Implications of a 21-cm signal for dark matter annihilation and decay

Hongwan Liu[*] and Tracy R. Slatyer[†]

Center for Theoretical Physics, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

Measurements of the temperature of the baryons at the end of the cosmic dark ages can potentially set very precise constraints on energy injection from exotic sources, such as annihilation or decay of the dark matter. However, additional effects that lower the gas temperature can substantially weaken the expected constraints on exotic energy injection, whereas additional radiation backgrounds can conceal the effect of an increased gas temperature in measurements of the 21-cm hyperfine transition of neutral hydrogen. Motivated in part by recent claims of a detection of 21-cm absorption from $z \sim 17$ by the EDGES experiment, we derive the constraints on dark matter annihilation and decay that can be placed in the presence of extra radiation backgrounds or effects that modify the gas temperature, such as dark matter (DM)–baryon scattering and early baryon-photon decoupling. We find that if the EDGES observation is confirmed, then constraints on light DM decaying or annihilating to electrons will in most scenarios be stronger than existing state-of-the-art limits from the cosmic microwave background, potentially by several orders of magnitude. More generally, our results allow mapping any future measurement of the global 21-cm signal into constraints on dark matter annihilation and decay, within the broad range of scenarios we consider.

I. INTRODUCTION

Between thermal decoupling of baryons from the cosmic microwave background (CMB) and star formation at $z \sim 20$, the universe can function as a sensitive calorimeter for exotic sources of energy injection. In the conventional ΛCDM model, the matter temperature T_m and hydrogen ionization fraction as a function of redshift are simple and well understood [1,2]. After recombination, Compton scattering between the residual free electrons and CMB photons keeps T_m at the CMB temperature $T_{\rm CMB}$ down to $z \sim 150$. Subsequently, the energy transfer rate becomes too small to prevent thermal decoupling, and soon after, T_m is determined solely by adiabatic expansion, evolving as $T_m(z) \propto (1+z)^2$. Deviations in temperature from this well-understood standard history are thus a strong indication of new sources of heating or cooling in the universe.

The recent measurement of an absorption profile at 78 MHz in the sky-averaged spectrum by the Experiment to Detect the Global Epoch of reionization Signature (EDGES) Collaboration [3] opens a new window into the cosmic dark ages, shedding new light on the ionization and thermal history at precisely this period of interest. Radiation with a frequency near the hyperfine transition of hydrogen illuminates neutral hydrogen clouds

during this epoch, and gets absorbed as they redshift into the transition frequency. The brightness temperature of the 21-cm hydrogen absorption line relative to the background radiation is given by [4]

$$T_{21}(z) \approx x_{\rm HI}(z) \left(\frac{0.15}{\Omega_m}\right)^{1/2} \left(\frac{\Omega_b h}{0.02}\right) \left(\frac{1+z}{10}\right)^{1/2}$$
$$\times \left[1 - \frac{T_R(z)}{T_S(z)}\right] 23 \text{ mK,} \qquad (1)$$

where $x_{\rm HI}$ is the neutral hydrogen fraction, Ω_m and Ω_b are the matter and baryon energy density as a fraction of the critical density, h is the Hubble parameter today in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and $T_R(z)$ is the effective temperature of the background 21-cm radiation at redshift z. $T_S(z)$, the spin temperature, determines the ratio of neutral hydrogen in the higher-energy spin-triplet state to the lower-energy spin-singlet state.

The expected value of T_S as a function of redshift has been studied extensively (see e.g., [5] for a review). At $z \sim 30$, we expect $T_S = T_R$, with the radiation temperature commonly assumed to be $T_{\rm CMB}$. Once the first stars start forming at $z \sim 20$ and begin to emit UV radiation, downward transitions from the spin-triplet to the spin-singlet state through the Wouthuysen-Field effect [6–8] start to occur, driving the spin temperature toward T_m. The combination of the background 21-cm radiation, UV

---

[*]hongwan@mit.edu
[†]tslatyer@mit.edu

radiation from stars, and collisional hyperfine excitation/deexcitation ensures that well before reionization,

$$T_m \lesssim T_S \lesssim T_R. \tag{2}$$

A measurement of a negative $T_{21}(z)$ at this time indicates that T_S lies below T_R, and also sets an upper bound on T_m if T_R is known.

The EDGES Collaboration measured a strong 21-cm absorption trough in the redshift range $14 < z < 20$, reporting a value of T_{21} at $z \sim 17.2$ of $T_{21} = -500^{+200}_{-500}$ mK [9], with 99% confidence limits specified. This result, together with Eq. (2), sets the following constraint on the matter and radiation temperature at $z = 17.2$ at the 99% confidence level:

$$\frac{T_m}{T_R}(z = 17.2) \lesssim 0.105. \tag{3}$$

Precise calculations of the temperature evolution after recombination assuming the ΛCDM model [1,2] give $T_m(z = 17.2) \sim 7$ K; however, assuming $T_R = T_{\text{CMB}}$ in Eq. (3), we obtain $T_m \lesssim 5.2$ K, which lies well below the expected value.

Since the publication of the EDGES result, this discrepancy has been explained by either a colder-than-expected gas temperature or an additional source of 21-cm photons at $z \sim 20$. In both cases, the effect is to reduce the expected value of the ratio T_m/T_R. Models with interactions between baryons and cold dark matter (DM) with a Rutherford-like cross section have been explored [9] as a mechanism to cool the gas, particularly in the context of millicharged DM models [10–13]. These models have been shown to be highly constrained, with millicharged DM likely to make up only a subdominant component of DM. Modifications to the redshift of thermal decoupling of baryons from the CMB can also result in a cooler-than-expected gas temperature. Such a scenario can occur due to an imbalance between the proton and electron number densities [14] or early dark energy [15] (although the latter scenario appears difficult to reconcile with other observations). The possibility that interacting dark energy or other effects could modify the evolution of the Hubble parameter and change the 21-cm brightness temperature was proposed in [16], but the change to the Hubble parameter required at $z \lesssim 20$ is large. Finally, models which inject additional 21-cm radiation through light DM decays [12,17] or radio emission from black holes [18,19] have been studied as a means of raising T_R.

In any model of DM, the annihilation and decay rates into Standard Model (SM) particles are important quantities to understand. Models with DM-baryon scattering are likely to imply the existence of DM annihilation to SM particles by crossing symmetry, and these annihilation processes could potentially set the relic abundance of DM via thermal freeze-out at early times. Even if DM-baryon scattering does not occur or is not strong enough to markedly affect the matter temperature, new constraints on annihilation and decay can be set using the information on the thermal history provided by 21-cm measurements of this epoch.

Previous studies [20–24] have explored such constraints under the assumption that there are no other modifications to the conventional thermal history. However, any attempt to explain the EDGES result mandates the presence of additional effects, and such modifications could also be present even if the EDGES result is not confirmed.

In this paper, we will study the implications that a confirmed 21-cm absorption measurement from $z \sim 20$ would have for DM annihilation and decay, in conjunction with three general mechanisms that could deepen an absorption signal: (i) nonstandard recombination histories; (ii) baryon-DM scattering; and (iii) an additional source of 21-cm photons at $z \sim 20$. We will use the EDGES result as a benchmark; if it is confirmed, the forecast limits in this work can be applied as constraints on the DM parameter space.

Throughout this paper, all algebraic expressions will be written in natural units with $\hbar = c = k_B = 1$, and we adopt cosmological parameters that are equal to the Planck 2015 TT, TE, EE + lowP central values [25].

II. IONIZATION AND THERMAL HISTORY

We first review the standard thermal and ionization history, and the modifications necessary to include DM energy injection.

The conventional understanding of the thermal and ionization history of the universe during and after recombination is well approximated by the three-level atom (TLA) model first derived in [26,27] (a review of the model is given in e.g., [1]). In this model, the evolution of the matter temperature T_m and the ionization fraction are

$$\dot{T}_m^{(0)} = -2HT_m + \Gamma_C(T_{\text{CMB}} - T_m), \tag{4}$$

$$\dot{x}_e^{(0)} = -\mathcal{C}[n_{\text{H}}x_e^2\alpha_{\text{B}} - 4(1 - x_e)\beta_{\text{B}}e^{-E_{21}/T_{\text{CMB}}}], \tag{5}$$

where H is the Hubble parameter, n_{H} is the total number density of hydrogen (both neutral and ionized), $x_e \equiv n_e/n_{\text{H}}$ where n_e is the number density of free electrons, and $E_{21} = 10.2$ eV is the Lyman-α transition energy. α_{B} and β_{B} are case-B recombination and photoionization coefficients, respectively, and $\mathcal{C}$ is the Peebles-C factor [1,26] that represents the probability of a hydrogen atom in the $n = 2$ state decaying to the ground state before photoionization can occur. Γ_C is the Compton scattering rate, given by

$$\Gamma_C = \frac{x_e}{1 + f_{\text{He}} + x_e} \frac{8\sigma_T a_r T_{\text{CMB}}^4}{3m_e}, \tag{6}$$

where σ_T is the Thomson cross section, a_r is the radiation constant, m_e is the electron mass, and $f_{\mathrm{He}} \equiv n_{\mathrm{He}}/n_{\mathrm{H}}$ is the relative abundance of helium nuclei by number.

DM annihilation or decay into SM particles injects energy into the universe, leading to additional ionization, excitation, and heating of the gas. Given a velocity-averaged s-wave annihilation cross section $\langle \sigma v \rangle$ or a decay lifetime τ (assumed to be much longer than the age of the universe), the rate of energy injection for these processes is given by

$$\left(\frac{dE}{dVdt}\right)^{\mathrm{inj}} = \begin{cases} f_{\chi,\mathrm{ann}}^2 \rho_{\chi,0}^2 (1+z)^6 \frac{\langle \sigma v \rangle}{m_\chi}, & \text{annihilation,} \\ f_{\chi,\mathrm{dec}} \rho_{\chi,0} (1+z)^3 \frac{1}{\tau}, & \text{decay,} \end{cases} \quad (7)$$

where $\rho_{\chi,0}$ is the mass density of DM today, and m_χ is the DM mass. $f_{\chi,\mathrm{ann}}$ and $f_{\chi,\mathrm{dec}}$ are the fractions by mass of the DM that annihilates or decays, respectively; our constraints are presented with these quantities set to 1 unless explicitly stated otherwise, and a straightforward rescaling of $\langle \sigma v \rangle$ or τ will give the limits for other values of these fractions. An additional factor of $1/2$ is required if the annihilation occurs between distinct particles (e.g., Dirac fermions), but we will assume throughout this paper that the two particles in an annihilation are indistinguishable unless otherwise stated: constraints for distinguishable particles can be obtained by a simple rescaling of $\langle \sigma v \rangle$ by a factor of 2.

The effect of the injected energy on the thermal and ionization history may not be instantaneous, and can in fact be significantly delayed: these effects are captured by a deposition efficiency $f_c(z)$ into some energy deposition channel c (ionization, excitation, or heating), which is defined as

$$\left(\frac{dE}{dVdt}\right)_c^{\mathrm{dep}} = f_c(z)\left(\frac{dE}{dVdt}\right)^{\mathrm{inj}}. \quad (8)$$

Values for $f_c(z)$ into the relevant channels have been previously computed in [28], and were extended to include structure formation boosts to the annihilation rate in [29]. We note that $f_c(z)$ is computed assuming the standard ionization history, whereas new sources of energy injection or cooling will no doubt change the ionization history. However, the atomic processes that determine these deposition efficiencies remain relatively unaffected by small changes in the ionization fraction. This is a valid assumption in this study, where we will ultimately constrain DM energy injection rates small enough that $x_e \lesssim 0.1$ at $z \sim 20$.

The resulting modifications to the temperature and ionization history for DM energy injection are

$$\dot{T}_m^\chi = \frac{2f_{\mathrm{heat}}(z)}{3(1+f_{\mathrm{He}}+x_e)n_{\mathrm{H}}}\left(\frac{dE}{dVdt}\right)^{\mathrm{inj}},$$
$$\dot{x}_e^\chi = \left[\frac{f_{\mathrm{ion}}(z)}{\mathcal{R}n_{\mathrm{H}}} + \frac{(1-\mathcal{C})f_{\mathrm{exc}}(z)}{0.75\mathcal{R}n_{\mathrm{H}}}\right]\left(\frac{dE}{dVdt}\right)^{\mathrm{inj}}, \quad (9)$$
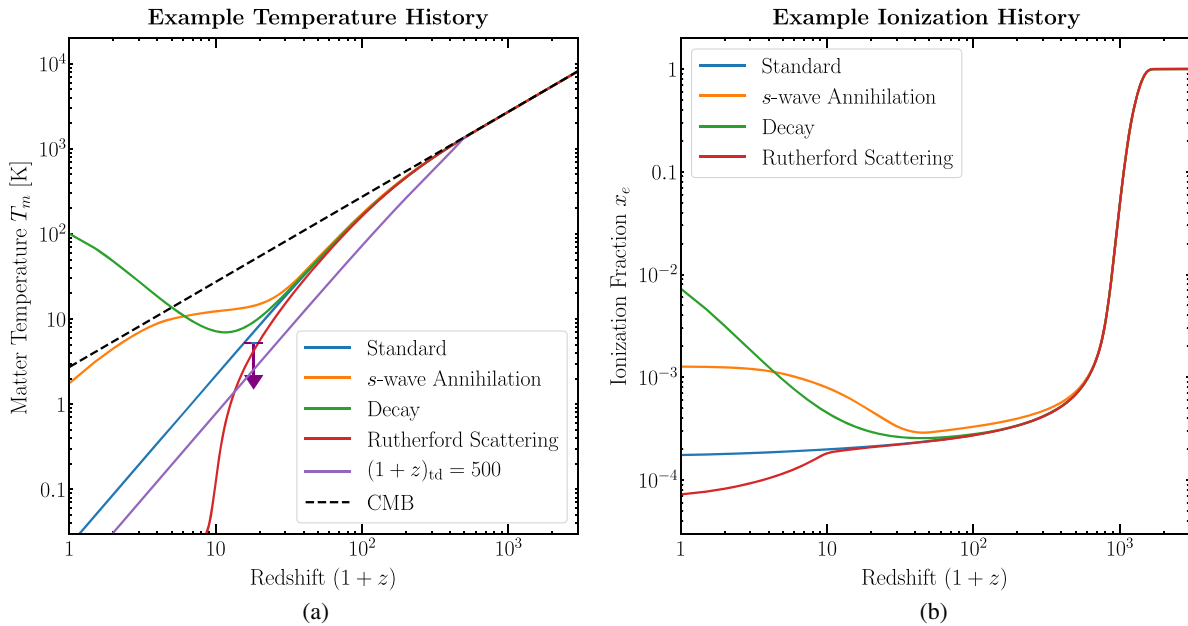

FIG. 1.   Example thermal (left) and ionization (right) histories, for $m_\chi = 100$ MeV. The standard history (blue lines), with DM s-wave annihilation ($\langle \sigma v \rangle = 10^{-29}$ cm^3 s^{-1}) (orange lines), with DM decays ($\tau = 10^{28}$ s) (green lines), and with DM-baryon Rutherford scattering (red lines) are shown. The thermal history with an earlier redshift of thermal decoupling at $(1+z)_{\mathrm{td}} = 500$ (light purple line) is also shown, with the CMB temperature (black dashed line) plotted for reference. The purple arrow indicates the EDGES temperature limit.

where $\mathcal{R} = 13.6$ eV is the ionization potential of hydrogen. The contribution from excitation to ionization is given by the number of excitation events (each event deposits an energy of $E_{21} = 0.75\mathcal{R}$), multiplied by the probability of an excited hydrogen atom getting ionized, given by $1 - C$.

