Effects of dark matter decay and annihilation on the high-redshift 21 cm background

Steven R. Furlanetto*
Yale Center for Astronomy and Astrophysics, Yale University, 260 Whitney Avenue, New Haven, Connecticut 06511, USA

S. Peng Oh
Department of Physics, University of California, Santa Barbara, California 93106, USA

Elena Pierpaoli
Physics and Astronomy Department, University of Southern California, Los Angeles, California 90089-0484, USA
and Theoretical Astrophysics, California Institute of Technology, Mail Code 130-33, Pasadena, California 91125, USA

(Received 10 August 2006; published 1 November 2006)

The radiation background produced by the 21 cm spin-flip transition of neutral hydrogen at high redshifts can be a pristine probe of fundamental physics and cosmology. At \( z \sim 30–300 \), the intergalactic medium (IGM) is visible in 21 cm absorption against the cosmic microwave background (CMB), with a strength that depends on the thermal (and ionization) history of the IGM. Here we examine the constraints this background can place on dark matter decay and annihilation, which could heat and ionize the IGM through the production of high-energy particles. Using a simple model for dark matter decay, we show that, if the decay energy is immediately injected into the IGM, the 21 cm background can detect energy injection rates \( \geq 10^{-24} \text{ eV cm}^{-3} \text{ sec}^{-1} \). If all the dark matter is subject to decay, this allows us to constrain dark matter lifetimes \( \leq 10^{27} \text{ sec} \). Such energy injection rates are much smaller than those typically probed by the CMB power spectra. The expected brightness temperature fluctuations at \( z \sim 50 \) are a fraction of a mK and can vary from the standard calculation by up to an order of magnitude, although the difference can be significantly smaller if some of the decay products free stream to lower redshifts. For self-annihilating dark matter, the fluctuation amplitude can differ by a factor \( \leq 2 \) from the standard calculation at \( z \sim 50 \). Note also that, in contrast to the CMB, the 21 cm probe is sensitive to both the ionization fraction and the IGM temperature, in principle allowing better constraints on the decay process and heating history. We also show that strong IGM heating and ionization can lead to an enhanced \( \text{H}_2 \) abundance, which may affect the earliest generations of stars and galaxies.

DOI: 10.1103/PhysRevD.74.103502 PACS numbers: 95.35.+d, 98.62.Ra, 98.70.Vc

I. INTRODUCTION

The cosmic “dark ages,” stretching from the last scattering surface of the cosmic microwave background (CMB) at \( z \sim 1100 \) to the formation of the first luminous sources at \( z \sim 30 \), are one of the last frontiers for observational cosmology. The problem is the difficulty of finding any probes: because the intergalactic medium (IGM) has mostly decoupled from the CMB, there is no background radiation field against which we can study it, and (by definition) there are no local light sources.

This is unfortunate because, at least in principle, the physics of the dark ages is sufficiently simple that we can hope to understand it in detail. The only factors that enter are the CMB, the expanding Universe, recombinations, Compton scattering (which couples the CMB to the IGM), and gravitational growth—almost entirely in the linear regime. Thus, the IGM properties during the dark ages (and especially its fluctuations) constitute a strong test of fundamental cosmological parameters in an entirely analogous way to the CMB. Conversely, if we take the physics as well understood, probes of the dark ages would offer stringent tests of exotic physics.

Two examples are dark matter decay and annihilation. If either of these processes happens (even for only a fraction of the dark matter), they could inject high-energy photons (or other particles) into the IGM. These would then scatter and deposit some or all of their energy as ionizations and heat. An altered reionization history may affect the total optical depth and, as a consequence, the CMB power spectrum. Decaying dark matter particles were in fact initially advocated as a possible explanation for the high optical depth detected in the first year of WMAP data [1–6]. Such data allowed an upper limit on the energy injection rate and consequently on the decay time \( (t \geq H_0^{-1}) \). (These constraints, and others mentioned throughout this paper, were obtained with the first-year WMAP data, but we do not expect the newer data [7] to improve them substantially given the many uncertainties on the various reionization processes affecting the total optical depth.)

However, the CMB is only affected if a substantial fraction of the CMB photons interact with the IGM—or in other words if the optical depth to electron scattering is substantially altered by the decay products [1,5,8]. Compared to the IGM temperatures of \( \sim 10–10^3 \) K at \( z \sim 30–300 \), ionization requires a substantial energy input, and decays with long timescales (or low-energy injection rates) cannot be ruled out with the CMB. Nevertheless, because

---

*Electronic address: steven.furlanetto@yale.edu
the IGM temperature is so small, these models can still significantly affect the thermal history of the IGM [1,8] and hence the history of star formation. A number of recent particle physics models motivate such scenarios (see the discussion below in Sec. III A).

Dark matter annihilation has also been considered recently, primarily as an explanation for gamma-ray sources in the local Universe (e.g., [9]). But any such scenario will also predict an annihilation background from encounters between IGM particles; because the physical density increases like \((1 + z)^3\), such events could be relatively common in the early Universe and so affect the CMB [8,10]. Of course, these scenarios will continue to affect the IGM throughout the dark ages as well.

The 21 cm background can offer powerful constraints on dark matter decay and annihilation (or indeed any exotic process that injects energy into the IGM during the dark ages). The hyperfine level populations of the ground state of neutral hydrogen are determined through competition between absorption of (and emission stimulated by) CMB photons (and possibly UV photons) and collisions. Since the IGM is dense at high-redshifts, collisions dominate, driving the level populations into equilibrium with the kinetic temperature of the gas. Because the latter is colder than the CMB, the IGM is a net absorber of CMB photons at \(z \sim 30–300\). Thus, the 21 cm transition can be used to map fluctuations in the IGM during the dark ages, offering a pristine probe of cosmology [11]. Previous work has emphasized the possibility of constraining the matter power spectrum with this tool [11–14]. Here we point out that the fundamental properties of the fluctuations—on all scales—depend sensitively on the thermal history (see also Ref. [15]). Thus the 21 cm background can be used to constrain dark matter decay and annihilation. It is much more sensitive than the CMB to dark matter decay because (i) it depends on the thermal history, not just the ionized fraction, and (ii) it is directly sensitive to the late time behavior (when most of the energy is injected). It is less useful for constraining annihilation scenarios because in that case a large fraction of the energy is injected at early times.

The excess heating and ionization induced by dark matter decay and annihilation can also affect the chemistry of the IGM. In the simple chemical environment of the primordial IGM, their most important effect will be on the \(\text{H}_2\) abundance, which is significant because that molecule is an important coolant for low-temperature gas [16,17] and is thought to play a key role in the formation of the first stars [18–21]. Unfortunately, the implications—and even magnitude—of any possible boost to early structure formation are controversial [10,22,23]. Here, we provide a more detailed look at \(\text{H}_2\) formation and incorporate some hitherto neglected effects in the calculation.

The rest of this paper is organized as follows. We briefly review the physics of the 21 cm background in Sec. II. We describe our simple model for dark matter decay and annihilation, and review some particle physics motivation, in Sec. III. We then present our results for the 21 cm background in Sec. IV and for the \(\text{H}_2\) abundance in Sec. V. Finally, we discuss their implications in Sec. VI.

In our numerical calculations, we assume a cosmology with \(\Omega_m = 0.26\), \(\Omega_{\Lambda} = 0.74\), \(\Omega_b = 0.044\), \(H = 100h\) km s\(^{-1}\) Mpc\(^{-1}\) (with \(h = 0.74\)), \(n = 0.95\), and \(\sigma_8 = 0.8\), consistent with the most recent measurements [7], although we have increased \(\sigma_8\) from the best-fit WMAP value to improve agreement with weak lensing. We quote all distances in comoving units, unless otherwise specified.

