

**Directional detection of dark matter streams**Ciaran A. J. O’Hare<sup>\*</sup> and Anne M. Green<sup>†</sup>*School of Physics and Astronomy, University of Nottingham,  
University Park, Nottingham NG7 2RD, United Kingdom  
(Received 15 October 2014; published 8 December 2014)*

Directional detection of weakly interacting massive particles, in which the energies and directions of the recoiling nuclei are measured, currently presents the only prospect for probing the local *velocity* distribution of Galactic dark matter. We investigate the extent to which future directional detectors would be capable of probing dark matter substructure in the form of streams. We analyze the signal expected from a Sagittarius-like stream and also explore the full parameter space of stream speed, direction, dispersion and density. Using a combination of nonparametric directional statistics, a profile likelihood ratio test and Bayesian parameter inference we find that within acceptable exposure times [ $\mathcal{O}(10)$  kg yr for cross sections just below the current exclusion limits] future directional detectors will be sensitive to a wide range of stream velocities and densities. We also examine and discuss the importance of the energy window of the detector.

DOI: 10.1103/PhysRevD.90.123511

PACS numbers: 95.35.+d

**I. INTRODUCTION**

Attempts to directly detect weakly interacting massive particles (WIMPs) using liquid and solid-state detectors have a long history. A key goal of this field is the detection of one (or more) of the “smoking gun” signals of direct detection: annual modulation [1,2], material dependence (e.g. Ref. [3]) and direction dependence [4]. Much theoretical work has been carried out investigating how the WIMP particle physics (mass and cross section) and astrophysics (local density and velocity distribution) could be inferred from the energies (e.g. Refs. [3,5,6]) or energies and directions [7–11] of WIMP induced nuclear recoil events.

The directionality of the WIMP event rate is due to the motion of the Sun with respect to the Galactic halo. We are moving toward the constellation Cygnus, and hence the nuclear recoils are expected to be strongly peaked in the direction opposite to this. The strength of the signal is expected to be large; an anisotropic set of recoils can be discriminated from isotropic backgrounds with as few as 10 events [7,12], while the peak recoil direction can be measured, and the Galactic origin of the scattering particle confirmed, with around 30–50 events [13,14].

Practically directional detection is still in its early stages, with a number of prototype detectors in the process of development. Directional detection is typically achieved using gas time projection chambers (TPCs) which contain low pressure gases such as CF<sub>4</sub>, CS<sub>2</sub>, C<sub>4</sub>H<sub>10</sub> or <sup>3</sup>He (see Ref. [15] for a review). After an interaction with a WIMP, the recoiling nucleus leaves an ionization track in its wake. The three-dimensional reconstruction of the recoil track

then allows the full energy and direction dependence of the WIMP recoil spectrum to be measured and, in principle, the WIMP velocity distribution can be inferred [16]. The detectors currently in the prototype phase include DMTPC [17], DRIFT [18], MIMAC [19], and NEWAGE [20]. Directional detection offers several theoretical advantages over its nondirectional counterpart. First and foremost, there are no known backgrounds able to mimic the WIMP signal in its characteristic peak direction. Furthermore it offers the only prospect for constraining the local *velocity* distribution of the dark matter.

The dependence of the experimental signals on the form of the local WIMP velocity distribution has attracted a lot of attention in the literature, as there is significant uncertainty in the form of the velocity distribution, and the parameters on which it depends (for a review see Ref. [21]). Data from direct detection experiments are typically compared using the standard halo model (SHM), for which the velocity distribution is an isotropic Maxwell-Boltzmann distribution. The shape of the true local velocity distribution is expected to depart significantly from this simple model [22–25]. However, the use of the SHM is at least somewhat justified for several reasons. First there is no fully agreed upon alternative parametrization of the velocity distribution, and second the expected deviations from the SHM are unlikely to affect the analysis of data from the *current* generation of nondirectional detectors, provided the free parameters are appropriately marginalized over [26].

Nevertheless, results from *N*-body and hydrodynamical simulations of galaxy formation show deviations from the putative smooth and isotropic SHM, including tidal streams [27], a dark disk [28,29] and debris flows [30,31], which could be detectable by future experiments. In this paper we focus on substructure in the form of tidal streams, which result from the disruption of subhalos accreted by the

<sup>\*</sup>ciaran.ohare@nottingham.ac.uk  
<sup>†</sup>anne.green@nottingham.ac.uk

Milky Way. There are hints that such a feature may pass through the Solar neighborhood. A contiguous group of stars moving with a common velocity, which are possibly part of a tidal tail from the Sagittarius dwarf galaxy, have been observed nearby [32–35]. Moreover it has been argued that the dark matter content of the stream is likely to be significantly more extended than the stellar content and to have an offset of as much as a few kpc [36].

We examine the capabilities of future directional dark matter detectors to detect a dark matter stream. In Sec. II, we overview the calculation of the directional event rate, including the relevant particle physics and astrophysics input. We will begin our statistical analysis in Sec. III with nonparametric directional tests, studying first the example of a Sagittarius-like stream and then extending the analysis to streams in general. Then in Sec. IV we perform likelihood analyses, using Bayesian inference to reconstruct the parameters of a Sagittarius-like stream and a profile likelihood ratio test to assess the detectability of general streams in a parametric way. Finally we conclude with a discussion of our results in Sec. V

## II. DIRECTIONAL DETECTION

### A. Particle physics

The directional event rate is typically written in the form of a double differential [4,16,37],

$$\frac{d^2R}{dE d\Omega_q} = \frac{\rho_0 \sigma_N}{4\pi m_\chi \mu^2} F^2(E) \hat{f}_{\text{lab}}(v_{\min}(E), \hat{\mathbf{q}}). \quad (1)$$

Here,  $R$  is the event rate per unit time per unit mass of detector which is expressed as a function of the direction and energy of the recoiling nucleus  $\mathbf{E} = E\hat{\mathbf{q}}$  in the laboratory frame, where  $d\Omega_q$  is the solid angle element around the recoil direction  $\hat{\mathbf{q}}$ . The local dark matter density is denoted by  $\rho_0$ , the WIMP-nucleus scattering cross section by  $\sigma_N$ , the WIMP mass by  $m_\chi$  and  $\hat{f}_{\text{lab}}(v_{\min}(E), \hat{\mathbf{q}})$  is the Radon transform of the local WIMP velocity distribution in the lab frame [16], where

$$v_{\min} = \frac{\sqrt{2m_N E}}{2\mu} \quad (2)$$

is the minimum WIMP speed that can scatter to create a recoil of energy  $E$ . The WIMP-nucleus reduced mass is  $\mu = m_\chi m_N / (m_\chi + m_N)$ , where  $m_N$  is the nucleus mass.

The nuclear form factor,  $F(E)$ , is the Fourier transform of the nuclear density distribution. It accounts for the decrease in the effective cross section for scattering events with nonzero momentum transfer. We use the Helm form factor,

$$F^2(E) = \left[ \frac{3j_1(qr_n)}{qr_n} \right]^2 e^{-q^2 s^2}, \quad (3)$$

where  $r_n = 1.2A^{1/3}$  fm, where  $A$  is the mass number of the nucleus and  $s = 0.9$  fm [37,38]. For light targets such as  $^{19}\text{F}$ , considered in this work, the form factor can be approximated as a simple exponential

$$F^2(E) = e^{-\alpha(qr_n)^2}, \quad (4)$$

where the value of  $\alpha$  depends on the nucleus under consideration. For  $^{19}\text{F}$ , we find  $\alpha = 1.14$ .

The WIMP-nucleus cross section  $\sigma_N$  can be written in terms of the experiment-independent WIMP-nucleon cross section  $\sigma_{p,n}$  which for spin-dependent scattering is

$$\sigma_N = \frac{4\mu_N^2 J + 1}{3\mu_p^2 J} \left[ \frac{a_p \langle S_p \rangle + a_n \langle S_n \rangle}{a_{p,n}} \right]^2 \sigma_{p,n}, \quad (5)$$

where  $\mu_p$  is the WIMP-proton reduced mass,  $J$  the total spin of the nucleus,  $a_{p,n}$  the model-dependent couplings to protons and neutrons, and  $\langle S_{p,n} \rangle$  are the expectation values of proton and neutron spins inside the nucleus [39].

### B. Astrophysics

The velocity distribution of WIMPs in the rest frame of the laboratory is obtained through a Galilean transformation of the Galactic frame distribution,  $f_{\text{gal}}(\mathbf{v})$ , by the laboratory velocity  $\mathbf{v}_{\text{lab}}$ ,

$$f_{\text{lab}}(\mathbf{v}) = f_{\text{gal}}(\mathbf{v} + \mathbf{v}_{\text{lab}}). \quad (6)$$

We ignore the effects of gravitational focusing of the distribution by the Earth and Sun, which is known to be small and takes place below the threshold energies of most detectors [40] and also the annual modulation due to the Earth's orbit, which is also only detectable with large numbers of events [41]. The lab velocity is then the sum of the velocity of the local standard of rest (LSR) and the peculiar velocity of the Sun:  $\mathbf{v}_{\text{lab}} = \mathbf{v}_{\text{LSR}} + \mathbf{v}_{\odot}$ . The velocities are represented in Cartesian coordinates, defined using the local Galactic coordinate system, oriented such that the LSR points in the  $y$  direction. The speed of the LSR is typically taken to be  $v_{\text{LSR}} = 220$  km s $^{-1}$  although there is some uncertainty in this value [42,43]. The most recent measurement of the peculiar velocity of the Sun is  $\mathbf{v}_{\odot} = (6.0, 10.6, 6.5) \pm (0.5, 0.8, 0.3)$  km s $^{-1}$  [44].

The SHM (which has a density profile  $\rho \sim r^{-2}$ , and hence a flat rotation curve) has an isotropic Maxwell-Boltzmann (MB) velocity distribution which is typically truncated by hand at the Galactic escape speed,  $v_{\text{esc}}$ ,

$$f_{\text{gal}}^{\text{MB}}(\mathbf{v}; \sigma_v, v_{\text{esc}}) = \frac{1}{N_{\text{esc}} (2\pi\sigma_v^2)^{3/2}} \exp\left(-\frac{|\mathbf{v}|^2}{2\sigma_v^2}\right) \theta(v_{\text{esc}} - |\mathbf{v}|), \quad (7)$$

where the normalization constant,  $N_{\text{esc}}$ , is given by

$$N_{\text{esc}} = \text{erf}\left(\frac{v_{\text{esc}}}{\sqrt{2}\sigma_v}\right) - \sqrt{\frac{2}{\pi}} \frac{v_{\text{esc}}}{\sigma_v} \exp\left(-\frac{v_{\text{esc}}^2}{2\sigma_v^2}\right). \quad (8)$$

The dispersion  $\sigma_v$  is related to the speed of the LSR via  $\sigma_v = v_{\text{LSR}}/\sqrt{2}$ , and we take the Galactic escape speed to be  $v_{\text{esc}} = 533 \text{ km s}^{-1}$  [45].