Figure 1 shows a number of representative ionization and thermal histories, including the standard history obtained by integrating Eqs. (4) and (5), as well as two examples with DM s-wave annihilation and decay, which includes Eq. (9). The EDGES upper limit on the matter temperature if we take $T_R = T_{\text{CMB}}$ in Eq. (3) is also indicated. We have also included two of the new interactions that we will examine later: Rutherford-like interactions between the dark sector and hydrogen, as well as a temperature history with early decoupling of the photon and baryon temperatures. Throughout the paper, no star-formation or reionization models are included in this analysis: excluding these effects, which would only raise the matter temperature near $z \sim 20$, leads to annihilation cross section or decay lifetime limits that are less constraining and thus conservative. The impact of s-wave annihilation becomes significantly enhanced beginning at $z \sim 40$ due to structure formation, which greatly increases the local DM density. We discuss the systematics associated with structure formation in Appendix A.

The authors of [30] have recently pointed out that Lyman-α radiation at $z \sim 20$ is able to mediate a transfer of energy from the 21-cm CMB photons to the thermal motion of the gas, providing an additional and significant source of heating during this epoch. Although the inclusion of this effect would ultimately be important in setting precise DM annihilation and decay constraints, we neglect this effect in this paper, and leave a proper treatment of this process to future work. This is consistent with our omission of the process of reionization, and leads to limits that are less constraining than they would be in a more complete treatment.

III. ADDITIONAL 21-CM SOURCES

A large absorption trough can be explained by the existence of an additional 21-cm source, which would raise T_R, the effective radiation temperature at a wavelength of 21 cm at $z \sim 17.2$, above the CMB temperature. If T_R is large enough so that Eq. (3) is satisfied with $T_m \gtrsim 7$ K, no additional sources of cooling are required to explain the EDGES result. For $T_m = 7$ K, we require $T_R = 67$ K, compared to the CMB temperature at this redshift, $T_{\text{CMB}} = 50$ K.

This required value of T_R is large: since $T_R \gg \nu_{21}$, where $\nu_{21} = 1.42$ GHz is the hyperfine transition frequency, the effective temperature is directly proportional to the intensity, and so the additional 21-cm source must have approximately 35% of the intensity of the CMB at this frequency, at $z = 17.2$. The additional source must be confined to frequencies well below the peak of the CMB

spectrum in order to be consistent with FIRAS observations of the CMB blackbody spectrum in the present day, which rule out distortions to the CMB spectrum greater than one part in 10^4–10^5 for frequencies 60 GHz and greater [31,32]. It is interesting to note that the ARCADE 2 experiment has reported a deviation from the CMB blackbody spectrum in the (present day) 22 MHz–10 GHz range [33].

Models where such a strong photon emission comes directly from DM decay or deexcitation run into stringent electroweak precision constraints [12], but models where DM decays into dark photons which oscillate into 21-cm photons are viable, and can provide an order one or larger contribution to the intensity compared to the CMB [17]. This large temperature also cannot be explained by uncertainties in the radio emission from astrophysical sources [9], but may be possible to explain with optimistic black hole formation rates [19].

We shall therefore set constraints on DM annihilation and decay as a function of T_R, assuming that $1.3T_{\text{CMB}} \lesssim T_R \lesssim 10T_{\text{CMB}}$ at $z = 17.2$. If we take the EDGES 99% confidence limit on T_m shown in Eq. (3), this corresponds to gas temperatures in the range $7\text{K} \lesssim T_m \lesssim 50$ K at the same redshift. Lower values of T_R lead to values of T_m that are below the standard matter temperature at this redshift, in which case no additional heating would be allowed.

The evolution equations we solve are

$$\dot{T}_m = \dot{T}_m^{(0)} + \dot{T}_m^{\chi},$$
$$\dot{x}_e = \dot{x}_e^{(0)} + \dot{x}_e^{\chi}. \tag{10}$$

Note that the CMB temperature used in these equations remains unchanged, as we do not expect significant modifications to the overall energy density of the CMB. Since the evolution equations are essentially the same as the TLA with DM annihilation or decay, these constraints are related to those derived in [34], but are broadly applicable to measurements with $T_{21} \lesssim -200$ mK, including the EDGES measurement. We also use the structure formation prescription described in [29,35], with the boost factor included in $f_c(z)$, to account for any delayed deposition of energy.

Figures 2 and 3 show the constraints on the minimum decay lifetime and maximum annihilation cross section with an additional source of 21-cm radiation. The limits are presented as a function of m_χ and the ratio $(T_S/T_R)_{\text{obs}}/(T_S/T_R)_{\text{EDGES}} \times T_R$, with $(T_S/T_R)_{\text{EDGES}} = 0.105$ as given in Eq. (3); these limits can be rescaled if future 21-cm measurements alter or improve the measurement of $(T_S/T_R)_{\text{obs}}$. The constraints found in [34] for $T_{21} = 100$ mK and 50 mK are equivalent to the constraints obtained with $(T_S/T_R)_{\text{obs}} = 0.26$ and 0.41, respectively, and by setting $T_R = T_{\text{CMB}}$ at $z = 17.2$. Zoomed-in versions of these plots for lower temperatures are shown in Figs. 15 and 16.
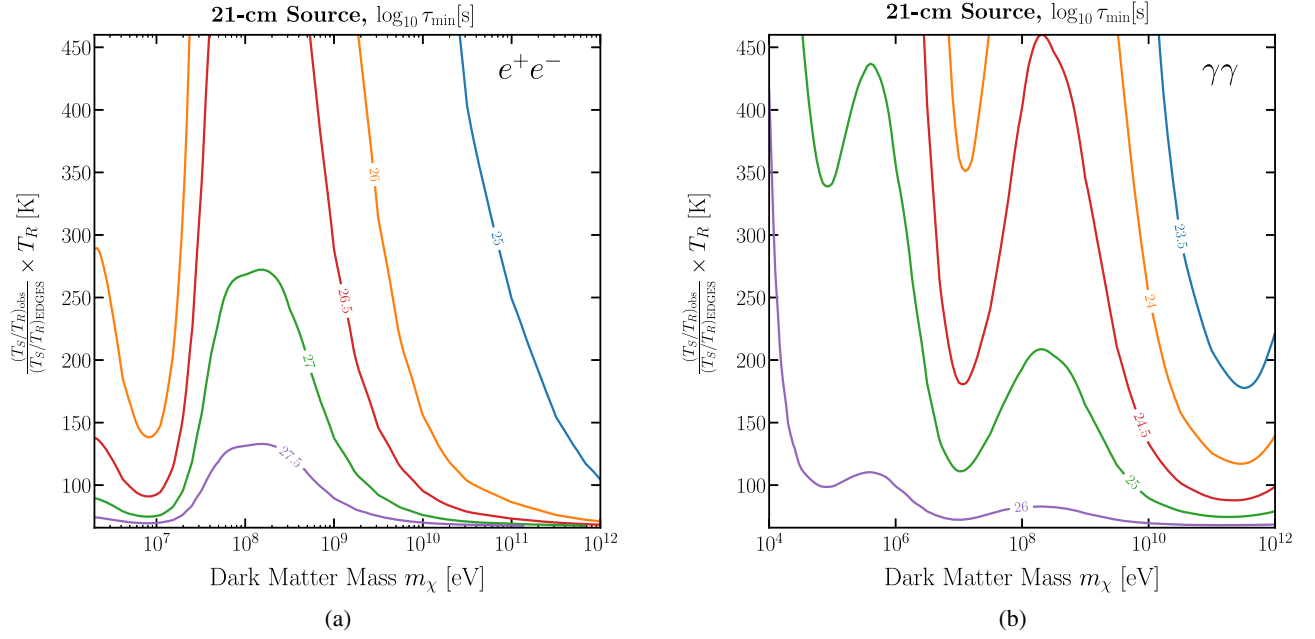
FIG. 2. Decay lifetime constraints with an additional 21-cm source with $\chi \to e^+ e^-$ (left) and $\chi \to \gamma\gamma$ (right), as a function of m_χ and $(T_S/T_R)_{\rm obs}/(T_S/T_R)_{\rm EDGES} \times T_R$. Contour lines of constant minimum $\log_{10} \tau$ (in seconds) are shown.
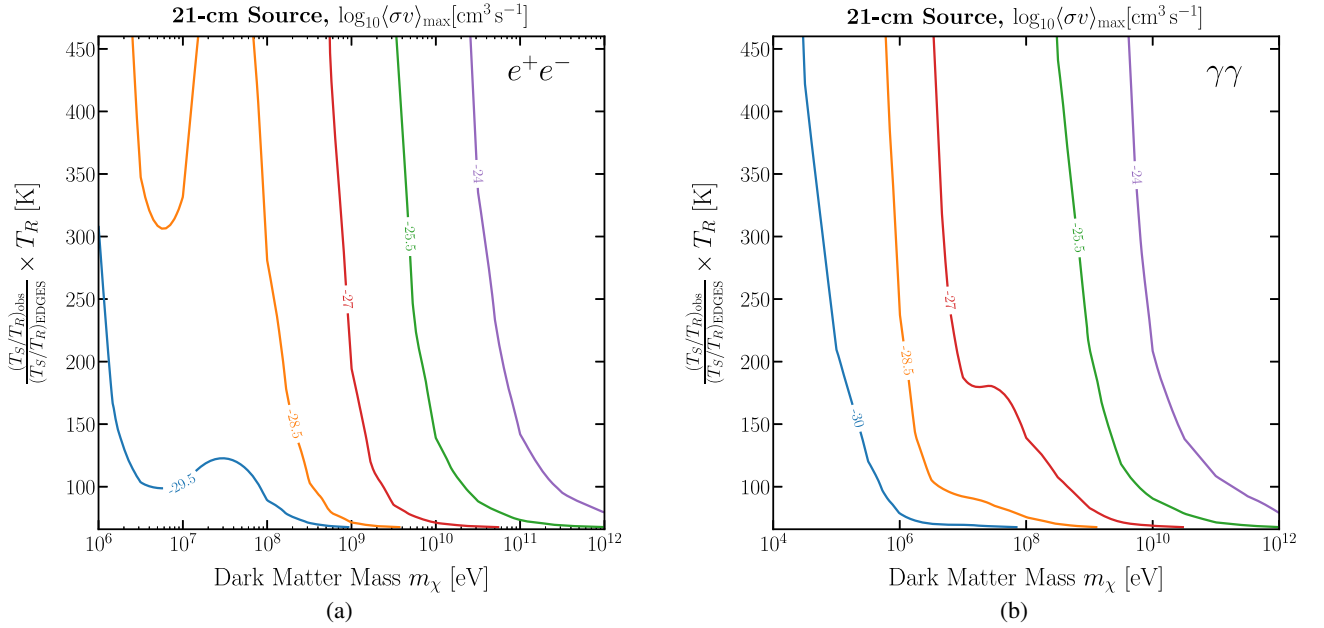

FIG. 3. Annihilation cross section constraints with an additional 21-cm source with $\chi\chi \to e^+ e^-$ (left) and $\chi\chi \to \gamma\gamma$ (right), as a function of m_χ and $(T_S/T_R)_{\rm obs}/(T_S/T_R)_{\rm EDGES} \times T_R$. Contour lines of constant maximum $\log_{10}\langle\sigma v\rangle$ (in cm³ s⁻¹) are shown. The green contour corresponds to the canonical relic abundance cross section of 3×10^{-26} cm³ s⁻¹.

For a given measurement of T_{21}, Eq. (3) permits a higher matter temperature for larger values of T_R, which weakens the constraints that can be set. Taking the observed EDGES measurement of T_S/T_R, a radiation temperature of $T_R \sim$ 100 K constrains the decay lifetime for $\chi \to e^+ e^-$ to more than 10^{25} s across all DM masses, which is significantly stronger than the existing Planck power spectrum limits [36]. Cross section constraints similarly strengthen considerably with respect to the Planck limits for $T_R < 100$ K.

IV. NONSTANDARD RECOMBINATION

Thermal decoupling occurs when the Compton scattering rate becomes comparable to the adiabatic cooling rate, marking the point where the matter temperature transitions from $T_m \propto (1 + z)$ to $T_m \propto (1 + z)^2$. The standard redshift of thermal decoupling without additional sources of heating or cooling $(1 + z)_{\text{td},0}$ is therefore obtained by setting $2HT_m = \Gamma_C T_m$, giving

$$(1 + z)_{\text{td},0} \approx \left[\frac{45 m_e H_0 \sqrt{\Omega_m}}{4\pi^2 \sigma_T x_e T_{\gamma,0}^4} \right]^{2/5}. \qquad (11)$$

Substituting a value of $x_e = 3 \times 10^{-4}$, a typical value for x_e during the dark ages, we get $(1 + z)_{\text{td},0} \approx 155$.

In nonstandard models of recombination, the ionization history can be altered in such a way that the x_e evolution equation, Eq. (5), is modified while leaving the temperature evolution unchanged; this can happen, for example, if the background radiation at energies on the order of the ionization potential for hydrogen deviates significantly from a blackbody distribution during recombination [37]. Another example will be discussed in Sec. V: if a small fraction of DM couples strongly to baryons, it can act as an additional heat sink and likewise modify the ionization and thermal history. Other mechanisms for early thermal decoupling, which have recently been proposed to explain the EDGES measurement, include the influence of early dark energy [15], or charge sequestration [14], where the number density of protons and electrons are unequal owing

to the presence of an additional dark charged species. In this work, we remain agnostic as to the cause of early decoupling, parametrizing it by the modified redshift of decoupling.

While the ionization history in such a situation would be model dependent, once thermal decoupling occurs, the evolution of the thermal history without DM energy injection is completely specified by $\dot{T}_m = -2HT_m$. The full evolution equation that we will thus solve is

$$\dot{T}_m = -2HT_m + \dot{T}_m^\chi, \qquad (12)$$

starting from the redshift of thermal decoupling. In reality, the nonzero value of x_e would still provide some additional Compton heating, but limits set by ignoring this effect are less constraining and thus conservative.

The effect of such modifications on the thermal history can therefore be parametrized by the redshift of decoupling $(1 + z)_{\text{td}}$. An earlier redshift of decoupling, when the gas temperature first diverges from the photon temperature, results in a lower gas temperature at later times: The EDGES result can be explained, for example, by a modified ionization and thermal history of this sort [3]. Without considering specific models for increasing $(1 + z)_{\text{td}}$, we can set constraints on DM energy injection processes as a function of this quantity, as long as heating from these processes are unimportant relative to adiabatic expansion prior to thermal decoupling.

