Relic density and CMB constraints on dark matter annihilation with Sommerfeld enhancement

Jesús Zavala,1 Mark Vogelsberger,1,2 and Simon D. M. White1

1Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85740 Garching bei München, Germany
2Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

We calculate how the relic density of dark matter particles is altered when their annihilation is enhanced by the Sommerfeld mechanism due to a Yukawa interaction between the annihilating particles. Maintaining a dark matter abundance consistent with current observational bounds requires the normalization of the s-wave annihilation cross section to be decreased compared to a model without enhancement. The level of suppression depends on the specific parameters of the particle model, with the kinetic decoupling temperature having the most effect. We find that the cross section can be reduced by as much as an order of magnitude for extreme cases. We also compute the $\mu$-type distortion of the CMB energy spectrum caused by energy injection from such Sommerfeld-enhanced annihilation. Our results indicate that in the vicinity of resonances, associated with bound states, distortions can be large enough to be excluded by the upper limit $|\mu| \leq 9.0 \times 10^{-5}$ found by the COBE/FIRAS experiment.

PACS numbers: 95.35.+d, 98.80.Es

I. INTRODUCTION

If dark matter annihilates, the byproducts of the annihilation (positrons, neutrinos, gamma-rays, etc.) can leave non-gravitational signatures that, if observed, would be crucial for clarifying the nature of dark matter.

In recent years, a number of observations have highlighted anomalies that might be explained by invoking dark matter annihilation in our Galactic halo. Among them are: (i) an anomalous abundance of positrons in cosmic rays above 10 GeV according to the PAMELA experiment [1] confirming and extending previous measurements by experiments such as AMS-01 [2]; (ii) an excess of microwave emission from the galactic center as measured by the WMAP experiment, known as the “WMAP haze” [3]; (iii) the apparent excess of diffuse galactic gamma-rays with energies above 1GeV inferred from observations by the EGRET satellite [4]; (iv) analyses on the balloon experiments ATIC and PPB-BETS [5, 6] have reported an excess in the total flux of electrons and positrons in cosmic rays. For the last two anomalies, we note that recent observations by FERMI seem inconsistent with the claimed gamma-ray excess in the EGRET data [7], and that the excess in the electron and positron flux is smaller than previously thought [8]. The latter is actually consistent with a modified cosmic ray propagation model that does not require additional primary sources of electrons and positrons [9].

Although other astrophysical sources could explain these anomalies (see for example [10, 12] for an explanation for the PAMELA excess based on particle acceleration by pulsars, and [13, 14] for one based on supernova remnants), dark matter annihilation offers an attractive solution. Large annihilation rates are needed, however, to explain the observations. In particular for the PAMELA data, the required annihilation rate is typically a few orders of magnitude larger than the value obtained by assuming the standard cross section inferred from the observed abundance of dark matter together with a smooth local distribution of dark matter [15]. Thus, an additional hypothesis is needed to boost the annihilation rate to the required levels and this must not change the present-day abundance of dark matter.

Such a boost is difficult to obtain from the effects of substructures in the local dark matter distribution. Recent numerical simulations predict this to be remarkably smooth [16]. A detailed analysis of the impact of substructure on the production of positrons by [17] came to a similar conclusion. The possibility of a nearby “spike” of dark matter produced by an intermediate mass black hole seems also a priori implausible [18].

An alternative that has produced a plethora of papers in recent years is that of a Sommerfeld enhancement to the cross section produced by the mutual interaction of the annihilating dark matter particles. Such an interaction could be produced by a force carrier which might be any of the standard model weak force gauge bosons [19] or a new force carrier [20]. For certain values of the parameters of these models, the enhancement is easily large enough to boost the cross section to the required values.

However, a large cross section has a significant impact on other observables and may violate other constraints. For instance, [21] showed that for the case where the cross section increases as $1/v$ (a particular case of Sommerfeld-enhancement models), where $v$ is the relative velocity of the annihilating particles, there are severe constraints from measurements of the diffuse extragalactic gamma-ray background radiation and from CMB constraints on ionization and heating of the intergalactic medium (IGM) by annihilation in the first generation of halos. In this scenario, the boost factors required to fit
the above anomalies would be inconsistent with current constraints. This problem is avoided, however, by more general cases of the Sommerfeld enhancement where the effect saturates at low velocities [19, 20].

Recently, [23] and [24] analyzed constraints on the annihilation cross section from perturbations to the CMB angular power spectra resulting from heating and ionization of the photon-baryon plasma at recombination. They found interesting upper limits to models with Sommerfeld enhancement (see fig. 5 of [22]) that already rule out some extreme cases.

At higher redshifts, the effects of the Sommerfeld enhancement have scarcely been treated. For example, the annihilation of dark matter impacts the predictions from big-bang nucleosynthesis on the abundance of light elements (e.g. [24]). This could put constraints on certain models with Sommerfeld enhancement. However, this has only been studied in passing [25, 26].

Also, the abundance of dark matter today is commonly assumed to be unaltered by this effect, apparently because the typical dark matter particle velocities at freeze-out are very large making the enhancement very close to one at that epoch [23, 27]. As we will show, this reasoning is flawed.

The thermodynamic equilibrium between matter and radiation in the early Universe would be perturbed by energy released during a certain process, dark matter annihilation for example. This equilibrium tends to be restored by different interaction mechanisms: Compton scattering, double Compton emission and bremsstrahlung radiation. The efficiency of these to fully restore equilibrium varies with redshift. For \( z \gtrsim 2 \times 10^6 \) they are efficient enough to restore distortions in the energy spectrum and thus, the photon distribution is that of a black body with a slightly higher temperature than the one in the case of no energy injection. For lower redshifts these mechanisms can not restore the black body spectrum. In particular, for \( 5.1 \times 10^3 \lesssim z \lesssim 2 \times 10^6 \), the spectrum is perturbed into a Bose-Einstein distribution with a chemical potential \( \mu \) [28].

