

Parametrizing the local dark matter speed distribution: a detailed analysis

Bradley J. Kavanagh*

School of Physics & Astronomy, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

(Dated: 9th December, 2013)

In a recent paper, a new parametrization for the dark matter (DM) speed distribution $f(v)$ was proposed for use in the analysis of data from direct detection experiments. This parametrization involves expressing the *logarithm* of the speed distribution as a polynomial in the speed v . We present here a more detailed analysis of the properties of this parametrization. We show that the method leads to statistically unbiased mass reconstructions and exact coverage of credible intervals. The method performs well over a wide range of DM masses, even when finite energy resolution and backgrounds are taken into account. We also show how to select the appropriate number of basis functions for the parametrization. Finally, we look at how the speed distribution itself can be reconstructed, and how the method can be used to determine if the data are consistent with some test distribution. In summary, we show that this parametrization performs consistently well over a wide range of input parameters and over large numbers of statistical ensembles and can therefore reliably be used to reconstruct both the DM mass and speed distribution from direct detection data.

PACS numbers: 07.05.Kf, 14.80.-j, 95.35.+d, 98.62.Gq

I. INTRODUCTION

The dark matter (DM) paradigm has enjoyed much success in explaining a wide range of astronomical observations (for a review, see e.g. Ref. [1]). As yet no conclusive evidence has been provided for the identity of particle DM, though there are a variety of candidates, including the supersymmetric neutralino [2], sterile neutrinos [3], axions [4] and the lightest Kaluza-Klein particle [5]. Here, we focus on the search for particles which belong to the generic class of Weakly Interacting Massive Particles (WIMPs). Direct detection experiments [6, 7] aim to measure the energies of nuclear recoils induced by WIMP DM in the Galactic halo. Under standard assumptions about the DM halo, this data can be used to extract the WIMP mass and interaction cross section, allowing us to check for consistency with other search channels (such as indirect detection [8] and collider experiments [9]) and to probe underlying models of DM.

Direct detection experiments are traditionally analyzed within the framework of the Standard Halo Model (SHM), in which WIMPs are assumed to have a Maxwell-Boltzmann speed distribution in the Galactic frame. The impact of uncertainties in the WIMP speed distribution has been much studied (see e.g. Refs. [10–12]), leading to the conclusion that such uncertainties may introduce a bias into any reconstruction of the WIMP mass from direct detection data. As yet, the speed distribution is unknown, while a number of proposals have been put forward for its form, including analytic parametrizations (e.g. Ref. [13]) and distributions reconstructed from the potential of the Milky Way [14] or from N-body simulations [15–18]. Recent results from N-body simulations which attempt to include the effects of baryons on structure formation also report the possible presence of a dark

disk in the Milky Way [19–21]. With such a wide range of possibilities, we should take an agnostic approach to the speed distribution, not only to avoid introducing bias into the analysis of data, but also with the hope of measuring the speed distribution and thereby probing the formation history of the Milky Way.

Several methods of evading these uncertainties have been proposed. These include simultaneously fitting the parameters of the SHM and dark matter properties [22, 23], fitting to empirical forms of the speed distribution (e.g. Ref. [24]) and fitting to a self-consistent distribution function [25]. However, these methods typically require that the speed distribution can be well fitted by a particular functional form. More model-independent methods, such as fitting the moments of the speed distribution [26, 27] or using a step-function speed distribution [11], have also been presented. However, these methods can still introduce a bias into the measurement of the WIMP mass and perform less well with the inclusion of realistic experimental energy thresholds.

In a recent paper [28] (hereafter referred to as Paper 1), a new parametrization of the speed distribution was presented, which allowed the WIMP mass to be extracted from hypothetical direct detection data without prior knowledge of the speed distribution itself. Paper 1 demonstrated this for a WIMP of mass 50 GeV, using several underlying distribution functions. In the present paper, we extend this analysis to a wider range of masses. We also aim to demonstrate the statistical properties of the method and show how realistic experimental parameters affect its performance. Finally, we will also elaborate on some of the technical details of the method and assess its ability to reconstruct the underlying WIMP speed distribution.

Section II of this paper explains the direct detection event rate formalism and presents the parametrization of the speed distribution introduced in Paper 1. In Sec. III, the methodology for testing the parametrization is out-

* Electronic address: ppxbk2@nottingham.ac.uk

lined. In Section IV, we consider the choice and number of basis functions for the method. We then study the performance of the method as a function of input WIMP mass (Sec. V) and when Poisson fluctuations in the data are taken into account (Sec. VI). In Sec. VII, we demonstrate how the speed distribution can be extracted from this parametrization and examine whether or not different distribution functions can be distinguished. Finally, we summarize in Sec. VIII the main results of this paper.

II. DIRECT DETECTION EVENT RATE

Dark matter direct detection experiments aim to measure the energies E of nuclear recoils induced by interactions with WIMPs in the Galactic halo. Calculation of the event rate at such detectors has been much studied (e.g. Refs. [2, 6, 7, 29]). For a target nucleus with nucleon number A , interacting with a WIMP of mass m_χ , the event rate per unit detector mass is given by:

$$\frac{dR}{dE} = \frac{\rho_0 \sigma_p}{2m_\chi \mu_{\chi p}^2} A^2 F^2(E) \eta(v_{\min}), \quad (1)$$

where ρ_0 is the local dark matter mass density, σ_p is the WIMP-proton spin-independent cross section and the reduced mass is defined as $\mu_{AB} = m_A m_B / (m_A + m_B)$. The Helm form factor $F^2(E)$ [30] describes the loss of coherence of spin-independent scattering due to the finite size of the nucleus. A wide range of possible interactions have been considered in the literature, including inelastic [31], isospin-violating [32] and more general non-relativistic interactions [33–35]. We focus here on the impact of the WIMP speed distribution on the direct detection event rate. We therefore restrict ourselves to considering only spin-independent scattering, which is expected to dominate over the spin-dependent contribution for heavy nuclei, due to the A^2 enhancement in the rate.

Information about the WIMP velocity distribution $f(\mathbf{v})$ is encoded in the function η , sometimes referred to as the mean inverse speed,

$$\eta(v_{\min}) = \int_{v > v_{\min}} \frac{f(\mathbf{v})}{v} d^3\mathbf{v}, \quad (2)$$

where \mathbf{v} is the WIMP velocity in the reference frame of the detector. The integration is performed only over those WIMPs with sufficient speed to induce a nuclear recoil of energy E . The minimum required speed for a target nucleus of mass m_N is

$$v_{\min}(E) = \sqrt{\frac{m_N E}{2\mu_{\chi N}^2}}. \quad (3)$$

We distinguish between the directionally averaged velocity distribution

$$f(v) = \oint f(\mathbf{v}) d\Omega_{\mathbf{v}}, \quad (4)$$

and the 1-dimensional speed distribution

$$f_1(v) = \oint f(\mathbf{v}) v^2 d\Omega_{\mathbf{v}}. \quad (5)$$

The distribution function should in principle be time-dependent, due to the motion of the Earth around the Sun. However, this is expected to be a percent-level effect (for a review, see e.g. Ref. [36]) and we therefore assume that $f_1(v)$ is time independent in the present work.

We consider several benchmark speed distributions in this work, including the SHM and the SHM with the addition of a moderate dark disk which accounts for 23% of the total WIMP density [21]. We model the speed distributions as combinations of Gaussian functions in the Earth frame

$$g(\mathbf{v}) = N \exp\left(-\frac{(\mathbf{v} - \mathbf{v}_{\text{lag}})^2}{2\sigma_v^2}\right) \Theta(v_{\text{esc}} - |\mathbf{v} - \mathbf{v}_{\text{lag}}|), \quad (6)$$

where \mathbf{v}_{lag} specifies the peak velocity of the distribution in the Earth frame and σ_v the velocity dispersion. We truncate the distribution above the escape speed v_{esc} in the Galactic frame and the factor N is required to satisfy the normalization condition (Eq. 9). We use the value $v_{\text{esc}} = 544 \text{ km s}^{-1}$, which lies within the 90% confidence limits obtained from the RAVE survey [37, 38]. In addition, we also use the speed distribution of Lisanti et al. [13], which has the following form in the Earth's frame:

$$f(\mathbf{v}) = N \left[\exp\left(\frac{v_{\text{esc}}^2 - |\mathbf{v} - \mathbf{v}_0|^2}{kv_0^2}\right) - 1 \right]^k \Theta(v_{\text{esc}} - |\mathbf{v} - \mathbf{v}_0|). \quad (7)$$

We use the parameter values $k = 2$ and $v_0 = 220 \text{ km s}^{-1}$ in this work. We summarize in Tab. I the different speed distributions considered. We also plot several of these in Fig. 1 for reference.

