



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Stochastic properties of ultralight scalar field gradients

Mariangela Lisanti, Matthew Moschella, and William Terrano

Phys. Rev. D **104**, 055037 — Published 24 September 2021

DOI: [10.1103/PhysRevD.104.055037](https://doi.org/10.1103/PhysRevD.104.055037)

Stochastic Properties of Ultralight Scalar Field Gradients

Mariangela Lisanti,¹ Matthew Moschella,¹ and William Terrano¹

¹*Department of Physics, Princeton University, Princeton, NJ 08544*

Ultralight axion-like particles are well-motivated dark matter candidates that are the target of numerous direct detection efforts. In the vicinity of the Solar System, such particles can be treated as oscillating scalar fields. The velocity dispersion of the Milky Way determines a coherence time of about 10^6 oscillations, beyond which the amplitude of the axion field fluctuates stochastically. Any analysis of data from an axion direct detection experiment must carefully account for this stochastic behavior to properly interpret the results. This is especially true for experiments sensitive to the gradient of the axion field that are unable to collect data for many coherence times. Indeed, the direction, in addition to the amplitude, of the axion field gradient fluctuates stochastically. We present the first complete stochastic treatment for the gradient of the axion field, including multiple computationally efficient methods for performing likelihood-based data analysis, which can be applied to any axion signal, regardless of coherence time. Additionally, we demonstrate that ignoring the stochastic behavior of the gradient of the axion field can potentially result in failure to discover a true axion signal.

I. INTRODUCTION

If dark matter (DM) consists primarily of bosons with mass $m_a \ll \text{eV}$, then the DM density near the Sun guarantees that the number density of particles is large enough to be treated as an oscillating classical field. The wavelike nature of such ultralight scalar fields—often referred to as axions—can lead to distinctive signatures that impact experimental searches. In particular, interference among these waves results in a locally stochastic field that can potentially enhance or suppress the signal recorded by a laboratory experiment [1, 2]. In this work, we present a complete treatment of the stochastic properties of the gradient of the axion field, and explore its experimental consequences.

Axions are common in many extensions of the Standard Model, and correspond to Goldstone bosons that are produced when a global symmetry is broken. The QCD axion remains one of the most compelling examples as it both provides a plausible DM candidate and resolves the strong CP problem [3–13]. For the case of the QCD axion, there is a direct relationship between its mass and the symmetry-breaking scale in the theory. More generally, axion-like particles (ALPs) can exist with a mass that is not set entirely by the symmetry-breaking scale. Such ALPs can be copiously produced in string theories, for example [14]. Throughout this paper, we will use the word ‘axion’ to refer to any ultralight scalar boson that couples to the axial current.

Axions can couple to the Standard Model, opening the possibility of discovering them in the laboratory or beyond. The axion coupling to fermions takes the form of a derivative term in the Lagrangian of the form $\mathcal{L} \propto g_{aff} \partial_\mu a \bar{f} \gamma^5 \gamma^\mu f$, which in the non-relativistic limit becomes $\mathcal{L} \propto g_{aff} \nabla \mathbf{a} \cdot \mathbf{S}_f$ with \mathbf{S}_f being the fermion spin. Searches for these gradient interactions typically rely on sophisticated atomic-molecular-optical (AMO) or nuclear magnetic resonance (NMR) techniques, which are sensitive to small changes in, e.g., neutron spins [15–19]. DM searches for the axion-fermion couplings have

been performed using data from neutron electric dipole moment (nEDM) experiments [20], Penning-trap systems [21], CASPER [22, 23], the Eöt-Wash spin-polarized torsion balance [24], and various atomic magnetometers [25–28], including the NASDUCK experiment [29]. Additional laboratory constraints come from searches for new long-range forces [30, 31] and invisible meson decays [32]. The strongest constraints on the axion-fermion couplings are astrophysical, coming from axions produced in the Sun [33], as well as neutron star [34], supernova [35], and white dwarf [36] cooling. These astrophysical constraints may however be subject to large uncertainties (see, e.g., Ref. [37]).

The pseudoscalar axion can also couple to photons through the Lagrangian operator $\mathcal{L} \propto \frac{1}{4} g_{a\gamma\gamma} a \mathbf{F} \cdot \mathbf{F}$, which in the non-relativistic limit reduces to $\mathcal{L} \propto g_{a\gamma\gamma} a \mathbf{E} \cdot \mathbf{B}$. This coupling has been the target of axion DM searches with experiments such as ADMX [38–40], HAYSTAC [41, 42], ABRACADABRA [43–45], and SHAFT [46]. Additional constraints on $g_{a\gamma\gamma}$ that do not rely on the local axion DM density come from direct laboratory searches [47–49], axion helioscopes such as CAST [50], and other indirect astrophysical searches [51–55].

Any experiment searching for axion DM must contend with the unique phenomenology that arises from its wavelike behavior. At a fixed position in space, the axion field oscillates with frequency $\omega \approx m_a + \frac{1}{2} m_a v^2$, where v is the DM velocity. If the DM is virialized in the Milky Way with a velocity dispersion $\sigma_v \sim 10^{-3}$, the observable classical field will be the superposition of many oscillating fields with a frequency dispersion $\Delta\omega \sim \frac{1}{2} m_a \sigma_v^2$. The effect of this frequency dispersion is to introduce a coherence timescale, $\tau_c \sim 1/m_a \sigma_v^2$, to the oscillations. On timescales $\tau \ll \tau_c$, the DM classical field is well-described by oscillations at the Compton frequency m_a with a single amplitude and phase; however, on timescales $\tau \gtrsim \tau_c$, interference effects between the different frequency modes cause the amplitude and phase of oscillations to vary stochastically. One consequence of this is the presence of localized regions where total destructive interference

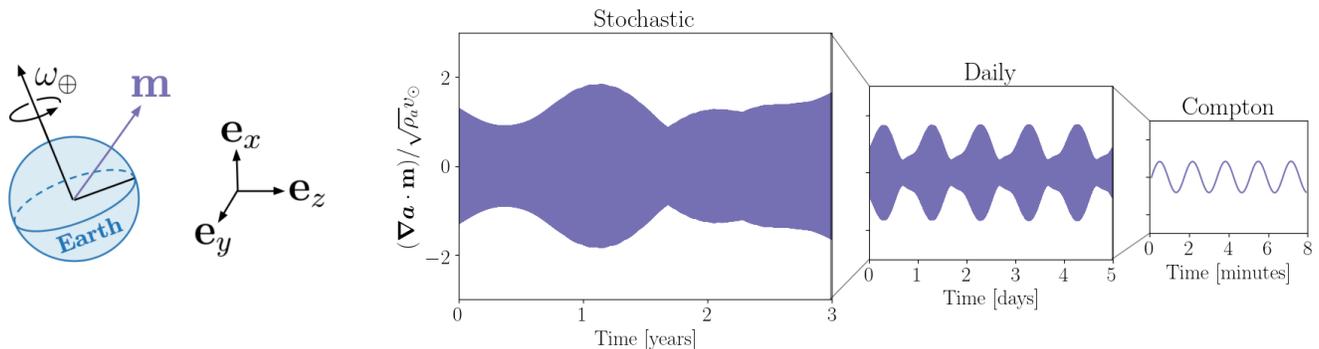


FIG. 1. **Left:** A schematic illustration of the orientation of the laboratory measurement axis \mathbf{m} , which we assume has a fixed orientation relative to the surface of the Earth at some latitude and longitude, and our Galactic-frame basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$. As the Earth rotates about its axis relative to the inertial Galactic frame, the Galactic-frame components of the measurement axis \mathbf{m} oscillate with period of 1 day, thereby inducing daily modulation into the signal $\nabla \mathbf{a} \cdot \mathbf{m}$. Note that the daily modulation of $\nabla \mathbf{a} \cdot \mathbf{m}$ is periodic but not necessarily sinusoidal and that its exact form is stochastic, since it depends on the instantaneous direction of the axion gradient. **Right:** A single randomly generated time series for an axion gradient signal $\nabla \mathbf{a} \cdot \mathbf{m}$ with mass $m_a/2\pi = 0.01$ Hz over the course of 3 years. On the smallest timescales, the signal undergoes rapid coherent oscillations with period $2\pi/m_a = 100$ s. The amplitude of these oscillations modulates with period of 1 day due to the rotation of the Earth, and evolves randomly on long timescales due to the coherence time $\tau_c \sim 6$ months.

drives the axion field to zero, resulting in vortices [56].

For experiments sensitive to the gradient of the axion field, there is an additional source of time dependence. This arises because such experiments do not measure the entire vector $\nabla \mathbf{a}$, but the projection onto a sensitive axis, which is fixed in the laboratory frame.¹ Because the direction of $\nabla \mathbf{a}$ is uniquely determined in the Galactic frame, this introduces an additional modulation of the signal due to the rotation of the Earth, as illustrated in the far left of Fig. 1. From an experimental perspective, this daily modulation is particularly useful when the Compton period is longer than 1 day, as it up-modulates the signal to a manageable frequency.

An axion gradient signal is therefore characterized by three timescales: the Compton period, $2\pi/m_a$; the coherence time, $\tau_c \sim 10^6/m_a$; and the period of the Earth's rotation, 1 day. The relative size of these timescales and the length of experimental observation can dramatically affect the phenomenology of the signal. Fig. 1 provides a visualization of each of these timescales for an experiment that is sensitive to the axion field gradient. The figure shows the measured time series over the span of 3 years for an axion of mass $m_a/2\pi = 0.01$ Hz. For the particular axion mass shown in the figure, the three timescales obey the following hierarchy: $2\pi/m_a \ll 1 \text{ day} \ll \tau_c$. On very short timescales (right panel), the signal oscillates coherently with frequency m_a . The amplitude of these coherent oscillations undergoes daily modulations with a

period of 1 day (middle panel). On very long timescales (left panel), the signal decoheres and the overall amplitude fluctuates stochastically.

The stochastic nature of the axion field can have a profound effect on experimental observations—potentially leading to a suppression or enhancement in the signal, depending on the nature of the interference and its time dependence. These effects must be properly accounted for in a complete data analysis to obtain reliable limits or properly recover a signal. Although many experimental analyses have simply ignored the stochastic nature of the axion field, such effects have been well understood for experiments that are sensitive to the amplitude of the axion field itself [1, 2, 57]. However, these results do not immediately generalize to the case of the gradient of the axion field, where the *direction* and not just the *amplitude* of the axion field fluctuates stochastically. Although there have been some attempts in recent literature to address the stochastic nature of the gradient of the axion field [28, 57], they have relied on inadequate assumptions that do not capture the parametric freedom of an axion gradient signal.

In this paper, we present a general stochastic description for the gradient of the axion field as well as provide a likelihood formalism for the statistical analysis of experimental data. This stochastic formalism can be used to analyze axion signals of any frequency, regardless of the coherence time. We demonstrate, contrary to the assumptions of Ref. [57], that the gradient of the axion field has both a random amplitude and direction, and that an analysis that ignores this effect has a non-negligible chance of failing to correctly identify an axion signal. Sec. II introduces the basic formalism we use to model the superposition of axion states. Then, in Sec. III, we

¹ It is possible for an experiment to have multiple sensitive axes. Although we do not consider this case specifically in this paper, the procedure outlined here is straightforward to generalize. See Ref. [29] for an example.

demonstrate how to perform an analysis for data taken in the time domain. Sec. IV specializes to the specific case of a coherent axion signal and summarizes the main conclusions of our work using some examples on mock data. We conclude in Sec. V. Three appendices are also included, which discuss how to derive the Central Limit Theorem (Appendix A), how to estimate uncertainties using Ordinary Least Squares (Appendix B), and how to evaluate the two-point correlation functions presented in the main text for the case of a Maxwellian velocity distribution (Appendix C).

