derived in Section 5.3.1, the cross section for a single-single interaction with pericenter distance less than $r_{\mathrm{p,max}}$ can be written as

$$\sigma_{\mathrm{SS}}(r_{\mathrm{min}} < r_{\mathrm{p,max}}) \simeq \frac{2\pi G m_{\mathrm{tot}} r_{\mathrm{p,max}}}{v_{\infty}^2}. \tag{5.40}$$

The capture cross section can be estimated by inserting $r_{\mathrm{p,max}} = r_{\mathrm{cap}}$ in equation (5.40),

$$\sigma_{\mathrm{cap}} = 2\pi\left(\frac{85\pi}{6\sqrt{2}}\right)^{2/7} \frac{G^2 m_1^{2/7} m_2^{2/7} (m_1+m_2)^{10/7}}{c^{10/7} v_{\infty}^{18/7}}. \tag{5.41}$$

This cross section can then be compared directly with the cross section for inspirals arising from binary-single encounters. The ratio between the two cross sections can be approximated using equation (5.36),

$$\frac{\sigma_{\mathrm{insp}}}{\sigma_{\mathrm{cap}}} \propto \left(\frac{a_0 v_{\infty}^2}{m}\right)^{2/7}. \tag{5.42}$$

The number of inspirals relative to single-single captures is then expected to increase with a_0 and v_{∞}, but decrease as the mass increases.

Figure 5.11 shows the numerically derived ratio of binary-single inspirals to single-single captures based on 8×10^5 binary-single scatterings. The two mechanisms have similar cross sections for tight binaries and typical cluster velocity dispersions. For binary SMA larger than 10^{-3} AU, binary-single inspiral interactions clearly dominate. This implies that inspirals resulting from binary-single interactions may contribute substantially to the inspiraling and eccentric merging binary population in globular clusters. In the next section, we will explore the particularly interesting case of binaries that pass through the *LIGO* detector frequency band with high eccentricity.

5.7 ECCENTRIC INSPIRALS IN THE *LIGO* BAND

Compact merging binaries will be observed by advanced *LIGO* in the near future (42; 78; 2; 77). To detect these inspirals, templates must be convolved with the timeseries data from the interferometer (2; 123; 90; 14). The waveforms of relatively high eccentricity differ from those of circular binaries. For example, Huerta & Brown (52) find that for eccentricities greater than about $e \approx 0.2$, the match to circular templates is degraded by more than 50%. An understanding of the quantity and origin of eccentric binaries that pass through the *LIGO* band is therefore extremely important for future GW searches.

In the GW inspirals and mergers, one might expect that the majority of binaries will be nearly circular when entering the *LIGO* band, since GWs carry away both energy and angular momentum at a rate such that the circularization time is similar to the merging time (97; 39; 40). However, as we show in this paper, the dynamical inspiral states formed in binary-single encounters are formed with very high initial eccentricity and rapid merger times. As a result, most of these dynamical formed inspirals will be directly observable in the *LIGO* band at the time of formation, i.e. when they are still highly eccentric. In what follows, we explore in detail the fraction of highly eccentric *LIGO* sources one expects to come from binary-single interactions as well as making a direct comparison to highly eccentric inspirals formed via single-single interactions.

To quantify the number of eccentric binary mergers in our scattering experiments, we use an approximate form for the gravitational peak frequency (128),

$$f_{\mathrm{GW}} = \frac{1}{\pi}\sqrt{\frac{Gm_{\mathrm{tot}}}{a^3}} \frac{(1+e)^{1.1954}}{(1-e^2)^{1.5}}, \tag{5.43}$$

where $\sqrt{a^3/Gm_{\mathrm{tot}}}$ is the orbital time, T_{orb}.

5.7.1 ECCENTRIC BINARIES FROM BINARY-SINGLE INTERACTIONS

The eccentricity distribution of binaries resulting from binary-single interactions includes binaries that *evolve* into the *LIGO* band and binaries that are *born* in the *LIGO* band. Figure 5.12 shows the results from binary-single interactions between NSs with $1.4\ M_{\odot}$ masses and $12\ \mathrm{km}$ radius for different initial SMA of the target binary. The *top panel* shows the distribution of all binaries in the $\log(a, 1-e^2)$ plane immediately after final-state identification. Inspirals are shown with *large square* symbols. The distribution of inspirals is not static. Instead, each binary evolves due to GW radiation according to equation (5.22). The *dotted black* lines show a few of these evolutionary tracks. The two *dashed-black* lines show constant gravitational peak frequencies $f_{\mathrm{GW}} = 10^1, 10^4$ Hz, which have been chosen to illustrate the sensitivity window range for advanced *LIGO* (42; 77).

By comparing the orbit evolution trajectories in Figure 5.12 with the lines of constant f_{GW}, we can see that they are parallel for $\log(1-e^2) \ll 0$. This is because the evolution of a for both scales as $\left(1-e^2\right)^{-1}$. This implies that high eccentricity mergers that are not born in the *LIGO* band cannot evolve into it with high eccentricity. The binaries that are identified inside the *LIGO* band are thus the only ones that are able to be detected with high eccentricity. This set of binaries is the dynamically formed inspirals. From the (a, e) distributions shown in the *top* panel in Figure 5.12 one can calculate the corresponding f_{GW} distributions by making use of equation (5.43) (*bottom* panel in Figure 5.12). The values of f_{GW} are observed to change only slightly during inspiral, since the binaries spiral in with almost constant peak frequency. As observed in Figure 5.12, target binaries with $a \sim 10^{-2} - 10^{-3}$ AU produce inspirals with

$f_{\rm GW}$ distributions that peak around the most sensitive *LIGO* frequency ≈ 200 Hz. The relative normalizations of the distributions shown in the *bottom* panel of Figure 5.12 can be derived from Figure 5.10.

5.7.2 ECCENTRIC BINARIES FROM SINGLE-SINGLE CAPTURE

Once a binary is formed via single-single GW capture, its subsequent evolution can be followed in the (a, e) plane according to equation (5.22). By analogy with arguments presented above for the binary-single capture case, we can conclude that if binaries formed through single-single capture are not formed with $f_{\rm GW}$ that places them in the *LIGO* band, they will circularize before *LIGO* can observe them as eccentric binaries.

To estimate the cross section for highly eccentric *LIGO* sources resulting from single-single captures, we first rewrite equation (5.43) in the equal mass case and in the high eccentricity limit ($e \sim 1$),

$$r_0 \simeq \left(\frac{2^{2.3908}}{4\pi^2}\frac{Gm}{f_0^2}\right)^{1/3}, \tag{5.44}$$

where m is the mass of each of the objects, and r_0 is the required pericenter distance for an eccentric binary to have a peak frequency f_0. It then follows that all encounters with pericenter distance $r_{\rm min} < r_0$ will have $f_{\rm GW} > f_0$. Therefore, the cross section for a single-single encounter having $f_{\rm GW} > f_0$ can be simply calculated by setting $r_{\rm p,max} = r_0$ in equation (5.40),

$$\sigma_{\rm SS}(f_{\rm GW} > f_0) = \frac{4\pi Gm}{v_\infty^2}\left(r_0 - 2r_{\rm obj}\right). \tag{5.45}$$

To account for the object's finite size ($r_{\rm obj}$), we have subtracted the cross section for direct collisions in equation (5.45). The velocity dependence (v_∞^{-2}) in equation (5.45) implies that the cross section for high eccentricity single-single captures scales as the gravitational focusing cross section. The single-single capture cross section scales as $v_\infty^{-18/7}$ such that $\sigma_{\rm SS}(f_{\rm GW} > f_0)/\sigma_{\rm cap} \propto v_\infty^{4/7}$. As the velocity increases, the single-single high eccentricity cross section relative to the capture cross section will also increase. The *dashed-black* line in the *bottom* panel in Figure 5.12 shows the eccentricity distribution given by equation (5.45) for single-single encounters, which we confirmed using scattering experiments of single-single objects. In Figure 5.13, we show the different cross sections and corresponding scalings for the various outcomes expected from single-single and binary-single encounters.

5.7.3 COMPARISON BETWEEN BINARY-SINGLE AND SINGLE-SINGLE

In previous sections, we have computed the scalings for the cross sections of binary-single interactions and single-single captures; a summary of our results is given in Figure 5.13. We now turn our attention to the relative normalization of eccentric inspirals arising from binary-single and single-single capture as a function of binary SMA and GW frequency threshold. Figure 5.14 shows the normalization of the numerically computed inspiral cross sections for interacting NSs given three frequency thresholds $f_0 = 10, 30,$ and 100 Hz as a function of the initial binary SMA. The *upper* panel shows the resulting cross sections in AU^2 for encounters with $v_\infty = 10$ km s^{-1}. The *lower* panel shows these cross sections normalized to the corresponding single-single cross sections. Inspirals become increasingly dominant relative to the number of single-single eccentric binaries as the frequency threshold and the SMA increases. The ratio between the two cross sections is independent of velocity because both cross sections scale with the gravitational focusing cross section, v_{∞}^{-2}. This general behavior can be understood analytically by writing out the ratio

$$\frac{\sigma_{\rm insp}}{\sigma_{\rm SS}(f_{\rm GW} > f_0)} \simeq \frac{3}{4}\frac{P_{\rm insp}a_0}{r_0 - 2r_{\rm obj}} \propto a_0^{2/7} f_0^{2/3}. \tag{5.46}$$

The estimation of $P_{\rm insp}$ in this limit is given by equation (5.35). Our numerical and analytical results strongly suggest that the cross section for the formation of eccentric compact binary inspirals is significantly larger in the binary-single case than in the single-single case even when the fraction of compact objects in binaries is relatively modest.

5.8 DISCUSSION

We have discussed the formation of eccentric inspirals in the context of binary-single interactions and compared them to the more widely discussed single-single capture scenario. The expected outcomes for binary-single and single-single interactions of equal mass NSs are shown in Figure 5.15. The *solid-black* line shows the binary-single CI cross section. Other outcomes shown are sub-categories of the CI cross section. The *solid-red* line shows exchange, the *solid-grey* inspirals, and the *solid-purple* collisions. The *green* line shows binaries with merger lifetimes less than a Hubble time, which will be discussed in Section 5.8.2. Similarly, the *dashed-black* line shows the total cross section for single-single capture, while the *dashed-red* line shows only eccentric captures for which $f_{\rm GW} > 10$ Hz, and the *dashed-purple* line shows the collision cross section. As we emphasized in the previous section, most inspirals occur with $f_{\rm GW} \gtrsim 10$ Hz, so the inspiral cross section may be directly compared to the eccentric component of the single-single cross section. The upper *x*-axis label shows the GW inspiral lifetime for binaries separated by a given initial SMA (bottom *x*-axis labels).

Here we turn our attention to the implications of our results and illustrate how they change with the inclusion of a more extended binary companion by calculating scatterings for WD-NS binaries in Section 5.8.1. We discuss the merger lifetime and resulting center-of-mass kicks in Sections 5.8.2 and 5.8.3, respectively. We provide a simple estimate of typical event rates in dense stellar systems in Section 5.8.4. Finally, we present our conclusions in Section 5.8.5.

5.8.1 TARGET BINARIES CONTAINING WHITE DWARFS

We have seen that wider binary SMAs lead to an enhancement in the cross section for inspiral outcomes in the case of binaries comprised of NSs and BHs. In widely separated binaries, the binary members need not be compact objects. In this section, we consider the case where the target binary contains a white dwarf (WD) companion (125).

WDs have a well defined mass-radius relationship, which takes the following form for lower-mass WDs,

$$\begin{aligned} r_{\rm WD} &\simeq \frac{1}{m_{\rm WD}^{1/3}}\frac{(18\pi)^{2/3}}{10}\frac{\hbar^2(m_{\rm p}/0.5)^{-5/3}}{Gm_{\rm e}}, & (5.47)\\ &\approx 2.9\times10^9\,(m_{\rm WD}/M)^{-1/3}\,{\rm cm}, & (5.48)\end{aligned}$$

where $m_{\rm e}$ is the electron mass and $m_{\rm p}$ the proton mass (15).

Another characteristic scale imposed by the size of the WD is the separation at which the WD fills its Roche lobe,

$$a_{\rm MT} \simeq r_{\rm WD}\frac{0.6q^{2/3} + \ln(1+q^{1/3})}{0.49q^{2/3}}, \tag{5.49}$$

where $q = m_{\rm WD}/m_{\rm NS}$ (24). In WD-NS binaries containing moderately massive WDs, the resulting mass transfer is stable, and the binary overcomes the destabilizing effects produced by GW radiation due to the ongoing mass transfer (e.g. 79; 96).

The phase space of NS-NS binary outcomes that result from NS scatterings including a companion WD are shown in the upper panel of Figure 5.16, which can be directly compared to the upper panel of Figure 5.12. These experiments involve a $1.4\ M_{\odot}$ NS encountering a WD-NS binary containing a $0.5\ M_{\odot}$ WD and a $1.4\ M_{\odot}$ NS. A comparison to Figure 5.15 shows the increased importance of collisions in the WD-NS target case when compared to NS-NS targets. However, we see that inspiral outcomes between two NSs are still possible, despite the presence of the WD. By contrast, inspirals between the WD and the NS typically do not occur due to the extended radius of the WD (see e.g. 129, for double WDs seen by *LISA*[4]). However, the cross-section for inspirals is reduced somewhat as compared to NS-NS target binaries. This is partially due to the fact that there is one (rather than three) possible pairwise combination that can result in double NS binaries. Additionally, in tight binaries with $a \approx a_{\rm MT}$, collisions with the WD play an important role in depleting inspiral outcomes (74). The hierarchy of masses in the system also likely plays a role by somewhat reducing the typical number of resonances (116). Despite these effects which tend to deplete the number of inspiral outcomes, we find that NS-NS inspirals have a larger cross section than single-single captures with $f_{\rm GW} > 10$ Hz as long as the binary SMA $a_0 \gtrsim 10^{-3}$ AU. Thus we still expect wide binaries containing WDs to contribute meaningfully to the eccentric inspiral channel, in particular if they dominate the NS-hosting binary population as in Grindlay et al. (37). A concern for systems containing extended objects is that tidal dissipation may play in important role in modifying the dynamics (e.g. 80), an effect we ignore here and hope to implement in future work.

5.8.2 BINARY LIFETIMES

Even if the initial binary lifetime is greater than a Hubble time, $t_{\rm Hubble}$, a fraction of binaries that undergo a scattering will be either deposited or exchanged into orbits with very short lifetimes (17). Thus a fraction of even very widely separated binaries can produce mergers with $t_{\rm life} < t_{\rm Hubble}$. Figure 5.17 shows the distribution of final binary lifetimes realized following binary-single scatterings with varying binary SMA. In the classical point-mass limit, we see that an approximate power-law distribution is produced. The inclusion of GW radiation and finite radii introduces two physical scales that break the self-similarity of the problem. The hard cutoff corresponds to the scales of the objects themselves and depletion by collisions. The inspiral population manifests itself as a knee at scales corresponding to the typical pericenter distances of the rapid inspiral outcomes.

The cross section for creation of binary products whose lifetime are less than a Hubble time is plotted in Figure 5.15 for encounters involving NS. The key feature of this cross section is that it does not vanish when $a_0 \gtrsim 10^{-1.7}$ AU, where $t_0 > t_{\rm Hubble}$. Instead this cross section remains approximately flat. The reason for this is that resultant binaries generally have a much smaller pericenter distance than the target binary and therefore also a shorter lifetime as seen in Figure 5.17.

---

[4] http://lisa.nasa.gov/

5.8.3 RETENTION OR EJECTION OF BINARY-SINGLE OUTCOMES

A remaining question is whether final binaries resulting from binary-single interactions are kicked out, or whether they merge in-situ. Kicks relative to the initial center of mass occur when a fraction of the initial binary's binding energy is transferred to the relative motion of the binary and the single (98). We denote the resulting binary kick velocity as v_{kick}. The associated hardening of these binaries leads to a shorter binary lifetime (since $t_{\mathrm{life}} \propto a^4$) and one therefore expects that a high kick velocity is associated with a short lifetime. A binary that receives a high-velocity kick will therefore not necessarily merge outside of its environment.

This tradeoff between lifetime and kick velocity is evident in Figure 5.18. The Figure shows a scatter plot of kick velocity v_{kick} and survival distance, defined as $v_{\mathrm{kick}} \times t_{\mathrm{life}}$ for all endstate NS binaries with respect to the initial center of mass. We use the survival distance to estimate where the binary will merge. Radius and escape velocity for a typical globular cluster are shown with dashed lines. In this simple calculation only final binaries in the upper right quadrant merge outside the cluster. If we now assume that binary SMAs are lognormally distributed and we only consider binaries that merge in less than a Hubble time (below the *dash-dotted* line in Figure 5.18), we calculate that $\sim$ 30% (10%) of all merging binaries arising from NS-NS (NS-WD) targets are kicked out with a median distance of $\sim$ 80 (50) kpc. While there is little direct evidence that close double neutron star binaries can form and merge in globular clusters, the double neutron star system PSR B2127+11C in the Galactic GC M15 (3) is an example of such a system and has $t_{\mathrm{life}} \approx 2 \times 10^8$ years.

The retention or ejection of binaries has implications for cluster dynamics and merger-induced transients such as e.g. short gamma-ray bursts (8; 73). If binaries are retained, they participate in the continued cluster evolution acting as a heat source or sink depending on their SMA. In some cases the binary distribution may reach a steady-state (e.g. 58). Merging binaries are expected to show environmental dependance in their electromagnetic signatures (95; 110; 83; 64; 109).

If a relativistic (short-gamma-ray burst) or a mildly relativistic mass ejection resulted from the merger of two compact objects, the resulting afterglow could then, at least in part, be due to the interaction of the ejecta with the stellar winds of the red giant cluster members (18). Due to the large stellar density in the cluster core, the external shock would then take place within a more dense medium than the IGM (74). In addition, the merger sites of compact binaries will determine whether we expect the electromagnetic signatures of binary mergers to statistically trace the globular cluster distribution around galaxies (37; 74; 16) or the galactic potential (11; 112; 8; 131; 130; 26; 65; 27).

5.8.4 RATES

Given distributions of target binaries and single encounters, we can convert the calculated cross sections into event rates. In this section we present some simple order-of-magnitude estimates of the rates of dynamical NS-NS inspirals achieved in globular cluster environments. We denote the total number of NSs by N_{NS}, and assume that some fraction f_{b} are in N_{bin} binary systems (target binaries). The remaining fraction remains single (encounter population), $f_{\mathrm{s}} = 1 - f_{\mathrm{b}}$. The target binaries are distributed according to their SMA dN_{bin}/da, which we assume is lognormal, $dN_{\mathrm{bin}}/da \propto a^{-1}$. The differential rate of inspirals per SMA can then be written

$$\frac{d\Gamma_{\mathrm{insp}}}{da} = \frac{dN_{\mathrm{bin}}}{da} n_{\mathrm{s}} \sigma_{\mathrm{insp}} v_{\infty}, \tag{5.50}$$

where n_s is the number density of single NSs, $n_\mathrm{s} = f_\mathrm{s} N_\mathrm{NS}/V_\mathrm{core}$, and V_core is the volume of the cluster core over which both single and binary objects are distributed. To obtain the total rate of inspirals, we integrate over the binary distribution,

$$\Gamma_\mathrm{insp} = \int \frac{d\Gamma_\mathrm{insp}}{da} da. \tag{5.51}$$

We note here that while we need to evaluate this integral for a given binary distribution and inspiral cross-section as a function of SMA, it will generally scale as $\Gamma_\mathrm{insp} \propto N_\mathrm{NS}^2 f_\mathrm{b}(1 - f_\mathrm{b}) v_\infty^{-1}$. Below we provide some rate estimates based on simple examples that describe the distribution of NSs in globular clusters.

In a typical globular cluster, there may be as many as $N_\mathrm{NS} \sim 10^3$, for example, as modeled in the case of M15 by Murphy et al. (88) whose best fit model has 1500 NSs with a half-mass radius of 0.17 pc. In what follows, we take $V_\mathrm{core} = (0.17\mathrm{pc})^3$, a typical relative velocity $v_\infty = 10\mathrm{km~s}^{-1}$, and $N_\mathrm{NS} = 10^3$. If 30% of these NSs are in NS-NS binaries distributed between 10^{-3} and 1 AU in SMA ($f_\mathrm{b} = 0.3$), the rate of NS inspirals will be

$$\Gamma_\mathrm{insp}^{(\mathrm{NS-NS})} \approx 0.7~\mathrm{yr}^{-1}~\mathrm{Gpc}^{-3}. \tag{5.52}$$

To express the above rate in units of volume, we have assumed that the density of galaxies is $n_\mathrm{gal} = 0.1~\mathrm{Mpc}^{-3}$ and each galaxy has 100 globular clusters, $N_\mathrm{GC} = 10^2$, implying 10 GC/Mpc^{-3} (13).

If we instead assume NSs are in WD-NS binaries distributed between 10^{-3} and 1 AU in SMA ($f_\mathrm{b} = 0.3$) and treat our WD-NS scattering cross section as representative for these binaries, we find

$$\Gamma_\mathrm{insp}^{(\mathrm{WD-NS})} \approx 0.3~\mathrm{yr}^{-1}~\mathrm{Gpc}^{-3}. \tag{5.53}$$

As with the NS-NS case, this numeric result scales $\propto N_\mathrm{NS}^2 f_\mathrm{b}(1 - f_\mathrm{b}) v_\infty^{-1}$. This estimate should be treated as an upper limit, because, if, for example the NS is in a binary with a main sequence star, the effects of collisions will be more significant than those with a WD companion.

These same assumptions imply a rate of single-single NS captures in globular clusters,

$$\Gamma_\mathrm{cap} = f_\mathrm{s} N_\mathrm{NS} n_\mathrm{s} \sigma_\mathrm{cap} v_\infty \approx 0.5~\mathrm{yr}^{-1}~\mathrm{Gpc}^{-3} \tag{5.54}$$

where we note that the velocity dependence in this case is $v_\infty^{-11/7}$. By the same token, we can calculate the rate of eccentric binaries in the *LIGO* band arising from single-single encounters

$$\Gamma_\mathrm{SS}(f_\mathrm{GW} > 10~\mathrm{Hz}) \approx 0.15~\mathrm{yr}^{-1}~\mathrm{Gpc}^{-3}, \tag{5.55}$$

which has a velocity dependence v_∞^{-1}. Thus, if the binary fraction $f_\mathrm{b} > 0.18$ (for WD-NS binaries) or $f_\mathrm{b} > 0.08$ (for NS-NS binaries), the binary-single channel will dominate the formation of eccentric NS inspirals over the widely discussed single-single channel.

We can also compare to the number of non-eccentric mergers which occur from dynamical interactions. These are defined in our scattering experiments as those binaries arising from either an exchange or flyby interaction whose lifetime is less than a Hubble time, $t_\mathrm{life} < t_\mathrm{Hubble}$. If we take our NS-NS target binary simulations as representative, non-eccentric merger outcomes have a rate of approximately

$$\Gamma_\mathrm{merge}^{(\mathrm{NS-NS})} \approx 120~\mathrm{yr}^{-1}~\mathrm{Gpc}^{-3}. \tag{5.56}$$

Binaries with $t_\mathrm{life} < t_\mathrm{Hubble}$ are thus more common by a factor of approximately 160 than inspirals. Grindlay et al. (37), whose rate estimate is in rough agreement with equation (5.56), concludes that $\sim 10\%$ of all mergers may be dynamically assembled in globular clusters. The remainder of mergers are expected to arise from binaries assembled in the field (e.g. 19; 20). However, the exact fraction of mergers in clusters depends sensitively on the distribution of wide binaries containing compact objects which is difficult to constrain observationally. If this estimate is correct, then the inspiral rate represents a $\sim 1\%$ fraction of the anticipated total compact object merger rate assembled in cluster.[5]

Normalized to the rate of eccentric NS mergers from single-single capture for which $f_{\rm GW} > 10$ Hz, We can write a hierarchy of rates as

$$\begin{aligned}\Gamma_{\rm SS}(f_{\rm GW} > 10\,{\rm Hz}) : \Gamma_{\rm insp}^{\rm (WD-NS)} : \Gamma_{\rm insp}^{\rm (NS-NS)} : \Gamma_{\rm merge}^{\rm (NS-NS)} \\ \approx 1 : 2 : 5 : 800.\end{aligned} \tag{5.57}$$

The expected number, and correspondingly the number density, of BHs in globular clusters remains uncertain. Mass-segregation, for example, has been argued to give rise to a BH-dominated subsystem that collapses and dynamically decouples from the remainder of the stellar system (118; 69; 92). In this case, very high BH number densities can be achieved, leading to the formation of a binary population through GW capture. Binary-single and single-single BH interactions are expected to rapidly eject BHs from the cluster after the formation of binaries (69; 115). However, these binary interactions may also produce inspirals and mergers, perhaps even leading to the runaway formation of a massive black hole (99). Even if the number of BH binaries is small, the number density of single black holes may be high enough to produce an inspiral rate comparable to the NS inspiral rate (92; 91). However, it is probably unreasonable to expect that a fraction of order unity of globular clusters might undergoing such an extreme phase at a given time. We therefore expect NS-NS inspirals rather than BH-BH inspirals to dominate the inspiral rate.