Figure 4 shows the constraints set on the decay lifetime of a DM particle χ decaying to e^+e^- and $\gamma\gamma$, respectively, as
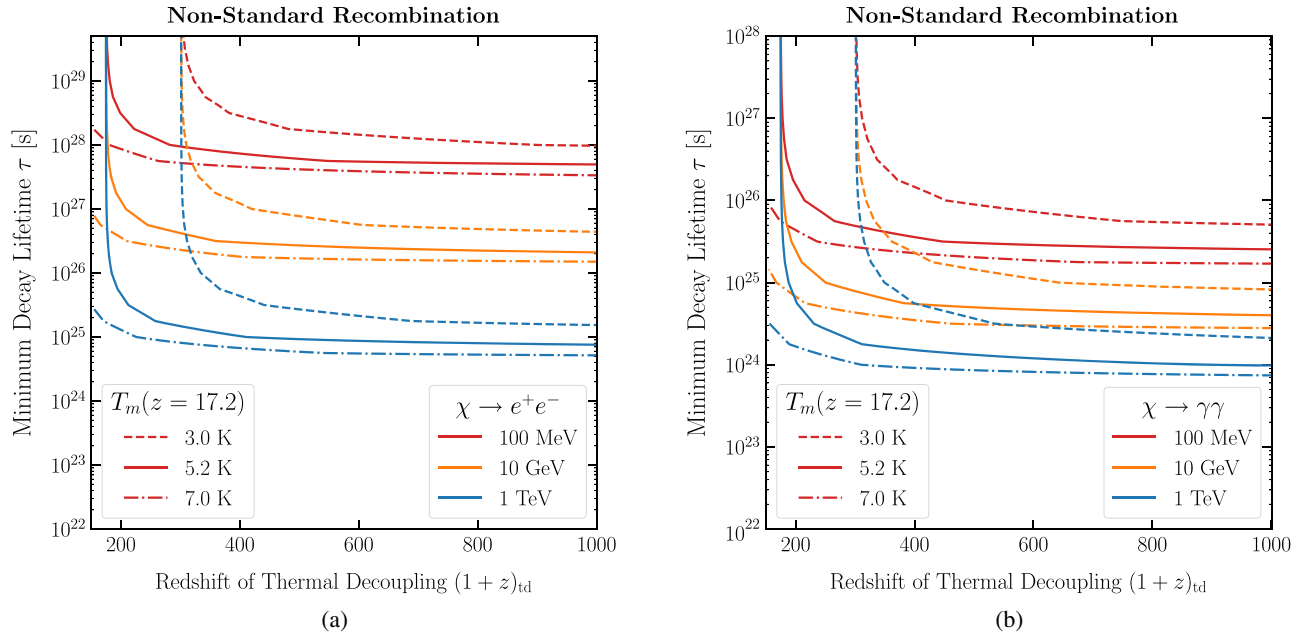

FIG. 4.   Decay lifetime constraints for nonstandard recombination as a function of the redshift of thermal decoupling, with $\chi \to e^+e^-$ (left) and $\chi \to \gamma\gamma$. In both plots, we show the limits for 100 MeV (red lines), 10 GeV (orange lines), and 1 TeV (blue lines) DM, assuming a measured upper limit on the matter temperature at $z = 17.2$ of 3.0 K (dashed lines), 5.2 K (solid lines), and 7.0 K (dot-dashed lines). The 5.2 K value corresponds to the EDGES limit.
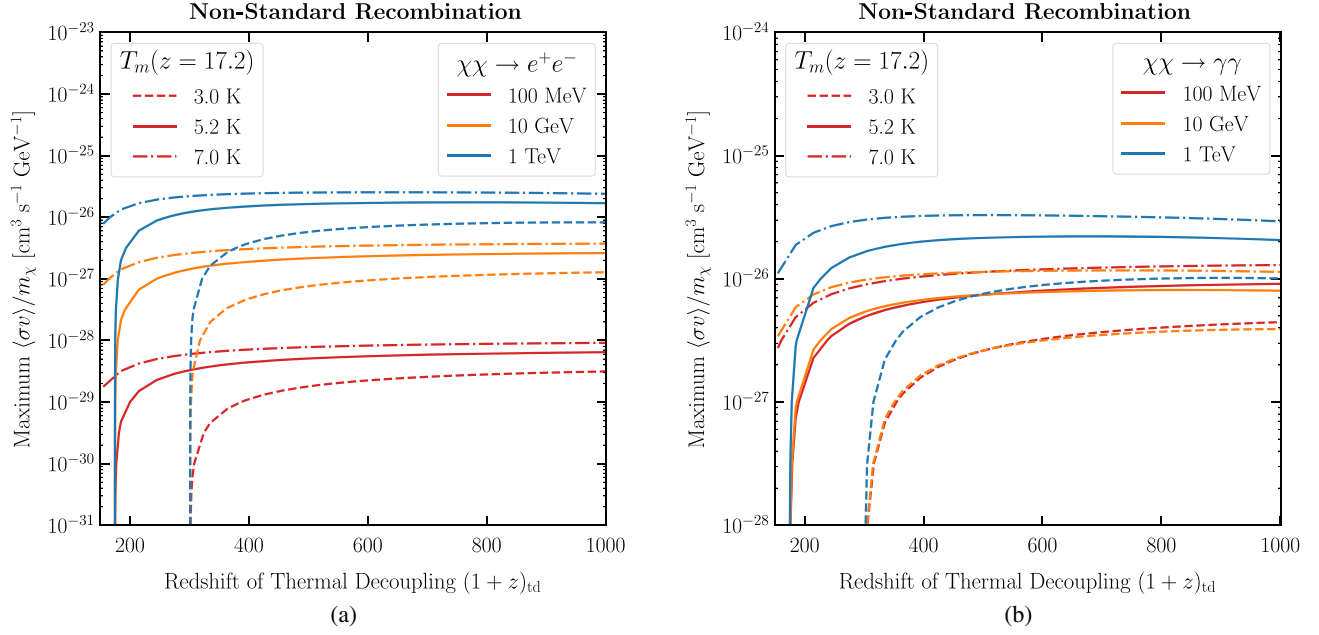
FIG. 5. Annihilation cross section constraints for nonstandard recombination as a function of the redshift of thermal decoupling, with $\chi\chi \rightarrow e^+e^-$ (left) and $\chi\chi \rightarrow \gamma\gamma$. In both plots, we show the limits for 100 MeV (red lines), 10 GeV (orange lines), and 1 TeV (blue lines) DM, assuming a measured upper limit on the matter temperature at $z = 17.2$ of 3.0 K (dashed lines), 5.2 K (solid lines), and 7.0 K (dot-dashed lines). The 5.2 K value corresponds to the EDGES limit.

a function of $(1 + z)_{\text{td}}$, for different possible values of T_m at $z = 17.2$. For temperatures below 7 K, the temperature for standard recombination, thermal decoupling must occur at a sufficiently high redshift before adiabatic cooling can bring T_m to that value. For the EDGES value of 5.2 K, this corresponds to $(1 + z)_{\text{td}} \sim 175$; additional heating from DM energy injection is allowed only when the thermal decoupling occurs above this value.

Once $(1 + z)_{\text{td}}$ exceeds the critical value for sufficient cooling, the constraints on the minimum decay lifetime depend only weakly on $(1 + z)_{\text{td}}$. To understand this, note that if the baryon temperature in the absence of heating is well below the observed temperature limit, then the temperature including heating is solely determined by the energy injection rate, and is relatively independent of the baseline baryon temperature without heating. Consequently, the temperature (with heating) is also nearly independent of $(1 + z)_{\text{td}}$.

With $T_m(z = 17.2) = 5.2$ K, the constraints set by this temperature measurement for $m_\chi = 100$ MeV is $\sim 5 \times 10^{27}$ s for decays to e^+e^- and $\sim 10^{25}$ s for $\gamma\gamma$, which is both at least an order of magnitude stronger than limits set by the Planck CMB power spectrum measurement [36]. These limits are valid assuming only no additional sources of cooling for the matter temperature after thermal decoupling.

Figure 5 shows a similar plot for the constraints on the annihilation cross section, with the main features of these constraints being similar to the result for decays. The constraints set by $T_m(z = 17.2) = 5.2$ K are once again

stronger than the current Planck constraints [25,38] for $m_\chi = 100$ MeV annihilating into e^+e^- by about an order of magnitude, with little dependence on $(1 + z)_{\text{td}}$.

Figure 6 shows the constraints for several decay and annihilation channels into SM particles for the case where $(1 + z)_{\text{td}} = 500$, and $T_m(z = 17.2) = 5.2$ K. These constraints apply across a large range of $(1 + z)_{\text{td}} \gtrsim 250$ due to the weak dependence on $(1 + z)_{\text{td}}$. To obtain these constraints, electron and photon spectra were computed using the PPPC4DMID [39], and the corresponding $f_c(z)$ were computed by taking an energy-weighted sum of the spectra [38]. The behavior of these limits is set mainly by the ability of the secondary photons and electrons of these decays to heat baryons at $z \sim 20$. The universe between $z \sim 20$ and recombination is mostly transparent to high energy photons, which explains the relatively weak constraints for $\chi \rightarrow \gamma\gamma$. The limits of the other channels relative to $\chi \rightarrow e^+e^-$ weaken when more neutrinos are produced during the cascade and strengthen when soft electrons/positrons with energy ~ 100 MeV are produced: electrons and positrons in this energy range are particularly effective at depositing their energy into baryons during this epoch [29,36].

V. RUTHERFORD COOLING

A. Evolution equations

To ensure that the matter temperature at $z \sim 17$ satisfies Eq. (3) while taking $T_R = T_{\text{CMB}}$, several groups [9–13]

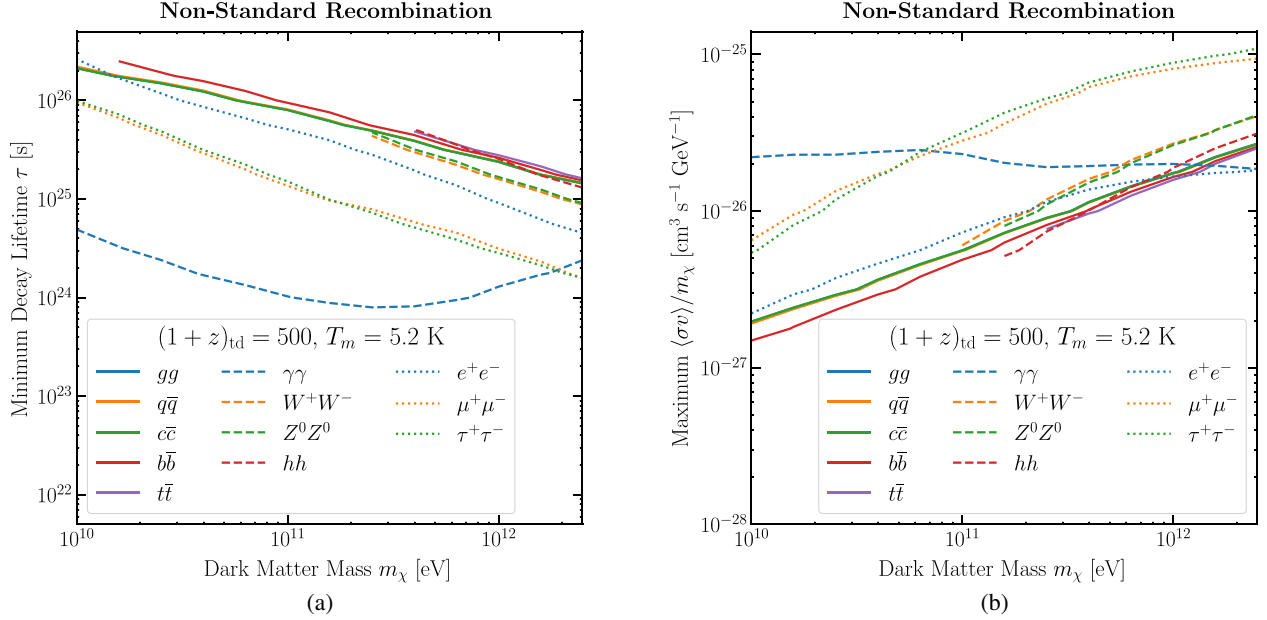Non-Standard Recombination

Non-Standard Recombination

FIG. 6.   Minimum decay lifetime (left) and maximum annihilation cross section (right) constraints for nonstandard recombination as a function of the redshift of thermal decoupling for several SM channels. In both plots, $(1 + z)_{\mathrm{td}} = 500$, and $T_m(z = 17.2) = 5.2$ K.

have examined the possibility of a new DM-baryon or DM-electron scattering cross section that has a Rutherford-like behavior, i.e., $\sigma = \sigma_0 v^{-4}$. This interaction may occur between only a fraction of DM and the SM. Both the difference in temperature between matter and DM as well as their bulk relative velocity $V_{\chi b}$ from earlier DM clustering can affect T_m, which evolves according to [10,40]

$$\dot{T}_m^c = \sum_j \frac{2}{3(1 + f_{\mathrm{He}} + x_e)n_{\mathrm{H}}} \frac{f_{\chi,\mathrm{int}}\rho_{\mathrm{DM}}\rho_j}{(m_\chi + m_j)^2}$$
$$\times \frac{\sigma_{0,j}}{u_j}\left[\sqrt{\frac{2}{\pi}}\frac{e^{-r_j^2/2}}{u_j^2}(T_\chi - T_m) + m_\chi \frac{F(r_j)}{r_j}\right], \quad (13)$$

where the sum is over all species j that can interact with the DM: this may be over all baryons [9], or over free protons and electrons in millicharged DM models [10]. ρ_{DM} and ρ_j are the mass densities of all DM and species j, respectively, with $f_{\chi,\mathrm{int}}$ being the fraction of DM interacting with the SM by mass. m_χ and T_χ are the mass and temperature of the interacting DM, respectively, $u_j \equiv (T_m/m_j + T_\chi/m_\chi)^{1/2}$ is the thermal sound speed of the DM-j fluid, and $r_j \equiv V_{\chi b}/u_j$. The function $F(r)$ is

$$F(r) \equiv \mathrm{erf}\left(\frac{r}{\sqrt{2}}\right) - \sqrt{\frac{2}{\pi}}e^{-r^2/2}r. \quad (14)$$

To solve for the full evolution, we must also evolve the temperature of the interacting DM [40],

$$\dot{T}_\chi = -2HT_\chi + \sum_j \frac{2}{3}\frac{m_\chi\rho_j}{(m_\chi + m_j)^2}$$
$$\times \frac{\sigma_{0,j}}{u_j}\left[\sqrt{\frac{2}{\pi}}\frac{e^{-r_j^2/2}}{u_j^2}(T_m - T_\chi) + m_j\frac{F(r_j)}{r_j}\right], \quad (15)$$

as well as the bulk relative velocity [10]

$$\dot{V}_{\chi b} = -HV_{\chi b} - \left(1 + \frac{f_{\chi,\mathrm{int}}\rho_{\mathrm{DM}}}{\rho_b}\right)\sum_j \frac{m_j n_j \sigma_{0,j}}{m_\chi + m_j}\frac{F(r_j)}{V_{\chi b}^2}. \quad (16)$$

When $f_{\chi,\mathrm{int}} < 1$, Eq. (15) assumes that the interacting component of DM has a temperature that is separate from the rest of the dark sector; relaxing this assumption would mean that the energy flow into the dark sector is distributed among more particles, with the exact effect on the thermal history determined by the masses of both the interacting and noninteracting components.

To set constraints on DM energy injection in the presence of this scattering process, the full set of rate equations which should be integrated are Eqs. (15) and (16), together with

$$\dot{T}_m = \dot{T}_m^{(0)} + \dot{T}_m^\chi + \dot{T}_m^c,$$
$$\dot{x}_e = \dot{x}_e^{(0)} + \dot{x}_e^\chi. \quad (17)$$

For simplicity, we will restrict our discussion to the case of DM-hydrogen scattering (both neutral and ionized, with no scattering on helium or free electrons) until we discuss the

millicharged DM model, where scattering occurs between DM and free charged particles. From a model building perspective, the existence of a DM-baryon interaction with a Rutherford-like scattering cross section is hard to accomplish without invoking millicharged DM or introducing a new long-range force; the latter scenario is extremely constrained by fifth force experiments [41,42], and the former leads to scattering on charged particles rather than neutral hydrogen. However, constraints on DM-hydrogen interactions from the CMB power spectrum [43–47] or from forecasts of 21-cm measurements [40] have been derived for a range of velocity-dependent cross sections, including Rutherford scattering, and it is informative to compare our constraints with the existing literature.

B. Weak- and strong-coupling regimes

The magnitude of σ_0 determines how tightly coupled the interacting DM and neutral hydrogen are. In the weakly coupled regime, the DM temperature T_χ remains well below the matter temperature T_m, and the interacting DM component is able to collapse into structures well before recombination, leading to a nonzero bulk relative velocity $V_{\chi b}$. However, for a sufficiently large σ_0, the temperature of the interacting DM becomes close to the matter temperature, and collapse into structures becomes impossible. For $f_{\chi,\mathrm{int}} = 1$, i.e., all of the DM interacts with the SM, this scenario is highly constrained by the damping effect this would have on the CMB power spectrum [45–48]. However, a subdominant component ($f_{\chi,\mathrm{int}} \lesssim 0.01$) of DM can have significant interactions with the SM at recombination without contradicting precision CMB measurements: the interacting DM component would essentially be an additional, small contribution to the baryon fluid, while leaving structure formation due to the bulk of DM unaffected [49].