II. SUPERPOSITION OF AXION WAVES

This section presents the basic formalism for axion wave superposition. We generalize the treatment of the axion field presented in Ref. [2] to the case of the axion field gradient. The formalism introduced here will be applied in Sec. III to construct a likelihood to analyze time-domain data.

In general, due to the time-dependent motion of the terrestrial laboratory relative to the Galactic frame, it is best to work with Galactic-frame velocities, \mathbf{w} , rather than laboratory-frame velocities, \mathbf{v} , as the underlying state variables. This transformation is given by the Galilean boost $\mathbf{v}(t) = \mathbf{w} + \mathbf{v}_{\text{obs}}(t)$, where

$$\mathbf{v}_{\text{obs}}(t) = \mathbf{v}_{\odot} + \mathbf{v}_{\oplus}(t) + \mathbf{v}_{\text{lab}}(t) \quad (1)$$

is the velocity of the terrestrial laboratory relative to the Galactic frame, which depends on the Sun's velocity relative to the Galactic Center, \mathbf{v}_{\odot} , the Earth's velocity relative to the Sun, $\mathbf{v}_{\oplus}(t)$, and the experiment's velocity relative to the Earth's center, $\mathbf{v}_{\text{lab}}(t)$.

We treat the axion as a classical field of mass m_a consisting of a very large number of states, $N_a \gg 1$, and assume that the local axion density, ρ_a , comprises the entirety of the local DM. The field contribution from an individual state with Galactic-frame velocity \mathbf{w}_λ is therefore

$$a_\lambda(\mathbf{x}, t) = \frac{\sqrt{2\rho_a/N_a}}{m_a} \cos(E_\lambda(t)t + \mathbf{k}_\lambda(t) \cdot \mathbf{x} + \phi_\lambda), \quad (2)$$

where $\lambda \in 1, 2, \dots, N_a$ is an index that runs over all states, $E_\lambda(t) \approx m_a + \frac{1}{2}m_a v_\lambda(t)^2$ is the energy of the state, $\mathbf{k}_\lambda(t) \approx m_a \mathbf{v}_\lambda(t)$ is the momentum of the state, $\mathbf{v}_\lambda(t) = \mathbf{w}_\lambda + \mathbf{v}_{\text{obs}}(t)$ is the laboratory-frame velocity of the state, and $\phi_\lambda \in [0, 2\pi)$ is the phase of the state. The factor of $1/\sqrt{N_a}$ in the amplitude ensures that the total axion field, which is the superposition of all states, has root mean square density ρ_a .

The gradient of an individual axion state at a fixed point in space is therefore

$$\nabla \mathbf{a}_\lambda(t) = \sqrt{\frac{2\rho_a}{N_a}} (\mathbf{w}_\lambda + \mathbf{v}_{\text{obs}}(t)) \cos(E_\lambda(t)t + \phi_\lambda). \quad (3)$$

Moving forward, the spatial dependence will be ignored because we only consider experiments at a fixed spatial location.

The total gradient of the axion field is obtained by summing Eq. (3) over all individual states

$$\nabla \mathbf{a}(t) = \sum_\lambda \sqrt{\frac{2\rho_a}{N_a}} (\mathbf{w}_\lambda + \mathbf{v}_{\text{obs}}(t)) \cos(E_\lambda(t)t + \phi_\lambda), \quad (4)$$

where each ϕ_λ is independent and identically distributed (IID) from the uniform distribution on the interval $[0, 2\pi)$ and each \mathbf{w}_λ is IID from the Galactic-frame velocity distribution $f(\mathbf{w})$.²

Because the number of terms in the summation of Eq. (4) is very large, and each term is an IID random variable, the Central Limit Theorem guarantees that $\nabla \mathbf{a}(t)$ is a Gaussian process. That is, $\nabla \mathbf{a}(t)$ is a normally distributed random vector at any particular time with a random time evolution obeying a particular two-point correlation function (see Appendix A).

Experiments that are sensitive to the gradient of the axion field typically have a single axis of sensitivity. That is, they do not measure the vector field $\nabla \mathbf{a}(t)$ but rather its projection onto some measurement axis, $\mathbf{m}(t)$. We assume that this axis is fixed in the laboratory frame, and therefore that it is time-dependent in the Galactic frame due to the rotation of the Earth about its axis, as illustrated in Fig. 1. Because the direction of the vector field $\nabla \mathbf{a}(t)$ is determined by the velocity of the axion particles, which are inertial in the Galactic frame, this projection causes the signal to undergo daily modulation with frequency $\omega_\oplus = 2\pi/(\text{sidereal day})$. The signal $\nabla \mathbf{a}(t) \cdot \mathbf{m}(t)$ can be decomposed in the Galactic frame by writing

$$\nabla \mathbf{a}(t) \cdot \mathbf{m}(t) = \sum_{i=x,y,z} [\nabla \mathbf{a}(t) \cdot \mathbf{e}_i] m_i(t), \quad (5)$$

where $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ are an arbitrary set of orthonormal basis vectors fixed in the Galactic frame. The components of $\mathbf{m}(t)$ in this basis, $m_i(t) = \mathbf{e}_i \cdot \mathbf{m}(t)$, undergo daily modulation due to the rotation of the Earth. For an axis that is fixed in the laboratory frame, this modulation can be parametrized as

$$m_i(t) = C_i \cos(\omega_\oplus t) + D_i \sin(\omega_\oplus t) + E_i, \quad (6)$$

for $i = x, y, z$. Here, C_i, D_i, E_i are constants that depend on the position and orientation of \mathbf{m} relative to the Earth's rotational axis. For concreteness in the rest of this paper, we take $\mathbf{m}(t)$ to be the upward direction at 40° N latitude and 75° W longitude, and we always measure time from J2000.

² Note that this construction assumes that the Galactic-frame DM velocity distribution is time-independent. If not, then Eq. (4) would no longer be valid for all times, but only at a particular moment of time. Fortunately, the dynamical timescale for the Milky Way is ~ 300 Myr, which is much longer than the observation time of any human-timescale experiment, and the time dependence of the Galactic-frame velocity distribution function can be safely ignored.

III. ANALYSIS STRATEGY

Next, we present a procedure for analyzing an observable signal in the time domain, using the construction of the axion field gradient in Eq. (4). We begin with a description of the likelihood formalism in Sec. III A before working through the general time-binned stochastic analysis in Sec. III B.

A. Likelihood Formalism

Consider a time series of N data points $\mathbf{D} = \{D(t_n) \mid n \in 1, \dots, N\}$, where each observation is the sum of a signal and background contribution. To write down a likelihood function for the observed data set, we must understand the probability distributions for both the signal and the background, which are assumed to be independent of each other.

The signal is related to the gradient of the axion field,

$$S(t_n) = g_{\text{eff}} \nabla \mathbf{a}(t_n) \cdot \mathbf{m}(t_n), \quad (7)$$

where g_{eff} is the effective axion coupling constant that is proportional to g_{aff} , but depends on the particular experiment at hand. For example, for a Helium-Potassium comagnetometer and an axion that couples only to neutrons, the observable can be taken to have dimensions of magnetic field with $g_{\text{eff}} = g_{aNN}/\gamma_n$, where γ_n is the gyromagnetic ratio of the neutron [28]. As discussed in Sec. II, the Central Limit Theorem guarantees that the components of $\nabla \mathbf{a}(t)$ are jointly-distributed Gaussian random variables and thus, from Eq. (7), that $S(t_n)$ forms a Gaussian process. The uniform random phases guarantee that the mean of this process is always zero. Therefore, the likelihood of the signal (ignoring background) is an N -dimensional zero-mean Gaussian function.

Throughout this work, the background is modeled by stationary Gaussian white noise with zero mean, although the generalization to other background probability distributions is straightforward. Taking this together with the signal model, the total likelihood is

$$\mathcal{L}(\mathbf{D} \mid \boldsymbol{\theta}_{\text{sig}}, \boldsymbol{\theta}_{\text{bkg}}) = \frac{1}{\sqrt{(2\pi)^{6N} \det \boldsymbol{\Sigma}}} e^{-\frac{1}{2} \mathbf{D}^\top \boldsymbol{\Sigma}^{-1} \mathbf{D}}, \quad (8)$$

where $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\text{sig}}(\boldsymbol{\theta}_{\text{sig}}) + \boldsymbol{\Sigma}_{\text{bkg}}(\boldsymbol{\theta}_{\text{bkg}})$ is the combined covariance matrix for the signal and background and $\boldsymbol{\theta}_{\text{sig}}$ and $\boldsymbol{\theta}_{\text{bkg}}$ are the signal and background model parameters, respectively. For the white noise scenario, the background covariance matrix is proportional to the identity matrix, $\boldsymbol{\Sigma}_{\text{bkg}} = \sigma_{\text{bkg}}^2 \mathbf{I}$, where σ_{bkg} is the root-mean-square (RMS) noise. The signal model parameters are the axion mass, m_a , and the coupling constant, g_{eff} .

The final ingredient for defining the likelihood function is the covariance matrix, or two-point correlator, of the signal:

$$\Sigma_{\text{sig}, nm} = \langle S(t_n) S(t_m) \rangle. \quad (9)$$

Although in principle it is possible to compute this object, the result is not particularly useful because the $N \times N$ covariance matrix is unwieldy with N typically being extremely large. However, we will show below that the effective size of the data set can be significantly reduced and that a computationally feasible analysis in the time domain is always possible.

The analysis could also be done in the frequency domain, which may be more efficient if a Fast Fourier Transform can be performed on the data and the integration time is much longer than the coherence time, $T_{\text{exp}} \gg \tau_c$. However, the daily modulation of the axion gradient signal, which leads to three peaks in frequency space at m_a and $m_a \pm \omega_\oplus$, adds significant complication to a frequency domain analysis compared to an analysis of a signal proportional to the scalar axion field [2]. We will present an example of a full frequency domain analysis in upcoming work (see also Ref. [29]).

B. Time-Binned Stochastic Analysis

Next, we discuss an efficient manner of performing a time-binned analysis for an axion signal with a relatively short coherence time compared to the experimental integration time, $\tau_c \lesssim T_{\text{exp}}$. In this regime, the signal is not coherent over the lifetime of the experiment and it is necessary to fully characterize the stochastic effects of the axion field. To begin, the expression for the signal in Eq. (7) is rewritten to separate out the deterministic and stochastic time dependencies. Specifically,

$$S(t) = \sum_{i=x,y,z} [A_i(t) \cos(m_a t) - B_i(t) \sin(m_a t)] m_i(t), \quad (10)$$

where

$$\begin{aligned} A_i(t) &= g_{\text{eff}} \sqrt{\frac{2\rho_a}{N_a}} \sum_{\lambda} v_{\lambda,i}(t) \cos\left(\frac{1}{2} m_a v_{\lambda}(t)^2 t + \phi_{\lambda}\right) \\ B_i(t) &= g_{\text{eff}} \sqrt{\frac{2\rho_a}{N_a}} \sum_{\lambda} v_{\lambda,i}(t) \sin\left(\frac{1}{2} m_a v_{\lambda}(t)^2 t + \phi_{\lambda}\right) \end{aligned} \quad (11)$$

and $v_{\lambda,i}(t)$ is the \mathbf{e}_i component of the vector $\mathbf{v}_{\lambda}(t) = \mathbf{w}_{\lambda} + \mathbf{v}_{\text{obs}}(t)$. The utility of this decomposition is immediately apparent. The signal manifestly oscillates at the Compton frequency m_a , with daily modulations at frequency ω_\oplus due to the $m_i(t)$ terms. The $A_i(t)$ and $B_i(t)$ coefficients, which vary on the coherence timescale, are Gaussian random variables that encode the stochastic behavior of the axion field.