5.8.5 SIGNIFICANCE OF ECCENTRIC INSPIRALS

We have demonstrated that binary-single scatterings are likely to dominate the production of eccentric binaries. In such GW-driven inspirals, the energy change is much more rapid than the angular momentum change, such that the circularization time and inspiral time are similar, $t_{\rm insp} \approx t_{\rm circ}$ (97). One consequence of this is that binaries whose peak frequency, equation (5.43), is at lower frequency than the *LIGO* band will enter the *LIGO* band with relatively low eccentricity since these objects tend to circularize as they inspiral. This can be seen most clearly in the trajectories drawn in Figures 5.12 and 5.16. For a binary to be seen as eccentric in a given waveband, it must have been formed with high eccentricity in that band. Eccentric inspirals produce gravitational waveforms which are distinct from those of circularly inspiraling binaries (68; 120; 21; 35; 23; 34; 52). These may be so distinct that non-circular binaries will go undetected without uniquely created waveform templates (22; 52), and the timing between pre-merger GW bursts will contain valuable information about the equation of state. Close encounters in these systems can also lead to tidal deformations strong enough to crack the crust of the NS and tap into the $\sim 10^{46}$ erg stored in elastic energy, potentially generating flaring activity prior to the merger (126; 127). In contrast to quasi-circular NS-NS mergers, eccentric binary mergers can also result in massive disks even for equal mass binaries (23).

Neutron stars that merge with high eccentricity have potentially unique gravitational and post-merger electromagnetic signatures (e.g. 74; 23). The merger of these binaries may eject

---

[5] This estimate neglects other channels that could lead to eccentric binaries mergers, such as Kozai resonance in a triple systems (87; 124).

copious neutron rich material in tidal tails that will synthesize significantly larger masses of r-process rich material (74; 109) than the widely discussed, non-eccentric binary mergers (70; 107; 106; 103; 73; 84; 101; 7; 63; 4; 121; 38).

Multi-messenger astronomy offers tantalizing prospects for probing the nature of compact objects, their binary assembly, evolution, and eventual merger (105; 104; 12; 74; 108; 25; 83; 75; 64; 89; 93; 94; 9; 122; 5), in addition to possible insights into the origin of r-process nucleosynthetic elements and short gamma-ray bursts (70; 72; 111; 102; 86; 101; 7). An eccentric GW signal detection might be one of the most exciting prospects, as it would provide a clear signature of the dynamical binary assembly process. In the explicit absence of such detection, the use of eccentric waveform template searches could help exclude a significant dynamically assembled population of merging compact binaries in dense stellar systems.

It is a pleasure to thank J. Goldstein, J. Guillochon, S. H. Hansen, J. Hjorth, D. Kasen, L. Kelley, W. Lee, L. Lehner, I. Mandel, C. Miller, F. Pretorius, S. Rosswog, and D. Tsang for helpful discussions. M.M. and E.R-R. thank the DARK cosmology centre for its hospitality. We acknowledge support from the David and Lucile Packard Foundation, NSF grant: AST-0847563 and the NSF Graduate Research Fellowship (M.M.). The Dark Cosmology Centre is funded by the Danish National Research Foundation.

5.9 N-BODY INTEGRATOR WITH GW ENERGY LOSS CORRECTION

We use a Fourth-Order Hermite Integrator with a variable time step to evolve the N-body system. The dynamical effect from GW radiation is included using the Post-Newtonian (PN) formalism (10) by modifying the Newtonian acceleration term from $\mathbf{a}_0$ to $\mathbf{a}_0 + c^{-5}\mathbf{a}_5$ as described in Section 5.5.1. This modified PN expansion of the acceleration is strictly valid only for two isolated objects. However, one can still make use of this approach without introducing significant errors for $N > 2$ objects since the 2.5PN term has a much steeper dependence on the distance r than the Newtonian acceleration ($r^{-9/2}$ vs r^{-2} for a circular binary). The contribution from the closest pair will therefore always dominate. Further justification for this formalism can be found in (40). The 2.5PN term is the first term in the expansion that acts like an energy sink, i.e. carries energy out of the system. The energy loss from this term is, when orbit averaged, equivalent to the loss calculated from the quadrupole formalism described in (97). A comparison between the two approaches is shown in Figure 5.19, which plots the orbital evolution in the (a, e) plane for a binary that inspirals (*top panel*) because of GW radiation and for a single object that captures another single one by emitting GW (*bottom panel*). The *black-solid* lines are from our N-body code where the *red* dots show the result from solving for (a, e), using the quadrupole formalism: equations (5.18) and (5.19). Very good agreement in these tests was found, as can be seen in Figure 5.19.

To speed up the binary-single scattering experiments we have propagated the encounter from infinity to a distance r_{proj} from the center-of-mass (COM) of the target binary by modeling the binary-single system as a two-body system. The distance r_{proj} was chosen to be a fraction of the maximum value of either r_{bs} or a_{b}, where r_{bs} is the minimum distance between the COM of the binary and the interloper in the two-body frame and a_{b} is the SMA of the binary. This approach ignores the effect from the binary's dipole gravitational field on the encounter for $r > r_{\mathrm{proj}}$, but the error is insignificant. Further details on the errors related to this strategy can be found in (55).

5.10 IDENTIFYING STATES

5.10.1 BINARY-SINGLE STATE

Following (29) we state that the three interacting objects are in a binary-single state if the binary objects are bound to each other and the tidal force from the single at the binary's apocenter ($F_{\rm tid}$) is smaller than the relative force at apocenter ($F_{\rm rel}$) by some fraction $\delta_{\rm tid}$, i.e. if $F_{\rm tid}/F_{\rm rel} < \delta_{\rm tid}$. The two force terms are simply given by

$$F_{\rm rel} = \frac{m_{\rm bin,1}m_{\rm bin,2}}{[a(1+e)]^2} \tag{5.58}$$

and

$$F_{\rm tid} \simeq \frac{2(m_{\rm bin,1}+m_{\rm bin,2})m_{\rm s}}{r^3}a(1+e), \tag{5.59}$$

where $m_{{\rm bin},i}$ is the mass of binary object i, $m_{\rm s}$ the mass of the single object, r the distance between the single object and the center-of-mass of the binary and a, e are the semi-major axis and eccentricity of the binary, respectively.

If a three-body state is identified as a binary-single state and the single object is unbound from the binary, the state is labeled either as an *exchange* or a *fly-by* depending on which objects the binary is composed of. If the single object is instead bound to the binary, the state is denoted as an *intermediate binary-single state* (IMS). In this case, the bound single is chosen to have a finite minimum distance to the binary. The chosen threshold, $\delta_{\rm tid}$, will thus have an influence on the identified number of IMS and the corresponding distribution in (a, e). There is no dependence on $\delta_{\rm tid}$ if the single is unbound. For this work we use a $\delta_{\rm tid} = 0.5$ for identifying IMS and $\delta_{\rm tid} = 0.1$ for identifying exchange or a fly-by.

5.10.2 INSPIRALS

A binary with a bound single companion that inspirals due to GW radiation is denoted an *inspiral*. Since the binaries that inspirals have a bound companion, the inspiral state is a subclass of the IMS discussed above. In these cases, the (a, e) values for the orbital parameters of the inspiraling binary are set at initial identification, when the three-body state is identified as an IMS. The value for this first set of (a, e) depends strongly on the threshold $\delta_{\rm tid}$ since a smaller $\delta_{\rm tid}$ allows more time for the binary to spiral in. However, the total number of inspirals is not affected, and therefore the resulting cross sections are also not sensitive to the choice of $\delta_{\rm tid}$.

5.10.3 COLLISIONS

We assume in all scattering experiments that the objects are rigid spheres with radius r_i. We say that object i and j have collided if these spheres ever overlap, $r_{ij} < r_i + r_j$. To distinguish collisions from inspirals we say that collisions are colliding objects that are not in an IMS binary. This definition is practical, but there is some gray-zone between collisions and inspirals. One can for example have an IMS binary with initial pericenter distance $r_{\rm min} < r_i + r_j$, or a configuration where enough GW energy is radiated away such that two objects collide before an IMS is identified by the code. In general, this overlap is only important at the very smallest binary SMAs, in which the SMA begins to become comparable to the size of the objects, of order 10^{-5} AU for solar mass compact objects. At larger separations, any sensitivity is lost because the number of inspirals greatly dominates over the number of direct impacts.

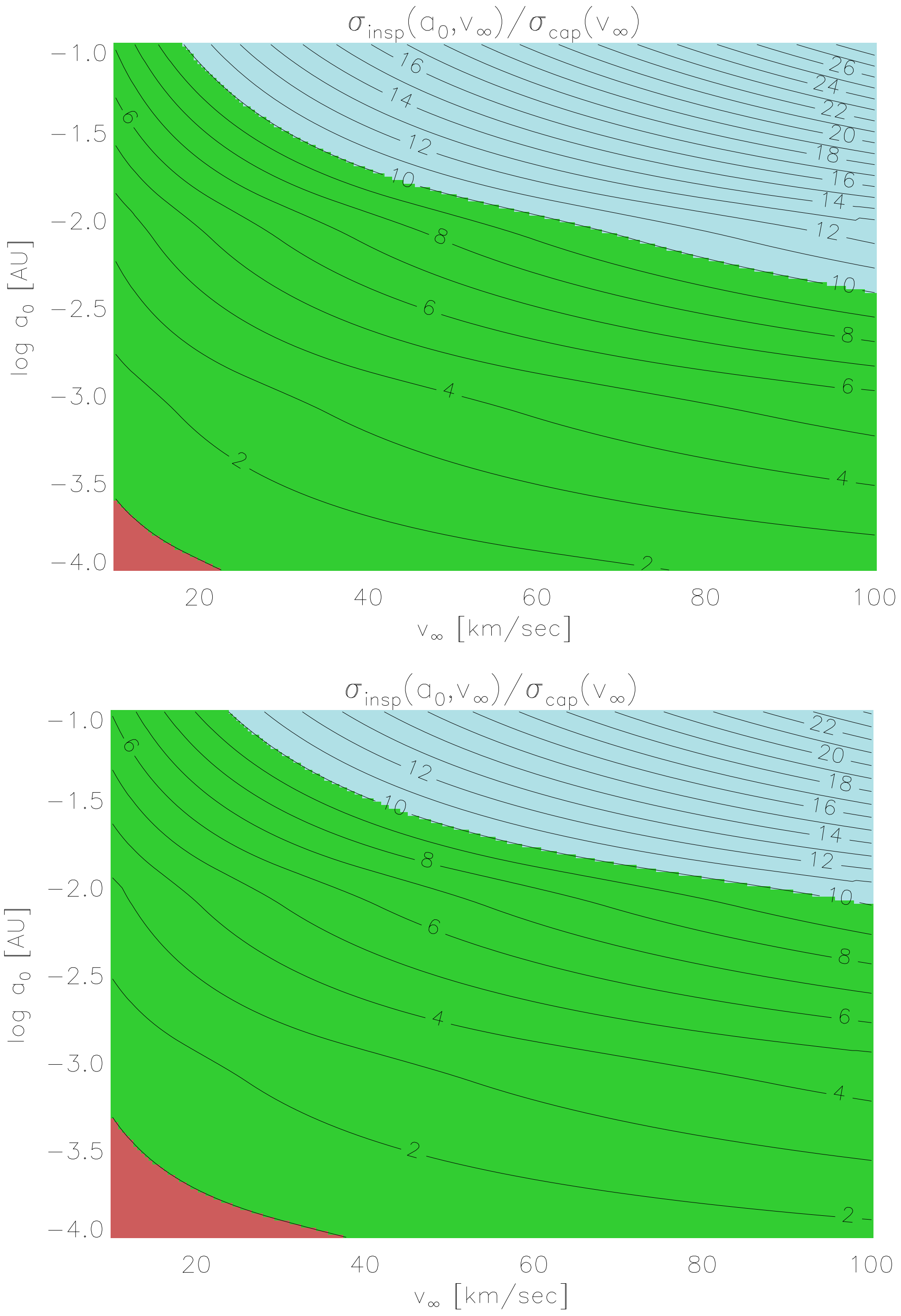


Figure 5.11 Ratio between GW single-single capture cross section and binary-single inspirals cross section, $\sigma_{\rm insp}/\sigma_{\rm cap}$, as a function of relative velocity at infinity v_∞ and SMA, a_0, of the target binary. All interactions are between equal mass objects. *Top:* $\sigma_{\rm insp}/\sigma_{\rm cap}$ for $1\ M_\odot$ BHs. *Bottom:* $\sigma_{\rm insp}/\sigma_{\rm cap}$ for $1.4\ M_\odot$ NSs with radius 12 km. The *red*, *green* and *light blue* colors respectively mark the regions where $\sigma_{\rm insp} < \sigma_{\rm cap}$, $\sigma_{\rm cap} < \sigma_{\rm insp} < 10\sigma_{\rm cap}$ and $10\sigma_{\rm cap} < \sigma_{\rm insp}$.

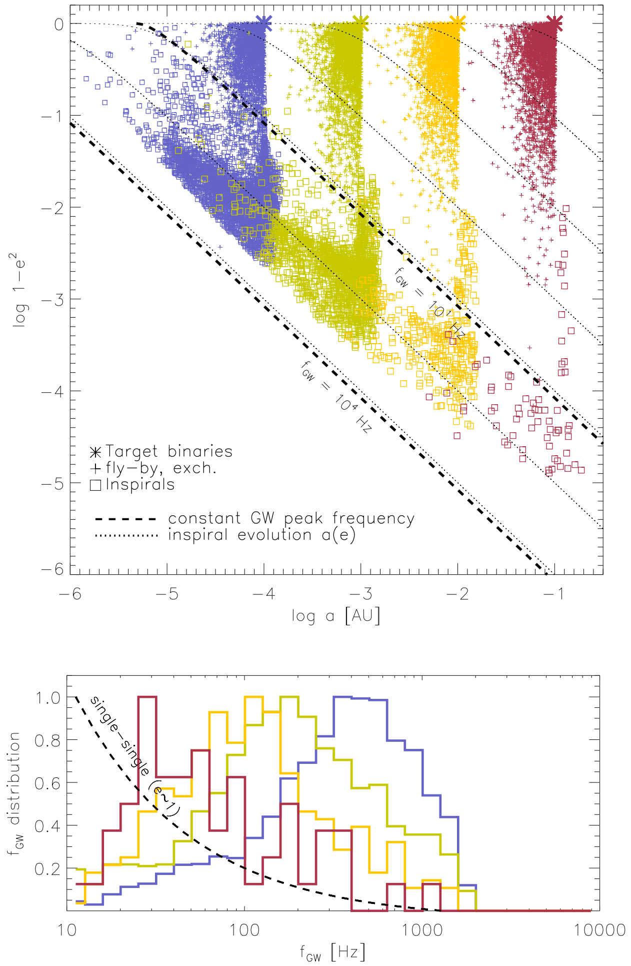


Figure 5.12 Distribution of orbital parameters in the $(a, 1-e^2)$ plane and the corresponding gravitational peak frequency $f_{\mathrm{GW}}(a, e)$ for all endstate binaries resulting from binary-single interactions between NSs with $1.4\ M_{\odot}$ masses and 12 km radii. The relative velocity between the encounter and the target binary is $v_{\infty} = 10\ \mathrm{km\ s^{-1}}$. The plot includes the classical outcomes exchange and fly-by (*plus symbols*) and the GR outcome inspirals (*squares*). Different colors denote different initial SMA of the target binary. *Top:* Orbital parameters at the time of final state identification. The inspirals fade away as the SMA increases. The *dashed-black lines* show the GW peak frequencies $10^1, 10^4$ Hz that are approximately representative of the advanced *LIGO* window. The *dotted black lines* show a few examples of the inspiral orbital evolution due to GW radiation given by equation (5.22). When $(1-e^2) \ll 1$, these evolutionary tracks are parallel to the gravitational peak frequency lines. This implies that if a binary with high eccentricity is *not* formed in the *LIGO* band, then it will *never* evolve into it with high eccentricity. Inspirals are therefore the only states arising from a binary-single interaction that will have the potential of being observable as high eccentric mergers. *Bottom:* Distributions of gravitational peak frequencies from all identified inspirals. These distributions stay almost unchanged during the inspiral since the binaries evolve with approximately constant GW frequency. The sensitivity of *LIGO* peaks around ~ 200 Hz. The *dashed-black line* shows the eccentricity distribution expected from merging binaries resulting from single-single captures. For illustration purposes, all histograms have been normalized to their peak values.

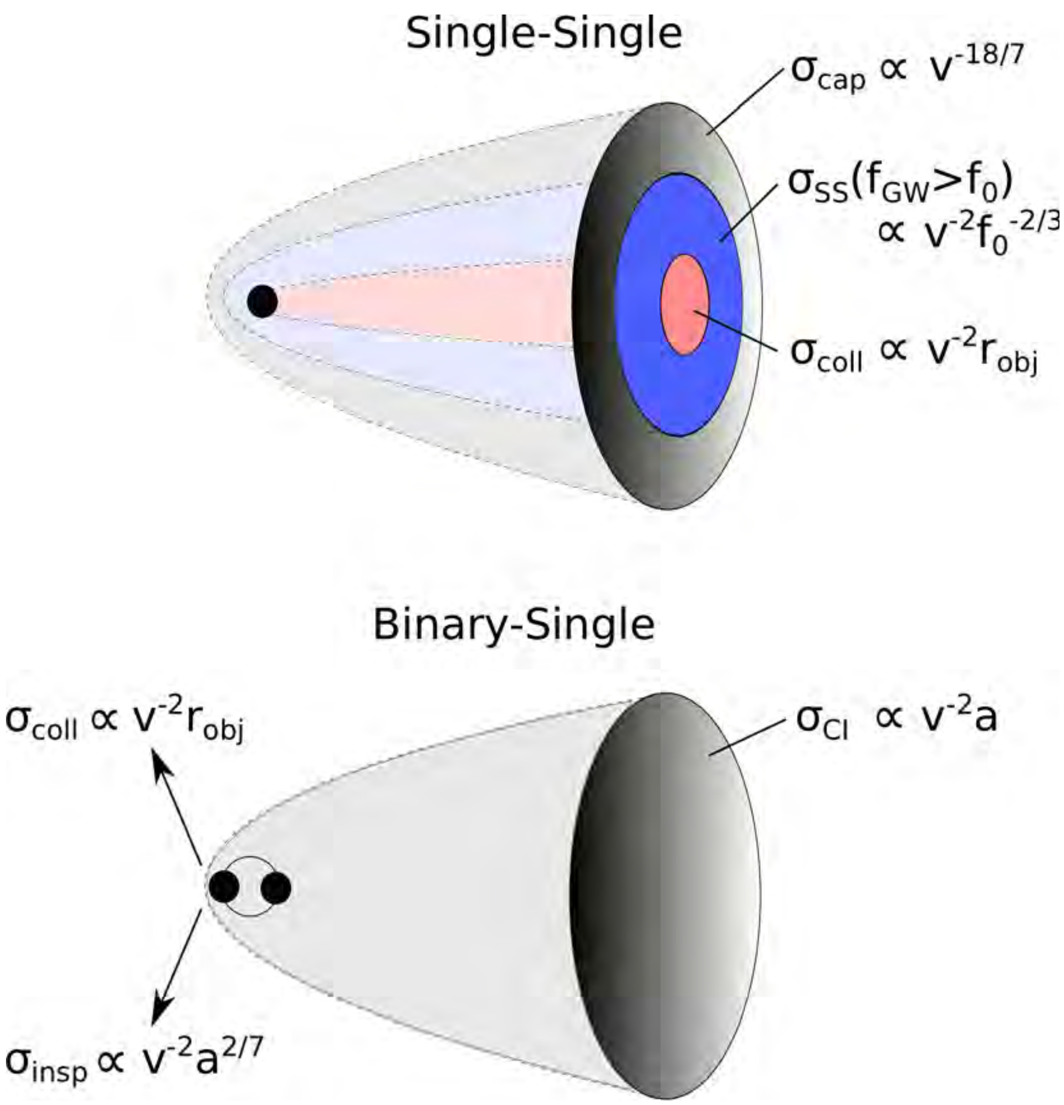


Figure 5.13 Diagram illustrating different outcome cross sections arising from single-single (*top*) and binary-single (*bottom*) interactions. Also shown are the approximate dependence of the various cross sections on the encounter velocity, v, the SMA of the target binary, a, the object radius, $r_{\rm obj}$, and the gravitational peak frequency, $f_{\rm GW}$. The single-single capture cross section is denoted by $\sigma_{\rm cap}$ (equation 5.41), the high eccentric single-single capture cross section with $f_{\rm GW} > f_0$ by $\sigma_{\rm SS}(f_{\rm GW} > f_0)$ (equation 5.45), the direct collision cross section by $\sigma_{\rm coll}$ (equation 5.40), the CI cross section by $\sigma_{\rm CI}$ (equation 5.3) and the binary-single inspiral cross section by $\sigma_{\rm insp}$ (equation 5.36). It is particularly interesting to compare $\sigma_{\rm SS}(f_{\rm GW} > f_0)$ with $\sigma_{\rm insp}$ because the inspiralling binaries formed in each of these cases give very similar observational signatures. For example, both channels can form inspirals that enter the *LIGO* band with high eccentricity, an event that is not observed when field binaries merge. Since both channels scale as $\propto v^{-2}$, their ratio is independent of v as shown in equation (5.46).

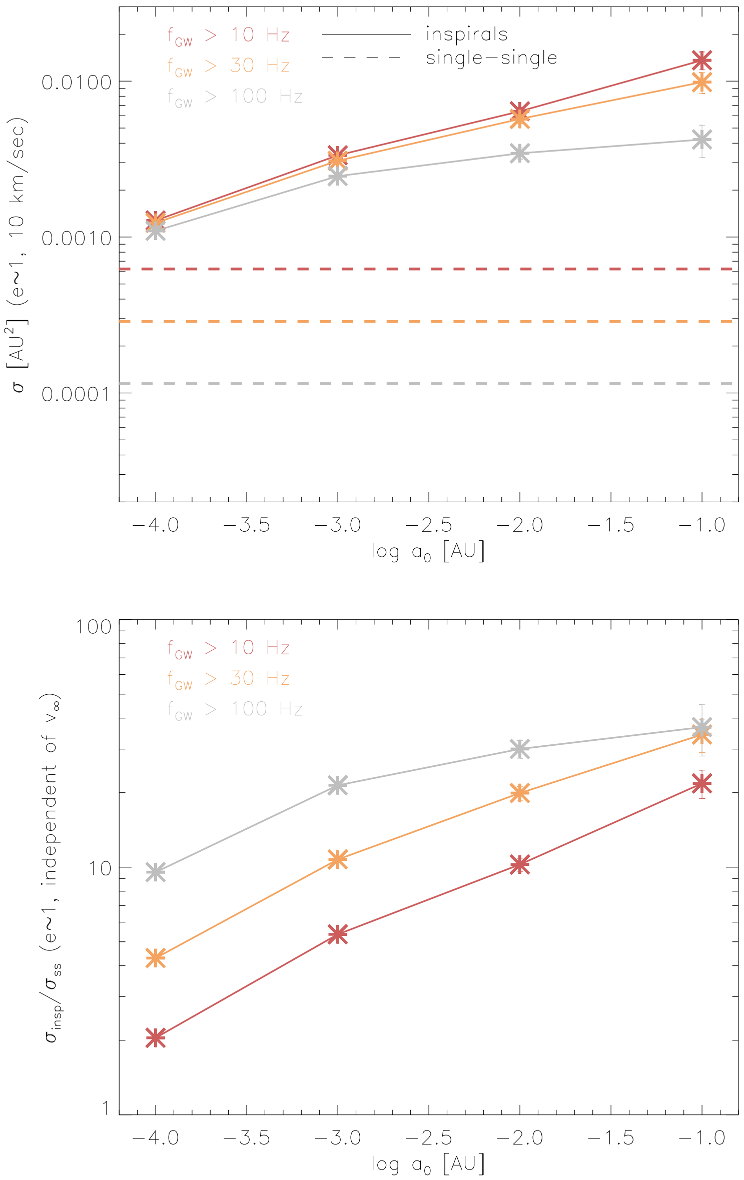


Figure 5.14 Numerically calculated cross sections for high eccentricity binaries ($e \sim 1$) arising from binary-single encounters and their relative importance when compared to those produced by single-single encounters. We divide the merging binaries based on their gravitational peak frequency at formation: $f_{\rm GW} > 10, 30, 100$ Hz. During the inspiral, the orbital parameters (a, e) change according to equation (5.22) but $f_{\rm GW}$ remains relatively constant, which means that the results are not altered significantly as the binary evolves. All results are for scatterings between NSs with $1.4\ M_{\odot}$ masses and 12 km radii. *Top:* Inspiral cross sections. Solid lines show inspirals formed by binary-single interactions and dashed lines show inspirals formed by single-single captures. The resultant high eccentricity binaries formed via binary-single and single-single encounters have different gravitational peak frequencies at formation as shown in Figure 5.12. Each line defined by $f_{\rm GW}$ denotes a cross section that only includes inspirals that are born with a gravitational frequency above the given threshold. *Bottom:* Ratio between the single-single and binary-single cross sections shown in the *top* panel. As described in the text, both high eccentricity single-single and binary-single inspirals scale as $\propto v_{\infty}^{-2}$. This makes the ratio *independent* of velocity.