In the weak-coupling regime, the interacting component of DM remains cold and collapses efficiently, and $V_{\chi b}$ is expected to have an rms velocity of 29 km s^{-1} at photon decoupling, $z = 1010$, the value expected for cold, non-interacting DM [50]. From Eq. (13), while $T_m \gg T_\chi$, the effect of a nonzero value of $V_{\chi b}$ is generally to increase T_m. This additional source of heating forces the energy injection from DM annihilation or decay to be smaller than if we set $V_{\chi b} = 0$, leading to tighter cosmological constraints. For the rest of the results in this section, we will show only results with $V_{\chi b} = 0$, which leads to the most robust constraints: the effect of fully evolving $V_{\chi b}$ starting at a nonzero value at recombination will be shown in Appendix A. We integrate Eqs. (15) to (17), with $T_\chi = 0$, $T_m = T_{\mathrm{CMB}}$, and $x_e = 1$ starting from before recombination.

In the strong-coupling regime, the interacting component of DM is in thermal equilibrium with baryons and the CMB, and cannot collapse into structures. In this case, $V_{\chi b} = 0$ at the point of recombination; furthermore, the

strong coupling between the two sectors ensures $T_m = T_\chi$ throughout. We can therefore integrate Eqs. (15) and (17), with $T_\chi = T_m = T_{\mathrm{CMB}}$ and $x_e = 1$ starting from before recombination, with $V_{\chi b} = 0$.

We delineate the two regimes by requiring the rate of DM heating due to DM-hydrogen scattering to be larger than HT_χ at recombination in the strong-coupling limit, so that DM and baryons remain at the same temperature at this point. This leads to the criterion

$$\sigma_0^{\mathrm{strong}} \gtrsim \frac{H}{n_H} \frac{(m_\chi + m_p)^2}{m_\chi m_p} \left(\frac{T_{\mathrm{CMB}}}{m_p} + \frac{T_{\mathrm{CMB}}}{m_\chi} \right)^{3/2} \quad (18)$$

at recombination for strong coupling to be valid, and we take the weak-coupling regime to be $\sigma_0 < 0.1\sigma_0^{\mathrm{strong}}$.

C. CMB power spectrum limits

DM annihilation and decay during the cosmic dark ages increase the residual ionization of the universe after recombination as compared to the standard history, and this change to the ionization history can be constrained by considering its impact on the CMB power spectrum. The presence of an additional source of cooling of the matter temperature, however, also modifies the ionization history during this time. If the rate of cooling is sufficiently large to decouple baryons from the CMB at a time earlier than $(1 + z)_{\mathrm{td,0}}$ given in Eq. (11), then T_m becomes smaller than
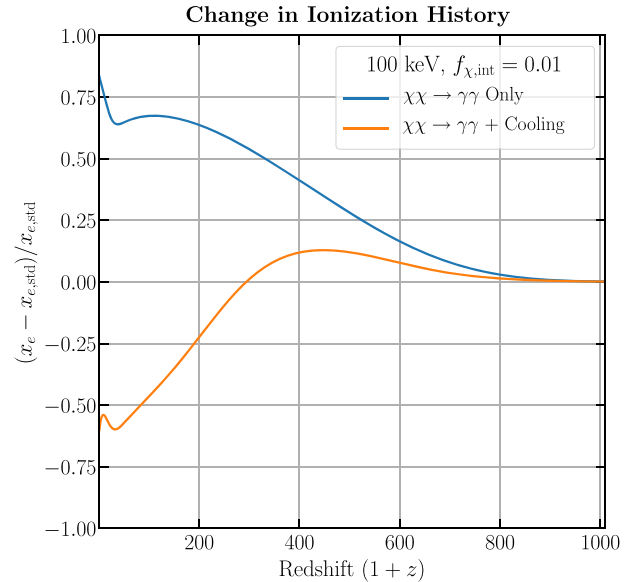

FIG. 7.   The change in ionization histories for $\chi\chi \to \gamma\gamma$ annihilation, with (yellow line) and without (blue line) Rutherford cooling, with respect to the standard ionization history (with no DM energy injection), $x_{e,\mathrm{std}}$. Here, $m_\chi = 100\,\mathrm{keV}$ and $f_{\chi,\mathrm{int}} = 0.01$. The chosen value of $\langle\sigma v\rangle = 6.6 \times 10^{-32}$ cm^3 s^{-1} is the maximum allowed from the Planck CMB limits in the absence of scattering; this scenario with scattering may evade these limits.

expected, which in turn increases the recombination rate, decreasing the residual ionization.

Figure 7 shows an example of the change in ionization history with respect to the standard ionization history. We have chosen an annihilation cross section for $\chi\chi \to \gamma\gamma$ that is naively ruled out by the Planck CMB limits. Due to the presence of Rutherford cooling, however, the ionization at $z \sim 600$ ($z \sim 300$) relative to the standard history is reduced by a factor of 2 (more than 10). Since the *s*-wave annihilation and decay constraints from *the CMB power spectrum are most sensitive to energy injection at $z \sim 600$ and $z \sim 300$, respectively [36], we conclude that the CMB power spectrum constraints on energy injection during this epoch can be significantly relaxed if additional sources of cooling lead to thermal decoupling of baryons during or before these redshifts.*

To estimate when thermal decoupling of baryons occurs in the presence of Rutherford cooling, we can compare the heat transfer rate due to cooling from DM to the Compton scattering term. For DM-hydrogen scattering, this gives the following condition for thermal decoupling to occur at $(1 + z)_{\mathrm{td}}$:

$$\sigma_{0,\mathrm{td}} \lesssim \sigma_T (1 + z)_{\mathrm{td}}^{5/2} \frac{x_e}{f_{\chi,\mathrm{int}}} \left[\frac{T_{\mathrm{CMB},0}^{11/2} m_\chi^2}{\mu_{\chi p}^2 m_p^{1/2} m_e \rho_{\mathrm{DM},0}} \right], \quad (19)$$

where $\mu_{\chi p}$ is the reduced mass of DM and protons, $T_{\mathrm{CMB},0}$ is the CMB temperature today, and $\rho_{\mathrm{DM},0}$ is the DM density today. Numerically,

$$\sigma_{0,\mathrm{td}} \lesssim \left(\frac{m_\chi}{\mu_{\chi p}} \right)^2 \left(\frac{(1 + z)_{\mathrm{td}}}{600} \right)^{5/2} \frac{10^{-40} \ \mathrm{cm}^2}{f_{\chi,\mathrm{int}}}, \quad (20)$$

where we have taken $x_e \approx 3 \times 10^{-4}$. Thus, for $(1 + z)_{\mathrm{td}} = 300$ and $(1 + z)_{\mathrm{td}} = 600$, the CMB power spectrum constraints for decays and *s*-wave annihilation may become inapplicable for $\sigma_0 > \sigma_{0,\mathrm{td}}$ due to the enhanced recombination from cooling at these redshifts. A sufficiently large T_χ can relax this condition, but we neglect this effect; CMB constraints on all plots are therefore shown only in regions where their validity is assured. A comprehensive study of how CMB constraints on DM annihilation relax under these circumstances is left to future work.

D. Weak-coupling results

Figure 8 shows a typical ionization and temperature history in the weak-coupling limit with both cooling and DM annihilation. Thermal decoupling of matter from the CMB occurs slightly earlier than $(1 + z)_{\mathrm{td}} \sim 155$ due to the additional cooling, but not significantly earlier. Since the matter temperature is locked to the radiation temperature until well after $z \sim 600$, the ionization history, even in the presence of DM annihilation, differs very little from the expected history without cooling. As a result, constraints on *s*-wave annihilation set by the CMB spectrum, which is most sensitive to energy injection at $z \sim 600$, are still applicable.

Figure 9 shows the constraints for DM decays to $e^+ e^-$ and $\gamma\gamma$, respectively as a function of σ_0 for DM-hydrogen
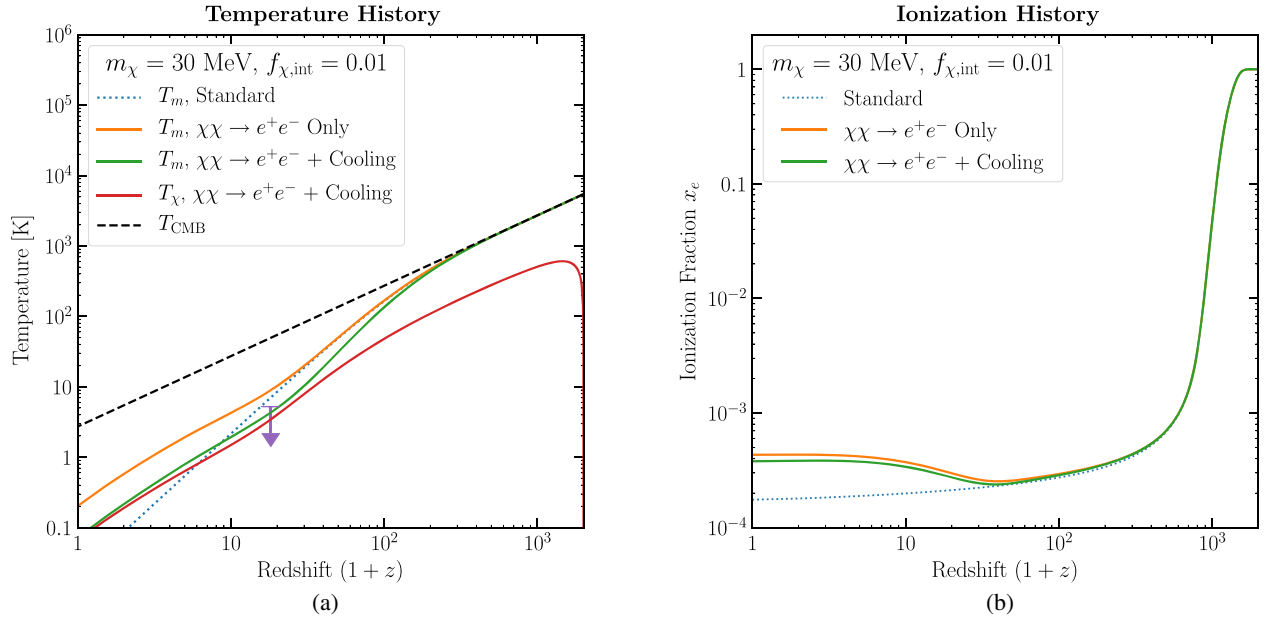

FIG. 8.   Thermal (left) and ionization (right) histories with $\chi\chi \to e^+ e^-$ annihilation and Rutherford cooling, $m_\chi = 30$ MeV and $f_{\chi,\mathrm{int}} = 0.01$. The standard history (blue dotted lines), DM annihilation only with $\langle \sigma v \rangle = 10^{-30}$ cm^3 s^{-1} (orange lines), and DM annihilation and DM-baryon Rutherford scattering with $\sigma_0 = 10^{-38}$ cm^2 (green lines) are shown. The DM temperature evolution (red line), CMB temperature (black dashed line), and the EDGES temperature limit (purple arrow) are also shown.
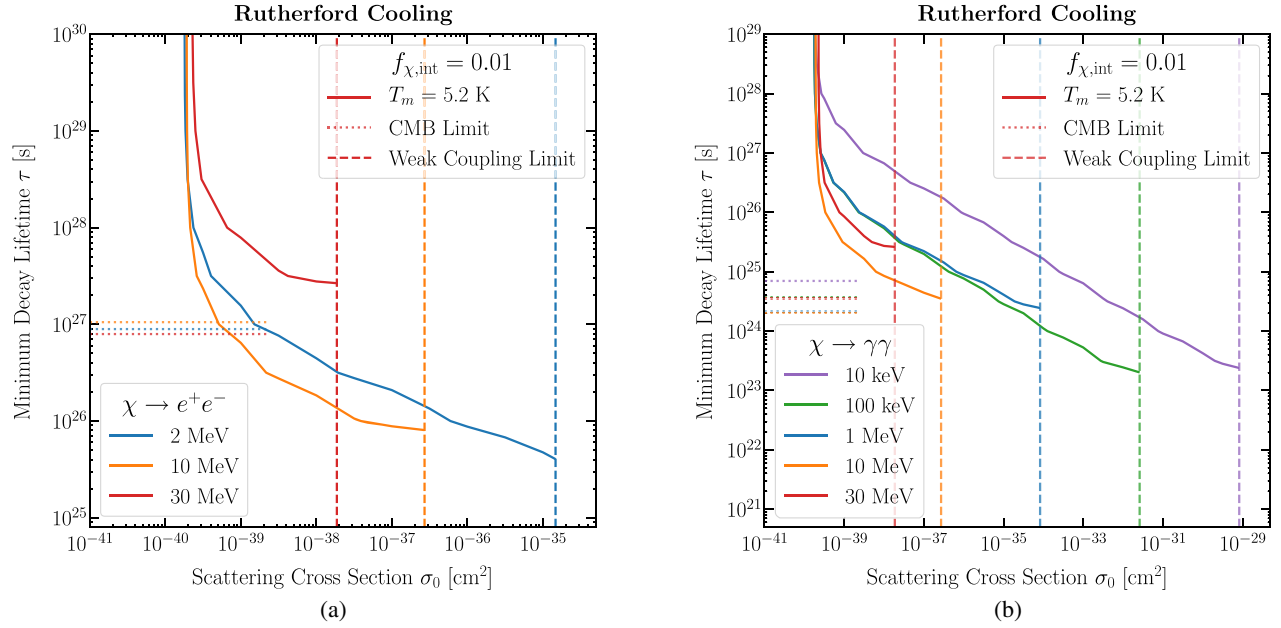
FIG. 9. Rutherford cooling constraints on the minimum decay lifetime for $\chi \to e^+ e^-$ (left) and $\chi \to \gamma\gamma$ (right) from the matter temperature $T_m(z = 17.2) = 5.2$ K (solid lines), $f_{\chi,\mathrm{int}} = 0.01$. Limits from the Planck measurement of the CMB power spectrum are also shown up to $\sigma_0 = \sigma_{0,\mathrm{td}}(z = 300)$ (dotted lines), together with the maximum scattering cross sections for the weak-coupling limit to hold (dashed lines). The vertical part of the contours marks the minimum value of σ_0 required to cool the gas down to 5.2 K in the absence of any source of energy injection.

scattering in the weak-coupling limit (set by the dashed lines), for the case where $f_{\chi,\mathrm{int}} = 0.01$ and $f_{\chi,\mathrm{dec}} = 1$. We also assume that the decaying DM component has the same mass as the interacting DM for simplicity. The CMB power spectrum constraints are shown up to $\sigma_0 = \sigma_{0,\mathrm{td}}$, after which the constraints may not be applicable. Even without any energy injection from decay, a minimum scattering cross section of $\sigma_0 \sim 10^{-40}$ cm^2 is required for sufficient cooling to bring T_m down to 5.2 K, absent any additional heat source. This minimum value is marked by the vertical contours of constant σ_0 at large decay lifetimes. Over a large range of σ_0, the temperature constraint set by the EDGES 21-cm measurement is more constraining than the CMB limits for parts of parameter space. For 10–100 keV DM decaying to photons, thermal decoupling as given in Eq. (20) occurs earlier than $z \sim 300$ even in the weak-coupling regime, and at large scattering cross sections, only the temperature measurement can effectively constrain the decay lifetime.

Figure 10 shows similar constraints on the s-wave annihilation cross section. The temperature limits in both cases are relatively insensitive to the actual value of T_m at $z \sim 20$. The exact value of T_m sets the minimum scattering cross section for cooling with no energy injection, but at higher cross sections/shorter decay lifetimes, the constraints are essentially set by having the large amount of heating almost entirely canceled by Rutherford cooling.

Analogous plots for the case where $f_{\chi,\mathrm{int}} = 1$ are shown in Figs. 19 and 20 in Appendix B.