\section{II. The 21 cm Background}

We review the relevant characteristics of the 21 cm transition here; we refer the interested reader to Ref. [24] (and references therein) for a more comprehensive discussion. The 21 cm brightness temperature (relative to the CMB) of a patch of the IGM is

\[
\delta T_b = 27 x_{\text{HI}} (1 + \delta) \left( \frac{\Omega_b h^2}{0.023} \right) \left( \frac{0.15}{10} \frac{1 + z}{T_{\text{CMB}}} \right)^{1/2} \left( \frac{T_S - T_{\text{x}}}{T_S} \right) \times \left[ \frac{\frac{H(z)(1 + z)^{3/2}}{3}}{dV/d\eta} \right] \text{mK},
\]

where \(\delta\) is the fractional overdensity, \(x_{\text{HI}} = 1 - x_i\) is the neutral fraction, \(x_i\) is the ionized fraction, \(T_{\text{x}}\) is the spin temperature, \(T_{\gamma}\) is the CMB temperature, and \(dV/d\eta\) is the gradient of the proper velocity along the line of sight. When \(T_S < T_{\gamma}\), the IGM appears in absorption. The last factor accounts for redshift-space distortions [12,25].

Before the first luminous sources turn on, the spin temperature \(T_S\) is determined by competition between scattering of CMB photons, collisions [26], and scattering of Ly\(\alpha\) photons [27–29]. In equilibrium (which is achieved rapidly),

\[
T_S^{-1} = T_{\gamma}^{-1} + x_e T_{\text{K}}^{-1} + x_e T_{\text{K}}^{-1} \frac{T_{\text{K}}}{1 + x_e + x_{\alpha}}.
\]

Here \(x_e\) is the total collisional coupling coefficient, including both H–H interactions [30,31] and H–e\(^-\) collisions [32,33]. We will denote the separate coefficients via \(x_{e_{\text{H}}H}\) and \(x_{e_{\text{H}}H}\), respectively. The last term describes Wouthuysen-Field coupling: \(x_{\alpha}\) is the coupling coefficient and \(T_{\text{K}}\) is the effective color temperature of the radiation field [29,34]. Typically \(T_{\text{K}} \approx T_{\text{K}}\) in the IGM, and

\[
x_{\alpha} = 1.81 \times 10^{11} \frac{(1 + z)^{-1} S_{\alpha} J_{\alpha}}{S_{\alpha} J_{\alpha}},
\]

where \(S_{\alpha}\) is a factor of order unity describing the detailed scattering process [29,34–36] and \(J_{\alpha}\) is the radiation background at the Ly\(\alpha\) frequency, in units of photons cm\(^{-2}\) s\(^{-1}\) Hz\(^{-1}\).

From Eq. (1) it is obvious that fluctuations in the density, temperature, ionized fraction, radiation background, and velocity all source fluctuations in the brightness tempera-
ture. Because, to linear order in \( k \) space, velocity perturbations are simply proportional to density perturbations, we can write the Fourier transform of the fractional 21 cm brightness temperature perturbation as

\[
\delta_{21}(k) = (\beta + \mu^2) \delta + \beta_H \delta_H + \beta_a \delta_a + \beta_T \delta_T,
\]

where \( \mu \) is the cosine of the angle between the line of sight and the wavevector \( k \) and each \( \delta_i \) describes the fractional variation in a particular quantity: \( \delta_a \) for the Ly\( \alpha \) coupling coefficient \( x_a \), \( \delta_H \) for the neutral fraction, and \( \delta_T \) for \( T_K \). The expansion coefficients \( \beta \) are

\[
\beta = 1 + \frac{x_c}{x_{tot}(1 + x_{tot})},
\]

\[
\beta_H = 1 + \frac{x_c^{HH} - x_c^H}{x_{tot}(1 + x_{tot})},
\]

\[
\beta_a = \frac{x_a}{x_{tot}(1 + x_{tot})},
\]

\[
\beta_T = \frac{T_T}{T_K - T_T} + \frac{x_c}{x_{tot}(1 + x_{tot})} \frac{d \ln x_c}{d \ln T_K},
\]

where \( x_{tot} \equiv x_c + x_a \). These terms are all relatively easy to understand [24]. The first terms in \( \beta \) and \( \beta_H \) simply reflect the proportionality between \( \delta T_K \) and the density of neutral atoms, while the second terms account for the variation in the collisional coupling rate with density. \( \beta_a \) takes this form because, of course, only the fraction of the coupling due to the Ly\( \alpha \) background matters for it. Finally, the first term in \( \beta_T \) parameterizes the speed at which \( T_S \) responds to fluctuations in \( T_K \), while the second takes into account the explicit dependence of the collision rates on temperature.

### III. DARK MATTER AND THE IGM

The implications of dark matter decay (and/or annihilation) for the thermal and ionization histories of the IGM have been considered by a number of authors [1–6,10]. Here we first review some particle physics scenarios that can affect the dark ages, and then we show how to compute the resulting IGM histories.

#### A. Some example particle physics scenarios

Dark matter particles can ionize and heat the IGM through either decay or annihilation.

For decay, short lifetimes (\( t_X \lesssim 3 \times 10^{18} \) sec, with \( X \) denoting the dark matter particle) are disfavored, as they cause the CMB power spectra to disagree with existing data [1,5]. On the other hand, particles with exceptionally long lifetimes, such as gravitinos with mass \( m_X c^2 \approx 10–100 \) MeV (and hence \( t_X \approx 10^{31} \) sec [37]) only affect the IGM temperature at low redshift (\( z \lesssim 2 \)) [8]. We wish to consider particles with lifetimes in the range \( t_X \sim 10^{22}–10^{27} \) sec.

One such example is an axino with mass \( m_X c^2 \approx 1–100 \) MeV; positrons produced as their decay products could explain the 511 keV emission excess from the Galactic center [37]. Axinos in this mass range have \( t_X \approx 3 \times 10^{24}–10^{26} \) (MeV/\( m_X c^2 \)) sec. The ionized fraction produced by such particles is too low to be observable with the CMB [8], but the IGM temperature departs from the standard case by \( z \approx 100 \) and at \( z \approx 20 \) is \( \sim 20 \) times higher, if energy transfer is perfectly efficient [8]. However, in these particular models a detailed accounting of the interactions between decay products and the IGM substantially decreases their effect and pushes it to lower redshifts [22].

Another possible candidate is a sterile neutrino with \( m_X c^2 \sim 2–4 \) keV. Again, even for the highest mass in this range, the ionized fraction produced by these particles is small (\( x_i \lesssim 0.1 \) at \( z = 0 \)), but the present-day IGM temperature increases by several orders of magnitude (although again the energy transfer is likely not perfectly efficient; [8,22]). Such a neutrino is unlikely to be the sole dark matter component [38], but it may still constitute a non-negligible fraction of the dark matter. Such neutrinos have decay times \( t_X \approx 3 \times 10^{27} \) sec.

Decays of superheavy dark matter particles (\( m_X c^2 \gtrsim 10^{22} \) GeV) have been considered as possible sources of reionization [39,40] and also as a production mechanism for ultra-high-energy cosmic rays [41,42]. With such energetic particles, the precise ionization and heating strongly depend on the type of reactions assumed.

