Overcoming gamma ray constraints with annihilating dark matter in Milky Way subhalos

Aaron C. Vincent,1,* Wei Xue,1,† and James M. Cline1,2,‡

1Department of Physics, McGill University, 3600 Rue University, Montréal, Québec, Canada H3A 2T8
2Perimeter Institute for Theoretical Physics, 31 Caroline St N, Waterloo, Ontario N2L 2Y5, Canada

(Received 12 October 2010; published 17 December 2010)

We reconsider Sommerfeld-enhanced annihilation of dark matter (DM) into leptons to explain PAMELA and Fermi electron and positron observations, in light of possible new effects from substructure. There is strong tension between getting a large enough lepton signal while respecting constraints on the fluxes of associated gamma rays; we show how DM annihilations within subhalos can get around these constraints. Specifically, if most of the observed lepton excess comes from annihilations in a nearby (within 2 kpc) subhalo along a line of sight toward the galactic center, it is possible to match both the lepton and gamma ray observations. We demonstrate that this can be achieved in a simple class of particle physics models in which the DM annihilates via a hidden leptophilic U(1) vector boson, with explicitly computed Sommerfeld enhancement factors. Gamma ray constraints on the main halo annihilations (and cosmic microwave background constraints from the era of decoupling) require the annihilating component of the DM to be subdominant, of order $10^{-2}$–$10^{-3}$ of the total DM density.

DOI: 10.1103/PhysRevD.82.123519 PACS numbers: 95.35.+d

I. INTRODUCTION

Nongravitational signals of dark matter (DM) have been sought after for some time now by the astrophysical and particle physics communities. At the same time results from the Payload for AntiMatter Matter Exploration and Light-nuclei Astrophysics (PAMELA) experiment and from the Fermi space telescope suggest a local excess positron fraction $e^+/(e^++e^-)$ at energies above 10 GeV as well as an excess of $e^++e^-$ peaking around 500 GeV. Standard cosmic ray propagation models do not account for these excesses. An attractive explanation is that a DM WIMP (weakly interacting massive particle) is present in our galaxy at large enough concentrations to self-annihilate into standard model leptons. A TeV-scale WIMP annihilating to electron-positron pairs could produce such signals. In order to be consistent with the observed relic abundance of DM, the annihilation cross section $(\sigma v)_0 \sim 3 \times 10^{-26}$ cm$^3$ s$^{-1}$ would have to be enhanced by a factor of order 100, for example, by a velocity-dependent Sommerfeld enhancement.

Many authors [1–10] have explored this possibility, and have constrained the allowable mass versus boost factor parameter space. However these papers assume that the dominant source of indirect signals is from annihilations in the main DM halo. In a previous work [11] we considered the possibility of adding the effects of dark matter substructure to the theoretical model and we found examples where annihilations in subhalos could provide a significant fraction of the observed lepton excesses. We showed that one could find a better overall fit to the electron-positron data from the Fermi and PAMELA experiments, and we suggested that gamma ray constraints which are now putting considerable pressure on these models could be alleviated. Our purpose in the present work was to ascertain whether this is indeed the case.

The constraints mentioned come from recent gamma ray observations of the galaxy and from cosmic microwave background (CMB) measurements. As high energy electron-positron pairs are produced and diffuse throughout the galaxy, they will emit final-state radiation as well as scatter on the ambient photon field, giving rise to $\sim 1$–100 GeV gamma rays that should be detectable. Given the large expected concentrations of both DM and radiation near the galactic center (GC), gamma rays from inverse Compton scattering (ICS) near the GC are particularly constraining. The Fermi Large Area Telescope (LAT) is specifically designed to detect gamma rays in this range, and its latest results have been used to rule out large regions of parameter space for annihilating WIMP models [5,12–14].

However in this work we will show that if a sizeable proportion of the leptons from DM annihilation originates from nearby subhalos, the constraints from GC gamma rays can be relieved. Final-state (bremsstrahlung) radiation from subhalos has been examined by other authors [8,15–20], and Ref. [21] has studied the $e^++e^-$ spectrum from a nearby subhalo. In this follow-up work we extend our previous findings to a prediction of the gamma ray spectrum including a full calculation of ICS radiation in the galaxy, which we compare to the full-sky data from the Fermi LAT. We include the expected contribution to the gamma ray background coming from background electrons and positrons. Using a fully-numerical approach, we find that there is less room for new contributions from the annihilation products of the DM, making the
constraints on the DM models more severe. This is a serious issue even for less cuspy and cored DM profiles, that have been shown to satisfy the constraints in previous semianalytic treatments which ignored the background gamma ray fluxes.

In our previous work we focused on the contributions of distant subhalos to the flux of leptons at Earth. Even though these new contributions can improve the fit to the lepton data alone, here we show that they do not soften the gamma ray constraints sufficiently to be viable. Instead, we focus on the possibility that an accidentally nearby subhalo could provide the bulk of the lepton flux. The associated gamma rays would be sufficiently hidden by strong backgrounds if this subhalo happened to lie between us and the galactic center. The effects of nearby subhalos have been previously considered by Ref. [22], but only allowing for purely astrophysical boost factors, due to the density of the subhalos. Here we find that velocity-dependent Sommerfeld enhancement is crucial for obtaining a positive outcome. It is precisely because of the larger boost factor available within subhalos (which have orders of magnitude smaller velocity dispersion) relative to the main halo that we are able to soften the gamma ray constraint due to the main halo near the GC, yet have a large enough lepton signal from a nearby subhalo. In addition, we must assume that the leptophilic component of the DM responsible for these processes is subdominant to the main inert (for our purposes) component, in order to sufficiently reduce the effective boost factor for annihilations in the main halo [14]. This gives rise to the interesting possibility that different kinds of DM are responsible for the cosmic ray anomalies than those which might manifest themselves in direct detection experiments.

Using a modified version of the cosmic ray propagation code GALPROP and the data from the recent Via Lactea II simulation of dark matter evolution and collapse in a Milky Way-sized galaxy, we modeled the two-dimensional axisymmetric distribution of electrons and positrons in the galaxy. These results were combined with simulated interstellar radiation field (ISRF) data in order to compute a realistic skymap of the gamma ray spectrum expected from DM annihilation in the Galaxy, which was in turn compared with a year’s worth of diffuse gamma ray observation from the Fermi LAT.

We start with a summary the cosmic ray model and results of our previous work in Sec. II, before discussing the relevant ICS and gamma ray physics in Sec. III. In Sec. IV we describe our methodology, and present model-independent fits to the data in several scenarios for the distribution of subhalos and the halo profiles. In particular, we show that an accidentally nearby subhalo can provide a promising loophole to the gamma ray constraints on cuspy profiles. We also predict the gamma ray flux from the subhalo, which could provide a test of the model if future measurements and understanding of backgrounds are improved. In Sec. V we then demonstrate that the boost factors required for this scenario can be explicitly realized in a simple class of hidden sector particle physics models. We conclude with a discussion of the overall viability of this picture in Sec. VI.

II. COSMIC RAY PROPAGATION

Inside the galactic diffusion zone, particles and nuclei propagate according to the diffusion-loss equation [23], which applies to electrons and positrons as follows:

$$\frac{d}{dt} \psi_e(x, p, t) = Q_e(x, E) + \nabla \cdot (D(E) \nabla \psi_e(x, p, t)) + \frac{\partial}{\partial E}[b(x, E) \psi_e(x, p, t)].$$

(1)

Here $\psi_e(x, p, t)$ denotes the particle number density per unit momentum $|p|$, $Q$ represents the source function, $D(E)$ is the spatial diffusion coefficient and $b(x, E)$ is the energy loss coefficient. We seek the steady-state solution of Eq. (1): $d\psi_e(x, p, t)/dt = 0$.

Since (1) is linear, the leptons from DM annihilation travel independently in the astrophysical background. The source $Q_{e^-}$ comes from DM annihilation which depends on the particle physics and the local density of the dark matter:

$$Q_{e^-} = \frac{1}{2} \left( \frac{\rho(x)}{M} \right)^2 (\sigma v) \frac{dN_{e^-}}{dE} = \frac{n_{DM}^2}{2} BF(\sigma v)_0 \frac{dN_{e^-}}{dE},$$

(2)

where the prefactor 1/2 is a symmetry factor for self-annihilation, $n_{DM}(x, E)$ is the DM energy density, $\langle \sigma v \rangle_0 = 3 \times 10^{-26} \text{ cm}^3 \text{s}^{-1}$ is the benchmark value for standard cosmology to explain the relic density of DM, and $dN_{e^-}/dE$ is the energy spectrum of the annihilation products. Neglecting the effect of soft photons, the spectrum can be approximated by the simple form $dN_{e^-}/dE = 2M_{DM}^{-1} \Theta(M_{DM} - E)$, where $\Theta(x)$ is the usual Heaviside step function, and the factor 2 arises because that the final state has two electrons or two positrons. The latter has the correct qualitative shape, and is easier to implement in GALPROP than would be a more exact spectrum. $BF$ denotes the boost factor due to Sommerfeld enhancement, originating from a nonperturbative $\sim 1/\nu$ correction due to the slow ($v/c < \alpha$) motion of the DM particles.