In Paper 1, a parametrization for the WIMP speed distribution was introduced, for use in the analysis of direct detection data. The parametrization of Paper 1 has the form:

$$f_1(v) = v^2 \exp\left\{-\sum_{k=0}^{N-1} a_k P_k(v)\right\} \quad (8)$$

where $P_k(v)$ is some basis of polynomial functions of v . We fit the coefficients $\{a_1, \dots, a_{N-1}\}$ using data, and fix a_0 by normalization

$$a_0 = \ln\left(\int_0^\infty v^2 \exp\left\{-\sum_{k=1}^{N-1} a_k P_k(v)\right\} dv\right). \quad (9)$$

This form of parametrization ensures that the distribution function $f_1(v)$ is everywhere positive and can be used to fit an arbitrary underlying directionally-averaged distribution function (given a sufficiently large number of polynomial basis functions). We explore in Sec. IV which basis functions should be used in the parametrization, as well as how many basis functions are required.

Speed distribution benchmark	Fraction	$v_{\text{lag}}/\text{km s}^{-1}$	$\sigma_v/\text{km s}^{-1}$
SHM	1	220	156
SHM+DD	0.77	220	156
	0.23	50	50
Stream	1	400	20
Bump	0.97	220	156
	0.03	500	20
Double-peak	0.5	200	20
	0.5	400	20
Lisanti et al.	$v_0 = 220 \text{ km s}^{-1} \quad k = 2$		

TABLE I. Summary of speed distribution benchmarks used in this work. Some benchmarks are modelled as mixtures of two gaussian components (defined in Eq. 6), for which we give the fractional contribution of each component (labelled ‘Fraction’). The remaining parameters are defined in Eqs. 6 and 7 and the accompanying text. The ‘bump’ and ‘double-peak’ distributions are discussed in Sec. IV.

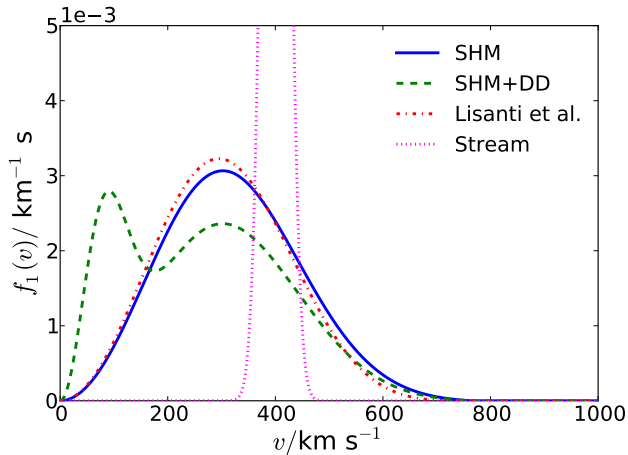


FIG. 1. Several of the benchmark speed distributions used in this work. They are defined in Eqs. 6 and 7 with parameters from Tab. I. These distributions are the SHM (solid blue), SHM+DD (dashed green), Lisanti et al. (dot-dashed red) and the stream (dotted magenta).

III. PARAMETER RECONSTRUCTION

In order to assess the performance of the parametrization method, we attempt to reconstruct the WIMP mass m_χ and polynomial coefficients $\{a_1, \dots, a_{N-1}\}$ using the nested sampling software MULTINEST [39–41]. We also include the WIMP-proton spin-independent cross section σ_p as a free parameter. However, we are forced to treat the cross section as a nuisance parameter. As has previously been noted [28, 42], taking an agnostic approach to the DM speed distribution means that we do not know what fraction of WIMPs lie above the energy thresholds

Experiment	Target Mass, A	Detector Mass (fid.), m_{det}/kg	Efficiency, ϵ	Energy Range/keV
Xenon	131	1100 [48]	0.7 [49]	7-45 [50]
Argon	40	1000	0.9 [51]	30-100 [52]
Germanium	73	150 [53]	0.6 [54]	8-100 [54]

TABLE II. Summary of experimental parameters used in this work, defined in Sec. III A. An exposure of $t_{\text{exp}} = 2$ years is used for all 3 experiments.

of the experiments. While this does not adversely impact the reconstruction of the WIMP mass, it does result in a strong degeneracy, such that only lower limits can be placed on the cross section using such methods. In any case, the cross section appears in the event rate (Eq. 1) in the degenerate combination $\rho_0 \sigma_p$. Uncertainties on the local DM density ρ_0 are at present on the order of a factor of 2 (see e.g. [43–46]) and thus any reconstruction of the cross section would be subject to the same systematic uncertainty. In this work, we focus instead on reconstructing the WIMP mass and the shape of the speed distribution. For concreteness, we use the values $\sigma_p = 10^{-45} \text{ cm}^2$ and $\rho_0 = 0.3 \text{ GeV cm}^{-3}$ throughout this work.

A. Experimental benchmarks

In order to generate mock data sets, we consider three idealized mock experiments, loosely based on detectors which are currently in development. As previous work has shown [42, 47], the WIMP mass and speed distribution are degenerate when data from only a single experiment is considered. However, this degeneracy can be broken by including data from additional experiments with different nuclear target masses. The three target materials we consider here are Xenon, Argon and Germanium. We describe each experiment in terms of its nucleon number A , fiducial detector mass m_{det} , efficiency ϵ and energy sensitivity window $[E_{\text{min}}, E_{\text{max}}]$. We incorporate the effects of detector sensitivity, analysis cuts and detector down-time into the value of the efficiency ϵ , which we take to be energy independent for simplicity. We consider a total exposure time for all experiments of $t_{\text{exp}} = 2$ years. The experimental parameter values used in this work are summarized in Tab. II.

The exact parameter values we used in this work do not strongly impact the results we present. However, it is important to note that the total mass and exposure of the experiments will affect the total number of events observed. This in turn will affect the precision of the reconstructions. For example, we have chosen a total Argon mass of 1000 kg. This is the stated target for Argon-based experiments which are in development (e.g. Ref. [55]), though at present typical fiducial masses for

Parameter	Prior type	Prior range
m_χ/GeV	log-flat	$[10^0, 10^3]$
σ_p/cm^2	log-flat	$[10^{-46}, 10^{-42}]$
$\{a_k\}$	linear-flat	$[-50, 50]$
R_{BG}/dru	log-flat	$[10^{-12}, 10^{-5}]$

TABLE III. Summary of the priors on the parameters used in this work. The background rate R_{BG} is defined in Sec. V, while the remaining parameters are defined in Sec. II.

Parameter	Value
N_{live}	10000
efficiency	0.25
tolerance	10^{-4}

TABLE IV. Summary of the MultiNest sampling parameters used in this work.

Argon prototypes are of the order of 100 kg [52]. The data we have generated does not represent the ‘high-statistics’ regime: across all three experiments the total number of events observed is roughly 200-300 with as few as 10 events in the Germanium detector for some scenarios. Using a smaller exposure (or equivalently a smaller interaction cross section) will reduce the precision of the results, but should not introduce any additional bias. We also briefly consider the impact of a *larger* number of events in Sec. VII.

B. Parameter sampling

We make parameter inferences using a combination of Bayesian and frequentist statistics. Bayes theorem for the probability of a particular set of theoretical parameters Θ given the observed data \mathbf{D} is:

$$P(\Theta|\mathbf{D}) = \frac{P(\Theta)P(\mathbf{D}|\Theta)}{P(\mathbf{D})}, \quad (10)$$

where $P(\Theta)$ is the prior on the parameters and $P(\mathbf{D})$ is the Bayesian evidence, which acts as a normalizing factor and has no impact on parameter inference. We summarize the priors used in this work in Tab. III. We also summarize in Tab. IV the MultiNest sampling parameters used.

The factor $P(\mathbf{D}|\Theta)$ is simply the likelihood of the data given the parameters Θ . In Sec. IV and Sec. V, we consider the effects of varying the form of the parametrization and of varying the input WIMP mass. In order to eliminate the effects of Poisson noise, we use Asimov data [56] for these sections. This means that we divide the energy window of each experiment into bins of width 1 keV. We then set the observed number of events $N_{o,i}$

in bin i equal to the expected number of events $N_{e,i}$. In this case, we use the binned likelihood, calculated for N_b energy bins:

$$\mathcal{L}_b = \prod_{i=1}^{N_b} \frac{N_{e,i}^{N_{o,i}} e^{-N_{e,i}}}{N_{o,i}!}. \quad (11)$$

In Sec. VI and Sec. VII, we consider many realisations of data, including the effects of Poisson noise. We therefore use the extended likelihood which has previously been used by both the Xenon [57] and CDMS [58] collaborations, which for a single experiment is given by:

$$\mathcal{L} = \frac{N_e^{N_o} e^{-N_e}}{N_o!} \prod_{i=1}^{N_o} P(E_i), \quad (12)$$

where the expected number of events is given by:

$$N_e = \epsilon m_{\text{det}} t_{\text{exp}} \int_{E_{\text{min}}}^{E_{\text{max}}} \frac{dR}{dE} dE, \quad (13)$$

and the normalised recoil spectrum is given by:

$$P(E) = \frac{\epsilon m_{\text{det}} t_{\text{exp}}}{N_e} \frac{dR}{dE}. \quad (14)$$

The total likelihood is then the product over all experiments considered.