Two kinds of annihilating dark matter have been investigated in connection with nonstandard reionization: neutralino annihilation (e.g. [43,44]), involving particles with \( m_X c^2 \gtrsim 30 \) GeV, and light dark matter annihilation (\( m_X c^2 \approx 1–100 \) MeV [45]). While the expected ionization fraction is small, it extends to high redshifts, greatly affecting the CMB power spectrum. These models can already be constrained with current CMB data and will be better constrained by future CMB polarization measurements [8,10]. In the most reasonable models, the IGM temperature can be several times larger than normally expected at \( z \sim 20–30 \).

The effects of light dark matter annihilation on CMB power spectra have also been considered. CMB data still allow models with small ionization fractions but with IGM temperatures a few times larger than in the standard scenario (e.g., \( \sim 2 \) times larger at \( z = 60 \); [22,46]).

#### B. Energy deposition

Rather than examine each of these models individually, we will follow the method of Ref. [1], who presented a simple but general model for the effects of dark matter decay. We first suppose that the dark matter particle has a decay rate \( \Gamma_X = t_X^{-1} \), so that the physical number density of particles is

\[
n_X(z) = n_X^0 (1 + z)^{3} e^{-\Gamma_X z},
\]

where \( n_X^0 \) is the comoving number density. In our calcu-
simple models the rules energy and the ionized fraction \[48\], to the accuracy of our will write than attempt to compute the relevant processes in detail, we

\[ Q^{\text{thick}} = f \Gamma_X n_X m_X c^2, \]

where \( f \) is the fraction of decay energy that is potentially available to the IGM (e.g., excluding particles that free stream without interacting, such as neutrinos); the meaning of the superscript will become apparent momentarily. It is convenient to define \( \xi_X = \Gamma_X \Omega_X / \Omega_b = 5.9 \Gamma_X \) to be the decay rate normalized for notational convenience to the baryon fraction (i.e., the equivalent decay rate if it were the baryons that decayed). Then \( Q^{\text{thick}} = f \xi_X n_0^0 m_p c^2 \), where \( n_0^0 \) is the comoving number density of baryons.

For concreteness, we will assume that the particle decays to high-energy photons (but see below), which then either scatter through the IGM or free stream to lower redshifts. High-energy photons can interact with their surroundings through photoionization, Compton scattering, pair production (via scattering off neutral atoms, free electrons, CMB photons, or ions), and scattering with CMB photons [47]. At the high redshifts of interest here, photoionization and scattering off of free electrons and atoms dominates energy loss at low photon energies (\( E \lesssim 3 \text{ keV} \); in this regime, the IGM is opaque. At high energies (\( E \gtrsim 10^{-10^3} \text{ GeV}, \) depending on redshift), pair production off CMB photons dominates, and the total optical depth is also large. The resulting pairs rapidly deposit their energy in the IGM. However, between these thresholds lies a “transparency window” of relatively low optical depth; photons in this regime experience optical depths (over a Hubble length) \( \tau \sim 10^{-2-1} \) at \( z = 100 \) [1]. Thus they lose only a fraction of their energy to the IGM, with the rest free streaming to the present or being lost to cosmic expansion.

The decay energy can be deposited in heating, collisional excitation, or ionization; while the fractions \( \chi_b, \chi_e, \) and \( \chi_i \), in each process actually depend on the initial photon energy and the ionized fraction [48], to the accuracy of our simple models the rules \( \chi_i \sim \chi_e \sim (1 - x_i)/3 \) and \( \chi_b \sim (1 + 2x_i)/3 \) suffice [1].

When the decay products experience a large optical depth, the energy deposition rate is simply \( Q^{\text{thick}} \) (hence the superscript). However, if the energy is injected in the transparency window, we have instead

\[ Q^{\text{thin}} = Q^{\text{thick}}[1 - e^{-c\tau(z)}] = Q^{\text{thick}} \tau(z), \]

where \( \tau(z) = c n_b(z) \sigma_{\text{eff}} / H(z) \) is the effective optical depth for interactions over a Hubble length. Here \( \sigma_{\text{eff}} \) is the energy-averaged cross section for interactions. Rather than attempt to compute the relevant processes in detail, we will write

\[ \tau(z) = \tau_{100} \left( \frac{1 + z}{100} \right)^{3/2}, \]

where \( \tau_{100} \) is the optical depth at \( 1 + z = 100 \) and the exponent assumes a constant \( \sigma_{\text{eff}} \) (so that \( \tau \propto n_b / H \)). More detailed calculations of \( \tau(z) \) can take into account the specific interactions with the IGM for a given particle physics model (e.g., [22]), but we forego these here to keep our calculations generic. Thus decay into the transparency window can be described in the same way as instantaneous energy injection provided we make the replacement \( \xi_X \rightarrow \xi_X [1 - \exp(-\tau)] \). In this case, most of the energy would free-stream to the present-day and contribute to the x-ray and γ-ray backgrounds. If the decay photons are monoengetic, the observed background provides fairly powerful constraints: \( \xi_X \lesssim 10^{-25} \text{ s}^{-1} \) for \( m_X c^2 = 10 \text{ MeV} \) or \( \xi_X \lesssim 10^{-30} \text{ s}^{-1} \) for \( m_X c^2 = 100 \text{ GeV} \) [1]. However, in the general decay case for which the photons have a wide range of energies, these limits become unimportant and much larger energy deposition rates are allowed. Thus our fiducial models below cannot yet be ruled out.

We will also consider the case of dark matter annihilation. Neglecting the evolving clumpiness of the dark matter field, the comoving energy deposition rate is

\[ Q^{\text{ann}} = 2 f m_X c^2 (n_0^0)^2 \langle \sigma v \rangle (1 + z)^3, \]

where the factor of 2 occurs because two particles annihilate in each event and \( \langle \sigma v \rangle \) is the velocity-averaged annihilation cross section. In this case, we can define an effective baryon-normalized “lifetime” via

\[ \xi_X^{\text{ann}} = \frac{\Omega X 000^0}{m_X} \langle \sigma v \rangle (1 + z)^3 \left( \frac{\Omega_X}{\Omega_b} \right). \]

Like decay into the transparency window, annihilation can be thought of as a redshift-dependent decay rate. Obviously, annihilation injects relatively more energy at high redshifts. For particle decay, the free parameter is \( m_X \Gamma_X \), which is easy to interpret. For annihilation, it is the less intuitive quantity \( \langle \sigma v \rangle / m_X \). Reasonable values for neutralino annihilation are \( \langle \sigma v \rangle = 2 \times 10^{-26} \text{ cm}^3 \text{ s}^{-1} \) and \( m_X c^2 = 100 \text{ GeV} \) (e.g., [10]). For illustrative purposes, we will consider a range of cross sections up to \( \langle \sigma v \rangle = 2.5 \times 10^{-25} \text{ cm}^3 \text{ s}^{-1} \) (with \( f = 1 \))—somewhat above the upper limit provided by existing CMB data [8]. (Note that our value of \( \langle \sigma v \rangle \) is 4 times smaller than that used by Ref. [8], because they assumed \( f = 0.25 \).)

Although we have phrased this discussion in terms of photons, our model is actually much more general. Generically, the primary decay products will not be photons; however, most of the energy that is ultimately transferred to the IGM will pass through an intermediate photon stage. Most decay products (e.g., muons, heavy quarks, etc.) are unstable and will eventually cascade to electron-positron pairs, photons, and neutrinos. Energetic electrons (and positrons) lose most of their energy by inverse-
Compton scattering CMB photons [1]; the resulting energetic photons will be rapidly absorbed unless they lie within the transparency window (which occurs if the electron energy is in the range $1 \text{ GeV} - 50 \text{ TeV}$). Thus, even in this case photon interactions with the IGM are probably the most important ones. We hide our ignorance of the non-interacting component, including neutrinos, in the factor $f$. Note that decay into protons, which are highly penetrating, must also be included in this noninteracting component.