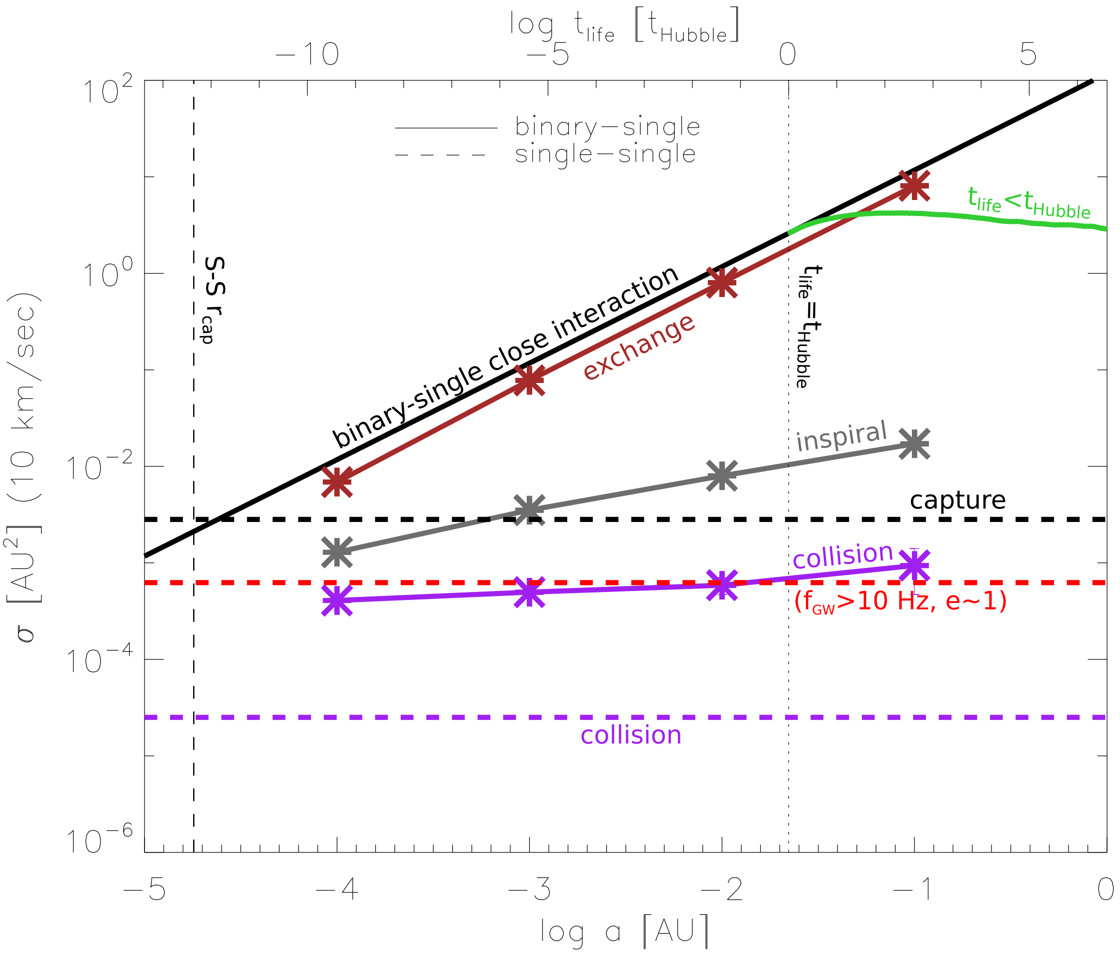


Figure 5.15 Summary of relevant outcome cross sections arising from binary-single and single-single encounters between equal mass NSs. Each NS has a mass of $1.4\ M_{\odot}$ and radius of 12 km. The *dashed* lines show results from single-single encounters while the *solid* lines show results from binary-single interactions. The *black* solid line shows the CI crossection, the *dark-grey* line the inspiral cross section and the *purple* and *brown* lines the cross sections for collisions and exchanges, respectively. The *green* line shows the cross section for binaries that merge in less than a Hubble time. The *black-dashed* shows the single-single capture cross section and the *red-dashed* line shows the cross section for single-single high eccentric ($e \sim 1$) binary with gravitational peak frequency $f_{\mathrm{GW}} > 10$ Hz. The *vertical-black-dashed* line shows the single-single pericenter distance for a capture r_{cap}. We note that the scaling between lines depends on velocity, here assumed to be 10 km s^{-1}.

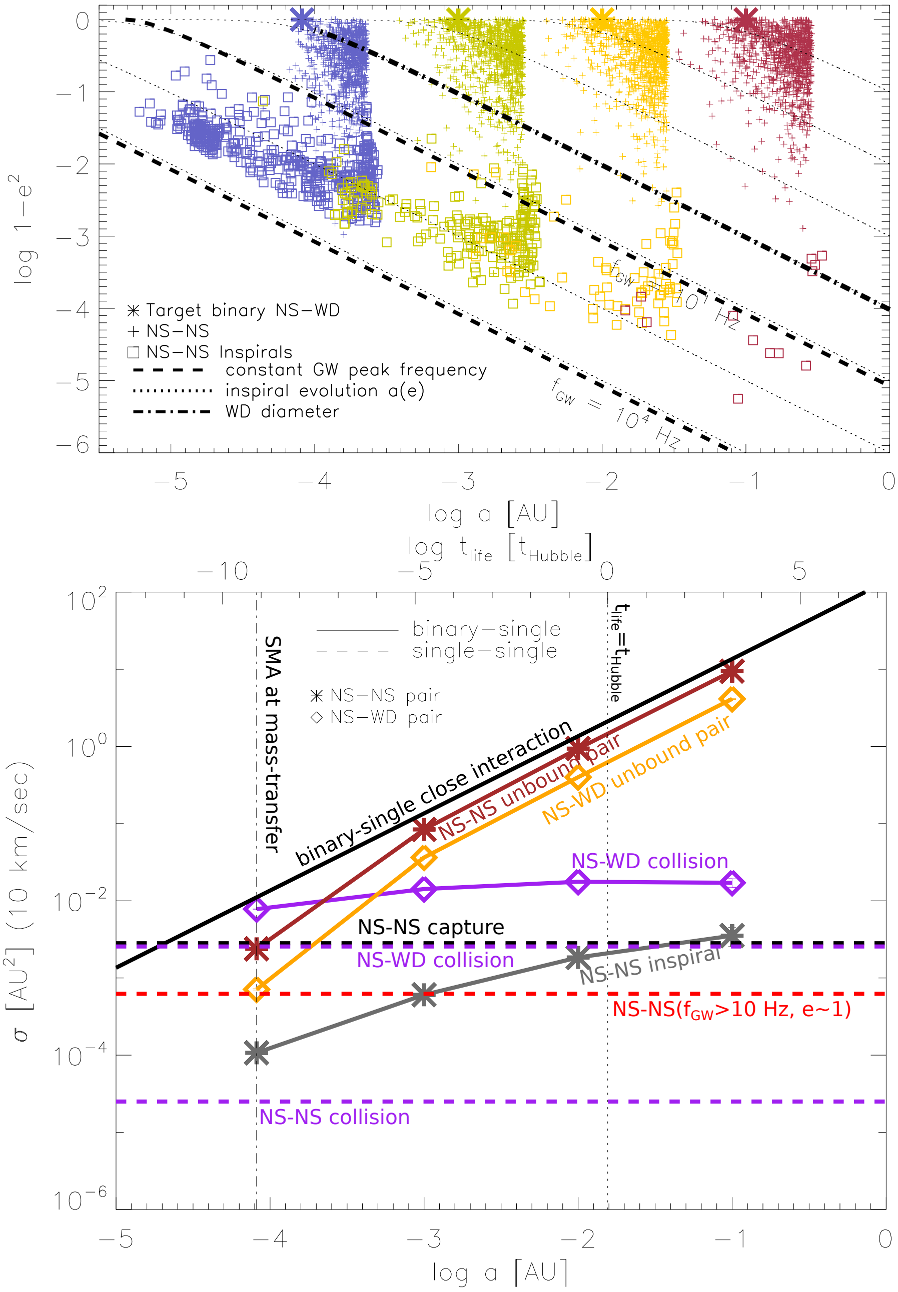


Figure 5.16 Results from scatterings between a NS$(1.4\ M_{\odot}, 12\ \mathrm{km})$ − WD$(0.5\ M_{\odot})$ binary and a NS$(1.4\ M_{\odot}, 12\ \mathrm{km})$ encounter with $v_{\infty} = 10\ \mathrm{km\ s^{-1}}$. *Top:* Scatter plot of the orbital parameters $(a, 1-e^2)$ for all endstate NS-NS binaries (similar to Figure 5.12). The resulting inspirals are shown with *square* symbols. Each color show results for a given SMA. The radius of the WD is shown as a *dashed-dotted line*. As can be seen, this line is well above the region where inspirals form implying that WD inspirals are very unlikely. *Bottom:* Similar to Figure 5.15 but for target binaries including a WD companion with $0.5\ M_{\odot}$. The cross section for inspirals is significantly smaller here than in the equal mass NS case. The three main reasons for this are that only 1 out of 3 endstates can result in a NS-NS inspiral, collisions with the WD deplete inspiral outcomes, and the relatively small mass of the WD suppresses resonances which could otherwise form inspirals.

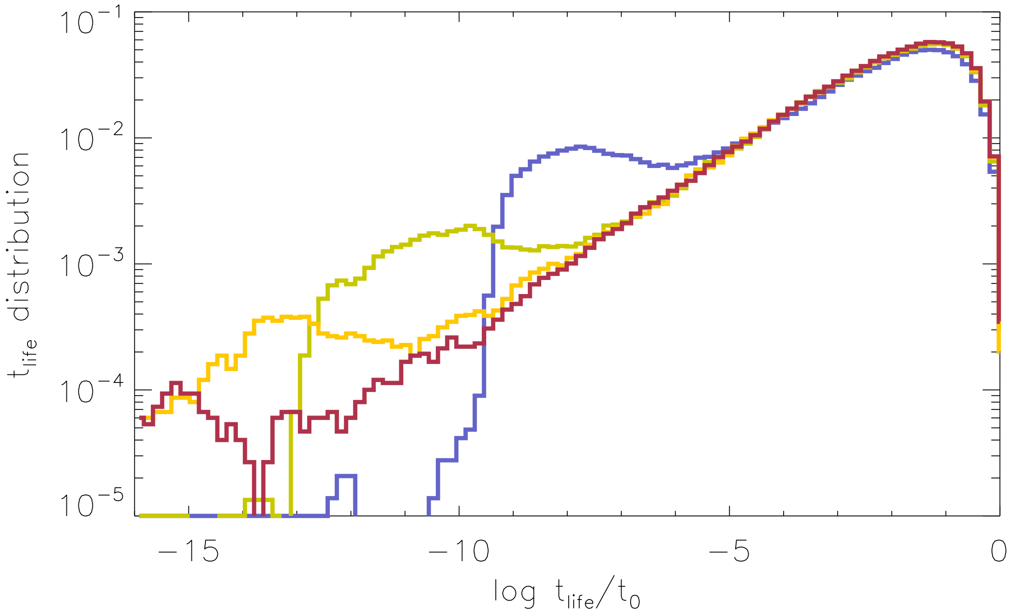


Figure 5.17 Lifetime distributions including all endstate binaries. Colors denote initial binary SMA, a_0, from $10^{-4} - 10^{-1}$ AU (blue to red). In the Newtonian point-mass case, all initial SMAs would follow the same distribution, but when collisions and GR are included then the initial SMA a_0 plays a role in forming the final distribution. The knee that appears for each distribution is the fast merging inspirals.

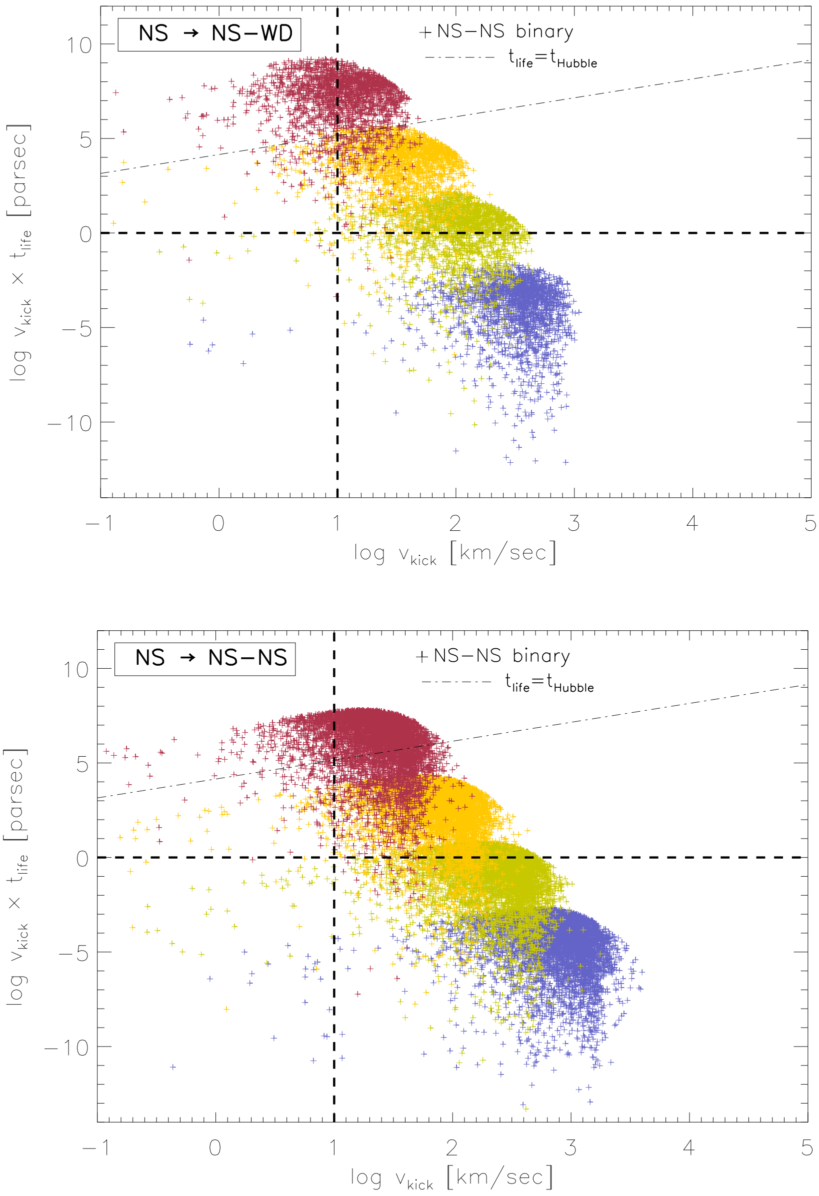


Figure 5.18 Resulting kick velocities $v_{\rm kick}$ and travelled distance before merger defined as $v_{\rm kick} \times t_{\rm life}$ for endstate NS-NS binaries. The kick velocity is with respect to the initial center of mass of the interaction. In all scatterings, the encounter comes from infinity with $v_{\infty} = 10$ km s^{-1}. *Top:* Results from the scattering NS $\rightarrow$ NS $-$ WD $(0.5\ M_{\odot})$. *Bottom:* Results from the scattering NS $\rightarrow$ NS $-$ NS. The *dashed* lines show characteristic values for a typical globular cluster. In this simple picture all binaries in quadrants II-IV will merge within the cluster while binaries in the upper right corner will merge outside. The corresponding single object will be kicked in the opposite direction with a fraction $m_{\rm bin}/m_{\rm sin}$ of the binary's kick velocity. The *dash-dot* line shows where the binary lifetime is equal to the Hubble time. All binaries below the line will have a lifetime less than a Hubble time. Different colors indicate different initial SMA.

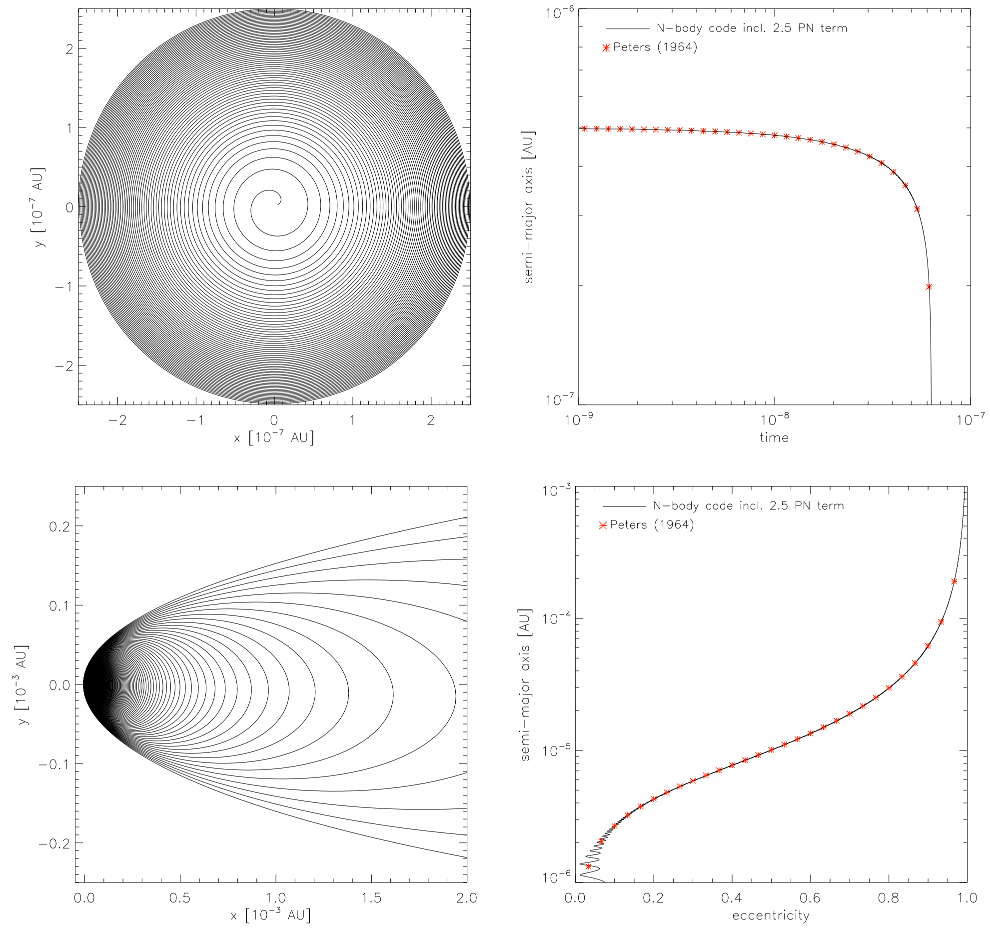


Figure 5.19 Comparison between our N-body code (*solid-black lines*) and the analytical solution from (97) (*red points*). *Top:* A circular binary that spirals in due to GW radiation. The upper left plot shows the trajectory of one of the objects. The upper right plot shows how the distance between the two objects decreases with time. *Bottom:* Evolution of an initial highly eccentric binary. The two objects are initially not bound to each other, but enough energy is radiated away in terms of GW to make the system bound after the first orbit. This is an illustration of a single-single capture. The lower left plot shows the evolution of the incoming single in the rest-frame of the target object. To the right is shown the evolution in the (a, e) plane. The wiggles in the lower left corner (for low a and e) illustrate the limitation of the integration scheme. As seen, we find good agreement between our code and the analytical prediction in both cases.

BIBLIOGRAPHY

[1] Aarseth, S. J., & Lecar, M. 1975, ARA&A, 13, 1

[2] Abadie, J., et al. 2010, Class. Quantum Grav., 27, 173001

[3] Anderson, S. B., Gorham, P. W., Kulkarni, S. R., Prince, T. A., & Wolszczan, A. 1990, Nature, 346, 42

[4] Barnes, J., & Kasen, D. 2013, eprint arXiv:1303.5787

[5] Bartos, I., Brady, P., & Márka, S. 2013, Classical and Quantum Gravity, 30, 123001

[6] Baumgardt, H., Hut, P., & Heggie, D. C. 2002, MNRAS, 336, 1069

[7] Bauswein, A., Goriely, S., & Janka, H. T. 2013, eprint arXiv:1302.6530

[8] Belczynski, K., Perna, R., Bulik, T., Kalogera, V., Ivanova, N., & Lamb, D. Q. 2006, ApJ, 648, 1110

[9] Berger, E., Fong, W., & Chornock, R. 2013, ArXiv e-prints

[10] Blanchet, L. 2006, Living Reviews in Relativity, 9, 4

[11] Bloom, J. S., Sigurdsson, S., & Pols, O. R. 1999, MNRAS, 305, 763

[12] Bloom, J. S., Holz, D. E., Hughes, S. A., et al. 2009, arXiv:0902.1527

[13] Brodie, J. P., & Strader, J. 2006, ARA&A, 44, 193

[14] Brown, D. A., Kumar, P., & Nitz, A. H. 2013, Phys. Rev. D, 87, 082004

[15] Carroll, B. W., & Ostlie, D. A. 1996, An Introduction to Modern Astrophysics

[16] Church, R. P., Levan, A. J., Davies, M. B., & Tanvir, N. 2011, MNRAS, 413, 2004

[17] Clausen, D., Sigurdsson, S., & Chernoff, D. F. 2012, MNRAS, 428, 3618

[18] De Colle, F., Ramirez-Ruiz, E., Granot, J., & Lopez-Camara, D. 2012, ApJ, 751, 57

[19] Dominik, M., Belczynski, K., Fryer, C., Holz, D. E., Berti, E., Bulik, T., Mandel, I., & O'Shaughnessy, R. 2012, ApJ, 759, 52

[20] —. 2013, arXiv:1308.1546

[21] East, W., Pretorius, F., & Stephens, B. 2012, Phys. Rev. D, 85, 124009

[22] East, W. E., McWilliams, S. T., Levin, J., & Pretorius, F. 2013, Phys. Rev. D, 87, 043004

[23] East, W. E., & Pretorius, F. 2012, ApJ, 760, L4

[24] Eggleton, P. P. 1983, ApJ, 268, 368

[25] Faber, J. A., & Rasio, F. A. 2012, Living Reviews in Relativity, 15, 8

[26] Fong, W., Berger, E., & Fox, D. B. 2010, ApJ, 708, 9

[27] Fong, W.-f., & Berger, E. 2013, arXiv:1307.0819

[28] Fregeau, J. M. 2008, ApJ, 673, L25

[29] Fregeau, J. M., Cheung, P., Portegies Zwart, S. F., & Rasio, F. A. 2004, MNRAS, 352, 1

[30] Fregeau, J. M., Gürkan, M. A., Joshi, K. J., & Rasio, F. A. 2003, ApJ, 593, 772

[31] Fregeau, J. M., Ivanova, N., & Rasio, F. A. 2009, ApJ, 707, 1533

[32] Fregeau, J. M., Joshi, K. J., Portegies Zwart, S. F., & Rasio, F. A. 2002, ApJ, 570, 171

[33] Fregeau, J. M., & Rasio, F. A. 2007, ApJ, 658, 1047

[34] Gold, R., Bernuzzi, S., Thierfelder, M., Brügmann, B., & Pretorius, F. 2012, Phys. Rev. D, 86, 121501

[35] Gold, R., & Bruegmann, B. 2012, eprint arXiv:1209.4085

[36] Goodman, J., & Hut, P. 1993, ApJ, 403, 271

[37] Grindlay, J., Zwart, S. P., & McMillan, S. 2006, Nat Phys, 2, 116

[38] Grossman, D., Korobkin, O., Rosswog, S., & Piran, T. 2013, arXiv:1307.2943

[39] Gültekin, K., Miller, M. C., & Hamilton, D. P. 2004, ApJ, 616, 221

[40] —. 2006, ApJ, 640, 156

[41] Hansen, R. 1972, Phys. Rev. D, 5, 1021

[42] Harry, G. M., & the LIGO Scientific Collaboration. 2010, Class. Quantum Grav., 27, 084006

[43] Heggie, D. C. 1975, MNRAS, 173, 729

[44] Heggie, D. C., & Hut, P. 1993, ApJS, 85, 347

[45] Heggie, D. C., Hut, P., & McMillan, S. L. W. 1996, ApJ, 467, 359

[46] —. 1996, ApJ, 467, 359

[47] Hills, J. G. 1975, AJ, 80, 1075

[48] —. 1975, AJ, 80, 809

[49] —. 1976, MNRAS, 175, 1P

[50] Hills, J. G., & Fullerton, L. W. 1980, AJ, 85, 1281

[51] Hopman, C., Guetta, D., Waxman, E., & Portegies Zwart, S. 2006, ApJ, 643, L91

[52] Huerta, E. A., & Brown, D. A. 2013, eprint arXiv:1301.1895

[53] Hut, P. 1983, ApJ, 268, 342

[54] —. 1993, ApJ, 403, 256

[55] Hut, P., & Bahcall, J. N. 1983, ApJ, 268, 319

[56] Hut, P., et al. 1992, Astronomical Society of the Pacific, 104, 981

[57] Ivanova, N., Belczynski, K., Fregeau, J. M., & Rasio, F. A. 2003, arXiv:astro-ph/0312497

[58] —. 2005, MNRAS, 358, 572

[59] Ivanova, N., Chaichenets, S., Fregeau, J., Heinke, C. O., Lombardi, J. C. J., & Woods, T. E. 2010, ApJ, 717, 948

[60] Ivanova, N., Fregeau, J. M., & Rasio, F. A. 2005, Binary Radio Pulsars, 328, 231

[61] Ivanova, N., Heinke, C. O., Rasio, F. A., Belczynski, K., & Fregeau, J. M. 2008, MNRAS, 386, 553

[62] Ivanova, N., Heinke, C. O., Rasio, F. A., Taam, R. E., Belczynski, K., & Fregeau, J. 2006, MNRAS, 372, 1043

[63] Kasen, D., Badnell, N. R., & Barnes, J. 2013, eprint arXiv:1303.5788

[64] Kelley, L. Z., Mandel, I., & Ramirez-Ruiz, E. 2013, Phys. Rev. D, 87, 123004