Using nested sampling, we can extract the full posterior probability distribution of the parameters $P(\Theta|\mathbf{D})$, as well as the likelihood $\mathcal{L}(\Theta)$. However, we often want to make inferences not jointly for all parameters but for only a subset (treating the remaining as nuisance parameters). If we conceptually partition the parameter space into the parameters of interest ψ and the remaining nuisance parameters ϕ , we would like to make inferences about the values of ψ , without reference to the values of ϕ . One option for doing this is to calculate the marginalized posterior distribution, obtained by integrating the posterior probability over the parameters we are not interested in:

$$P_m(\psi) = \int P(\psi, \phi) d\phi. \quad (15)$$

This method performs well for small numbers of observations (compared to the number of free parameters in the fit). We take the mode of the distribution to be the reconstructed parameter value and construct p% *minimal* credible intervals, which include those parameter values with $P_m(\psi) \geq h$, where h is chosen such that p% of the probability distribution lies within the interval. The marginalized posterior method is used in Sec. VI and Sec. VII, where in some cases the number of events observed in an experiment is less than 10.

An alternative method is to calculate the profile likelihood. This is obtained by maximizing the full likelihood function over the nuisance parameters:

$$\mathcal{L}_p(\psi) = \max_{\phi} \mathcal{L}(\psi, \phi). \quad (16)$$

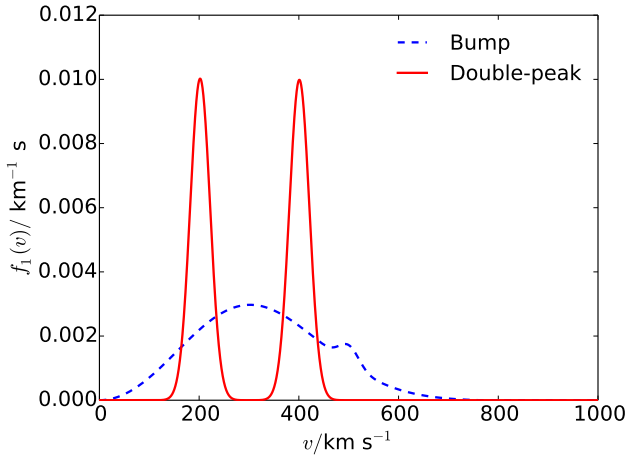


FIG. 2. Benchmark speed distributions used in Sec. IV to test the performance of the parametrization as a function of the number and type of basis functions.

For a large number of observations, we can take the value which maximizes \mathcal{L}_p as the reconstructed value and construct confidence intervals using the asymptotic properties of the profile likelihood. We use the profile likelihood for parameter inferences in Sec. IV and Sec. V, as the Asimov data sets provide a large number of measurements of $N_{e,i}$ over a large number of bins. The profile likelihood can also lead to less noisy reconstructions than the marginalized posterior, especially when the dimensionality of the parameter space becomes high, as in Sec. IV and Sec. V.

IV. TESTING THE PARAMETRIZATION

We now consider the two questions: 1. how many basis functions are required and 2. which polynomial basis should be used? In order to answer these questions, we use the two benchmark distribution functions illustrated in Fig. 2. We have chosen these benchmarks not because they are necessarily realistic distribution functions but because they should be difficult to fit using standard techniques and fitting functions (e.g. [13]). The first distribution (referred to as ‘bump’) is a SHM distribution with the addition of a small bump, which contributes just 3% of the total WIMP population and could correspond to a small sub-halo or stream [15]. This should be difficult to fit because it represents only a very small deviation from the standard scenario. The second distribution (referred to as ‘double-peak’) has a sharp and rapidly varying structure, which we anticipate should be difficult to capture using a small number of basis functions.

A. Varying the number of basis functions

We first investigate how the reconstructed WIMP mass m_{rec} and uncertainty varies with the number of basis functions N . For now, we fix our choice of basis to shifted Legendre polynomials, as used in Paper 1:

$$P_k(v) = L_k\left(2\frac{v}{v_{\text{max}}} - 1\right), \quad (17)$$

where L_k is the Legendre polynomial of order k , and v_{max} is a cut off for the parametrization. We should choose v_{max} to ensure that $f_1(v)$ is negligible above the cut off. However, too high a choice of v_{max} will result in $f_1(v)$ being close to zero over a large range of the parametrization, making fitting more difficult. We use the value $v_{\text{max}} = 1000 \text{ km s}^{-1}$, while lies significantly above the Galactic escape speed.

The lower panel of Fig. 3 shows the best fit mass and 68% confidence intervals as a function of N , using as input a WIMP of mass 50 GeV and the ‘bump’ distribution function. The reconstructed mass very rapidly settles close to the true value, using as few as three basis functions. This is because adding the bump near $v \sim 500 \text{ km s}^{-1}$ still leaves the mean inverse speed relatively smooth, so a large number of basis function are not required. The correct mass is reconstructed and we emphasize in the lower panel of Fig. 3 that the reconstruction is stable with the addition of more basis functions.

We should also consider how the quality of the fit changes as a function of N . We would expect that adding fit parameters should always lead to a better fit. Eventually, the fit should be good enough that adding additional basis functions will no longer improve it significantly. We can then be confident that our reconstruction is accurate and not an artifact of using too few basis functions. In order to investigate this, we utilise the Bayesian Information Criterion (BIC) [59], which is given by:

$$BIC = 2N_p \ln(N_m) - \ln(\mathcal{L}_{\text{max}}), \quad (18)$$

where N_p is the number of free parameters, N_m is the number of measurements or observations and \mathcal{L}_{max} is the maximum likelihood value obtained in the reconstruction. For the case of binned data, N_m corresponds simply to the total number of energy bins across all experiments. This criterion penalises the inclusion of additional free parameters and in comparing several models, we should prefer the one which minimises the BIC.

The upper panel of Fig. 3 shows the BIC (in arbitrary units) as a function of the number of basis functions for the ‘bump’ distribution function. The BIC is comparable for the cases of $N = 2$ and $N = 3$, indicating that the quality of the fit is improved slightly by the addition of another basis function. However, adding further basis functions does not have a significant impact on the maximum likelihood, leading to an increase in the BIC. This

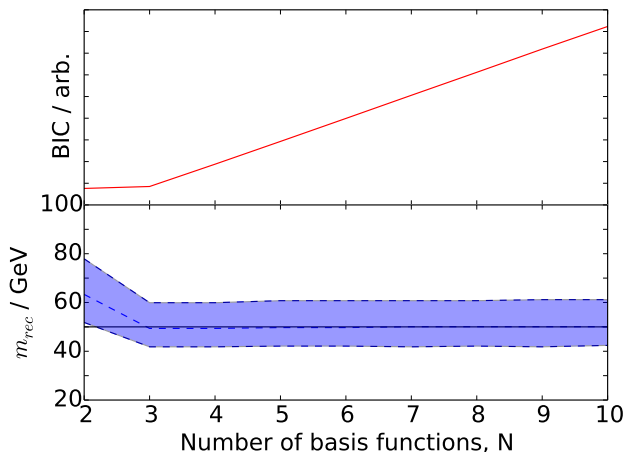


FIG. 3. Bayesian information criterion (BIC) as a function of the number of basis functions for an underlying ‘bump’ distribution function, 50 GeV WIMP and using Legendre polynomial basis functions (upper panel). Also shown (lower panel) are the reconstructed WIMP mass (dashed blue line), 68% confidence interval (shaded blue region) and underlying WIMP mass (solid horizontal black line).

coincides with the stabilization of the reconstructed mass around the true value and we conclude that only two or three basis functions are required to provide a good fit to the data.

Figure 4 shows the corresponding results for the ‘double-peak’ distribution function. Here, we note that the bias induced by using too small a number of basis functions is larger than for the case of the ‘bump’ distribution, due to the more complicated structure in this case. The BIC is minimized for $N = 7$, indicating that additional basis functions do not significantly improve the quality of the fit to data. This suggests that the shape of the speed distribution can be well fit by $N \geq 7$ basis functions. As shown in the lower panel of Fig. 4, the reconstruction of the WIMP mass is stable around the true mass for these values of N .

We propose that such a procedure should be used in the case of real data should a dark matter signal be observed at multiple detectors. We have shown that by analyzing the reconstructed mass as a function of N we can recover the true mass and that by using the BIC we can be confident that we have obtained an adequate fit to data.