In this paper, we revisit the impact of the Sommerfeld enhancement on the relic particle abundance and we show that its effect is not negligible. We also study, for the first time, \( \mu \)-type distortions of the CMB spectrum due to energy deposition by dark matter annihilation in models with Sommerfeld enhancement.

The paper is organized as follows. In section 2 we summarize the Sommerfeld enhancement and describe how we include it in our calculations. The relic density calculation is set out in detail in section 3. In section 4, the \( \mu \)-type distortion to the CMB from annihilation with Sommerfeld enhancement is calculated. Finally we present a summary and our conclusions in section 5.

II. THE SOMMERFELD ENHANCEMENT

The Sommerfeld enhancement of dark matter annihilation is a nonrelativistic quantum effect occurring when annihilating particles interact through a potential of some sort. If their kinetic energy is low enough, their wave function is distorted, and a significant enhancement (or suppression if the force is repulsive) to the annihilation cross section occurs [19, 21, 29]. The product of the relative velocity times the annihilation cross section will therefore be boosted: \( \sigma v = S(\sigma v)_0 \), where \( (\sigma v)_0 \) is the standard product of the tree level cross section times the relative velocity and \( S \) is the so-called Sommerfeld boost.

The simplest interaction is that for an attractive Yukawa potential with coupling strength \( \alpha_c \) mediated by a scalar boson of mass \( m_\phi \) (more complicated models exist where there are multiple force carriers [20]). In this case, the Sommerfeld enhancement can be computed by solving the radial Schrödinger equation for s-wave annihilation in the non-relativistic limit (we use natural units \( \hbar = c = 1 \)):

\[
\frac{1}{m_\chi} \frac{d^2\Psi(r)}{dr^2} + \frac{\alpha_c}{r} e^{-m_\phi r} \Psi(r) = -m_\chi \beta^2 \Psi(r)
\]

(1)

where \( m_\chi \) and \( \beta \) are the mass and velocity of the dark matter particle, respectively, and \( \Psi(r) \) is the reduced two-body wave function. Eq. (1) should be solved with the boundary condition: \( d\Psi/dr = im_\chi \beta \Psi \) as \( r \to \infty \). The Sommerfeld boost \( S \) is simply given by:

\[
S = \frac{||\Psi(\infty)||^2}{||\Psi(0)||^2}
\]

(2)

In the limit of a massless force carrier (\( m_\phi \to 0 \)), the potential reduces to a Coulomb interaction and Eq. (1) can be solved analytically yielding a value of \( S \) that is independent of the dark matter particle mass:

\[
S = \frac{\pi \alpha_c}{\beta} \left( 1 - e^{-\pi \alpha_c / \beta} \right)^{-1}
\]

(3)

In this limit, and for small velocities \( (\beta \ll \pi \alpha_c) \), \( S \sim \pi \alpha_c / \beta \). This is the reason why the Sommerfeld enhancement is usually associated with a “1/\( \nu \)” enhancement.

In the more general case, where \( m_\phi \neq 0 \), the “1/\( \nu \)” behavior is no longer valid for very small velocities because the Yukawa interaction has a finite range. This can also be shown by expanding the exponential term in Eq. (1) in powers of \( m_\phi r \), then the necessary condition to recover the Coulomb-like interaction solution (Eq. (3)) is \( \beta^2 \gg \alpha_c m_\phi / m_\chi \) [19], which, for fixed values of \( \alpha_c \) and \( m_\phi / m_\chi \), will not be fulfilled for arbitrarily low velocities. In the opposite regime, for \( \beta^2 \ll \alpha_c m_\phi / m_\chi \), a series of resonances appear for specific values of \( \alpha_c m_\chi / m_\phi \). These resonances are associated with bound states, the so-called WIMPonium [30, 31]. In them, the enhancement greatly increases with smaller relative velocity, as \( S \sim 1/\beta^2 \). For even lower velocities and for fixed values
of \( \alpha_c \) and \( m_\phi/m_\chi \), the enhancement saturates reaching a value that can be estimated by solving Eq. (1) with \( \beta = 0 \). Away from the resonances, an order of magnitude estimate of this value is \( S_{\text{max}} \sim 6\alpha_cm_\chi/m_\phi \), reached for a threshold velocity of \( \beta \sim 0.5m_\phi/m_\chi \).

On the other hand, for large velocities \( (\beta \sim \pi\alpha_c) \), the Coulomb approximation is valid and the enhancement is effectively reduced to values very close to 1.

To include in detail the influence of the Sommerfeld enhancement in our study we need to solve numerically the Schrödinger equation.

### III. RELIC DENSITY CALCULATION

The evolution of the phase space distribution per particle of a given species is given by the Boltzmann equation. For Majorana fermions that are stable, for example, neutralinos if they are the Lightest Supersymmetric Particle, the Boltzmann equation is [e.g. 32]:

\[
\frac{dn_\chi}{dt} + 3Hn_\chi = -\langle \sigma v \rangle \left( n_\chi^2 - (n_\chi^{\text{EQ}})^2 \right) \tag{4}
\]

where \( H = \dot{a}/a \) is the Hubble parameter, \( \langle \sigma v \rangle \) is the thermally averaged product of the total pair annihilation cross section times the Moller velocity and \( n_\chi^{\text{EQ}} \) is the equilibrium dark matter number density. In the remainder of the paper we assume that dark matter is made of neutralinos.