In any case, so long as the IGM is optically thick to the decay products, their precise composition does not strongly affect our results: they depend primarily on the net rate of energy deposition, regardless of the source. There are, however, two subtle ways in which the decay products could ultimately be diagnosed using the 21 cm background. First, as we will argue below, the heating rate will vary with the local density and so depends on the detailed interactions between the energetic particles and the IGM. Second, if the decay products emerge in some optically thin component other than photons, they may experience a different scaling of energy deposition with redshift (compared to the $[1 + z]^{3/2}$ dependence for photons). However since there is no leading candidate for such decay products we do not speculate on this possibility further.

C. The mean evolution of the IGM

We now wish to describe how these processes affect the IGM; our approach in this section is similar to Ref. [15]. First consider the effects on the globally-averaged temperature and ionization fraction. Without dark matter, the mean IGM temperature $T_K$ is determined by the competition between adiabatic cooling and Compton scattering of CMB photons off of the residual free electrons. We obtain [1,5,49]

$$\frac{d \tilde{T}_K}{dt} = -2 \tilde{T}_K H(z) + \frac{x_i(z)}{\eta_1} (T_y - \tilde{T}_K) - 2 \frac{\eta_2}{\eta_1} \frac{m_p c^2}{k_B} \tilde{\xi} x_i \frac{\dot{n}_b}{n_b},$$

(15)

where $T_y = (3m_e c)/(8\sigma_T u_{\text{CMB}})$, $u_{\text{CMB}}$ is the CMB energy density at redshift $z$, $\eta_1 = (1 + f_{\text{He}} + \xi)$, $\eta_2 = (1 + 4f_{\text{He}})$, and $f_{\text{He}}$ is the helium fraction by number (we assume helium remains neutral for simplicity). The mean ionized fraction evolves according to

$$\frac{dx_i}{dt} = -\alpha x_i^2 n_H + \eta_2 \frac{m_p c^2}{E_{\text{ion}}} \tilde{x}_i \dot{\xi} x_i,$$

(16)

where the first term describes recombinations with an effective coefficient $\alpha$ (taken from Ref. [49]) and $n_H$ is the total mean density of ionized and neutral hydrogen nuclei (note that gas clumping is negligible at the redshifts of interest here). The second term results from dark matter decay; $E_{\text{ion}} = 13.6 \text{ eV}$ is the hydrogen ionization threshold. Most ionizations are caused by collisions with hot secondary photoelectrons rather than by direct photoionization, with a total number of ionizations $\sim E_{\nu}/(37 \text{ eV})$ for a photon of energy $E_{\nu}$ at $\tilde{x}_i \ll 1$ [48]—which agrees well with our assumption that $x_i \sim (1 - x_i)/3$. Note that, by increasing $\tilde{x}_i$, dark matter decay also makes Compton scattering more efficient in Eq. (15), which provides a second channel for heat input. This modification to the reionization history affects the CMB temperature and polarization power spectra through Thomson scattering during the dark ages. The WMAP experiment requires $\xi_X \lesssim 10^{-22} \text{ s}^{-1}$[1,5] for energy deposition in the optically thick regime.

A more subtle effect of the energy injection is to set up a global Ly$\alpha$ background, which can couple the spin and kinetic temperatures through the Wouthuysen-Field effect. This occurs through collisional excitations induced by fast photoelectrons (see the analogous discussions in the context of x-ray heating in [50–52]). Each hydrogen excitation will produce a cascade of line photons; the Lyman-series photon produced in the cascade will then scatter through the (extremely optically thick) IGM. Higher Ly$\nu$ photons will eventually cascade further, with $\sim 1/3$ of them being “recycled” into Ly$\alpha$ photons [29,53]. If a fraction $\chi_\alpha$ of the input energy eventually goes into Ly$\alpha$ photons, the background at the line is

$$J_\alpha = \frac{c}{4\pi} \frac{m_p c^2}{h \nu_\alpha} \frac{x_i \tilde{\xi} \chi h_b}{H(z)},$$

(17)

where $\nu_\alpha$ is the Ly$\alpha$ line frequency and $n_b$ is the number density of baryons. Interestingly, this can induce relatively large coupling:

$$x_\alpha \sim 4.5 \xi_{-24} S_{\alpha}(\frac{\chi_\alpha}{1/6}) \left(\frac{1 + z}{100}\right)^{1/2},$$

(18)

where $\xi_{-24} = \tilde{\xi}_X/(10^{-24} \text{ sec})$. As we shall see, this becomes potentially important at lower redshifts, when collisional coupling between $T_S$ and $T_K$ weakens. Estimating the precise coupling requires a detailed examination of the collisional excitation processes, which we will forego here given our generic and crude approach. We will simply assume that one-half of the total excitation energy goes into Ly$\alpha$ photons, i.e. $\chi_\alpha = \chi_\nu/2$ (see [52] for a more comprehensive discussion). This Ly$\alpha$ background also affects the IGM temperature, but the rates are completely negligible in the models we consider [34,36,54].

D. Fluctuations in the IGM

Eqs. (15)–(17) suffice to describe the evolution of the global 21 cm background. But a more interesting observable is how this signal fluctuates, which is easier to measure because the signal monopole is severely contaminated by foregrounds (see [24] and references therein). Equation (4) shows that these fluctuations have a variety of sources. How do variations in the energy deposition rate source 21 cm fluctuations? We can imagine some simple
cases for perturbations to $Q$. The first is uniform deposition. This would be appropriate if, for example, the decays produce particles with long mean-free paths that then deposit energy after interacting with a uniform background (such as the CMB). This is probably not realistic, because in most cases the majority of energy is ultimately transferred to the IGM through relatively low-energy photons or particles with short mean-free paths. In such a scenario, the local energy injection rate will be proportional to the density (because the energy deposition rate per particle is constant), even if the intermediate decay products have reasonably long mean-free paths. To distinguish these scenarios, we define a variable $\theta_u$ such that the energy deposition rate is proportional to $(1 + \theta_u \delta)$ (i.e., $\theta_u = 0$ for uniformly deposited energy). For annihilating particles, $Q \propto n_X^2 \propto (1 + \delta)^2$. If the energy is deposited immediately, we would have $\theta_u = 2$ (assuming that $\delta \ll 1$ and Taylor expanding); in the perhaps more plausible case that the intermediate products have long mean-free paths, $\theta_u = 1$ would be more appropriate.

The fractional perturbation to the gas temperature, $\delta_T$, obeys (cf., [12,14])

$$\frac{d\delta_T}{dt} = \frac{2}{3} \frac{d\delta}{dt} - \frac{\bar{x}_i}{T_\gamma} \left( T_\gamma \delta_T + \bar{x}_i \delta_T + \bar{\delta}_T \right),$$

where $\bar{\delta}_T = -\delta_{Th}(1 - \bar{x}_i)/\bar{x}_i$ is the fractional perturbation to the ionized fraction. The first term describes adiabatic compression or expansion, the second and third describe fluctuations in the rate at which energy is transferred through Compton scattering, and the last term describes how dark matter decay affects the fluctuations. (Note that we have assumed $\bar{x}_i \delta_T \ll \delta_T$, because the ionized fraction remains small in all of the models we consider. We have also ignored variations in $\chi_h$ and $\chi_i$ with the ionized fraction.) The corresponding equation for $\delta_i$ is

$$\frac{d\delta_i}{dt} = -\alpha \bar{x}_i \bar{n}_H (\delta_i + \delta + \alpha' \delta_T) - \eta_{\chi} \left( \frac{m_p c^2}{E_{ion}} \right) \bar{x}_i \chi_i [\delta_i + \delta(1 - \theta_u)],$$

where $\alpha' \equiv d \ln n_e/d \ln T_K$ is the second accounts for dark matter decays. Note that, in the absence of extra energy injection, $\delta_i = 0$ is a reasonable approximation at $z \lesssim 150$ (cf. [13]), because recombinations are slow in the largely neutral gas, but it must be included in the more general case.