The function  $\hat{f}$  that appears in the differential recoil spectrum, Eq. (1), is the radon transform of this distribution at  $w = v_{\text{min}}$  which is defined as [46]

$$\hat{f}(w, \hat{\mathbf{q}}) = \int \delta(\mathbf{v} \cdot \hat{\mathbf{q}} - w) f(\mathbf{v}) d^3 v. \quad (9)$$

It is most efficient to calculate the Radon transform of the velocity distribution in the lab frame by first calculating it in the Galactic frame and then transforming to the lab frame via

$$\hat{f}_{\text{lab}}(v_{\text{min}}, \hat{\mathbf{q}}; \sigma_v, v_{\text{esc}}) = \hat{f}_{\text{gal}}(v_{\text{min}} + \mathbf{v}_{\text{lab}} \cdot \hat{\mathbf{q}}, \hat{\mathbf{q}}). \quad (10)$$

The calculation for the SHM velocity distribution, Eq. (7), yields [16]

$$\hat{f}_{\text{lab}}^{\text{MB}}(v_{\text{min}}, \hat{\mathbf{q}}; \sigma_v, v_{\text{esc}}) = \frac{1}{N_{\text{esc}}(2\pi\sigma_v^2)^{1/2}} \times \left[ \exp\left(-\frac{|v_{\text{min}} + \mathbf{v}_{\text{lab}} \cdot \hat{\mathbf{q}}|^2}{2\sigma_v^2}\right) - \exp\left(-\frac{v_{\text{esc}}^2}{2\sigma_v^2}\right) \right]. \quad (11)$$

The simplest way to incorporate substructure into a halo model is by combining two distributions, a background smooth halo model and an additional component corresponding to the substructure. We adopt the MB velocity distribution function for the background halo model, as deviations from this are likely to be too small to have an appreciable effect on the detection of streams. We assume

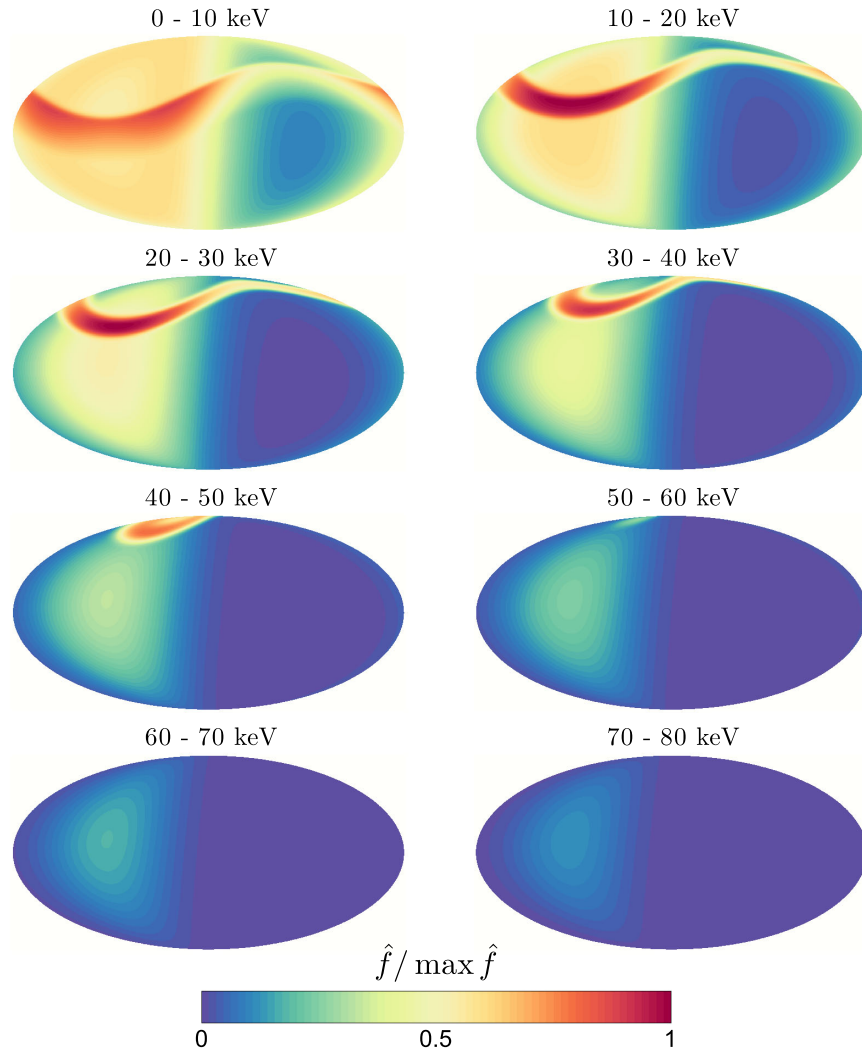


FIG. 1 (color online). Mollweide projection sky maps of the rescaled radon transform of the velocity distribution, integrated over 10 keV energy bins, for the halo + stream model described in Sec. II B, for a Sagittarius-like stream with density fraction  $\xi = 0.1$  for an  $m_\chi = 50 \text{ GeV}$  WIMP scattering off  $^{19}\text{F}$ .

that a single tidal stream makes up a fraction  $\xi = \rho_{\text{str}}/\rho_0$  of the local dark matter density, so that the total velocity distribution of the halo + stream model is given by

$$f_{\text{gal}}^{\text{h+s}}(\mathbf{v}) = (1 - \xi)f_{\text{gal}}^{\text{halo}} + \xi f_{\text{gal}}^{\text{str}}, \quad (12)$$

where

$$f_{\text{gal}}^{\text{halo}}(\mathbf{v}; \sigma_v, v_{\text{esc}}) = f_{\text{gal}}^{\text{MB}}(\mathbf{v}; \sigma_v, v_{\text{esc}}), \quad (13)$$

$$f_{\text{gal}}^{\text{str}}(\mathbf{v}; \mathbf{v}_{\text{str}}, \sigma_{\text{str}}, v_{\text{esc}}) = f_{\text{gal}}^{\text{MB}}(\mathbf{v} - \mathbf{v}_{\text{str}}; \sigma_{\text{str}}, v_{\text{esc}}). \quad (14)$$

For simplicity we have taken the velocity distribution of the stream to be a MB distribution too, but with dispersion  $\sigma_{\text{str}}$  and a mean velocity  $\mathbf{v}_{\text{str}}$  in the Galactic frame.

The full radon transform for the halo + stream model is therefore

$$\begin{aligned} \hat{f}_{\text{lab}}^{\text{h+s}}(v_{\text{min}}, \hat{\mathbf{q}}) &= (1 - \xi)\hat{f}_{\text{gal}}^{\text{MB}}(v_{\text{min}} + \mathbf{v}_{\text{lab}} \cdot \hat{\mathbf{q}}; \sigma_v, v_{\text{esc}}) \\ &+ \xi\hat{f}_{\text{gal}}^{\text{MB}}(v_{\text{min}} + (\mathbf{v}_{\text{lab}} - \mathbf{v}_{\text{str}}) \cdot \hat{\mathbf{q}}; \sigma_{\text{str}}, v_{\text{esc}}), \end{aligned} \quad (15)$$

where  $\hat{f}_{\text{gal}}$  is the radon transform of the Galactic frame Maxwell-Boltzmann distribution. Figure 1 shows the angular and energy dependence of the radon transform of the halo + stream distribution. We have taken the stream velocity and dispersion to correspond to a Sagittarius-like stream,  $\sigma_{\text{str}} = 10 \text{ km s}^{-1}$  and  $\mathbf{v}_{\text{str}} = 400 \times (0, 0.233, -0.970) \text{ km s}^{-1}$  [47,48]; however, we have set the stream density fraction to  $\xi = 0.1$  in order to make the feature more prominent in the image. This is larger than the maximum stream density found in simulations [27,49].

The top panel of Fig. 2 displays more clearly the energy dependence of the signal and the dependence on the WIMP mass. The energy dependence is found by integrating Eq. (1) over angles,

$$\frac{dR}{dE} = \int \frac{d^2R}{dE d\Omega_q} d\Omega_q. \quad (16)$$

We have fixed the values of the local density and the cross section to  $\rho_0 = 0.3 \text{ GeV cm}^{-3}$  and  $\sigma_p = 10^{-3} \text{ pb}$  in this and subsequent plots. In the bottom panel we show the difference between the differential event rate for the halo + stream and halo-only models defined as

$$\Delta_{\text{str}} = \left. \frac{dR}{dE} \right|_{\xi=0} - \frac{dR}{dE}. \quad (17)$$

The stream manifests itself in the energy dependence of the event rate as a steplike feature that appears at higher energies for more massive WIMPs (this effect has been studied further in the context of annual modulation in Ref. [48]). The deviation from the smooth halo event rate is sharper and larger in size for lighter WIMPs.

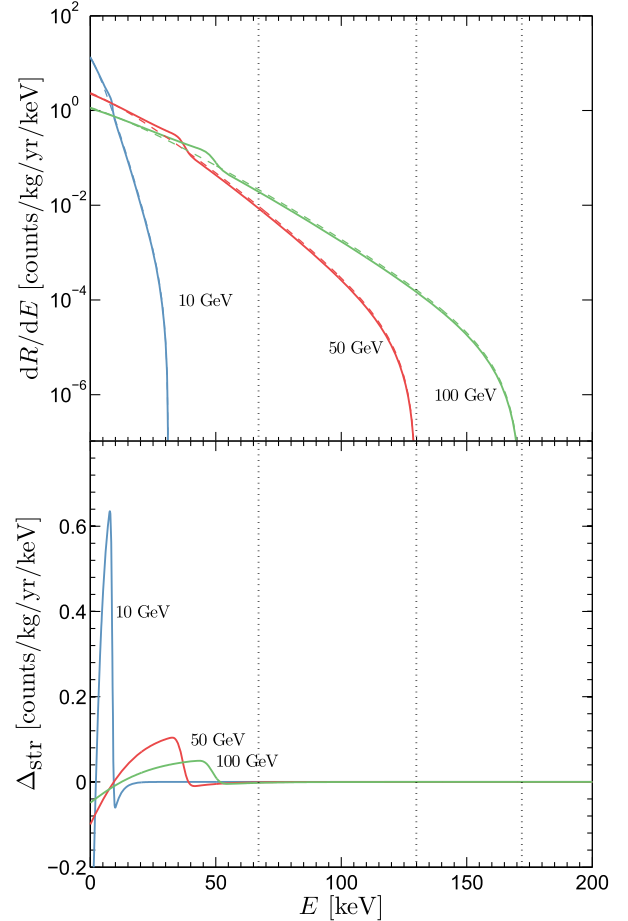


FIG. 2 (color online). Top panel: Nondirectional differential event rate for the halo (dashed lines) and halo + stream models (solid lines) as a function of recoil energy and  $m_\chi = 20$  (blue curve), 50 (red curve) and 100 (green curve) GeV. The dotted vertical lines indicate the cutoffs due to the escape speed. Bottom panel: difference between the event rate in the halo and halo + stream models,  $\Delta_{\text{str}}$  [defined in Eq. (17)], as a function of energy for the same WIMP masses. The input parameters are the same as in Fig. 1.

### C. Scattering simulation

To study the detectability of a stream we simulate mock experiments by Monte Carlo generating nuclear recoils. To do this we build the velocity flux distribution, sample velocities from this distribution and scatter the resulting velocities off the chosen nucleus. We simulate a MIMAC-like experiment with target  $\text{CF}_4$ , in which the scattering is dominated by  $^{19}\text{F}$ .

The energy of a recoil is given by

$$E = \frac{2\mu^2 v^2 \cos^2 \theta}{m_N}, \quad (18)$$

where  $\theta$  is the scattering angle between the initial WIMP direction and the direction of the recoiling nucleus. The scattering angle is related to the center of mass scattering angle,  $\theta_{\text{com}}$ , by

$$\theta = \frac{\pi}{2} - \frac{\theta_{\text{com}}}{2}. \quad (19)$$

The scattering process is isotropic in the center of mass (com) frame so the com angle is taken to be isotropically distributed, i.e.  $\theta_{\text{com}} = \cos^{-1}(2u - 1)$  where  $u \in [0, 1)$  is a uniformly distributed random variate. The recoil vector is generated by deflecting the initial WIMP direction by the elevation angle  $\theta$  and then rotating the deflected vector by a uniformly random angle  $\phi \in [0, 2\pi)$  around the initial WIMP direction. The correction to the event rate due to the nuclear form factor is taken into account by calculating  $F^2(E)$  for each recoil and then discarding each recoil with a probability  $1 - F^2(E)$ .

Finally we must account for the exposure and background of the detector. We split the total number of events seen into WIMP and background events,  $N_{\text{tot}} = N_{\text{wimp}} + N_{\text{bg}}$ . The  $N_{\text{bg}}$  events are generated from an isotropic distribution with a flat energy spectrum. Reference [50] showed that the background model assumed has little effect in the initial detection phase, provided the detector has good sense recognition. We show in the Appendix that this is also the case for the reconstruction of the stream parameters using the profile likelihood ratio test statistic outlined in Sec. IV. To specify the number of events we vary the exposure,  $\mathcal{E}$ , measured in kg yr. This is the most physical way of adjusting the number of background events as it is independent of the WIMP mass or halo parameters. The total number of events is related to the exposure time and the total rate by

$$N_{\text{tot}} = \mathcal{E}R_{\text{tot}} = \mathcal{E}(R + R_{\text{bg}}), \quad (20)$$

where  $R_{\text{bg}}$  is the background event rate and  $R$  is the total WIMP rate

$$R = \int_{E_{\text{th}}}^{E_{\text{max}}} \int_{\Omega_q} \frac{d^2R}{dE d\Omega_q} d\Omega_q dE. \quad (21)$$

#### D. Parameter values

As seen in Eq. (1), the local WIMP density  $\rho_0$  and spin-dependent (SD) WIMP-nucleon cross section  $\sigma_p$  do not affect the shape of the recoil spectrum, only its amplitude, i.e. the total number of events seen in a given exposure time. We fix these parameters to  $\rho_0 = 0.3 \text{ GeV cm}^{-3}$  and  $\sigma_p = 10^{-3} \text{ pb}$ . This value of  $\sigma_p$  lies just below current exclusion limits from XENON100 [51], so our results are optimistic; however, they still hold if  $\sigma_p$  is smaller, but for larger exposure times.