To simplify our analysis, we take the boost factor $BF$ to be constant throughout the main halo, and tune it to provide the best possible fit to available electron and positron data. Since the Sommerfeld effect depends strongly on velocity, typical subhalos, which have a much smaller velocity
dispersion, have a much higher $BF$, and we treat it as an additional free parameter. Although each subhalo has different values of $BF$, we represent the subhalo $BF$ by a single average value in this first part of our analysis, where the $BF$s are treated as being uncorrelated and best-fit values are sought. This is not a limitation in the case we will eventually focus upon, namely, domination of the excess lepton signal by a single nearby subhalo. A further complication is that in fact $BF$ has a radial dependence within each halo, because the velocity dispersion is a function of $r$, which has been fitted by many-body simulations such as Via Lactea II [19]. We will take this into account in Sec. VA by averaging $BF$ over the phase space of DM in the halos, in order to make contact with the results obtained in this model-independent part of our analysis.

The spatial diffusion coefficient can be parametrized as follows [24]:

$$D(E) = D_0 \left( \frac{E}{4\text{ GeV}} \right)^\delta.$$  (3)

Two widely used approaches exist for solving the diffusion equation in the Galaxy: semianalytic and fully numerical. We chose the latter for Galaxy-scale propagation, in part because a numerical approach allows for better control over the spatial dependence of the astrophysical input, such as energy loss due to inverse Compton scattering. GALPROP 50.1p [25] is a publicly available software package that solves Eq. (1) with an implicit-in-time 2D or 3D Crank-Nicholson scheme. In 2D mode, it provides a $(r, z)$ map in cylindrical coordinates of the number density of each species within the Galactic diffusion zone. To constrain the diffusion parameters, the ratio of measured secondary-to-primary species such as B/C or sub-Fe/Fe can be simulated and fit to observations. This was done to a very high degree of accuracy in Ref. [24]. We used results from their best fits: $D_0 = 6.04 \times 10^{28} \text{ cm}^2 \text{ s}^{-1}(0.19 \text{ kpc}^2/\text{Myr})$, and $\delta = 0.41$.

The full energy loss rate is due to synchrotron radiation and inverse Compton scattering:

$$b(x, E) = -\frac{dE_x}{dt} = \frac{32\pi c_\alpha_{em} E^2}{3m^4_e} \left[ u_B + \sum_{i=1}^{3} u_{yi} \cdot R_i(E_x) \right].$$  (4)

$\alpha_{em}$ is the fine structure constant and $u_B = B^2/2$ is the energy density of the galactic magnetic field, for which we used the standard parametrization:

$$B(r, z) \approx 11 \mu G \cdot \exp \left( -\frac{r}{10 \text{ kpc}} - \frac{|z|}{2 \text{ kpc}} \right).$$  (5)

$u_{yi}$ are the energy densities of the three main components of the interstellar radiation field (ISRF): CMB radiation, thermal radiation from dust and starlight, which lie mainly in the microwave, infrared and optical regions of the electromagnetic spectrum, respectively. GALPROP uses position-dependent maps of ISRF compiled by [26], rather than using a constant energy-loss coefficient computed from a local average. The latter approach (explained in Sec. 3 of [27]) is commonly used in the semianalytic model. While it is indeed quite accurate when dealing with electrons from a smooth Galaxy-wide distribution of dark matter, it is an approximation that is less precise when considering the propagation into the Galaxy of electrons from DM subhalos outside of the diffusion zone. We will nonetheless make use of the semianalytic method in Sec. IV B, when only local propagation will be relevant. The position dependence of the ISRF in the Galaxy is presented in Fig. 1. Further details will be discussed in Sec. III B.
A. Via Lactea II and GALPROP

We assumed that the DM was composed of a single Dirac fermion $\chi$ of mass $M_{DM}$ annihilating through the channel $\chi\chi \rightarrow BB$, followed by the decay $B \rightarrow e^+ e^-$. where $B$ is some dark sector gauge boson which could also be responsible for the Sommerfeld enhancement. We considered two astrophysical models for the DM distribution: a main halo-only (MH) scenario, in which only a large, spherical halo contributed annihilation products; and a subhalo (MH + SH) scenario, where the overdensities formed by DM substructure were responsible for extra annihilation of DM into electrons and positrons. In both cases, we used a spherically symmetric Einasto profile for the DM density distribution:

$$\rho_{Ein}(r) = \rho_s \exp\left[\frac{-2}{\alpha}\left(\left(\frac{r}{r_s}\right)^{1/\alpha} - 1\right)\right]$$

where $r$ is the radial coordinate from the center of the halo, $\rho_s$ is the density at $r = r_s$, the distance at which the slope $\rho'/\rho = -2$. These parameters are simply related to the radius and rotational velocity of a given subhalo as explained in Ref. [19]. The shape parameter $\alpha$ can be read off from curve-fitting the distributions from $N$-body simulations such as [28,29]. It is generally taken to be around $\alpha \approx 0.17$. We took $r_s = 25$ kpc for the main galactic halo, with a local dark matter density $\rho_0 = 0.37 \text{ GeV cm}^{-3}$ in agreement with Via Lactea II and with other recent estimates, e.g., [30]. It should be noted that many authors use the convention $\rho_0 = 0.3 \text{ GeV cm}^{-3}$. This leads to a factor of $(0.3/0.37)^2 = 0.66$ difference in the constraints on the annihilation cross sections, but it is of no consequence when it comes to excluding models, since constraints come from the ratio of gamma rays lepton fluxes, which both scale linearly with $\rho_0^2(\sigma v)$.

It has been argued that direct observations of rotation velocities in the Milky Way are consistent with cored DM profiles (see, for example, Ref. [31]). Two such examples are the isothermal and Burkert [32] ansätze. The Burkert profile has been fitted to the rotation curves of galaxies other than our own, but we are not aware of references which attempt to fit the Milky Way. To allow for the alternative possibility of a cored main halo, we will therefore restrict our attention to the isothermal profile

$$\rho_{iso}(r) = \frac{\rho_s}{1 + (r/r_s)^2}$$

adopting the values $r_s = 3.2$ kpc and $\rho_s = 3.0 \text{ GeV cm}^{-3}$ similar to those used by Ref. [12]. These values are motivated by the constraint on the observed solar density $\rho_0$ (which we take to be somewhat higher than in [12]) and on the mass of the Galaxy with 50 kpc as determined from circular velocity measurements. However for the subhalos we will in all cases assume the Einasto form that is suggested by Via Lactea II.
Electrons from an extragalactic source have a very particular density profile. While the annihilation products from the main halo follow a roughly symmetric distribution about the GC, SH electrons sourced from the diffusion zone boundary tend to form a diffuse “shell” near the edge of the diffusion zone, as illustrated in Fig. 2. Ambient radiation prevents high-energy particles from reaching the GC, trapping them near the edge of the Galaxy. The large number of subhalos combined with a large boost factor can allow some particles to make their way to earth, albeit with a fraction of their initial energy.

We compared the best-fit combination of DM mass and boost factor for the MH scenario with the best fits for the MH + SH scenario in [11]. The results are summarized in Table I: a much better fit could be obtained by including subhalos and a dark matter particle with \( M_{\text{DM}} = 2.2 \text{ TeV} \), rather than the standard MH-only \( M_{\text{DM}} = 1 \text{ TeV} \). Of course, the fits are further improved by allowing the normalizations of the background electrons and positrons to be additional free parameters, denoted as the “freely varying background,” as opposed to the standard backgrounds resulting from GALPROP simulations which include the effects of heavier nuclear species. Assuming this extra freedom has been advocated or used by numerous authors [4,5,12,33]. In Table I we also show the fit we obtain in the present analysis for the main-halo-only case with an isothermal profile and fixed background. It is significantly worse than the corresponding one for an Einasto profile.

### B. Annihilation channels

While we have mostly focused on the \( 4e \) final state, there is no reason for other, heavier particles not to be produced if the mass of the intermediate gauge boson is large enough. Since the amount of Sommerfeld enhancement ultimately depends on this mass, it is important to include the decays to muons and pions. The possible final states are all the four-particle combinations of \( 2e, 2\mu \) and \( 2\pi \). The muon and pion spectra are given by Ref. [34], whose authors were kind enough to provide us with the appropriate GALPROP implementation.