We now consider the behavior of the signal on small timescales. We divide the time series $\mathbf{S} = \{S(t_n) \mid n \in 1, \dots, N\}$ into P bins of size $\Delta t \ll 2\pi/\omega_\oplus, \tau_c$ centered at times t_p . Within each bin, the daily modulations and stochastic fluctuations can be neglected, so the signal in

the p^{th} bin can be written as a coherent oscillation

$$S_p(t) = \tilde{A}_p \cos(m_a t) - \tilde{B}_p \sin(m_a t), \quad (12)$$

where

$$\begin{aligned} \tilde{A}_p &= \sum_{i=x,y,z} m_i(t_p) A_i(t_p) \\ \tilde{B}_p &= \sum_{i=x,y,z} m_i(t_p) B_i(t_p) \end{aligned} \quad (13)$$

are constant within each bin and vary smoothly (and stochastically) in time from bin to bin.

Because $\tau_c \sim 10^6/m_a$ and we need only consider $m_a \gg \omega_\oplus$,³ it is always possible to choose the bin size Δt such that there are many coherent oscillations within each bin. As a result, the coefficients \tilde{A}_p and \tilde{B}_p can be measured by fitting the data within each bin to the form of Eq. (12) with, e.g., an ordinary least squares procedure. This motivates an analysis strategy where the $2P$ coefficients \tilde{A}_p and \tilde{B}_p are analyzed instead of the full time series. Because there are many coherent oscillations within each bin, this method achieves a significant reduction in the effective size of the data set.⁴

We now consider the statistical analysis of such a compressed data set consisting of P measurements of the two coefficients $\{\tilde{A}(t_p), \tilde{B}(t_p)\}$ at the binned times t_p for $p = 1, \dots, P$. Eq. (11) and the Central Limit Theorem guarantee that the time series of these coefficients form joint stationary Gaussian processes. Assuming Gaussian white noise, the uncertainty on the recovered ordinary least squares coefficients is also Gaussian distributed, but with uncertainty $\tilde{\sigma} \approx \sigma_{\text{bkg}} \sqrt{2P/N}$ (see Appendix B for further details). Therefore, a likelihood of the form in Eq. (8) is still valid, and we need only compute the two-point correlation functions in order to construct the covariance matrix:

$$\begin{aligned} \langle \tilde{A}(t) \tilde{A}(t') \rangle &= \sum_{i,j} m_i(t) m_j(t') \langle A_i(t) A_j(t') \rangle \\ \langle \tilde{B}(t) \tilde{B}(t') \rangle &= \sum_{i,j} m_i(t) m_j(t') \langle B_i(t) B_j(t') \rangle \\ \langle \tilde{A}(t) \tilde{B}(t') \rangle &= \sum_{i,j} m_i(t) m_j(t') \langle A_i(t) B_j(t') \rangle. \end{aligned} \quad (14)$$

The expectation values in Eq. (14) can be evaluated using Eq. (A3), which, after integrating over the random

phases, gives

$$\begin{aligned} \langle A_i(t) A_j(t') \rangle &= g_{\text{eff}}^2 \rho_a \int d^3 \mathbf{w} v_i(t) v_j(t') f(\mathbf{w}) \cos(\Delta \varpi) \\ \langle A_i(t) B_j(t') \rangle &= g_{\text{eff}}^2 \rho_a \int d^3 \mathbf{w} v_i(t) v_j(t') f(\mathbf{w}) \sin(\Delta \varpi) \\ \langle A_j(t) A_j(t') \rangle &= \langle B_j(t) B_j(t') \rangle, \end{aligned} \quad (15)$$

where $\Delta \varpi = \frac{1}{2} m_a (v^2(t') t' - v^2(t) t)$ and $\mathbf{v}(t) = \mathbf{w} + \mathbf{v}_{\text{obs}}(t)$. This is the most general solution for the covariance matrix, and can be computed for any velocity distribution of the axion DM.

While the discussion until now has remained general, it is useful to consider the specific case where the Galactic-frame DM velocity distribution is Maxwellian,

$$f(\mathbf{w}) = \frac{1}{(2\pi\sigma_v^2)^{3/2}} e^{-w^2/2\sigma_v^2}, \quad (16)$$

with velocity dispersion $\sigma_v \approx 155$ km/s [58–60]. Note that to compute the integrals in Eq. (15) analytically, we do not introduce a cutoff of the distribution at the escape velocity.

Additionally, we make the simplifying assumption that $\mathbf{v}_{\text{obs}}(t) \approx \mathbf{v}_\odot$, ignoring the (comparatively small) time-dependent contributions of the Earth's velocity around the Sun, and the laboratory's velocity around the Earth's center. We take $\mathbf{v}_\odot = (11, 232, 7)$ km/s in the standard Galactic (U, V, W) coordinate basis⁵ [61]. This assumption allows the integrals in Eq. (15) to be computed analytically via a change of variables from $\mathbf{w} \rightarrow \mathbf{v} = \mathbf{w} + \mathbf{v}_\odot$. We find that the two-point correlation functions evaluated using this approximation differ from the full solution by less than 10%.

Because $f(\mathbf{w})$ is an isotropic function and \mathbf{v}_\odot is constant, it is useful to align our otherwise arbitrary basis vectors $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ with the unique direction \mathbf{v}_\odot . Specifically, we choose $\mathbf{e}_z \parallel \mathbf{v}_\odot$. This guarantees that all two-point correlation functions are proportional to δ_{ij} , i.e., the x , y , and z components are statistically independent of one another, and that there is an $x \leftrightarrow y$ symmetry. Substituting Eq. (16) into Eq. (15), we find that

$$\begin{aligned} \langle A_z(t) A_z(t') \rangle &= \mathcal{A}_\parallel(\xi) \cos \Psi_\parallel(\xi) \\ \langle A_z(t) B_z(t') \rangle &= \mathcal{A}_\parallel(\xi) \sin \Psi_\parallel(\xi) \\ \langle A_x(t) A_x(t') \rangle &= \mathcal{A}_\perp(\xi) \cos \Psi_\perp(\xi) \\ \langle A_x(t) B_x(t') \rangle &= \mathcal{A}_\perp(\xi) \sin \Psi_\perp(\xi), \end{aligned} \quad (17)$$

where we have defined $\xi \equiv m_a \sigma_v^2 (t' - t)$ and the amplitudes and phases are:

³ In this subsection, we assume that the coherence time is relatively short, $\tau_c \lesssim T_{\text{exp}}$. Assuming a human-timescale experiment with $T_{\text{exp}} \lesssim 10$ yr, this implies that $m_a \gtrsim 40 \omega_\oplus$. When this is not satisfied, the analysis in Sec. IV can be used.

⁴ The coefficients \tilde{A}_p and \tilde{B}_p , and indeed the bin sizes themselves, have to be recomputed for each axion mass m_a ; however, so does the inverse and determinant of the covariance matrix. Since the bandwidth of any experiment is not larger than $\mathcal{O}(N)$ frequency

points, it will always be more efficient to recompute the ordinary least squares coefficients \tilde{A}_p and \tilde{B}_p than to work with the full $N \times N$ covariance matrix.

⁵ The (U, V, W) coordinate basis is aligned such that \mathbf{e}_U is the direction from the Solar System barycenter towards the Galactic center, \mathbf{e}_V is the direction of the local Milky Way disk rotation, and \mathbf{e}_W points towards the Galactic North Pole.

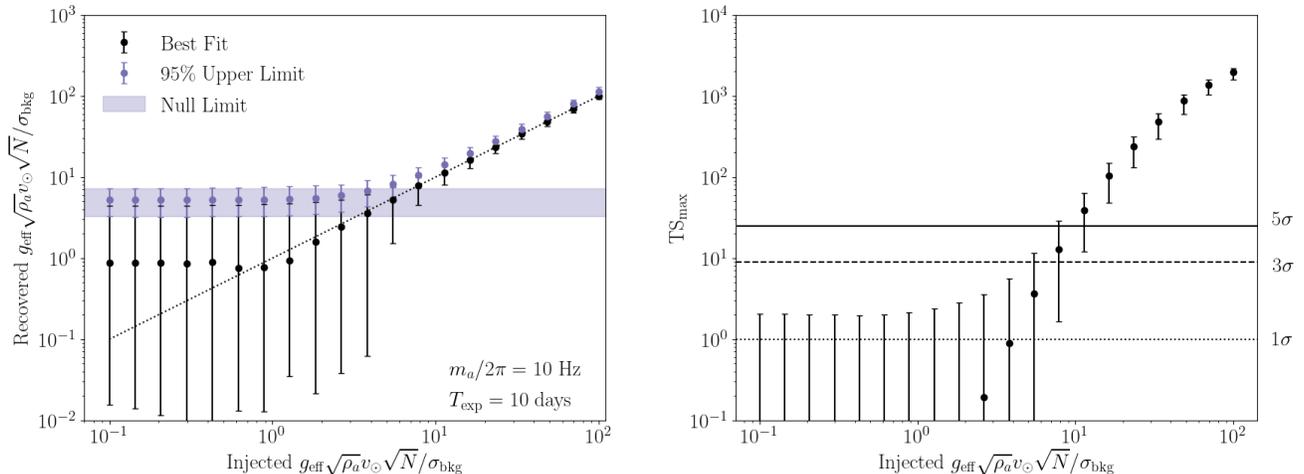


FIG. 2. **Left:** Signal injection plot using mock data to validate the analysis procedure outlined in Sec. III B. The mock data were generated with a sampling rate of 50 Hz for 10 days with stationary Gaussian white noise background with standard deviation $\sigma_{\text{bkg}} = 10^{-8} \sqrt{\rho_a} v_\odot \text{ GeV}^{-1}$. The signal was generated with an axion mass of $m_a/2\pi = 10 \text{ Hz}$ for 20 different couplings g_{eff} , shown on the horizontal axis. For each value of injected coupling, we simulate 100 random mock data sets, and analyze each data set using the Gaussian likelihood procedure described in the text. The black (purple) points indicate the median best-fit (95% upper limit) couplings over the 100 mock data sets as a function of the injected signal strength. The horizontal band indicates the 95% upper limit with zero injected signal. **Right:** The median discovery test statistic, TS_{max} , obtained in the analysis of the mock data as a function of the injected signal strength. The horizontal lines indicate 1-, 3-, and 5- σ local significance. In both panels, the error bars indicate the 5th and 95th percentiles of the best-fit coupling, recovered limit, or test statistic. The couplings are reported as a dimensionless signal-to-noise ratio $g_{\text{eff}} \sqrt{\rho_a} v_\odot \sqrt{N} / \sigma_{\text{bkg}}$, where $\sigma_{\text{bkg}} / \sqrt{N}$ is the relative scaling of the background for an infinitely coherent signal.

$$\begin{aligned}
 \mathcal{A}_{\parallel}(\xi) &= \frac{g_{\text{eff}}^2 \rho_a}{(1 + \xi^2)^{7/4}} \exp \left[-\frac{v_\odot^2}{2\sigma_v^2} \frac{\xi^2}{(1 + \xi^2)} \right] \sqrt{(\sigma_v^2 + v_\odot^2)^2 + \sigma_v^4 \xi^2} & \mathcal{A}_{\perp}(\xi) &= \frac{g_{\text{eff}}^2 \rho_a \sigma_v^2}{(1 + \xi^2)^{5/4}} \exp \left[-\frac{v_\odot^2}{2\sigma_v^2} \frac{\xi^2}{(1 + \xi^2)} \right] \\
 \Psi_{\parallel}(\xi) &= \frac{v_\odot^2}{2\sigma_v^2} \frac{\xi}{(1 + \xi^2)} + \frac{7}{2} \arctan(\xi) - \arctan \left(\frac{\sigma_v^2 \xi}{\sigma_v^2 + v_\odot^2} \right) & \Psi_{\perp}(\xi) &= \frac{v_\odot^2}{2\sigma_v^2} \frac{\xi}{(1 + \xi^2)} + \frac{5}{2} \arctan(\xi).
 \end{aligned} \tag{18}$$

Further details on the evaluation of these correlation functions are provided in Appendix C.