[65] Kelley, L. Z., Ramirez-Ruiz, E., Zemp, M., Diemand, J., & Mandel, I. 2010, ApJ, 725, L91

[66] Kelly, B. C., & Shen, Y. 2013, ApJ, 764, 45

[67] Kocsis, B., & Levin, J. 2012, Phys. Rev. D, 85, 123005

[68] Königsdörffer, C., & Gopakumar, A. 2006, Phys. Rev. D, 73, 124012

[69] Kulkarni, S. R., Hut, P., & McMillan, S. J. 1993, Nature, 364, 421

[70] Lattimer, J. M., & Schramm, D. N. 1974, ApJ, 192, L145

[71] Lee, M. H. 1993, ApJ, 418, 147

[72] Lee, W. H., & Ramirez-Ruiz, E. 2002, ApJ, 577, 893

[73] Lee, W. H., & Ramirez-Ruiz, E. 2007, New J. Phys., 9, 17

[74] Lee, W. H., Ramirez-Ruiz, E., & van de Ven, G. 2010, ApJ, 720, 953

[75] Lehner, L., Palenzuela, C., Liebling, S. L., Thompson, C., & Hanna, C. 2012, Phys. Rev. D, 86, 104035

[76] Lightman, A., & Shapiro, S. 1978, Rev. Mod. Phys., 50, 437

[77] LIGO Scientific Collaboration, Virgo Collaboration, Aasi, J., et al. 2013, arXiv:1304.0670

[78] Mandel, I., & O'Shaughnessy, R. 2010, Class. Quantum Grav., 27, 114007

[79] Marsh, T. R., Nelemans, G., & Steeghs, D. 2004, MNRAS, 350, 113

[80] McMillan, S. L. W. 1986, ApJ, 306, 552

[81] —. 1991, In: The formation and evolution of star clusters (A93-48676 20-90), 13, 324

[82] McMillan, S. L. W., & Hut, P. 1996, Astrophysical Journal v.467, 467, 348

[83] Metzger, B. D., & Berger, E. 2012, ApJ, 746, 48

[84] Metzger, B. D., et al. 2010, MNRAS, 406, 2650

[85] Meylan, G., & Heggie, D. C. 1997, A&A Rev., 8, 1

[86] Miller, M. C. 2005, ApJ, 626, L41

[87] Miller, M. C., & Hamilton, D. P. 2002, ApJ, 576, 894

[88] Murphy, B. W., Cohn, H. N., & Lugger, P. M. 2011, ApJ, 732, 67

[89] Nissanke, S., Kasliwal, M., & Georgieva, A. 2013, ApJ, 767, 124

[90] Nitz, A. H., Lundgren, A., Brown, D. A., et al. 2013, arXiv:1307.1757

[91] O'Leary, R. M., O'Shaughnessy, R., & Rasio, F. A. 2007, Phys. Rev. D, 76, 061504

[92] O'Leary, R. M., Rasio, F. A., Fregeau, J. M., Ivanova, N., & O'Shaughnessy, R. 2006, ApJ, 637, 937

[93] Palenzuela, C., Lehner, L., Ponce, M., et al. 2013, arXiv:1301.7074

[94] Palenzuela, C., Lehner, L., Liebling, S. L., et al. 2013, arXiv:1307.7372

[95] Panaitescu, A., Kumar, P., & Narayan, R. 2001, ApJ, 561, L171

[96] Paschalidis, V., MacLeod, M., Baumgarte, T. W., & Shapiro, S. L. 2009, Phys. Rev. D, 80, 24006

[97] Peters, P. 1964, Phys. Rev., 136, B1224

[98] Phinney, E. S., & Sigurdsson, S. 1991, Nature (ISSN 0028-0836), 349, 220

[99] Portegies Zwart, S. F., Baumgardt, H., Hut, P., Makino, J., & McMillan, S. L. W. 2004, Nature, 428, 724

[100] Rasio, F. A., et al. 2007, Highlights of Astronomy, 14, 215

[101] Roberts, L. F., Kasen, D., Lee, W. H., & Ramirez-Ruiz, E. 2011, ApJ, 736, L21

[102] Rosswog, S. 2004, arXiv.org

[103] —. 2005, ApJ, 634, 1202

[104] —. 2007, MNRAS, 376, L48

[105] —. 2007, Triggering Relativistic Jets (Eds. William H. Lee & Enrico Ramírez-Ruiz) Revista Mexicana de Astronomía y Astrofísica (Serie de Conferencias) Vol. 27, 27, 57

[106] Rosswog, S., & Liebendörfer, M. 2003, MNRAS, 342, 673

[107] Rosswog, S., Liebendörfer, M., Thielemann, F. K., Davies, M. B., Benz, W., & Piran, T. 1999, A&A, 341, 499

[108] Rosswog, S., Piran, T., & Nakar, E. 2012, eprint arXiv:1204.6240

[109] —. 2013, MNRAS, 430, 2585

[110] Rosswog, S., & Ramirez-Ruiz, E. 2002, MNRAS, 336, L7

[111] —. 2003, MNRAS, 343, L36

[112] Rosswog, S., Ramirez-Ruiz, E., & Davies, M. B. 2003, MNRAS, 345, 1077

[113] Shakeshaft, R., & Spruch, L. 1979, Rev. Mod. Phys., 51, 369

[114] Shapiro, S. L., & Teukolsky, S. A. 1983, Research supported by the National Science Foundation. New York, Wiley-Interscience, 1983, 663 p., -1

[115] Sigurdsson, S., & Hernquist, L. 1993, Nature, 364, 423

[116] Sigurdsson, S., & Phinney, E. S. 1993, ApJ, 415, 631

[117] —. 1995, ApJS, 99, 609

[118] Spitzer, L. J. 1969, ApJ, 158, L139

[119] Steiner, A. W., Lattimer, J. M., & Brown, E. F. 2010, ApJ, 722, 33

[120] Stephens, B. C., East, W. E., & Pretorius, F. 2011, ApJ, 737, L5

[121] Tanaka, M., & Hotokezaka, K. 2013, eprint arXiv:1306.3742

[122] Tanvir, N. R., Levan, A. J., Fruchter, A. S., Hjorth, J., Wiersema, K., Tunnicliffe, R., & de Ugarte Postigo, A. 2013, ArXiv e-prints

[123] The NRAR Collaboration et al. 2013, ArXiv e-prints

[124] Thompson, T. A. 2011, ApJ, 741, 82

[125] Thompson, T. A., Kistler, M. D., & Stanek, K. Z. 2009, arXiv:0912.0009

[126] Tsang, D., Read, J. S., Hinderer, T., Piro, A. L., & Bondarescu, R. 2012, Physical Review Letters, 108, 011102

[127] Tsang, D. 2013, arXiv:1307.3554

[128] Wen, L. 2003, ApJ, 598, 419

[129] Willems, B., Kalogera, V., Vecchio, A., Ivanova, N., Rasio, F. A., Fregeau, J. M., & Belczynski, K. 2007, ApJ, 665, L59

[130] Zemp, M., Ramirez-Ruiz, E., & Diemand, J. 2009, ApJ, 705, L186

[131] Zheng, Z., & Ramirez-Ruiz, E. 2007, ApJ, 665, 1220

6

GENERATING HIGH VELOCITY OBJECTS

A Gravitational Double Scattering Mechanism for Generating High Velocity Objects

Johan Samsing[1]

[1] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

Abstract

We present a dynamical model describing how halo particles can receive a significant energy kick from the merger between their own host halo and a target halo. This is relevant for understanding the growth of cosmological halos and could also provide an explanation for high velocity objects ranging from stars to galaxies. The model we present is a *double scattering mechanism* where a halo particle is given a significant energy kick by undergoing two subsequent gravitational deflections during the merger: The first deflection is by the potential of the target halo, whereas the second is by the particle's original host halo potential. The kick is generated because the two deflection frames are moving relative to each other. We derive an analytical solution for this dynamical energy kick in case of a head-on merger between two Hernquist halos. From this we find that the particles receiving a significant positive energy kick are all located in a cone pointing along the velocity vector of their host halo, where the maximum kick is given to particles located $\sim 0.1R_{\rm vir}$ from the center. We estimate for a $1:10$ merger that a kicked particle can be observed with a radial velocity exceeding ~ 2 times the virial velocity of the target halo at its virial sphere. This motivate us to explain the high velocity of the recently discovered globular cluster HVCG-1 (19). Our proposed double scattering mechanism is not only applicable to halo mergers, but can be applied to nearly any merging system. For instance, within the field of heavy nuclei scatterings it is well known that an electron can be ejected or captured by undergoing a similar double scattering process.

6.1 INTRODUCTION

Several high energy objects on seemingly unbound orbits have been observed, ranging from stellar objects (16; 117), supernovae (SNe) (88; 37), gamma ray burst (GRBs) (105; 35; 12) to extended objects like globular clusters (GCs) (19; 73) and dwarf galaxies (65; 23). In many of these observed cases the kick energy mechanism is unknown and a range of mechanisms have been suggested. One of the proposed ways is binary-single interactions where the binding energy of a binary is dynamically released into to a third object which thereby can escape with high velocity (44). These interactions can either be between stellar objects (96; 40) or between stars encountering either single or binary massive black holes (BHs) (47; 115; 14). Several observations indicate in fact that stellar interactions with the massive BH at the center of our galaxy is a very likely explanation for at least some high velocity stars (15; 38). More extended objects like GCs are probably not kicked by BH binary interactions due to the high probability for disruption, however the outcome from such an interaction is still uncertain (19). Dark matter (DM) subhalo interactions on the other hand can dynamically kick extended objects up to ~ 3 times the virial velocity of the host halo without major disruptions as indicated by numerical simulations (63). High velocity stars can also be created in isolated binaries if the heavier member undergoes a violent mass loss, a channel first suggested by (11) to explain the high number of "run away" O-B stars. More exotic kick mechanisms to describe hostless stellar remnants, pulsars and possible hyper velocity BHs have also been suggested from the role of asymmetric GW radiation (28; 74; 86; 34; 8) to the asphericity of supernovae explosions (53; 18; 17).

Unbound particles have also been discussed from a cosmological perspective. Recent studies (7) illustrate that $\sim 10\%$ of all the DM at the virial radius are in fact unbound. Luminous matter with no specific host halo have also been observed in especially galaxy clusters, a component also known as intra cluster light (ICL). This has been extensively studied both through observations (e.g. (83; 39; 118)) and numerically (111) and is believed to be a direct consequence of the dynamical evolution of galaxies which includes e.g. tidal stripping and mergers (68). Theoretical attempts have also been made to understand the final distribution of particles in DM halos. This include models from spherical collapse (e.g. (27; 9)) to statistical mechanics (e.g. (71; 64; 99; 41; 79; 48)). Especially concerning the unbound and high velocity component, recent work by (103; 55; 20) show that high velocity particles are likely generated in mergers due to rapid mean field changes in the potential. This was also noted by (1) who further proposed a direct connection to the observed population of high velocity B-type stars.

Data from upcoming surveys like *LSST*[1] and especially *Gaia*[2] will in the near future also measure positions and velocities for more than ~ 150 million stars with unprecedented precision. This not only offers unique possibilities for mapping out the current Milky Way potential and its past evolution (e.g. (84; 89; 72; 116)), but will also make it possible to make detailed studies of the past dynamical interactions (38). A central question could here be if the Milky Way in its past had a SMBH binary dynamically interacting with the environment. Detections of high velocity objects are here again playing a central role.

In this paper we present a new dynamical mechanism in which high velocity particles are created in the merger between two halos consisting of DM particles. It is well known that halo mergers produce an unbound component (46), but no clear dynamical explanation has been given yet. In this work we seek to present such an explanation. The mechanism we propose is a double scattering mechanism where a given subject particle gets a significant energy kick by undergoing two deflections during the merger. We derive analytically the energy kick for two

---

[1] http://www.lsst.org/lsst/

[2] http://sci.esa.int/gaia/

merging Hernquist halos (45), but the mechanism is not limited to this scenario. For instance, we note that a very similar mechanism has been described within heavy nuclei interactions where an electron can be ejected into the continuum (unbound orbit) or captured by a passing nucleus (dynamical capture) by undergoing a double collision[3](95).

The paper is organized in the following way: In section 6.2 we give a description of our proposed double scattering mechanism and introduce a model describing the orbit of a particle undergoing a double scattering. From this model we derive in section 6.3 an analytical solution for estimating the kick energy of the particle as a function of its position in the incoming halo prior to the merger. In section 6.4 we describe observational signatures. Here we especially discuss the possibility that the high velocity of HVGC-1 recently discovered by (19) was generated by our double scattering mechanism during a merger between a dark matter halo and Virgo. Conclusions are given in section 6.5.

6.2 KINEMATICS OF THE DOUBLE SCATTERING MECHANISM

We will now describe the kinematics of the double scattering mechanism. In section 6.2.1 we provide a simple explanation for how the kick energy is generated during the merger, and in section 6.2.2 we introduce a slightly more detailed picture of the merger process. This picture will play a central role when we use it for deriving an analytical solution.

6.2.1 ORIGIN OF THE KICK ENERGY

The double scattering mechanism is a process where a particle is gravitationally deflected two times during the merger between its own host halo (H_2) and a target halo (H_1). The first deflection is by the potential of H_1 which is momentarily dominating as the two halos overlap, whereas the second is by the potential of H_2 which dominates at a later stage in the merger process. We will refer to the deflection by H_1 as the 'first deflection' and the subsequent deflection by H_2 as the 'second deflection'. The two merging halos are moving relative to each other during the merger and the two deflections are therefore happening in two different velocity frames. The energy of the particle during each deflection is approximately conserved in the frame of deflection, but because of the velocity difference between the frames, the deflection in one frame can result in an energy change in the other. The second deflection by H_2 therefore provides the explanation for how a given particle can change its energy during the merger relative to H_1.

An N-body simulation of a merger is shown in Figure 6.1, where we have highlighted the orbits of two particles, one that receives a positive kick and thereby escapes (green) and one that gets bound as a result of a negative kick (red). The orange dots illustrate all the particles which prior to the merger are located within 10% of the virial radius of the incoming DM halo (black). We can think about these as the luminous 'galactic component' of the halo. We clearly see that before the merger the green particle is only on a slightly larger orbit than the galactic orange particles, whereas after the merger they have greatly separated as a result of the dynamical kick mechanism. The green particle is now likely to be observed and classified as a 'hostless high velocity object' as e.g. the HVGC-1 (19).

---

[3] The quantum-mechanical solution to this interaction was not found until 1955 (30) due to the fascinating fact that the second Born term is here dominating over the first due to the double scattering nature, or two-step process, of the problem.

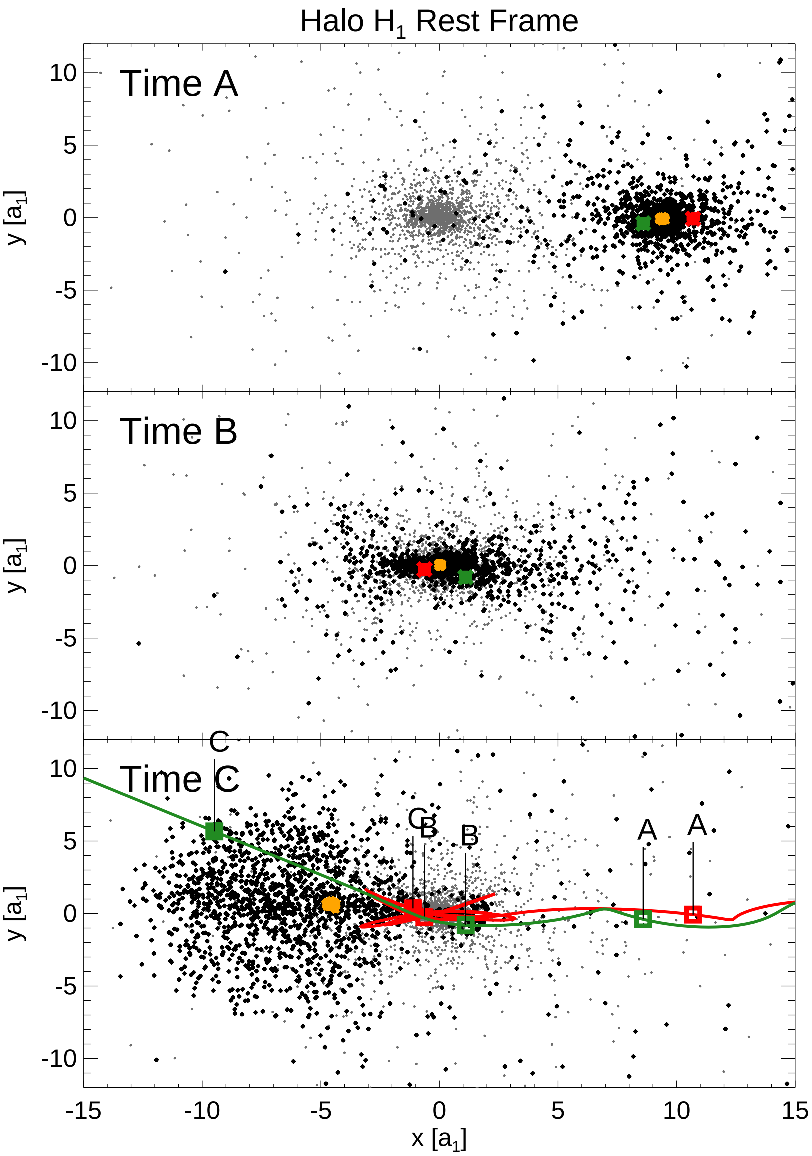


Figure 6.1 Illustration of a $1:10$ merger between two DM halos merging with the escape velocity of the target halo. The particles in the smaller incoming halo are shown in *black* while the particles of the larger target halo are shown in *grey*. The panels from *top* to *bottom* show three different times (A,B,C) during the merger. In the bottom panel the full trajectories of two selected particles are also shown. The *green* particle gains a positive energy kick during the merger and as a result kicked away, whereas the *red* particle looses energy and becomes bound to the target halo. The *orange* dots illustrate the particles which are located within 10% of the virial radius of the smaller halo prior to the merger. These could illustrate a luminous 'galactic component'. The merger has clearly separated the green particle from the orange particles and the green particle could therefore be observed as a 'hostless high velocity object'. This separation and clear velocity difference are due to our proposed double scattering mechanism. Properties of the ejected green particle are shown in Figure 6.3 and 6.4.

6.2.2 A PARTICLE UNDERGOING TWO DEFLECTIONS

We now turn to a more detailed kinematical description of a particle receiving an energy kick by undergoing two deflections. For this we consider a particle initially bound to halo H_2 with an initial orbital velocity $\mathbf{v_0}$ and polar position (l, θ) measured in the rest frame (RF) of H_2. Details are shown in Figure 6.2. During the merger between H_2 and a target halo H_1 the particle receives an energy kick by undergoing two deflections. We consider the case where the mass of H_2 is much smaller than the mass of H_1 and the two halos merge with zero impact parameter. For developing an analytical model for the double scattering mechanism we now approximate the full orbit of the particle by a simpler orbital model which we divide into two phases:

In the first phase, H_2 is passing through H_1 on a radial orbit from $-l$ to $+l$ with radial dependent velocity $w(r)$, where r is the distance between the two halos. The potential of H_1 is greatly dominating and the particle is therefore not moving by the influence of H_2, but instead following an orbit in the potential of H_1 with velocity $\sim w$ and impact parameter $b \sim lsin(\theta)$. The energy is conserved along any orbit in a static potential (10), the energy of the particle in the RF of H_1 is therefore approximately conserved during this phase. In the RF of H_2 the strong perturbation from H_1 results in a large velocity kick $\mathbf{\Delta v}$ of the particle relative to the center of mass (CM) of H_2. This relative velocity arises because the particle and the CM of H_2 are displaced relative to each other and will therefore experience different accelerations during the merger. In our analytical model we assume the particle is not moving relative to H_2 during this first phase. The particle is therefore still at its initial position in H_2 at the end of the phase, but its velocity has changed by an amount $\mathbf{\Delta v}$. This first phase is shown in Figure 6.2 from number $(1) - (3)$.

In the second phase, the two merging halos are moving away from each other at a distance $r > l$. Halo H_1 is now only influencing the particle by a tidal force and the potential of H_2 can therefore again affect the motion. In the RF of H_2 is the particle moving from its initial position (l, θ) with a velocity $\mathbf{v_1}$ composed of its initial velocity and its kick velocity, i.e. $\mathbf{v_1} = \mathbf{v_0} + \mathbf{\Delta v}$. In the limit where the kick velocity is completely dominating the motion, the particle is moving on an almost straight line through H_2 until it reaches the central part where it slightly deflects. This last deflection rotates the velocity vector of the particle $\mathbf{v_1}$ by an angle α into a new vector $\mathbf{v_2}$, as illustrated in Figure 6.2. The energy of the particle is conserved during this deflection in the RF of H_2, but due to the relative velocity between H_1 and H_2 the energy is not conserved in the RF of H_1. The particle can therefore gain or lose energy from the second deflection by H_2, where a forward deflection along the motion of H_2 results in a positive energy kick or vice versa. This second phase is shown in Figure 6.2 from number $(3) - (5)$.

6.3 ANALYTICAL MODEL

We now develop an analytical expression for the dynamical energy kick a particle can receive from our proposed double scattering mechanism. We assume the two merging halos H_1 and H_2 are described by Hernquist (HQ) profiles (45) with an anisotropy parameter $\beta = 0$ and a mass profile given by

$$M_i(r) = M_i \frac{(r/a_i)^2}{(1 + r/a_i)^2}, \tag{6.1}$$

and a corresponding potential

$$\Phi_i(r) = \frac{GM_i}{a_i} \frac{1}{(1 + r/a_i)}, \tag{6.2}$$

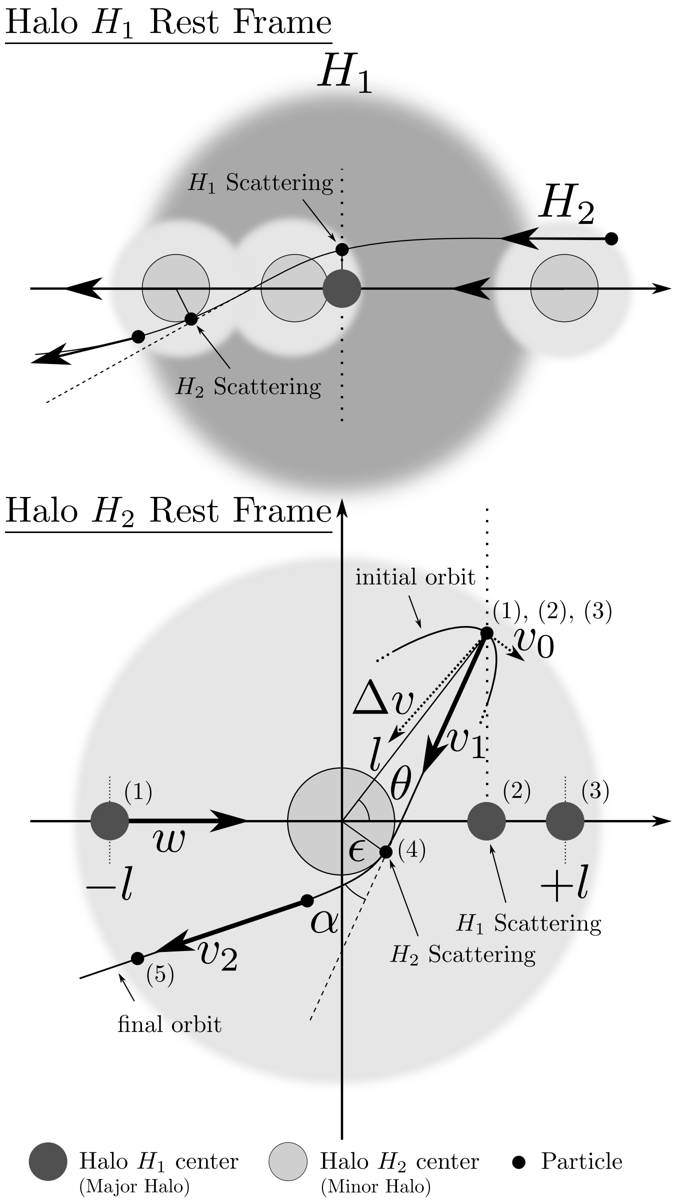


Figure 6.2 Schematic illustration of a particle (black dot) gaining an energy kick from the merger between its original host halo H_2 (light grey) and a target halo H_1 (dark grey). The *top plot* shows the orbital trajectory of the particle in the RF of H_1 while the *bottom plot* shows the trajectory in the RF of H_2. As illustrated, the particle undergoes two clear deflections during its orbit; the first is by the momentarily dominating potential of H_1 while the second is by the potential of its original host halo H_2. The first deflection by H_1 results in a strong velocity change $\mathbf{\Delta v}$ of the particle in the RF of H_2 as illustrated in the bottom plot. The energy of the particle is conserved during this first deflection in the RF of H_1, but not in H_2. The particle now travels through H_2 on its new perturbed orbit where it scatters off the central parts of H_2 at a peri-center distance $\sim \epsilon$. The energy of the particle during this second deflection is conserved in the RF of H_2, but not in the RF of H_1 since H_1 and H_2 are moving relative to each other. As a result of the two deflections the particle gains energy in the RF of H_1 and thereby emerges with a higher velocity. The numbers from (1-5) denote five important moments in the orbital history of the particle and is used for reference in the text.

where M_i is the total mass of halo i, r is the distance from the center, $\Phi(r)$ is the potential at distance r and a is a characteristic scale radius. In the following we occasionally use units of a_1 and we use a 'prime' to denote this, e.g. $x\prime \equiv x/a_1$. By energy conservation we can also write down the relative radial velocity between halo H_i and a particle moving in its potential on a radial orbit

$$w^2(r) = 2\Phi_i(r) + (w^2(0) - 2\Phi_i(0)), \tag{6.3}$$

where $w(r)$ is the radial velocity of the particle at distance r and $\Phi_i(0)$ is the value of the potential at the center. We will use this relation to calculate the relative velocity between halo H_1 and the incoming halo H_2. The estimate for $w(r)$ in the above equation (6.3) ignores the effect from dynamical friction which causes H_2 to loose orbital energy by exchanging momentum with the surrounding particles in H_1 (22). In our case dynamical friction will actually play a minor role when describing particles on large orbits, but for now we will ignore it to keep the analysis as simple as possible. We further assume that the two halos merge with zero impact parameter and that the mass of the target halo H_1 is much larger than the incoming halo H_2, i.e. $M_1 \gg M_2$. This mass hierarchy is relevant for the growth of cosmological halos that are believed to build hierarchically by hundreds of minor mergers (32).