We solve Eq. (4) by using the standard freeze-out approximation [e.g. 32]. The freeze-out epoch is defined as the time where the annihilation rate is equal to the expansion rate of the Universe (\( \Gamma \sim H \)). In the standard case where \( \langle \sigma v \rangle \) is velocity independent, the annihilation process becomes subdominant after freeze-out(chemical decoupling) and the “relic” abundance of particles is roughly fixed [47].

We define \( \Delta = X - X_{\text{EQ}} \), where \( X \) and \( X_{\text{EQ}} \) are the comoving number density out of and at equilibrium, then we can write Eq. (4) as:

\[
\frac{d\Delta}{dt} + \frac{dX_{\text{EQ}}}{dt} = -a^{-3} \langle \sigma v \rangle \Delta (2X_{\text{EQ}} + \Delta) \tag{5}
\]

Since before freeze-out the number density follows closely the equilibrium solution then: \( \frac{d\Delta}{dt} \ll \frac{dX_{\text{EQ}}}{dt} \), and at freeze-out: \( \Delta(t_f) \simeq cX_{\text{EQ}}(t_f) \), where \( c \) is a constant of order unity [46]. Under these approximations, the value of the freeze-out time can be found by solving the following equation:

\[
\left. \frac{d\ln X_{\text{EQ}}}{dt} \right|_{t_f} \simeq -a^{-3} \langle \sigma v \rangle_t c(2 + c)X_{\text{EQ}}(t_f) \tag{6}
\]

It is convenient to change the variable \( X \) to the ratio \( Y = n_\chi/s \) of the number density to the total entropy density of the Universe. The entropy is dominated by the contribution of relativistic particles \( s = (2\pi^2/45)g_\ast T^3 \), where \( g_\ast \) are the effective degrees of freedom for the entropy density, and \( T \) is the radiation temperature. In thermal equilibrium, the entropy per comoving volume is conserved \( sa^3 = s_0 \sim 2918 \text{ cm}^3 \) (taking \( T_0 = 2.728^\circ \text{K} \) and \( g_\ast S(T_0) \approx 3.915 \)).

It is also useful to change the time variable to \( x = m/T \) using the equation: \( t = 0.301g^{-1/2}m_{\text{pl}}(x/m)^2 \equiv G_s(m,T)x^2/2 \), where \( m_{\text{pl}} = 1.221 \times 10^{19} \text{ GeV} \) is the Planck mass and \( g_s \) are the effective degrees of freedom for the total energy density of the Universe, which are generally a function of temperature. We have taken into account the temperature dependence of \( g_\ast \) in our calculations, whose values vary from \( g_\ast \sim 100 \) at \( T \sim 100 \text{ keV} \) to \( g_\ast \sim 3 \) at \( T \sim 10 \text{ keV} \). This equation is valid during the radiation-dominated era. If we neglect the variations in \( g_\ast(T) \) around the freeze-out temperature [47], we can finally write Eq. (6) as:

\[
\frac{d\ln Y_{\text{EQ}}}{dx} \mid_{x_f} \simeq -G_s(m,T_f)s_f \langle \sigma v \rangle_f c(2 + c) Y_{\text{EQ}}(x_f) x_f \tag{7}
\]

In the non-relativistic limit, the equilibrium solution takes a simple form: \( Y_{\text{EQ}}(x) = 0.145(g_\ast s_x)^{-1}e^{-x} \), where \( g \) is the degeneracy factor for the particle species \( (g = 2 \text{ for neutralinos}) \). Therefore, \( x_f \) is simply given by the solution to the implicit equation:

\[
e^{x_f} \simeq 0.038 \left( \frac{g}{g_{\ast}^{1/2}(T_f)} \right) m_{\text{pl}} c(2 + c) \left( \frac{m_\chi}{x_f} \right) \langle \sigma v \rangle_f \tag{8}
\]

where we have used the approximation \( 3x_f^{-1/2} \ll 1 \), since we expect \( x_f \gg 20 \).

In the non-relativistic limit, the thermal average \( \langle \sigma v \rangle \) reduces to an average over a Maxwell-Boltzmann distribution function [32]:

\[
\langle \sigma v \rangle = \frac{x^{3/2}}{2\pi^{1/2}} \int_0^1 (\beta^2 e^{-x \beta^2/2}) d\beta \tag{9}
\]

where \( v \) is the relative velocity of the annihilating particles. The expansion of \( \langle \sigma v \rangle \) in powers of \( \beta^2 \) leads to the usual expansion of the non-relativistic thermal average in powers of \( x \), where the first term of the expansion corresponds to the constant s-wave annihilation term \( \langle \sigma v \rangle_S \).

For the case when the cross section is enhanced by the Sommerfeld mechanism the s-wave annihilation thermal average is given by:

\[
\langle \sigma v \rangle = \langle \sigma v \rangle_S \left( \frac{x^{3/2}}{2\pi^{1/2}} \int_0^1 S(\beta) \beta^2 e^{-x \beta^2/2} d\beta \right) = \langle \sigma v \rangle_S S(x) \tag{10}
\]

where \( S(x) \) is the thermally averaged Sommerfeld enhancement. The thermal average maintains the qualitative behavior of \( S \) with velocity. For example, when \( S(\beta) \propto \beta^{-1} \), \( S(x) \propto x^{1/2} \propto \sigma_c^{-1} \), where \( \sigma_c \) is the velocity dispersion of the dark matter particles [33]. We have omitted the notation of \( S(x) \) the dependence on the parameters of the Yukawa interaction; in general \( S \equiv S(x, \alpha_c, m_\phi/m_\chi) \).
At freeze-out, although the dark matter particles are already non-relativistic, the typical velocities are still very large, and we can safely assume $S(\beta) \sim 1$, i.e., $\langle \sigma v \rangle_f \sim \langle \sigma v \rangle_S$. We have checked that this is indeed a good approximation by solving fully the Schrödinger equation, there is no significant change in the results that follow.