The appearance of $\delta$, $\delta_T$, and $\delta_i$ in each of these equations implies that the different density modes source independent temperature and ionization modes. This introduces a nontrivial scale dependence into the calculation [13,14]. But for a simple estimate it suffices to follow only the growing density mode, $\delta \propto a$, which dominates on the large scales that may ultimately be observable [12].

Writing $\delta_T = g_T(z)\delta$ and $\delta_i = g_i(z)\delta$, we have

$$\frac{dg_T}{d\tau} = \frac{g_T - 2/3}{1 + z} + \frac{\bar{x}_i}{\eta_{1/2} \gamma} \left[ g_T(T_\gamma - \bar{T}) \right]$$

$$\frac{2}{3} \frac{\eta_{m_p c^2}}{\beta_k T_\gamma} \xi_{X} \chi_h \left[ (1 - \theta_u) + g_T \right] \frac{dT}{d\tau},$$

$$\frac{dg_i}{d\tau} = \frac{g_i + \alpha \bar{x}_i \bar{n}_H (1 + g_i + \alpha' g_T)}{1 + z} \frac{\bar{T}}{\bar{n}_H} \left( 1 + \frac{1}{z} \right),$$

$$\frac{\xi_{X} \chi_h}{(1 - \theta_u) + g_i},$$

These equations make the effects of dark matter perturbations obvious. First consider $g_T$. Adiabatic expansion and compression tend to drive $\delta_T \rightarrow 26/3$. But when Compton cooling is efficient (and if $\delta_i = 0$), it inputs a constant amount of energy per particle, driving the gas toward isothermality ($g_T \rightarrow 0$). Decay or annihilation energy that is injected in the same way (proportional to the local gas density) also tends to drive the gas toward isothermality. On the other hand, if the energy is injected uniformly, it drives $g_T \rightarrow -1$ because it must be shared between more particles in denser regions.

In Eq. (22), the first term holds $\delta_i$ constant in the absence of recombinations. The second term shows that recombinations tend to drive $g_i \rightarrow -1$, because denser gas combines more quickly. The dark matter decay term behaves similarly to its analog in Eq. (21): injection at a constant rate per particle damps out perturbations in $x_i$, while uniform injection preferentially decreases the ionized fraction in dense regions.

Finally, we must consider fluctuations in the Ly$\alpha$ background. If the energy is injected uniformly, $\delta_a$ is constant and $\delta_a = 0$ if $\theta_u = 1$, $\delta_a = \delta$. Note the contrast with $\delta_T$ and $\delta_i$: the coupling efficiency depends on the total Ly$\alpha$ background, not the background per particle, because each Ly$\alpha$ photon scatters many times.

Using the baryonic power spectrum $P_{\delta \delta}(k)$ and integrating Eqs. (21) and (22), we can estimate the fluctuation amplitude for $\delta_T$, $\delta_i$, and $\delta_a$ as a function of scale at any redshift. The power spectrum of the 21 cm fluctuations is then [see Eq. (4)] [24]

$$P_{21}(k, \mu) = \delta T^2_b(\beta' + \mu^2) P_{\delta \delta}(k),$$

where

$$\beta' = \beta + \beta_T g_T - \beta_n \bar{x}_i/g_i/(1 - \bar{x}_i) + \theta_u \beta_n.$$

For simplicity, we will average over the $\mu$ dependence when presenting our results. Note that, because the 21 cm power spectrum is simply proportional to the matter power spectrum on large scales, we will only present results at a single wavenumber $k = 0.04$ Mpc$^{-1}$. The
shape is straightforward to compute [11,13], but we are concerned with the differences induced by dark matter decay and annihilation. Because we have used only the growing mode, these differences are scale-independent. We have chosen \( k = 0.04 \) Mpc\(^{-1}\) because it is in the range of scales most accessible to observations (see Sec. VI).

IV. RESULTS: THE 21 CM BACKGROUND

A. Dark matter decay: optically thick regime

We have integrated the temperature and ionization equations with initial values taken from RECFAST [49] at \( z = 350 \) (dark matter decay at higher redshifts does not significantly affect the history). Figure 1 shows the resulting thermal and ionization histories for \( \xi \approx 1, 0.1, 10^{-2}, 10^{-3} \) and 0 from top to bottom. These correspond to lifetimes \( t_X \approx 6 \times (10^{24} - 10^{27}) \) sec, assuming that the decaying particle makes up a large fraction of the dark matter. The uppermost curve would significantly affect the CMB and can be ruled out by WMAP [1,5], but the others have almost no effect on it (the total optical depths to electron scattering from \( z > 10 \) are \( \tau_{\text{es}} = 0.083, 0.016, 0.0060, 0.0048 \), and 0.0046 in these models).

The most important point of Fig. 1 is that, even if \( \bar{x}_i \) remains small, \( \bar{T}_K \) can still increase significantly. This is of course simply because the IGM is so cold at these redshifts: if roughly equal amounts of energy go toward ionization and heating, we would expect \( \bar{T}_K \approx 10^3 (\bar{x}_i/0.1) \) K, far above the usual temperature. For the larger values of \( \xi_X \), the increased efficiency of Compton scattering in the more highly-ionized Universe exaggerates the effectiveness of heating. This aspect is invisible so far as the CMB is concerned, but the 21 cm background measures it directly.

Figure 2 shows the resulting 21 cm signals. Panel (a) shows the angle-averaged fluctuation amplitude at \( k = 0.04 \) Mpc\(^{-1}\) in the different decay scenarios of Fig. 1. In this (and in all figures unless otherwise specified), we assume \( \theta_\gamma = 1 \) (i.e., the energy deposition rate is proportional to the local density). Panel (b) shows the sky-averaged 21 cm brightness temperature as a function of redshift.

First consider the solid curves, which present the standard calculation without dark matter decay (e.g., [11,12]). In this case, Compton scattering becomes inefficient at \( z \approx 200 \), and the gas begins to cool below \( T_K \). At \( z \approx 80 \), the density is large enough for collisions to drive \( T_S \approx T_K \), so the IGM becomes visible in absorption. Below this redshift, \( \delta T_b \) returns to zero because collisional coupling becomes inefficient as the density decreases. The fluctuation amplitude behaves similarly, except with a peak at \( z \approx 50 \), because it is weighted by the growing density mode, \( \delta \propto (1 + z)^{-1} \) [11,12].