The simulations we compare with along the way are performed using Gadget II (100) with the HQ halos set up in equilibrium by Eddington's Method (31) using a well tested code previously used for studying the anisotropy in HQ halo mergers (97; 98). The halo concentration $c_i \equiv R_{i,\mathrm{vir}}/a_i$ is set to 5 for both halos where the virial radius $R_{i,\mathrm{vir}}$ is calculated by requiring the mean density inside the halo to be 200 times the mean matter density of the universe at redshift $z = 2$ (67; 109). We fix the merger mass ratio at $1 : 10$ for all simulations and the incoming halo H_2 is set to have zero kinetic energy at infinity with respect to the target halo H_1. These ICs are typical in a cosmological perspective (82), however a wide range of both encounter velocities and impact parameters are of course seen in full cosmological simulations (110).

The calculations below are divided into steps based upon our simplified orbital picture described in section 6.2.2. We therefore first calculate the kick velocity $\mathbf{\Delta v}$ the particle receives from its 'first deflection' by H_1. We then use this kick to model the new orbit of the particle through H_2 where it makes a 'second deflection' and is deflected by an angle α. From this deflection we then calculate the resultant energy change of the particle ΔE in the RF of H_1. We further show that the change in angular momentum ΔJ also can be calculated from our model. We can therefore in principle reconstruct the resultant mass profile of the disrupted halo which is particular interesting for the understanding of the growth of DM halos. In the end we discuss the final energy of a dynamically kicked particle and how this energy is affected by the particle's initial kinetic and potential energy.

6.3.1 VELOCITY KICK $\mathbf{\Delta v}$ FROM HALO H_1

The particle receives a velocity kick $\mathbf{\Delta v}$ relative to H_2 because the CM of H_2 and the particle experience different accelerations during the merger. This is illustrated in the top plot of Figure 6.4 which shows the velocity of the particle from Figure 6.3 as a function of time. We can analytically estimate the velocity kick in the impulsive limit where one assumes the particle is not moving during the encounter (see e.g. (50)). In this approximation, the velocity kick is found by integrating over the time dependent force from H_1

$$\mathbf{\Delta v} \approx \int_{-T}^{+T} a(t)\hat{\mathbf{a}}(t)dt = \int_{-R}^{+R} \frac{1}{w(x)} \frac{GM_1(d)}{d^2} \hat{\mathbf{a}}(t)dx, \tag{6.4}$$

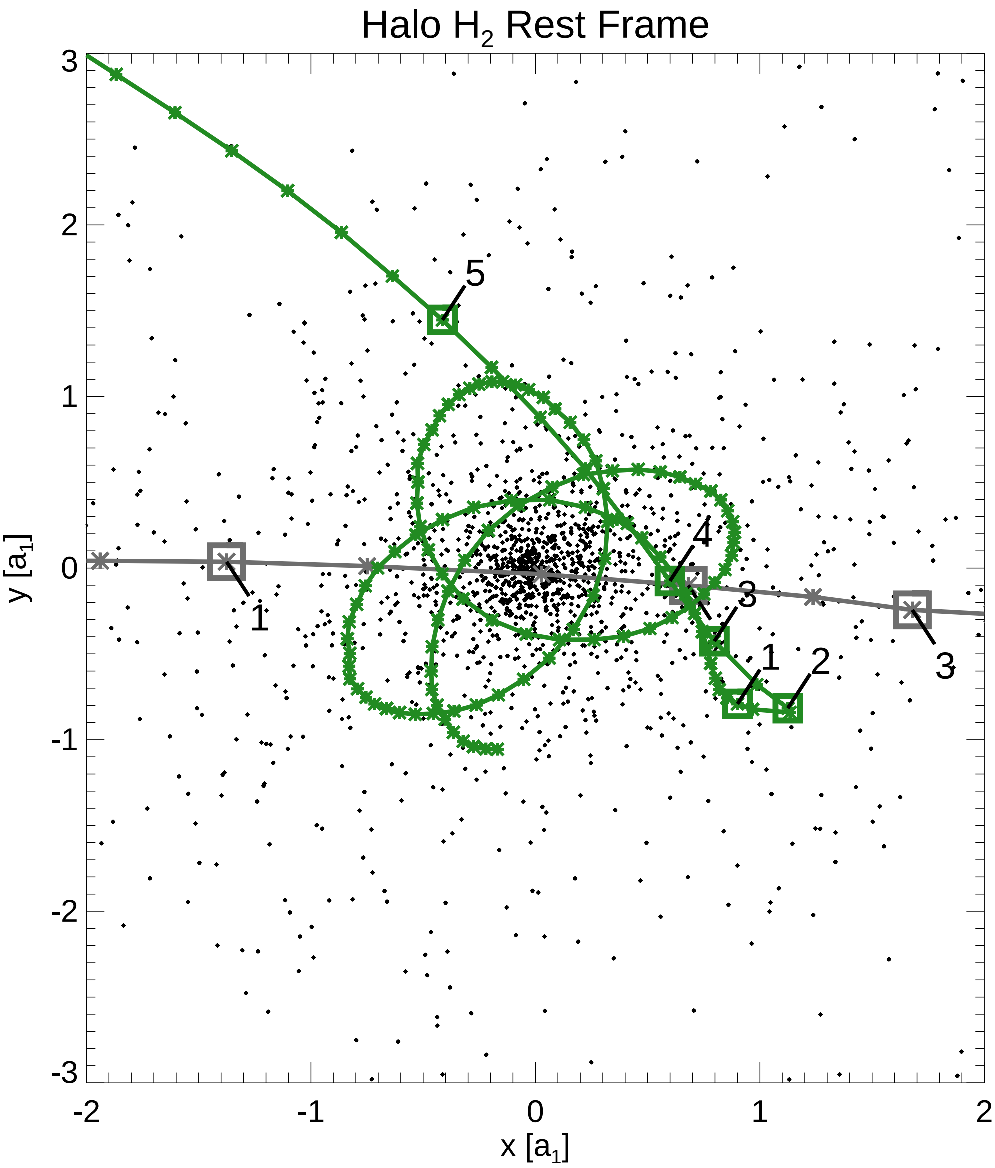


Figure 6.3 Orbital trajectory of a halo particle (*green*) gaining a significant energy kick from the merger between its own original host halo H_2 (*black*) and a target halo H_1 (*grey*). The particle is the same as the green one shown in Figure 6.1, but is here plotted in the RF of H_2 instead of H_1. The smaller black dots are H_2 particles at the time of merger. The horizontal thick grey line shows the orbit of the target halo H_1 which in this frame is moving from left to right. The smaller *stars* on the plotted orbits indicate equal time intervals, where the numbered *squares* indicate five important moments as explained in Figure 6.2. The first deflection by H_1 (phase 1 from 1-3) gives the particle a huge velocity kick $\mathbf{\Delta v}$ pointing towards the upper left corner. Along this new perturbed orbit the particle will undergo a second deflection (phase 2 from 3-5) by the central potential of H_2. This second deflection gives the particle the resultant energy kick as explained in section 6.2.2. Plots of the corresponding time dependent velocity and energy of the particle are shown in Figure 6.4.

where $a(t)$ is the acceleration, $\mathbf{\hat{a}}(t) = (\hat{a}_x(t), \hat{a}_y(t))$ is the corresponding unit vector, d is the distance between the particle and the CM of H_1 and $w(x)$ is the relative velocity between H_1 and H_2. From choosing symmetric integration limits from $-R$ to $+R$, the resulting force on the CM of H_2 is equal to zero. The expression above therefore directly give us the kick velocity of the particle in the RF of H_2. The horizontal and the vertical components of the kick velocity, $\mathbf{\Delta v} = (\Delta v_x, \Delta v_y)$, will be calculated below.

6.3.2 HORIZONTAL KICK VELOCITY Δv_x

Following the orbital model described in section 6.2.2 we are interested in the velocity kick Δv_x the particle receives in the first phase where H_2 is passing through H_1 from $-l$ to $+l$. By integrating equation (6.4) over this range and using the relation $\hat{a}_x = a(d)(x - x_p)/d$, where $x_p = lcos(\theta)$ is the x-coordinate of the particle in the frame of H_2, we find

$$\Delta v_x \approx \frac{\Phi_1(0)}{w(l)} \left(\frac{1}{1 + l'\sqrt{2}\sqrt{1 - cos(\theta)}} - \frac{1}{1 + l'\sqrt{2}\sqrt{1 + cos(\theta)}} \right) \tag{6.5}$$

where we have assumed that w is constant. By comparing with equation (6.2) we see that the expression, except for the $1/w$ term, is exactly equal to the difference in potential energy of the particle between the initial configuration, where H_1 is at $-l$, and the final configuration, where H_1 is at $+l$. This is consistent from the perspective of energy conservation where the particle must receive a kinetic energy kick to 'compensate' for the potential energy difference $\Delta\Phi_{-l,+l}$. To illustrate this we note that in the RF of H_1 the kinetic energy of the particle is $E_{kin}(l) \approx w(l)\Delta v_x$ and from energy conservation the kick must therefore be $\Delta v_x \approx \Delta\Phi_{-l,+l}/w(l)$ as was also found in equation (6.5). The horizontal velocity kick is therefore not due to a real dynamical deflection, but it arises purely from an energy difference. This difference can be calculated exactly and as a result is our estimate for Δv_x also relative accurate. In practice it is useful to approximate equation (6.6) by $\Delta v_x(\theta) \approx \Delta v_x(0) cos(\theta)$, where $\Delta v_x(\theta)$ denotes the solution including the full θ dependence. Using this approximation we find

$$\Delta v_x \approx \frac{\Phi_1(0)}{w(l')} \frac{2l'}{1 + 2l'} cos(\theta) \tag{6.6}$$

In the limit where H_2 passes with the escape velocity of H_1 this reduces to the simple form: $\Delta v_x \approx \sqrt{2\Phi_1(0)} l' cos(\theta) \sqrt{1 + l'}/(1 + 2l')$.

6.3.3 VERTICAL KICK VELOCITY Δv_y

The particle receives a vertical velocity kick Δv_y in the RF of H_2 because it briefly follows an orbit in the potential of H_1 which momentarily dominates when the two merging halos pass each other. The velocity kick Δv_y can therefore be estimated from writing down the orbital solution for a particle with encounter velocity $\sim w$ and impact parameter $\sim lsin(\theta)$ moving in the HQ potential of H_1. However, there are no analytical solutions for the majority of DM density profiles including the HQ profile (10) and we must therefore use the impulsive approximation presented in equation (6.4). Assuming the particle is only deflected by the mass of H_1 enclosed by a sphere with radius $b = lsin(\theta)$, we find

$$\Delta v_y \approx \frac{\Phi_1(0)}{w(x_p')} \frac{2b'}{(1 + b')^2}, \tag{6.7}$$

where w again is assumed constant and $w(x'_p)$ is the relative velocity at the time H_1 has it first closest distance to the particle (at time '(2)' shown in Figure 6.2). In the limit where H_2 passes with the escape velocity of H_1 the above expression reduces to $\Delta v_y \approx \sqrt{2\Phi_1(0)}b'\sqrt{1+x'_p}/(1+b')^2$. In contrast to the horizontal kick Δv_x, the vertical kick Δv_y arises from a real dynamical deflection which makes it hard to estimate precisely. By comparing with simulations we find that our above estimation for Δv_y unfortunately is about a factor of ~ 1.5 too low. One reason for this is that we only include the mass of H_1 enclosed by a finite radius b. However, including the full HQ profile in the integration leads to a divergent result which clearly illustrates the limits of the impulsive approximation.

6.3.4 THE 'SECOND DEFLECTION' BY HALO H_2

After the particle has received the velocity kick $\mathbf{\Delta v}$, it will start to move along a new orbit towards the central region of H_2 where it undergoes a 'second deflection' by the potential of H_2. This changes the velocity vector of the particle from $\mathbf{v_1}$ to $\mathbf{v_1} + \delta\mathbf{v}$ which results in an energy change in the RF of halo H_1. To estimate the components of $\mathbf{v_2}$ we first calculate the impact parameter, ϵ, for the deflection by H_2, as illustrated in Figure 6.2. Assuming $|\Delta v_x/\Delta v_y| tan(\theta) < 1$ we find from simple geometry

$$\epsilon = |x_p(\Delta v_y - \Delta v_x tan(\theta))/\Delta v| \tag{6.8}$$

where $\Delta v^2 \equiv \Delta {v_x}^2 + \Delta {v_y}^2$. Using the relation $\alpha \approx \delta v/\Delta v$ and equation (6.4) to estimate δv, we can now write down an expression for the deflection angle α,

$$\alpha \approx \frac{2GM_2(\epsilon)}{\epsilon \Delta v^2} = \frac{2\Phi_2(0)}{\Delta v^2}\frac{(\epsilon/a_2)}{(1+\epsilon/a_2)^2} \tag{6.9}$$

In the last equality we have inserted the HQ mass profile of H_2 and we assume that the particle is only perturbed by the mass of H_2 enclosed by ϵ. This deflection by H_2 conserves the length of $\mathbf{v_1}$ in the RF of H_2, i.e. $|v_1| = |v_2|$, but rotates it by the angle α into a new vector $\mathbf{v_2} = \mathbf{v_1} + \delta\mathbf{v}$ with components given by

$$\begin{aligned} v_{2,x} &= v_{1,x} cos\alpha + v_{1,y} sin\alpha \\ v_{2,y} &= v_{1,y} cos\alpha - v_{1,x} sin\alpha \end{aligned} \tag{6.10}$$

As we will shortly illustrate, the particle will only receive a positive energy kick in the forward direction if $v_{2,x} > v_{1,x}$, i.e. if the kick velocity Δv and deflection angle α fulfill the inequality $\Delta v_y/\Delta v_x > tan(\alpha/2)$ in the limit where $\Delta v \gg v_0$.

From the definition of $\delta\mathbf{v} = \mathbf{v_2} - \mathbf{v_1}$ we now find the change in velocity due to the second deflection,

$$\begin{aligned} \delta v_x &\approx |\Delta v_y|\alpha \\ \delta v_y &\approx |\Delta v_x|\alpha \end{aligned} \tag{6.11}$$

where we have assumed that $\alpha \ll 1$ and that the kick velocity dominates the motion of the particle along its new perturbed orbit, i.e. $v_1 \approx \Delta v$. This last assumption is not only necessary for the double scattering mechanism to effectively work, but also a very good approximation as we confirmed by simulations.

6.3.5 CHANGE OF ENERGY ΔE

To calculate the change in energy ΔE of the particle we first assume that the second deflection by H_2 happens instantaneously, i.e. the velocity vector of the particle changes from $\mathbf{v_1}$ to $\mathbf{v_2}$

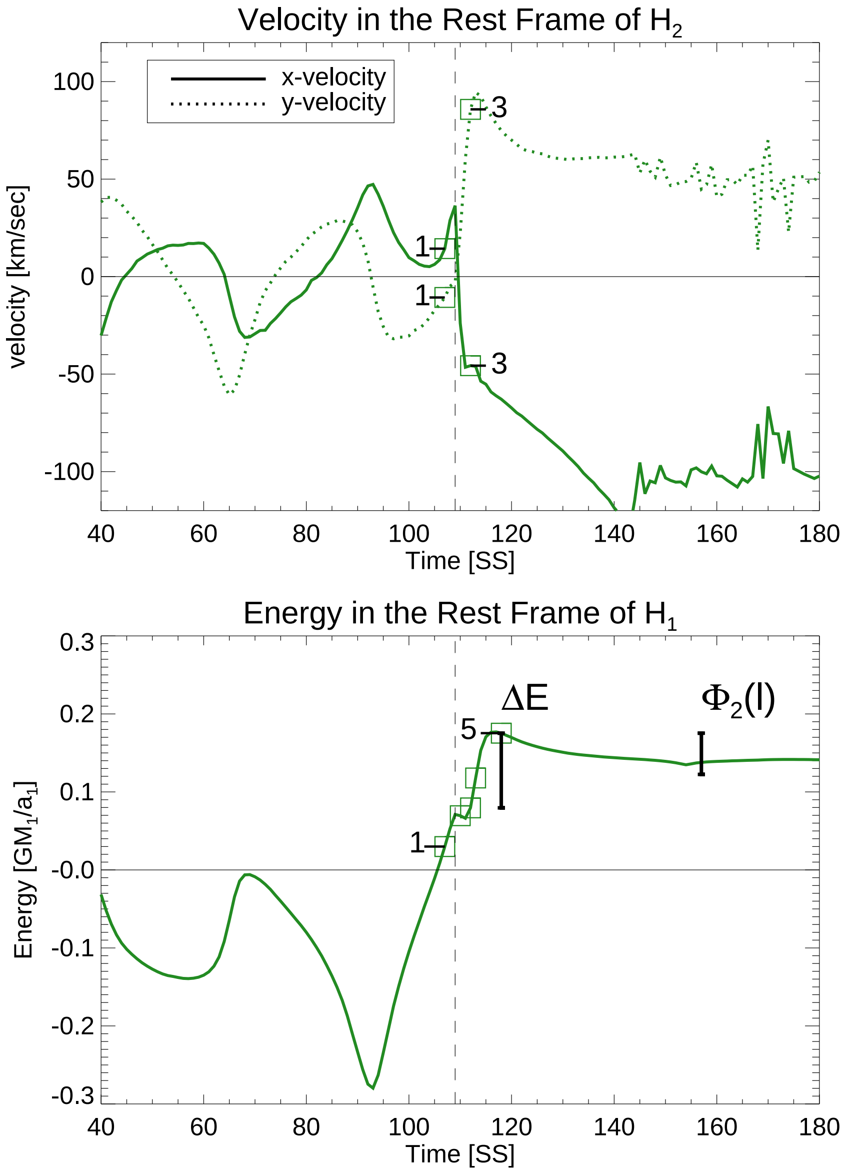


Figure 6.4 Velocity and energy as a function of time for a particle gaining a positive energy kick. The particle is the same as shown in Figures 6.1 and 6.3. The vertical dashed line indicates the time the two merging halos pass each other. *Top:* Horizontal and vertical velocity of the particle in the RF of its original host halo H_2 as a function of time. The numbered squares (1,3) indicate when H_2 passes H_1 from $-l$ to $+l$ as illustrated in Figure 6.2. The particle gains a significant velocity kick $\mathbf{\Delta v}$ during this passage as discussed in section 6.3.1. *Bottom:* Energy of the particle in the RF of H_1 as a function of time. The energy of the particle increases just after the dashed line (time 3-5) by an amount ΔE. Comparing with Figure 6.3, we see that the increase happens when the particle undergoes its second deflection by H_2. This is in complete agreement with our double scattering model. On the same plot we also show the potential of H_2 at distance l, $\Phi_2(l)$. As seen is the particle approximately losing this energy in the RF of H_1 as it travels away from H_2. However, the disruption of H_2 during the merger weakens the potential which makes it slightly easier for the particle to escape.

at a single point. This deflection point will occur when the particle passes the H_2 center at a distance $\sim \epsilon$ as shown in Figure 6.2. From this assumption it naturally follows that the potential energy of the particle is approximately constant during the deflection and the change in total energy will therefore be dominated by the change in kinetic energy. The kick energy ΔE can therefore be estimated by,

$$\Delta E(l,\theta) \approx \frac{1}{2}(\mathbf{v_2} + \mathbf{w}(r'_\epsilon))^2 - \frac{1}{2}(\mathbf{v_1} + \mathbf{w}(r'_\epsilon))^2 = w(r'_\epsilon)\delta v_x, \tag{6.12}$$

where δv_x is the x-component of the velocity change in the RF of H_2 given by equation (6.11), r'_ϵ is the distance between H_1 and H_2 at the time the particle undergoes its second deflection by H_2 and $w(r'_\epsilon)$ is the corresponding relative velocity between H_1 and H_2. In the limit where the two halos pass each other with the escape velocity of H_1, r'_ϵ is found by solving the differential equation $w(r) = dr/dt = \sqrt{2\Phi_1(r)}$. The solution can be written in the form $r'_\epsilon = (3\Delta t\sqrt{2\Phi_1(0)}/(2a_1) + (1+l')^{3/2})^{2/3} - 1$, where $\Delta t \approx l/\Delta v$ is the time from the first deflection by H_1 to the second deflection by H_2.

From our analytical solution given by the above equation (6.12), we can now estimate the dynamical kick energy a halo particle receives from the merger between its own original host halo and a target halo. Figure 6.5 shows results for two $1:10$ merger examples. As seen in the top panel predicts our model that the particles which receive a positive energy kick are all located in a cone with two wings pointing along the velocity vector of H_2. The overall kick energy is also clearly increasing with concentration which is somewhat expected since a larger concentration results in more mass in the center and thereby a larger forward deflection α. From the bottom panel we see that the maximum energy is given to particles located at a distance $\sim 10\% R_{2,\text{vir}}$ from the center of H_2. Stars are typically located within $1-5\% R_{\text{vir}}$ of their DM host halo (58), we therefore expect only the outer parts of a possible stellar halo to be effectively kicked by our proposed mechanism. These outer parts are usually populated by loosely bound stars and stellar systems such as GCs and dwarf galaxies (81). We especially note that GCs are possible to observe at large distances due to their large number of stars ($\sim 10^6$), we therefore suggest these could be observable tracers of our proposed kick mechanism.

A comparison between simulations and our analytical calculation for the kick energy ΔE is shown in the top plot of Figure 6.6. For the analytical calculation we have used equation (6.12) with numerical values for Δv and w to completely isolate the prediction from the double scattering mechanism itself. We see a good agreement and perfect convergence as the kick velocity, Δv, becomes increasingly dominating over the initial motion of the particle, v_0.

6.3.6 CHANGE OF ANGULAR MOMENTUM ΔJ

The change in angular momentum can also be estimated from our model. Using the same approximations as in section 6.3.5 and the definition of angular momentum $\mathbf{J} \equiv \mathbf{r} \times \mathbf{v}$, the change in angular momentum ΔJ from the merger is given by,

$$\Delta J(l,\theta) \approx \mathbf{r}_\epsilon \times (\mathbf{v_2} + \mathbf{w}(r'_\epsilon)) - \mathbf{r}_\epsilon \times (\mathbf{v_1} + \mathbf{w}(r'_\epsilon)) = \mathbf{r}_\epsilon \times \delta\mathbf{v} \tag{6.13}$$

where $\mathbf{r}_\epsilon$ is the vector from the CM of H_1 to the point where the particle is deflected by H_2 at the time of deflection and $\delta\mathbf{v}$ the change in velocity given by equation (6.11). The components of the position vector $\mathbf{r}_\epsilon$ are simply given by $r_x = r_\epsilon - \epsilon\Delta v_y/\Delta v$ and $r_y = \epsilon\Delta v_x/\Delta v$.

We are now in a position to actually predict the orbit of individual particles after the merger since we know both the change in energy and angular momentum, i.e., we can in principle model the whole resultant mass profile of the particles originally bound to H_2. This is highly interesting for understanding how DM profiles are build up over time which has for decades been an active field of research. However, it is beyond the scope of this paper to extend our model for applying it to better understand the growth of DM halos, our focus will therefore only be on explaining the kick energy.

6.3.7 FINAL ENERGY OF THE PARTICLE

The final energy of a particle is generally given by a sum of its dynamical kick energy ΔE and a component coming from its initial kinetic and potential energy. To see this we now consider the lower plot in Figure 6.4. We first notice that the energy of the particle in the RF of H_1 oscillates before the merger. This is because the particle moves on an orbit inside H_2 with a velocity v which translates into an energy change in the RF of H_1 of order $\sim vw$. This energy variation will be comparable to the dynamical kick energy from the double scattering if $v \sim \alpha\Delta v_y$. A slight orbital motion at the time of merger can therefore easily give rise to a significant energy excess. This is also seen in the bottom plot of Figure 6.6, which shows the the final energy (y-axis) as a function of the dynamical kick energy (x-axis) for particles undergoing a double scattering. The majority of the particles escape with a higher energy than their kick energy due to their initial orbital motion.

Right after the particle has been dynamically kicked from the double scattering, we see on the bottom plot in Figure 6.4 that its energy slightly decreases before it reaches a final value. This decrease is due to the fact that the particle has to use energy for traveling out of the potential of H_2 from a distance $\sim l$. The potential the particle experiences is actually slightly lower than the initial unperturbed potential $\Phi_2(l)$ because H_2 is getting disrupted during the merger. This weakens the potential and the particle will therefore need less energy to escape. This explains why the particle energy shown in the bottom plot in Figure 6.4 approaches a slightly higher value than predicted from using the unperturbed $\Phi_2(l)$.