The branching ratios are given by \( r_i = f_i/\sum f_j \), where the \( f_j \) are given by

\[
f_i = \sqrt{\mu^2 - 4m_i^2} \left\{ \begin{array}{c} 4(\mu^2 + 2m_i^2), & i = e, \mu \\ (\mu^2 - 4m_i^2), & i = \pi \end{array} \right. \tag{8}
\]

In each \( f_i \), the square root factor comes from the phase space, while the rest is from the squared matrix element for the decay. Below threshold, \( f_i \) is defined to be zero. For a gauge boson with a mass \( \mu \gtrsim 1 \text{ GeV} \), we find \( r_e = r_\mu = 0.45 \) and \( r_\pi = 0.1 \). In this case the electrons produced from the final decay of the \( \mu \)-s and \( \pi \)-s peak at a lower energy, thus requiring a slightly higher mass of \( M_{\text{DM}} = 1.2 \text{ TeV} \) in order to fit the Fermi and PAMELA data. This is much smaller than the well-known \( M_{\text{DM}} \approx 2 \text{ TeV} \) best fit in the pure-muon final state [4,5,33] because of the large fraction of gauge bosons still decaying directly to high-energy electrons. These results are also shown in Table I.

### III. GAMMA RAY COMPUTATION FROM INVERSE COMPTON SCATTERING AND BREMSSTRAHLUNG

#### A. “Prompt” gamma ray emission (bremsstrahlung)

Prompt gamma ray emission appears in the final stage of DM annihilation, softening the lepton spectrum. The flux can be divided into main halo and subhalo parts:

\[
\frac{d\Phi}{dE\,d\Omega} = \frac{d\Phi_{\text{main}}}{dE\,d\Omega} + \frac{d\Phi_{\text{sub}}}{dE\,d\Omega} \tag{9}
\]

The astrophysical and particle physics dependences of each flux can be factorized as
\[
\frac{d\Phi_{\text{main}}}{dE_\gamma d\Omega} = \frac{1}{2} \frac{(\sigma v)}{4\pi} \rho_0^2 \frac{dN}{m_\chi^2 dE_\gamma} \langle J \rangle_{\text{main}}
\]

(10)

and

\[
\frac{d\Phi_{\text{sub}}}{dE_\gamma d\Omega} = \frac{1}{2} \frac{(\sigma v)}{4\pi} \frac{dN}{dE_\gamma} \langle J \rangle_{\text{sub}}.
\]

(11)

In each case, the \( \langle J \rangle \) factor depends only upon astrophysical inputs. The main halo \( J \) factor is defined as a line-of-sight (l.o.s.) integral of flux at each pixel:

\[
\langle J \rangle_{\text{main}} = \frac{1}{\Delta \Omega} \int d\Omega \int_{\text{l.o.s.}} ds \left( \frac{\rho_{\text{main}}[r(s, \psi)]}{\rho_0} \right)^2.
\]

(12)

In the case of flux originating from many distant subhalos, we may treat each one as a point source of radiation. In this case, the diffuse flux per solid angle requires a sum over each contributing source with density \( \rho_i \) and distance \( d_i \) within the observed solid angular region \( \Delta \Omega \):

\[
\langle J \rangle_{\text{sub}} = \frac{1}{\Delta \Omega} \sum_{\Delta \Omega} \left( \frac{\rho_{\text{sub}}[r_i]}{\rho_0} \right)^2 \int dV \frac{d^2j_i}{m_\chi^2}.
\]

(13)

This clearly depends not only on the density profiles, but also on the distribution of subhalos in the Galaxy. We will not present the results of the distant subhalo calculation of final-state radiation here, since it has been thoroughly explored by other authors in similar contexts. We direct the interested reader to Refs. [8,19,35].

Finally, if a particular subhalo is close enough to subtend an angle larger than the detector’s pixel size, it can no longer be treated as a point source: Eq. (12) must be used, including the angular dependence of the projected density profile of the given subhalo, \( \rho_{\text{SH}}(R, \theta, \phi) \). We will return to this case in Sec. IV B.

The particle physics contribution to (10) and (11) comes from the photon spectrum, defined as

\[
\frac{dN}{dE_\gamma} = \frac{1}{\langle \sigma v \rangle_{\text{total}}} \frac{d(\sigma v)}{dE_\gamma}.
\]

(14)

In the case of a two-lepton final state [36]:

\[
\frac{dN}{dx} = \frac{\alpha}{\pi} \frac{1 + (1 - x)^2}{x} \ln \left( \frac{s(1 - x)}{m_\gamma^2} \right)
\]

(15)

where \( x = 2E_\gamma/\sqrt{s} \) and \( s \) is the standard Mandelstam variable. We are interested in the case of TeV dark matter \( \chi \) annihilating to a four-lepton final state, with an \( O(1) \) GeV leptophilic gauge boson \( B \) as the messenger. The annihilation is dominated by \( \chi\chi \rightarrow BB \), where the \( B \)’s are on shell. The cross section can be obtained by first computing in the rest frame of the \( B \) using the decay \( B \rightarrow e^+ + e^- \) and then boosting to the lab frame, in which the slowly moving DM particles are approximately at rest. This can easily be done numerically. We present the resulting spectrum in Fig. 3. Since we will not make use of the final-state bremsstrahlung for other annihilation channels (\( 4\mu \) or \( 4\pi \)) we will not discuss their spectra.

### B. Inverse Compton scattering

Charged particles travelling through the interstellar medium scatter off ambient photons of the interstellar radiation field (ISRF), which is composed of microwave (\( \sim 10^{-9} \) eV) radiation from the cosmic microwave background (CMB), infrared (\( \sim 10^{-7} \) eV) radiation from dust, and optical (\( \sim 10^{-2} \) eV) photons from starlight. Along with the galactic magnetic fields, this is the main source of energy loss for electrons diffusing within the Galaxy. We will show that ISRF photons that have scattered with TeV-scale electrons have spectra that peak at several hundred GeV, which should fall squarely within the measurement window of diffuse gamma rays by the Fermi Large Area Telescope (LAT).

Once integrated over scattering angles, the well-known Klein-Nishina formula for the Compton scattering process \( e^\pm \gamma \rightarrow e^\pm \gamma' \) can be integrated along the line of sight to give the total flux of scattered photons per solid angle arriving on a detector [5,37]:

\[
\frac{d\Phi_{\gamma'}}{dE_{\gamma'} d\Omega} = \frac{1}{2} \frac{\alpha^2}{4\pi^2} \frac{e^2}{c^3} \int_{\text{l.o.s.}} ds \int \frac{dE_{\gamma}}{dE_{\gamma'} dE_{\gamma}} \frac{dE_{\gamma}}{E_{\gamma'} E_{\gamma}^2} f_{\text{IC}}.
\]

(16)

\( \int_{\text{l.o.s.}} ds \) represents the line-of-sight integral from the observer’s position to infinity (practically speaking, to the edge of the diffusion zone). We have used the definitions:

\[
f_{\text{IC}} = 2q \log q + (1 + 2q)(1 - q) + \frac{1}{2} \frac{(\epsilon q)^2}{1 + \epsilon q}(1 - q)
\]

(17)
and
\[ \epsilon = \frac{E_\gamma}{E_\gamma}, \quad \Gamma = \frac{4E_\gamma E_\nu}{m_e c}, \quad q = \frac{\epsilon}{\Gamma(1 - \epsilon)}. \] (18)

We numerically integrated Eq. (16) along the line of sight, as well as over the incoming particle energies. All the quantities in the integrand are known: we used the two-dimensional \((r, z)\) distribution of electrons and positrons \(dn_e/dE_e\) from DM annihilations produced with GALPROP, as discussed in Sec. II. For the ISRF, we used a realistic two-dimensional photon energy density distribution \(du_\gamma/dE\) from [26], which is publicly available on the GALPROP website. Both distributions assumed cylindrical symmetry around the Galactic axis. For each galactic latitude-longitude pair, the line-of-sight integration was performed in a three-dimensional sky from the Sun’s position to the edge of the diffusion zone which was taken to extend to a radius \(r_{\text{max}} = 20\) kpc and to a height \(z_{\text{max}} = 5\) kpc above and below the galactic plane. A trapezoidal integration step size of 0.1 kpc was found to be numerically converged. The values of \(dn_e/dE_e\) and \(du_\gamma/dE\) at each step were found in the heliocentric coordinate system by using a bilinear interpolation scheme. On top of the DM annihilation products, we used the densities of primary and secondary electrons as well as secondary positrons to compute the ICS contribution of the background lepton field. This had the effect of further constraining the gamma ray background.

We performed the integration once per grid point on an equally-spaced \(20^\circ \times 20^\circ\) latitude-longitude grid of the quarter-sky in the ranges \(\theta = [0, \pi/2], \phi = [0, \pi]\). This was sufficient to reconstruct the entire sky, given the symmetry of the data input.

C. Fermi all-sky diffuse gamma ray measurements

The Fermi Large Area Telescope (LAT) is a high-sensitivity gamma ray instrument capable of detecting photons in the \(\sim 30\) MeV to \(>300\) GeV range. It has an effective detector area of \(\sim 8000\) cm², a 2.4 sr field of view and can resolve the angle of an incident photon to 0.15° at energies above 10 GeV. Data from the first year of observation are publicly available from the Fermi Collaboration.