B. Choice of basis functions

We now consider the second question posed at the start of Sec. IV: which polynomial basis should be used? We see immediately that a naive power series of the form

$$\ln f(v) \approx a_0 + a_1 v + a_2 v^2 + a_3 v^3 + \dots, \quad (19)$$

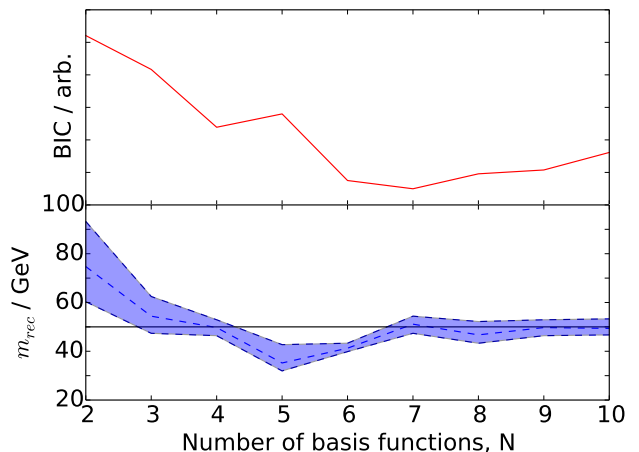


FIG. 4. As Fig. 3 but for an underlying ‘double-peak’ distribution function.

is not practical for the purposes of parameter estimation. Higher powers of v will have rapidly growing contributions to $\ln f$, meaning that the associated coefficients must be rapidly decreasing in order to suppress these contributions. Fitting to the SHM using just 5 terms, the range of values for the a_k in the case of a simple power series would span around 13 orders of magnitude. Ideally, we would like to specify an identical prior on each of the coefficients. However, in this scenario this would result in a highly inefficient exploration of the parameter space when some of the terms are so small.

This problem can be significantly improved by rescaling v . We choose to rescale by a factor of $v_{\text{max}} = 1000 \text{ km s}^{-1}$, and cut off the distribution function at v_{max} . The basis functions $(v/v_{\text{max}})^k$ are now less than unity by construction and the coefficients a_k are now dimensionless:

$$\ln f(v) \approx a_0 + a_1(v/v_{\text{max}}) + a_2(v/v_{\text{max}})^2 + a_3(v/v_{\text{max}})^3 + \dots \quad (20)$$

We now address the problem of *conditioning* of the polynomial basis (see e.g. Refs. [60, 61]). Conditioning is a measure of how much the value of a polynomial changes, given a small change in the coefficients. For a well-conditioned polynomial, small changes in the coefficient are expected to lead to small changes in the value of the polynomial. This is ideal for parameter estimation as it leads to a more efficient exploration of the parameter space. Orthogonal polynomial basis functions typically have improved conditioning [60] and we consider two specific choices: the Legendre polynomials which have already been considered and the Chebyshev polynomials. The Chebyshev polynomials are used extensively in polynomial approximation theory [62] and are expected to be well conditioned [60].

We have checked that the reconstruction results using Chebyshev polynomials are largely indistinguishable

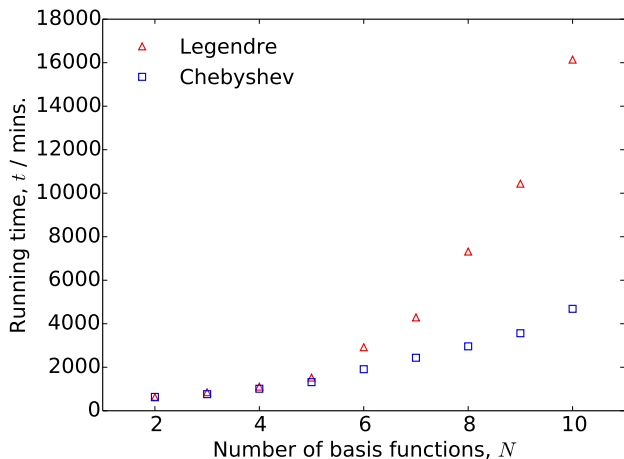


FIG. 5. Time taken (using 4 processors in parallel) for the reconstruction of the ‘bump’ benchmark, as a function of number of basis functions. The time taken using the Chebyshev basis (blue squares) grows more slowly with N than for the Legendre basis (red triangles).

from the case of Legendre polynomials for both the ‘bump’ and ‘double-peak’ distributions and as a function of N . This leads us to conclude that the accuracy of the reconstruction is independent of the specific choice of basis. However, the reconstruction was much faster in the case of the Chebyshev basis. This is illustrated in Fig. 5, which shows the time taken for reconstruction of the ‘bump’ benchmark as a function of N . The time taken grows much more slowly for the Chebyshev basis (roughly as N^2) than for the Legendre basis (roughly as N^3). We have also checked that this difference is not an artifact of how we calculate the basis functions. These results indicate that this choice of basis provides both reliable and efficient reconstruction for the WIMP mass and we therefore use the Chebyshev basis in the remainder of this work.

V. VARYING m_χ

In previous work [28], this parametrization method was only tested for a single WIMP mass of 50 GeV. Here, we extend this analysis to a wider range of WIMP masses. We generate Asimov data for WIMP masses of 10, 20, 30, 40, 50, 75, 100, 200 and 500 GeV and reconstruct the best fit WIMP mass m_{rec} and 68% and 95% confidence intervals from the profile likelihood. We use the SHM as a benchmark distribution function and use a fixed number of $N = 5$ basis functions. The results are shown in Fig. 6, along with the line $m_{\text{rec}} = m_\chi$ for reference.

For large values of m_χ , the shape of the event spectrum becomes independent of m_χ [63], which results in a widening of the confidence intervals as the WIMP mass increases. For low mass WIMPs, fewer events are ob-

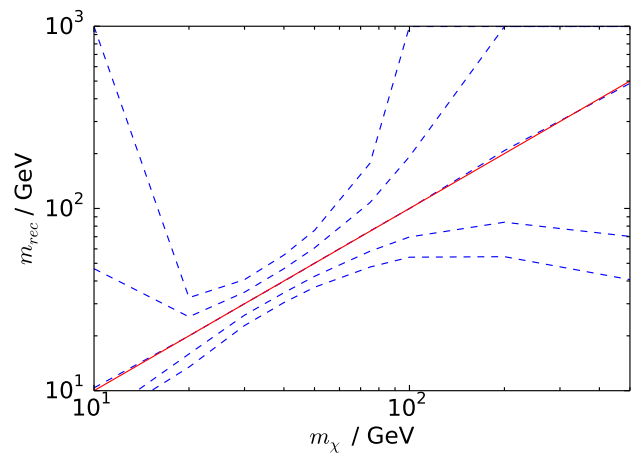


FIG. 6. Reconstructed WIMP mass m_{rec} (central dashed blue line) as a function of input WIMP mass m_χ as well as 68% and 95% intervals (inner and outer blue dashed lines respectively). The line $m_{\text{rec}} = m_\chi$ (solid red line) is also plotted for reference.

served in each bin, again resulting in wider confidence intervals. It should be noted that for this analysis we have used Asimov data, in which the exact (non-integer) number of events is recorded in each bin. For low mass WIMPs, this means that the spectrum (and therefore the correct WIMP mass) is still well reconstructed using Asimov data, in spite of the small number of events. The tightest constraints are obtained when the input WIMP mass is close to the masses of several of the detector nuclei (in the range 30-80 GeV). There also appears to be no bias in the WIMP mass: the reconstruction matches the true mass across all values considered.

An alternative parametrization method was proposed in Ref. [42], in which the *momentum* distribution of halo WIMPs was parametrized. For a given speed distribution, the corresponding momentum distribution may be broad and easily reconstructed for high mass WIMPs. However, for low mass WIMPs the momentum distribution would be much narrower, owing to their lower momenta. The momentum parametrization method therefore performs poorly for low mass WIMPs. The parametrization presented in this paper does not suffer from similar problems.

So far, we have only considered idealized direct detection experiments. We now apply the method to more realistic mock detectors, taking into account the effects of finite energy resolution, as well as unrejected background events. We assume here that each experiment has a gaussian energy resolution with fixed width $\sigma_E = 1$ keV, such that the observed event rate for recoils of energy E is given by:

$$\frac{dR}{dE} = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_E} \exp\left\{-\frac{(E-E')^2}{2\sigma_E^2}\right\} \frac{dR'}{dE'} dE', \quad (21)$$

where the primed event rate is the underlying (perfect resolution) rate. We also assume a constant flat background rate for each experiment $R_{\text{BG}} = 10^{-6}$ events/kg/keV/day (which has been suggested as a possible background rate for Xenon1T [50] and WArP-100L [52]) when generating mock data sets. However, we allow the flat background rate in each experiment to vary as free parameters during the fit.

We have chosen relatively generic resolution and background parameters in this work, because the precise details of energy resolution and background shape and rate will depend on the specific experiment under consideration. Instead, we hope to show that the inclusion of more realistic experimental setups does not introduce an additional bias or otherwise spoil the good properties of the method presented here.

Figure 7 shows the reconstructed mass as a function of input mass in this more realistic scenario. The 68% and 95% confidence intervals are now wider and the reconstructed mass does not appear to be as accurate. For input masses above ~ 100 GeV, the uncertainties become very wide, with only a lower limit of $m_{\text{rec}} > 20$ GeV being placed on the WIMP mass. Due to the poorer energy resolution the shape of the energy spectrum is less well-determined. In addition, a flat background contribution can mimic a higher mass WIMP, as it leads to a flatter spectrum. This leads to a strong degeneracy, as a wide range of mass values can provide a good fit to the data. For high input masses, the profile likelihood is approximately constant above $m_{\text{rec}} \sim 20$ GeV, indicating that there is no sensitivity to the underlying WIMP mass.