The dark matter relic density can be calculated by solving the Boltzmann equation in the late times regime ($t > t_f$) where the number density at equilibrium $n_\chi^E$ is considerably lower than the out-of-equilibrium solution and can therefore be dropped from the Boltzmann equation. Thus, Eq. (11) was written as a function of the variable $x = m_\chi / T$ as [e.g. 32]:

$$\frac{1}{Y(x_0)} = \frac{1}{Y(x_f)} \sqrt{\frac{\pi}{45}} \frac{m_\chi m_{Pl} \langle \sigma v \rangle_S}{x_0^3} \int_{x_f}^{x_0} \frac{g_s^{1/2}(x)S(x)}{x^2} \, dx$$

(11)

where we have used Eq. (10), and the limits of integration corresponding to freeze-out $x_f = m_\chi / T_f$ and present-day values $x_0 = m_\chi / T_0$. The value of $Y(x_0)$ is connected to the value of the ratio of the dark matter density to the critical density today ($\Omega_\chi,0 = m_\chi n_{\chi,0} / \rho_{crit,0}$):

$$\Omega_\chi,0 h^2 = \frac{8\pi}{3M_{Pl}^2 (H_0 / h)^2} m_\chi S(x_0) Y(x_0)$$

$$\approx 2.757 \times 10^8 \left( \frac{m_\chi}{\text{GeV}} \right) Y(x_0)$$

(12)

The integral in the r.h.s of Eq. (11) can be divided in two. Before kinetic decoupling, the dark matter particles are still coupled to the photon-baryon plasma through scattering with standard model particles and therefore we can treat the temperatures of the radiation and of the neutralinos interchangeably ($T_\chi = T$) for $x < x_{kd} = m_\chi / T_{kd}$, where $T_{kd}$ is the temperature at kinetic decoupling.

After kinetic decoupling, the temperature of the dark matter particles drops as $a^{-2}$ instead of $a^{-1}$, as the temperature of radiation does. Thus, for this second part of the integral in Eq. (11), one must be careful in the definition of the variable $x$ in the thermal average (Eq. 18) and in the Boltzmann equation (Eq. 11). The thermal average uses a Maxwell-Boltzmann distribution function with the temperature of the dark matter as a function $S(x_\chi)$, whereas Eq. (11) was written as a function of $x = m_\chi / T$, where $T$ is the radiation temperature. We can easily account for this by evaluating $S(x_\chi)$ in Eq. (10) as a function of the radiation temperature by using the relation: $x_\chi = x^2 / x_{kd}$. As pointed out by [44], in this regime ($x > x_{kd}$), for the case where $S(x_\chi) \propto x_\chi^{1/2} \propto x$, the relic density decays logarithmically: $Y \propto 1 / \text{ln} x$. In section 2 we mentioned that for particular values of $\alpha_c$ and $m_\phi / m_\chi$, the Sommerfeld enhancement has resonances that for very low velocities produce very large values of $S$, in these cases, $S(x_\chi) \propto x_\chi \propto x^2$, thus, the relic density decreases as $1 / x$.

However, these large suppressions of the relic density are eventually cut-off by the saturation of $S$, which occurs before matter-radiation equality. Consider, for example, a 100 GeV neutralino. Since $z_E \approx 4 \times 10^{10}$, and because $T_\chi \sim (z_E / z_{kd})^2$ with $z_{kd} \sim 4 \times 10^{10}$ (the exact value of $z_{kd}$ depends on the specific supersymmetric model, this is a typical value for a neutralino of this mass), we have a dark matter velocity dispersion of the order of $10^{-8}$ by $z = z_E$, which is sufficiently low to reach saturation, unless the combination of $\alpha_c$ and $m_\phi / m_\chi$ is fine-tuned to be close to a resonance.

In summary, a proper calculation of the relic density needs to include the effects of kinetic decoupling. We do so by choosing a model with neutralino dark matter, and by using the results of [35], who made a scan of the parameter space in mSUGRA (minimal supergravity) and more general MSSM models, and give a range for the values of the kinetic decoupling temperature as a function of neutralino mass (see fig. 2 of their paper). Typical values for $m_\chi \in (100, 5 \times 10^6)$ GeV are in the range $x_{kd} \in (200, 2 \times 10^4)$.

For a fixed value of $\Omega_\chi h^2$, we can solve Eqs. (8) and (11) simultaneously to get the values of $x_f$ and $\langle \sigma v \rangle_S$ that give the correct relic abundance. The value of the present-day dark matter density according to the WMAP-5 year data, at the 2σ level, is $\Omega_\chi h^2 = 0.1143 \pm 0.0068$ [36]. Fig. 1 shows the values of $\langle \sigma v \rangle_S$ consistent with $\Omega_\chi h^2 = 0.1143$ for a scan in the parameter space $(m_\phi / m_\chi, \alpha_c)$ with $m_\chi = 100$ GeV and $T_{kd} = 8$ MeV. The values of the thermally averaged cross section are color-coded in units of $1 \times 10^{-26} \text{cm}^2 \, \text{s}^{-1}$. For compari-
son, in the absence of any Sommerfeld enhancement the required cross section is $\sim 2.4 \times 10^{-26} \text{cm}^3\text{s}^{-1}$.