We used the all-sky diffuse photon file from the Fermi weekly LAT event data webpage [38]. This covered observations from mission elapsed time (MET) 239557417 to MET 272868753 (seconds), corresponding to 55 weeks of observation between August 8, 2008 and August 25, 2009. We processed the photon data with the Fermi LAT science tool software, available from the Fermi Science Support Center (FSSC) website. We first removed all events with a zenith angle greater than 105° to eliminate Earth albedo. The data were further trimmed to only keep the photons measured during “good” time intervals. We then created an exposure cube from the spacecraft data for the corresponding period, to account for effective instrument exposure.

The data were separated into 0.25° \(\times\) 0.25° latitude and longitude bins spanning the entire sky, and into 16 logarithmically separated bins from 100 MeV to 200 GeV. Uncertainties were assigned according to Ref. [39]. We compared our results to the August–December 2008 \(10^9 \leq |b| \leq 20^\circ\) spectrum presented by the Fermi Collaboration [39]. The half-year data agreed exactly, while adding the extra 8 months to the full 55-week data set changed the picture only very slightly. We rebinned the data into a \(40 \times 40\) grid, in correspondence with the ICS computation.

Before proceeding to the results of our numerical analyses, we should note that many factors contribute to the theoretical uncertainty. While we were able to reproduce the results of Simet et al. [24] quite closely, there are substantial discrepancies between the results of GALPROP and other methods of solving the transport equation. This lack of agreement is further discussed in [11]. There is an additional uncertainty in the injection spectrum of primary electrons, which serve, along with secondary electrons and positrons from spallation, as the astrophysical background to our results.

IV. EMPIRICAL FITS

As expected, we found that allowing subhalos to contribute to the overall flux of DM annihilation products reduced the flux of expected gamma rays from the galactic center, while increasing fluxes at higher galactic latitudes. The most stringent constraints were from the low-latitude regions just above and just below the galactic plane, where astrophysical sources of gamma rays are less prominent, but the DM distribution is still quite dense. Specifically, we used the lower right-hand region \((-9^\circ < b < -4.5^\circ, 0^\circ < \ell < 9^\circ\) in Galactic coordinates) which was found to be the most constraining, in agreement with Ref. [33].

After including the ICS from background electrons and positrons, we found that the boost factor of a main halo \(1\) TeV DM annihilation process cannot violate the bound \(BF \leq 25\) if the signal is to remain below the top Fermi LAT error bars. If we extend the constraint to \(\Phi_\gamma < \text{Fermi} + 2\sigma\), this condition is only slightly relaxed to \(BF \leq 30\). In the case of a \(2.2\) TeV DM candidate, these bounds become \(BF \leq 42\) and \(BF \leq 52\) at \(1\sigma\) and \(2\sigma\), respectively. While this agrees qualitatively with other works [5,33], we attribute our more stringent upper bounds mainly to our higher \(\rho_\phi\), as discussed in Sec. II A, to our inclusion of the ICS contribution from background electrons and positrons, but mainly to the different method used to solve the diffusion Eq. (1).

Using the best-fit scenario of Ref. [11], the reduction of flux was however not enough to overcome the constraints from the Fermi observations. This is illustrated in Fig. 4, which shows that the MH + SH scenario still violates constraints from the data by as much as \(4\sigma\). On its own, the predicted flux exceeded the data at energies above
of Table I) made no appreciable difference with respect to region. Allowing the background to freely vary (top section model even farther outside of the observationally allowed constraints.

gamma rays, and was not enough to satisfy the observed fluxes in this region of the sky are still well below included. It should however be emphasized that the previous works studying the ICS constraints have compared the effects of cored versus cuspy DM profiles, noting that the constraints are weaker for cored profiles. To better quantify exactly how much cuspiness can be tolerated, it is interesting to vary the parameters for Eq. (6) while keeping the local density constant at $\rho_0 = 0.37$ GeV cm$^{-3}$. This is illustrated in Fig. 6. Flatter profiles with $\alpha = 0.20$ or 0.25, $r_s = 30$ kpc reduce the gamma ray fluxes somewhat, but not enough to bring the predicted flux to within the observations in the offending energy bins between 10 and 100 GeV. The same is true for the isothermal profile, whose corresponding results are shown in Fig. 7. For both cases, the problem arises because the predicted background parameters for Eq. (6) while keeping the local density constant at $\rho_0 = 0.37$ GeV cm$^{-3}$. This is illustrated in Fig. 6. Flatter profiles with $\alpha = 0.20$ or 0.25, $r_s = 30$ kpc reduce the gamma ray fluxes somewhat, but not enough to bring the predicted flux to within the observations in the offending energy bins between 10 and 100 GeV. The same is true for the isothermal profile, whose corresponding results are shown in Fig. 7. For both cases, the problem arises because the predicted background parameters for Eq. (6) while keeping the local density constant at $\rho_0 = 0.37$ GeV cm$^{-3}$. This is illustrated in Fig. 6. Flatter profiles with $\alpha = 0.20$ or 0.25, $r_s = 30$ kpc reduce the gamma ray fluxes somewhat, but not enough to bring the predicted flux to within the observations in the offending energy bins between 10 and 100 GeV. The same is true for the isothermal profile, whose corresponding results are shown in Fig. 7. For both cases, the problem arises because the predicted background

100 GeV by at least 2$\sigma$, while we expect that additional constraints from $\pi^0 \rightarrow 2\gamma$ decays should also be large in this energy range [40] and push predictions from this model even farther outside of the observationally allowed region. Allowing the background to freely vary (top section of Table I) made no appreciable difference with respect to gamma rays, and was not enough to satisfy the observational constraints.

Figure 5 illustrates how the ICS gamma ray flux is increased at higher galactic latitudes when subhalos are included. It should however be emphasized that the predicted fluxes in this region of the sky are still well below the level of Fermi observations.

A. Less cuspy dark matter profiles

In Sec. II A we mentioned the motivations for considering less cuspy DM profiles. Many previous works studying the ICS constraints have compared the effects of cored versus cuspy DM profiles, noting that the constraints are weaker for cored profiles. To better quantify exactly how much cuspiness can be tolerated, it is interesting to vary the parameters of the Einasto profile that control this [14, 41]. In particular, larger values of $\alpha$ and $r_s$ correspond to less concentrated halos. We ran simulations of the lepton distribution and gamma ray fluxes with slightly different

FIG. 6 (color online). $-9^\circ < b < -4.5^\circ$, $0^\circ < \ell < 9^\circ$ region. Similar to previous figures, showing how reducing the cuspiness of the Einasto profile (Eq. (6)) reduces predicted total gamma ray signal (magenta line). Here $\alpha = 0.20, 0.25$, respectively, and $r_s = 30$ kpc.
gamma flux is not far below the observed flux in the most constraining bins. This leaves very little room for the additional contribution from the DM decay products ICS signal.

Increasing the intermediate gauge boson mass to 1 GeV, and thus allowing a decay to muons and pions according to the branching ratios described in Sec. II B does not alleviate the problem. Indeed, the $1 \sigma$ ($2 \sigma$) bounds become $BF < 23(<28)$ for an Einasto profile, and $BF < 63(<72)$ in the isothermal case. These fall well short of the required $BF = 118$ to explain the Fermi and PAMELA excesses, as long as the DM mass is increased to $M_{DM} = 1.2$ TeV. These results are summarized in the bottom of Table II. The reason ICS constraints are stronger when muons are included is due to the nature of the data. Indeed, the peak of the ICS spectrum lines up with the most constraining data point when $M_{DM} = 1.2$ TeV. This provides a stronger than expected constraint, relative to the $4e$ final state at $M_{DM} = 1$ TeV.

### B. Close subhalo

The above analyses implicitly assume that no single subhalo dominates the lepton signal. But if a subhalo happens to be very close (within a kpc) to the solar system, the picture changes significantly, since the electrons and positrons from the close subhalo can dominate the observed flux, and its gamma ray emissions can come from a sizable solid angle in the sky. We treat this case separately from the previous subhalo scenario, since a larger DM mass is no longer required to produce the observed lepton signal; rather, the small amount of ICS energy loss during propagation from a local subhalo means that a 1 TeV-scale DM particle appropriately conforms to the Fermi $e^+ + e^-$ measurements. We concentrate on the $4e$ final-state channel, although previous results allow this to be generalized. The solution depends linearly on the spectrum $dN_e/dE$, so that the boost factor required to explain the observed lepton excess should scale in the same way that it does in the main halo scenario: $BF_{(e,\mu,\pi)/BF_{4e}} \approx 118/110$, as read from Table I.