In spite of this, the true mass values still lie within the 68% and 95% confidence intervals. In addition, the poor values for the reconstructed mass for heavy WIMPs are a side effect of the loss of sensitivity. Because the profile likelihood is approximately flat, the maximum likelihood point is equally likely to be anywhere within the 68% interval. These effects would be present even if we had considered a fixed form for the speed distribution. However, when we allow for a range of possible speed distributions, the effects become more pronounced. These results show that for more realistic experimental scenarios, the method presented in this paper remains reliable over a range of masses, though its precision may be significantly reduced.

VI. STATISTICAL PROPERTIES

We now consider the impact of statistical fluctuations on the reconstruction of the WIMP mass. In reality, the number of events observed N_o at a given experiment will be Poisson distributed about the expected value N_e , while the observed distribution of recoil energies will not exactly match that expected from the calculated event rate. The fundamental statistical limitations of future direct detection experiments have been studied in detail in Ref. [64]. In this work, we generate 250 realisations

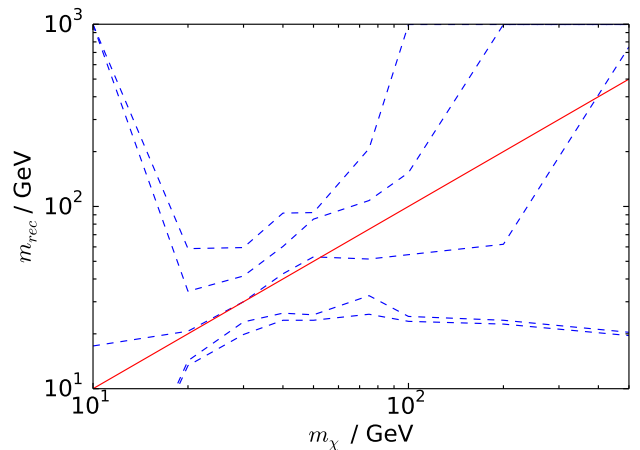


FIG. 7. As fig. 6 but including the effects of finite energy resolution and non-zero backgrounds, as described in the text.

of data from the mock experiments described in Tab. II. Each realisation of the mock data is generated as follows:

1. Calculate the number of expected events N_e , given $\{m_\chi, \sigma_p, f(v)\}$, using Eq. 13,
2. Pick the number of observed events N_o from a Poisson distribution with mean N_e ,
3. Pick recoil energies $\{E_1, E_2, \dots, E_{N_o}\}$, from the distribution $P(E)$ in Eq. 14,
4. Repeat for all three experiments.

For each realisation, we then use the method described in Sec. III (using $N = 5$ basis functions) to reconstruct the WIMP mass and 68% and 95% credible intervals. Figure 8 shows the distribution of reconstructed masses for an input mass of 50 GeV for three benchmark speed distributions: SHM, SHM+DD and Lisanti et al. as described in Sec. II. In all three cases, the reconstructions are peaked close to the true value, regardless of the underlying distribution. For the SHM+DD distribution, the spread of reconstructions is slightly wider (with more reconstructions extending up to higher masses). This is due to the smaller number of events for this benchmark, making the data sets more susceptible to Poisson fluctuations.

In order to assess the accuracy of the reconstructed value of the mass m_{rec} , we also calculate the bias b for each realisation:

$$b = \ln(m_{\text{rec}}/\text{GeV}) - \ln(m_{\text{true}}/\text{GeV}). \quad (22)$$

We compare the logarithms of the mass values because we have used logarithmically-flat priors on the WIMP mass. In Tab. V we show the average bias across all 250 realisations for each of the three benchmark distributions. In

Benchmark speed distribution	Mean bias $\langle b \rangle$
SHM	0.002 ± 0.008
SHM+DD	0.005 ± 0.007
Lisanti et al.	0.01 ± 0.01

TABLE V. Mean bias $\langle b \rangle$ in the reconstructed log WIMP mass (Eq. 22). This was calculated over 250 realisations using three different benchmark speed distributions.

Benchmark speed distribution	68% coverage	95% coverage
SHM	$71 \pm 3 \%$	$94 \pm 3 \%$
SHM+DD	$68 \pm 3 \%$	$91 \pm 4 \%$
Lisanti et al.	$70 \pm 3 \%$	$95 \pm 3 \%$

TABLE VI. Coverage of 68% and 95% credible intervals calculated from 250 data realisations each for three benchmark speed distributions. The concept of coverage is described in the text of Sec. VI.

all three cases, the average bias is consistent with zero. Even in the SHM+DD case, which shows larger fluctuations away from the true value, there is no statistical bias.

We also test the *coverage* of the credible intervals which have been constructed. For a $p\%$ credible interval, we expect that the true parameter value of the WIMP mass will lie within the interval in $p\%$ of realisations. In this case, we say that the method provides *exact coverage*. However, if the true parameter lies within the interval in fewer than $p\%$ of realisations, our reconstructed credible intervals are too narrow and provide *undercoverage*. Alternatively, we obtain *overcoverage* when the true parameter lies within the interval more often than $p\%$ of the time. Table VI shows the coverage values for the 68% and 95% intervals obtained in this section. In each case, there is very close to exact coverage. We have also checked that these intervals only provide exact coverage for the true WIMP mass of 50 GeV. Other values of m_{rec} are contained within the intervals less frequently than the true value, again indicating that this parametrization allows for unbiased and statistically robust reconstructions of the WIMP mass.

VII. RECONSTRUCTING $f_1(v)$

Using the method described in this paper, we can obtain the posterior probability distribution for the coefficients $\{a_1, \dots, a_{N-1}\}$ given the data, which we refer to as $P(\mathbf{a})$. We would like to be able to present this information in terms of the distribution function $f_1(v)$ in order to compare with some known distribution or look

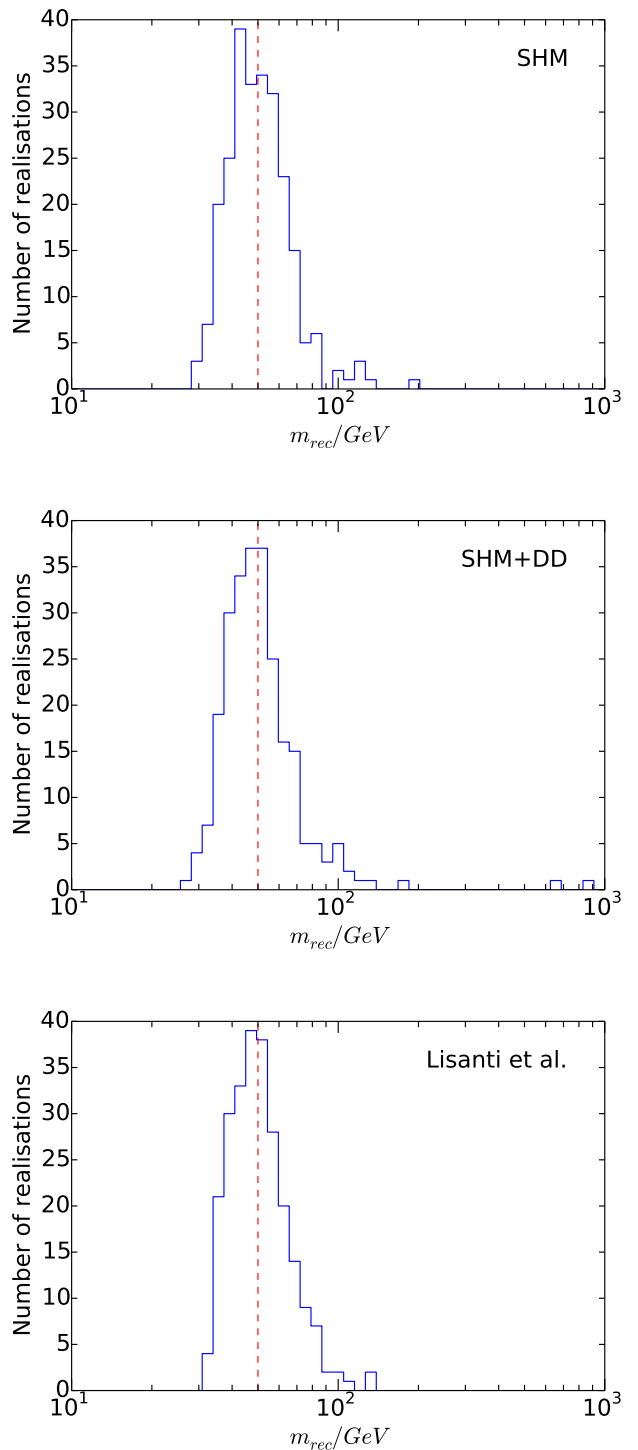


FIG. 8. Distribution of the reconstructed mass m_{rec} for 250 mock data sets generated using several benchmark speed distributions, defined in Sec. II. These are the SHM (top), SHM+DD (middle) and Lisanti et al. (bottom) distributions. The input WIMP mass of $m_\chi = 50$ GeV is shown as a vertical dashed red line.