We validate this likelihood and analysis procedure by generating 10 days of mock data sampled at 50 Hz with Gaussian white noise and injected signal with an axion mass of $m_a/2\pi = 10 \text{ Hz}$ and 20 different values of the coupling g_{eff} . We repeat this for 100 different random iterations of noise and axion signal. For each mock data set, we bin the data in bins of length $\Delta t = 10^5 / m_a$ ⁶ and obtain the \tilde{A} and \tilde{B} coefficients in each bin from ordinary least squares fitting.

For the likelihood procedure, we use a multivariate Gaussian likelihood of the form in Eq. (8), but with the covariance matrix given by the two-point correlation functions in Eq. (14), which are computed using Eq. (17). The uncertainty of the \tilde{A} and \tilde{B} coefficients, $\tilde{\sigma}$, is treated

as a nuisance parameter and accounted for using the profile likelihood method [62, 63]. The best-fit coupling for a given axion mass, g_{best} , is obtained by maximizing the likelihood $\mathcal{L}(\mathbf{D} | m_a, g_{\text{eff}})$ for data set \mathbf{D} . The test statistic is defined as

$$\text{TS} \equiv 2 [\log \mathcal{L}(\mathbf{D} | m_a, g_{\text{best}}) - \log \mathcal{L}(\mathbf{D} | m_a, g_{\text{eff}})], \tag{19}$$

from which it follows that the 95% upper confidence limit on the coupling corresponds to the value of g_{eff} where $\text{TS} = 2.71$. To characterize the significance of a potential discovery of an axion with g_{best} , we use TS_{max} , defined as in Eq. (19), but where the maximized likelihood is compared to the null hypothesis with $g_{\text{eff}} = 0$.

The results of our mock analysis are displayed in Fig. 2, and serve as a validation for the analysis procedure. The value of the true injected axion coupling is shown on the horizontal axis of both panels. In the left panel, the black (purple) points indicate the median best-fit (95% upper limit) couplings recovered over the 100 random mock datasets for each value of injected coupling. The right panel shows the median discovery test-statistic ob-

⁶ Recall that, in general, the bin size must satisfy $\Delta t \ll \tau_c$, 1 day. In practice, we find that choosing $\Delta t = \min(10^5 / m_a, 2.4 \text{ hr})$ is enough to ensure that the essential features appear in Fig. 2.

tained for each value of injected coupling. At small injected couplings, the discovery significance is low and the 95% upper limit couplings asymptote to a constant value, as expected. At large injected couplings, the discovery significance increases dramatically, the best-fit coupling becomes centered on the true injected value, and the error bars decrease, as expected in the regime where the signal-to-noise ratio is large. In this regime, the 95% upper limit is always above the best-fit coupling and above the true injected coupling 95% of the time.

IV. THE COHERENT LIMIT

This section focuses on the limit of very long axion coherence times, $\tau_c \gg T_{\text{exp}}$, where the signal can be treated as coherent for the entire runtime of the experiment. For simplicity, we continue to assume a Maxwellian velocity distribution function and take $\mathbf{v}_{\text{obs}} \approx \mathbf{v}_{\odot}$. As seen in Eqs. (17) and (18), there is an analytic expression for the two-point correlations in this limit. In this case, $\xi \rightarrow 0$ and we have $\mathcal{A}_{\parallel} = g_{\text{eff}}^2 \rho_a (\sigma_v^2 + v_{\odot}^2)$, $\mathcal{A}_{\perp} = g_{\text{eff}}^2 \rho_a \sigma_v^2$, and $\Psi_{\parallel} = \Psi_{\perp} = 0$. Therefore, all cross-correlations vanish and each auto-correlation matrix becomes singular. This is because, in the limit of long coherence times, each coefficient $A_z(t)$, $B_z(t)$, ... does not vary over the course of the experiment and can be treated as a single Gaussian random variable, rather than a Gaussian process. This is consistent with our intuition that the coefficients should only vary on timescales of order the coherence time, and can be treated as constants on much shorter timescales. Taking these six Gaussian random variables that fully specify the signal and rescaling them so that they obey a standard normal distribution, we obtain the compact expression:

$$\begin{aligned} S(t) = & g_{\text{eff}} \sqrt{\rho_a (\sigma_v^2 + v_{\odot}^2)} \alpha_z \cos(m_a t + \phi_z) m_z(t) \\ & + g_{\text{eff}} \sqrt{\rho_a \sigma_v^2} \alpha_y \cos(m_a t + \phi_y) m_y(t) \\ & + g_{\text{eff}} \sqrt{\rho_a \sigma_v^2} \alpha_x \cos(m_a t + \phi_x) m_x(t), \end{aligned} \quad (20)$$

where α_i are three Rayleigh-distributed amplitudes, and ϕ_i are three uniformly-distributed random phases.⁷ We stress that this result is only valid for $\tau_c \gg T_{\text{exp}}$ in a coordinate system defined such that $\mathbf{e}_z \parallel \mathbf{v}_{\odot}$ and for the Maxwellian velocity distribution. Importantly, Eq. (20) contains contributions from the overlap of all three Galactic basis vectors with the sensitive axis \mathbf{m} , and each component in the Galactic basis has an independent random amplitude and phase.

To analyze a data set in the coherent limit, we can simply fit the data to the form of Eq. (20), rather than constructing a multivariate Gaussian likelihood function

with a covariance matrix that encapsulates the stochastic fluctuations in the signal. Although the amplitudes and phases in Eq. (20) are random variables, they can be treated as nuisance model parameters. The true likelihood would then be obtained in the usual way by marginalizing over the nuisance parameters. Unfortunately, with six nuisance parameters, the marginalization integrals quickly become analytically intractable and numerically cumbersome, especially for relatively large data sets. Fortunately, we can again take advantage of ordinary least squares fitting to dramatically reduce the effective size of the data set, thereby enabling a modified likelihood function to be used.

Eq. (20) is equivalent to an ordinary least squares problem of the form

$$S(t) = \sum_{i=x,y,z} [A_i \cos(m_a t) - B_i \sin(m_a t)] m_i(t), \quad (21)$$

and the data set can be reduced from a time series of size N to the six measured least-squares coefficients $\{(A_i, B_i) \mid i = x, y, z\}$. Each measured coefficient comes with an associated uncertainty from the fitting procedure. For Gaussian white noise with scale σ_{bkg} , the resulting uncertainty on the coefficients is also Gaussian. Unlike in Sec. III B, however, this uncertainty is different for each coefficient. By symmetry, the uncertainty on each A_i and B_i should be the same, which we will denote $\tilde{\sigma}_i$, with each uncertainty scaling roughly as $\tilde{\sigma}_i \sim \sigma_{\text{bkg}}/\sqrt{N}$, but with the exact form given by Eq. (B5). Since each A_i and B_i are independent Gaussian random variables, the total likelihood is the product of the individual probabilities of observing the given coefficients, marginalized over the underlying probability distributions. Because the underlying noise distribution is Gaussian as well, the result is a Gaussian likelihood

$$\mathcal{L}(A_i, B_i, \tilde{\sigma}_i | g_{\text{eff}}) = \prod_{i=x,y,z} \frac{e^{-(A_i^2 + B_i^2)/2\sigma_i^2(g_{\text{eff}}, \tilde{\sigma}_i)}}{2\pi\sigma_i^2(g_{\text{eff}}, \tilde{\sigma}_i)}, \quad (22)$$

where

$$\begin{aligned} \sigma_x^2(g_{\text{eff}}, \tilde{\sigma}_x) &= g_{\text{eff}}^2 \rho_a \sigma_v^2 + \tilde{\sigma}_x^2 \\ \sigma_y^2(g_{\text{eff}}, \tilde{\sigma}_y) &= g_{\text{eff}}^2 \rho_a \sigma_v^2 + \tilde{\sigma}_y^2 \\ \sigma_z^2(g_{\text{eff}}, \tilde{\sigma}_z) &= g_{\text{eff}}^2 \rho_a (\sigma_v^2 + v_{\odot}^2) + \tilde{\sigma}_z^2. \end{aligned} \quad (23)$$

Note that there is a near-degeneracy between g_{eff} and σ_{bkg} (through each of the $\tilde{\sigma}_i$). In Sec. III B, we treated σ_{bkg} as unknown and marginalized over $\tilde{\sigma}$ as a nuisance parameter. However, due to the degeneracy between g_{eff} and σ_{bkg} in Eq. (22), we cannot repeat the same procedure and must assume that σ_{bkg} is measured independently. For example, σ_{bkg} can be determined via $\sigma_{\text{bkg}} \approx \sqrt{\text{SSR}/N}$, where SSR is the minimum sum of the squares of the residuals obtained in the fitting procedure. We will therefore assume that σ_{bkg} and the fit coefficient uncertainties, $\tilde{\sigma}_i$, are known.

We validate this likelihood and analysis procedure by generating 10 days of mock data sampled at 10 mHz with

⁷ If X and Y are standard normal random variables, then $\sqrt{X^2 + Y^2}$ follows a Rayleigh distribution, and $\arctan(Y/X)$ follows a uniform distribution.

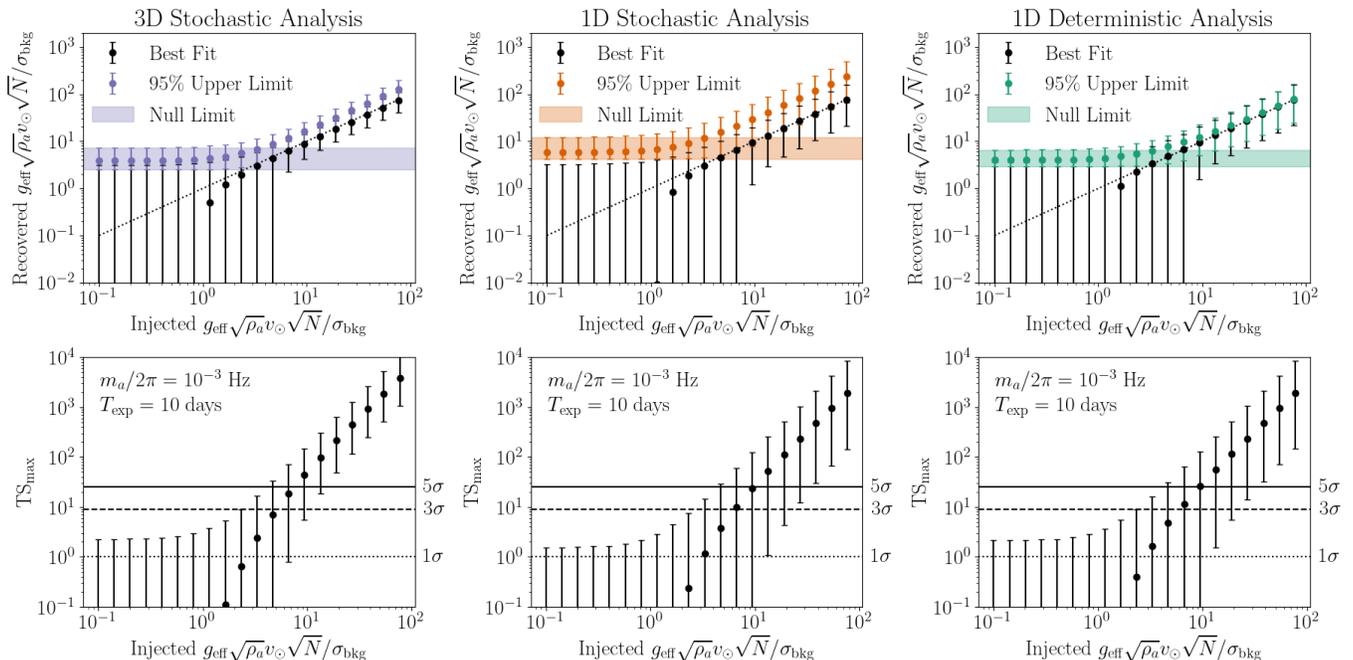


FIG. 3. Signal injection plots using mock data to test the benchmark analysis procedures considered in Sec. IV. Each column is analogous to Fig. 2 but for the 3D Stochastic (**Left**), 1D Stochastic (**Center**), and 1D Deterministic (**Right**) models. The same mock data were used for all three analyses. The mock data were generated with a sampling rate of 1 mHz for 10 days with Gaussian white noise background with standard deviation $\sigma_B = 10^{-8} \sqrt{\rho_a v_\odot}$ GeV $^{-1}$. The signal was generated with an axion mass of $m_a/2\pi = 1$ mHz for 20 different couplings g_{eff} . For each value of injected coupling, we simulate 10,000 random mock data sets, and analyze each data set using the corresponding procedure, as described in the text. The 1D Stochastic and 1D Deterministic models set an incorrect 95% upper limit on g_{eff} and exhibit more variability in their TS_{max} , as compared to the 3D Stochastic model. Additionally, for large injected coupling, the 95% upper limit for the 1D Deterministic model falls below the true value about 50% time, when this should only occur about 5% of the time for a correct model.