Since the typical velocities of the dark matter particles between $v_F$ and $v_{kd}$ are still very high, the Coulomb regime is valid in this range for most of the parameter space (recall that the Yukawa part of the potential can be ignored when $\beta^2 \gg \alpha_c m_\phi/m_\chi$), thus, if we cut the integral in the r.h.s of Eq. (11) at $x = x_{kd}$, the relic density calculation would depend only in the strength of the coupling $\alpha_c$. This dependence can be seen in the overall trend from left to right in Fig. 1. However, after kinetic decoupling, the WIMPs get colder much faster than the radiation and the Sommerfeld enhancement depends not only on $\alpha_c$ but also on $m_\phi/m_\chi$. This is particularly significant near resonances where the enhancement is large enough to suppress the dark matter abundance by a factor of a few, therefore, the cross section normalization needs to be suppressed by this same factor in order to be consistent with the observed relic density.

As shown by the discussion in the previous paragraph, the particular value of the kinetic decoupling temperature has an impact on Fig. 1. The closer this temperature is to the freeze-out temperature, the stronger the impact will be. To appreciate this better, we show in Fig. 2 the ratio of two maps analogous to the map in Fig. 1 but for two extreme values of the kinetic decoupling temperature for $m_\chi = 100 \text{ GeV}$. $T_{kd} = 150 \text{ MeV}$ and $T_{kd} = 4 \text{ MeV}$. For large values of the coupling strength and low values of $m_\phi/m_\chi$, and particularly near resonances, the cross section at freeze-out needs to be lower for a higher kinetic decoupling temperature to produce the correct relic density. From the figure we see that uncertainties in the kinetic decoupling temperature introduce changes to $\langle \sigma v \rangle_S$ at the most of order 30% for a 100 GeV neutralino.

Since current uncertainties on the abundance of dark matter are in the percent level, the impact of the particular value of $\Omega_\chi h^2$ on Fig. 1 is not so large. By exploring changes in $\Omega_\chi h^2$ within the 2$\sigma$ bounds we found changes in the values of $\langle \sigma v \rangle_S$ of 20% at the most.

These results are also not very sensitive to changes in the neutralino mass, because the relic density is nearly independent of it. This can be seen by looking at Eqs. (11) and (12). If we neglect the term $1/Y(x_F)$, then, for a fixed value of $\langle \sigma v \rangle_S$, $\Omega_\chi h^2$ depends on the neutralino mass just through the integral in the r.h.s of Eq. (11). Moreover, since the Sommerfeld enhancement only depends on the neutralino mass through the ratio $m_\phi/m_\chi$, a change on $m_\chi$ can always be compensated by choosing a different mass for the force carrier. The remaining dependence is through the values of $x_F$ and $x_{kd}$.

Of course, to get the precise impact on the value of the cross section required for a particular particle model, it is necessary to do the calculation for that model. Uncertainties on the kinetic decoupling temperature and on the abundance of dark matter will have a minor-to-mild impact on the quantitative results. However, the general behavior and order of magnitude of the expected values of $\langle \sigma v \rangle_S$ can be directly read off from Fig. 1.

In summary, the relevant difference between the relic density calculation for the standard case where $\langle \sigma v \rangle$ is constant and the case we study here is given by the onset of the Sommerfeld enhancement after kinetic decoupling. Before this phase, the velocities of dark matter particles are too large for the enhancement to be relevant, but once the decoupling happens the typical velocities decrease quickly since the temperature of the dark matter fluid drops as $a^{-2}$. This causes a rapid increase on the enhancement values and the annihilation process takes a relevant role once more decreasing the dark matter density. On the contrary, the comoving density remains constant after freeze-out for the standard case. The aforementioned effect continues to be important until the enhancement saturates fixing the value of $\langle \sigma v \rangle$ to a maximum. Afterwards, the annihilation rate becomes subdominant compared to the expansion rate of the Universe and thus, the comoving dark matter density remains fixed.

Owing to this difference, the value of the annihilation cross section before the freeze-out epoch needs to be smaller for the case with Sommerfeld enhancement than for the standard one in order to produce the observed dark matter density. With a lower annihilation rate, the dark matter particles go through chemical decoupling at earlier times, and therefore the dark matter density is higher at these epoch than in the standard case. This higher value compensates the subsequent decrease of the density due to the Sommerfeld enhancement. In this way, the final relic density can be consistent with the expected value for the abundance of dark matter today.
IV. $\mu$-TYPE DISTORTION OF THE CMB SPECTRUM

Measurements of the CMB spectrum by the FIRAS instrument onboard COBE have shown that is nearly that of a perfect blackbody with temperature $T_0 \sim 2.728^{+6}_{-5}$ K. Deviations from a pure black-body spectrum are, however, expected if there is any energy input in the early Universe. If this energy is injected between $5.1 \times 10^4 \lesssim z \lesssim 2 \times 10^6$ (12 eV $\lesssim T \lesssim$ 470 eV), the spectral distortion is a Bose-Einstein $\mu$-type distortion where $\mu$ is the chemical potential [28]. If the number of photons injected during the energy release is small, compared to the number of photons in the radiation plasma, then $\mu \sim 1.4 \delta \rho_\gamma / \rho_\gamma$, where $\delta \rho_\gamma$ is the injected energy and $\rho_\gamma$ is the energy density of the CMB photons.