6.4 OBSERVATIONAL SIGNATURES

A clear observational consequence of the double scattering mechanism is that energetically kicked particles will move around with a relatively high velocity compared to the virialized particles bound to H_1. From energy conservation we can easily find the velocity of a particle moving on a radial orbit in H_1 from its total energy,

$$u_{rad}(r) = \sqrt{2(E_i + \Delta E + \Phi_1(r))} \tag{6.14}$$

where $u_{rad}(r)$ is the radial velocity of the particle at distance r, E_i is the initial energy of the particle, $\Phi_1(r)$ is the radial dependent potential of H_1 and ΔE any additional energy contributions. All quantities are here defined in the RF of H_1.

The question is now what the velocity difference is between dynamically kicked particles, particles receiving no kick and the virialized particles of H_1 here characterized by the virial velocity $V_{1,\mathrm{vir}} = \sqrt{GM_{1,\mathrm{vir}}/R_{1,\mathrm{vir}}}$. Figure 6.7 illustrates these velocity differences relative to H_1 when measured at the virial sphere and infinity, respectively. As seen can the energy contribution from the double scattering mechanism generate particles with a velocity $\sim 2V_{1,\mathrm{vir}}$ at the virial sphere of H_1 and $\sim 1V_{1,\mathrm{vir}}$ at infinity. In the section below we will discuss the possibility that that the recently discovered HVGC-1 has been kicked by our proposed double scattering mechanism.

6.4.1 IS HVGC-1 KICKED BY A DOUBLE SCATTERING?

The first detection of a high velocity globular cluster (HVGC-1) was recently reported by (19). This high velocity object was identified as a GC from spectroscopy and *uiK* photometry and was found between GC candidates collected over several years by Keck/DEIMOS, LRIS and MMT/Hectospec (87; 102). The GC is located in the Virgo cluster at a projected distance of ~ 84 kpc from M87 with a radial velocity relative to Virgo and M87 of about 2100 and 2300 km s^{-1}, corresponding to ~ 2 and ~ 7 times their virial velocities, respectively. The interesting question is now, how did this GC get this high velocity? As discussed in the report by the authors, the GC could have been kicked by a SMBH binary located in the center of M87. However, it is very uncertain if a GC can survive this due to the possibility of disruption. Subhalo interactions near M87 could also be an explanation, but no subhalos have been observed in its close vicinity yet. The nature of the kick is therefore still unsolved.

We here suggest that the GC could have been kicked by our presented double scattering mechanism. This would be a possible explanation if the GC was initially bound to a DM halo merging nearly head-on with Virgo. For a $1:10$ mass ratio we have shown that this scenario can generate objects with a radial velocity of ~ 2 times the virial velocity of the target halo at its virial radius. In case of Virgo, this would mean a velocity of about $\sim 2 \times 1100 = 2200$ km s^{-1} (The virial velocity of Virgo is somewhere between $\sim 900 - 1300$ km s^{-1} (102)), which is completely consistent with the observed value.

We also note that HVGC-1 is at present observed to be hostless and isolated. This picture is easily explained by our mechanism if the GC initially was on a slightly larger orbit in its original host halo than the central galactic stellar component. This is a very likely scenario which have been observed for many galaxies (51; 81). Due to its larger orbit the GC could therefore receive a significantly larger energy kick than the central stellar component as we analytically derived and discussed in section 6.3.5. This kick would lead to a large spatial separation between the two systems at the time of observation and thereby explain its observed isolation. This was numerically illustrated in Figure 6.1. The original host galaxy of HVGC-1 should therefore now be among the Virgo galaxies according to this model. Full numerical simulations can of course be used for exploring this in more detail including the role of encounter velocity, halo concentrations and mass profiles, impact parameter and mass ratio. We leave that for a future study.

6.5 CONCLUSION

We present a dynamical double scattering mechanism describing how halo particles can get a significant energy kick from the merger between their original host halo and a target halo. According to this mechanism a particle can gain a significant energy kick by performing two deflections during the merger. The first deflection is by the mass of the target halo while the second is by the mass of the particle's original host halo. The resultant energy kick of the particle arises because the two frames of deflection, i.e. the two halos, are moving relative to each other during the merger.

From our proposed mechanism we analytically derive the dynamical kick energy of a particle as a function of its position in its original host halo just prior to the merger. In case of a $1:10$ head-on merger we illustrate that the particles which receive the largest energy kick are all located in a cone tracing the motion of their host halo. We further find that the largest kick is not given to particles in either the center or in the outskirts, but to particles located at around $\sim 0.1 R_{\mathrm{vir}}$ from their host halo center.

For a $1:10$ merger we estimate the maximum kick energy of a particle to be $\sim 0.3\Phi_1(0)$ which can be translated into a velocity ~ 2 times the virial velocity of the target halo measured at the virial sphere. Our dynamical mechanism is therefore a natural channel for generating high velocity objects. We show e.g. that the high velocity of the recently discovered globular cluster HVGC-1 (19) can straight forward be explained by our double scattering mechanism. We believe this serves as a more natural explanation than other proposed ideas which include 3-body interactions with a possible SMBH binary in M87.

We further note that it is possible from our model to not only calculate the change in energy of individual particles, but also the change in orbital momentum. It is therefore in principle possible to estimate the full resultant mass profile of the incoming disrupted halo, which is highly interesting in the perspective of structure formation and evolution of DM halos. However, the merger history of cosmological DM halos is highly non trivial and we must make major extensions to our model to make any real predictions for this case.

It is pleasure to thank S. H. Hansen, J. Zavala, R. Wojtak, M. Sparre, J. Hjorth, S. Pedersen, J. Fynbo, J. Zabl, Kristian Finlator, Sebastian Honig, M. MacLeod for comments and helpful discussions. The Dark Cosmology Centre is funded by the Danish National Research Foundation.

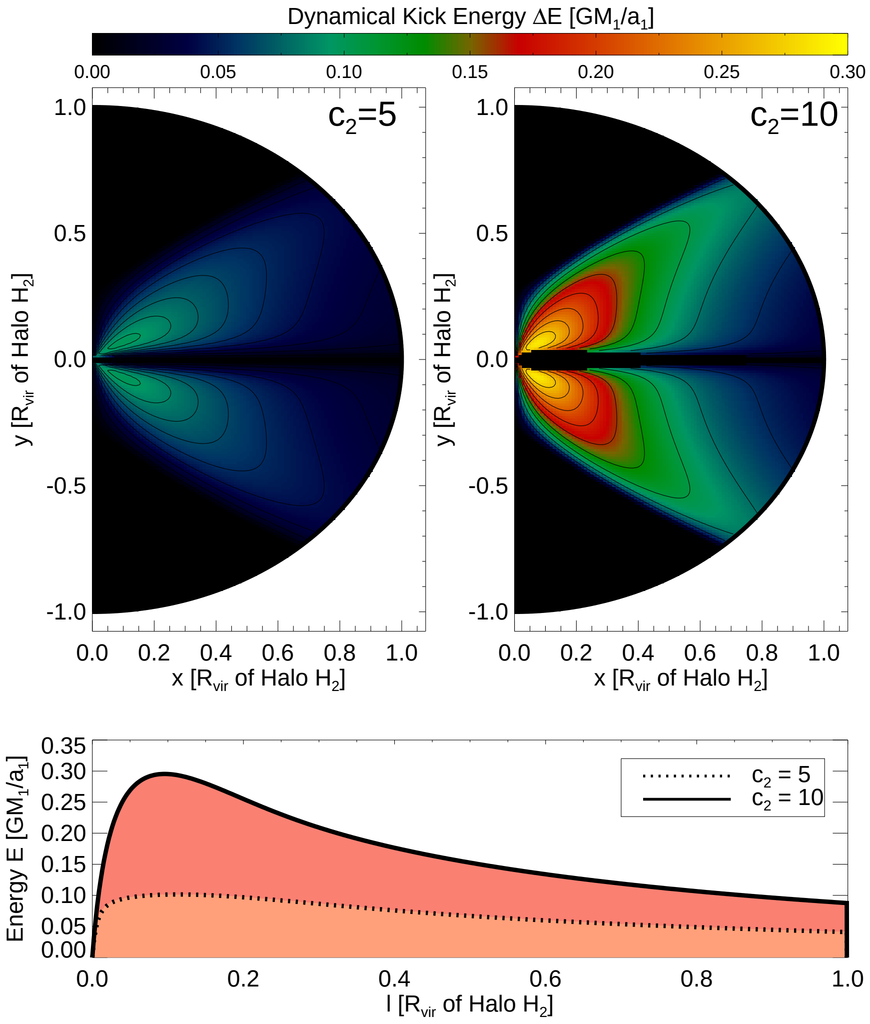


Figure 6.5 Dynamical kick energy ΔE calculated from our proposed double scattering model for a $1:10$ head-on merger between two HQ halos passing each other with the escape velocity of the larger halo. We show the result of two examples, one where the halo concentrations are equal $(c_1 = 5, c_2 = 5)$ and one where the incoming halo H_2 is much more compact, $(c_1 = 5, c_2 = 10)$. We see that a higher concentration leads to a higher kick energy. *Top*: Contour plots showing our theoretical calculated kick energy $\Delta E(x, y)$ given by equation (6.12), as a function of the position (x, y) of the particle in H_2 prior to the merger. The plots only show the right hand side of the incoming halo H_2 that in this picture is moving from right to left towards H_1. The particles which receive a positive energy kick are all located in a cone pointing along the motion of H_2. The left plot is for $c_2 = 5$ and the right is for $c_2 = 10$. Both contour plots only include regions where the kick energy is positive. *Bottom*: The maximum kick energy $\Delta E(l)$ as a function of the particle distance $l = \sqrt{x^2 + y^2}$ from the CM of H_2. The black solid line shows the maximum ΔE for $c_2 = 10$ where the black dotted line is for $c_2 = 5$. In both plots we have used $1.5\Delta v_y$ to correct for the known bias as explained in section 6.3.3.

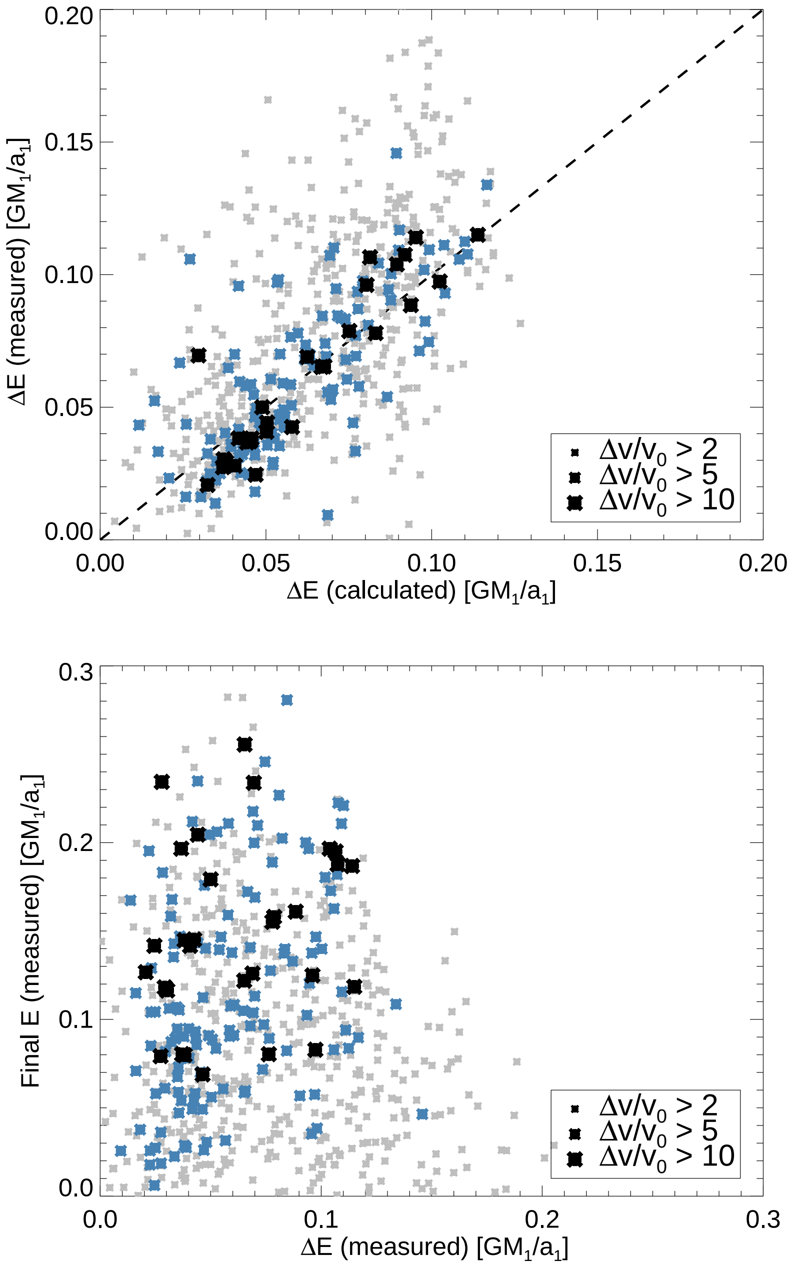


Figure 6.6 Dynamical kick energy and final escape energy. *Top:* Comparison between our analytical calculation for the dynamical energy kick (x-axis) and the 'true' simulation values (y-axis). The three symbol sizes indicate different thresholds between the dynamical kick velocity, Δv, and the peculiar motion of the particle, v_0, at the time of merger. We see that our model performs successfully and converges as expected as Δv becomes increasingly dominating over v_0. For the analytical calculation we have used equation (6.12) with measured values for Δv and $w(r)$ to completely focus on how the mechanism performs. *Bottom:* Scatter plot of the numerical measured kick energy (x-axis) and the final energy of the particle (y-axis). The majority of the particles which clearly undergo a double scattering (blue/black) end up with an energy equal or larger than their kick energy. As explained in section 6.3.7, the extra energy contribution is coming from slight peculiar motions of the particles of order $v_0 \sim \alpha \Delta v_y$. The data used for this figure is the same as used in Figure 6.1.

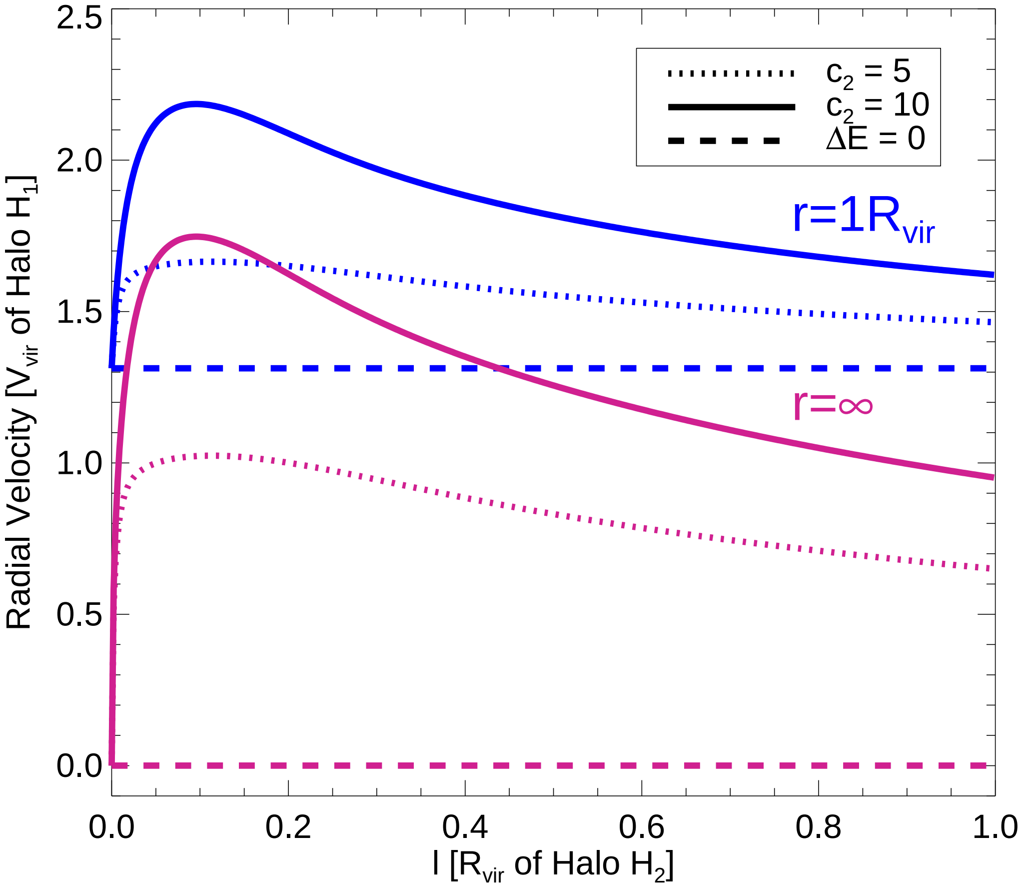


Figure 6.7 The solid (dotted) lines show the maximum radial velocity a dynamical kicked particle will have after the merger at distance r relative to H_1 for a $1:10$ merger with $E_i = 0$ and concentration $c_2 = 5$ ($c_2 = 10$). The dashed lines show the velocity if the particle do not receive any kick, i.e. when $\Delta E = 0$, and therefore just fall unperturbed through the target halo H_1. The velocity is calculated from equation (6.14) using our analytical estimate for the kick energy ΔE given by equation (6.12). The x-axis is the radial position of the kicked particle in its original host halo H_2 just prior to the merger. The figure is similar to the lower plot in Figure 6.5 except that here is the energy kick converted to a radial velocity which is a more direct measurable quantity. As seen can our proposed double scattering mechanism generate a population of dynamical 'hot' objects that far away from the target halo will move around with a velocity about 1-2 times the virial velocity of the target halo, $V_{1,\text{vir}}$. At distances closer to the target halo H_1 can kicked particles be observed with velocities exceeding ~ 2 times $V_{1,\text{vir}}$.

7

THE SHAPE OF GALAXY CLUSTERS AND X-RAY OBSERVATIONS

MEASURING THE 3D SHAPE OF X-RAY CLUSTERS

Johan Samsing[1], **Andreas Skielboe**[1], **Steen H. Hansen**[1]

[1] Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark

Abstract

Observations and numerical simulations of galaxy clusters strongly indicate that the hot intracluster x-ray emitting gas is not spherically symmetric. In many earlier studies spherical symmetry has been assumed partly because of limited data quality, however new deep observations and instrumental designs will make it possible to go beyond that assumption. Measuring the temperature and density profiles are of interest when observing the x-ray gas, however the spatial shape of the gas itself also carries very useful information. For example, it is believed that the x-ray gas shape in the inner parts of galaxy clusters is greatly affected by feedback mechanisms, cooling and rotation, and measuring this shape can therefore indirectly provide information on these mechanisms. In this paper we present a novel method to measure the three-dimensional shape of the intracluster x-ray emitting gas. We can measure the shape from the x-ray observations only, i.e. the method does not require combination with independent measurements of e.g. the cluster mass or density profile. This is possible when one uses the full spectral information contained in the observed spectra. We demonstrate the method by measuring radial dependent shapes along the line of sight for `CHANDRA` mock data. We find that at least 10^6 photons are required to get a $5-\sigma$ detection of shape for an x-ray gas having

realistic features such as a cool core and a double powerlaw for the density profile. We illustrate how Bayes' theorem is used to find the best fitting model of the x-ray gas, an analysis that is very important in a real observational scenario where the true spatial shape is unknown. Not including a shape in the fit may propagate to a mass bias if the x-ray is used to estimate the total cluster mass. We discuss this mass bias for a class of spacial shapes.

7.1 SUMMARY

The largest relaxed dark matter (DM) structures we see today in our universe are galaxy clusters containing baryons in the form of stars and gas. The gas represents around $15-20\%$ of the total mass and is heated to a virial temperature of $\sim 10^7$ K resulting in strong X-ray emission. If the cluster is relaxed the gas will be in hydrostatic equilibrium, which make it possible to estimate the cluster mass profile and its spacial shape through observations of the X-ray (90). Both the cluster mass and the shape play an important role in constraining the cosmological parameters. However, including a non-spherical shape is problematic mainly due to degeneracies in the observations between the shape and the density profile. The shape is therefore often assumed spherical, an assumption generally leading to a wrong cluster mass estimation at the $\sim 10\%$ level and no measurement of the shape.

To address this problem, I developed a new numerical method to measure the cluster mass and shape. With this method, I am in particular able to measure radial shape variations, which is in contrast to previous studies. This becomes possible by using the full spectral information, taking into account that the observed spectrum is a sum of different spectra along the line of sight.

To first study the possibility for measuring the shape, I emulated an observation of a realistic X-ray cluster gas with cool core, double power law density profile and a radial dependent shape with ellipticity $\epsilon(r)$ following the form $\epsilon(r) = \alpha log(r) + \beta$. For this I created a full numerical package to fit CHANDRA X-ray spectra including X-SPEC as spectrum generator. By performing an MCMC analysis I estimated a minimum of $\sim 10^6$ photons are needed for a 5σ detection of all cluster parameters including the shape. This suggests a radial shape could be measured for e.g. A1689 with slightly improved data. I also investigated how the choice for the input parametrizations that my method relies on impacts the estimation for shape and mass. For this I did a Bayesian model comparison analysis and showed, among other things, how an input function with slightly insufficient freedom to model the true density profile easily results in a fairly biased shape estimation.

7.2 INTRODUCTION

Galaxy clusters are the largest bound objects in the universe and they provide unique and independent information on the cosmological evolution. The standard LCDM parameters and a possible redshift varying dark energy component has accurately been measured and constrained from cluster observations in a variety of ways ((108), (2), (66), (3), (107)). They reveal the distant universe behind them through gravitational magnification ((57), (4), (13)), and they are even sensitive to the initial perturbations of our universe ((33), (25), (91)). Clusters not only serve as excellent laboratories for constraining the standard cosmology, but because of their relative high mass and cosmological size they also provide a unique possibility to test general relativity itself in several independent ways, e.g. from measurements of cosmic growth ((85)) to gravitational redshift ((112)) and gravitational waves ((114)). Other probes have also been suggested such as lensing, cluster abundance and the integrated Sachs-Wolfe effect ((52)). Despite their importance in modern cosmology, basic properties such as spatial shape is still not well measured for individual clusters. One reason is simply that the main part of a cluster is composed of dark matter which can only be measured indirectly by its gravitational interaction. The indirect measurements of the dark matter and its radial distribution are usually done using either lensing ((80), (101)), by studying the dynamics of the intracluster galaxies ((113), (62), (61)) or by the hot baryonic x-ray emitting gas located in the inner regions of all clusters (for a review of x-ray physics and applications see e.g. (90)). Especially observations of the intracluster x-ray gas in terms of spacial shape, density and temperature profiles, play a key role for estimating local properties of the cluster. Many earlier studies assume a spherical shape of the gas ((78), (49), (56), (75), (42), (85)), however there are several strong motivations why a precise estimation of the shape is interesting. One is a precise estimation of the cluster mass profile. This profile can directly be measured if the radial shape, temperature and density profiles of the gas are known and the gas is in hydrostatic equilibrium. Only recently it was shown that allowing the gas to have a triaxial shape is necessary for the estimated mass profile from x-ray to agree with the mass estimated from lensing ((70), (94), (69)), a result in good agreement with numerical simulations ((43)). This overall triaxiality is mostly due to the underlying shape of the dark matter potential. However, in the central cluster regions it is believed that a possible non-spherical x-ray shape is more affected by microphysical processes such as radiative cooling, turbulence and different feedback mechanisms ((60)) than the dark matter potential shape is. These mechanisms change the gas shape into having relatively high ellipticity towards the center compared to the underlying dark matter potential shape. It is therefore possible to infer properties of these mechanisms if the shape of the gas, temperature and density profiles are known to high precision.

In this paper we suggest and develop a method from which a possible radial dependent shape of an x-ray gas can be extracted from the x-ray observations only. We explicitly demonstrate the possibilities for measuring the shape by fitting to `CHANDRA` mock data and we estimate the mass bias if a shape is not treated correctly in the fitting. The method we use is a parametrized approach, i.e. we assume that the shape and profiles can be described by a set of well defined functional forms. We also discuss the complications of choosing the best set of functions, i.e. a model, to describe the data.

The paper is organized in the following way; The method for measuring shape is explained in section 7.3.1. We apply the method in section 7.4 on `CHANDRA` mock data. We discuss how to quantify the goodness of fit in section 7.4.3. Mass bias from not including the shape in the fitting is discussed in section 7.5.

7.3 EXTRACTING 3D X-RAY INFORMATION FROM 2D OBSERVATIONS

An intracluster x-ray emitting gas has a three dimensional extension, spherical or not, but an observer will only see the two dimensional projected image on the sky. Therefore, a given observed spectrum is a sum of all emission spectra along the line of sight through the gas (for a discussion see figure 7.1).