As we have displayed in Fig. 2 the WIMP mass affects the energy dependence of the differential event rate. However  $m_\chi$  also appears in the prefactor of Eq. (1); hence the amplitude of the signal will also depend on the WIMP mass. In Fig. 3 we show the dependence of the signal event rate,  $R$ , on the WIMP mass for three different experimental energy

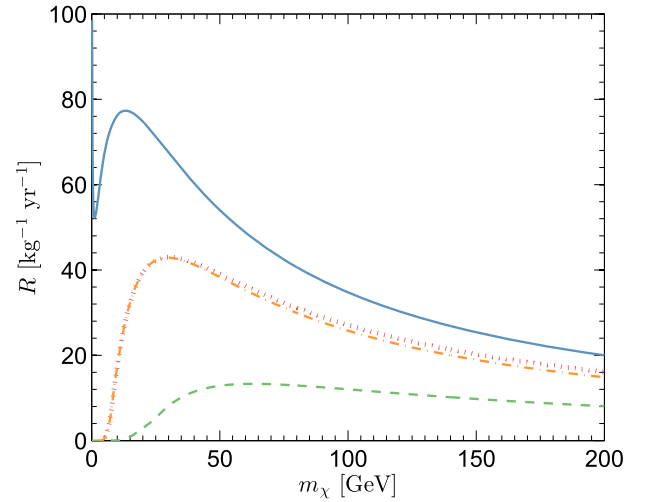


FIG. 3 (color online). Total WIMP scattering rate,  $R$ , for a  $^{19}\text{F}$  target as a function of WIMP mass,  $m_\chi$ , for a range of experimental energy windows  $[E_{\text{th}}, E_{\text{max}}]: [0, \infty]$  keV (blue line),  $[5, 100]$  keV (red dotted line),  $[5, 50]$  keV (orange dash-dotted line), and  $[20, 100]$  keV (green dashed line).

windows,  $[E_{\text{th}}, E_{\text{max}}]$ . We include the dependence on the energy window as this will play a crucial role in the detectability of streams. A change in the input WIMP mass will affect the results of the statistical tests due to both the change in the number of events and also the prominence of the stream in the data. For heavier WIMPs one would expect fewer total events and a stream signal that is more dispersed in energy, making it harder to detect a stream. When studying the effects of other parameters and for the most computationally demanding calculations we fix the input WIMP mass to  $m_\chi = 50 \text{ GeV}$ ; however, where appropriate, we also consider input WIMP masses of 10 and 100 GeV.

In the following sections we will compare the power of several different statistical tests using the benchmark WIMP, halo, experimental and stream parameter values listed in Table I. For the other background halo parameters the values used are taken from the literature (as outlined in Sec. II B) but the precise values are not expected to impact the results significantly. The parameters used to describe the stream are the stream density fraction,  $\xi$ , velocity,  $\mathbf{v}_{\text{str}}$ , and velocity dispersion,  $\sigma_{\text{str}}$ . When describing the properties of the stream we will often parametrize the direction in which it is moving in terms of the angle between  $\mathbf{v}_{\text{str}}$  and  $\mathbf{v}_{\text{lab}}$ ,

$$\Delta\theta = \cos^{-1}(\hat{\mathbf{v}}_{\text{lab}} \cdot \hat{\mathbf{v}}_{\text{str}}). \quad (22)$$

From symmetry considerations the performance of the statistical tests is only sensitive to this quantity. For the stream density and dispersion, unless otherwise stated, we will use  $\xi = 0.1$  and  $\sigma_{\text{str}} = 10 \text{ km s}^{-1}$ .

The experimental energy window,  $[5, 50]$  keV, reflects a plausible forecast for a  $\text{CF}_4$  experiment such as MIMAC [19]. As we will subsequently see, the threshold energy,

TABLE I. Benchmark WIMP, halo, experimental and stream parameters used in Sec. III.

WIMP:	$m_\chi$	50 GeV
	$\sigma_p$ (SD)	$10^{-3}$ pb
Halo:	$\rho_0$	$0.3 \text{ GeV cm}^{-3}$
	$\sigma_v$	$v_{\text{LSR}}/\sqrt{2}$
	$v_{\text{esc}}$	$533 \text{ km s}^{-1}$
	$\mathbf{v}_{\text{lab}}$	$(6.0, 230.6, 6.5) \text{ km s}^{-1}$
Experiment:	$m_N$	18.998 amu (F)
	$E_{\text{th}}$	5, 20 keV
	$E_{\text{max}}$	50, 100 keV
	$R_{\text{bg}}$	$10 \text{ kg}^{-1} \text{ yr}^{-1}$
Sgt. stream:	$\mathbf{v}_{\text{str}}$	$400 \times (0, 0.233, -0.970) \text{ km s}^{-1}$
	$\sigma_{\text{str}}$	$10 \text{ km s}^{-1}$
	$\xi_{\text{str}}$	0.1

$E_{\text{th}}$ , and the maximum energy,  $E_{\text{max}}$ , play an important role in determining the detectability of streams. Therefore we also study the effects of (optimistically) raising  $E_{\text{max}}$  to 100 keV and (less optimistically) raising  $E_{\text{th}}$  to 20 keV. Last we assume that the scattering is only off  $^{19}\text{F}$  with pure proton coupling, so in Eq. (5) we take  $J = 1/2$ ,  $a_p = 1$ ,  $a_n = 0$  and  $\langle S_p \rangle = 0.5$ .

### III. DIRECTIONAL STATISTICS

There are two steps in characterizing a stream. First the stream must be detected, and then its parameters (e.g. density and velocity) can be measured. First we consider nonparametric statistical tests. These tests use the direction information only and have been used in the past to determine the number of events required to distinguish a WIMP signal from isotropic backgrounds [12,52]. The advantage of a nonparametric analysis is that it is not necessary to assume a particular model for the smooth

component of the halo, and the results are valid provided the basic hypotheses that define the statistical tests are satisfied. However, a notable disadvantage of nonparametric analysis is that it is less powerful than parametric analysis (in the sense that, for a given data set, the detection significance will be lower). Therefore we will subsequently perform a parametric analysis using a likelihood in Sec. IV

The question we wish to answer is: what exposure is needed to reject, at a given confidence level, a smooth and isotropic Galactic WIMP distribution in the event that the true distribution contains an additional stream component? Figure 4 illustrates the problem; how does one extract information from a small quantity of spherical data? The left hand panels show the normalized directional signal observed in a perfect detector with infinite exposure, for the halo and stream parameters given in Table I. The top row shows the signal with no stream present, and the bottom with the inclusion of a stream. The right hand panels show the signal expected in a real detector with an exposure of  $\mathcal{E} = 10 \text{ kg yr}$  including events from experimental backgrounds with a total rate of  $R_{\text{bg}} = 10 \text{ kg}^{-1} \text{ yr}^{-1}$  distributed isotropically. Since the directional statistics only use information about the directions of the nuclear recoils, and not their energies, their performance is independent of the background energy spectrum. The data consist of 304 signal events in the halo-only case and 316 in the halo + stream case. The data have been binned on a sphere using a HEALPix [53] equal angular area discretization with  $N_{\text{pix}} = 768$ . While detecting the isotropy of a WIMP signal requires few events, finding deviations from a smooth halo will require a great deal more [52].

We construct test statistics,  $\mathcal{T} > 0$ , that are distributed as  $p_0(\mathcal{T})$  under some null hypothesis that is satisfied in the presence of a distribution without a stream and  $p_1(\mathcal{T})$  when a stream is present in the data. These distributions may be

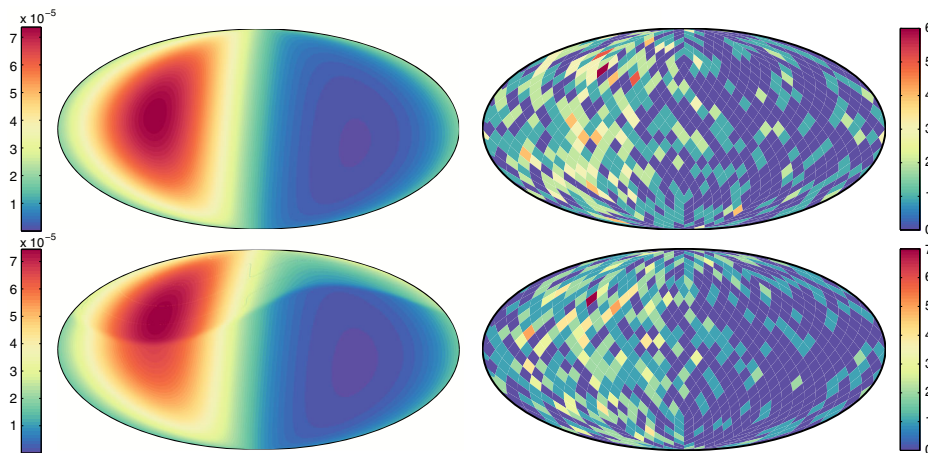


FIG. 4 (color online). Halo (top row) and halo + stream (bottom row) directional signals for an idealized detector with an infinite exposure (left, normalized to unity) and for a realistic detector (right, showing event counts) with finite exposure, angular resolution and experimental backgrounds present. The signals have been generated using the parameters listed in Table I with an energy window of [5, 100] keV.

known analytically for some test statistics in the null case, but if not they can be Monte Carlo generated. For a particular measured value of the test statistic,  $\mathcal{T}_{\text{obs}}$ , the significance is the probability of measuring  $\mathcal{T} < \mathcal{T}_{\text{obs}}$  if the null hypothesis is true

$$S = \int_0^{\mathcal{T}_{\text{obs}}} p_0(\mathcal{T}) d\mathcal{T}. \quad (23)$$

The statistical power is the probability of rejecting the null hypothesis if the null hypothesis is false. In other words, it is the probability of measuring  $\mathcal{T} > \mathcal{T}_{\text{obs}}$  when  $\mathcal{T}$  is distributed according to  $p_1(\mathcal{T})$ ,

$$\mathcal{P} = \int_{\mathcal{T}_{\text{obs}}}^{\infty} p_1(\mathcal{T}) d\mathcal{T}. \quad (24)$$

For detection of a stream we require  $\mathcal{P} = 0.95$  and define  $S_{95}$  as the detection significance achievable in 95% of experiments.