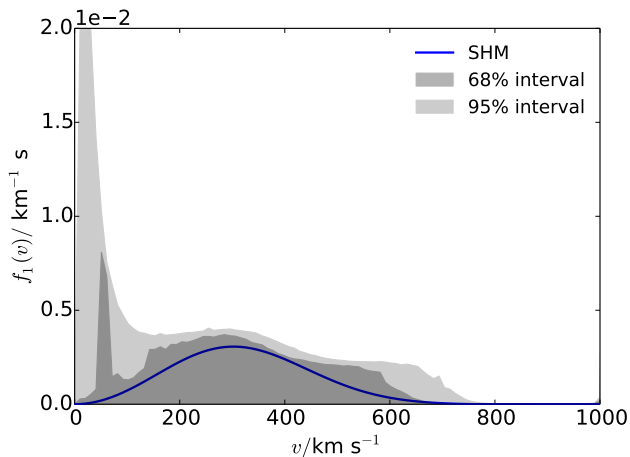


FIG. 9. Reconstructed speed distribution for a single realisation of data, generated for a 50 GeV WIMP. 68% and 95% credible intervals are shown as dark and light shaded regions respectively, while the underlying SHM distribution function is shown as a solid blue line.

for particular features in the distribution. However, due to the fact that the distribution function is normalized, the values of f_1 at different speeds will be strongly correlated. We illustrate here how robust comparisons with benchmark distributions can be made.

As a first step, we can attempt to sample from the $P(\mathbf{a})$, in order to obtain $P(f_1(v))$. This is the probability distribution for the value of f_1 at a particular speed v , marginalizing over the values of f_1 at all other speeds. We can repeat for a range of speeds to obtain 68% and 95% credible intervals for the whole of $f_1(v)$. The result of this procedure is presented in Fig. 9, for a randomly selected realisation from the SHM ensemble of Sec. VI. The underlying SHM distribution is shown as a solid line, while the 68% and 95% marginalized intervals are shown as dark and light shaded regions respectively. In this naive approach, we see that there is little shape information which can be recovered from the reconstruction, with only upper limits being placed on the speed distribution.

This method performs poorly because, as initially mentioned in Sec. III, we have no information about the fraction of dark matter particles below the energy threshold of our experiments. If this fraction is large, the event rate for a given cross-section is suppressed. However, increasing the cross-section will increase the total event rate. There is thus a degeneracy between the shape of the speed distribution and the cross-section, meaning that we can only probe the shape of $f_1(v)$, rather than its overall normalization. This degeneracy has not been accounted for in Fig. 9. We can attempt to correct for this by adjusting the normalization of $f_1(v)$. If we fix $f_1(v)$ to be normalized to unity above v_a (where $v_a \approx 171 \text{ km s}^{-1}$ is the lowest speed probed by the experiments for a WIMP of mass 50 GeV), we can compare the shapes of the un-

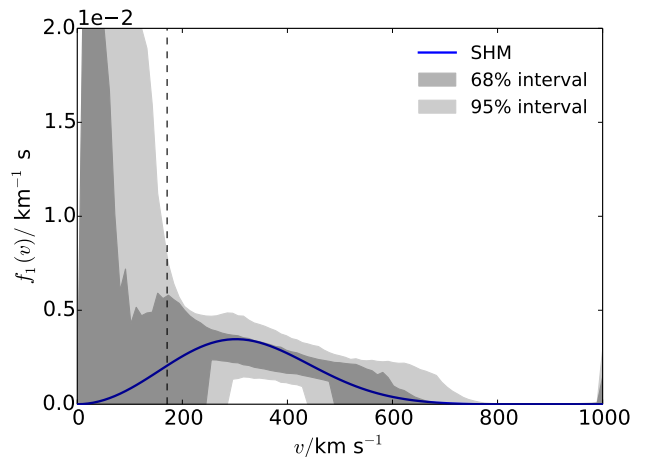


FIG. 10. Reconstructed speed distribution for the same realisation of data as Fig. 9. In this case, we have also normalized $f_1(v)$ to unity above $v_a \approx 171 \text{ km s}^{-1}$ (vertical dashed line). This is the lowest speed accessible to the experiments for a WIMP of mass 50 GeV. 68% and 95% credible intervals are shown as dark and light shaded regions respectively, while the underlying SHM distribution function is shown as a solid blue line.

derlying and reconstructed distribution functions. This is illustrated in Fig. 10, which shows that we now broadly reconstruct the correct shape of $f_1(v)$. Below v_a , the value of $f_1(v)$ is poorly constrained, because the experiments provide no information about the shape of the distribution below threshold.

There remain several issues with this approach. In order to utilize this method, we must know the approximate value of the lowest speed probed by the experiments. However, this value is set by the WIMP mass. We could determine v_a using the reconstructed WIMP mass, but this would be subject to significant uncertainty. In addition, direct reconstructions of the speed distribution are easily biased. The upper limit of the energy windows of the experiments corresponds to a particular WIMP speed (for a given WIMP mass). WIMPs above this speed still contribute to the total event rate, but contribute no spectral information. The reconstructed shape of the high speed tail of the distribution is therefore not constrained by the data, but may affect the reconstructed value of f_1 at lower speeds.

An alternative approach is to reconstruct the mean inverse speed $\eta(v)$ (defined in Eq. 2) at some speed v . Because $\eta(v)$ is an integral function of f_1 , it is less prone to bias as it takes into account the full shape of the distribution at speeds greater than v . However, we do not know the normalization of f_1 and so we must normalize η appropriately. For each point sampled from $P(\mathbf{a})$, we calculate η . We then divide by $\alpha(v)$, the fraction of WIMPs above speed v , calculated using the same param-

eter point:

$$\alpha(v) = \int_v^\infty f_1(v') dv'. \quad (23)$$

We will write this rescaled mean inverse speed as $\eta^*(v) = \eta(v)/\alpha(v)$. The value of $\eta^*(v)$ is a measure of the shape of the distribution function above v . However, information about the normalization of the distribution has been factored out by dividing by $\alpha(v)$. We no longer need to know the value of v_a in order to obtain information about the shape of the distribution at higher speeds. We may still need to decide the speed down to which we trust our reconstruction, but this no longer relies on an arbitrary choice of v_a to normalize the reconstructions at all speeds.

In Fig. 11, we plot the mean reconstructed value of η^* at several values of v , using 250 realisations of the 50 GeV SHM benchmark. We also show the mean upper and lower limits of the 68% credible intervals as errorbars. The form of η^* for the SHM is shown as a solid blue line. In all cases except for $v = 100 \text{ km s}^{-1}$, the mean reconstructed value is close to the true value, indicating that η^* can be reconstructed without bias using this method. At low speeds, the reconstructed value deviates from the true value. In addition, the credible intervals lead to *undercoverage* in the $v = 100 \text{ km s}^{-1}$ case. However, this point lies below the lowest speed to which the experiments are sensitive and therefore we cannot trust the reconstruction at this low speed. We have checked that for the remaining values of v the method provides exact or overcoverage, indicating that at higher speeds we can use η^* as a reliable and statistically robust measure of the shape of the distribution.

In the case of a single realisation of data, we would like to compare the probability distribution for $\eta^*(v)$ (obtained from $P(\mathbf{a})$) to the value calculated from some test distribution. We note that several distributions may produce the same value of $\eta^*(v)$ at a given value of v . Thus, we may fail to reject a distribution which is not the true distribution. However, if the calculated value of $\eta^*(v)$ does lie outside the $p\%$ interval, we can reject it at the $p\%$ level.

We can increase the discriminating power of this method by repeating this reconstruction over all speeds and checking to see if the benchmark value of η^* is rejected at any value of v . The result of this procedure is shown in Fig. 12 for a single realisation of data generated using an SHM distribution (the same data as in Figs. 9 and 10). We plot the 68%, 95% and 99% credible intervals as shaded regions, as well as the values of $\eta^*(v)$ calculated from several benchmark speed distribution. We will focus on the intermediate speed range ($v \gtrsim 200 \text{ km s}^{-1}$), as we do not know *a priori* the lowest speed to which the experiments are sensitive.

The reconstructed intervals are consistent with a range of possible distribution functions. The SHM and SHM+DD distributions are identical over a wide range of speeds. This is because above $\sim 200 \text{ km s}^{-1}$, the two

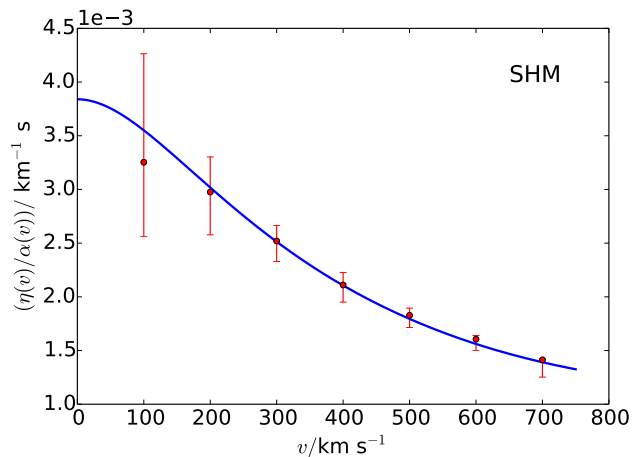


FIG. 11. Mean reconstructed values of the rescaled mean inverse speed $\eta(v)/\alpha(v)$ at several values of v , calculated over 250 realisations of data using a 50 GeV WIMP and underlying SHM distribution function. Errorbars indicate the mean upper and lower limits of the 68% credible intervals. The underlying form of $\eta(v)/\alpha(v)$ obtained from the SHM is shown as a solid blue line.

distributions differ in normalization but not in shape. Differences appear between the two at low speeds where their shapes diverge. The Lisanti et al. distribution results in a larger deviation from the SHM, but not sufficiently large to differentiate between the two distributions given the size of the uncertainties. Finally, the stream distribution results in a significantly different form for $\eta^*(v)$. At approximately 400 km s^{-1} , the curve for the stream distribution lies outside the reconstructed 99% credible interval. We can therefore use this method to reject the stream distribution at the 99% confidence level.