E. Strong-coupling results

The ionization history with both cooling and DM annihilation in the strong-coupling limit, on the other hand, exhibits important differences from the history with no cooling, especially when the interacting component of DM is light. Since the transfer of energy between baryons and the interacting DM is efficient, Compton heating from the CMB must also be able to efficiently heat all of the interacting DM particles in order to keep the matter temperature at the CMB temperature. The additional heating needed means that thermal decoupling between the CMB and baryons can occur at a much higher redshift, if the DM mass is sufficiently light. After decoupling, since $T_m = T_\chi$, Eq. (13) shows that both of these temperatures simply evolve through adiabatic cooling in the absence of DM energy injections. The strong-coupling limit therefore reduces to a nonstandard recombination history with early thermal decoupling, discussed in Sec. IV.

We can obtain the redshift of thermal decoupling between photons and the coupled baryon-DM fluid by replacing the Compton heating term in Eqs. (4) and (6) by

$$\Gamma_C \to \Gamma_C \frac{n_{\mathrm{H}}}{n_{\mathrm{H}} + n_\chi}, \tag{21}$$

since energy from Compton heating must be redistributed into the dark sector as well. This gives
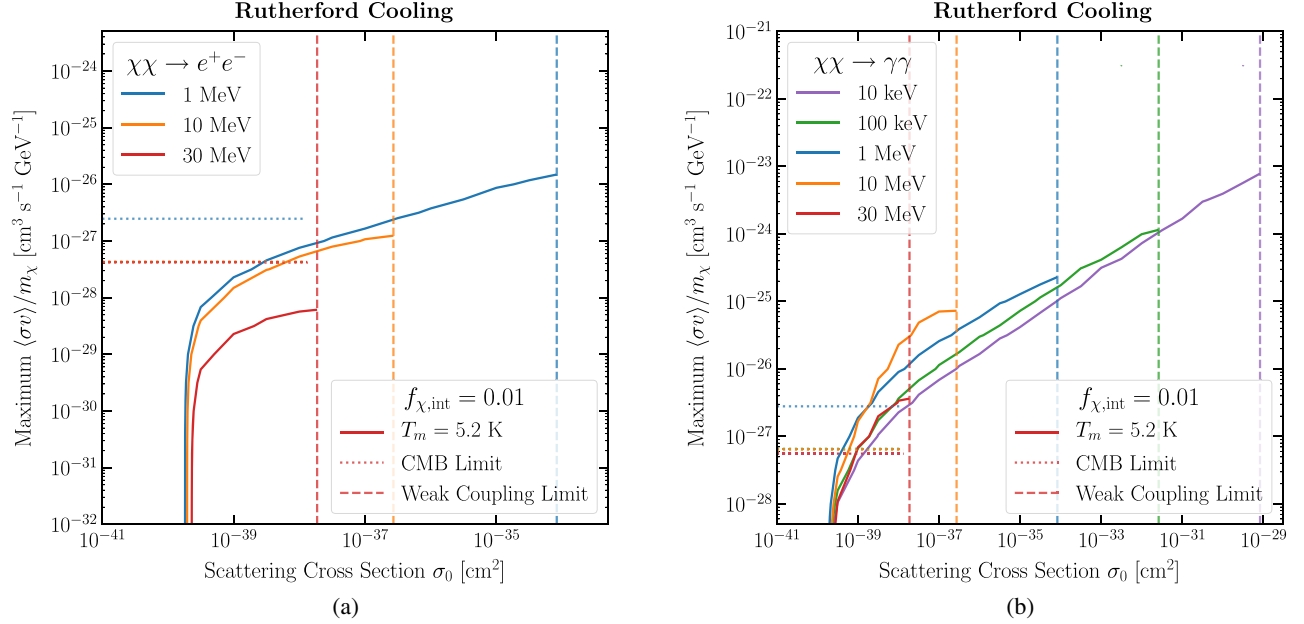
FIG. 10. Rutherford cooling s-wave annihilation constraints for $\chi\chi \to e^+e^-$ (left) and $\chi\chi \to \gamma\gamma$ (right) from the matter temperature $T_m(z = 17.2) = 5.2$ K (solid lines), $f_{\chi,\text{int}} = 0.01$. Limits from the Planck measurement of the CMB power spectrum are also shown up to $\sigma_0 = \sigma_{0,\text{td}}(z = 600)$ (dotted lines), together with the maximum scattering cross section for the weak-coupling limit to hold (dashed lines). The vertical part of the contours marks the minimum value of σ_0 required to cool the gas down to 5.2 K in the absence of any source of energy injection.

$$(1+z)_{\text{td}}^{\text{strong}} \approx \left[\frac{45 m_e H_0 \sqrt{\Omega_m}}{4\pi^2 \sigma_T x_e T_{\gamma,0}^4}\left(1 + \frac{f_{\chi,\text{int}}\rho_\chi}{n_H m_\chi}\right)\right]^{2/5}$$

$$\approx 155\left(1 + 5 f_{\chi,\text{int}}\frac{m_p}{m_\chi}\right)^{2/5}, \qquad (22)$$

with the redshift of thermal decoupling being independent of the scattering cross section. Note that we limit $(1+z)_{\text{td}}^{\text{strong}}$ to a maximum value of 1090, corresponding to the redshift of recombination, since thermal decoupling cannot occur before that, owing to the strong coupling between the fully ionized plasma and the CMB. In the limit of strong coupling, the exact details of how this coupling comes about is not important in determining the thermal and ionization history of the baryons.

For a given $\langle\sigma v\rangle$, the heating rate of baryons in the presence of a strongly coupled DM is less than without DM, since some amount of the heating is transferred to the dark sector. If x_e is small, we can account for this difference by replacing

$$\langle\sigma v\rangle \to \frac{n_H}{n_H + n_\chi}\langle\sigma v\rangle, \qquad (23)$$

and similarly for $1/\tau$ with decays. The constraints for the strong-coupling limit can easily be determined from the nonstandard recombination constraints: if $\langle\sigma v\rangle_{\text{max},(1+z)_{\text{td}}}$ is the maximum annihilation cross section from early thermal decoupling at redshift $(1+z)_{\text{td}}$, then the corresponding constraint from the strong-coupling limit is

$$\langle\sigma v\rangle_{\text{max,strong}} = \left(1 + \frac{f_{\chi,\text{int}}\rho_\chi}{n_H m_\chi}\right)\langle\sigma v\rangle_{\text{max},(1+z)_{\text{td}}^{\text{strong}}}$$

$$\approx \left(1 + 5 f_{\chi,\text{int}}\frac{m_p}{m_\chi}\right)\langle\sigma v\rangle_{\text{max},(1+z)_{\text{td}}^{\text{strong}}}. \qquad (24)$$

This has been explicitly checked by directly solving Eqs. (15) and (17) with $V_{\chi b} = 0$.

F. Millicharged DM

We now turn our attention to the millicharged DM model discussed in [10,11], focusing on the case where $f_{\chi,\text{int}} = 0.01$, which evades the DM-baryon scattering CMB limits. We emphasize that we allow only the millicharged DM to scatter off free electrons and protons. The charge of the proton in neutral hydrogen is screened throughout the cosmic dark ages from any millicharged DM, since the typical momentum transfer between DM and neutral hydrogen lies well below the Bohr momentum of the electron in a hydrogen atom.

For DM masses of interest ($\gtrsim 1$ MeV), a symmetric, Dirac fermion[1] millicharged DM has an unavoidable

---

[1]The cross section for annihilation of complex scalars to e^+e^- pairs is p-wave suppressed, and while p-wave annihilation also leads to significant heating at $z \sim 20$, we defer a proper treatment of this process to future work.
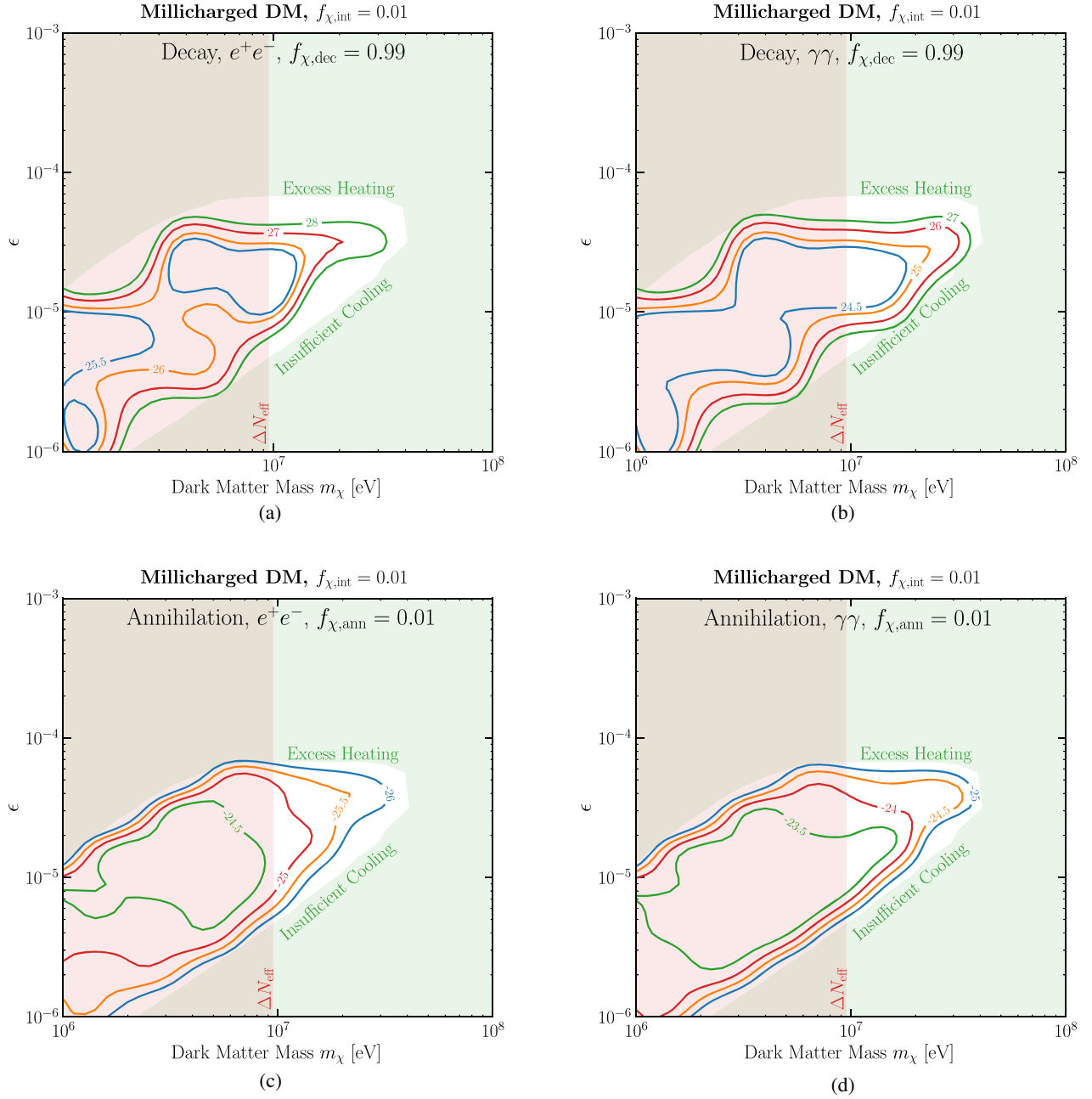
FIG. 11. Constraints on the millicharged DM, with an additional source of DM decay (upper panels) or annihilation (lower panels) to e^+e^- (left panels) and $\gamma\gamma$ (right panels), with $f_{\chi,\mathrm{dec}} = 0.99$ and $f_{\chi,\mathrm{ann}} = 0.01$, respectively. The region of parameter space ruled out by changes to N_{eff} from the CMB power spectrum measurement (red) is shown. Charges ϵ that are not large enough for efficient cooling of baryons (green, below) or so large that excess heating occurs at $z \sim 20$ are excluded. Contours of constant minimum $\log_{10} \tau$ in seconds for decay and maximum $\log_{10}\langle \sigma v \rangle$ in cm^3 s^{-1} so that $T_m(z = 17.2) < 4$ K are drawn.

s-wave annihilation channel into e^+e^-, with a velocity-averaged cross section given by

$$\langle \sigma v \rangle = \frac{\pi \alpha_{\mathrm{EM}}^2 \epsilon^2}{m_\chi^2} \sqrt{1 - \frac{m_e^2}{m_\chi^2}} \left(1 + \frac{m_e^2}{2m_\chi^2} \right), \qquad (25)$$

where ϵ is the charge of the millicharged DM, and α_{EM} is the electromagnetic fine-structure constant. The tree-level annihilation to $\gamma\gamma$ also exists, but this is suppressed by a factor of ϵ^2 relative to the annihilation to e^+e^- and can be safely neglected. Raising ϵ therefore increases both the rate of DM-baryon scattering and the rate of DM annihilation to electrons, and the opposing effects on T_m should be

properly taken into account when considering the viability of this model. Since annihilation takes place between Dirac millicharged particles, the annihilation rate given in Eq. (7) must have the additional factor of $1/2$.

Figure 11 shows a plot of the $m_\chi - \epsilon$ parameter space of this model with several relevant constraints. Sufficiently light millicharged DM remains in thermal equilibrium with electrons and photons until after neutrinos decouple from the SM, altering the $N_{\rm eff}$ measurement from the CMB power spectrum [51]. Since the irreducible annihilation to e^+e^- heats the baryons, by requiring $T_m(z = 17.2) \lesssim 4\ {\rm K}$,[2] we can set an upper limit of $\epsilon \lesssim 5 \times 10^{-5}$, cutting the remaining parameter space down to a narrow window between $m_\chi \sim 10$–100 MeV and $\epsilon \sim 5 \times 10^{-6}$ to 5×10^{-5}. The region of parameter space that is ruled out is labeled "Excess Heating" in Fig. 11, and is a new constraint set for the first time on this model. These limits are stronger than the conventional CMB power spectrum limits, since at large values of ϵ, DM and baryons become strongly coupled early on, and the temperature evolution is mostly dominated by adiabatic cooling until structure formation; once structure formation starts, a small perturbation on the order of a few kelvins from millicharged DM annihilation to e^+e^- is all that is required to raise the temperature above the EDGES measurement. Other experimental constraints set by the SLAC millicharge experiment [52] and observations of the cooling of SN1987a [53] set limits that are already ruled out by a combination of the two limits shown and have been left out.

We now consider an additional source of DM-related energy injection through s-wave annihilation or decay. This need not come from annihilation or decay of the millicharged DM itself; in principle, other particles in the dark sector could contribute such an energy injection. However, the existence of an additional annihilation channel for the millicharged DM could potentially allow it to obtain its correct relic abundance through thermal freeze-out, since the cross section of the irreducible annihilation to e^+e^- is too small in the allowed region for this to happen.

We set $f_{\chi,{\rm ann}} = 0.01$ in Eq. (7) when including a new source of energy injection from annihilation. If the millicharged DM has an additional annihilation channel, the contours set an upper limit on the annihilation cross section of this channel. These results can also be rescaled for other values of $f_{\chi,{\rm ann}}$, which may be useful for models where the dominant component of DM has an annihilation channel to the SM instead, assuming the dominant component has the same mass as the millicharged component. For decays, since the millicharged DM is expected to be stable, we choose $f_{\chi,{\rm dec}} = 0.99$ corresponding to the remaining, dominant component of DM; the constraints would apply to decays of this component into the SM.