The current observational limit for the $\mu$-type distortion at the 95% confidence level is $|\mu| \lesssim 9.0 \times 10^{-5}$ [29]. In principle, this limit could already have been improved by nearly two orders of magnitude by now given recent technological advances [38]. In the case of energy deposited by annihilation of neutralinos, the $\mu$-type distortion was studied by [35] for the case of s-wave and p-wave cross sections. More recently it was also mentioned by [10]. The value of $\mu$ is given by:

$$\mu = 1.4 \frac{\delta \rho_\gamma}{\rho_\gamma} = 1.4 \int_{t_1}^{t_2} \frac{\rho_\gamma}{\rho_\gamma} dt = 1.4 \int_{t_1}^{t_2} \frac{m_\chi \langle \sigma v \rangle n_\chi^2}{\rho_\gamma \alpha^4} dt$$

(13)

where, $\rho_\gamma$ is the present-day energy density of the CMB ($\Omega_\gamma h^2 \sim 2.47 \times 10^{-5}$), $t_1 - t_2$ is the time interval corresponding to the energy injection and $f$ is the efficiency in which the injected energy is transformed into heat [38]. In principle, $f$ depends on the channel of annihilation and on time. However at the relevant redshifts ($z \gg 1000$), $f$ is basically given by the annihilation channel [23]. For electrons and photons $f \sim 1$, but for annihilation into $\tau$'s for example, some of the energy is lost in neutrinos. Nevertheless, for all relevant channels, [23] found that $f > 0.25$ for $z > 2500$. We will adopt $f = 1$, our results can be reinterpreted easily for a different value of $f$.

Since the relevant redshift range of energy injection is in the radiation dominated era and after kinetic decoupling, we can write $t = t_{kd}(1 + z_{kd})^2 / (1 + z)^2 = t_{kd} x_\chi / x_{kd}$. It is also convenient to write Eq. (13) in terms of the variable $x_\chi = m_\chi / T_\chi$:

$$\mu = 1.4 f \left( \frac{\langle \sigma v \rangle_S}{m_\chi \rho_\gamma}, \frac{t_{kd}}{x_{kd}(1 + z_{kd})^2}, \right) \int_{x_{kd}^1}^{x_{kd}^2} S(x_\chi) x\chi^{-1} \rho_\chi(x_\chi) dx_\chi$$

(14)

where:

$$x_{kd}^{1,2} = \frac{m_\chi T_{kd}}{T_0^2 (1 + z_{1,2})^2}$$

(15)

FIG. 3: Scan of the parameter space ($m_\phi/m_\chi$, $\alpha_c$) with the expected values of the $\mu$-type distortion to the CMB spectrum for $m_\chi = 100$ GeV, $T_{kd} = 8$ MeV and cross section values satisfying the constraint in the relic density: $\Omega_{\chi} h^2 = 0.1143$. The 2$\sigma$ observational upper limit on $|\mu|$ is $9 \times 10^{-5}$. The plot shows the values of the logarithm of $\mu$ color-coded according to the scale on the right.

The density of neutralinos changes with time (temperature through the variable $x_\chi$) according to the Boltzmann equation. Its value at a given radiation temperature is related to the ratio $Y$ given by Eq. (11) (recall that $\rho_\chi = m_\chi S Y$) by replacing $x_0$ with the corresponding $x = m_\chi / T$ and using the relation between the radiation temperature and that of the neutralino gas:

$$\rho_\chi(x_\chi) = m_\chi s(x = (x_{kd} x_\chi)^{1/2}) Y(x = (x_{kd} x_\chi)^{1/2})$$

(16)

Note that for the redshift range of interest $T < 470$ eV and thus $g_{s,Y}$, which enters in the value of the entropy $s(x)$, is a constant equal to $g_{s,Y}(x_0) = 3.915$.

Now, $t_{kd}(1 + z_{kd})^2 = t_E(1 + z_E)^2 = t_0(1 + z_E)^{1/2}$, where $z_E$ is the redshift of matter and radiation equality and $t_0$ is the age of the Universe. In the last equality we have used the relation between the scale factor and time for an Einstein-de Sitter Universe. Since $(1 + z_E) = \Omega_{\chi,0}/\Omega_{\gamma,0}$, we can finally write:

$$\mu = 1.4 f \left( \frac{\Omega_{\chi,0} h^2}{\Omega_{\gamma,0} h^2}, \frac{\langle \sigma v \rangle_S}{m_\chi \rho_\gamma}, \frac{t_0}{x_{kd}^{1/2}} \right) \int_{x_{kd}^1}^{x_{kd}^2} S(x_\chi) x\chi^{-1} \left( \frac{Y(x = (x_{kd} x_\chi)^{1/2})}{Y(x_0)} \right)^2 dx_\chi$$

(17)

Note that the function in parentheses inside the integral is equal to one if the comoving density of neutralinos is no longer evolving (has frozen) by $z = 2 \times 10^8$. This is
the case in the majority of the parameter space of the Yukawa interaction because by this redshift the velocity dispersion of dark matter particles is low enough for the Sommerfeld enhancement to be saturated. As we mentioned before, this means that the comoving density is essentially frozen afterwards.

Fig. 4 shows the values of $\mu$ found by solving Eq. (17) for the same scan of the parameter space $(m_\phi/m_\chi, \alpha_c)$ as in previous figures. We took as fiducial values $m_\chi = 100$ GeV, $T_{kd} = 8$ MeV and $\Omega_\chi h^2 = 0.1143$, that is, the values of the normalization to the cross section are required to give the observed dark matter abundance. The values of the logarithm of $\mu$ appear color-coded according to the scale in the right of the figure. The resonances inherited from $S(x)$ and shown in Fig. 1 are also clear in this figure. Most of the region on the lower right corner is above, or very close to, the observational upper limit $|\mu| \leq 9 \times 10^{-5}$. Points of the parameter space very near resonances also give values of $\mu$ that can be excluded. To show in more detail the behavior near a resonance, we present in Fig. 4 a plot of the values of $\mu$ as a function of the perpendicular distance, in parameter space, to the first resonance in the upper left of Fig. 3. Lines with different colors show the magnitude of the change produced by different values of the kinetic decoupling temperature and of $\Omega_\chi h^2$, corresponding to the legends in the figure. The upper limit on $\mu$ according to the COBE/FIRAS experiment is also marked as a horizontal line. The figure shows how close one can get to a resonance without violating the current constraint.