### A. Median direction

The median direction,  $\hat{\mathbf{x}}_{\text{med}}$ , of a set of  $N$  directions,  $\{\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_N\}$ , is found by minimizing the quantity [54],

$$M = \sum_{i=1}^N \cos^{-1}(\hat{\mathbf{x}}_{\text{med}} \cdot \hat{\mathbf{x}}_i). \quad (25)$$

The median recoil direction of the smooth halo distribution is  $-\hat{\mathbf{x}}_{\text{lab}}$ , i.e. the inverse of the direction of Solar motion. To test whether recoils are consistent with a hypothesized median direction,  $\hat{\mathbf{x}}_0$ , we use the  $\chi^2$  test statistic which is calculated as follows [54]. First the recoil vectors  $\hat{\mathbf{x}}_i$  are rotated so that they are measured relative to a north pole at the sample median given by  $(\theta_{\text{med}}, \phi_{\text{med}})$ , and this is done using the rotation

$$\hat{\mathbf{x}}'_i = \mathbf{R}_y(\pi/2 - \theta_{\text{med}}) \mathbf{R}_z(-\phi_{\text{med}}) \hat{\mathbf{x}}_i, \quad (26)$$

where  $\mathbf{R}_y$  and  $\mathbf{R}_z$  are the Cartesian rotation matrices for rotations around the  $y$  and  $z$  axes. After the recoil vectors have been rotated, the azimuthal angles  $\phi'_i$  are then measured in this new coordinate system. Then the matrix,

$$\Sigma = \frac{1}{2} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}, \quad (27)$$

where

$$\sigma_{11} = 1 + \frac{1}{N} \sum_{i=1}^N \cos 2\phi'_i, \quad (28)$$

$$\sigma_{22} = 1 - \frac{1}{N} \sum_{i=1}^N \cos 2\phi'_i, \quad (29)$$

$$\sigma_{12} = \sigma_{21} = \frac{1}{N} \sum_{i=1}^N \sin \phi'_i, \quad (30)$$

is constructed. Next, the recoil vectors are rotated again using Eq. (26) but now so that they are measured relative to a north pole at the hypothesized median direction  $(\theta_0, \phi_0)$ . Finally the test statistic is calculated as

$$\chi^2 = U^T \Sigma^{-1} U, \quad (31)$$

where  $U$  is defined as

$$U = \frac{1}{\sqrt{N}} \begin{pmatrix} \sum \cos \phi_i^0 \\ \sum \sin \phi_i^0 \end{pmatrix}. \quad (32)$$

If the hypothesized median direction is correct and the number of events is  $N > 25$ , then the test statistic is distributed according to a  $\chi^2_2$  distribution. The statistical significance of a particular observed value  $\chi^2_{\text{obs}}$  is then equal to the cumulative distribution function for  $\chi^2_2$  at  $\chi^2_{\text{obs}}$  according to Eq. (23).

### B. Modified Kuiper test

The modified Kuiper test is a test for rotational symmetry around some hypothesized direction  $\hat{\mathbf{x}}_0$ . The test is performed by first rotating all recoil direction vectors  $\hat{\mathbf{x}}_i$  so that their spherical angles  $(\theta_i, \phi_i)$  are measured relative to a north pole with angles  $(\theta_0, \phi_0)$  prior to rotation, using the same procedure as outlined in the previous section in Eq. (26). After the recoil vectors have been rotated, the azimuthal angles  $\phi_i$  are then measured in this new coordinate system in units of  $2\pi$ , reorganized in ascending order and denoted  $X_i$  such that they define a cumulative distribution  $F(X)$ . In the case that the data possess rotational symmetry the cumulative distribution follows  $F(X) = X$ . The modified Kuiper statistic quantifies deviations from this and is defined as

$$\mathcal{V}^* = \mathcal{V} \left( N^{1/2} + 0.155 + \frac{0.24}{N^{1/2}} \right), \quad (33)$$

where  $\mathcal{V} = D^+ + D^-$  is the unmodified Kuiper statistic and

$$D^+ = \max \left( \frac{i}{N} - X_i \right), \quad (34)$$

$$D^- = \max \left( X_i - \frac{i-1}{N} \right), \quad (35)$$

with  $i = 1, \dots, N$ . The modification factor in Eq. (33) allows the distribution of the Kuiper statistic to be independent of the sample size for  $N \geq 8$ . There is no analytic form for this distribution in the null case but there are published critical values [54] or the distribution can be Monte Carlo generated using any set of vectors with

rotational symmetry satisfied, for example, the isotropic experimental background recoils.

### C. Results

Figure 5 shows the distributions of the two test statistics when the null hypothesis is true (generated with isotropic backgrounds) and when performed on data containing recoils from a Sagittarius-like stream with varying density fraction,  $\xi$ . The distributions in each case were generated from  $10^5$  mock experiments with an exposure of 10 kg yr. As  $\xi$  is increased the degree of rotational asymmetry is increased, the median direction becomes more displaced from  $-\hat{v}_{\text{lab}}$  and the observed distributions of the test statistics become further separated from the null distributions.

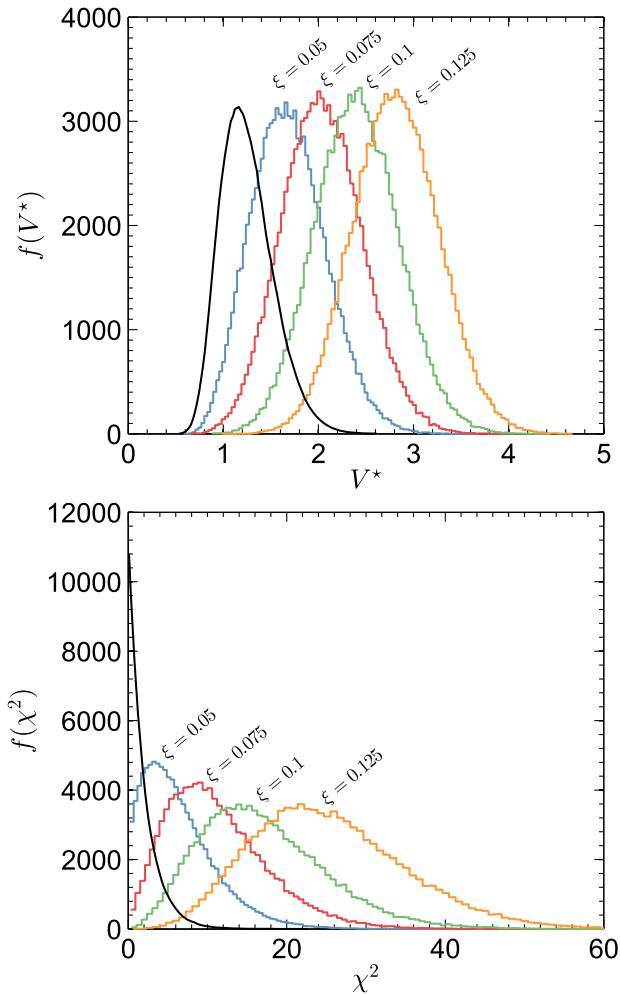


FIG. 5 (color online). The distributions of the Kuiper statistic,  $f(V^*)$  (top panel), and Median direction  $\chi^2$  statistic,  $f(\chi^2)$  (bottom panel), built from  $10^5$  Monte Carlo experiments. The distributions were generated for isotropic background recoils (black curve) and with a Sagittarius-like stream with variable density fraction  $\xi$  present (colored histograms), with increasing  $\xi$  going from left to right. The stream velocity and velocity dispersion, as well as the parameters for the MB distribution of the smooth halo are given in Table I.

### 1. Sagittarius stream

To demonstrate the performance of the tests as a function of parameters other than the stream velocity we will take as a concrete example, a Sagittarius-like tidal stream. We will assume a velocity of  $\mathbf{v}_{\text{str}} = 400 \times (0, 0.233, -0.970)$  [47,48] for the stream. There is sizable uncertainty in both the speed and direction of the velocity (and even whether or not dark matter from this stream passes through the Solar neighborhood), therefore in the subsequent subsections we will vary the stream velocity.

In Fig. 6 we plot  $S_{95}$ , the value of the significance  $S$  given a required power  $\mathcal{P} = 0.95$  (i.e. the value of significance obtainable by 95% of hypothetical experiments) for the Kuiper and median direction  $\chi^2$  tests. We show how this quantity varies with exposure time, stream dispersion, stream density, experimental background rate, WIMP mass and escape speed. We also show the dependence of the results on the energy window of the detector, considering three examples, [5, 100] keV, [5, 50] keV and [20, 100] keV. The benchmark experimental and halo parameter values that are fixed when the other parameters are varied are indicated in the figure by vertical lines.

The results are intuitive, with longer exposure times or for streams which make up a larger fraction of the local WIMP density, the signal is stronger and hence the tests perform better. The number of stream WIMPs scales linearly with exposure time and stream density fraction; hence the significance of the result scales as roughly the square root of those quantities. To detect a Sagittarius-like stream at 90%–95% confidence in 95% of experiments, one would need exposures between 10 and 20 kg yr for stream densities around  $\xi \sim 0.1$ .

Experiments with larger background rates perform predictably poorly compared to those with fewer backgrounds to contaminate the WIMP signal. Note that the quoted value of  $R_{\text{bg}}$  is the rate observed in the energy window [5, 100] keV and the value for the other ranges has been scaled appropriately to account for the smaller sensitivity window. The dispersion of the stream has no effect on the performance of the test, because, for dispersions of at most tens of  $\text{km s}^{-1}$ , the WIMPs scatter into the same angular area independent of the dispersion. The significance achieved decreases weakly as the escape speed is increased. This is because increasing the escape speed does not affect the recoils from stream WIMPs, but slightly increases the number of recoils from the smooth halo.

The tests perform more or less equally well; however, there is a notable difference between the performance of the tests when the threshold energy is 20 keV. The lowest energy recoils are those which scatter through the largest angles. Removing the low energy recoils from the signal by increasing the threshold energy thus removes the stream WIMPs scattering farthest from the peak stream direction. As can be seen in Fig. 6, the performance of the median



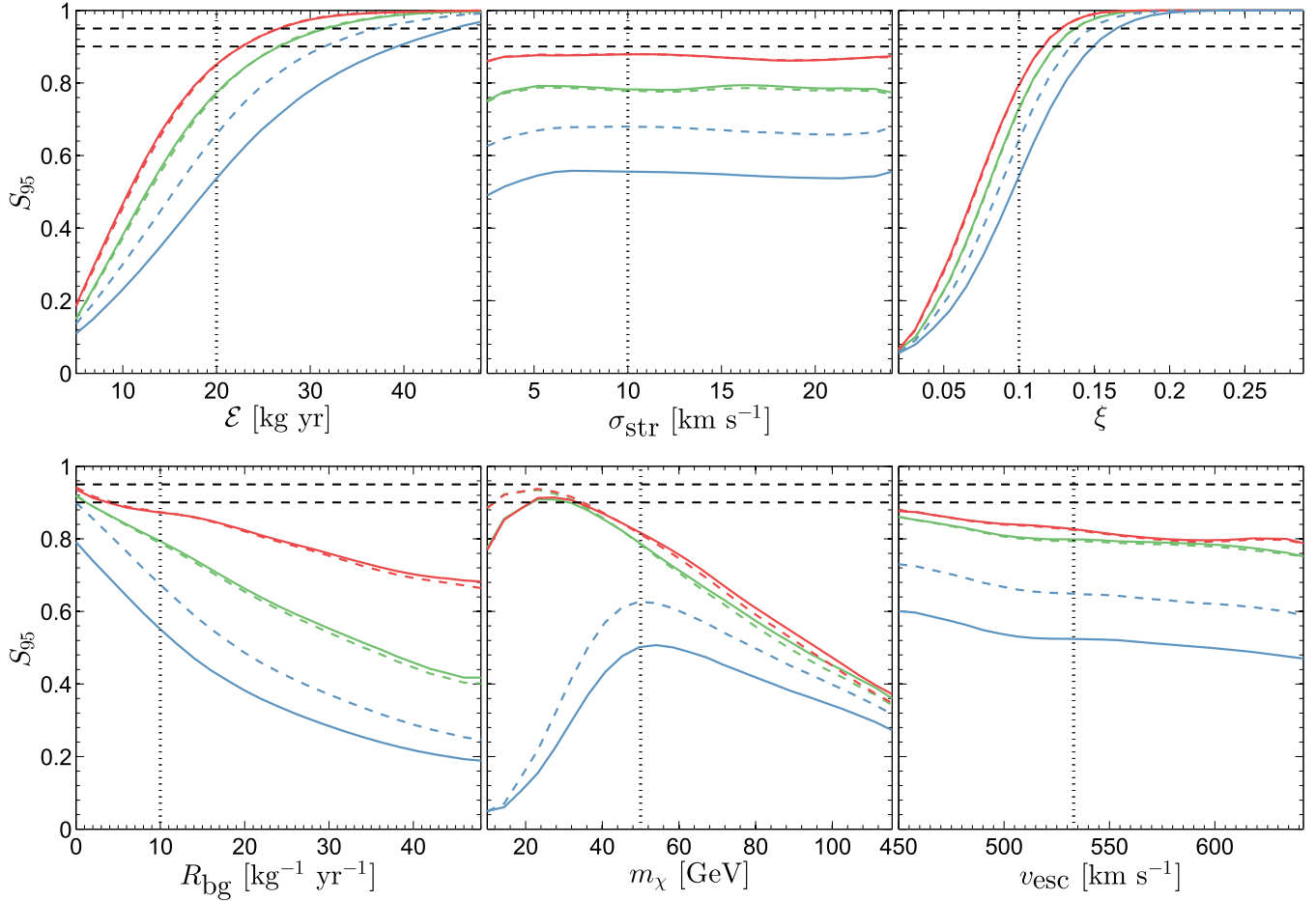


FIG. 6 (color online). Significance obtainable by 95% of experiments,  $S_{95}$ , as a function of exposure time,  $\mathcal{E}$ , stream dispersion,  $\sigma_{\text{str}}$ , stream density fraction,  $\xi$ , background rate,  $R_{\text{bg}}$ , WIMP mass,  $m_\chi$ , and escape speed,  $v_{\text{esc}}$ , for the median direction (solid lines) and Kuiper (dashed lines) tests for the Sagittarius-like stream. The results are shown for energy windows of [5, 100] keV (green lines), [5, 50] keV (red lines) and [20, 100] keV (blue lines). The horizontal dashed lines indicate the desired 0.9–0.95 significance levels and the vertical dotted lines in each plot indicate the parameter values used in neighboring plots.

direction test is weakened compared to the Kuiper test under the removal of these low energy recoils.