Since GALPROP is not easily adapted in its 2D mode to include the effects of a highly localized additional source term, we adopt a semianalytic approach to solve the diffusion Eq. (1) for leptons produced in the nearby subhalo. Given that the leptons and gamma rays in this scenario would be from a local origin, the spatial dependence of the interstellar radiation and magnetic fields becomes much less important. We used the method described in Ref. [42],

| Subhalo | $r_s$ (kpc) | $\rho_s$ | $\log BF$ | $d_{\text{min}}$ (pc) | $V_{\text{max}}$ (km/s) |
|---------|------------|----------|------------|---------------------|------------------|
| 1       | 0.01       | 69       | 4.74       | 33.9                | 2.9              |
| 2       | 0.1        | 3.46     | 4.34       | 95.5                | 6.7              |
| 3       | 3.2        | 0.04     | 3.76       | 178                 | 22               |
| 4       | 0.9        | 1.27     | 2.35       | 165                 | 36               |
| 5       | 1.1        | 2.0      | 1.70       | 170                 | 55               |

Main halo, $4e$ channel

| Subhalo | $r_s$ (kpc) | $\rho_s$ | $\log BF$ | $d_{\text{min}}$ (pc) | $V_{\text{max}}$ (km/s) |
|---------|------------|----------|------------|---------------------|------------------|
| Einasto | 25         | 0.048    | $< 1.40$   | $\cdots$           | 201–277          |
|         | 3.2        | 2.32     | $< 1.81$   | $\cdots$           | 201–277          |

Main halo, $4e + 4\mu + 4\pi$ channel

| Subhalo | $r_s$ (kpc) | $\rho_s$ | $\log BF$ | $d_{\text{min}}$ (pc) | $V_{\text{max}}$ (km/s) |
|---------|------------|----------|------------|---------------------|------------------|
| Einasto | 25         | 0.048    | $< 1.36$   | $\cdots$           | 201–277          |
|         | 3.2        | 2.32     | $< 1.80$   | $\cdots$           | 201–277          |

| Subhalo | $r_s$ (kpc) | $\rho_s$ | $\log BF$ | $d_{\text{min}}$ (pc) | $V_{\text{max}}$ (km/s) |
|---------|------------|----------|------------|---------------------|------------------|
| Isothermal | 25         | 0.048    | $< 1.45$   | $\cdots$           | 201–277          |
|         | 3.2        | 2.32     | $< 1.86$   | $\cdots$           | 201–277          |
with the same diffusion parameters as presented in Sec. II (of the present work), but with an energy-loss coefficient parametrized by

$$b(x, E) = -\frac{dE_{\ell}}{dt} = E_{\ell}^2 \frac{\tau_{E}}{\tau_{E}}$$  \hspace{1cm} (19)$$

with $\tau_{E} = 10^{16}$ s GeV characterizing the local energy-loss rate.

We sampled subhalos from the Via Lactea II simulation to identify examples that could allow for simultaneously fitting the PAMELA/Fermi lepton fluxes and the Fermi gamma ray fluxes. Four such examples are labeled as SH1 - SH4 in Table II, and a fifth (SH5) is one that we have “engineered” by choosing parameters that are close to those of SH4, but with a higher density and hence higher circular velocity, dynamically related to each other by Eq. (13) of [19],

$$V_{\text{max}}^2 = f_{V} 4 \pi G \rho_{\star} r_{s}^2$$  \hspace{1cm} (20)$$

with $f_{V} = 0.897$. Because of the higher density, SH5 requires a lower boost factor to produce the observed lepton signal, and so it represents a kind of best-case scenario. The distribution of Via Lactea II subhalos in the space of $(r_{s}, \rho_{\star})$ is shown as a scatter plot in Fig. 8, and the five subhalos of interest are highlighted on this plot. They are atypical in the sense of needing a higher-than-average central density. A further caveat is that such a large $r_{s}$ is unlikely at small distances from the GC due to tidal disruption. Indeed, subhalos within the visible galaxy in the Via Lactea II simulation were of the order $r_{s} = 0.05 \sim 0.85$ kpc, falling below the $0.9 \sim 1.1$ kpc compatible with the most plausible particle physics scenario discussed in Sec. V.

Each subhalo was situated along an optimal axis, namely, that connecting the earth to the GC. Such an accidental alignment makes it easier to “hide” the gamma rays originating from the subhalo since they are coming primarily from the same direction as the GC, where the background emissions are strongest. This is also the reason that the most stringent ICS constraints on the main halo arise from the regions $4.5^\circ < |b| < 9^\circ$ of galactic longitude instead of the most central region. However in this case we find that the biggest contribution to the emission is from final-state bremsstrahlung rather than ICS. The latter is found to produce gamma ray fluxes that are 3 orders of magnitude smaller than observed. This is consistent with the fact that the main source of ICS is IR radiation and starlight, which is concentrated far from the vicinity of the solar system.

Results were then compared to the Fermi lepton and gamma ray data in order to establish constraints. The strictest gamma constraints were at the largest energy data point from the Fermi LAT analysis of $E = 162$ GeV, because of the shape of the FSR spectrum, which rises steadily until $\sim 1$ TeV. We used a slightly different region of the sky than in our previous ICS analysis, $4.5^\circ < |b| < 9^\circ, 9^\circ < |\ell| < 18^\circ$, because there were not enough good data points in this energy bin at lower longitudes to constrain the data. We compared the lepton prediction to the Fermi measurements at 559 GeV, where the observed $e^+ + e^-$ spectrum is at a maximum deviation from a

![FIG. 9 (color online). Fluctuations of gamma rays and $e^+ + e^-$ from the five subhalos presented in Table II. The gamma ray fluxes (curve labeled by the number of the corresponding subhalo) are at $E_{\gamma} = 137$ GeV, whereas the leptons are at an energy of 559 GeV (the peak of the observed Fermi spectrum). In both cases, the amplitude is the predicted flux divided by the observed flux from the Fermi satellite, such that a value of $10^0$ means that the predicted flux is equal to the observed value. Boost factors in each case (as given in Table II) were fixed to allow the Fermi lepton signal to be explained entirely by the subhalo. The allowed position of each subhalo with respect to earth is therefore the region to the right of each gamma ray curve, up to $\sim 2$ kpc where the lepton flux starts to fall.](image-url)
power law. In both cases we included the additional constraints from astrophysical backgrounds computed by GALPROP and by our ICS routine.

Results are shown in Fig. 9. If the single subhalo is allowed to saturate the observed lepton signal, Fig. 9 gives clear bounds (summarized in Table II) on the proximity of each subhalo, providing a minimum distance from the solar neighborhood to such a subhalo. So long as the boost factor for the main halo remains sufficiently small, this scenario can therefore overcome the ICS constraints that restricted the standard MH-only model.

C. Astrophysical prediction and extragalactic constraints

In Fig. 10 we provide an example of the gamma ray flux predicted by the close subhalo scenario, as compared to the main halo scenario. The gamma ray flux comes predominantly from final-state radiation rather than inverse Compton scattering of the annihilation products. We chose the energy bin $E = 23$ GeV, which is the most constraining for the main halo case. Although both scenarios converge at high latitudes, low latitude measurements have already ruled out the main halo scenario, and provide a way to constrain the model. With more exposure and precise removal of point sources, the Fermi LAT may provide a diffuse background low enough to rule out these predictions. As a further test, census experiments such as the upcoming Gaia satellite may provide a precise enough map of the local gravitational potential to confirm or rule out the presence of such a DM overdensity [43]. Direct measurement of such an overdensity would however be difficult: a subhalo such as SH5, located at a distance that would not saturate gamma ray bounds, would contribute less than 0.1% of the local DM density.

From previous works, we infer that extragalactic bounds on this scenario are not as strong as the ones we have computed above. Bounds from dwarf spheroidal galaxies could plausibly be important since the velocity dispersions are of the same order as what is required for our subhalo enhancement, i.e. $\sim 10$–$50$ km s$^{-1}$ [44]. However, the most stringent Fermi LAT bounds [45] from such galaxies put the upper limit on DM annihilation into a $2\mu$ final state at around $BF = 3000$ if only final-state radiation is considered, and around 300 if ICS bounds are included as well. [46] computed the cosmological dark matter annihilation bounds for the same $2\mu$ final-state scenario, and find that $BF$ larger than 300 is excluded at the 90% confidence level. This is using the results of the Millennium II structure formation simulation, and is indeed model-dependent. Extrapolation to the $4\mu$ scenario is independent of astrophysics. We can therefore take the results of [5,33] who have constructed bounds on both channels. They show that FSR bounds are consistently an order of magnitude weaker in the $4\mu$ case, given the softer photon spectrum in this scenario. We can therefore take these extragalactic results to be far less constraining than the stringent bounds from the center of our own galaxy.

Finally, we verify that this model does not saturate bounds on dipole anisotropy of the cosmic ray $e^+ + e^-$ spectrum. The dipole anisotropy can be defined as

$$\delta = 3\sqrt{\frac{C_1}{4\pi}}$$

FIG. 10 (color online). Dependence of predicted gamma ray fluxes on galactic latitude $b$, in the region $-9^\circ < \ell < 9^\circ$ at $E = 23$ GeV, the most constraining energy bin for the main halo scenario. Black: main halo scenario (Einasto profile, $BF = 110$) Dashed: subhalo 5, as specified in Table II. Background ICS is included in both predictions, but the signal is dominated by final-state radiation. Dots are the Fermi data for that region and energy.