For s-wave annihilation, we find that $\langle \sigma v \rangle \lesssim 2 \times 10^{-25}\ {\rm cm}^3\,{\rm s}^{-1}$ for $\chi\bar{\chi} \to e^+e^-$ and $\langle \sigma v \rangle \lesssim 7 \times 10^{-24}\ {\rm cm}^3\,{\rm s}^{-1}$ for $\chi\bar{\chi} \to \gamma\gamma$. Since the cross section to produce the correct relic abundance of the millicharged DM with $f_{\chi,{\rm int}} = 0.01$ is $\langle \sigma v \rangle \sim 6 \times 10^{-24}\ {\rm cm}^3\,{\rm s}^{-1}$, it is unlikely that any additional source of s-wave annihilation to e^+e^- (on top of the irreducible s-channel annihilation through the SM photon) can produce the correct relic abundance while remaining consistent with the EDGES T_m measurement at $z \sim 20$. There is a small parameter space allowed for annihilation to photons to get the correct relic abundance without late-time suppression, but this requires a small branching ratio to electrons at the same time.

VI. CONCLUSION

We have computed the constraints that can be set on annihilating/decaying DM by a measurement of the 21-cm line of neutral hydrogen from the end of the cosmic dark ages. The recent claimed observation of an absorption trough by EDGES motivates the inclusion of some mechanism beyond the simplest scenario to explain the unexpectedly low inferred gas temperature; however, even if a future experiment found a weaker absorption signal, such additional mechanisms could still potentially be present and should be included to obtain conservative constraints.

We have considered three general scenarios that could weaken constraints from 21-cm observations on exotic energy injection from heating in the cosmic dark ages: (1) additional radiation backgrounds in the frequency range surrounding 21 cm, (2) nonstandard recombination allowing the gas to decouple thermally from the CMB earlier, and (3) cooling of the gas through DM-baryon scattering. We have demonstrated that the strong-coupling limit of scenario (3) implements scenario (2) as a corollary, and that scenario (3) can generically weaken previously studied constraints on exotic energy injections from modifications to the ionization history during the cosmic dark ages.

We have mapped out the constraints on DM annihilation/decay in these three scenarios. We have found that in cases (2) and (3), there is an asymptotic behavior where the constraints become nearly independent of the redshift of decoupling [in case (2)] or the interaction cross section [in case (3)] for sufficiently early decoupling/large cross sections (see Figs. 4 and 5). In these scenarios, we can thus present robust constraints that do not depend on the exact redshift of decoupling in case (2) or the size of the cross section in case (3).

In the case where a small fraction of light DM (below 100 MeV) is millicharged and scatterings on this component are responsible for cooling of the gas, we have demonstrated that if this component has additional annihilation channels sufficient to obtain its relic density through thermal freeze-out, then the energy injection from those channels will generically overheat the gas. Thus such a component would likely need to possess a nonthermal

---

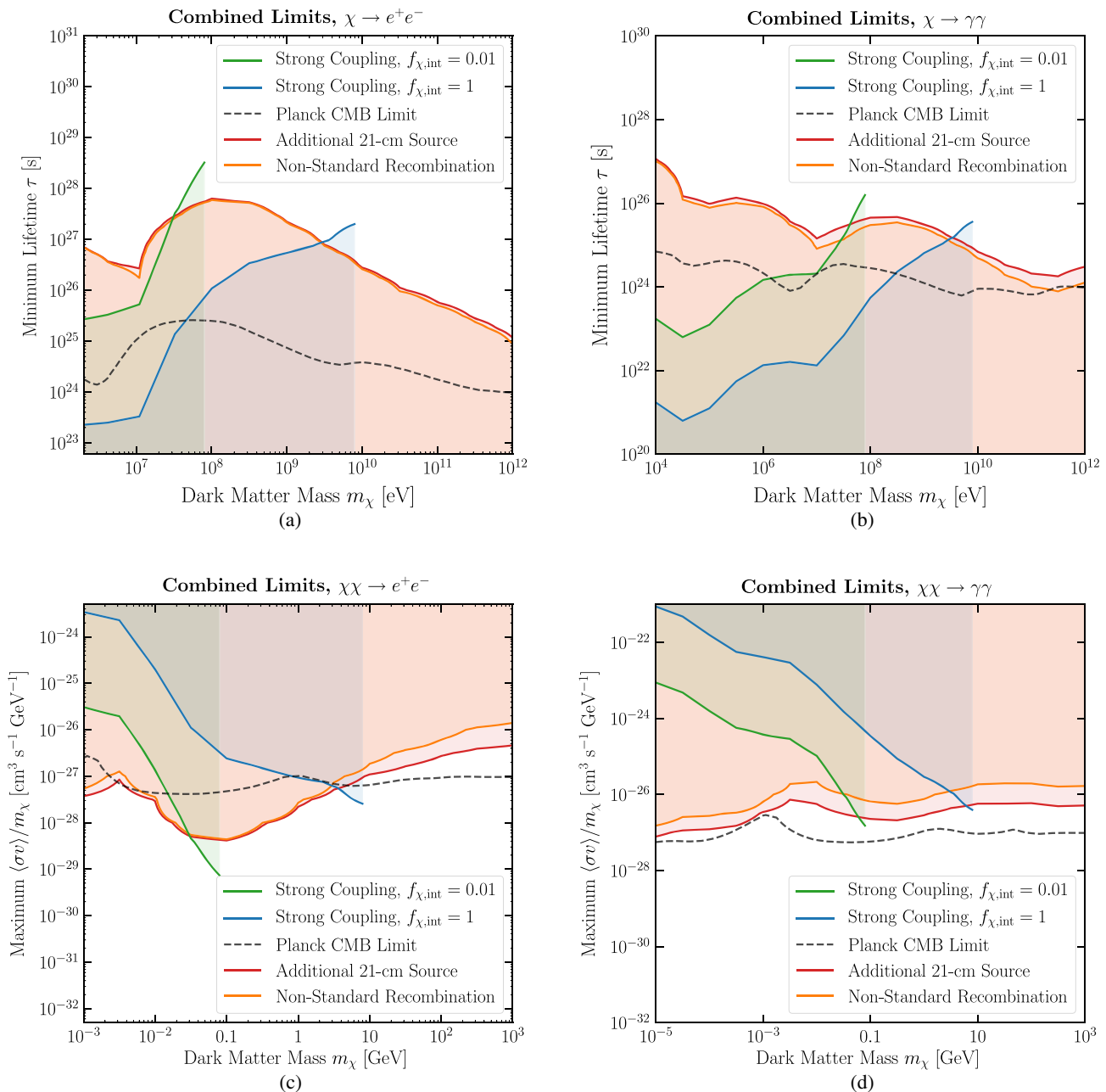[2]We choose 4 K in this section for consistency with existing results in the literature.

FIG. 12. Lower limits on the DM decay lifetime (upper panels) and upper limits of annihilation cross section (lower panels) from requiring $T_m/T_R(z = 17.2) \leq 0.105$ [3], for decay to e^+e^- pairs (left panels) and photons (right panels). In the presence of an additional 21-cm radiation source with number density (at that frequency) smaller than or equal to that of the CMB number density, constraints are shown by the red solid line. In the limit of early baryon-photon decoupling, constraints are shown by the orange solid line. The solid green and solid blue lines represent the constraints in the presence of DM-baryon scattering, in the limit of large cross section, for, respectively, 1% and 100% of the DM interacting with the baryons (these mechanisms cannot sufficiently cool baryons to match the data above critical mass scales, represented by the vertical cutoffs on the right-hand side of the green/blue regions). The black dashed line represents previously derived constraints on the decay lifetime [36] (upper panels) or annihilation cross section (lower panels) [38] from measurements of the CMB.

origin or, if a thermal relic, have annihilation channels in the early universe that are suppressed at late times (or have a large branching ratio for annihilation directly to neutrinos).

In Fig. 12 we summarize the constraints that can be obtained on keV–TeV DM annihilation or decay into e^+e^- pairs or photons, in these three scenarios, if the EDGES result is confirmed, and compare these limits with the

Planck CMB constraints; other limits from indirect detection also exist for both channels, and may be more constraining at higher DM masses (e.g., [54,55]). These particles are the main stable, electromagnetically interacting by-products of more general annihilation/decay channels (other than annihilation/decay directly to neutrinos), and consequently the constraints on more general channels can be estimated by combining these results. To set a limit in the case of additional radiation backgrounds, we assume that the effective radiation temperature T_R at the 21-cm wavelength is not more than twice the temperature of the CMB at $z = 17.2$.

In the case of DM-baryon scattering [scenario (3)], the cooling is only sufficient to reduce the gas temperature below 5.2 K if the DM mass is below a certain critical scale (depending on the fraction of the DM that is interacting); consequently, the constraints cut off above a certain mass scale because even for zero energy injection from annihilation/decay, the proposed mechanism cannot explain the data. The other two scenarios are in principle viable at all DM mass scales. We find that in these scenarios, for decaying DM, these constraints would generically be stronger than previously derived early-universe bounds, and in the case of decay primarily to electrons (as is expected for sub-100-MeV DM), these limits are stronger by up to 2 orders of magnitude. For DM annihilating to electrons, the constraints in these scenarios are generally stronger than CMB-based limits for sub-GeV DM (without taking into account that the CMB constraints may weaken due to a modified ionization history).

Simultaneous with and slightly after the release of this work, several other authors also studied the sensitivity of 21-cm measurements to DM annihilation or decays [56–58]. In particular, the authors of [57,58] set decay lifetime limits in a similar manner as [34], assuming an absorption signal that is smaller than the EDGES signal, with either $T_{21} = 100$ mK or 50 mK. We reiterate that these limits are equivalent to our additional 21-cm source limits, with $(T_S/T_R)_{\rm obs} = 0.26$ and 0.41, respectively, and setting $T_R = T_{\rm CMB}$ at $z = 17.2$. Our work is more general than these other studies as we consider new effects that must be present to account for the large negative value of T_{21} for the EDGES measurement. Consequently, our results are not merely a sensitivity study, and are immediately applicable to the various scenarios that have been suggested to explain the claimed EDGES detection. Even if future 21-cm measurements report a less negative value of T_{21}, the effects that we study here could potentially be present and are important to consider in setting future limits on DM annihilation and decay.

ACKNOWLEDGMENTS

APPENDIX A: ASTROPHYSICAL SYSTEMATICS

1. Uncertainties from annihilation in DM halos

Once structure formation begins, the DM annihilation rate is no longer set purely by the well-measured cosmological average DM density, but instead becomes dominated by annihilation in overdense regions, which can enhance $\langle \rho^2 \rangle$ greatly over $\langle \rho \rangle^2$. A model for this "boost factor" $\langle \rho^2(z) \rangle / \langle \rho(z) \rangle^2$ is included in the $f_c(z)$ factors that determine the amount of heating and ionization from s-wave DM annihilation [as defined in Eq. (8)], as discussed in [29].

On the one hand, the size of this enhancement is quite uncertain, due mostly to large uncertainties in the abundance and concentration of low-mass halos that cannot be resolved by cosmological simulations. On the other hand, at $z \sim 17$ the enhancement factor is expected to still be relatively modest; furthermore the heating and ionization at that epoch are determined by the integral over DM annihilation at all previous times, not only at $z \sim 17$, which also dilutes the effect of late-time overdensities (e.g., [35]). This last effect is stronger for secondary particles that take longer to cool and deposit their energy; in particular, for most energies of injection, photons take longer to deposit their energy than electrons, and thus the systematic uncertainty in $f_c(z)$ due to structure formation is expected to be smaller (as the typical photon contributing to late-time ionization/heating was injected at an earlier epoch where structure formation was less important).

To quantitatively estimate the uncertainty in our quoted annihilation limits due to uncertainties in the contribution from DM overdensities, we repeat the analysis using three different models from the literature for the boost factor. The first two are limiting cases from [29]: they correspond to (1) DM halos with Einasto density profiles [59], where we also include an estimate of substructure within main halos, and (2) DM halos with Navarro-Frenk-White profiles [60], without substructure included. The second model, with the lowest boost factor of those considered in [29], is the benchmark we use for the plots in the main text. Finally, the third model is (3) a simple analytic form proposed as a conservative model for the boost factor by [34]. We label these models as "high", "benchmark", and "low", respectively.

We show the limits on the annihilation cross section for DM annihilating to e^+e^- in the presence of Rutherford cooling, for these three models, in Fig. 13. Since the boost factor is approximately degenerate with the annihilation cross section (changing it can also lead to a slight
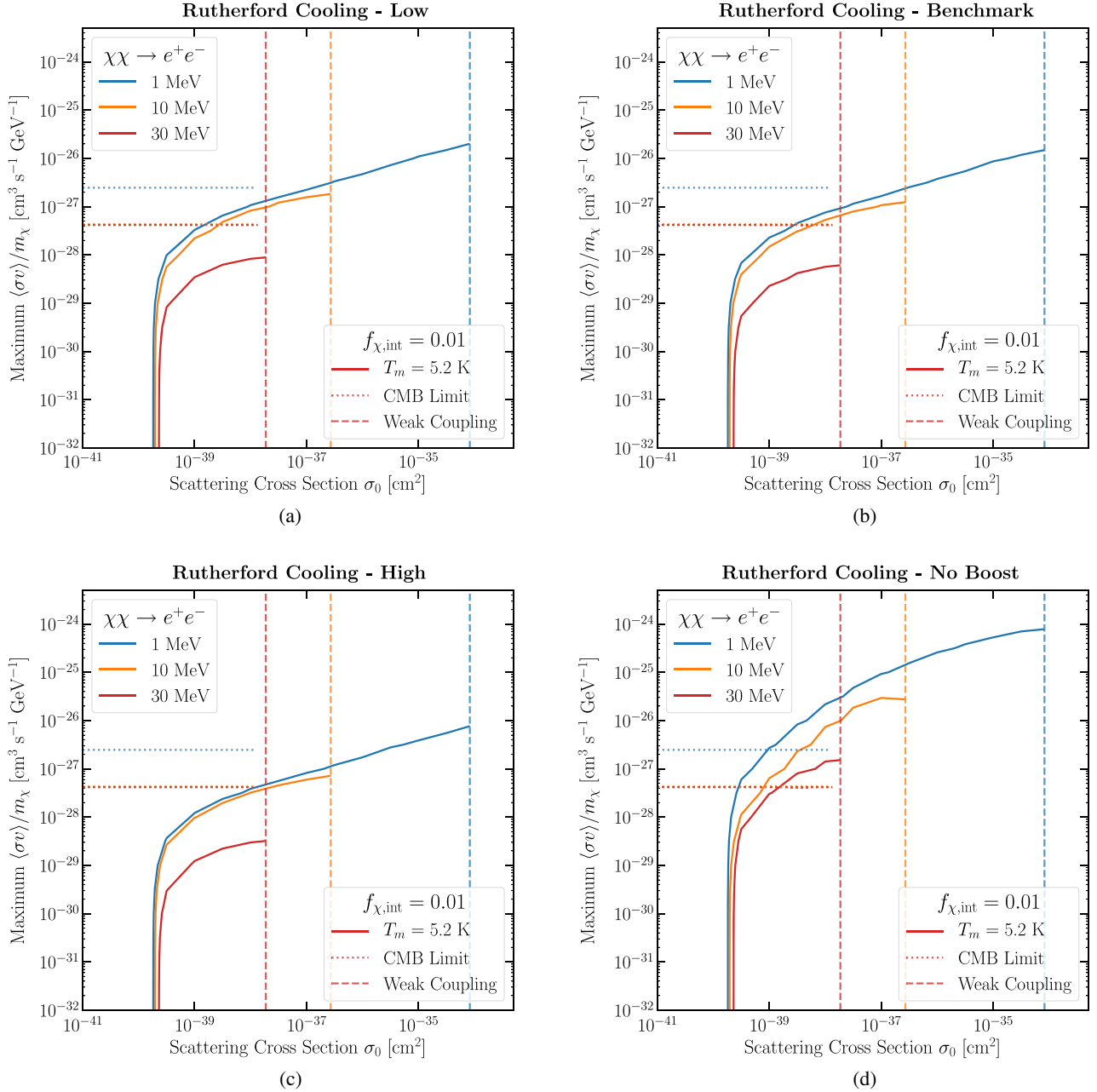
FIG. 13. Comparison of DM annihilation constraints when the gas is cooled by Rutherford scattering, where the "high" (upper left), "benchmark" (upper right), "low" (lower left), and "no boost" (lower right) models for the DM structure formation history are employed (see text for details).

modification of the redshift dependence of the annihilation rate), we expect the changes in the limits on the cross section to be similar for the other scenarios (early decoupling and additional radiation).