The performance of the tests also depends on the mass of the WIMP, due mainly to the variation in the total WIMP scattering rate. As shown in Fig. 3, typically higher mass WIMPs yield fewer events (assuming fixed exposure and local density  $\rho_0$ ), and hence the significance achieved by the tests decreases with increasing WIMP mass. However, for light WIMPs the majority of the recoils from stream WIMPs are below threshold and the significance achieved decreases. This occurs for  $m_\chi \lesssim 50(20)$  GeV for a 20 (5) keV energy threshold. As we will see in Fig. 11 in Sec. IV, the number of recoils from stream WIMPs also depends on the stream velocity.

## 2. Stream velocity

It is possible that a DM stream without a stellar component could pass through the Solar neighborhood. Given this, and also the uncertainties in the velocity of the Sagittarius

stream, we now examine how the test statistics perform when the stream velocity is varied. As mentioned previously, the stream velocity can be described by two parameters, the speed of the stream,  $v_{\text{str}}$ , and the angle between the lab and stream velocities,  $\Delta\theta$ , defined in Eq. (22).

Figure 7 shows how the significance varies with the stream direction and speed, and also the energy window of the detector. The tests do not perform equally for all stream directions. In particular the tests return a low significance when the stream is antialigned with the Solar velocity, as in this case the null hypotheses of rotational symmetry and inverse-Solar median direction are in fact correct. The number of events from stream WIMPs plays an important role in determining the significance which can be achieved; for streams with small  $\Delta\theta$  this number is very low, and for low streams speeds or high threshold energies it can even be zero. The symmetry in the plots is due to this dependence on both the sample size and the positioning of the stream recoils with respect to the smooth halo recoils. For the fastest streams, increasing the threshold energy

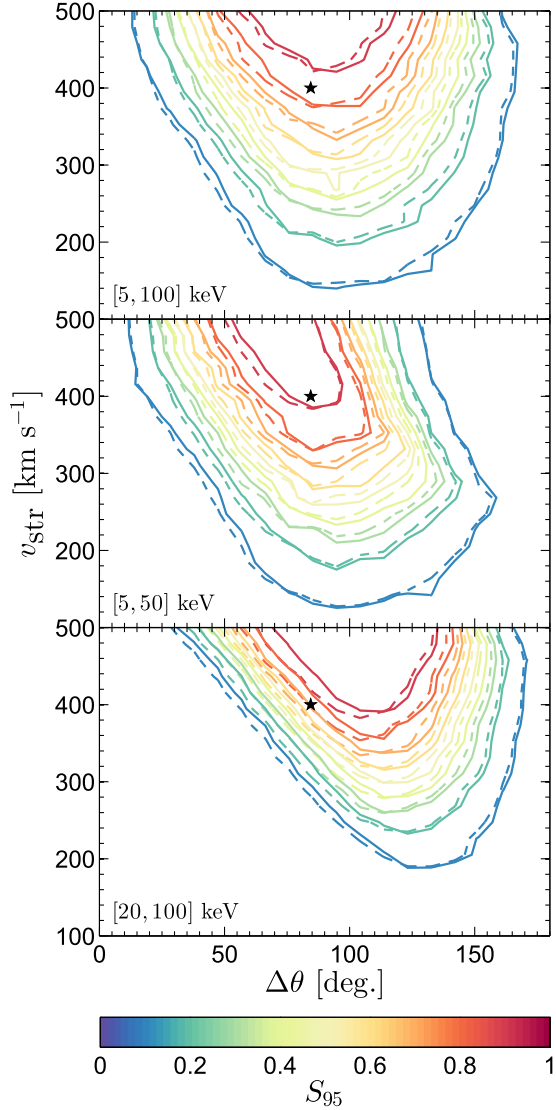


FIG. 7 (color online). Significance obtainable by 95% of experiments with a 10 kg yr exposure as a function of stream speed,  $v_{\text{str}}$ , and the angle between the stream velocity and the direction of Solar motion,  $\Delta\theta$ , for a  $m_\chi = 50$  GeV WIMP. The results for the Kuiper test are shown by the dashed lines and the median direction  $\chi^2$  test by the solid lines. The three panels are for energy windows [5, 100] keV, [5, 50] keV and [20, 100] keV, from top to bottom. The star denotes the speed and direction of the Sagittarius stream.

results in a higher significance, and this is essentially a background rejection effect; it preferentially removes low energy smooth halo recoils, and the stream is more prominent, even with fewer overall events. The dropoff at large  $\Delta\theta$  for the [5, 50] keV window is due to the stream being boosted so that a portion of the recoils falls above the energy window. This is the reverse effect to that at low values of  $\Delta\theta$  in the [20, 100] keV window; in this case the stream recoils lie below the threshold energy.

The Kuiper and median direction tests studied in this section only use the direction of the recoils and not their

energies. As has been shown in Ref. [11], use of the full energy and direction data allows better constraints to be placed on the WIMP speed distribution. It may be possible to improve the tests we have studied by incorporating the energy dependence of the signal by performing the test successively on energy ordered recoils or by binning the recoils in energy. In the case where there is no stream present the hypotheses of rotational symmetry and inverse-Solar median direction are satisfied for all energies. However, with a stream in the signal there will be a range of energies where the hypothesis is false. The degree to which the test is failed will initially increase with energy, and then above a certain value, determined by the stream speed, the test would return a value closer to the null case. Accounting for this effect in the test statistic would decrease the overlap between the null and alternative distributions and hence increase the significance of a particular result. However, it is likely to be a small effect; in Fig. 2 one can see that the energy dependent effect of a stream is quite small and highly dependent on mass. Furthermore splitting the recoils in energy would result in a smaller number of events, and hence a loss of information, for each individual evaluation of the test statistic. Therefore it is likely that such a modified energy dependent nonparametric test would perform no better than the standard direction-only nonparametric tests we have studied, given the low numbers of recoils expected.

## IV. LIKELIHOOD ANALYSIS

### A. Likelihood function

We now turn our attention toward statistical tests capable of placing constraints on the properties of substructure present in the local dark matter velocity distribution as well as distinguishing between the halo and halo + stream models. For placing constraints on the parameters of the model we use Bayesian inference, whereas the test that is most appropriate for model comparison is a profile likelihood ratio test, as it uses the assumption that the null hypothesis is obtained by applying a constraint to the alternative hypothesis. The likelihood ratio test has the added advantage of being able to conveniently account for astrophysical uncertainties by treating them as nuisance parameters [50].

The model we describe has 11 free parameters,  $\theta = \{m_\chi, \sigma_p, \rho_0, \sigma_v, v_{\text{esc}}, \sigma_{\text{str}}, v_{\text{str},1}, v_{\text{str},2}, v_{\text{str},3}, \xi, R_{\text{bg}}\}$ , where we split the velocity of the stream into its three Galactic coordinate components. For simplicity we will again assume the smooth component of the halo has a MB distribution, but in principle the approach could be used with other velocity distributions, e.g. the form of Mao *et al.* [23].

We define the likelihood function as the product of the probabilities for obtaining recoils with energies and directions  $(E_i, \hat{\mathbf{q}}_i)$  ( $i = 1, \dots, N_o$ ) multiplied by a Poisson factor

accounting for the probability of obtaining  $N_o$  observed events given an expected number  $N_e(\boldsymbol{\theta})$ ,

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{N_e(\boldsymbol{\theta})^{N_o}}{N_o!} e^{-N_e(\boldsymbol{\theta})} \prod_{i=1}^{N_o} [\lambda P_{\text{wimp}}(E_i, \hat{\mathbf{q}}_i; \boldsymbol{\theta}) + (1-\lambda) P_{\text{bg}}]. \quad (36)$$

The expected number of events is a function of the parameters  $\boldsymbol{\theta}$  and is equal to the sum of the expected number of WIMP and background events,  $N_e^{\text{wimp}}$  and  $N_e^{\text{bg}}$ , respectively,

$$N_e(\boldsymbol{\theta}) = N_e^{\text{wimp}} + N_e^{\text{bg}}, \quad (37)$$

$$= \mathcal{E} \left[ \int_{E_{\text{th}}}^{E_{\text{max}}} \int_{\Omega_q} \frac{d^2 R}{dE d\Omega_q} \Big|_{\boldsymbol{\theta}} d\Omega_q dE + R_{\text{bg}} \right]. \quad (38)$$

The probabilities  $P_{\text{wimp}}$  and  $P_{\text{bg}}$  are the probabilities for an event to occur at  $(E_i, \hat{\mathbf{q}}_i)$  in the signal and background (no WIMP) cases, respectively, i.e.

$$P_{\text{wimp}}(E_i, \hat{\mathbf{q}}_i; \boldsymbol{\theta}) = \frac{1}{R} \frac{d^2 R}{dE d\Omega_q} \Big|_{E_i, \hat{\mathbf{q}}_i; \boldsymbol{\theta}}, \quad (39)$$

and, assuming a flat background energy spectrum,

$$P_{\text{bg}} = \frac{1}{4\pi(E_{\text{max}} - E_{\text{th}})}, \quad (40)$$

and  $\lambda$  is the signal fraction,

$$\lambda = \frac{R}{R + R_{\text{bg}}}. \quad (41)$$

It will be difficult to constrain the escape speed,  $v_{\text{esc}}$ , as it only affects the  $^{19}\text{F}$  recoil spectrum at energies beyond the maximum energies we are considering. This can be overcome, as in Ref. [50], by treating  $v_{\text{esc}}$  as a nuisance parameter and accounting for its uncertainty by hand by including an additional multiplicative term to the likelihood in the form of a Gaussian with mean and standard deviation of  $v_{\text{esc}} = 533 \pm 54 \text{ km s}^{-1}$  [45,55].

The WIMP density,  $\rho_0$ , and cross section,  $\sigma_p$ , are also difficult to constrain as they only appear in the amplitude of the recoil spectrum and hence only affect the number of events seen for a given exposure. Even for the quite large stream densities that we are considering here, the difference between the number of events in the stream case and in the null case is small. Moreover, the two parameters are degenerate with one another meaning there is no single set of values for  $\rho_0$  and  $\sigma_p$  that maximize the likelihood function. Therefore we also employ a Gaussian contribution to the likelihood function for  $\rho_0$  with mean and standard deviation  $\rho_0 = 0.3 \pm 0.1 \text{ GeV cm}^{-3}$  [56].

## B. Parameter reconstruction

While the questions that we are trying to address are frequentist in nature, there is much to be gained from parameter estimation by Bayesian inference. We can test this approach by reconstructing the input parameters used to generate a set of mock data, again initially studying the example case of a Sagittarius-like stream and a  $m_\chi = 50 \text{ GeV}$  WIMP, using the same benchmark parameters and experimental configuration as in Sec. III.