The effects of dark matter decay on \( \delta T_b \) are straightforward and compare well to Ref. [15]: by continually heating the IGM, it decreases the intensity of absorption or even turns it into emission. In the strongest decay models, the extra heating is sufficient to render the IGM visible at \( z \approx 10 \) even without luminous sources: this is because the collisional coupling rate \( x_c \) is a sensitive function of

![FIG. 1. IGM histories for long-lived dark matter. In each panel, the curves take \( \xi = 1, 0.1, 10^{-2}, 10^{-3}, \) and 0, from top to bottom. (a): Ionization histories. (b): Thermal histories. Here the thin solid curve shows \( T_\gamma \).](image1)

![FIG. 2. 21 cm signals for long-lived dark matter. Curves take the same parameters as in Fig. 1. (a): Fluctuation amplitude at \( k = 0.04 \) Mpc\(^{-1}\) (note that this scale is arbitrary). (b): Mean (sky-averaged) signal. The thin horizontal dotted line shows \( \delta T_b = 0 \).](image2)
The peak signal still occurs at much higher redshifts. Interestingly, the fluctuations evolve in a nontrivial way. At the highest redshifts, $P_{21}$ is always small. But it is weakest when dark matter decay is strongest, because the increased ionized fraction helps keep $T_K = T_X$. In all of these scenarios, the rms amplitude reaches mK levels by $z \sim 50$. With strong heating, it remains large at lower redshifts even though $\delta T_b$ decreases, because (i) the fluctuations continue growing, (ii) the increase in $T_K$ makes collisional coupling more efficient [31], and (iii) the Ly$\alpha$ background continues to contribute.

The cases with moderately strong heating show the most interesting structure, because in such scenarios the 21 cm signal changes from absorption to emission. Near to (but slightly before) the crossover point (at which $T_K = T_X$), $P_{21}$ also goes to zero. (When $\delta T_b$, overdense regions have $T_K > T_X$ and hence are visible.) This transition point would be a clear signature of strong heating from some exotic process. Unfortunately, in many models the clearest differences occur at $z \lesssim 40$, when confusion with (rare) luminous sources may make it difficult to separate the signal.

Figure 3 illustrates how the different processes shape the curves in Fig. 2 for a fiducial model with $\xi_{-24} = 0.1$ (solid curve). Here the dotted curve assumes the standard recombination history ($\xi_{-24} = 0$). The short-dashed curve ignores the Ly$\alpha$ photons created through collisional excitation. This makes no difference at high redshifts, where collisional coupling is already efficient, but decreases the mean signal by $\sim 25\%$ and the rms fluctuations by $\sim 20\%$ at low redshifts. In this regime the Ly$\alpha$ background helps to maintain contact between $T_X$ and $T_K$ (even though $x_e \propto (1 + z)^{1/2}$). The long-dashed curve assumes $\theta_a = 0$, so that the energy is deposited uniformly. (Spatial fluctuations in the dark matter decay rate obviously have no effect on $\delta T_b$, so we do not show this curve in the bottom panel.) This increases the fluctuation amplitude at higher redshifts by decreasing the heating rate in dense gas so that it absorbs more strongly. For the same reason, the “zero-point” for the fluctuations actually follows that of $\delta T_b$. Of course, once the gas appears in emission, uniform energy injection tends to damp out the fluctuations (because dense regions remain colder and hence less luminous). These kinds of variations will depend on the physics of the decay itself and so offer the possibility of distinguishing, for example, between photon-mediated decay and other types of products (see Sec. III B above).

**B. Dark matter decay: the optically thin regime**

Figure 4 shows the thermal and ionization histories for several scenarios where energy is deposited in the transparency window. We take $\xi_{-24} = 1$ and assume a roughly constant cross section. The dotted, dashed-dotted, short-dashed, and long-dashed curves take $\tau_{100} = 1$, $10^{-1}$, $10^{-2}$, and $10^{-3}$, respectively; note that at any given redshift $\tau$ and $\xi_X$ are essentially degenerate in our simple model: the net energy deposition rates at $z = 100$ in these models pre-
cisely equal the corresponding curves in Fig. 1. It is only the redshift evolution that changes. We note that $\tau_{ov} = 0.047, 0.018, 0.0062,$ and $0.0046$ for these models (including only gas at $z > 10$). Thus the dotted curve substantially affects the CMB (and can already be ruled out), but the others have quite weak effects on it.

For a fixed effective energy deposition rate, the major difference within the transparency window is that a larger fraction of the heating and ionization occurs at relatively high redshifts, with $\tau_i$ and $\bar{T}_K$ increasing more slowly at lower redshifts. As a result the crossover point $\bar{T}_K > T_{\gamma}$ occurs earlier (if it occurs at all, of course), so the features in the 21 cm signal (shown in Fig. 5) occur at significantly lower frequencies. Nevertheless, they have the same general structure compared to the standard calculation, with a reduced fluctuation amplitude at higher redshifts but stronger fluctuations at lower redshift. The overall magnitude of the effect is comparable to the optically thick models in Fig. 2, although here the observable consequences tend to be stronger at higher redshifts and weaker at lower redshifts.

C. Dark matter annihilation

Figs. 6 and 7 show results for several models of dark matter annihilation, with $m_X e^2 = 100$ GeV and $\langle \sigma v \rangle = 1.25, 5.6,$ and $25 \times 10^{-26}$ cm$^3$ s$^{-1}$ (from bottom to top). The uppermost curve corresponds to the strong annihilation case considered by Ref. [8] and is near the upper limit
set by WMAP. As expected from $\xi_X(z) \propto (1 + z)^3$, most of the energy is deposited at high redshifts. In Fig. 6(a), this manifests itself as an elevated ionization fraction at $z \gg 100$, which decreases slowly toward lower redshift because $\xi_X(z)$ becomes almost negligible and recombinations take over. The effect on $T_K$ is much smaller at high redshifts because most of the energy is deposited while Compton scattering is still efficient, which is much stronger than dark matter annihilation. On the other hand, at lower redshifts the gas has cooled sufficiently that even the tiny energy injection rate from annihilations suffices to increase the temperature by up to a factor of a few.

Figure 7 shows that the effects on $\delta T_K$ and the 21 cm fluctuations are also relatively modest. The strong annihilation case decreases both the mean signal and the fluctuation amplitude at high redshifts by a factor of several, because the IGM temperature remains closer to $T_\gamma$. It modestly increases the fluctuations at lower redshifts because of the increased $T_K$; at temperatures $\approx 100$ K, the collisional coupling efficiency is quite sensitive to temperature [31]. Unfortunately, more realistic scenarios with collisional coupling efficiency is quite sensitive to temperature [31]. Unfortunately, more realistic scenarios with weaker ($\sigma v$) have considerably smaller effects that will be difficult to observe given the challenges posed by the observations; they are at best comparable to some of the slower decay models above.

As with dark matter decay, perturbations in the energy injection rate affect $\delta T_\gamma$ and $\delta n_i$ (and hence the 21 cm fluctuation signal). Most of the curves in Fig. 7(a) assume $\theta_a = 1$, or in other words that the energy injection rate is proportional to the local density. But of course the annihilation rate is actually proportional to $n_X^2$; if the products have a short mean-free path and interact with the local IGM, we would have $\theta_a = 2$. This is shown for $\langle \sigma v \rangle = 25 \times 10^{-26}$ cm$^3$ s$^{-1}$ by the dashed-dotted curve. Comparison to the long-dashed curve shows that this scenario further suppresses the high-redshift fluctuations, because it injects even more energy into the dense spots. Their temperatures therefore approach $T_\gamma$. On the other hand, the effects at low redshifts are negligible, because the IGM is approaching isothermality anyway.

In summary, we do not expect the 21 cm background to be as useful for studying dark matter annihilation (as opposed to decay). This is because a large fraction of the energy is injected at high redshifts, when the 21 cm background vanishes. Fortunately, this is also precisely the regime to which the CMB is most sensitive, so together these techniques will offer a useful window into dark sector processes.