Figure 13 shows the results of a reconstruction using a larger exposure. In this case, we generate data using the Lisanti et al. distribution and an exposure increased by a factor of 2.5, resulting in approximately 1000 events across the three detectors. As expected, the resulting credible intervals are now substantially narrower. The stream distribution now lies significantly outside the 99% interval. In Fig. 14, we show the same results, but focusing in on the region around $v \sim 400 \text{ km s}^{-1}$. At certain points, the SHM and SHM+DD distributions now lie outside the 95% credible interval, suggesting that with a number of events of the order of 1000, we may be able to reject these benchmarks.

While the method displayed in Fig. 10 allows the approximate shape of the speed distribution to be reconstructed, reconstructions of $\eta^*(v)$ allow more statistically robust statements to be made about the underlying speed distribution. In particular, Fig. 14 illustrates that with larger exposures deviations from Maxwellian speed distributions can be detected in a model-independent fashion.

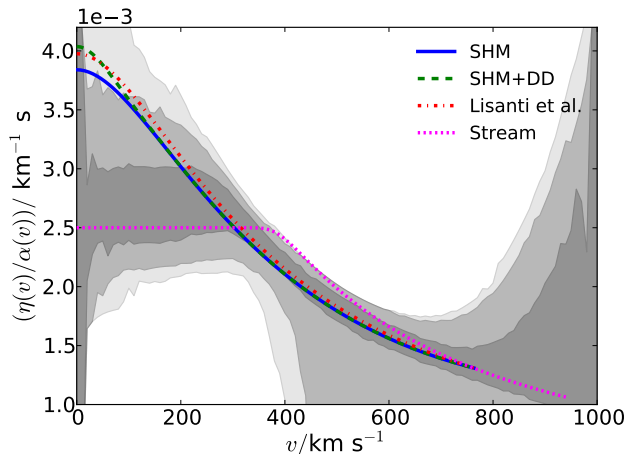


FIG. 12. Rescaled mean inverse speed $\eta(v)/\alpha(v)$, reconstructed from a single realisation of data using a 50 GeV WIMP and underlying SHM distribution function. At each value of v we calculate 68%, 95% and 99% credible intervals (shown as shaded intervals). We also show the calculated values of $\eta(v)/\alpha(v)$ for several possible benchmark speed distributions: SHM (solid blue), SHM+DD (dashed green), Lisanti et al. (dot-dashed red) and stream (dotted magenta). The benchmark curves are truncated when the underlying distribution function goes to zero.

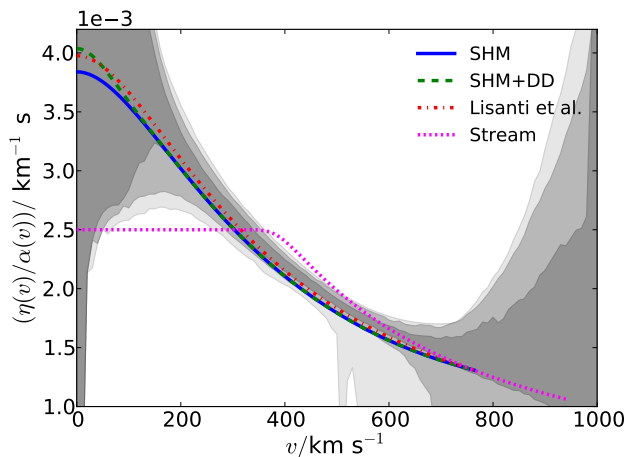


FIG. 13. As Fig. 12, but using as input a Lisanti et al. speed distribution and an exposure time which is 2.5 times longer.

VIII. CONCLUSIONS

We have studied in detail the parametrization for the local dark matter speed distribution introduced in Paper 1. This method involves writing the logarithm of the speed distribution as a polynomial in speed v and fitting the polynomial coefficients (along with the WIMP mass and cross section) to the data. We have attempted to disentangle in this paper the influence of different bench-

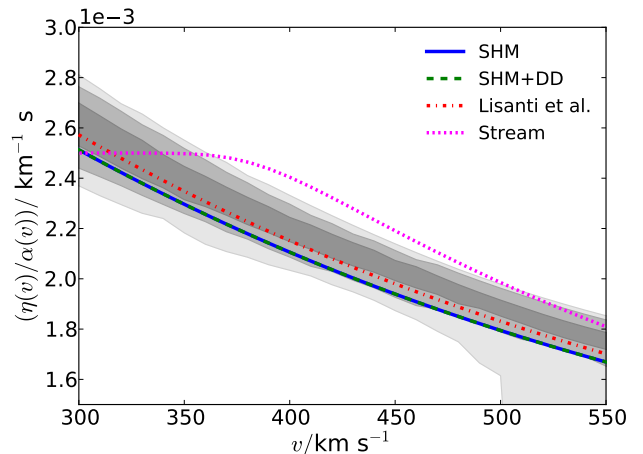


FIG. 14. As Fig. 13, but focusing on the region around $v \sim 400 \text{ km s}^{-1}$. Notice that in the range $400 - 550 \text{ km s}^{-1}$, both the SHM and SHM+DD curves lie at or below the lower limit of the 95% credible interval.

mark speed distributions, different benchmark WIMP masses and different forms for the parametrization. We summarize our conclusions as follows:

- We have shown that the reconstruction of the WIMP mass is robust under changes in the number of basis functions N . We have used the Bayesian Information Criterion (BIC) to compare models with different values of N and have shown that minimizing the BIC allows us to determine how many basis functions are required for a reliable reconstruction. We have also demonstrated that the results of the method do not depend strongly on the choice of basis functions, but that the speed of reconstructions may be improved by using the Chebyshev polynomial basis.
- We have shown that the method leads to unbiased reconstructions of the WIMP mass for masses in the range 10-500 GeV. Including realistic experimental parameters, including non-zero backgrounds and finite energy resolution, reduces the precision of these reconstructions. In particular, for large values of the input mass, we can only place a lower limit of approximately 20 GeV on the reconstructed mass. This is significantly lower than in the idealized case, where we can typically constrain the WIMP mass to be heavier than around 50 GeV.
- We have used several ensembles of data realisations to demonstrate the statistical properties of the method, including unbiased reconstructions and exact coverage of the WIMP mass.
- We have presented several ways of displaying the reconstructed WIMP speed distribution using this

method. In order to make robust statistical inferences about the speed distribution, we calculate the probability distribution of $\eta(v)/\alpha(v)$. This is the mean inverse speed $\eta(v)$, which appears in the direct detection event rate (eq. 1), rescaled by the fraction of WIMPs $\alpha(v)$ above speed v . This can be used as a measure of the *shape* of the distribution function, from which the unknown normalization has been factored out. We can then compare to the expected value of $\eta(v)/\alpha(v)$ from a given benchmark speed distribution, allowing us to distinguish between different underlying models.

We have shown that this parametrization method is statistically robust and works well over a large range of input parameters, both in terms of particle physics and astrophysics. The inclusion of more realistic experimen-

tal parameters does not introduce any additional bias, but does reduce the precision of reconstructions. We obtain unbiased estimates of the WIMP mass over large numbers of data sets. Finally, we have shown that we can distinguish different forms of the speed distribution. With around 1000 events, it may be possible to detect minor deviations from the Standard Halo Model and begin to search for more interesting structure in the speed distribution of the Milky Way.

ACKNOWLEDGMENTS

The author thanks Anne M. Green and Mattia Fornasa for helpful comments. BJK is supported by STFC. Access to the University of Nottingham High Performance Computing Facility is also gratefully acknowledged.