We find that, as expected, the constraints are more stringent for the "high" model, by roughly a factor of 2. However, the "benchmark" and "low" models agree closely, with the constraints differing on the 15%–20% level. We have performed the same check for DM annihilating to photons, and the difference between the models

is even smaller. Thus we expect our benchmark constraints to be similar to others set using a conservative structure formation model.

If we completely ignore structure formation and consider only annihilations in the smooth DM density, then the constraints weaken considerably, by about a factor of 50, as shown in the fourth panel of Fig. 13. This is the maximally conservative case, and is probably unrealistic; we leave a detailed study of the minimum possible boost factor to further work.
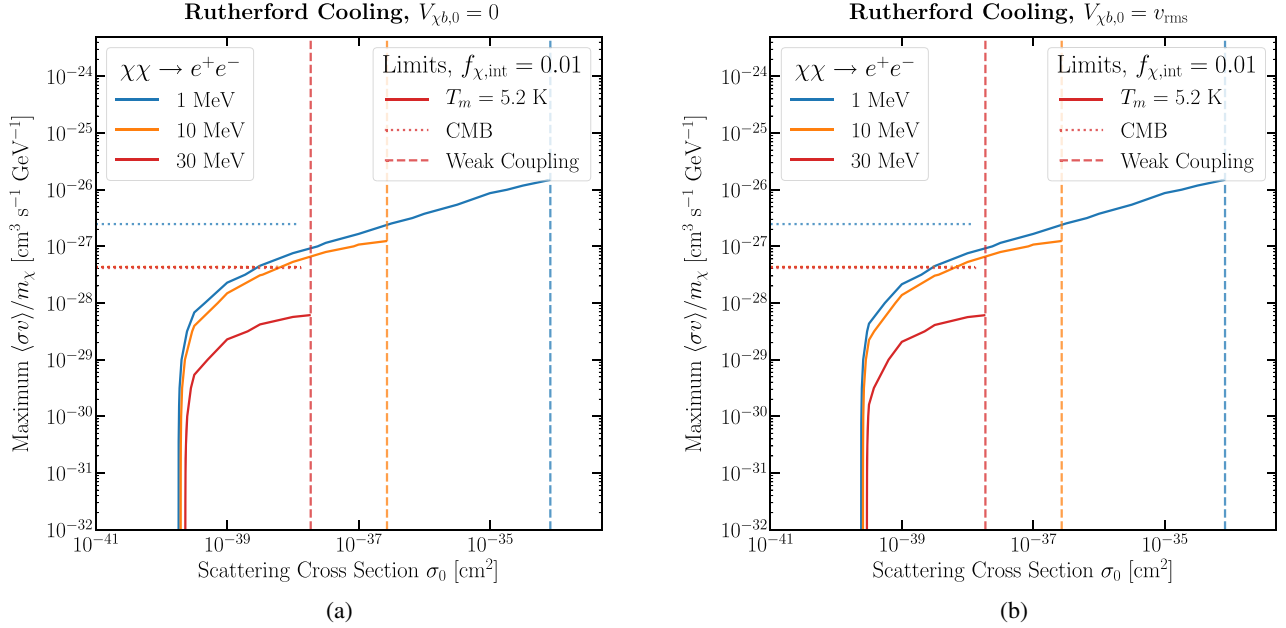
FIG. 14.   Comparison of DM annihilation constraints when the gas is cooled by Rutherford scattering, where the bulk DM-baryon relative velocity at recombination is taken to be $V_{\chi b,0} = 0$ (left) or $V_{\chi b,0} = v_{\rm rms}$ (right).

We note these uncertainties do not apply to DM decay, which probes the average DM density rather than the average DM density squared.

2. Uncertainties in the initial value of $V_{\chi b}$

In the scenario where Rutherford scattering cools the gas, the scattering rate depends on the relative velocity between the DM and baryons, and hence on the initial value of the bulk relative velocity $V_{\chi b,0}$ at recombination. As argued in Sec. V, we expect the cooling effect to be strongest for $V_{\chi b,0} = 0$, thus leaving the maximum amount of room for heating from annihilation/decay products.

We test this hypothesis in Fig. 14, comparing constraints on DM annihilating to e^+e^- for $V_{\chi b,0} = 0$ and $V_{\chi b,0} = v_{\rm rms} = 29 \text{ km s}^{-1}$, in the case where a 1% fraction of the DM participates in the scattering. We find that once the scattering cross section σ_0 is well above the value required to cool the gas to 5.2 K, the constraints on annihilation are unaffected by this change in the initial conditions, because the large baryon-DM scattering cross section induces a drag force that drives $V_{\chi b}$ to zero (for the interacting DM component). However, the minimum σ_0 needed to cool the baryons to that temperature does increase modestly (by a few tens of percent) when $V_{\chi b,0} = v_{\rm rms}$.

Accordingly we conclude that in the regime where the constraints are not very rapidly varying as a function of σ_0, away from the minimum σ_0 needed to achieve the required cooling, the systematic error due to neglecting the distribution of initial relative velocities $V_{\chi b}$ is small.

APPENDIX B: SUPPLEMENTAL PLOTS

In Figs. 15 and 16, we show zoomed-in versions of Figs. 2 and 3, to highlight the region where the additional radiation source is comparable to or smaller than the CMB, in terms of number density at a wavelength of 21 cm.

In Figs. 17 and 18, we show constraint plots for DM masses below 100 MeV in the presence of nonstandard recombination, for s-wave annihilation and decay, respectively. These analyses are otherwise performed as discussed in the main text.

Finally, Figs. 19 and 20 show the limits on the minimum decay lifetime and maximum annihilation cross section for Rutherford cooling with $f_{\chi,\rm int} = 1$. Values of σ_0 exceeding $\sim 10^{-42} \text{ cm}^2$ affect the CMB power spectrum significantly and are ruled out by Planck [47]; for models that are consistent with this limit, the value of T_m at $z \sim 20$ is a more powerful constraint on additional energy injection on models than the high-redshift CMB limits on annihilation and decay.
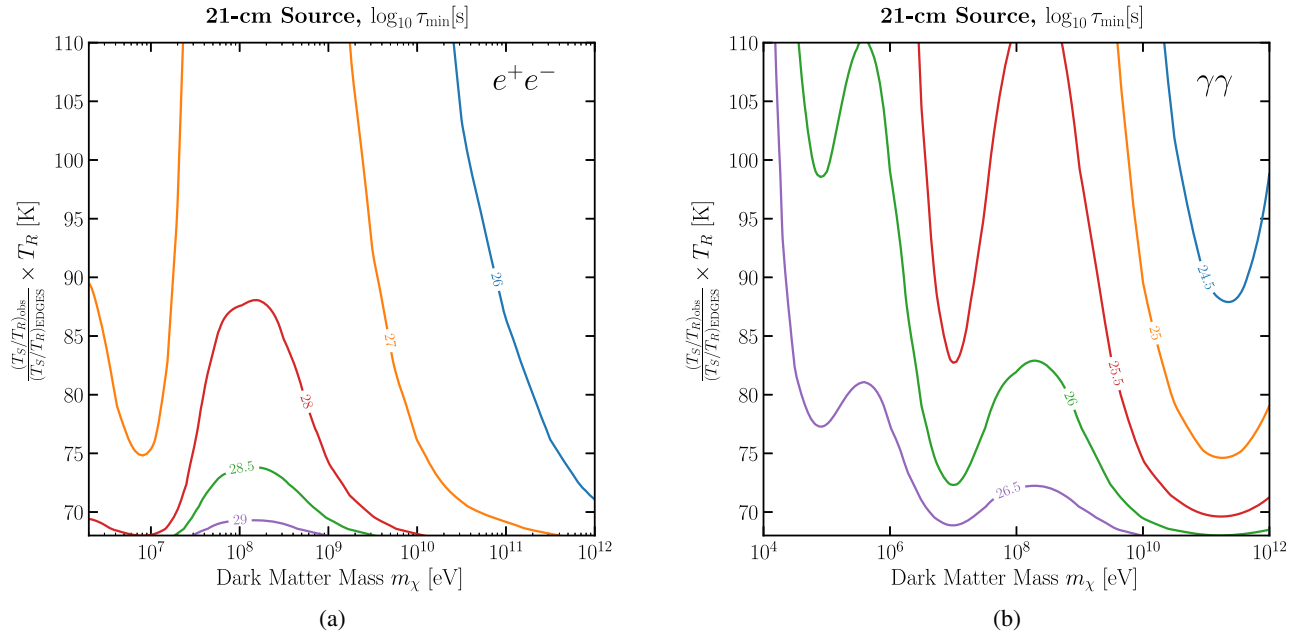
FIG. 15.   Decay lifetime constraints with an additional 21-cm source with $\chi \to e^+ e^-$ (left) and $\chi \to \gamma\gamma$ (right), as a function of m_χ and $(T_S/T_R)_{\mathrm{obs}}/(T_S/T_R)_{\mathrm{EDGES}} \times T_R$. This is a zoomed-in version of Fig. 2. Contour lines of constant minimum $\log_{10}\tau$ (in seconds) are shown.
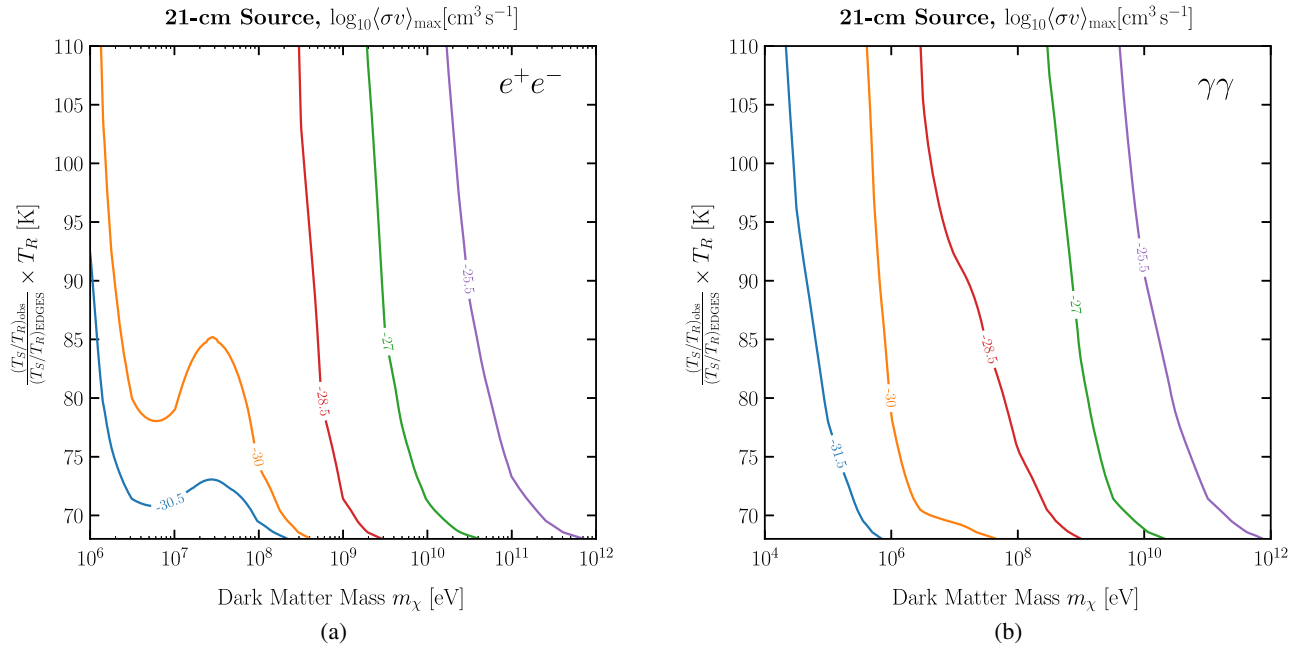

FIG. 16.   Annihilation cross section constraints with an additional 21-cm source with $\chi\chi \to e^+ e^-$ (left) and $\chi\chi \to \gamma\gamma$ (right), as a function of m_χ and $(T_S/T_R)_{\mathrm{obs}}/(T_S/T_R)_{\mathrm{EDGES}} \times T_R$. Contour lines of constant maximum $\log_{10}\langle\sigma v\rangle$ (in $\mathrm{cm}^3\,\mathrm{s}^{-1}$) are shown. This is a zoomed-in version of Fig. 3 The green contour corresponds to the canonical relic abundance cross section of $3 \times 10^{-26}\ \mathrm{cm}^3\,\mathrm{s}^{-1}$.
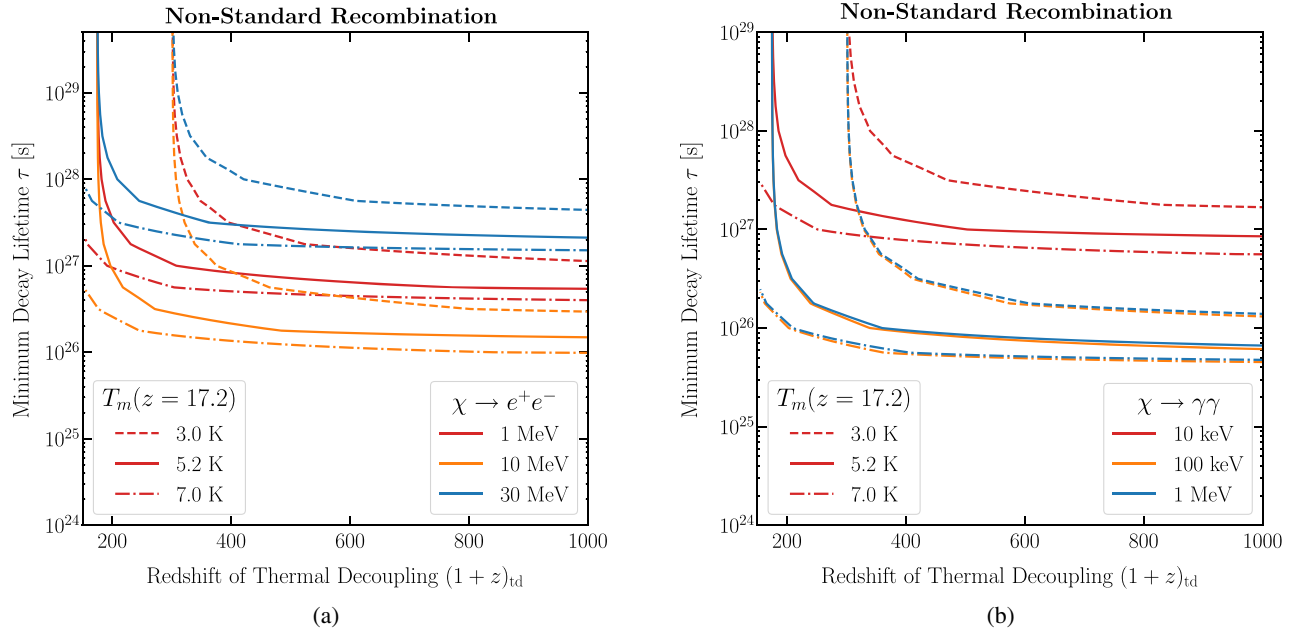
FIG. 17.   As Fig. 4, but extended to lower masses.