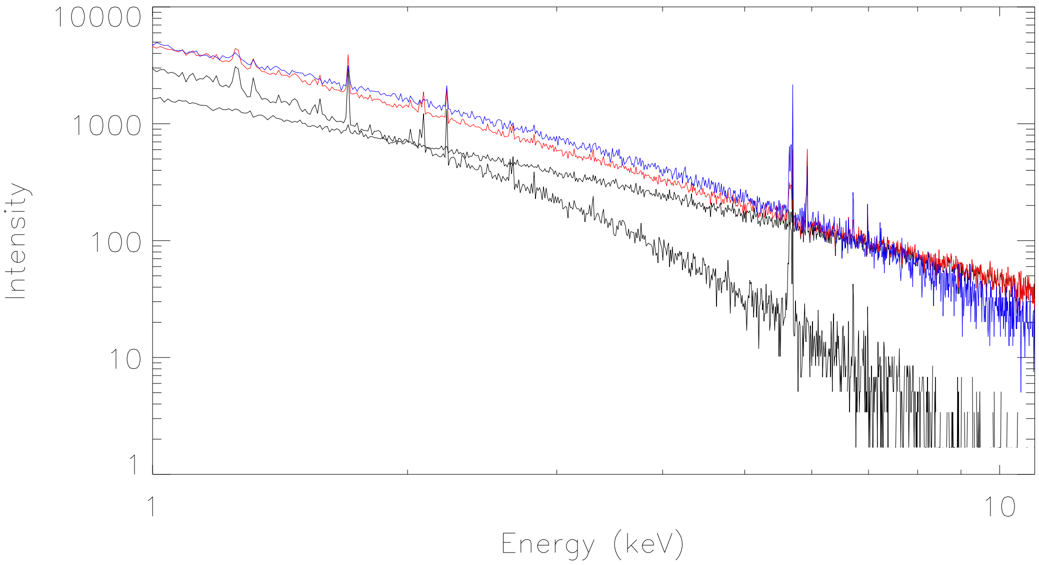


Figure 7.1 Simplified description of how information propagates from three dimensions to a two dimensional observation. The figure shows two free-free emission X-ray spectra (the two lowest spectra at 3 keV in black), the sum of these two spectra (second spectrum from the top at 3 keV in red), and a best fit free-free spectrum to the red spectrum (upper spectrum at 3 keV in blue). This could correspond to a part of a gas with two temperature components (one spectrum for each component) projected along the line of sight (observed data). By comparing the blue and the red spectra one notices that the red spectrum cannot simply be fitted accurately with a free-free emission X-ray spectrum. This is true in the general case; a sum of free-free spectra cannot in general be fitted by another free-free spectrum. If the observed red spectrum is not correctly fitted one could incorrectly conclude the presence of either a *non thermal hard X-ray excess* or *non thermal soft X-ray excess* component, depending on the shape of the red spectrum. In our example one would report a *hard excess* component since the red spectrum is above the blue in the tail (see e.g. (36) for a hard excess discussion for Abell 2256).

Each spectrum has a spectral shape determined by the local temperature and a scaling proportional to the local density squared ((90)). Mathematically, no unique mapping can construct the true three dimensional shape, density and temperature profiles using only the observed two dimensional image. However, if one makes prior assumptions it can be done. For instance by assuming that the gas is spherical the density and temperature profiles can be found. From this assumption several previous groups have measured the temperature and density profiles of the x-ray gas using either projection or de-projection techniques (see e.g. the `XSPEC` packages 'deproject' and 'projct')

The method we present in this paper for extracting three-dimensional information relies on the assumption that the x-ray gas shape, density and temperature profiles can be described by parametrizations. This means that shape and profiles are believed to be well described by a set of functions. In contrast to several previous studies we use the whole spectral information from the integrated observed picture of the x-ray gas. It means that we take into account that the actual observed spectra is a sum of spectra along the line of sight, and can not simply be fitted by a single free-free spectrum. See figure 7.1 for a discussion. We allow a radial dependent shape in contrast to previous studies. The tradeoff for including this extra freedom is that we limit our analysis to structures that are seen spherical on the sky. This is for purely practical reasons: in theory the fitting method we describe is not limited by this assumption, but with present day available data it is simply not possible to resolve a radial dependent shape if the 3D-shape and orientation of the gas is completely free to vary. In other words, the symmetry in the sky makes it possible to extract higher order corrections to the usual assumption about either sphericity or triaxiality with constant axis ratios. The method and procedure will be described in the following sections, and technical details are found in the appendix together with illustrations of generated spectra and an x-ray structure.

7.3.1 FITTING SHAPE AND PROFILES USING THE PARAMETERIZATION APPROACH

The procedure needed in order to measure spatial shape, temperature and density profiles of an observed x-ray gas using the parametrization approach is as following: First we choose a model, i.e. a set of parametrizations, that are believed to generally describe the form of density, temperature and spatial shape (along the line of sight) for the observed structure. The chosen parameterizations must be sufficiently general to accurately describe observations of real and simulated structures. We then calculate the agreement between an artificial generated dataset (see appendix section .1 for how we generate artificial datasets and mock data) created from the chosen model given a specific combination of parameter values and the observed dataset. In our case we quantify the agreement by a simple χ^2 statistic which simply can be related to a probability by $exp(-\chi^2/2)$ when the noise is gaussian. This routine of comparing artificial generated datasets with the observed dataset is then repeated for a wide range of parameter value combinations until a good estimate of the underlying probability distribution function (PDF) for our model has been made. For this we use standard Monte Carlo techniques as described in section .1 in the appendix. From the parameter combination having the maximum PDF value, the best estimate for profiles and shape, given our prior input parameterizations, can then be made. The overall procedure can then be repeated for different models, until the best model is found. We will discuss this in more detail in section 7.4.3.

7.4 RESULTS FROM FITTING SHAPE AND PROFILES OF SELECTED X-RAY MODELS

In the following we show the possibilities of measuring radial profiles of non-spherical x-ray structures with varying radial dependent shape along the line of sight. As briefly discussed in the end of section 7.3, we only consider structures that are spherical on the sky. We consider two simulated structures in our analysis; First a simple toy model to clearly illustrate the method, and second a more realistic model with features such as a cool core and a double powerlaw for the density profile. The shape parameterizations are described later. In this part of the

Table 7.1. Simulated profiles

Model	equation	n_0	r_c	β	$\alpha/2$	T_0	a	b	r_t	s_1	s_2
shM1	1,2,3	$\cdots$	0.11	0.6	0	5.0	0	0.14	0.09	0.3	0.83
shM2	4,6,7	$\cdots$	0.15	0.76	1.2	4.3	2.45	0.7	0.13	0.94	0.2

analysis we fit for profiles and shape using the same set of parameterizations that are used to generate the data. In this way we get the cleanest picture of how a shape signal propagates to observables.

We present results in terms of a virial radius r_v. The shape, temperature and density profiles we use, are consistent with a virial radius similiar to r_{500} ((106)).

7.4.1 A SIMPLE TOY MODEL

We consider a dataset denoted by 'shM1' where the density and temperature profiles are modeled by simple broken powerlaws

$$\rho(r) = n_0(1 + (r/r_c)^2)^{-3\beta/2} \tag{7.1}$$

$$T(r) = T_0(1 + (r/r_t)^2)^{-b} \tag{7.2}$$

known as beta-models. Parameter n_0 acts as a normalization factor and is regulated such that the artificial dataset has a fixed number of total (photon) counts. The shape parametrization we consider is a simple linear function for ellipticity

$$\epsilon_1(r) = s_2 \cdot r + s_1 \tag{7.3}$$

where $\epsilon \equiv b/a$ is defined as the ratio between the radius perpendicular to the observer (b) and the radius along the line of sight (a) of the observer. The parameter values for shM1 are listed in table 7.1, and figure 7.2 shows the corresponding shape and profiles. The chosen parameters for the density and temperature profiles are in fair agreement with typical observed values. The priors on the shape parametrization we use in this example are: a) $0.2 < b/a \leq 1$ and b) $\epsilon_1(r/r_v = 1) > 0.5$. In general, a structure could naturally have an axis ratio $b/a \geq 1$ and still be spherical on the sky, and therefore in a scenario where no prior shape information is available, shapes with $b/a \geq 1$ must be included in the fit as well.

The left plot of figure 7.3 shows the maximized PDFs for the fitted density, temperature and shape parameters for a total of $3 \cdot 10^4$ photon counts ($\approx$ 10 ks `CHANDRA` observation of A1689). The width of the projected PDFs, i.e. a measure of the fitting error for each parameter, is simply related to the number counts by $\sim 1/\sqrt{N}$ where N is the number of photons. The right plot of figure 7.3 shows the corresponding correlation matrix defined in the usual way as $CORR(X, Y) = \frac{\langle (X-\mu_X)(Y-\mu_Y) \rangle}{\sigma_X \sigma_Y}$ where X, Y are random variables with expectation values μ_X, μ_Y and standard deviations σ_X, σ_Y. In our case, to find e.g. the correlation coefficient $CORR(p_i, p_j)$ X must be replaced with the vector of MCMC sampled p_i values and Y of sampled p_j values. The correlation matrix is symmetric by construction and the shading goes from 0 (black) to 1 (white). The correlation matrix can be divided up in several regions. On the plot

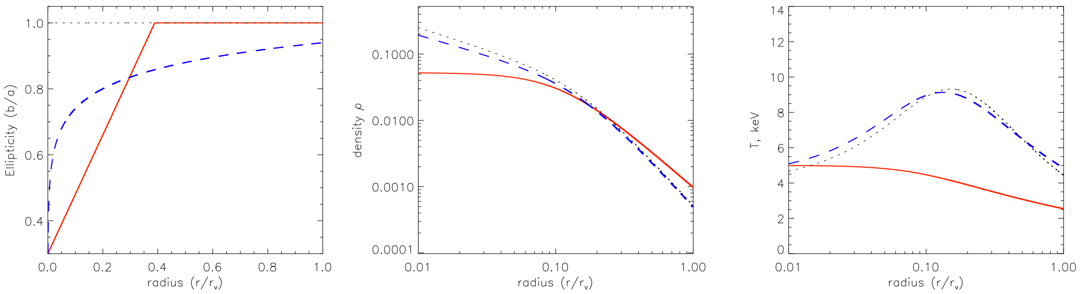


Figure 7.2 Shape and profiles for dataset shM1 (red, solid lines) and shM2 (blue, dashed lines). Left: Ellipticity of the x-ray gas along the line of sight. Central: x-ray gas density profile. Right: x-ray gas temperature profile. Black, dotted lines: Best estimate from a MCMC fitting to shM2 using the true parametrizations for temperature and density, but assuming spherical symmetry.

is highlighted a region bounded by a dotted line and a solid line. The dotted region is the part which shows the correlation between temperature and density and the solid is the region that shows the correlation between temperature and shape. In the lower left corner of the correlation matrix is the region showing the correlation between shape and density. In general we see that the temperature is weakly correlated with the rest of the parameters, especially when compared to the correlation between shape and density. The physical reason is simply that their individual contributions to a spectrum by nature are completely different; temperature affects the spectral form, but density and shape affect only the normalization. This is clearly seen in the analytic form for the bremsstrahlung spectrum $I(T, \rho, \nu) \propto \rho^2 T^{-1/2} exp(-\nu/T)$ ((90)).

Among our chosen priors, the prior on s_1 ($s_1 > 0.2$) is the one that affects the shape of the PDFs the most. Besides a trivial truncation on the s_1 parameter axis it is also responsible for especially the truncation (or skewness) of the n_0 distribution. The reason is the relative strong correlation between these two parameters. This correlation is clearly seen on the correlation matrix and can be understood in the following way: The degree of constant ellipticity captured by s_1 effectively acts as mass scaling term when the structure is projected along the line of sight. This is simply because an ellipticity "stretches" the structure and therefore "allows" more mass along the line of sight. This is exactly how n_0 affects the projected dataset too. So if we increase the overall scaling (increasing n_0) we can compensate by decreasing the ellipticity (increasing s_1), that means the lower truncation of s_1 also shows up as a lower truncation on n_0. In fact, a constant ellipticity along the line of sight ϵ is completely degenerate with the overall density scaling ρ_0 by $\rho_0^2\epsilon$. This is an intrinsic degeneracy and can only be broken by including other observations, e.g. SZ observations which effectively traces $\rho_0 \epsilon T$ (see e.g. (77), (29), (26), (93)).

The overall conclusion from the fitting results is that the parameter values specifying the true shape as well as temperature and density are exactly reconstructed. This is an ideal case, but it is clearly showing that temperature, density and shape in principle can be separated.

From the correlations we can conclude that the temperature profile is well and almost independently fitted. In perspective of optimizing the fit for shape, this also implies that independent measurements of the density will directly result in a better fit for the shape.

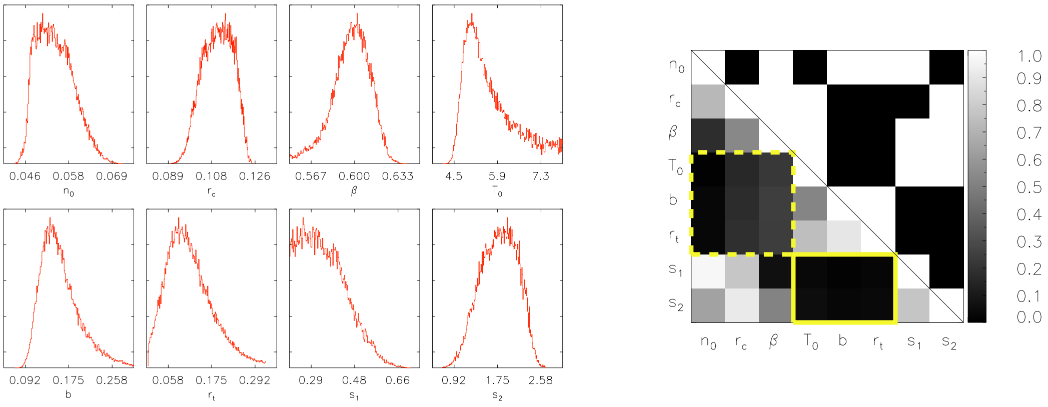


Figure 7.3 Fitting results for dataset 'shM1'. Left: Maximized PDFs along each of the 8 parameters in the model. Right: Correlation matrix for the 8 parameters in the model. The lower left part of the matrix shows the absolute value of the correlation coefficient where the upper right corner shows the sign of the coefficient in black (negative) and white (positive). The results are based on a dataset scaled to have a total of $3 \cdot 10^4$ photon counts.

7.4.2 A MORE REALISTIC MODEL

We now perform an analysis on a dataset, denoted by 'shM2', describing a structure with cool core and a double powerlaw for the density profile. Including these features are motivated by real observations ((106)). The temperature and density profiles are now parameterized as

$$\rho(r) = n_0 \frac{(r/r_c)^{-\alpha/2}}{(1+(r/r_c)^2)^{3\beta/2-\alpha/4}} \tag{7.4}$$

$$T(r) = T_0 \frac{1+a(r/r_t)}{(1+(r/r_t)^2)^b}, \tag{7.5}$$

and the shape is parameterized by

$$\epsilon_2(r) = s_2 \cdot log_{10}(r) + s_1 \tag{7.6}$$

This shape parametrization approximately describes the gas shape seen in the inner parts ($r \leq r_{500}$) of clusters in numerical simulations ((60)). We use the same shape priors as used in the previous toy model example. A list of temperature and density parameterizations are found in ((106)).

The true parameter values for 'shM2' are listed in table 7.1 and the corresponding shape and profiles are plotted in figure 7.2. Figure 7.4 shows the PDF and the correlation matrix for the 10 parameter model fitting.

An inner density slope captured by α is now one of the new parameters compared to the toy model. Since both the shape and this inner slope have a logarithmic dependence, there is a strong correlation between α and s_2. This is clearly seen in the correlation matrix and the PDF plot where the lower cut on s_2 directly relates to the skewness in the α distribution. This freedom in the inner slope is the main reason for the fitting to require many more photons than the toy model. This is discussed in more detail in section 7.4.3 below.

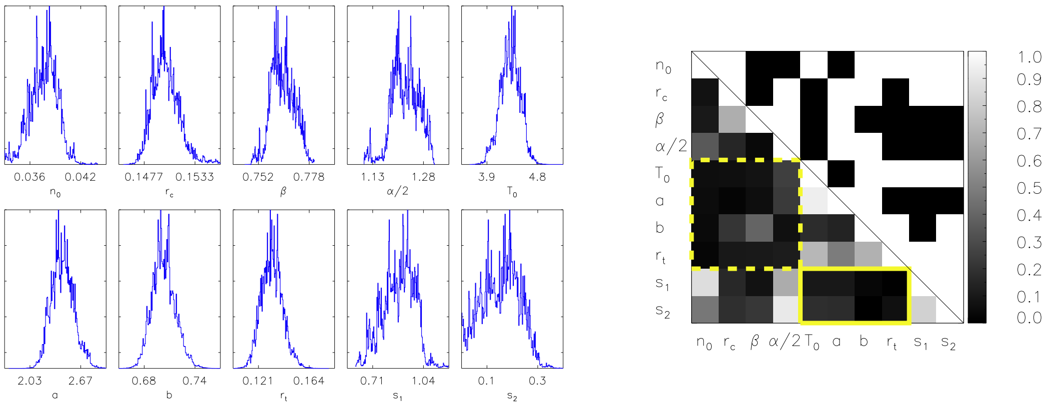


Figure 7.4 Fitting results for dataset 'shM2'. Left: Maximized PDFs along each of the 10 parameters in the model. Right: Correlation matrix for the 10 parameters in the model. The lower left part of the matrix shows the absolute value of the correlation coefficient where the upper right corner shows the sign of the coefficient in black (negative) and white (positive). The results are based on a dataset scaled to have a total of $1 \cdot 10^6$ photon counts.

The overall conclusion is that the true parameter values are reconstructed, but to keep down the statistical errors a relative high number of photons are required. This is mostly due to the similar parameterizations for shape and density. In agreement with intuition, we see that it is much harder to extract a logarithmic shape when the density is varying logarithmically too, compared to e.g. a linear dependent shape. From the correlation matrix we see that the temperature fitting is nearly unaffected as we also concluded in the previous toy model example.

7.4.3 QUANTIFYING THE GOODNESS OF FIT

In this section we will discuss how to quantify the goodness of fit for the parameters within a given model, as well as the goodness of fit for the model itself relative to other competitive models.

Individual parameters within a model

The best fit parameter values for a given model are located at the likelihood maximum, or the minimum χ^2 if the measurement noise is gaussian. To quantify the goodness of the fit is not unique in the same way. To quantify this one must often combine statistical estimators with prior knowledge. An often used estimator is the reduced chi square, $\chi^2_{red} = \chi^2/K$, where K is the number of degrees of freedom. However, this estimator has two major problems. First, χ^2 itself have a significant noise due to random noise of the data, and second, the number of degrees of freedom is not in general well defined ((5)). Another, maybe more intuitive, estimator is the the ratio $\hat{p}_i/\sigma_i$ where $\hat{p}_i$ is the best estimate for parameter p_i and σ_i the associated standard deviation. If we denote this ratio by n we can quantify the goodness of fit by reporting n for each parameter or the minimum n for the whole model. For the fitting examples we presented above, it is then of interest to know the number of photons required for e.g. a minimum $n = 5$ (or $5 - \sigma$) detection for all parameters. We will investigate this in the following. Figure 7.5 shows the ratio $\hat{p}_i/\sigma_i$ as a function of total photon counts for the parameter that have the largest

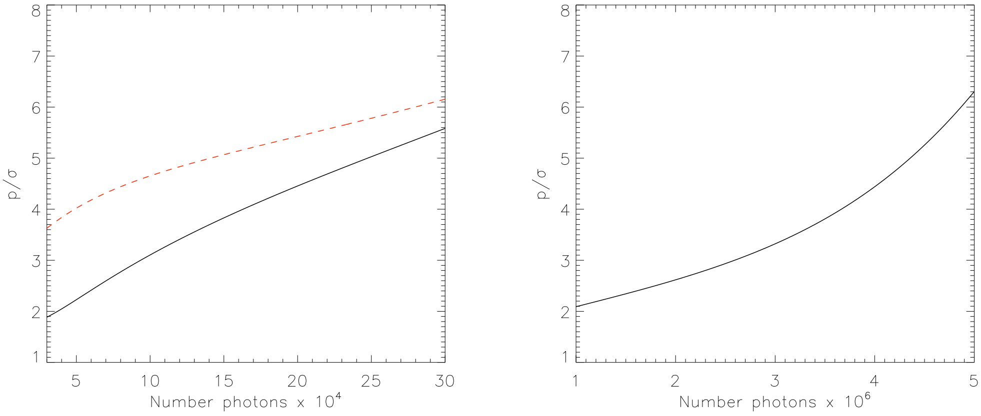


Figure 7.5 The plot shows the ratio $n = \hat{p}_i/\sigma_i$ for the most difficult parameters to estimate when fitting for structure 'shM1' (left figure) and 'shM2' (right figure) in black lines. The worst estimated parameter in terms of n when fitting to 'shM1' is r_t and s_2 when fitting to 'shM2'. The red dashed line in the left plot shows the ratio for parameter s_1 which is the worst determined of the two shape parameters when fitting to 'shM1'. It can be read of the figure that $\approx 4.2 \cdot 10^6$ number total photon counts are required for a minimum $5-\sigma$ detection on all parameters when fitting to 'shM2'. The radial and spectral binning is kept constant in the plot.

ratio, i.e. the parameter which is most difficult to estimate, for the two structures 'shM1' (left plot) and 'shM2' (right plot). In the 'shM1' example the most difficult parameter to estimate in terms of n is r_t, to reach a minimum $5-\sigma$ detection of this (and thereby for each parameter of the whole model) we find from the figure that more than $\approx 2.5 \cdot 10^5$ photons are required. If we instead only require that the shape parameters must be estimated with a minimum $5-\sigma$ each we find a limit of $\approx 1.4 \cdot 10^5$ photons, or roughly a factor of 2 less compared to an overall $5-\sigma$ detection. Following the same procedure for the more realistic example 'shM2' we find that a minimum of $\approx 4.2 \cdot 10^6$ photons are required for a minimum $5-\sigma$ detection on all parameters. The same number of photons are required for the shape fitting because s_2 is the most difficult parameter to estimate in terms of n.

Model comparison

Assuming that the quantities we try to measure for a gas can be parameterized, we still have the problem that we have no idea of how the "true" or best parametrization for the gas looks like in a real observation. This means e.g. that a set of shape parameters defined in a specific gas parametrization do not have to describe a real shape at all. The parameters could in principle just capture higher order corrections to the density profile, because of the general tight correlation between shape and density. In this case, the real problem is to realize that your model does not return information about the system in the way you believe. The question is therefore how to quantify how a specific model performs relative to one or several other competitive models. A useful measure of this can be found using Bayes' theorem. From this theorem it is possible to calculate the relative probability, also known as the *posterior odds*, of two competing models ((54), (104)). In the case where we assume flat parameter and model priors, the posterior odds ratio reduces to the simple ratio

$$\mathcal{F}(H_1, H_0) = \int \mathcal{L}(D \mid H_1, \beta) d\beta \Big/ \int \mathcal{L}(D \mid H_0, \alpha) d\alpha \tag{7.7}$$

where $\mathcal{L}(D \mid H, \beta)$ is the likelihood for getting the data D given the model H which depends on the parameterset β. This ratio $\mathcal{F}$ is often denoted the *evidence ratio* between model H_1 and H_0. Model H_0 is often a 'null' or default model where H_1 is a competing and often more complicated model. In our case, H_0 could be a model assuming spherical symmetry and H_1 a model allowing the shape to vary. The evidence threshold, or critical threshold, between rejecting or accepting a competitive model is often taken to be Jeffreys threshold 1:148 ((54)). Let us now go through a few examples.

First suppose we want to compare two models, M_1 and M_2, given the data set shM2. Both models are using the correct parameterizations for temperature and density, but not the same parametrization for shape; model M_1 includes the true parametrization of shape in the fitting, but model M_2 assumes spherical symmetry. We can now use Bayes' theorem to show if e.g. a $5 \cdot 10^5$ photon exposure carries enough information to distinguish between M_1 and M_2. Performing the two integrals in equation 7.7 for a $5 \cdot 10^5$ photon exposure we find $\mathcal{F} \sim 900$, i.e. we can correctly conclude that M_1 is strongly favored over M_2. The slightly biased estimations for the density and temperature when shM2 is fitted assuming M_2 is seen in figure 7.2.

Another scenario could be that we fit the shape with a parametrization that is different from the true one. In that case, suppose we fit dataset shM1 with two models M_1 and M_2. Both of them are using the true temperature and density parameterizations, but model M_1 is using equation 7.3 for the shape parameterization in contrast to model M_2 that is using equation 7.6. For a $3 \cdot 10^4$ photon exposure we find $\mathcal{F} \sim 6700$, concluding correctly that M_1 is strongly favored over M_2.

The last example is a case where the true structure has temperature and density profiles as 'shM2', but have a spherical shape. We now make a fit including shape, but we use equation 7.1, i.e. a simple beta-model, to describe the density instead of the true equation 7.4 that has one extra degree of freedom. The interesting thing is now that the best fit using the beta-model will show clear detection of shape away from spherical. This is seen on figure 7.6. The under fitted density profile is simply compensated by allowing a non-spherical shape in the inner parts. This is a false detection. In a real case where the true shape of the gas is not known, this can be very hard to realize. Comparing this fit using Bayes' theorem with a fit using the more general density profile in equation 7.4 we find $\mathcal{F} \sim 300$ for a $6 \cdot 10^4$ photon exposure. Which correctly means a spherical model is favored.