Gaussian white noise and injected signal with an axion frequency of $m_a/2\pi = 1$ mHz and 20 different values of the coupling g_{eff} . We repeat this for 10^4 different random iterations of noise and axion signal. For each mock data set, we obtain the six A_i and B_i coefficients as well as the uncertainty estimates $\tilde{\sigma}_i$ via ordinary least squares fitting. We then compute the best-fit coupling g_{best} , the discovery test statistic TS_{max} , and the 95% upper limit using the likelihood in Eq. (22), following the same procedure discussed in Sec. III B.

The results of the mock analysis are displayed in the left column of Fig. 3, which is analogous to Fig. 2, but for the coherent analysis discussed above. The discovery significance and 95% upper limit couplings behave as desired across all injected couplings. However, it is worth noting that the error bars on the recovered couplings and TS_{max} do not decrease in the limit of large injected coupling; this is expected behavior because, with much less than one coherence time of data, the stochastic fluctuations do not “average out.”

Next, we compare the data model and analysis procedure outlined in this section against two benchmark models that do not fully encapsulate the behavior of Eq. (20). Our goal is to motivate the use of the full stochastic axion model presented in this paper and show that certain sim-

plifying assumptions that have previously been made in the literature can fail to capture the crucial phenomenology of a true underlying axion signal. This can result in unreliable upper limits from experimental data and, more dramatically, missed discoveries of axion signals.

The first benchmark model we consider is purely deterministic and equivalent to the full stochastic model in Eq. (20) in the limit of zero velocity dispersion. In this case, the direction of $\nabla \mathbf{a}$ is the same as the average DM velocity $\mathbf{v}_{\text{obs}}(t) \approx \mathbf{v}_\odot$, and therefore, the signal is proportional to $\mathbf{v}_\odot \cdot \mathbf{m}(t) = v_\odot m_z(t)$:

$$S_{1\text{D,det.}}(t) = g_{\text{eff}} \sqrt{2\rho_a v_\odot^2} \cos(m_a t + \phi) m_z(t). \quad (24)$$

Here, ϕ is a random phase, but note that the amplitude of oscillations is determined solely by the axion coupling g_{eff} and astrophysical parameters. The factor of $\sqrt{2}$ in the amplitude ensures that the signal has the same RMS as Eq. (20) in the limit of zero velocity dispersion. We refer to this model as the 1D Deterministic model because it assumes that the axion gradient is always in one direction and is fully deterministic.

The second benchmark model is inspired by the treatment of the gradient of the axion field in Ref. [57]. This model again assumes that the direction of $\nabla \mathbf{a}$ is deter-

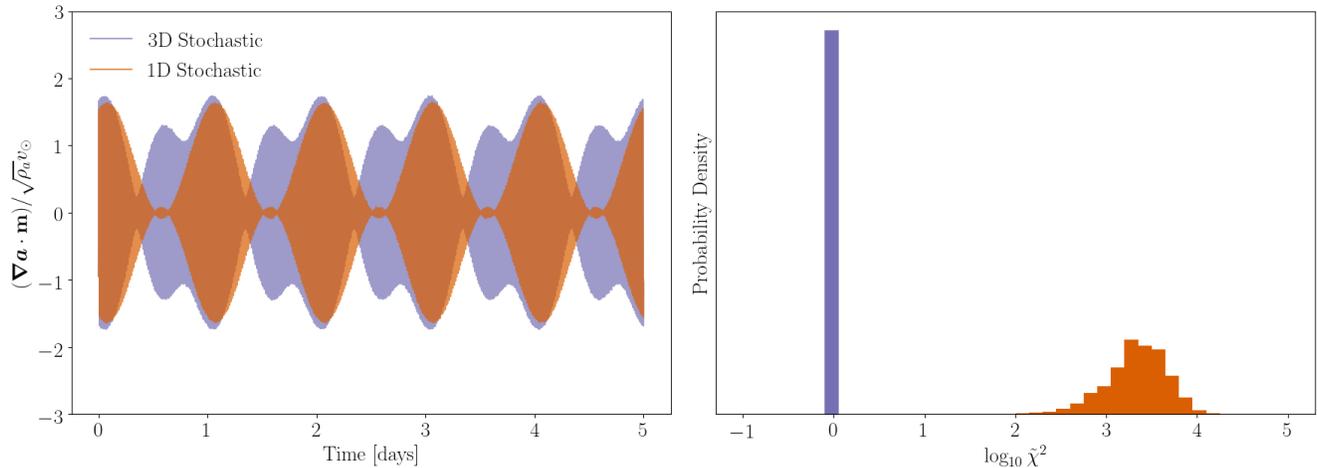


FIG. 4. **Left:** Randomly generated time series for an axion gradient signal $\nabla \mathbf{a} \cdot \mathbf{m}$ in the coherent limit. The data were generated with an axion mass of $m_a/2\pi = 10^{-3}$ Hz ($\tau_c \sim 5$ yr) for 5 days. The data include Gaussian white noise with standard deviation $\sigma_{\text{bkg}} = 10^{-2} \sqrt{\rho_a} v_\odot$. The purple curve is a random realization of a true axion signal, the 3D Stochastic model, given by Eq. (20) with a set of randomly generated amplitudes α_i and phases ϕ_i . The orange curve is a realization of the 1D Stochastic model given in Eq. (25) with the same amplitude and phase in the \mathbf{e}_z component as in the realization of the 3D model, but no contribution from the \mathbf{e}_x or \mathbf{e}_y components. The 1D model is an insufficient approximation of the full 3D model. Note, in particular, that the amplitude modulation in the orange curve is determined entirely by $m_z(t)$, while in the purple curve it is determined by a random linear combination of $m_x(t)$, $m_y(t)$, and $m_z(t)$, giving the purple curve parametrically more freedom than the orange curve. **Right:** The distribution of the reduced chi-squared statistic for the correct 3D Stochastic model and the 1D Stochastic model over many random iterations of the same time series. For each model, the best-fit amplitudes and phases are obtained by fitting the mock data to either Eq. (20) or Eq. (25). The reduced chi-squared statistic is defined as $\tilde{\chi}^2 = \sum_n \frac{S(t_n) - D_n}{\nu \sigma_{\text{bkg}}}$, where $S(t_n)$ is the predicted signal for the model given the least-squares amplitudes and phases and ν is the number of degrees of freedom equal to N minus the number of fit parameters. The reduced chi-squared statistic serves as a measure of goodness-of-fit for the models, with $\tilde{\chi}^2 \approx 1$ indicating a good fit, as is the case for the 3D model, and $\tilde{\chi}^2 \gg 1$ indicating a poor fit, as is the case for the 1D model.

mined solely by \mathbf{v}_\odot , so that the signal is still proportional to $v_\odot m_z(t)$; however, this model includes an additional random amplitude to account for stochastic interference effects:

$$S_{\text{1D, stoch.}}(t) = g_{\text{eff}} \sqrt{\rho_a v_\odot^2} \alpha \cos(m_a t + \phi) m_z(t). \quad (25)$$

Here, α is a Rayleigh-distributed random amplitude, and ϕ is a uniformly-distributed random phase. We refer to this model as the 1D Stochastic model.⁸ For the sake of comparison, we refer to the correct model given in Eq. (20) as the 3D Stochastic model, since it does not fix the direction of $\nabla \mathbf{a}$, but assigns a random amplitude and phase independently to each of the three components of $\nabla \mathbf{a}$, effectively giving the axion gradient a random direction as well as a random amplitude.

The 1D Stochastic model is motivated by previous studies, such as Refs. [57] and [28], that have incorrectly assumed that the axion signal contains a single amplitude and phase coming from the overlap of the sensitive axis with the direction of the average DM velocity,

⁸ Note that the 1D Stochastic model lacks self-consistency because there are no stochastic amplitude fluctuations in the limit of zero velocity dispersion.

\mathbf{e}_z . As illustrated in Fig. 4, this incorrect assumption implies a dramatic difference in the behavior of the signal time series. In the left panel of Fig. 4, the purple curve shows a randomly generated time series for an axion mass of $m_a/2\pi = 10^{-3}$ Hz ($\tau_c \sim 5$ yr) for 5 days, including Gaussian white noise with $\sigma_{\text{bkg}} = 10^{-2} / \sqrt{\rho_a} v_\odot$. This should be compared to the orange curve, which has the same amplitude and phase in the \mathbf{e}_z direction, but does not include the \mathbf{e}_x and \mathbf{e}_y components as in the 1D Stochastic model of Eq. (25). There are clear differences in the time evolution of the axion signal in the 1D and 3D Stochastic cases. Thus, it is not surprising that attempting to fit a real axion signal using the 1D Stochastic model assumption results in an extremely poor goodness-of-fit as quantified by the reduced chi-squared statistic in the right panel of the figure.

Both the 1D Deterministic and 1D Stochastic models contain no notion of a two-point correlation function. It is therefore not immediately clear how to apply these models to a more general scenario as we considered in Sec. III B. Although this point alone provides a strong motivation for the use of the full axion model, in the limit of long coherence times, the signal can be treated as coherent and these models can be directly compared to the correct model given by Eq. (20).

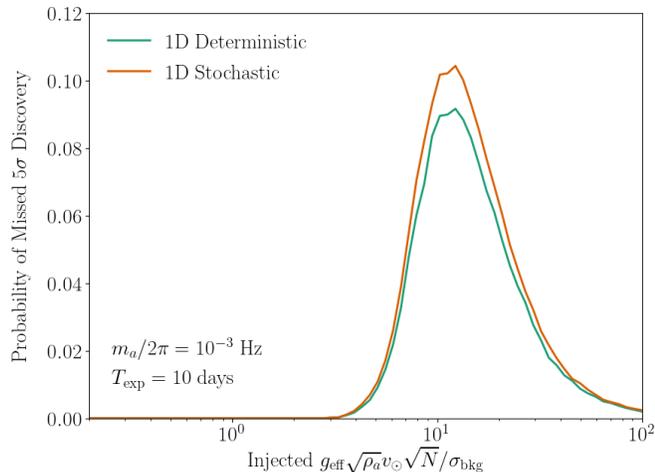


FIG. 5. The probability that an axion signal fluctuates such that the full 3D Stochastic analysis yields a local significance $\sqrt{\text{TS}_{\text{max}}} > 5\sigma$ and the 1D Stochastic (orange) or 1D Deterministic (green) analyses yield a local significance $\sqrt{\text{TS}_{\text{max}}} < 3\sigma$. This corresponds to the probability that a 5σ discovery, which could have been made using the 3D Stochastic analysis, would be “missed” in either of the other two benchmark analyses. This probability is plotted as a function of injected signal for the same mock data used in Fig. 3. For injected signals with signal-to-noise ratios of ~ 1 – 10 , the probability of a missed discovery is $\gtrsim 10\%$ for both the 1D Stochastic and 1D Deterministic analyses. This occurs because the 1D models do not sufficiently capture the time evolution of the axion signal. In particular, the daily modulation of a real axion signal need not follow the form of $m_z(t)$, as shown in Fig. 4.