We sample the likelihood using the nested sampling software MULTINEST [57,58] using 5000 live points, an evidence tolerance factor of 0.05 and a sampling efficiency of 0.3. The posterior distribution is the probability distribution of the parameters  $\boldsymbol{\theta}$  given a data set  $d$  and is defined by Bayes' theorem,

$$p(\boldsymbol{\theta}|d) = \frac{\pi(\boldsymbol{\theta})\mathcal{L}(d|\boldsymbol{\theta})}{E(d)}, \quad (42)$$

where  $\mathcal{L}$  is the likelihood function, i.e. the probability of the data given parameters  $\boldsymbol{\theta}$ ,  $\pi(\boldsymbol{\theta})$  is the prior distribution for the parameters and  $E(d)$  is the evidence (practically, the normalization factor for the posterior distribution). For each parameter we use a flat prior with ranges indicated by the axes of Fig. 8. We define minimum credible regions which give contours of the parameter space encompassing 95% or 68% of the posterior distribution and indicate how well the parameters are reconstructed from the data. Figure 8 shows the marginalized one- and two-dimensional posterior probability distributions of the parameters. The marginalized probability distributions are calculated by integrating the distribution over the remaining parameters, e.g. for parameters  $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_n\}$ ,

$$p(\theta_1, \theta_2|d) = \int p(\boldsymbol{\theta}|d) \prod_{i=3}^n d\theta_i. \quad (43)$$

While the stream parameters are recovered well, the largest uncertainties are present in the halo and WIMP parameters, in particular  $m_\chi$ . The effect of the Gaussian parametrization of  $v_{\text{esc}}$  and  $\rho_0$  is apparent and equivalent to using a Gaussian prior. There is still some correlation in the  $\rho_0$ - $\sigma_p$  plane but the degeneracy is broken. The halo parameters are reconstructed less accurately when the bare likelihood function is used; however, the Gaussian parametrization is representative of the astrophysical uncertainties.

Our conclusions are not sensitive to the form assumed for the background energy spectrum. In the Appendix we show that the maximum likelihood estimators for the stream parameters for a Sagittarius-like stream with an exponential background are almost identical to those with a flat background.

We have shown that good constraints can be placed on the parameters describing a stream, if the correct model is used in constructing the likelihood function for the data.

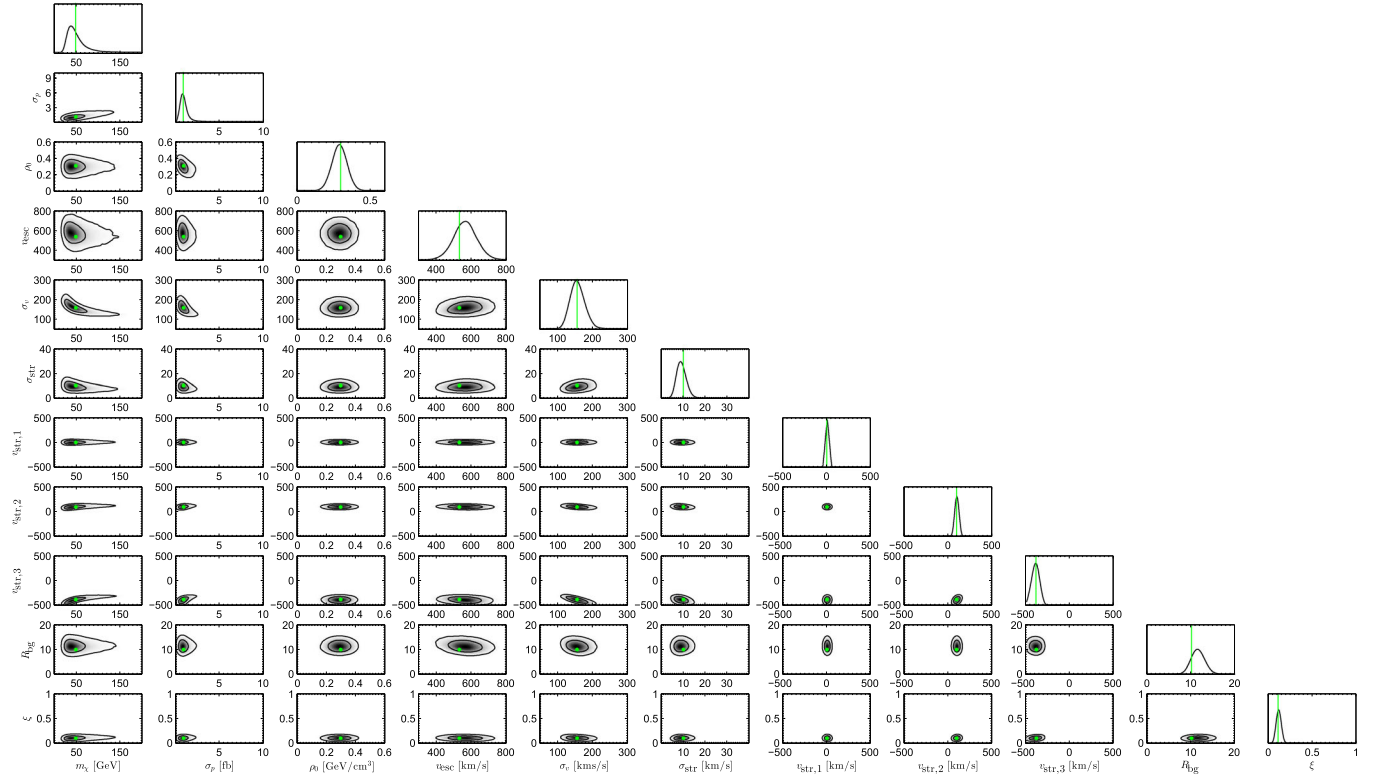


FIG. 8 (color online). The one- and two-dimensional marginalized posterior probability distributions for the 11 parameters of the MB + stream model. The contours indicate the 68% and 95% credible regions. The green dots/lines indicate the location of the input parameters.

Importantly, we note that the exposure times needed to make these constraints  $\sim 10$  kg yr are significantly shorter than are needed when the nonparametric directional tests are used. Our results agree with a similar work by Lee and Peter [11]. They likewise found that the constraints improved with the use of the full energy and direction information; however, they used a different parametrization, subsuming the degenerate local density and cross section into a single parameter, fixing the total number of WIMP and background events and normalizing out the amplitude of the signal. Here, to achieve independence under changing halo, WIMP and experimental parameters, we quote exposure time and not total number of events and must parametrize the likelihood function with local density and cross section and also use the background rate  $R_{bg}$  as an input parameter.

### C. Likelihood ratio

We now turn our attention toward the profile likelihood ratio test, a frequentist test that we will use to assess the detectability of streams with different velocities in a parametric way. Model comparison is performed by considering two hypotheses, a null hypothesis,  $H_0$ , that the WIMP distribution is described by a MB distribution alone with no stream component and an alternative hypothesis,  $H_\xi$ , where the density of the stream is nonzero. The profile likelihood ratio test utilizes the fact that the model best describing the

null hypothesis is contained within the model best describing the alternative hypothesis. We use this test as it is computationally easier than performing Bayesian inference over a range of input velocities. The likelihood ratio between the null and alternative hypotheses is

$$\Lambda = \frac{\mathcal{L}(\hat{\boldsymbol{\theta}}, \xi = 0)}{\mathcal{L}(\hat{\boldsymbol{\theta}})}, \quad (44)$$

where  $\hat{\boldsymbol{\theta}}$  are the maximum likelihood estimators in the alternative model, and  $\hat{\boldsymbol{\theta}}$  are the maximum likelihood estimators evaluated when the stream density fraction is fixed at zero,  $\xi = 0$ . The likelihood ratio test statistic is then defined as

$$\mathcal{D} = \begin{cases} -2 \ln \Lambda & 0 \leq \hat{\xi} \leq 1, \\ 0 & \hat{\xi} < 0, \quad \hat{\xi} > 1. \end{cases} \quad (45)$$

Next we require an expression for the statistical significance of a particular measured test statistic value. This requires knowledge of how the profile likelihood ratio test statistic  $\mathcal{D}$  is distributed in the case that the null hypothesis is true, i.e. if the observed value is  $D_{obs}$ ,

$$S = \int_0^{D_{obs}} f(\mathcal{D}|H_0) d\mathcal{D}. \quad (46)$$

It is known, however, from Wilk's theorem [59] that the distribution of the profile likelihood ratio test statistic in the

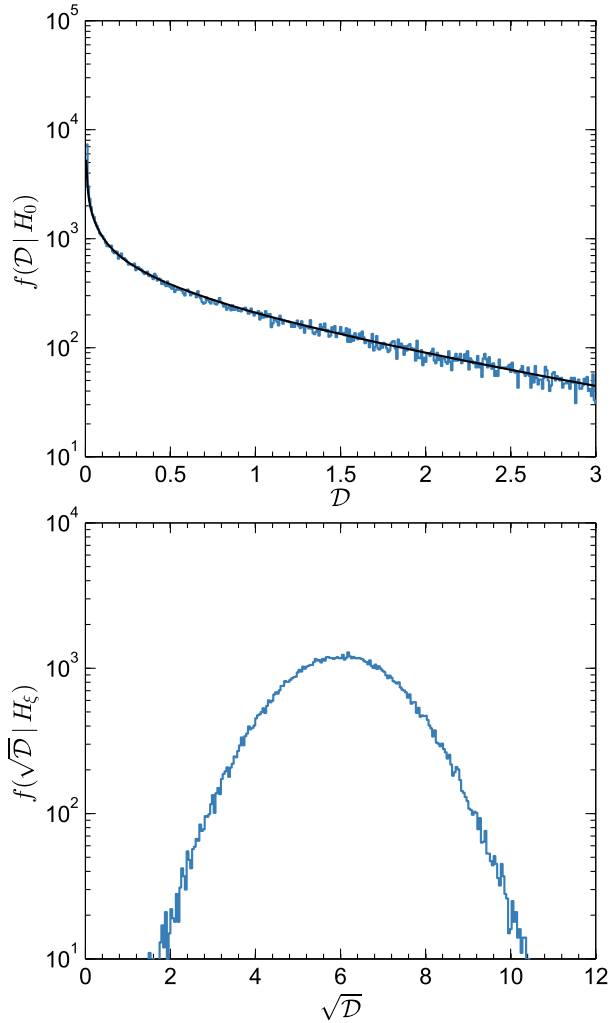


FIG. 9 (color online). Distributions of the profile likelihood ratio test statistic,  $f(\mathcal{D})$ , under the null hypothesis of zero stream density,  $\xi = 0$  (top) and the distribution of the significance in units of  $\sigma$ ,  $f(\sqrt{\mathcal{D}})$ , under the alternative hypothesis with  $\xi = 0.1$  (bottom) generated from  $10^5$  Monte Carlo experiments. The black curve in the top panel shows a fitted  $\chi_1^2$  distribution, demonstrating that Wilk's theorem holds. The distributions were generated from the input parameters in Table I with an exposure of 10 kg yr and energy window of [5, 100] keV.

case that the null hypothesis is true asymptotes toward a  $\chi_1^2$  distribution. Hence the discovery significance is defined as  $S = \text{erf}(\sqrt{\mathcal{D}_{\text{obs}}}/2)$ . However, as we will be quoting quite high values of significance, it is simpler to write them in units of standard deviation,  $\sigma$ , i.e.  $S = \sqrt{\mathcal{D}_{\text{obs}}}$ . So a value of  $\mathcal{D}_{\text{obs}} = 1$  corresponds to a  $1\sigma$  result or a significance of 68%. The significance obtainable by 95% of mock experiments,  $S_{95}$ , is then found by solving the equation

$$\int_0^{S_{95}} f(\sqrt{\mathcal{D}}) d\sqrt{\mathcal{D}} = 0.95. \quad (47)$$

In Fig. 9 we show the distributions of  $\mathcal{D}$  from  $10^5$  mock experiments under the null and alternative hypotheses, for