V. $H_2$ FORMATION AND IGM CHEMISTRY

Atomic species in gas of primordial composition lack low-lying energy levels. Thus, before the production of metals, molecular hydrogen is a critical coolant vital for star formation to proceed. An increased $H_2$ abundance could also in principle change the temperature of the gas and affect 21 cm observations. Computing the production of $H_2$ in the early universe has a long history stretching back to the first calculation by Saslaw and Zipoy [16], but more recently it has been the subject of intense study as computational resources and estimates of reaction cross sections have improved. Generally, the amount formed in pregalactic gas, $x_{H_2} \sim 10^{-6}$ [17,55–57], is insufficient to be of physical consequence, although the abundance $x_{H_2} \sim 10^{-4}$–$10^{-3}$ formed in the denser and hotter gas of collapsed halos (where reactions can proceed more quickly), is sufficient to trigger star formation [18,19]. If dark matter decay preheats and ionizes the IGM, this will boost $H_2$ formation both due to the increased temperature (since reaction rates are highly temperature dependent) and the increased abundance of free electrons (which are a key catalyst). It would be exceedingly interesting if this boost was sufficient to affect subsequent structure formation, or left an observational signature. Recent estimates differ as to the magnitude of this boost [10,22,23]. Here, we provide a more detailed look at $H_2$ formation and consider some hitherto neglected effects.

There are two main intermediaries by which $H_2$ is produced in the gas phase: $H^-$ and $H_3^+$. In contrast to previous studies, Ref. [57] resolved all 423 rotational-vibrational levels of the $H_2^+$ ion and showed that the $H_2^+$ pathway is greatly suppressed. This is because newly formed $H_2^+$ ions are photodissociated by the CMB before they can decay to the ground state or undergo charge transfer to become $H_2$ molecules. The key formation pathway is therefore through $H^-$, via the reactions: (1) $H + e^- \leftrightarrow H^- + \gamma$, and (2) $H^- + H \leftrightarrow H_2 + e^-$. A net sink of $H^-$ ions is mutual neutralization: (3) $H^- + H^+ \leftrightarrow 2H$. The $H_2$ production rate can then be obtained by assuming that the $H^-$ ion takes its equilibrium value [57]:

$$n_{H_2}^{\text{form}} = \frac{k_1 k_2 x_{H_2} x_{H_2}^2 n^2}{k_3 x_{H_3} + k_{-1} + k_3 [H^+] n}.$$ (25)

Here, $k_1$, $k_2$, $k_3$ refer to the forward reaction rates of reactions (1),(2),(3), $k_{-1}$ is the photodetachment rate of $H^-$ due to both the CMB and nonthermal spectral distortion photons from recombination, and $n$ is the total proper density of hydrogen nuclei. We use the fits for $k_1$, $k_2$, $k_3$ from [55] and calculate $k_{-1}$ as in [57]. In particular, we take care to compute the nonthermal spectral distortion to the CMB from HI two-photon decay and $\text{Ly}\alpha$ photons from cosmological recombination (the two have roughly comparable contributions) by running RECFAST [49] and performing the appropriate integrals [58,59]. We use the parametric fit to the two-photon profile given by [60]. This component generally becomes important at $70 < z < 120$, when CMB photons above the $H^-$ photodetachment threshold energy of 0.74 eV lie in the steeply declining Wien tail, and before a significant fraction of distortion photons redshift below threshold. As we shall see, this is
also the period of peak $H_2$ production, even in cases where the IGM is preheated and ionized. Finally, we also include collisional $H_2$ destruction processes, with reaction rates as given in [61], generally the most significant being the charge exchange reaction $H_2 + H^+ \rightarrow H_2^+ + H$. While normally unimportant, they can become significant when the IGM is heated above $\sim 3000$ K.

Before proceeding to calculate $H_2$ abundances in our energy injection scenarios, we need to consider the importance of other sources of radiation. The energy injected into the IGM from dark matter decay/annihilations introduces an additional nonthermal component to the radiation field, namely, hydrogen Lyman-series photons from atomic excitations (as well as two-phonon decay products from ionized hydrogen atoms). Photodissociation of $H^-$ and $H_2$ could in principle retard $H_2$ formation. In particular: (1) Ly$\alpha$ (and other) photons could photodissociate $H^-$ (which has a 0.74 eV threshold, and peaks at 1.4 eV). However, this effect is strongly subdominant to the CMB, and (at $z < 120$) the spectral distortion from cosmological recombination at $z = 1000$. The latter is easy to see: there must be at least one nonthermal photon per baryon from recombination. On the other hand, if $X_e \sim X_\gamma$, then as a crude first approximation there are $\sim x_e \ll 1$ nonthermal photons per baryon from dark matter decay, much less than the contribution from recombination. (2) Lyman-Werner (LW) photons in the 11.2–13.6 eV band could photodissociate $H_2$ molecules. Unlike $H^-$ photodissociation, this possibility only exists with high-redshift energy injection: the nonthermal distortion from recombination only extends up to 10.2 eV, and rapidly redshifts to lower energies. On the other hand, $Ly\beta$ and higher series photons from atomic excitations can photodissociate $H_2$. We can estimate this effect as follows. About $f_\beta \sim 15\%$ of excitation energy goes into $Ly\beta$ photons (or $X_\beta \sim f_\beta x_e \sim 5\%$ of injected energy), with a much smaller fraction going into higher order transitions (see [52]). Each $Ly\beta$ photon scatters $n_{\text{scat}} \sim 10$ times before being destroyed through a cascade to the 2s level [53]. Thus, $Ly\beta$ photons are produced at a rate $\dot{n}_\beta^+ = x_\beta \frac{m_p c^2 n_{\text{scat}} X_\beta}{(h v_\beta)}$, and destroyed at a rate $\dot{n}_\beta^- = \left(\frac{n_p \sigma_\beta}{n_{\text{scat}}}\right)n_\beta$. Since the destruction timescale is much shorter than the Hubble time, we can consider the $Ly\beta$ photons to assume their time-independent value and obtain $n_\beta$ by setting $\dot{n}_\beta^+ = \dot{n}_\beta^-$. The radiation field in the LW bands is then given by $J_{1L}^\text{W} = \frac{hc}{(4\pi)n_\beta}$, or

$$J_{1L}^\text{W} = \frac{10^{21} m_p c^2 X_\beta X_e}{4\pi} \frac{n_{\text{scat}}}{v_\beta \sigma_\beta} = 2 \times 10^{-5} \xi_{-24} \frac{X_\beta}{0.05} \left(\frac{n_{\text{scat}}}{10}\right) \left(\frac{10^{21}}{\text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1} \text{Hz}^{-1}}\right),$$

where $J_{21} = J/(10^{-21} \text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1} \text{Hz}^{-1})$. Note there is no explicit redshift dependence, apart from possible dependence of $\xi$ on redshift. Since $k_{\text{diss}} = 1.6 \times 10^{-12} J_{21}^\text{W} \text{ s}^{-1}$ [62], this implies a dissociation rate per Hubble time of

$$f_{\text{diss}} = \frac{k_{\text{diss}}}{H(z)} = 2.6 \times 10^{-2} \left(\frac{1 + z}{100}\right)^{-1.5} \xi_{-24} \left(\frac{X_\beta}{0.05}\right) \left(\frac{n_{\text{scat}}}{10}\right).$$

(27)

Thus, at most a few percent of $H_2$ molecules will be photodissociated by the ambient LW radiation field, and we will ignore this effect (of course, star formation seeded by $H_2$ production could produce a much larger UV background with much more significant consequences for $H_2$ chemistry).