We use the 1D Deterministic and 1D Stochastic models to analyze the same mock data that was used to validate the 3D Stochastic model. For the 1D Stochastic model, the data analysis procedure is exactly analogous to the 3D Stochastic model, except only the $i = z$ parts of Eq. (21) and Eq. (22) are used. For the 1D Deterministic model, the likelihood is different, because the signal amplitude is uniquely determined given the model parameters. In this case, the correct probability distribution for the observed amplitude is a Rice distribution. Changing variables from the observed amplitude and phase to the measured coefficients, A_z and B_z , the normalized likelihood is

$$\begin{aligned} \mathcal{L}_{\text{1D,det.}}(A_z, B_z, \tilde{\sigma}_z | g_{\text{eff}}) & \quad (26) \\ &= \frac{e^{-(R_z^2 + 2g_{\text{eff}}^2 \rho_a v_\odot^2) / 2\tilde{\sigma}_z^2}}{2\pi\tilde{\sigma}_z^2} I_0 \left(\frac{R_z g_{\text{eff}} \sqrt{2\rho_a v_\odot^2}}{\tilde{\sigma}_z^2} \right), \end{aligned}$$

where we have introduced the shorthand for the measured amplitude $R_z = \sqrt{A_z^2 + B_z^2}$, and $I_0(z)$ is the modified Bessel function of the first kind with order zero. With this likelihood, the data analysis for the 1D Deterministic model proceeds analogously to the 1D Stochastic and 3D Stochastic model.

We compare the results of the mock data analysis using the three benchmark models in Fig. 3. In particular, we find that neither the 1D Stochastic nor the 1D Deterministic models set the correct 95% upper limit on g_{eff} . Additionally, in the limit of large injected coupling, the 95% upper limit set by the 1D Deterministic analysis is below the true injected value about 50% of the time, while the expected frequency of this occurrence is 5% for a correct model. More concerning is the different level of local statistical significance as measured by TS_{max} for the three models. The lower panels in Fig. 3 show that there is considerable more variability in the TS_{max} in the 1D Stochastic and 1D Deterministic models than in the 3D Stochastic model.

We find that there can be considerable difference in TS_{max} across the three models for a given set of mock data. As shown in Fig. 5, for certain signal-to-noise ratios, both the 1D Deterministic and 1D Stochastic models have a $\sim 10\%$ chance of finding a test statistic with significance of less than 3σ on the same mock data set where the full analysis finds a significance of at least 5σ . Furthermore, even when the signal-to-noise ratio is large enough that 1D Deterministic and 1D Stochastic models would claim a significant detection based on the test statistic, these models would fit the data very poorly, as shown in the right panel of Fig. 4, and such a detection could easily be discarded as spurious.

Thus, using the 1D Stochastic or 1D Deterministic models to analyze a real data set incurs a risk of failing to discover a real axion signal. This occurs because the 1D models do not fully describe the time evolution of the axion signal. In particular, the daily modulation of a real axion signal need not follow the form of $m_z(t)$, as shown in the left panel of Fig. 4. The risk of failing to discover a real axion signal is ameliorated by using the 3D Stochastic model, or the more general model discussed in Sec. III.

V. CONCLUSIONS

Ultralight axion-like particles are well-motivated DM candidates, and their phenomenology in terrestrial direct detection experiments is stochastic. Any analysis of data from an axion DM direct detection experiment must carefully account for the stochasticity of the axion signal in order to properly interpret the results. We have shown that, contrary to previous assumptions in the literature, the signal in experiments sensitive to the gradient of the axion field does not always follow the average DM velocity and can be described as a Gaussian random vector field. We have provided a general framework in which to calculate the two-point correlation functions necessary to fully specify the likelihood for such an axion signal, and we have offered two computationally efficient methods for analyzing a real experimental data set in the time domain, which can be chosen based on the coherence of the axion signal over the lifetime of the exper-

iment. We illustrate that the analysis procedures that have been used in previous literature can fail dramatically on real data. Indeed, failing to properly account for the stochastic nature of the axion gradient signal when developing experimental, data acquisition, and data analysis procedures can potentially remove a real signal from the data or else fail to correctly identify a signal from background noise. As a resource to the interested reader, we have provided an open-source software package at <https://github.com/mtmoschella/axionpy> with some features for generating mock axion gradient signals and performing the analysis techniques described in this paper.

ACKNOWLEDGMENTS

We thank I. Bloch, K. Blum, J. Foster, Y. Hochberg, L. Hui, E. Kuffik, J. Lee, M. Romalis, B. Safdi, and T. Volansky for useful conversations. ML is supported by the DOE under Award Number DESC0007968. MM is supported by the Simons Foundation. WT is supported by the Simons Foundation and the Princeton University Dicke Fellowship. The work presented in this paper was performed on computational resources managed and supported by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of Information Technology's High Performance Computing Center and Visualization Laboratory at Princeton University. This research made use of the `numpy` [64], `scipy` [65], `matplotlib` [66], and `astropy` [67, 68] software packages.

APPENDIX

Appendix A: The Central Limit Theorem

Statement:

Let $\{\mathbf{x}_n | n = 1, \dots, N\}$ be a set of N independent and identically distributed random samples from the k -dimensional probability density function (PDF) $\mathcal{P}(\mathbf{x})$. Then, in the limit $N \rightarrow \infty$, the vector sum

$$\mathbf{S} = \frac{1}{\sqrt{N}} \sum_{n=1}^N \mathbf{x}_n \sim \mathcal{N}(\sqrt{N}\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (\text{A1})$$

is normally distributed with mean $\sqrt{N}\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, where

$$\mu_i = \mathbb{E}[x_i] = \int d^k x x_i \mathcal{P}(\mathbf{x}) \quad (\text{A2})$$

and

$$\begin{aligned} \Sigma_{ij} &= \mathbb{E}[(x_i - \mu_i)(x_j - \mu_j)] \\ &= \int d^k x (x_i - \mu_i)(x_j - \mu_j) \mathcal{P}(\mathbf{x}). \end{aligned} \quad (\text{A3})$$

Proof:

Consider the random variable

$$\begin{aligned} \tilde{\mathbf{S}} &\equiv \mathbf{S} - \sqrt{N}\boldsymbol{\mu} \\ &= \frac{1}{\sqrt{N}} \sum_{n=1}^N (\mathbf{x}_n - \boldsymbol{\mu}), \end{aligned} \quad (\text{A4})$$

which is equivalent to \mathbf{S} , but shifted by a constant. The characteristic function of $\tilde{\mathbf{S}}$, $\Phi_{\tilde{\mathbf{S}}}(\mathbf{t})$, which is defined in terms of the k -dimensional Fourier parameter \mathbf{t} , is

$$\begin{aligned} \Phi_{\tilde{\mathbf{S}}}(\mathbf{t}) &\equiv \mathbb{E} \left[\exp(i\mathbf{t} \cdot \tilde{\mathbf{S}}) \right] \\ &= \left(\mathbb{E} \left[\exp \left(\frac{i\mathbf{t} \cdot (\mathbf{x} - \boldsymbol{\mu})}{\sqrt{N}} \right) \right] \right)^N, \end{aligned} \quad (\text{A5})$$

where we have used Eq. (A4) and the fact that the \mathbf{x}_n random variables are independent and identically distributed. Now we expand the exponential in a multivariate Taylor series to leading order in large N , i.e., the small parameter $1/\sqrt{N}$, giving

$$\begin{aligned} \Phi_{\tilde{\mathbf{S}}}(\mathbf{t}) &= \left(1 - \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k \frac{\mathbb{E}[(x_i - \mu_i)(x_j - \mu_j)] t_i t_j}{N} \right)^N \\ &= \left(1 - \frac{1}{2N} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t} \right)^N. \end{aligned} \quad (\text{A6})$$

Note that the first-order term in the Taylor series vanishes due to the fact that $\mu_i = \mathbb{E}[x_i]$. As $N \rightarrow \infty$, Eq. (A6) becomes

$$\Phi_{\tilde{\mathbf{S}}}(\mathbf{t}) = \exp \left[-\frac{1}{2} \mathbf{t}^\top \boldsymbol{\Sigma} \mathbf{t} \right]. \quad (\text{A7})$$

Computing the Fourier inversion, we have the PDF for $\tilde{\mathbf{S}}$,

$$\begin{aligned} f_{\tilde{\mathbf{S}}}(\mathbf{x}) &= \frac{1}{(2\pi)^k} \int d^k t e^{-i\mathbf{t} \cdot \mathbf{x}} \Phi_{\tilde{\mathbf{S}}}(\mathbf{t}) \\ &= \frac{1}{\sqrt{(2\pi)^k \det \boldsymbol{\Sigma}}} e^{-\frac{1}{2} \mathbf{x}^\top (\boldsymbol{\Sigma})^{-1} \mathbf{x}}, \end{aligned} \quad (\text{A8})$$

which is the PDF for a multivariate normal distribution with zero mean. That is,

$$\tilde{\mathbf{S}} \sim \mathcal{N}(0, \boldsymbol{\Sigma}). \quad (\text{A9})$$

Switching back from $\tilde{\mathbf{S}}$ to \mathbf{S} by a constant additive transformation, we see that

$$\mathbf{S} \sim \mathcal{N}(\sqrt{N}\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (\text{A10})$$

as required.

Appendix B: Uncertainties from Ordinary Least Squares

In ordinary least squares (OLS), we fit a K -vector of observations \mathbf{y} to a linear function of the form

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}, \quad (\text{B1})$$

where $\boldsymbol{\beta}$ is an M -vector of coefficients, and \mathbf{X} is a $K \times M$ matrix. In particular, \mathbf{X} can be thought of as M different K -vectors \mathbf{x}_m that are determined by the basis functions that are being fit to. For example, suppose we want to fit a time series

$$y(t_k) = A \cos(\omega t_k) + B \sin(\omega t_k), \quad (\text{B2})$$

where ω is known and the coefficients A, B are constant with time. Then, the basis functions are $x_1(t_k) = \cos(\omega t_k)$ and $x_2(t_k) = \sin(\omega t_k)$ and the OLS coefficients are $\beta_1 = A$, $\beta_2 = B$. The *least squares* solution to this problem is the set of parameters $\boldsymbol{\beta}^*$ that minimize the sum of the squares of the residuals,

$$\text{SSR}(\boldsymbol{\beta}) = \sum_{k=1}^K \left(y(t_k) - \sum_m \beta_m x_m(t_k) \right)^2. \quad (\text{B3})$$

The solution $\boldsymbol{\beta}^*$ has the following closed-form expression:

$$\boldsymbol{\beta}^* = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}, \quad (\text{B4})$$

where the $M \times K$ matrix $\mathbf{X}^+ \equiv (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ is the Moore-Penrose pseudoinverse of \mathbf{X} .