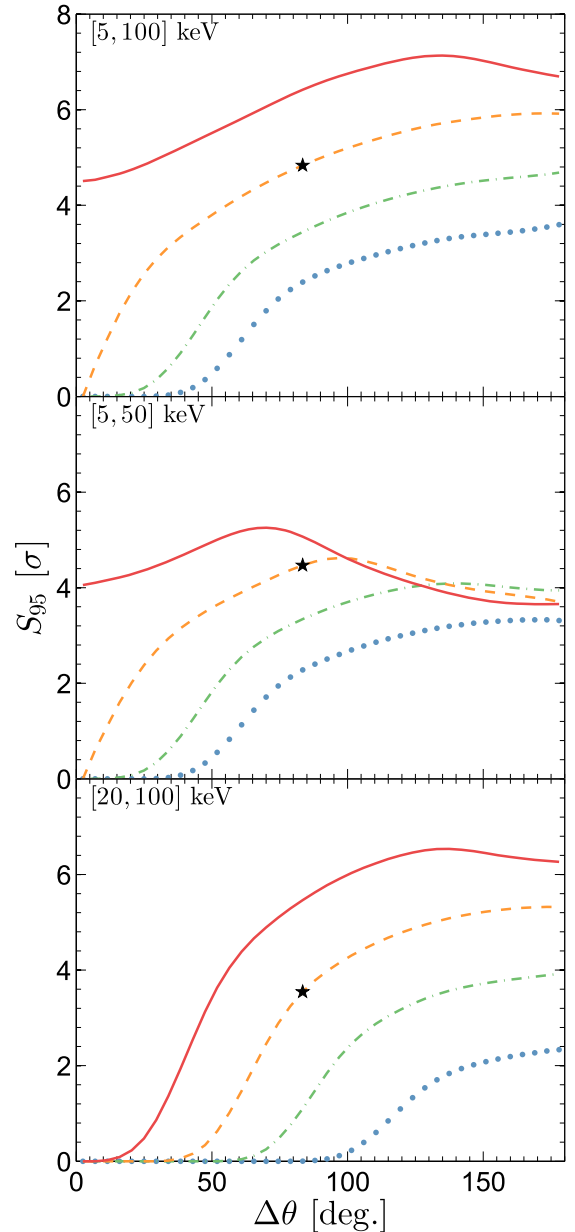


FIG. 10 (color online). The significance obtainable by 95% of experiments,  $S_{95}$ , in units of  $\sigma$ , as a function of the angle between the Solar and stream velocities,  $\Delta\theta$ , for a 50 GeV WIMP. The curves correspond, from top to bottom in each panel, to  $v_{\text{str}} = 500 \text{ km s}^{-1}$  (red solid line),  $400 \text{ km s}^{-1}$  (orange dashed line),  $300 \text{ km s}^{-1}$  (green dot-dashed line), and  $200 \text{ km s}^{-1}$  (blue dotted line). The three panels are for energy windows [5, 100] keV, [5, 50] keV and [20, 100] keV, from top to bottom. The star denotes the result for the Sagittarius stream.

the Sagittarius-like stream given in Table I. For a nonzero stream density fraction the test statistic has a Gaussian distribution which moves farther from  $\mathcal{D} = 0$  as the stream density is increased. Again, our conclusions here are not sensitive to the form assumed for the background energy spectrum. In Appendix A we show that the distribution of the likelihood ratio test statistic for a Sagittarius-like stream

with an exponential background is almost identical to that with a flat background.

As in Sec. III we now study the performance of the test over a range of stream velocities where we will again use the stream speed,  $v_{\text{str}}$ , and the angle between the Solar and stream velocities,  $\Delta\theta$ , to describe the stream velocity. In Fig. 10 we plot the significance, in units of  $\sigma$ , obtainable by 95% of hypothetical experiments,  $S_{95}$ , using the profile likelihood ratio test as a function of stream speed,  $v_{\text{str}}$ , and direction given by  $\Delta\theta$ , for a 50 GeV WIMP and the three energy windows considered previously. Figure 10, and also Fig. 11, was generated using exposure times of 5 kg yr, and the parameters not plotted were all taken to have the benchmark values given in Table I. The test, by virtue of being parametric, performs much better than the non-parametric tests. For the example of the Sagittarius stream, indicated in the figures by a star, the tests for the three energy windows detect the stream at  $4\sigma$ – $5\sigma$  in 95% of experiments. For the same exposure, the nonparametric tests could only reach a value of  $S_{95}$  between 0.1 and 0.2.

The enhancement in performance is also due to the use of the full energy and direction data, whereas before only the direction information was used. Furthermore the tests achieve high significance over a wide range of stream velocities, with the limiting factor being the number of WIMPs coming from the stream,  $N_{\text{wimp}}^{\text{str}}$ , as can be seen by comparing the two figures. For low values of  $\Delta\theta$ , where the number of stream WIMPs drops to zero, the significance can be seen to do likewise. There is similarly a dependence on the energy window of the detector which causes a reduction in the number of stream WIMPs when the stream becomes boosted past the maximum of the energy window. This can be seen clearly in the [5, 50] keV case. For a stream speed of  $500 \text{ km s}^{-1}$ , the significance begins to decrease for  $\Delta\theta > 70^\circ$  and drops by  $1.6\sigma$  up to  $\Delta\theta = 180^\circ$ . However, the significance for faster stream speeds is enhanced over what might be expected simply from the dependence on  $N_{\text{wimp}}^{\text{str}}$ . This is due to faster streams becoming more prominent because of the exponential dropoff with energy of the event rate for the smooth halo.

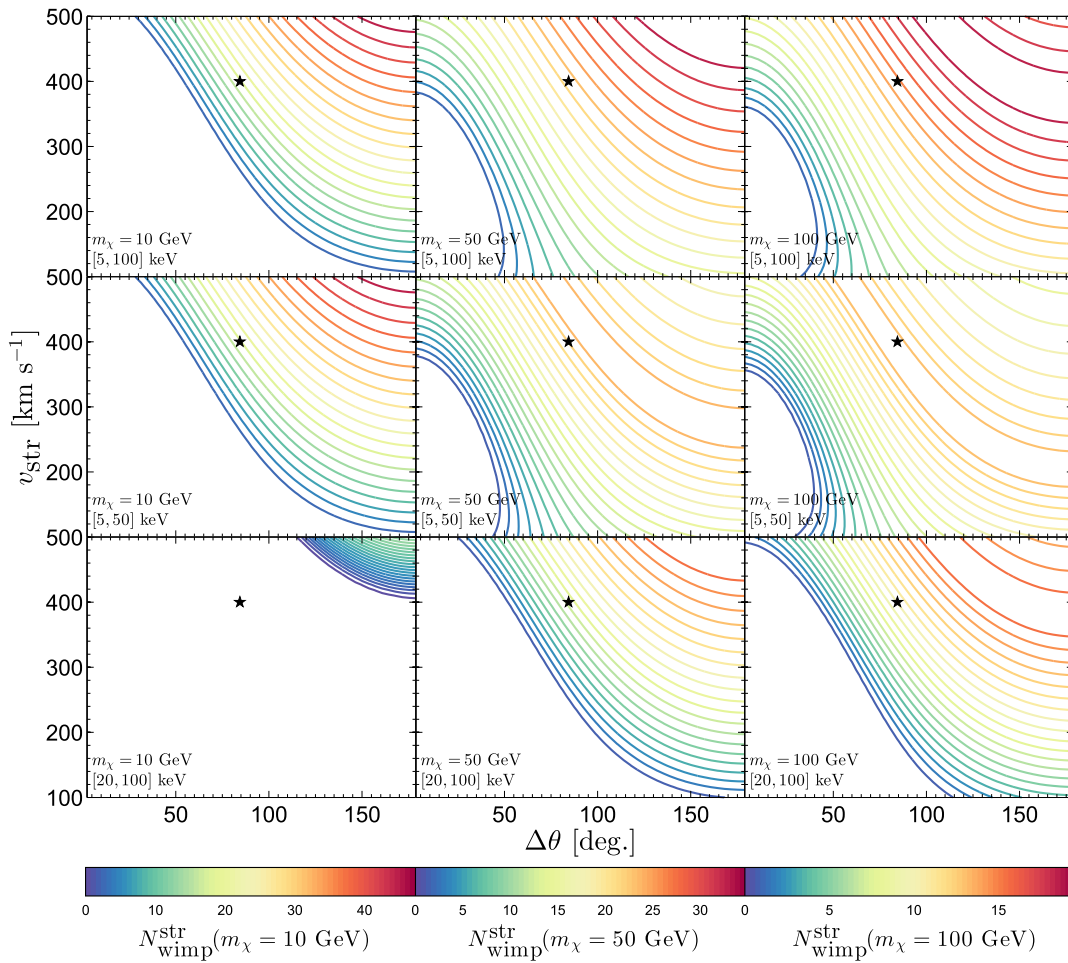


FIG. 11 (color online). The number of recoils from stream WIMPs,  $N_{\text{wimp}}^{\text{str}}$ , as a function of stream speed,  $v_{\text{str}}$ , and direction,  $\Delta\theta$ , for  $m_\chi = 10, 50$  and  $100$  GeV (from left to right) and energy windows  $[5, 100]$  keV,  $[5, 50]$  keV and  $[20, 100]$  keV (from top to bottom). The star denotes the speed and direction of the Sagittarius stream.

The number of events from WIMPs in the stream,  $N_{\text{wimp}}^{\text{str}}$ , plays a key role in determining the detectability of a stream. Therefore in Fig. 11 we show  $N_{\text{wimp}}^{\text{str}}$  as a function of the stream speed and direction for the three energy windows and  $m_\chi = 10, 50$  and  $100$  GeV. As discussed earlier, provided the WIMP mass is large enough that the stream recoils are above the energy threshold, then increasing the WIMP mass decreases the number of stream events. The energies of the stream recoils depend on the stream velocity as well as the WIMP mass. Increasing the threshold energy to 20 keV or reducing the maximum energy to 50 keV reduces the size of the region of stream speed-angle parameter space within which the number of stream events is large enough for the stream to be detected. For a 10 GeV WIMP and a 20 keV threshold a stream will be detectable only if it has a high speed and a large angle relative to the lab velocity.

## V. SUMMARY

Using both nonparametric directional tests and a profile likelihood test, we have shown that there are reasonable prospects for the detection of a moderately high density tidal stream by a future directional detector. We have looked at the concrete example of a Sagittarius-like stream and also explored the dependence on the parameters of the stream, namely its speed, direction, dispersion and density.

Using nonparametric directional statistics the detection of a Sagittarius-like stream would need a total of around 900 events, but using Bayesian parameter estimation good constraints can be placed on the stream parameters with around 300 events, independent of astrophysical uncertainties. For WIMP-nucleon spin-dependent cross sections just below the current exclusion limits [51] this corresponds to around 10 kg yr exposure, which is within the scope of future directional detectors [15].

We have shown that the detectability of a stream is highly dependent on its speed and position in relation to the lab velocity. Faster streams that are oriented at  $90^\circ$  with respect to the lab velocity are most easily detected when using directional statistics as the signal in this case deviates furthest from the smooth halo signal. Using a profile likelihood ratio test, there is weaker dependence on the position of the stream as the likelihood does not favor a particular directionality of the signal, but there is naturally a dependence on the number of WIMPs originating from the stream, which indirectly leads to a dependence on stream direction. Faster streams can be detected with higher significance, as they cause more events and are also more prominent in the energy spectrum, due to the exponential dropoff with energy of events from the smooth halo. A similar conclusion is likely for parameter estimation as weaker constraints would be made on parameters of the stream with fewer WIMPs from the stream. The estimation of halo and WIMP parameters would also be dependent on the stream velocity; if the stream is boosted out of the energy window of the detector, there would be fewer events overall which would weaken the relative constraints.

The advantage of using nonparametric tests is that one need not assume a model to describe the data, simply that the data satisfy either a null or an alternative hypothesis. However, nonparametric tests will always return a less significant result than parametric tests. The likelihood analyses also make use of both the energy and the direction information of the recoils, whereas the nonparametric tests are direction only. The Kuiper and median direction tests may be improved with the use of energy information; however, this would result in a greater dependence on the WIMP mass and would be expected to perform poorly with low numbers of recoils.

We have also investigated a number of experimental considerations. First, as might be expected, a larger experimental background rate results in poorer performance of the statistical tests, but the dependence on the background rate flattens at very large values. A crucial factor is the energy window of the detector, where for the detection of streams the window must cover the range of recoil energies from stream WIMPs. This is dependent on both the speed of the stream and its direction in relation to the lab velocity as we have demonstrated.

The statistical methods we have studied here may all be applied in a practical setting on real data; however, there remain some experimental issues that may affect the detectability of streams. The mock detector we have considered here has perfect angular resolution. In practice a finite angular resolution of around  $10^\circ$  is thought to be achievable for future detectors [19]. This will have some impact on the power of an experiment to detect fine substructure; however, the results here are likely to remain valid if finite angular resolution is taken into account either through a binned analysis or by adding in an uncertainty in the recoil data. Another experimental concern is that of sense recognition (i.e. distinguishing between  $\hat{\mathbf{q}}$  and  $-\hat{\mathbf{q}}$ ) which remains an issue for some current directional detectors. A lack of sense recognition has been shown to increase the number of events needed to reject an isotropic non-WIMP background in the initial detection phase [60] and is likely to be problematic for the detection of substructure; new tools may be needed to reconstruct the velocity distribution if it is still an issue in the future.