FIG. 11 (color online). Dipole anisotropy $\delta$ of the cosmic ray electron and positron flux predicted by SH5 if it saturates the Fermi excess. Background cosmic ray electrons and positrons are included, and taken to be isotropic. $\delta$ increases monotonically with energy from the red line (60 GeV) to the black line (500 GeV).
where \( C_1 \) is the standard dipole power of the measured electron and positron flux in the sky. The Fermi LAT Collaboration [47] has presented upper bounds on this quantity. These range from \( \delta \leq 3 \times 10^{-3} \) at \( E_e \approx 60 \) GeV up to \( \delta \leq 9 \times 10^{-2} \) at \( E_e \approx 500 \) GeV. Given a diffusive model, this can be computed [47]:

\[
\delta = \frac{3D(E)}{c} \left[ \frac{\bar{\epsilon}_e}{n_e} \right], \quad (22)
\]

where \( D(E) \) is the diffusion coefficient (3) and \( n_e \) is the density of cosmic ray electrons and positrons, including astrophysical backgrounds. Taking the background to be isotropic, we computed the dipole anisotropy in the case of a single close subhalo producing enough electrons to explain the Fermi excess. In every case \( \delta \) falls well below bounds. Results for SH5 are presented in Fig. 11. The anisotropy rises monotonically with energy, from 60 GeV (red line) to 500 GeV (black line).

**V. PARTICLE PHYSICS REALIZATIONS**

In the previous sections we have identified scenarios where subhalos could provide the observed excess PAMELA and Fermi leptons, from a purely phenomenological perspective. In particular, certain values for the annihilation cross section boost factors are needed for the subhalos, and upper bounds for that of the main halo (depending upon assumptions about its density profile) were derived. It is interesting to ask whether simple particle physics models with boost factors from Sommerfeld enhancement can be consistent with these requirements.

The simplest possibility for model building is dark matter that annihilates into light scalar or vector bosons, which subsequently decay into leptons. This class of models automatically gives a boost factor to the annihilation cross section, through multiple exchange of the boson, resulting in Sommerfeld enhancement. However it is not obvious that one can find models with the desired boost factors for the subhalos and main halo. One constraint that limits our freedom is to not exceed the measured density of dark matter. It will turn out that our mechanism works most naturally if the DM responsible for signals in the galaxy is a subdominant component comprising some fraction \( 1/f \) of the total DM population [14], with \( f > 1 \).

We focus on the case of a GeV-scale U(1) vector boson that kinetically mixes with the photon. Such models have the advantage of naturally explaining the coupling to light leptons, without producing excess antiprotons that would contradict PAMELA observations. Let us denote the vector’s mass by \( \mu \) and the coupling by \( g \), with \( \alpha_s = g^2/4\pi \). If \( M \) is the DM mass, then the Sommerfeld boost factor is controlled by two dimensionless parameters: \( \epsilon_\phi = \mu/(\alpha_s M) \) and \( \epsilon_v = v/(\alpha_s c) \), where \( v \) is the DM velocity in the center of mass frame. A reasonably accurate approximation to the exact Sommerfeld enhancement is given by the expression [48,49]

\[
S = \frac{\pi}{\epsilon_\phi \cosh X - \cos \sqrt{2\pi / \epsilon_\phi} - X^2} \quad (23)
\]

where \( \epsilon_\phi = (\pi/12)\epsilon_\phi \) and \( X = \epsilon_v/\epsilon_\phi \). (The cosine becomes \( \cosh \) if the square root becomes imaginary.)

To take into account leptophilic DM that is only a subdominant component of the total DM, suppose that \( \alpha_{e,\text{th}} \) is the value of \( \alpha_e \) that would give the correct thermal abundance, which scales like the inverse annihilation cross section \( \sigma^{-1} \propto \alpha_s^{-2} \); then we can parametrize \( \alpha_s = \sqrt{f}\alpha_{e,\text{th}} \). The rate of annihilations goes like \( \rho_1^2 \sigma \propto 1/f \) if \( \rho_1 \) stands for the leptophilic component of the DM. We accordingly define an effective boost factor

\[
\bar{S} = \frac{S}{f} \quad (24)
\]

where \( S \) is the intrinsic Sommerfeld enhancement factor. Thus any constraint on \( S \) in a theory with \( f = 1 \) becomes a constraint on \( \bar{S} \) in the more general situation.

### A. Averaging of boost factor

Of course, the DM velocity has no definite value; instead we need to average over the possible values within the subhalos and the main halo, weighted by the appropriate distribution function. We take it to be Maxwell-Boltzmann with a cutoff at some escape velocity,

\[
f(v) = N e^{-3v^2/2v^2} \theta(v_{\text{esc}} - v). \quad (25)
\]

This isotropic form is only an approximation since the true distribution has some small anisotropy between the radial and angular components; we will for simplicity ignore this complication. The velocity dispersion \( v_s = (v^2)^{1/2} \) depends upon the radial distance \( r \) from the center of the halo or subhalo. The dependence has been measured for the subhalos in the *Via Lactea II* simulation; see Fig. 12. The shape is universal, but is scaled along the respective axes by parameters \( V_{\text{max}} \) and \( r_{V_{\text{max}}} \) that depend upon the subhalo. The latter is related to the scale radius by \( r_{V_{\text{max}}} = 2.212r_s \); the former is given by (20) and also listed in Table II for the subhalos of interest. For numerical purposes we fit the sides of the curve passing through the points of Fig. 12 by lines (omitting the rightmost point), and the middle by an inverted parabola.\(^2\) We use the same form of \( v_s \) for the main halo, with \( r_s = 25 \) kpc and \( V_{\text{max}} = 201 \) km/s. Other authors have advocated higher values of the velocity dispersion, \( v_s = 309 \) km/s at \( r = r_0 \) [50], which would correspond to \( V_{\text{max}} = 277 \) km/s in the

\[\text{velocity dispersion curve is fit by}
\]

\[
y = \begin{cases} 
1.309 + 0.232x, & x < -0.841, \\
0.976 - 0.3437x, & x > -0.383, \\
0.9618 - 0.5475x - 0.4413x^2, & \text{in between}
\end{cases}
\]

where \( x = \log_{10}r/r_{V_{\text{max}}} \) and \( y = v_s/V_{\text{max}} \).
present parametrization. We will also consider the higher value to take account of this uncertainty.

The escape velocity can be computed explicitly for the subhalos from the standard result \( v_{\text{esc}}^2 = \frac{G}{r^2} \langle M(r)/r^2 \rangle dr \), where \( M(r) = 4\pi \rho \int_r^\infty r^2 \rho \, dr \) is the mass within radius \( r \). The result for an Einasto profile is

\[
v_{\text{esc}}^2 = G \rho \varepsilon^{2/\alpha} \frac{8\pi \alpha^3}{\alpha^3} \left[ \frac{2}{\alpha} \Gamma \left( \frac{2}{\alpha} \right) \left( \frac{2}{\alpha} - \frac{2}{\alpha} \frac{1}{\alpha} \left( \frac{r}{r_s} \right)^{\alpha} \right) \right] + \frac{r_s}{r} \left[ \frac{3}{\alpha} \frac{2}{\alpha} \left( \frac{r}{r_s} \right) \right]. \tag{27}
\]

where \( \Gamma(s, x) \) is the upper incomplete gamma function. For the main halo, this procedure would not be correct because of the significant contribution of baryons, not included here. We adopt the result for \( v_{\text{esc}} \) of Ref. [14] for the main halo (see Appendix C of that reference).

With these ingredients, we can compute an average Sommerfeld enhancement factor \( \langle S \rangle \) for each subhalo:

\[
\langle S \rangle = \frac{\int_r^\infty d^2r r^2 \rho^2 \int d^3v_1 d^3v_2 f(v_1) f(v_2) S_{12} |\tilde{v}_1 - \tilde{v}_2|}{\int_r^\infty d^2r r^2 \rho^2}. \tag{28}
\]

The factor of \( \frac{1}{2} \) in the argument of \( S \) occurs because the \( v \) appearing in Eq. (23) through \( \varepsilon_{v} \) is half of the relative velocity. \( \rho^2 \) is the appropriate weighting factor because the rate of annihilations is proportional to \( \langle \sigma v \rangle \rho^2 \). For the subhalos, the range of integration for \( r \) is from 0 to \( \infty \), but for the main halo we take lower and upper limits \( r_{1,2} \) that correspond to the angular region of the sky that is used to set the gamma ray constraints: \( r_1 = 0.67 \) kpc and \( r_2 = 1.34 \) kpc. The reason is that the bound \( \tilde{S} < 30 \) for the main halo comes from the gamma ray constraint rather than from lepton production. We are thus interested in the boost factor relevant to the region \( 4.5^\circ < |b| < 9^\circ \) of galactic latitude. The distances of closest approach to the galactic center, hence largest rate of \( \gamma \) ray production associated with these lines of sight, are given by \( r = r_0 \sin b \).