It is possible to write up a simple scaling relation between number photons and the evidence ratio given that the PDF approximately can be described by a multidimensional gaussian near its peak; Assume from a N_2 photon exposure we have calculated the evidence ratio $\mathcal{F}_{N_2}$ between two models M_A and M_B, from that we can simply calculate the ratio $\mathcal{F}_{N_1}$ for a N_1 photon exposure by $\mathcal{F}_{N_1} \approx \mathcal{F}_{N_2} (P^*_{M_A} / P^*_{M_B})^{(N_1/N_2 - 1)}$ where P^* is the value of the PDF at its maximum for the N_2 photon exposure. Here we have used the analytical solution to equation 7.7 (see e.g. (54) eq. 8). This scaling relation can be useful for forecasting the case where a correct integration is limited by, e.g. computational power. However, this estimator can be relative noisy because of its dependence on the value at the PDF maximum. One way to reduce this scatter could be to fit a gaussian to the PDF near its peak.

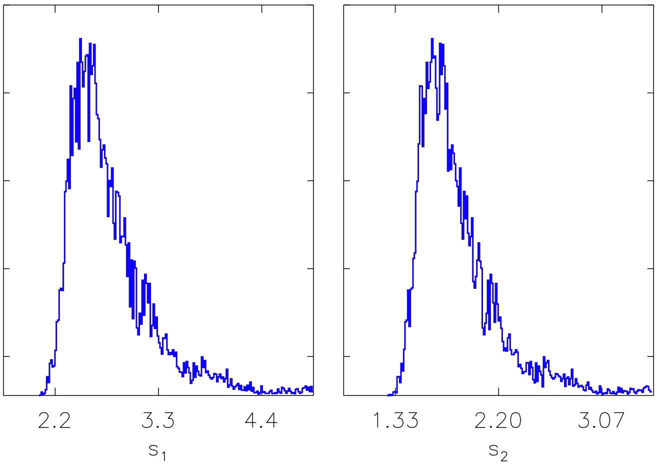


Figure 7.6 Example of a false detection of shape. Maximized PDFs for the two shape parameters s_1 and s_2 when fitting to a spherical version of 'shM2', but using a beta-model for the density instead of the true equation 7.4. The peaks in the shape parameters are real, but are not representing shape. They reflect that the parameterization for density used in the fitting is not having enough freedom to describe the variation of density in the inner parts. As described in the text, Bayes' theorem can be used to quantify if this is a true signal of shape or not. The plot is for a $6 \cdot 10^4$ photon exposure.

7.5 X-RAY GAS SHAPE AND CLUSTER MASS BIAS

By knowing the 3D x-ray gas temperature and density profiles one can calculate the underlying total cluster density, and hence mass, by combining the hydrostatic equilibrium (HE) equation

$$\nabla(\rho_{gas} T_{gas}) = -\rho_{gas} \nabla \Phi_{total} \tag{7.8}$$

with the poisson equation

$$\nabla^2 \Phi = 4\pi G \rho_{total} \tag{7.9}$$

where the index *'total'* indicates that the contribution is from both gas and dark matter. When x-ray observations are possible for a cluster and the x-ray gas is in HE, this method is one of the most precise ways to estimate the cluster mass as a function of radius within the visible x-ray region. However, as we can see, the estimated cluster mass will be wrong if the gas is not in HE or if ρ_{gas}, T_{gas} is not correctly known. One way of misestimating ρ_{gas} and T_{gas} is fitting a spherical model to data for an intrinsic non-spherical gas structure. Depending on the shape, this assumption will propagate to a bias in the estimated total cluster mass. In this section we will study the cluster mass bias as a function of different shapes along the line of sight.

7.5.1 MASS BIAS

The upper plot in figure 7.7 shows the shape along the line of sight for four different x-ray structures. We take the four structures to have temperature and density profiles similar to shM2, but different spacial shapes. Fitting temperature and density profiles to these four structures assuming spherical symmetry, will result in biased mass profiles. The ratio between the biased and the true mass profile is shown in the lower plot in figure 7.7. We have only included the mass contribution within r_v. Taking the rest of the mass of the cluster into account, requires an extrapolation of the dark matter potential form beyond the visible x-ray region. This is necessary when combining or comparing with other mass probes such as lensing.

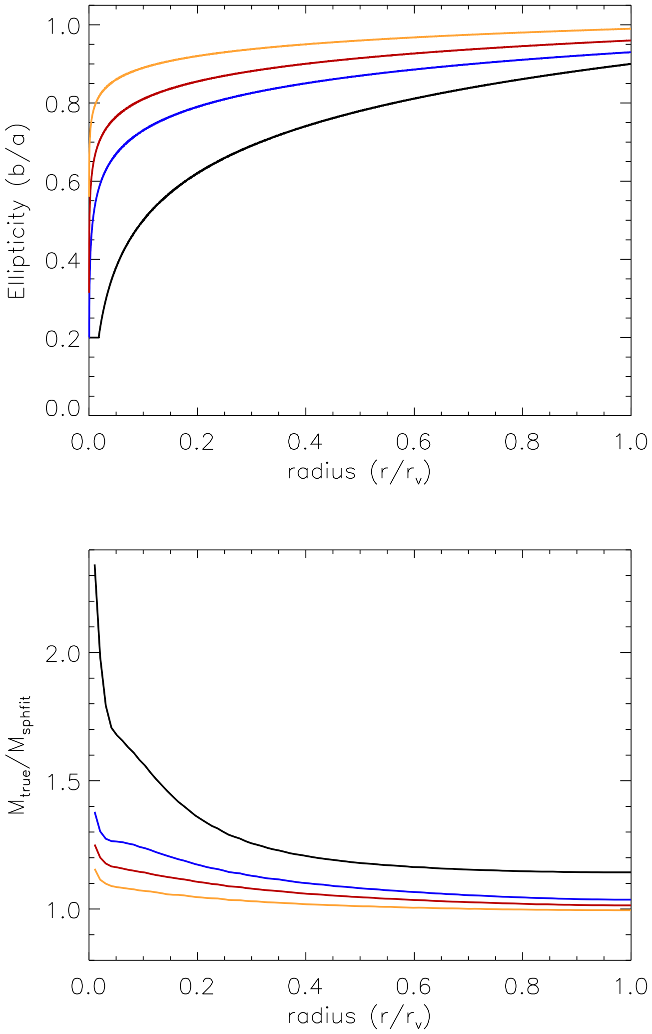


Figure 7.7 Mass bias from assuming spherical symmetry when fitting to non-spherical x-ray structures. Upper: Shape along the line of sight for four different prolate x-ray structures. Lower: Ratio between the true total mass profile and the mass profile estimated from a spherical fit to dataset 'shM2' having the shapes shown at the upper plot. Opposite bias (i.e. $M_{true}/M_{sphfit} < 1$) is expected for oblate structures.

As seen on the plot, the shapes we are considering leads to small biases at the 10 percent level, dependent on the radius. This difference can be important for doing future precision cosmology using clusters. However, at this level the degree of hydrostatic equilibrium may lead to higher uncertainties in the mass estimation ((59), (76), (21)).

7.6 CONCLUSIONS

We have presented a new method for measuring a radial dependent shape along the line of sight of the intracluster x-ray emitting gas. The method uses the assumption that the shape, temperature and density profiles can be described by parameterized functions. Compared to several previous studies, we use the whole spectral information. Using this method we have demonstrated the possibilities for measuring shape on `CHANDRA` mock data.

We find that around 10^6 photons are required to get a $5-\sigma$ detection of shape when fitting to a model showing realistic features of the gas, such as cool core and a double powerlaw for the density profile. We have seen, by presenting correlations matrices, that density and shape have a strong correlation, whereas temperature is essentially uncorrelated. This strong correlation indicates that independent measurements of the density profile can strongly improve the estimation of shape.

We demonstrated that Bayes' theorem very effectively can be used to compare different prior input models for our approach. This is of great importance since the actual science one extracts in the end has to be read off from the input model.

Finally we showed the effect on the mass profile estimation from assuming spherical symmetry when fitting structures with non-spherical shapes. Within our considered class of shapes, we found the mass estimation to be biased at the 10% level.

In a future paper we will use our framework on real data.

We warmly thank Martina Zamboni for useful discussions. The Dark Cosmology Centre is funded by the Danish National Research Foundation.

.1 APPENDIX

Creating artificial observations of an x-ray gas

In our analysis we have two different situations where we need to simulate a dataset. The first is as input to the MCMC routine when fitting to a given dataset. The second is where we actually simulate the dataset that has to be fitted, i.e. the mock data. The first steps for both are the same, and is described in the following; Given a set of parameterized profiles and shape we create a three dimensional x-ray gas on a grid. The local spectral information is calculated by `XSPEC`'s (see e.g. (6), (92)) model mekal (http://heasarc.nasa.gov/xanadu/`XSPEC`/manual/XSmodelMekal.html and references within) at redshift zero including galactic absorption. We use five times higher spatial resolution in the inner regions compared to the outer parts, to make sure no resolution effects propagate into the results. We then project all the spectral information onto the 2D observational plane defined such that the x-gas structure is spherical symmetric in that plane. The projected data is then convolved in `XSPEC` with an instrumental response function, here chosen to be from `CHANDRA`, to create a final observed picture. In an ideal world this is the picture read out from the instrument assuming pixelation from the CCD is unimportant. In a real world, a spacial and spectral rebinning is done at this step. When we create a dataset as input to the MCMC routine, the binning is done so that it matches the binning of the observed dataset. When generating a mock dataset we do the binning such that the radial bins have the same number photon counts and the spectral bins have more than a given threshold. This ensures equally statistical weights for each bin. For the fits in this paper, we fixed the number of radial bins to 12 for all datasets. Because an x-ray gas density profile usually have a logarithmic shape, the radial bins are therefore also approximately logarithmic linear spaced. Our spectral threshold is chosen such that the number of new spectral bins are around 200, of originally

1024. This corresponds to a threshold of 20 counts per spectral bin for a $6 \cdot 10^4$ number photons observation. It was not computationally possible to scan over different binning strategies, but the chosen binning is believed to match a real case scenario fairly well.

Figure 8 (left) illustrates a noise free generated x-ray gas map with a non-spherical shape and its temperature profile. The shape and the temperature profile is the one used for 'shM2' introduced in section 7.4.2. The right plot in figure 8 shows two spectra generated from the region between the two black lines shown in the left plot. The spectrum in red is a free-free spectrum generated with `XSPEC` using the mean projected temperature and the spectrum in blue is the true projected spectrum, i.e. the sum of many free-free spectra each generated locally in the x-gas. The difference seen in the lower part of the right plot is basically what give us information about shape and profiles.

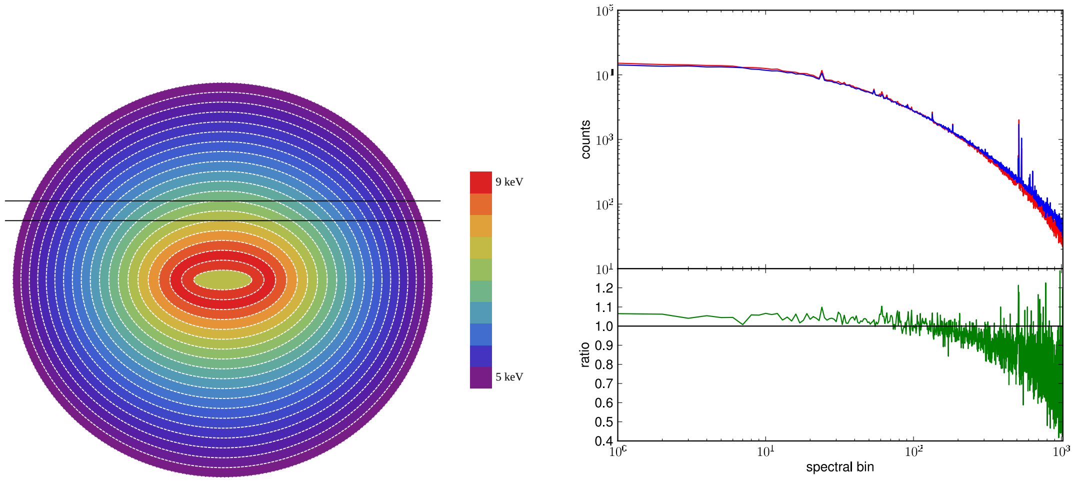


Figure 8 Left: A generated x-ray gas map with no noise added in the plane along the line of sight for structure 'shM2' introduced in section 7.4.2. The color indicates the temperature in keV. Right: Upper plot shows two spectra generated from the region between the two black lines shown in the left plot. The spectrum in red is a free-free spectrum generated with the mean x-ray temperature from the region between the two black lines in the left plot and the spectrum in blue is the true projected spectrum, i.e. the sum of many free-free spectra each generated locally in the x-gas. The lower plot shows the ratio between the red and the blue spectra.

Monte Carlo Technics used for this paper

We wrote a Monte Carlo Markov-Chain (MCMC) algorithm for fitting to a data set. The MCMC uses a Metropolis-Hastings sampling ((24)) with a flat and symmetric proposal density. The size of this proposal density was tuned to reach an acceptance rate of around 0.2-0.3 which has been shown to be the most optimal for sampling higher dimensional distributions. The width of the proposal density along each parameter axes was tuned in units of the root mean square for the individual PDF for each parameter. For all runs the sampling space was limited by bounds on each parameter axes and realizations with a temperature profile exceeding 15 kev or going below 0.5 kev was given zero probability. Among numerous tests of possible resolution, boundary or sampling effects we tested that the codes reproduced the theoretical expected degeneracy between an overall density scaling and a fixed axis ratio along the line of sight. We tested this up to a total number of 500.000 photons. A sample of tests was also done against an independently written code which generates artificial x-ray data using the “shell binning” approach (see e.g. http://cxc.harvard.edu/contrib/deproject/). We tested convergence by starting chains at random places and with different scalings (number photons) of the PDF. All distributions shown in the paper are based on $5 \cdot 10^6$ samplings. The fitting results presented are based on one realization of data, marginalizing over several realizations was not computationally possible. We assumed a diagonal covariance matrix for the observed photon measurements and the noise to be gaussian.

BIBLIOGRAPHY

[1] Abadi, M. G., Navarro, J. F., & Steinmetz, M. 2009, ApJ, 691, L63

[2] Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409

[3] Allen, S. W., Rapetti, D. A., Schmidt, R. W., Ebeling, H., Morris, R. G., & Fabian, A. C. 2008, MNRAS, 383, 879

[4] Amanullah, R., et al. 2011, ArXiv e-prints

[5] Andrae, R., Schulze-Hartung, T., & Melchior, P. 2010, ArXiv e-prints

[6] Arnaud, K. A. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes, 17–+

[7] Behroozi, P. S., Loeb, A., & Wechsler, R. H. 2013, , 6, 19

[8] Bekenstein, J. D. 1973, ApJ, 183, 657

[9] Bertschinger, E. 1985, ApJS, 58, 39

[10] Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton University Press)

[11] Blaauw, A. 1961, Bull. Astron. Inst. Netherlands, 15, 265

[12] Boylan, C., Li, Y., Fan, X. L., & Heng, I. S. 2014, ArXiv e-prints

[13] Bradley, L. D., et al. 2011, ArXiv e-prints

[14] Bromley, B. C., Kenyon, S. J., Geller, M. J., Barcikowski, E., Brown, W. R., & Kurtz, M. J. 2006, ApJ, 653, 1194

[15] Brown, W. R., Cohen, J. G., Geller, M. J., & Kenyon, S. J. 2012, ApJ, 754, L2

[16] Brown, W. R., Geller, M. J., & Kenyon, S. J. 2014, ApJ, 787, 89

[17] Burrows, A., & Hayes, J. 1996, in American Institute of Physics Conference Series, Vol. 366, High Velocity Neutron Stars, ed. R. E. Rothschild & R. E. Lingenfelter, 25–37

[18] Burrows, A., Livne, E., Dessart, L., Ott, C. D., & Murphy, J. 2007, ApJ, 655, 416

[19] Caldwell, N., Strader, J., Romanowsky, A. J., Brodie, J. P., Moore, B., Diemand, J., & Martizzi, D. 2014, ApJ, 787, L11

[20] Carucci, I. P., Sparre, M., Hansen, S. H., & Joyce, M. 2014, ArXiv e-prints

[21] Cavaliere, A., Lapi, A., & Fusco-Femiano, R. 2011, A&A, 525, A110+

[22] Chandrasekhar, S. 1943, ApJ, 97, 255

[23] Chapman, S. C., et al. 2007, ApJ, 662, L79

[24] Chib, S., & Greenberg, E. 1995, 49, 327

[25] Chongchitnan, S., & Silk, J. 2011, ArXiv e-prints

[26] Conte, A., de Petris, M., Comis, B., Lamagna, L., & de Gregori, S. 2011, A&A, 532, A14+

[27] Dalal, N., Lithwick, Y., & Kuhlen, M. 2010, ArXiv e-prints

[28] Davies, M. B., King, A., Rosswog, S., & Wynn, G. 2002, ApJ, 579, L63

[29] De Filippis, E., Sereno, M., Bautz, M. W., & Longo, G. 2005, ApJ, 625, 108

[30] Drisko, R. M. 1955, PhD thesis, CARNEGIE-MELLON UNIVERSITY.

[31] Eddington, A. S. 1916, MNRAS, 76, 572

[32] Fakhouri, O., Ma, C.-P., & Boylan-Kolchin, M. 2010, MNRAS, 406, 2267

[33] Fedeli, C., Moscardini, L., & Matarrese, S. 2009, MNRAS, 397, 1125

[34] Fitchett, M. J. 1983, MNRAS, 203, 1049

[35] Fong, W., et al. 2013, ApJ, 769, 56

[36] Fusco-Femiano, R., Landi, R., & Orlandini, M. 2005, ApJ, 624, L69

[37] Gal-Yam, A., Maoz, D., Guhathakurta, P., & Filippenko, A. V. 2003, AJ, 125, 1087

[38] Gualandris, A., Portegies Zwart, S., & Sipior, M. S. 2005, MNRAS, 363, 223

[39] Guennou, L., et al. 2012, A&A, 537, A64

[40] Gvaramadze, V. V., Gualandris, A., & Portegies Zwart, S. 2009, MNRAS, 396, 570

[41] Hansen, S. H., Egli, D., Hollenstein, L., & Salzmann, C. 2005, , 10, 379

[42] Hansen, S. H., & Piffaretti, R. 2007, A&A, 476, L37

[43] Hayashi, E., Navarro, J. F., & Springel, V. 2007, MNRAS, 377, 50

[44] Heggie, D. C. 1975, MNRAS, 173, 729

[45] Hernquist, L. 1990, ApJ, 356, 359

[46] —. 1992, ApJ, 400, 460

[47] Hills, J. G. 1988, Nature, 331, 687

[48] Hjorth, J., & Williams, L. L. R. 2010, ApJ, 722, 851

[49] Host, O., & Hansen, S. H. 2011, ApJ, 736, 52

[50] Hut, P. 1983, ApJ, 268, 342

[51] Huxor, A. P., et al. 2014, MNRAS, 442, 2165

[52] Jain, B., & Zhang, P. 2008, Phys. Rev. D, 78, 063503

[53] Janka, H.-T. 2012, Annual Review of Nuclear and Particle Science, 62, 407

[54] Jenkins, C. R., & Peacock, J. A. 2011, MNRAS, 413, 2895

[55] Joyce, M., Marcos, B., & Sylos Labini, F. 2009, MNRAS, 397, 775

[56] Kaastra, J. S., et al. 2004, A&A, 413, 415

[57] Kneib, J.-P., Ellis, R. S., Santos, M. R., & Richard, J. 2004, ApJ, 607, 697

[58] Kravtsov, A. V. 2013, ApJ, 764, L31

[59] Lau, E. T., Kravtsov, A. V., & Nagai, D. 2009, ApJ, 705, 1129

[60] Lau, E. T., Nagai, D., Kravtsov, A. V., & Zentner, A. R. 2011, ApJ, 734, 93

[61] Lemze, D., Broadhurst, T., Rephaeli, Y., Barkana, R., & Umetsu, K. 2009, ApJ, 701, 1336

[62] Łokas, E. L., & Mamon, G. A. 2003, MNRAS, 343, 401

[63] Ludlow, A. D., Navarro, J. F., Springel, V., Jenkins, A., Frenk, C. S., & Helmi, A. 2009, ApJ, 692, 931

[64] Lynden-Bell, D. 1967, MNRAS, 136, 101

[65] Majewski, S. R., et al. 2007, ApJ, 670, L9

[66] Mantz, A., Allen, S. W., Rapetti, D., & Ebeling, H. 2010, MNRAS, 406, 1759

[67] Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution

[68] Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler, A. 1996, Nature, 379, 613

[69] Morandi, A., Limousin, M., Rephaeli, Y., Umetsu, K., Barkana, R., Broadhurst, T., & Dahle, H. 2011, MNRAS, 416, 2567

[70] Morandi, A., Pedersen, K., & Limousin, M. 2010, ApJ, 713, 491

[71] Ogorodnikov, K. F. 1957, Soviet Ast., 1, 748

[72] Peñarrubia, J., Koposov, S. E., & Walker, M. G. 2012, ApJ, 760, 2

[73] Peng, E. W., et al. 2011, ApJ, 730, 23

[74] Pietilä, H., Heinämäki, P., Mikkola, S., & Valtonen, M. J. 1995, Celestial Mechanics and Dynamical Astronomy, 62, 377

[75] Piffaretti, R., Jetzer, P., Kaastra, J. S., & Tamura, T. 2005, A&A, 433, 101

[76] Piffaretti, R., & Valdarnini, R. 2008, A&A, 491, 71

[77] Planck Collaboration et al. 2011, ArXiv e-prints

[78] Pointecouteau, E., Arnaud, M., & Pratt, G. W. 2005, A&A, 435, 1

[79] Pontzen, A., & Governato, F. 2013, MNRAS, 430, 121

[80] Postman, M., et al. 2011, ArXiv e-prints

[81] Pota, V., et al. 2013, MNRAS, 428, 389

[82] Prada, F., Klypin, A. A., Cuesta, A. J., Betancort-Rijo, J. E., & Primack, J. 2012, MNRAS, 423, 3018

[83] Presotto, V., et al. 2014, A&A, 565, A126

[84] Price-Whelan, A. M., Hogg, D. W., Johnston, K. V., & Hendel, D. 2014, ArXiv e-prints

[85] Rapetti, D., Allen, S. W., Mantz, A., & Ebeling, H. 2010, MNRAS, 406, 1796

[86] Redmount, I. H., & Rees, M. J. 1989, Comments on Astrophysics, 14, 165

[87] Romanowsky, A. J., Strader, J., Brodie, J. P., Mihos, J. C., Spitler, L. R., Forbes, D. A., Foster, C., & Arnold, J. A. 2012, ApJ, 748, 29

[88] Sand, D. J., et al. 2011, ApJ, 729, 142

[89] Sanderson, R. E., Helmi, A., & Hogg, D. W. 2014, in IAU Symposium, Vol. 298, IAU Symposium, ed. S. Feltzing, G. Zhao, N. A. Walton, & P. Whitelock, 207–212

[90] Sarazin, C. L. 1988, X-ray emission from clusters of galaxies, ed. Sarazin, C. L.

[91] Sartoris, B., Borgani, S., Fedeli, C., Matarrese, S., Moscardini, L., Rosati, P., & Weller, J. 2010, MNRAS, 407, 2339

[92] Schafer, R. A. 1991, XSPEC, an x-ray spectral fitting package : version 2 of the user's guide, ed. Schafer, R. A.

[93] Sereno, M., Ettori, S., & Baldi, A. 2011, ArXiv e-prints

[94] Sereno, M., & Umetsu, K. 2011, MNRAS, 416, 3187

[95] Shakeshaft, R., & Spruch, L. 1979, Reviews of Modern Physics, 51, 369

[96] Sigurdsson, S., & Phinney, E. S. 1993, ApJ, 415, 631

[97] Sparre, M., & Hansen, S. H. 2012, , 7, 42

[98] —. 2012, , 10, 49

[99] Spergel, D. N., & Hernquist, L. 1992, ApJ, 397, L75

[100] Springel, V. 2005, MNRAS, 364, 1105

[101] Stark, D. P., Ellis, R. S., Richard, J., Kneib, J.-P., Smith, G. P., & Santos, M. R. 2007, ApJ, 663, 10

[102] Strader, J., et al. 2011, ApJS, 197, 33

[103] Teyssier, M., Johnston, K. V., & Shara, M. M. 2009, ApJ, 707, L22

[104] Trotta, R. 2008, Contemporary Physics, 49, 71

[105] Tunnicliffe, R. L., et al. 2014, MNRAS, 437, 1495

[106] Vikhlinin, A., Kravtsov, A., Forman, W., Jones, C., Markevitch, M., Murray, S. S., & Van Speybroeck, L. 2006, ApJ, 640, 691

[107] Vikhlinin, A., et al. 2009, ApJ, 692, 1060

[108] Vikhlinin, A., et al. 2009, in ArXiv Astrophysics e-prints, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey, 305–+

[109] Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, ApJ, 568, 52

[110] Wetzel, A. R. 2011, MNRAS, 412, 49

[111] Willman, B., Governato, F., Wadsley, J., & Quinn, T. 2004, MNRAS, 355, 159

[112] Wojtak, R., Hansen, S. H., & Hjorth, J. 2011, ArXiv e-prints

[113] Wojtak, R., & Łokas, E. L. 2010, MNRAS, 408, 2442

[114] Yoo, J., Fitzpatrick, A. L., & Zaldarriaga, M. 2009, Phys. Rev. D, 80, 083514

[115] Yu, Q., & Tremaine, S. 2003, ApJ, 599, 1129

[116] Zhao, H., Johnston, K. V., Hernquist, L., & Spergel, D. N. 1999, A&A, 348, L49

[117] Zheng, Z., et al. 2014, ApJ, 785, L23

[118] Zwicky, F. 1951, PASP, 63, 61