This is a significant point, because values very close to a resonance are often invoked to provide a large boost and so to match observations of the positron excess in cosmic rays.

According to Eq. (17), the value of $\mu$ decreases with increasing mass ($\mu \propto 1/m_\chi$) because $(\sigma v)_S$ is nearly independent of mass, and because, except very near resonances, $S(x_\chi)$ is already saturated by $z = 2 \times 10^6$. In this way, our results can be easily scaled to the desired neutralino mass. Care is needed very close to the resonances since the simple scaling $\mu \propto 1/m_\chi$ is no longer valid there.

Finally, we show in Fig. 5 the analogous result for the values of $\mu$ in the case where the interaction responsible of the Sommerfeld enhancement is Coulomb-like ($m_\phi \to 0$). In this case the enhancement depends only on $\alpha_c$. The plot shows, for the same fiducial case ($m_\chi = 100$ GeV, $T_{kd} = 8$ MeV and $\Omega_\chi h^2 = 0.1143$), that all cases with $\alpha_c \gtrsim 6 \times 10^{-2}$ are already excluded.

V. SUMMARY AND CONCLUSIONS

The prospects for dark matter detection have increased considerably in recent years. There is a continual advance in the sensitivity of detectors on Earth that look for direct dark matter elastic scattering with nuclei, and in experiments that search indirectly for dark matter by

![FIG. 4: The values of the $\mu$ distortion parameter near the first resonance in the upper left corner of Fig. 3 as a function of the perpendicular distance to the resonance. The lines with different colors correspond to different values of $T_{kd}$ and $\Omega_\chi h^2$ as marked in the legend. The solid black horizontal line shows the upper limit on $\mu$ according to the COBE/FIRAS experiment. As in Fig. 3 $m_\chi = 100$ GeV.](image)

![FIG. 5: The same as Fig. 4 but for the case where the interaction is strictly Coulomb-like. We have taken fiducial values of $m_\chi = 100$ GeV, $T_{kd} = 8$ MeV and $\Omega_\chi h^2 = 0.1143$. Values above the black solid horizontal line are excluded by the current limits of $\mu$ according to the COBE/FIRAS experiment.](image)
looking for non-gravitational signatures of the byproducts of its hypothetical annihilation. Such technological improvements may lead, in the near future, to a definite proof of the existence of dark matter.

Perhaps dark matter has produced non-gravitational signals that have already been detected. The recently reported excesses of positrons in cosmic rays by PAMELA and of electrons+positrons by FERMI/ATIC/PPB-BETS can be explained by dark matter annihilation. Although other explanations with a different astrophysical origin are also possible, a solution based on dark matter is an attractive possibility.

However, such solution seems to require an attractive force between the dark matter particles to enhance their annihilation through the Sommerfeld mechanism. For a Yukawa interaction via a single scalar, the magnitude of the enhancement depends on the coupling strength of the interaction $\alpha_c$, on the mass ratio of the force carrier to the dark matter particle $m_\phi/m_\chi$ and on the relative velocities of the annihilation particles $\beta$.

Several recent papers have computed the boost to the annihilation due to a Sommerfeld enhancement (in the Galactic halo and/or in its subhalos) [19, 20, 33]. One of the main aims of these works is to show that for certain values of $\alpha_c$ and $m_\phi/m_\chi$, this mechanism able to produce large enough boosts to explain the cosmic ray anomalies.

Only a handful of studies have addressed in detail the impact that such an enhancement has in the early Universe. In [21], the authors found that if the cross section increases with decreasing relative velocity as $1/\beta$ (which is valid in a certain regime for general Sommerfeld enhancement models), dark matter annihilation in the first halos would heat and ionize the IGM, violating current constraints from the CMB. As was later pointed out by [19] and [20], this problem is alleviated in more general models where the enhancement saturates at low velocities. In [22] and [23], a constraint on the annihilation cross section was obtained by considering limits on the energy deposition by annihilation at recombination. The constraint reported by [23] is $\langle \sigma v \rangle_{\text{REC}} < 3.6 \times 10^{-23} \text{cm}^3\text{s}^{-1}(m_\chi/1\text{TeV})/f_{\text{REC}}$, where $f_{\text{REC}}$ is an average efficiency of energy injection into the IGM by annihilation at recombination.

In the present paper, we have analyzed the impact of dark matter annihilation with Sommerfeld enhancement at higher redshift.

In the first place, we have revisited the calculation of the dark matter particle abundance by solving the Boltzmann equation from freeze-out through the epoch of kinetic decoupling, including a full solution to the Schrödinger equation. Contrary to previous claims [20, 27], we have found a significant suppression of the relic density, in agreement with a recent work by [34]. This suppression is particularly important near resonances, which are typically invoked to explain the cosmic ray anomalies.

We found that to fit the observed dark matter abundance, the normalization of the cross section needs to be lowered by up to a factor of 10 compared to the case without enhancement (see Fig. 1). The result depends on the coupling strength and the proximity to a resonance. Exploring a broad range of dark matter particle masses and kinetic decoupling temperatures, we found a minor-to-medium impact of these parameters on our results; variations on $T_{kd}$ have the strongest impact.

Secondly, we have calculated the amount of energy deposited by dark matter annihilation into the radiation plasma in the redshift range $5.1 \times 10^4 < z < 2 \times 10^6$. Energy injection at this epoch would create a distortion in the CMB energy spectrum first pointed out by [28]. The COBE/FIRAS experiment has put constraints on this Bose-Einstein $\mu$-type distortion. The upper limit on the chemical potential associated to the distortion is $|\mu| \leq 9.0 \times 10^{-5}$ at the $2\sigma$ level [37]. We have found that very near resonances, annihilation with Sommerfeld enhancement is already ruled out by this constraint on $\mu$ (see Fig. 3). Quantitatively, the “safe” range of proximity to a resonance depends on the particular model, higher values of dark matter mass and kinetic decoupling temperature allow a closer proximity to the resonance (see Fig. 4). In the case where the force carrier is massless, the Yukawa interaction reduces to a Coulomb one. Our findings are much more stringent in this case with values of $\alpha_c > 6 \times 10^{-2}$ already ruled out (see Fig. 5).