B. Relic density constraint

The enhancement factor (23) depends rather strongly on the gauge coupling \( \alpha_s \), therefore it is interesting to know what constraint the relic density places upon \( \alpha_s \). The effect of a Sommerfeld-enhanced DM model on the relic density has been discussed by [51]. Notice that DM transforming under a U(1) gauge symmetry as we have assumed must be Dirac and therefore could have a relic density through its asymmetry, similar to baryons. However, unless the DM was never in thermal equilibrium, then \( \alpha_s \) should not be less than the usual value \( \alpha_s, \text{th} \), leading to the correct relic density, since otherwise the thermal component will be too large.

There are two kinds of final states for annihilation of DM in this class of models: into a pair of gauge bosons \( B_\mu \), by virtual DM exchange in the \( t \) and \( u \) channels, or into dark Higgs bosons \( h \), by exchange of a gauge boson in the \( s \) channel. Assuming the DM (\( \chi \)) is much heavier than the final states, the respective squared amplitudes, averaged over initial and summed over final spins, are

\[
\frac{1}{4} \sum_i |\mathcal{M}_i|^2 = \left\{ 4g^4(1 + 2v^2), \quad \chi\chi \rightarrow BB \right\} \chi\chi \rightarrow h\bar{h} \tag{29}
\]

where \( q \) is the U(1) charge of \( h \) relative to \( \chi \) (replace \( q^2 \rightarrow \sum_i q_i^2 \) for multiple Higgs bosons), \( \theta \) is the scattering angle, and we have included the leading dependence on the initial velocity \( v \) in the center of mass frame. The factor \( \cos^2 \theta \) averages to \( 2/3 \) in the integral over \( \theta \). In computing the associated cross section, it must be remembered that the 2B final state consists of identical particles, while the Higgs channel does not. The total amplitude can therefore be written in the form \( \frac{1}{4} \sum_i |\mathcal{M}_i|^2 = 4g^4(a + b v^2) \), with

\[
a = 1 + \frac{1}{4} \sum_i q_i^2, \quad b = 2 \left( 1 - \frac{1}{12} \sum_i q_i^2 \right) \tag{30}
\]

if we use the phase space for identical particles.

To find the cross section relevant during freeze-out in the early universe, we thermally average the \( v \)-dependent \( \langle \sigma v \rangle \) following Ref. [52]. We include approximately the effect of Sommerfeld enhancement as described there, to obtain

\[
\langle \sigma v \rangle \equiv \frac{\pi \alpha_s^2}{2M^2} \left( a \left( 1 + \alpha_s \sqrt{\frac{M}{T}} \right) + \frac{T}{M} \left( b - \frac{4}{3} a \right) \right) \times \left( \frac{3}{2} + \alpha_s \sqrt{\frac{M}{T}} \right). \tag{31}
\]

The terms that are subleading in \( \alpha_s \), but enhanced by \( \sqrt{M/T} \), are due to the Sommerfeld correction. We approximate the freeze-out temperature as \( T \approx M/20 \), the usual result of solving the Boltzmann equation for DM in the TeV mass range, and equate \( \langle \sigma v \rangle \) to the value \( \langle \sigma v \rangle_0 = 3 \times 10^{-26} \text{ cm}^3/\text{s} \) usually assumed to give the correct relic density. This gives an implicit equation for \( \alpha_s, \text{th} \) of the
form $\alpha_g^2 = c_1 M^2 / (\sigma v)_0 / (1 + c_2 \alpha_g)$, which however quickly converges by numerically iterating. Figure 13 displays the resulting dependence of $\alpha_{g,\text{th}} / M$ on $\Sigma_i q_i^2$ for several values of $M$.

The bound that the density of the leptophilic DM component not exceed the total DM density is $\alpha_g > \alpha_{g,\text{th}}$. We parametrize the coupling by $\alpha_g = \sqrt{f} \alpha_{g,\text{th}}$ with $f > 1$ in what follows.

**C. Interpolation between $4e$ and mixed final states**

In our numerical computations with GALPROP, we considered two cases for the final-state annihilation channels: either $\chi \chi \rightarrow 4e$, applicable for gauge bosons with mass $\mu < 2m_e$, or to a mixture of electrons, muons and charged pions, appropriate for decays of gauge bosons with mass greater than $2m_e$. The relative abundance of $e$, $\mu$ and $\pi$ in the mixed final state can be computed from the branching fractions of the decays, discussed in connection with Eq. (8).

For intermediate values of the gauge boson mass, $2m_e < \mu < 2m_\pi$, we can use the branching ratios to interpolate between our maximum-allowed MH or best-fit SH boost factors for the $4e$ case and those of the fiducial $e + \mu + \pi$ case. The maximum allowed boost factors of the main halo complying with the ICS constraints are taken from Table II. To estimate the best-fit boost factors for the subhalos in the fiducial $e + \mu + \pi$ final state, we rescale the $4e$ results shown in Table II by the ratio of best-fit boost factors for the main halo, in the MH-only scenario. These ratios are $118/110$ for the Einasto profile and $119/113$ for the isothermal, quite close to unity, and so the best-fit values of the SH boost factors hardly depend upon this scaling. More significant is the change in the best-fit mass, from $M = 1.0$ to $1.2$ TeV, which enters into the computation of the Sommerfeld enhancement and the value of the gauge coupling ($\alpha_g \sim M$). We use the branching ratios to interpolate $M$ as well. For the MH upper bounds in the small- and large-$\mu$ regions, we use the values from Table II, and interpolate similarly for intermediate $\mu$.

**D. Theoretical fits**

For a given value of the gauge coupling $\alpha_g$, we can determine the predicted boost factors as well as the desired values for each subhalo, as a function of the gauge boson mass $\mu$, and similarly for the main halo, except here we have an upper bound on $\langle \bar{S} \rangle$ rather than a best-fit value. This bound in fact presents the biggest challenge to finding a working particle physics model. For $\alpha_g$ close to the thermal relic density value $\alpha_{g,\text{th}}$, the predicted boost factor of the main halo far exceeds the bound $\langle \bar{S} \rangle \leq 30$, even if we try to decrease $\langle \bar{S} \rangle$ by reducing $\alpha_g$ via a large hidden Higgs content or by increasing the dispersion of the main halo. Figure 14 illustrates the discrepancy for $\Sigma_i q_i^2 = 16$ and $V_{\text{max}} = 277$ km/s. Lower values of $V_{\text{max}}$ or $\Sigma_i q_i^2$ only make this tension worse.

As we mentioned above, even though it is not theoretically possible to make the gauge coupling weak enough to solve this problem, ironically one can rescue the scenario by increasing $\alpha_g$ beyond the thermal value, since this suppresses the relic density of the DM component we are interested in, and thus reduces the scattering rate. Allowing $\alpha_g = \sqrt{f} \alpha_{g,\text{th}}$ decreases both the density of the leptophilic component and the effective boost factor by $1/f$. With $f \sim 50–500$, depending upon the shape of the main halo DM density profile, we can satisfy the constraint on the MH and still have a large enough boost in certain hypothetical nearby subhalos for them to supply the observed lepton excess. The minimum value of $f$ that is needed is larger for a cuspy main halo.
We give two working examples in Fig. 15, one with $f = 500$ and $V_{\text{max}} = 277$ km/s (the larger value advocated in Ref. [50]) and assuming an Einasto profile for the main halo, and the other having $f = 50$ and $V_{\text{max}} = 201$ km/s (the more standard assumption for the velocity dispersion), with an isothermal halo. In these figures the averaged boost factor $\langle S \rangle$ of the relevant subhalos are plotted as solid lines, while the required values of $\langle S \rangle$ are the dashed curves. Wherever these intersect represents a possible value of the gauge boson mass to consistently explain the observed lepton excess. At the same time, the main halo boost factor (lowest solid curve in the small-$\mu$ region) must lie below the black dashed lines to satisfy gamma ray constraints. The rationale for taking the larger value of $V_{\text{max}}$ for the Einasto profile is that larger velocities help to suppress the boost factors and thus make it easier to satisfy the ICS constraint, so that we are not forced to choose an even larger value of $f$. The isothermal profile is less constrained.

In the first panel of Fig. 15 with the Einasto profile, only subhalos SH4 and SH5 have large enough boost factors to ever reach the required values. There are many points of intersection, but mainly those for SH5 and in the mass range $\mu < 750$ MeV are consistent with the gamma ray bounds on the main halo. For the isothermal halo, these constraints are less stringent, and it is possible to find points of intersection using $f = 50$ for all five of the sample subhalos, although they are much more rare for SH1 – SH3 than for SH4 and SH5. In this example, the intersection points that respect the ICS bound are restricted to $\mu \lesssim 1$ GeV. For larger values of $f$, all the boost factors will be further suppressed, and $\mu > 1$ GeV will become allowed for SH4 and SH5.