To conclude, we have shown that if a moderately high density WIMP stream is present in the Solar neighborhood, then there are good prospects for its detection using directional dark matter detectors.

## ACKNOWLEDGMENTS

The authors thank Mattia Fornasa and Tasos Avgoustidis for useful discussions. C. A. J. O. and A. M. G. are both supported by the STFC.

## APPENDIX: BACKGROUND MODEL DEPENDENCE

In this Appendix we study the effects of assuming an exponential, rather than flat, background energy spectrum

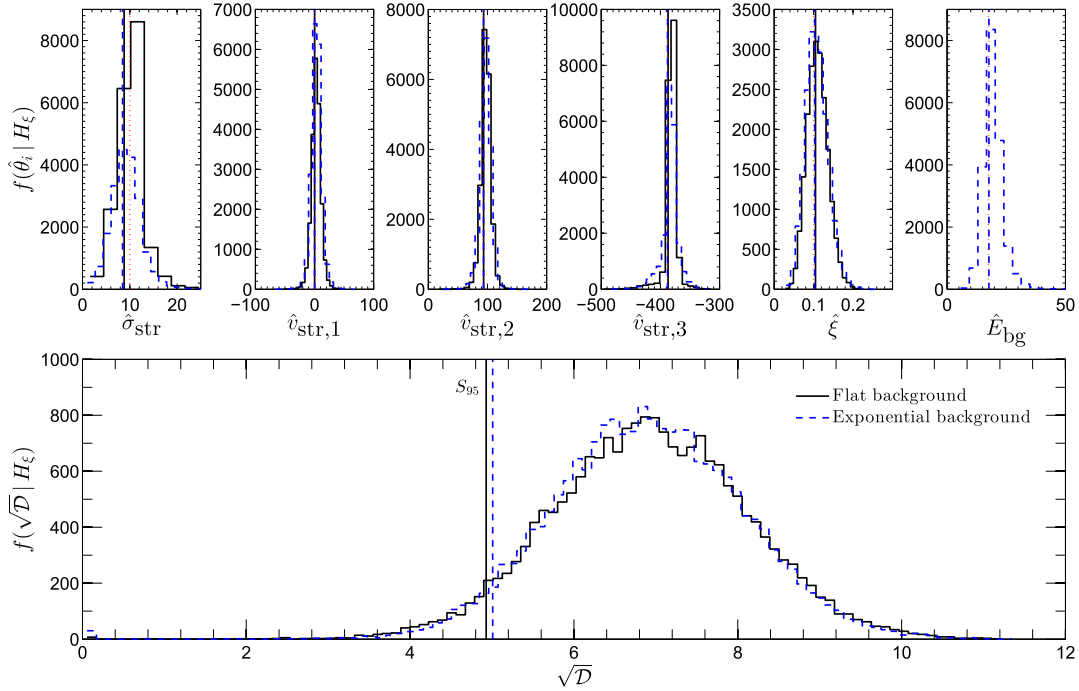


FIG. 12 (color online). Top row: The distributions of the maximum likelihood estimators (MLE) of the stream parameters and the background energy dependence parameter  $E_{\text{bg}}$  obtained over 20000 Monte Carlo runs for the flat and exponential background models (solid and dashed lines, respectively). The vertical lines indicate the mean MLE values and the red dotted vertical lines the input parameter values. Bottom plot: the distribution of the likelihood ratio test statistic for the flat and exponential background models (solid and dashed lines, respectively); the vertical lines mark the location of the significance achievable in 95% of experiments,  $S_{95}$ .

on the stream parameter estimation in Sec. IV B. The “worst possible” energy dependence would be an exponential with a characteristic scale that is similar to that of the WIMP signal, i.e.

$$\frac{d^2R}{dE d\Omega_q} = \frac{R_{\text{bg}}}{4\pi E_{\text{bg}}} \frac{e^{-E/E_{\text{bg}}}}{e^{-5 \text{ keV}/E_{\text{bg}}} - e^{-100 \text{ ke V}/E_{\text{bg}}}}, \quad (\text{A1})$$

where  $R_{\text{bg}}$  is defined as the rate observed in an experiment with a window of [5, 100] keV and  $E_{\text{bg}} = 17.5$  keV for a

background which mimics the energy dependence of a 50 GeV WIMP. The previous 11-dimensional parameter space is expanded to 12 dimensions with the addition of  $E_{\text{bg}}$ .

In Fig. 12 we plot the distribution of the likelihood ratio test statistic and the maximum likelihood estimators for the stream parameters of a Sagittarius-like stream for both flat and exponential backgrounds. It can be seen that the distribution of the test statistic and the accuracy with which the parameters are recovered are nearly identical for the two background models.

- 
- [1] A. K. Drukier, K. Freese, and D. N. Spergel, *Phys. Rev. D* **33**, 3495 (1986).
  - [2] K. Freese, M. Lisanti, and C. Savage, *Rev. Mod. Phys.* **85**, 1561 (2013).
  - [3] A. H. G. Peter, V. Glusevic, A. M. Green, B. K. Kavanagh, and S. K. Lee, arXiv:1310.7039.
  - [4] D. N. Spergel, *Phys. Rev. D* **37**, 1353 (1988).
  - [5] M. Pato, L. Baudis, G. Bertone, R. Ruiz de Austri, L. E. Strigari, and R. Trotta, *Phys. Rev. D* **83**, 083505 (2011).
  - [6] A. H. G. Peter, *Phys. Rev. D* **83**, 125029 (2011).
  - [7] C. J. Copi, J. Heo, and L. M. Krauss, *Phys. Lett. B* **461**, 43 (1999).
  - [8] C. J. Copi and L. M. Krauss, *Phys. Rev. D* **63**, 043507 (2001).
  - [9] J. Billard, F. Mayet, and D. Santos, *Phys. Rev. D* **82**, 055011 (2010).
  - [10] M. S. Alenazi and P. Gondolo, *Phys. Rev. D* **77**, 043532 (2008).
  - [11] S. Lee and A. H. G. Peter, *J. Cosmol. Astropart. Phys.* **04** (2012) 029.
  - [12] A. M. Green and B. Morgan, *Astropart. Phys.* **27**, 142 (2007).
  - [13] J. Billard, F. Mayet, J. Macias-Perez, and D. Santos, *Phys. Lett. B* **691**, 156 (2010).



- [14] A. M. Green and B. Morgan, *Phys. Rev. D* **81**, 061301 (2010).
- [15] S. Ahlen *et al.*, *Int. J. Mod. Phys. A* **25**, 1 (2010).
- [16] P. Gondolo, *Phys. Rev. D* **66**, 103513 (2002).
- [17] S. Ahlen *et al.* (DMTPC Collaboration), *Phys. Lett. B* **695**, 124 (2011).
- [18] E. Daw *et al.*, *EAS Publ. Ser.* **53**, 11 (2012).
- [19] C. Grignon *et al.* (MIMAC Collaboration), *JINST* **4**, P11003 (2009).
- [20] K. Miuchi *et al.*, *Phys. Lett. B* **686**, 11 (2010).
- [21] A. M. Green, *Mod. Phys. Lett. A* **27**, 1230004 (2012).
- [22] M. Kuhlen, N. Weiner, J. Diemand, P. Madau, B. Moore, D. Potter, J. Stadel, and M. Zemp, *J. Cosmol. Astropart. Phys.* **02** (2010) 030.
- [23] Y.-Y. Mao, L. Strigari, R. H. Weschler, H.-Y. Wu, and O. Hahn, *Astrophys. J.* **764**, 35 (2013).
- [24] Y.-Y. Mao, L. Strigari, and R. H. Weschler, *Phys. Rev. D* **89**, 063513 (2014).
- [25] M. Lisanti, L. E. Strigari, J. G. Wacker, and R. H. Weschler, *Phys. Rev. D* **83**, 023519 (2011).
- [26] M. Pato, L. E. Strigari, R. Trotta, and G. Bertone, *J. Cosmol. Astropart. Phys.* **02** (2013) 041.
- [27] M. Maciejewski, M. Vogelsberger, S. D. M. White, and V. Springel, *Mon. Not. R. Astron. Soc.* **415**, 2475 (2011).
- [28] J. I. Read, G. Lake, O. Agertz, and V. P. Debattista, *Mon. Not. R. Astron. Soc.* **389**, 1041 (2008).
- [29] J. I. Read, L. Mayer, A. M. Brooks, F. Governato, and G. Lake, *Mon. Not. R. Astron. Soc.* **397**, 44 (2009).
- [30] M. Kuhlen, M. Lisanti, and D. N. Spergel, *Phys. Rev. D* **86**, 063505 (2012).
- [31] M. Lisanti and D. N. Spergel, *Phys. Dark Univ.* **1**, 155 (2012).
- [32] B. Yanny *et al.*, *Astrophys. J.* **588**, 824 (2003).
- [33] S. R. Majewski, M. F. Skrutskie, M. D. Weinberg, and J. C. Ostheimer, *Astrophys. J.* **599**, 1082 (2003).
- [34] H. J. Newberg *et al.*, *Astrophys. J.* **596**, L191 (2003).
- [35] A. Sollima, E. Carretta, V. D'Orazi, R. G. Gratton, A. Bragaglia, and S. Lucatello, arXiv:1406.6198 [*Mon. Not. R. Astron. Soc.* (to be published)].
- [36] C. W. Purcell, A. R. Zentner, and M.-Y. Wang, *J. Cosmol. Astropart. Phys.* **08** (2012) 027.
- [37] J. D. Lewin and P. F. Smith, *Astropart. Phys.* **6**, 87 (1996).
- [38] J. Engel, *Phys. Lett. B* **264**, 114 (1991).
- [39] G. Jungman, M. Kamionkowski, and K. Griest, *Phys. Rep.* **267**, 195 (1996).
- [40] S. K. Lee, M. Lisanti, A. H. G. Peter, and B. R. Safdi, *Phys. Rev. Lett.* **112**, 011301 (2014).
- [41] N. Bozorgnia and T. Schwetz, *J. Cosmol. Astropart. Phys.* **08** (2014) 013.
- [42] F. J. Kerr and D. Lynden-Bell, *Mon. Not. R. Astron. Soc.* **221**, 1023 (1986).
- [43] J. Bovy *et al.*, *Astrophys. J.* **759**, 131 (2012).
- [44] V. V. Bobylev and A. T. Bajkova, arXiv:1404.6987.
- [45] M. C. Smith *et al.*, *Mon. Not. R. Astron. Soc.* **379**, 755 (2007).
- [46] S. R. Deans, *The Radon Transform and Some of Its Applications* (John Wiley & Sons, New York, 1983).
- [47] C. Savage, K. Freese, and P. Gondolo, *Phys. Rev. D* **74**, 043531 (2006).
- [48] K. Freese, P. Gondolo, and H. J. Newberg, *Phys. Rev. D* **71**, 043516 (2005).
- [49] 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).
- [50] J. Billard, F. Mayet, and D. Santos, *Phys. Rev. D* **85**, 035006 (2012).
- [51] E. Aprile *et al.* (XENON100 Collaboration), *Phys. Rev. Lett.* **111**, 021301 (2013).
- [52] B. Morgan, A. M. Green, and N. J. C. Spooner, *Phys. Rev. D* **71**, 103507 (2005).
- [53] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, *Astrophys. J.* **622**, 759 (2005).
- [54] N. I. Fisher, T. Lewis, and B. J. J. Embleton, *Statistical Analysis of Spherical Data* (Cambridge University Press, Cambridge, UK, 1987).
- [55] T. Piffl *et al.*, *Astron. Astrophys.* **562**, A91 (2014).
- [56] J. I. Read, *J. Phys. G* **41**, 063101 (2014).
- [57] F. Feroz and M. P. Hobson, *Mon. Not. R. Astron. Soc.* **384**, 449 (2008).
- [58] F. Feroz, M. P. Hobson, and M. Bridges, *Mon. Not. R. Astron. Soc.* **398**, 1601 (2009).
- [59] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Eur. Phys. J. C* **71**, 1554 (2011).
- [60] A. M. Green and B. Morgan, *Phys. Rev. D* **77**, 027303 (2008).