Improved upper limits on the $\mu$ parameter by a null distortion detection in the CMB spectrum would rule out a larger region of the parameter space for the Yukawa interaction. In [38], it was pointed out that an improvement close to two orders of magnitude has already been possible for a number of years, and [41] suggests that another order of magnitude is perhaps within reach. An upper limit on $|\mu|$ of the order of $10^{-7}$ would certainly exclude large regions of the parameter space. It would exclude to a large extent near-resonance regions, which are the ones that produce the largest boosts to the annihilation. On the other hand, the detection of a distortion could possibly tell us something about the parameters of the Yukawa interaction and the nature of the force carrier $\phi$. In summary, our results indicate that for a given set of parameters $(m_\chi, T_{kd}, m_\phi, \alpha_c)$, it is necessary to compute in detail the relic density to obtain the range of values for the normalization to the cross section that are compatible with current estimates of the dark matter abundance. Once this normalization is known, the energy input producing a $\mu$-type distortion in the CMB spectrum can be computed to check whether the particular model violates current constraints. Only for allowed models can a boost factor for local dark matter annihilation be computed and advocated.

Our findings show that the local boosts reported in the literature need to be renormalized to the proper value of $\langle \sigma v \rangle_{\phi}$ implied by the observed relic density. This renormalization can exceed a factor of 10 in extreme cases.

A recent analysis of the PAMELA and FERMI data by [42] suggests that boost factors, over a standard value
We find that these boosts are modified once the proper maximum boost factors are achieved only by invoking Sommerfeld-enhanced annihilation. For instance, [20] and [33] obtain maximum boost factors \( \sim 1000 \) by assuming a local Maxwell-Boltzmann velocity distribution with a velocity dispersion of \( \sim 150 \text{ kms}^{-1} \). Typical estimates on such an additional boost factor due to substructure in the galactic halo (albeit without Sommerfeld enhancement) are of the order of 1.4, unless a subhalo happens to be very close to Earth, in which case this boost could be larger than 10 [44]. These estimates do not include a possible further amplification due to the scaling of the Sommerfeld enhancement with velocity dispersion. However, the constraints found by [23] based on energy deposition at recombination suggest that the enhancement must be already close to saturation for \( \sigma_v = 150 \text{ kms}^{-1} \), thus such a further amplification seems implausible. Therefore, it is not clear that the inclusion of colder substructures can account for the additional boost.

In Fig. 6 we show the boost factors, i.e., the multiplicative factor to \( \langle \sigma v \rangle_S = 3 \times 10^{-26} \text{cm}^3\text{s}^{-1} \), that we find for a Maxwell-Boltzmann velocity distribution with \( \sigma_v = 5 \times 10^{-4} \) for the parameter space of the Yukawa interaction. The figure is for \( m_\chi = 100 \text{GeV} \), \( T_{kd} = 8 \times 10^{-3} \text{GeV} \) and is consistent with a relic abundance \( \Omega_\chi h^2 = 0.1143 \). Keep in mind, however, that changes in the neutralino mass have almost no impact on the normalization of the cross section. Different kinetic decoupling temperatures and variations on \( \Omega_\chi h^2 \) have also only a small effect (< 60% in combination).

This result suggests that additional assumptions are needed to account for the boost needed to explain the cosmic ray anomalies by dark matter annihilation alone. The inclusion of colder substructures to the overall smooth component with higher densities and lower velocity dispersions and thus, higher boost factors, could perhaps solve the issue. However, as found in recent high-resolution N-body simulations, the local dark matter distribution is rather smooth [16]. The inclusion of colder substructures to the overall smooth component with higher densities and lower velocity dispersions and thus, higher boost factors, could perhaps solve the issue. However, as found in recent high-resolution N-body simulations, the local dark matter distribution is rather smooth [16].

Acknowledgments

JZ is supported by the Joint Postdoctoral Program in Astrophysical Cosmology of the Max Planck Institute for Astrophysics and the Shanghai Astronomical Observatory. We thank the anonymous referee for helpful comments and suggestions.

[1] O. Adriani et al., Nature (London) 458, 607 (2009), arXiv:0810.4995.
[2] M. Aguilar et al. (AMS-01 Collaboration), Phys. Lett. B 646, 145 (2007), arXiv:astro-ph/0703154.
[3] D. Hooper, D. P. Finkbeiner, and G. Dobler, Phys. Rev. D 76, 083012 (2007), arXiv:0705.3655.
In Eq. (13) we used that the rate of annihilation events in Eq. (4) influences the relic abundance. The specific value of the parameter $c$ in the expected range is motivated by the values obtained in the constant plateau, see for example fig. 1 of [32]. This expected range of values: $c \in (0.1, 3.0)$, the changes are in the percent level. We have adopted $c = 1.0$ in the results reported here.

In the expected range of freeze-out temperatures for massive relics, is roughly in a constant plateau, see for example fig. 1 of [32]. This expected range is motivated by the values obtained in the standard calculations where $\langle \sigma v \rangle = cte$. We used that the rate of annihilation events per unit volume for neutralinos, being Majorana particles, is $f m_\chi/2$ and that the energy deposited in each event is $2 m_\chi$. We have adopted $c = 1.0$ in the results reported here.