One advantage of requiring large $f$ is that the corresponding dilution of the DM density by $1/f$ ensures that the model satisfies stringent CMB constraints from annihilations in the early universe changing the optical depth [53–55], as pointed out in [14]. The CMB constraint is shown in Fig. 16, along with the PAMELA/Fermi allowed regions from Ref. [33] for $4\mu$ and $4\mu$ final states. The $4\mu$ case is allowed by the CMB constraint, but $4\mu$ is ruled out. Because our model has at most a fraction of 0.45 of muons in the final state, it is probably already safe, but the additional weakening of the bound by the factor $1/f$ ensures that this will be the case. Similarly, our scenario overcomes the no-go result of Ref. [10], which pointed out that Sommerfeld enhanced annihilation in the early universe leads to constraints on the MH boost factor which are lower than those needed to explain the lepton anomalies. Our MH boost factor can satisfy these constraints since the MH is no longer considered to be the source of the excess leptons.

![FIG. 15 (color online). Predicted effective boost factors $\langle S \rangle$ as a function of gauge boson mass (solid curves) and target values (or upper limit in case of main halo, dashed curves) to explain PAMELA/Fermi lepton observations and Fermi gamma ray constraints. Pair of dashed curves for main halo (MH) correspond, respectively, to 1 and 2σ upper limits. Left panel is for $f = 100$, $V_{\text{max}} = 277$ km/s, Einasto main halo profile; right is for $f = 25$, $V_{\text{max}} = 201$ km/s, isothermal main halo profile. Subhalos are those of Table II. Points which satisfy all constraints are those where subhalo curves intersect their corresponding dashed line while the main halo curve lies below its dashed lines.](image1)

![FIG. 16 (color online). Allowed regions for PAMELA and Fermi excess leptons, and upper bounds from inverse Compton gamma rays, from Ref. [33], for Einasto profile with $\alpha = 0.17$ and $r_s = 20$ kpc. CMB constraint is from Ref. [53].](image2)
The Sommerfeld enhancement is nearly saturated for the low velocities of the subhalos at these large values of $\alpha_g \sim 0.1–0.35$, so their ($S$) curves are nearly overlapping except at the smallest gauge boson masses. The main halo boost factor is not saturated on the other hand, and lies below the FSR bound for most values of $\mu$. We have chosen the gauge couplings, parametrized by $f$, to nearly saturate the FSR bound. By taking larger $\alpha_g$ (larger $f$), the bounds could be satisfied by a larger margin. But this would also reduce the ($S$) values of the subhalos by a similar amount, making it more difficult to get a large enough lepton signal from SH1 – SH3. SH4 and SH5 would remain robust possible explanations.

VI. DISCUSSION AND CONCLUSIONS

We have shown that gamma ray constraints on leptophilic annihilating dark matter are significantly stronger than in previous studies, when we take into account the contributions to inverse Compton scattering from primary and secondary electrons and positrons, before including excess leptons from the DM annihilation. We attribute part of this difference to the method of solving the diffusion Eq. (1)—fully numerical rather than semianalytic—meaning that the $(r, z)$ space-dependence of the diffusion coefficient is taken into account. The difference between the predicted and observed spectra of gamma rays is greatly reduced, leaving little room for new contributions. Because of this, even cored halos, which were allowed by other analyses, become excluded. However, we find that these constraints can be weakened and possibly overcome if annihilations in a nearby subhalo are the dominant source of anomalous leptons, rather than annihilations in the main galactic DM halo. In this way, the PAMELA/Fermi cosmic ray excesses can be explained, without violating bounds from the recent Fermi LAT diffuse gamma ray survey.

It must be admitted that the subhalo loophole we present is rather special. First, only atypically dense subhalos, relative to the sample provided by Via Lactea II, give a large enough boost factor (see Fig. 8). Second, the subhalo would need to accidentally line up nearly with the galactic center in order for the ICS gamma rays associated with these leptons to be sufficiently hidden by the noisy background of the GC. Of course, had we neglected ICS contributions of background electrons, similar to previous studies, less fine-tuning of the subhalo properties would be necessary. Also we do not require the subhalo to be particularly close; Fig. 9 shows that the lepton flux only starts to fall at distances of $\sim 3$ kpc. Our finding could be regarded as a proof of concept. It is possible that the effects of unresolved substructure within the subhalo [15], which can increase the boost factor, would also make the scenario work more easily. On the positive side, there is the opportunity of testing whether there is such a nearby subhalo, since we predict the spectrum of ICS gamma rays it contributes (see Fig. 10). A better understanding of backgrounds, for example, from point sources, could make it possible to rule out the proposal.

On the particle physics side, we have shown in detail that the subhalo scenario can be made consistent with one of the simplest models of leptophilic dark matter, where the DM is in a hidden sector that communicates with the standard model only through kinetic mixing with hypercharge of a new gauge boson in the GeV mass range. The relative couplings to leptons and charged pions are completely specified and the model has only two free parameters, the gauge coupling $\alpha_g$ and gauge boson mass $\mu$ (the DM mass $M$ is fixed by fitting to the spectrum of anomalous $e^+ + e^-$). The gauge coupling is constrained by the relic density of the DM. The Sommerfeld enhancement factor is completely fixed by $(\alpha_g, \mu, M)$ and the kinematic halo properties. We find (similarly to Ref. [14]) that the predicted boost factor for the main halo is always too large to satisfy ICS constraints unless the leptophilic component of the DM is small, comprising a fraction of order $1/f = 0.02–0.002$ of the total DM. The small fraction can be achieved by assuming $\alpha_g$ is larger than the value required for the usual thermal abundance by the factor $f \sim 50–500$. This raises the interesting possibility that the DM that may be responsible for the cosmic ray anomalies is distinct from the dominant DM species that might be discovered by direct detection.

ACKNOWLEDGMENTS

We would like to thank Ilias Cholis for kindly giving us access to his modifications to the GALPROP code, and Troy Porter for valuable help with the Fermi data analysis. We thank Andrew Frey, Thomas Konstandin, Guy Moore, Natalia Toro and Mike Trott for helpful discussions. Our research is supported by NSERC (Canada) and FQRNT (Québec).

[1] N. Arkani-Hamed, D. P. Finkbeiner, T. R. Slatyer, and N. Weiner, Phys. Rev. D 79, 015014 (2009).
[2] M. Pospelov and A. Ritz, Phys. Lett. B 671, 391 (2009).
[3] I. Cholis, D. P. Finkbeiner, L. Goodenough, and N. Weiner, J. Cosmol. Astropart. Phys. 12 (2009) 007.
[4] I. Cholis, G. Dobler, D. P. Finkbeiner, L. Goodenough, and N. Weiner, Phys. Rev. D 80 123518 (2009).
[5] P. Meade, M. Papucci, A. Strumia, and T. Volansky, Nucl. Phys. B 831, 178 (2010).
[6] F. Chen, J. M. Cline, and A. R. Frey, Phys. Rev. D 80, 083516 (2009).
[7] G. Kane, R. Lu, and S. Watson, Phys. Lett. B 681, 151 (2009).
[8] M. Kuhlen, J. Diemand, and P. Madau, arXiv:0805.4416.
[9] M. Cirelli, M. Kadastik, M. Raidal, and A. Strumia, Nucl. Phys. B 813, 1 (2009).
[10] J. L. Feng, M. Kaplinghat, and H.-B. Yu, Phys. Rev. D 82, 083525 (2010).
[11] J. M. Cline, A. C. Vincent, and W. Xue, Phys. Rev. D 81, 083512 (2010).
[12] M. Cirelli, P. Panci, and P. D. Serpico, Nucl. Phys. B 840, 284 (2010).
[13] A. Strumia, Prog. Theor. Phys. Suppl. 180, 128 (2009).
[14] M. Cirelli and J. M. Cline, Phys. Rev. D 82, 023503 (2010).
[15] J. Bovy, Phys. Rev. D 79, 083539 (2009).
[16] M. Kuhlen, arXiv:0906.1822.
[17] M. D. Kistler and J. M. Siegal-Gaskins, Phys. Rev. D 81, 103521 (2010).
[18] S. Ando, Phys. Rev. D 80, 023520 (2009).
[19] M. Kuhlen, P. Madau, and J. Silk, Science 325, 970 (2009).
[20] G. Hutsi, A. Hektor, and M. Raidal, J. Cosmol. Astropart. Phys. 07 (2010) 008.
[21] M. Kuhlen and D. Malyshev, Phys. Rev. D 79, 123517 (2009).
[22] P. Brun, T. Delahaye, J. Diemand, S. Profumo, and P. Salati, Phys. Rev. D 80, 035023 (2009).
[23] A. W. Strong and I. V. Moskalenko, Astrophys. J. 509, 212 (1998).
[24] M. Simet and D. Hooper, J. Cosmol. Astropart. Phys. 08 (2009) 003.
[25] galprop.stanford.edu.
[26] T. A. Porter and A. W. Strong, arXiv:astro-ph/0507119.
[27] T. Delahaye et al., Astron. Astrophys. 501, 821 (2009).
[28] J. Diemand et al., Nature (London) 454, 735 (2008).