Our primary goal is to understand the uncertainty on the parameter estimate $\boldsymbol{\beta}^*$, assuming that \mathbf{y} has an independent identical Gaussian uncertainty δy on each measurement. Assuming δy is small enough to propagate errors in the usual way, we see that the uncertainty on the m^{th} estimated parameter, $\delta \beta_m^*$, is given by

$$(\delta \beta_m^*)^2 = \sum_k (X_{mk}^+ \delta y)^2. \quad (\text{B5})$$

To understand the matrix \mathbf{X}^+ , let us first consider the $M \times M$ matrix $\mathbf{X}^\top \mathbf{X}$, which has elements

$$(\mathbf{X}^\top \mathbf{X})_{m,m'} = \sum_k x_m(t_k) x_{m'}(t_k) \approx K \langle x_m x_{m'} \rangle, \quad (\text{B6})$$

where we have assumed that the sampling rate is rapid enough and the integration time is long enough to average over the entire domain of the basis functions. If the basis functions x_m satisfy the useful orthogonality property

$$\langle x_m x_{m'} \rangle = \langle x_m^2 \rangle \delta_{m,m'}, \quad (\text{B7})$$

then $\mathbf{X}^\top \mathbf{X}$ is approximately diagonal and the inverse matrix $(\mathbf{X}^\top \mathbf{X})^{-1}$ is trivial to compute. The pseudoinverse matrix is therefore

$$X_{mk}^+ = \sum_{m'} (\mathbf{X}^\top \mathbf{X})_{m,m'}^{-1} x_{m'}(t_k) \approx \frac{x_m(t_k)}{K \langle x_m^2 \rangle} \quad (\text{B8})$$

and so the uncertainties on the least-squares parameters are given by

$$(\delta \beta_m^*)^2 \approx \sum_k \frac{x_m^2(t) \delta y^2}{K^2 \langle x_m^2 \rangle^2} \approx \frac{\delta y^2}{K \langle x_m^2 \rangle}. \quad (\text{B9})$$

Solving for $\delta \beta_m^*$ explicitly, we obtain

$$\delta \beta_m^* \approx \frac{\delta y}{\sqrt{K \langle x_m^2 \rangle}}. \quad (\text{B10})$$

Appendix C: Evaluating Maxwellian Integrals

This Appendix demonstrates how to compute integrals of the form in Eq. (15) for the specific case of the velocity distribution function in Eq. (16) and under the assumption $\mathbf{v}_{\text{obs}}(t) \approx \mathbf{v}_\odot$. Changing variables of integration from $\mathbf{w} \rightarrow \mathbf{v} = \mathbf{w} + \mathbf{v}_\odot$, we obtain

$$\begin{aligned} \langle A_i(t) A_j(t') \rangle &= g_{\text{eff}}^2 \rho_a \int d^3 \mathbf{v} v_i v_j \frac{e^{-(\mathbf{v}-\mathbf{v}_\odot)^2/2\sigma_v^2}}{(2\pi\sigma_v^2)^{3/2}} \cos(\Delta\varpi) \\ \langle A_i(t) B_j(t') \rangle &= g_{\text{eff}}^2 \rho_a \int d^3 \mathbf{v} v_i v_j \frac{e^{-(\mathbf{v}-\mathbf{v}_\odot)^2/2\sigma_v^2}}{(2\pi\sigma_v^2)^{3/2}} \sin(\Delta\varpi) \end{aligned} \quad (\text{C1})$$

where $\Delta\varpi = \frac{1}{2} m_a v^2 (t' - t)$ and the indices $i, j = x, y, z$.

Choosing to evaluate the integral in spherical coordinates (v, θ, φ) with polar axis \mathbf{e}_z , the volume element becomes $d^3 \mathbf{v} = v^2 dv d\cos\theta d\varphi$, and the components of \mathbf{v} are expressed as

$$\begin{aligned} v_x &= v \sin\theta \cos\varphi \\ v_y &= v \sin\theta \sin\varphi \\ v_z &= v \cos\theta. \end{aligned} \quad (\text{C2})$$

Additionally, $(\mathbf{v} - \mathbf{v}_\odot)^2 = v^2 + v_\odot^2 - 2v v_\odot \cos\theta$ because \mathbf{e}_z is chosen to be parallel to \mathbf{v}_\odot .

Examining the $d\varphi$ integrals, it is clear that unless $i = j$, the integral vanishes, and that the $i = j = x$ integrals are equivalent to $i = j = y$. This leaves four unique non-zero integrals to compute:

$$\begin{aligned} \langle A_z(t) A_z(t') \rangle &= g_{\text{eff}}^2 \rho_a \int_0^\infty dv v^4 \tilde{f}_\parallel(v) \cos(\Delta\varpi) \\ \langle A_z(t) B_z(t') \rangle &= g_{\text{eff}}^2 \rho_a \int_0^\infty dv v^4 \tilde{f}_\parallel(v) \sin(\Delta\varpi) \\ \langle A_x(t) A_x(t') \rangle &= g_{\text{eff}}^2 \rho_a \int_0^\infty dv v^4 \tilde{f}_\perp(v) \cos(\Delta\varpi) \\ \langle A_x(t) B_x(t') \rangle &= g_{\text{eff}}^2 \rho_a \int_0^\infty dv v^4 \tilde{f}_\perp(v) \sin(\Delta\varpi), \end{aligned} \quad (\text{C3})$$

where we have ignored the escape velocity in order to carry out the dv integrals from 0 to ∞ . The $d\varphi$ and

d cos θ integrals are encapsulated in the effective speed distributions:

$$\begin{aligned}\tilde{f}_{\parallel}(v) &\equiv \int_{-1}^1 d \cos \theta \int_0^{2\pi} d\varphi \cos^2 \theta \frac{e^{-(v^2+v_{\odot}^2-2vv_{\odot} \cos \theta)/2\sigma_v^2}}{(2\pi\sigma_v^2)^{3/2}} \\ \tilde{f}_{\perp}(v) &\equiv \int_{-1}^1 d \cos \theta \int_0^{2\pi} d\varphi \sin^2 \theta \cos^2 \varphi \\ &\quad \times \frac{e^{-(v^2+v_{\odot}^2-2vv_{\odot} \cos \theta)/2\sigma_v^2}}{(2\pi\sigma_v^2)^{3/2}}\end{aligned}\quad (\text{C4})$$

Evaluating these integrals, we obtain the effective speed distributions

$$\begin{aligned}\tilde{f}_{\parallel}(v) &\equiv \frac{\sqrt{2}e^{-(v^2+v_{\odot}^2)/2\sigma_v^2}}{\sqrt{\pi}\sigma_v^3} \frac{(2+\eta^2) \sinh \eta - 2\eta \cosh \eta}{\eta^3} \\ \tilde{f}_{\perp}(v) &\equiv \frac{\sqrt{2}e^{-(v^2+v_{\odot}^2)/2\sigma_v^2}}{\sqrt{\pi}\sigma_v^3} \frac{\eta \cosh \eta - \sinh \eta}{\eta^3},\end{aligned}\quad (\text{C5})$$

where $\eta \equiv vv_{\odot}/\sigma_v^2$.

Using the results of Eq. (C5) and rewriting the trigonometric functions in Eq. (C3) as complex exponentials, we obtain

$$\begin{aligned}\langle A_z(t)A_z(t') \rangle &= \text{Re} \{ \mathcal{Z}_{\parallel} \} \\ \langle A_z(t)B_z(t') \rangle &= \text{Im} \{ \mathcal{Z}_{\parallel} \} \\ \langle A_x(t)A_x(t') \rangle &= \text{Re} \{ \mathcal{Z}_{\perp} \} \\ \langle A_x(t)B_x(t') \rangle &= \text{Im} \{ \mathcal{Z}_{\perp} \},\end{aligned}\quad (\text{C6})$$

where we have defined the following complex variables

$$\begin{aligned}\mathcal{Z}_{\parallel} &\equiv \sqrt{\frac{2}{\pi}} \frac{g_{\text{eff}}^2 \rho_a \sigma_v^3}{v_{\odot}^3} e^{-v_{\odot}^2/2\sigma_v^2} \left(2\mathcal{I}_1 - 2\frac{v_{\odot}}{\sigma_v^2} \mathcal{I}_2 + \frac{v_{\odot}^2}{\sigma_v^4} \mathcal{I}_3 \right) \\ \mathcal{Z}_{\perp} &\equiv \sqrt{\frac{2}{\pi}} \frac{g_{\text{eff}}^2 \rho_a \sigma_v^3}{v_{\odot}^3} e^{-v_{\odot}^2/2\sigma_v^2} \left(\frac{v_{\odot}}{\sigma_v^2} \mathcal{I}_2 - \mathcal{I}_1 \right).\end{aligned}\quad (\text{C7})$$

The complex Gaussian integrals are defined as:

$$\begin{aligned}\mathcal{I}_1 &= \int_0^{\infty} dv v \sinh(vv_{\odot}/\sigma_v^2) \exp \left[-\frac{v^2}{2\sigma_v^2} \zeta \right] \\ \mathcal{I}_2 &= \int_0^{\infty} dv v^2 \cosh(vv_{\odot}/\sigma_v^2) \exp \left[-\frac{v^2}{2\sigma_v^2} \zeta \right] \\ \mathcal{I}_3 &= \int_0^{\infty} dv v^3 \sinh(vv_{\odot}/\sigma_v^2) \exp \left[-\frac{v^2}{2\sigma_v^2} \zeta \right],\end{aligned}\quad (\text{C8})$$

with $\zeta \equiv 1 - i\xi$ and $\xi \equiv m_a \sigma_v^2 (t' - t)$. Evaluating the

Gaussian integrals, we find that

$$\begin{aligned}\mathcal{I}_1 &= \sqrt{\frac{\pi}{2}} \sigma_v v_{\odot} \zeta^{-3/2} e^{v_{\odot}^2/2\sigma_v^2 \zeta} \\ \mathcal{I}_2 &= \sqrt{\frac{\pi}{2}} \sigma_v v_{\odot}^2 \left(\zeta^{-5/2} + \frac{\sigma_v^2}{v_{\odot}^2} \zeta^{-3/2} \right) e^{v_{\odot}^2/2\sigma_v^2 \zeta} \\ \mathcal{I}_3 &= \sqrt{\frac{\pi}{2}} \sigma_v v_{\odot}^3 \left(\zeta^{-7/2} + 3 \frac{\sigma_v^2}{v_{\odot}^2} \zeta^{-5/2} \right) e^{v_{\odot}^2/2\sigma_v^2 \zeta}\end{aligned}\quad (\text{C9})$$

and therefore

$$\begin{aligned}\mathcal{Z}_{\parallel} &= g_{\text{eff}}^2 \rho_a \frac{\zeta \sigma_v^2 + v_{\odot}^2}{\zeta^{7/2}} \exp \left[-\frac{v_{\odot}^2}{2\sigma_v^2} \left(1 - \frac{1}{\zeta} \right) \right] \\ \mathcal{Z}_{\perp} &= g_{\text{eff}}^2 \rho_a \frac{\sigma_v^2}{\zeta^{5/2}} \exp \left[-\frac{v_{\odot}^2}{2\sigma_v^2} \left(1 - \frac{1}{\zeta} \right) \right].\end{aligned}\quad (\text{C10})$$