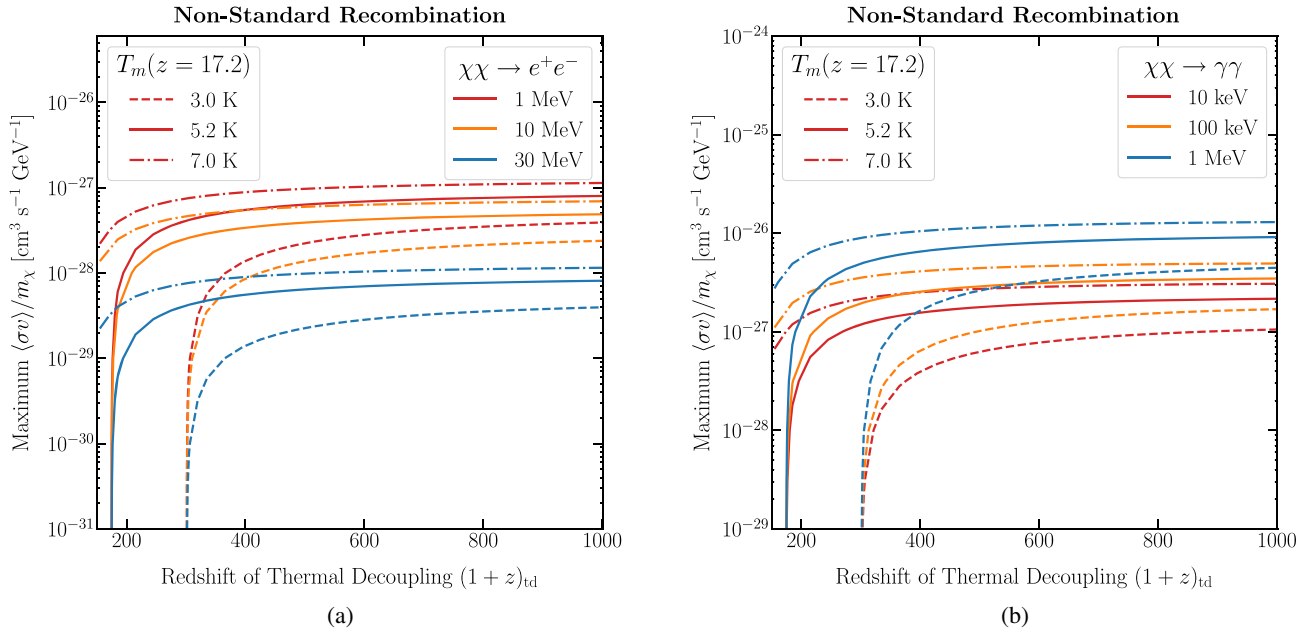FIG. 18.   As Fig. 5, but extended to lower masses.
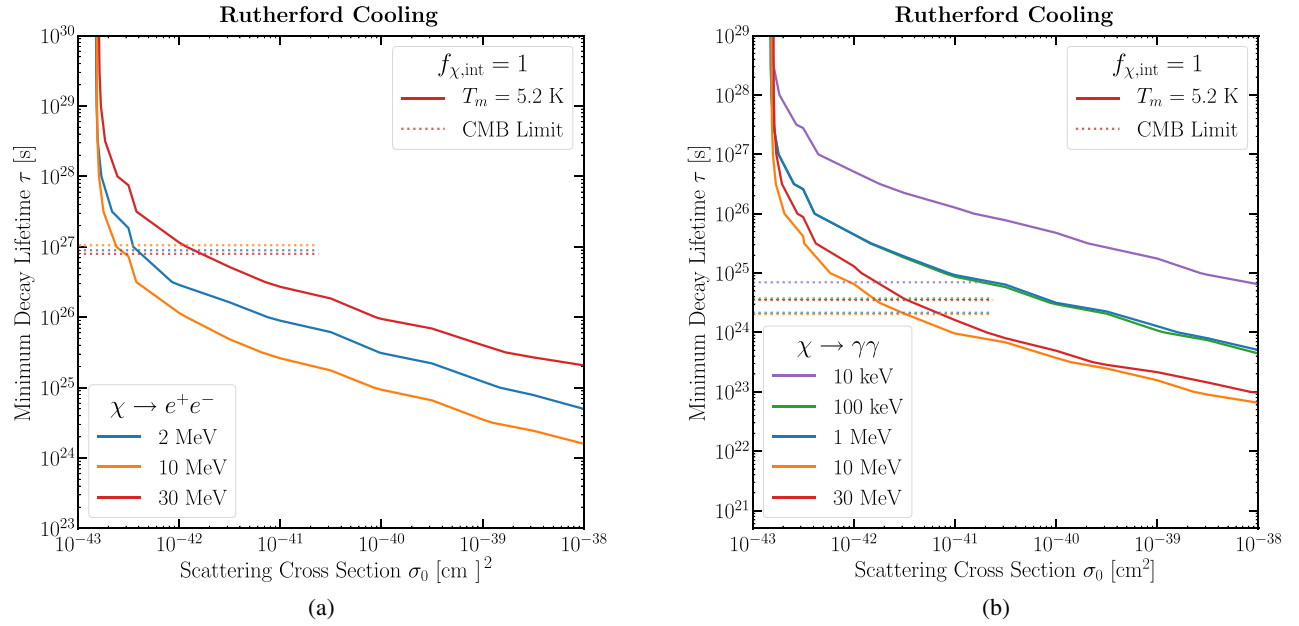
FIG. 19. Rutherford cooling constraints on the minimum decay lifetime for $\chi \to e^+ e^-$ (left) and $\chi \to \gamma\gamma$ (right) from the matter temperature $T_m(z = 17.2) = 5.2$ K (solid lines), $f_{\chi,\text{int}} = 1$. Limits from the Planck measurement of the CMB power spectrum are also shown up to $\sigma_0 = \sigma_{0,\text{td}}(z = 300)$ (dotted lines).
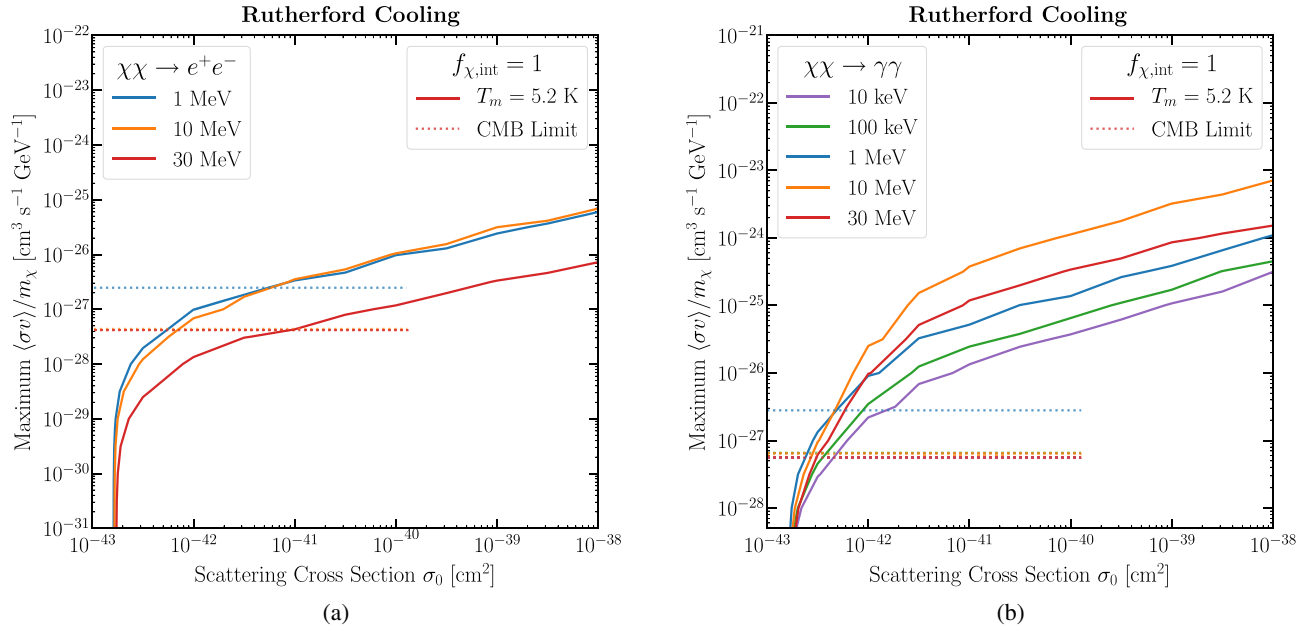

FIG. 20. Rutherford cooling s-wave annihilation constraints for $\chi\chi \to e^+ e^-$ (left) and $\chi\chi \to \gamma\gamma$ (right) from the matter temperature $T_m(z = 17.2) = 5.2$ K (solid lines), $f_{\chi,\text{int}} = 1$. Limits from the Planck measurement of the CMB power spectrum are also shown up to $\sigma_0 = \sigma_{0,\text{td}}(z = 600)$ (dotted lines).

[1] Y. Ali-Haimoud and C. M. Hirata, Phys. Rev. D **83**, 043513 (2011).

[2] J. Chluba and R. M. Thomas, Mon. Not. R. Astron. Soc. **412**, 748 (2011).

[3] J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen, and N. Mahesh, Nature (London) **555**, 67 (2018).

[4] M. Zaldarriaga, S. R. Furlanetto, and L. Hernquist, Astrophys. J. **608**, 622 (2004).

[5] S. Furlanetto, S. P. Oh, and F. Briggs, Phys. Rep. **433**, 181 (2006).

[6] S. A. Wouthuysen, Astron. J. **57**, 31 (1952).

[7] G. B. Field, Astrophys. J. **129**, 536 (1959).

[8] G. B. Field, Astrophys. J. **129**, 551 (1959).

[9] R. Barkana, Nature (London) **555**, 71 (2018).

[10] J. B. Muñoz and A. Loeb, Nature (London) **557**, 684 (2018).

[11] A. Berlin, D. Hooper, G. Krnjaic, and S. D. McDermott, Phys. Rev. Lett. **120**, 011102 (2018).

[12] S. Fraser et al., arXiv:1803.03245.

[13] R. Barkana, N. J. Outmezguine, D. Redigolo, and T. Volansky, arXiv:1803.03091.

[14] A. Falkowski and K. Petraki, arXiv:1803.10096.

[15] J. C. Hill and E. J. Baxter, arXiv:1803.07555.

[16] A. A. Costa, R. C. G. Landim, B. Wang, and E. Abdalla, arXiv:1803.06944.

[17] M. Pospelov, J. Pradler, J. T. Ruderman, and A. Urbano, arXiv:1803.07048.

[18] J.-O. Gong and N. Kitajima, arXiv:1803.02745.

[19] A. Ewall-Wice, T. C. Chang, J. Lazio, O. Dore, M. Seiffert, and R. A. Monsalve, arXiv:1803.01815.

[20] V. Poulin, J. Lesgourgues, and P. D. Serpico, J. Cosmol. Astropart. Phys. 03 (2017) 043.

[21] S. R. Furlanetto, S. P. Oh, and E. Pierpaoli, Phys. Rev. D **74**, 103502 (2006).

[22] L. Lopez-Honorez, O. Mena, A. Moliné, S. Palomares-Ruiz, and A. C. Vincent, J. Cosmol. Astropart. Phys. 08 (2016) 004.

[23] M. Valdes, A. Ferrara, M. Mapelli, and E. Ripamonti, Mon. Not. R. Astron. Soc. **377**, 245 (2007).

[24] C. Evoli, A. Mesinger, and A. Ferrara, J. Cosmol. Astropart. Phys. 11 (2014) 024.

[25] P. A. R. Ade et al. (Planck Collaboration), Astron. Astrophys. **594**, A13 (2016).

[26] P. J. E. Peebles, Astrophys. J. **153**, 1 (1968).

[27] Ya. B. Zeldovich, V. G. Kurt, and R. A. Sunyaev, Zh. Eksp. Teor. Fiz. **55**, 278 (1968); [Sov. Phys. JETP **28**, 146 (1969)].

[28] T. R. Slatyer, Phys. Rev. D **93**, 023521 (2016).

[29] H. Liu, T. R. Slatyer, and J. Zavala, Phys. Rev. D **94**, 063507 (2016).

[30] T. Venumadhav, L. Dai, A. Kaurov, and M. Zaldarriaga, arXiv:1804.02406.

[31] J. Chluba, in Proceedings, 49th Rencontres de Moriond on Cosmology: La Thuile, Italy, March 15-22, 2014 (Moriond, Paris, France, 2014), pp. 327–334.

[32] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright, Astrophys. J. **473**, 576 (1996).

[33] D. J. Fixsen et al., Astrophys. J. **734**, 5 (2011).

[34] G. D'Amico, P. Panci, and A. Strumia, Phys. Rev. Lett. **120**, 011103 (2018).

[35] T. R. Slatyer, Phys. Rev. D **87**, 123513 (2013).

[36] T. R. Slatyer and C.-L. Wu, Phys. Rev. D **95**, 023010 (2017).

[37] F. De Bernardis, R. Bean, S. Galli, A. Melchiorri, J. I. Silk, and L. Verde, Phys. Rev. D **79**, 043503 (2009).

[38] T. R. Slatyer, Phys. Rev. D **93**, 023527 (2016).

[39] M. Cirelli, G. Corcella, A. Hektor, G. Hutsi, M. Kadastik, P. Panci, M. Raidal, F. Sala, and A. Strumia, J. Cosmol. Astropart. Phys. 03 (2011) 051; 10 (2012) E01.

[40] J. B. Muñoz, E. D. Kovetz, and Y. Ali-Haïmoud, Phys. Rev. D **92**, 083528 (2015).

[41] E. G. Adelberger, B. R. Heckel, S. A. Hoedl, C. D. Hoyle, D. J. Kapner, and A. Upadhye, Phys. Rev. Lett. **98**, 131104 (2007).

[42] D. J. Kapner, T. S. Cook, E. G. Adelberger, J. H. Gundlach, B. R. Heckel, C. D. Hoyle, and H. E. Swanson, Phys. Rev. Lett. **98**, 021101 (2007).

[43] K. K. Boddy and V. Gluscevic, arXiv:1801.08609.

[44] V. Gluscevic and K. K. Boddy, arXiv:1712.07133.

[45] C. Dvorkin, K. Blum, and M. Kamionkowski, Phys. Rev. D **89**, 023519 (2014).

[46] W. L. Xu, C. Dvorkin, and A. Chael, Phys. Rev. D **97**, 103530 (2018).

[47] T. R. Slatyer and C.-L. Wu, arXiv:1803.09734.

[48] S. D. McDermott, H.-B. Yu, and K. M. Zurek, Phys. Rev. D **83**, 063509 (2011).

[49] A. D. Dolgov, S. L. Dubovsky, G. I. Rubtsov, and I. I. Tkachev, Phys. Rev. D **88**, 117701 (2013).

[50] Y. Ali-Haïmoud, P. D. Meerburg, and S. Yuan, Phys. Rev. D **89**, 083506 (2014).

[51] C. Boehm, M. J. Dolan, and C. McCabe, J. Cosmol. Astropart. Phys. 08 (2013) 041.

[52] A. A. Prinz et al., Phys. Rev. Lett. **81**, 1175 (1998).

[53] J. H. Chang, R. Essig, and S. D. McDermott, arXiv:1803.00993.

[54] T. Cohen, K. Murase, N. L. Rodd, B. R. Safdi, and Y. Soreq, Phys. Rev. Lett. **119**, 021102 (2017).

[55] A. Albert et al. (DES and Fermi-LAT Collaborations), Astrophys. J. **834**, 110 (2017).

[56] K. Cheung, J.-L. Kuo, K.-W. Ng, and Y.-L. S. Tsai, arXiv:1803.09398.

[57] S. Clark, B. Dutta, Y. Gao, Y.-Z. Ma, and L. E. Strigari, arXiv:1803.09390.

[58] A. Mitridate and A. Podo, J. Cosmol. Astropart. Phys. 05 (2018) 069.

[59] J. Einasto, Trudy Inst. Astrofiz. Alma-Ata **51**, 87 (1965).

[60] J. F. Navarro, C. S. Frenk, and S. D. M. White, Astrophys. J. **462**, 563 (1996).