-
- [1] G. Bertone, D. Hooper, and J. Silk, *Phys. Rep.* **405**, 279 (2005), hep-ph/0404175.
 - [2] G. Jungman, M. Kamionkowski, and K. Griest, *Phys. Rep.* **267**, 195 (1996), hep-ph/9506380.
 - [3] S. Dodelson and L. Widrow, *Phys. Rev. Lett.* **72**, 17 (1994), hep-ph/9303287.
 - [4] L. D. Duffy and K. v. Bibber, *New J. Phys.* **11**, 105008 (2009), arXiv:0904.3346.
 - [5] E. W. Kolb and R. Slansky, *Phys. Lett. B* **135**, 378 (1984).
 - [6] M. W. Goodman and E. Witten, *Phys. Rev. D* **31**, 3059 (1985).
 - [7] A. K. Drukier, K. Freese, and D. N. Spergel, *Phys. Rev. D* **33**, 3495 (1986).
 - [8] J. Lavalle and P. Salati, *C. R. Phys.* **13**, 740 (2012), arXiv:1205.1004.
 - [9] M. Battaglia and M. E. Peskin, in *Particle Dark Matter: Observation, Models and Searches*, edited by G. Bertone (Cambridge University Press, 2010) Chap. 14.
 - [10] A. M. Green, *J. Cosmol. Astropart. Phys.* **10**, 034 (2010), arXiv:1009.0916.
 - [11] A. H. G. Peter, *Phys. Rev. D* **83**, 125029 (2011), arXiv:1103.5145.
 - [12] M. Fairbairn, T. Douce, and J. Swift, *J. Astropart. Phys.* **47**, 45 (2013), arXiv:1206.2693.
 - [13] M. Lisanti, L. E. Strigari, J. G. Wacker, and R. H. Wechsler, *Phys. Rev. D* **83**, 023519 (2010), arXiv:1010.4300.
 - [14] P. Bhattacharjee, S. Chaudhury, S. Kundu, and S. Majumdar, *Phys. Rev. D* **87**, 083525 (2013), arXiv:1210.2328.
 - [15] M. Vogelsberger, A. Helmi, V. Springel, S. D. M. White, J. Wang, C. S. Frenk, A. Jenkins, A. Ludlow, and J. F. Navarro, *Mon. Not. R. Astron. Soc.* **395**, 797 (2009), arXiv:0812.0362.
 - [16] M. Kuhlen, N. Weiner, J. Diemand, P. Madau, B. Moore, D. Potter, J. Stadel, and M. Zemp, *J. Cosmol. Astropart. Phys.* **02**, 030 (2010), arXiv:0912.2358.
 - [17] M. Kuhlen, M. Lisanti, and D. N. Spergel, *Phys. Rev. D* **86**, 063505 (2012), arXiv:1202.0007.
 - [18] Y.-Y. Mao, L. E. Strigari, R. H. Wechsler, H.-Y. Wu, and O. Hahn, *Astrophys. J.* **764**, 35 (2012), arXiv:1210.2721.
 - [19] J. I. Read, L. Mayer, A. M. Brooks, F. Governato, and G. Lake, *Mon. Not. R. Astron. Soc.* **397**, 44 (2009), arXiv:0902.0009.
 - [20] J. I. Read *et al.*, *AIP Conf. Proc.* **1240**, 391 (2010), arXiv:0901.2938.
 - [21] M. Kuhlen, A. Pillepich, J. Guedes, and P. Madau, (2013), arXiv:1308.1703.
 - [22] L. E. Strigari and R. Trotta, *J. Cosmol. Astropart. Phys.* **11**, 019 (2009), arXiv:0906.5361.
 - [23] A. H. G. Peter, *Phys. Rev. D* **81**, 087301 (2010), arXiv:0910.4765.
 - [24] M. Pato, L. Baudis, G. Bertone, R. R. de Austri, L. E. Strigari, and R. Trotta, *Phys. Rev. D* **83**, 083505 (2011), arXiv:1012.3458.
 - [25] M. Pato, L. E. Strigari, R. Trotta, and G. Bertone, *J. Cosmol. Astropart. Phys.* **02**, 041 (2013), arXiv:1211.7063.
 - [26] M. Drees and C.-L. Shan, *J. Cosmol. Astropart. Phys.* **06**, 011 (2007), astro-ph/0703651.
 - [27] M. Drees and C.-L. Shan, *J. Cosmol. Astropart. Phys.* **06**, 012 (2008), arXiv:0803.4477.
 - [28] B. J. Kavanagh and A. M. Green, *Phys. Rev. Lett.* **111**, 031302 (2013), arXiv:1303.6868.
 - [29] J. D. Lewin and P. F. Smith, *J. Astropart. Phys.* **6**, 87 (1996).
 - [30] R. Helm, *Phys. Rev.* **104**, 1466 (1956).
 - [31] D. Smith and N. Weiner, *Phys. Rev. D* **64**, 043502 (2001), hep-ph/0101138.
 - [32] A. Kurylov and M. Kamionkowski, *Phys. Rev. D* **69**, 063503 (2003), hep-ph/0307185.
 - [33] J. Fan, M. Reece, and L.-T. Wang, *J. Cosmol. Astropart. Phys.* **11**, 042 (2010), arXiv:1008.1591.
 - [34] A. L. Fitzpatrick, W. Haxton, E. Katz, N. Lubbers, and Y. Xu, (2012), arXiv:1211.2818.
 - [35] A. L. Fitzpatrick, W. Haxton, E. Katz, N. Lubbers, and Y. Xu, *J. Cosmol. Astropart. Phys.* **02**, 004 (2013), arXiv:1203.3542.
 - [36] K. Freese, M. Lisanti, and C. Savage, (2013), arXiv:1209.3339.
 - [37] M. C. Smith *et al.*, *Mon. Not. R. Astron. Soc.* **379**, 755 (2007), astro-ph/0611671.

- [38] T. Piffi *et al.*, (2013), arXiv:1309.4293.
- [39] F. Feroz, M. P. Hobson, and M. Bridges, *Mon. Not. R. Astron. Soc.* **398**, 1601 (2008), arXiv:0809.3437.
- [40] F. Feroz and M. P. Hobson, *Mon. Not. R. Astron. Soc.* **384**, 449 (2007), arXiv:0704.3704.
- [41] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt, (2013), arXiv:1306.2144.
- [42] B. J. Kavanagh and A. M. Green, *Phys. Rev. D* **86**, 065027 (2012), arXiv:1207.2039.
- [43] F. Iocco, M. Pato, G. Bertone, and P. Jetzer, *J. Cosmol. Astropart. Phys.* **11**, 029 (2011), arXiv:1107.5810.
- [44] J. Bovy and S. Tremaine, *Astrophys. J.* **756**, 89 (2012), arXiv:1205.4033.
- [45] L. Zhang, H.-W. Rix, G. van de Ven, J. Bovy, C. Liu, and G. Zhao, *Astrophys. J.* **772**, 108 (2013), arXiv:1209.0256.
- [46] F. Nesti and P. Salucci, *J. Cosmol. Astropart. Phys.* **07**, 016 (2013), arXiv:1304.5127.
- [47] A. H. G. Peter, V. Gluscevic, A. M. Green, B. J. Kavanagh, and S. K. Lee, (2013), arXiv:1310.7039.
- [48] E. Aprile (XENON Collaboration), (2012), arXiv:1206.6288.
- [49] E. Aprile *et al.* (XENON100 Collaboration), *Phys. Rev. Lett.* **109**, 181301 (2012), online supplementary material, arXiv:1207.5988.
- [50] E. Aprile, “XENON1T at LNGS: Proposal April 2010,” www.bo.infn.it/xenon/docs/xe1t_proposal_v2.pdf (2010).
- [51] P. Benetti *et al.*, *J. Astropart. Phys.* **28**, 495 (2007), arXiv:astro-ph/0701286.
- [52] L. Grandi, *WARP: an argon double phase technique for Dark Matter search*, Ph.D. thesis, Department of Nuclear and Theoretical Physics, University of Pavia (2005).
- [53] D. Bauer, “SuperCDMS SNOLAB,” <https://indico.fnal.gov/getFile.py/access?sessionId=1&resId=0&materialId=0&confId=6584> (2013).
- [54] D. Bauer, “SuperCDMS update,” www.fnal.gov/directorate/program_planning/all_experimenter_meetings/special_reports/Bauer_supercdms_status_02_04_13.pdf (2013).
- [55] A. Badertscher *et al.* (ArDM Collaboration), (2013), arXiv:1307.0117.
- [56] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* **71**, 1 (2013), arXiv:1007.1727.
- [57] E. Aprile *et al.* (XENON100 Collaboration), *Phys. Rev. D* **84**, 052003 (2011), arXiv:1103.0303.
- [58] Z. Ahmed *et al.* (CDMS Collaboration), *Phys. Rev. Lett.* **102**, 011301 (2009), arXiv:0802.3530.
- [59] G. Schwarz, *Ann. Stat.* **6**, 461 (1978).
- [60] W. Gautschi, in *Recent advances in numerical analysis*, edited by C. De Boor and G. H. Golub (Academic Press, New York, 1978) pp. 45–72.
- [61] J. H. Wilkinson, in *Studies in numerical analysis*, edited by G. H. Golub (Mathematical Association of America, 1984) p. 3.
- [62] J. C. Mason and D. C. Handscomb, *Chebyshev polynomials*, 1st ed. (Chapman & Hall/CRC, 2002).
- [63] A. M. Green, *J. Cosmol. Astropart. Phys.* **06**, 005 (2008), arXiv:0805.1704.
- [64] C. Strege, R. Trotta, G. Bertone, A. H. G. Peter, and P. Scott, *Phys. Rev. D* **86**, 023507 (2012), arXiv:1201.3631.