Writing the complex variables as a single amplitude and phase, we have

$$\begin{aligned}\mathcal{Z}_{\parallel} &= \mathcal{A}_{\parallel} e^{i\Psi_{\parallel}} \\ \mathcal{Z}_{\perp} &= \mathcal{A}_{\perp} e^{i\Psi_{\perp}},\end{aligned}\quad (\text{C11})$$

where the amplitudes are given by

$$\begin{aligned}\mathcal{A}_{\parallel} &= g_{\text{eff}}^2 \rho_a \frac{\sqrt{(v_{\odot}^2 + \sigma_v^2)^2 + \sigma_v^4 \xi^2}}{(1 + \xi^2)^{7/4}} \exp \left[-\frac{v_{\odot}^2}{2\sigma_v^2} \frac{\xi^2}{1 + \xi^2} \right] \\ \mathcal{A}_{\perp} &= g_{\text{eff}}^2 \rho_a \frac{\sigma_v^2}{(1 + \xi^2)^{5/4}} \exp \left[-\frac{v_{\odot}^2}{2\sigma_v^2} \frac{\xi^2}{1 + \xi^2} \right],\end{aligned}\quad (\text{C12})$$

and the phases are given by

$$\begin{aligned}\Psi_{\parallel} &= \frac{v_{\odot}^2}{2\sigma_v^2} \frac{\xi}{1 + \xi^2} + \frac{7}{2} \arctan \xi - \arctan \left(\frac{\xi \sigma_v^2}{\sigma_v^2 + v_{\odot}^2} \right) \\ \Psi_{\perp} &= \frac{v_{\odot}^2}{2\sigma_v^2} \frac{\xi}{1 + \xi^2} + \frac{5}{2} \arctan \xi.\end{aligned}\quad (\text{C13})$$

Taking the real and imaginary parts according to Eq. (C6), we obtain

$$\begin{aligned}\langle A_z(t)A_z(t') \rangle &= \mathcal{A}_{\parallel}(\xi) \cos \Psi_{\parallel}(\xi) \\ \langle A_z(t)B_z(t') \rangle &= \mathcal{A}_{\parallel}(\xi) \sin \Psi_{\parallel}(\xi) \\ \langle A_x(t)A_x(t') \rangle &= \mathcal{A}_{\perp}(\xi) \cos \Psi_{\perp}(\xi) \\ \langle A_x(t)B_x(t') \rangle &= \mathcal{A}_{\perp}(\xi) \sin \Psi_{\perp}(\xi),\end{aligned}\quad (\text{C14})$$

consistent with Eq. (17).

- [1] A. Derevianko, *Phys. Rev. A* **97**, 042506 (2018), 1605.09717.
- [2] J. W. Foster, N. L. Rodd, and B. R. Safdi, *Phys. Rev. D* **97**, 123006 (2018), 1711.10489.
- [3] R. D. Peccei and H. R. Quinn, *Phys. Rev. Lett.* **38**, 1440 (1977).
- [4] R. D. Peccei and H. R. Quinn, *Phys. Rev. D* **16**, 1791 (1977).
- [5] S. Weinberg, *Phys. Rev. Lett.* **40**, 223 (1978).
- [6] F. Wilczek, *Phys. Rev. Lett.* **40**, 279 (1978).
- [7] J. E. Kim, *Phys. Rev. Lett.* **43**, 103 (1979).
- [8] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Nucl. Phys. B* **166**, 493 (1980).
- [9] A. R. Zhitnitsky, *Sov. J. Nucl. Phys.* **31**, 260 (1980).
- [10] M. Dine, W. Fischler, and M. Srednicki, *Phys. Lett. B* **104**, 199 (1981).
- [11] J. Preskill, M. B. Wise, and F. Wilczek, *Phys. Lett. B* **120**, 127 (1983).
- [12] L. F. Abbott and P. Sikivie, *Phys. Lett. B* **120**, 133 (1983).
- [13] M. Dine and W. Fischler, *Phys. Lett. B* **120**, 137 (1983).
- [14] A. Arvanitaki, S. Dimopoulos, S. Dubovsky, N. Kaloper, and J. March-Russell, *Phys. Rev. D* **81**, 123530 (2010), 0905.4720.
- [15] P. W. Graham and S. Rajendran, *Phys. Rev. D* **88**, 035023 (2013), 1306.6088.
- [16] Y. V. Stadnik and V. V. Flambaum, *Phys. Rev. D* **89**, 043522 (2014), 1312.6667.
- [17] Y. V. Stadnik, Ph.D. thesis, University of New South Wales (2017).
- [18] P. W. Graham, D. E. Kaplan, J. Mardon, S. Rajendran, W. A. Terrano, L. Trahms, and T. Wilkason, *Phys. Rev. D* **97**, 055006 (2018), 1709.07852.
- [19] D. Aybas, H. Bekker, J. W. Blanchard, D. Budker, G. P. Centers, N. L. Figueroa, A. V. Gramolin, D. F. J. Kimball, A. Wickenbrock, and A. O. Sushkov, *Quantum Sci. Technol.* **6**, 034007 (2021), 2103.06284.
- [20] C. Abel et al., *Phys. Rev. X* **7**, 041034 (2017), 1708.06367.
- [21] C. Smorra et al., *Nature* **575**, 310 (2019), 2006.00255.
- [22] A. Garcon et al., *Sci. Adv.* **5**, eaax4539 (2019), 1902.04644.
- [23] T. Wu et al., *Phys. Rev. Lett.* **122**, 191302 (2019), 1901.10843.
- [24] W. A. Terrano, E. G. Adelberger, C. A. Hagedorn, and B. R. Heckel, *Phys. Rev. Lett.* **122**, 231301 (2019), 1902.04246.
- [25] T. W. Kornack, Ph.D. thesis, Princeton University (2005).
- [26] G. Vasilakis, Ph.D. thesis, Princeton University (2011).
- [27] J. M. Brown, Ph.D. thesis, Princeton University (2011).
- [28] I. M. Bloch, Y. Hochberg, E. Kuflik, and T. Volansky, *JHEP* **01**, 167 (2020), 1907.03767.
- [29] I. M. Bloch, G. Ronen, R. Shaham, O. Katz, T. Volansky, and O. Katz (2021), 2105.04603.
- [30] E. G. Adelberger, B. R. Heckel, S. A. Hoedl, C. D. Hoyle, D. J. Kapner, and A. Upadhye, *Phys. Rev. Lett.* **98**, 131104 (2007), hep-ph/0611223.
- [31] W. A. Terrano, E. G. Adelberger, J. G. Lee, and B. R. Heckel, *Phys. Rev. Lett.* **115**, 201801 (2015), 1508.02463.
- [32] R. Essig, R. Harnik, J. Kaplan, and N. Toro, *Phys. Rev. D* **82**, 113008 (2010), 1008.0636.
- [33] D. S. Akerib et al. (LUX), *Phys. Rev. Lett.* **118**, 261301 (2017), 1704.02297.
- [34] M. V. Beznogov, E. Rrapaj, D. Page, and S. Reddy, *Phys. Rev. C* **98**, 035802 (2018), 1806.07991.
- [35] P. Carena, T. Fischer, M. Giannotti, G. Guo, G. Martínez-Pinedo, and A. Mirizzi, *JCAP* **10**, 016 (2019), [Erratum: *JCAP* 05, E01 (2020)], 1906.11844.
- [36] M. M. Miller Bertolami, B. E. Melendez, L. G. Althaus, and J. Isern, *JCAP* **10**, 069 (2014), 1406.7712.
- [37] N. Bar, K. Blum, and G. D'Amico, *Phys. Rev. D* **101**, 123025 (2020), 1907.05020.
- [38] S. J. Asztalos et al. (ADMX), *Phys. Rev. D* **64**, 092003 (2001).
- [39] N. Du et al. (ADMX), *Phys. Rev. Lett.* **120**, 151301 (2018), 1804.05750.
- [40] T. Braine et al. (ADMX), *Phys. Rev. Lett.* **124**, 101303 (2020), 1910.08638.
- [41] B. M. Brubaker et al., *Phys. Rev. Lett.* **118**, 061302 (2017), 1610.02580.
- [42] L. Zhong et al. (HAYSTAC), *Phys. Rev. D* **97**, 092001 (2018), 1803.03690.
- [43] Y. Kahn, B. R. Safdi, and J. Thaler, *Phys. Rev. Lett.* **117**, 141801 (2016), 1602.01086.
- [44] J. L. Ouellet et al., *Phys. Rev. Lett.* **122**, 121802 (2019), 1810.12257.
- [45] C. P. Salemi et al. (2021), 2102.06722.
- [46] A. V. Gramolin, D. Aybas, D. Johnson, J. Adam, and A. O. Sushkov, *Nature Phys.* **17**, 79 (2021), 2003.03348.
- [47] K. Ehret et al., *Phys. Lett. B* **689**, 149 (2010), 1004.1313.
- [48] M. Betz, F. Caspers, M. Gasiior, M. Thumm, and S. W. Rieger, *Phys. Rev. D* **88**, 075014 (2013), 1310.8098.
- [49] R. Ballou et al. (OSQAR), *Phys. Rev. D* **92**, 092002 (2015), 1506.08082.
- [50] V. Anastassopoulos et al. (CAST), *Nature Phys.* **13**, 584 (2017), 1705.02290.
- [51] A. Abramowski et al. (H.E.S.S.), *Phys. Rev. D* **88**, 102003 (2013), 1311.3148.
- [52] A. Ayala, I. Domínguez, M. Giannotti, A. Mirizzi, and O. Straniero, *Phys. Rev. Lett.* **113**, 191302 (2014), 1406.6053.
- [53] A. Payez, C. Evoli, T. Fischer, M. Giannotti, A. Mirizzi, and A. Ringwald, *JCAP* **02**, 006 (2015), 1410.3747.
- [54] M. Ajello et al. (Fermi-LAT), *Phys. Rev. Lett.* **116**, 161101 (2016), 1603.06978.
- [55] C. S. Reynolds, M. C. D. Marsh, H. R. Russell, A. C. Fabian, R. Smith, F. Tombesi, and S. Veilleux (2019), 1907.05475.
- [56] L. Hui, A. Joyce, M. J. Landry, and X. Li, *JCAP* **01**, 011 (2021), 2004.01188.
- [57] G. P. Centers et al. (2019), 1905.13650.
- [58] S. E. Kposov, H.-W. Rix, and D. W. Hogg, *The Astrophysical Journal* **712**, 260 (2010).
- [59] J. Bovy, C. Allende Prieto, T. C. Beers, D. Bizyaev, L. N. da Costa, K. Cunha, G. L. Ebelke, D. J. Eisenstein, P. M. Frinchaboy, A. E. García Pérez, et al., *The Astrophysical Journal* **759**, 131 (2012).
- [60] A.-C. Eilers, D. W. Hogg, H.-W. Rix, and M. K. Ness, *The Astrophysical Journal* **871**, 120 (2019).
- [61] R. Schönrich, J. Binney, and W. Dehnen, *Monthly Notices of the Royal Astronomical Society* **403**, 1829 (2010).

- [62] W. A. Rolke, A. M. Lopez, and J. Conrad, Nucl. Instrum. Meth. A **551**, 493 (2005), physics/0403059.
- [63] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011), [Erratum: Eur.Phys.J.C 73, 2501 (2013)], 1007.1727.
- [64] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, et al., Nature **585**, 357 (2020).
- [65] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al., Nature Methods **17**, 261 (2020).
- [66] J. D. Hunter, Computing in Science Engineering **9**, 90 (2007).
- [67] Astropy Collaboration, T. P. Robitaille, E. J. Tollerud, P. Greenfield, M. Droettboom, E. Bray, T. Aldcroft, M. Davis, A. Ginsburg, A. M. Price-Whelan, et al., Astronomy & Astrophysics **558**, A33 (2013), 1307.6212.
- [68] Astropy Collaboration, A. M. Price-Whelan, B. M. Sipőcz, H. M. Günther, P. L. Lim, S. M. Crawford, S. Conseil, D. L. Shupe, M. W. Craig, N. Dencheva, et al., The Astronomical Journal **156**, 123 (2018), 1801.02634.