Could we observe the nonthermal spectral distortion to the CMB produced by Ly$\alpha$ photons from energy injection at high-redshift? If $x_\gamma \sim x_e$, then the comoving number density of Ly$\alpha$ photons is $n_\gamma \sim x_e n_\beta$. The observed intensity today from photons produced at $z_e$ is then

$$\nu I_\nu \sim \frac{c}{4\pi} n_\gamma \left(\frac{E_\nu}{1 + z_e}\right) \sim 7.8 \times 10^{-13} \left(\frac{x_e}{10^{-2}}\right) \left(\frac{1 + z_e}{100}\right)^{-1} \text{erg s}^{-1} \text{cm}^{-2} \text{sr}^{-1} \text{Hz}^{-1}.$$  

(28)

at wavelengths $\lambda \sim 12[(1 + z_e)/100]$ $\mu$m. By contrast, the observed extragalactic background light at these wavelengths is $\sim 10^{-6}$ erg s$^{-1}$ cm$^{-2}$ sr$^{-1}$ (see Fig. 9 of [63]). Observing the Ly$\alpha$ and two-photon distortions to the CMB from cosmological recombination, which is brighter by a factor $\sim x_e^4[(1 + z_e)/1000] \sim 10$ but suffers from foregrounds of similar strength, is likewise still extremely challenging, although there have been proposals to do so [59,64].

Since these contributions to the radiation field are negligible, $H_2$ production in the case of an IGM heated/ionized by dark matter decays/annihilation can then be obtained by straightforwardly integrating Eq. (25) with the previously calculated temperature and ionization histories. The results are shown in Fig. 8. We see that the boost in $H_2$ production can be substantial in dark matter decay scenarios, with values approaching $x_{H_2} \sim 5 \times 10^{-4}$, which is comparable to the maximal asymptotic abundance of $x_{H_2} \sim 10^{-3}$ in gas cooling via atomic transitions from $T > 10^4$ K (the “maximally heated” case of long-lived dark matter with $\xi_{-24} = 1$ tends to approach this value, but $H_2$ is rapidly destroyed by charge exchange reactions as the IGM temperature climbs past 3000 K). The latter “freeze-out” value can be understood from simple timescale arguments [61]. Note that we have only computed $H_2$ formation at the mean density of the IGM—it will be more efficient and rapid in the higher temperature and denser environments of collapsed halos. Thus, close to the maximal amount of $H_2$ can be formed, certainly comparable to the abundance
H2 suffice to destroy any low enough densities that a small LW background would accompanied by a negligible LW background. Ref. [65] showed formation, reionization by decaying dark matter is accompanied by a negligible LW background. By contrast, in our present scenario the LW background is sufficiently large to seed early star formation is naturally seed a good deal of high-redshift star formation, because of the overall simplicity of the (known) physics at that time—the expanding Universe, hydrogen recombination, and linear gravitational growth—it presents a unique probe of both cosmology [11,12] and exotic processes such as dark matter decay and annihilation. The heating and ionization induced by the decay (or annihilation) products can significantly affect the IGM. These processes can modify the CMB power spectrum [1,5,8], but this is a relatively insensitive measure because it requires a large \( \tilde{\xi} \) for scattering to be significant.

For example, one possible caveat to the claim that early preheating/reionization would seed early star formation is that the entropy of the IGM will suppress gas accretion onto halos and instead exert a negative feedback effect [65]. However, early reionization by decaying dark matter differs in one crucial respect from early reionization by stars, the scenario envisaged by Ref. [65]: unlike early star formation, reionization by decaying dark matter is accompanied by a negligible LW background. Ref. [65] showed that the cores of gas accreted from a high entropy IGM had low enough densities that a small LW background would suffice to destroy any H2 formed—i.e., the photodissociation timescale was much shorter than the H2 cooling time. By contrast, in our present scenario the LW background is negligible and indeed, H2 can form and survive at the mean density of the IGM. Thus, as long as the gas can contract to sufficient density that the cooling time falls below the expansion timescale, the effect of the entropy floor is unimportant. Of course, once some star formation takes place, the LW background rises and H2 destruction in low density cores and the IGM proceeds. Therefore, detailed study is necessary to understand if an early epoch of preheating and copious H2 production will indeed result in extensive star formation in violation of WMAP optical depth bounds.

There are a few general features worth noting about H2 production in these scenarios. Most of the H2 in all scenarios is made at \( z \sim 100 \); this is the highest redshift at which H\(^-\) photodetachment from the high-energy tail of CMB becomes unimportant, yet where the IGM gas is still sufficiently dense that reactions can proceed rapidly. This era of peak H2 production is fairly independent of temperature or ionization history in the different scenarios. The boost in H2 production is primarily due to the increased free electron fraction; for low \( x_e \), we see from Eq. (25) that \( x_{H_2} \propto x_e \). (indeed, the peak H2 abundance can roughly be estimated from \( x_{H_2} \sim k_m n_{eq} x_e t_{cool}\), where \( k_m = k_k k_\perp / k_{-1} \), and \( z_f \sim 100 \) is the redshift at which \( k_m \) peaks). The H2 formed should not have a significant effect on the temperature of the IGM: since the H2 cooling function \( \Lambda_{H_2} \propto T^4 \) for \( T < 3000 \) K, the cooling time is

\[
t_{cool} = 9 \times 10^8 \left( \frac{1 + z}{100} \right)^{-3} \delta \left( \frac{x_{H_2}}{10^{-4}} \right)^{-1} \left( \frac{T}{1000 \text{ K}} \right)^{-3} \text{yr},
\]

(29)

where \( \delta \) is the gas overdensity. This is substantially greater than the Hubble time and the Compton cooling time \( t_C = 1.2 \times 10^9 \left( [1 + z]/100 \right)^{-4} (x_e/10^{-2})^{-1} \) yr. The effects of H2 cooling are only important in dense virialized halos.

Apart from the possible effect of seeding high-redshift star formation, there are few observable consequences of this large amount of early H2 formation. It could potentially increase fluctuations in Ly\( \alpha \) coupling (due to the consumption of LW photons in photodissociation regions), but this is likely difficult to detect.

### VI. DISCUSSION

The 21 cm transition is, at least in principle, a window into the dark ages of structure formation at \( z \approx 50 \). We have argued that, because of the overall simplicity of the (known) physics at that time—the expanding Universe, hydrogen recombination, and linear gravitational growth—it presents a unique probe of both cosmology [11,12] and exotic processes such as dark matter decay and annihilation. The heating and ionization induced by the decay (or annihilation) products can significantly affect the IGM. These processes can modify the CMB power spectrum [1,5,8], but this is a relatively insensitive measure because it requires a large \( \tilde{\xi} \) for scattering to be significant.
Here we have shown that the 21 cm history is a sensitive measure of decay and annihilation during the dark ages (see also [15]), because it directly measures the thermal history of the IGM. In the standard calculation, adiabatic cooling drives \( T_k \) to such low levels that heating the IGM requires much less energy than ionizing it. We have shown that dark matter with lifetimes \( \sim 10^{24} - 10^{27} \) sec can substantially affect the 21 cm background (see Figs. 2 and 5). These timescales are about three to four orders of magnitude longer than those probed by the CMB. The improvement is considerably smaller for annihilation scenarios, because those tend to inject most of their energy nearer the surface of last scattering, when the CMB is more sensitive and the 21 cm background vanishes.

We have made predictions for both the sky-averaged signal—which measures the total energy deposition rate [15]—and the 21 cm power spectrum. Because the temperature, ionization fraction, and Ly\( \alpha \) background all affect the 21 cm signal, the overall amplitude and redshift evolution (at *any* scale) of \( P_{21} \) can provide powerful constraints on such exotic processes: reasonable scenarios produce order unity effects on the power spectrum and also introduce nontrivial redshift dependence if the IGM ever becomes hotter than the CMB. Thus *any* measurement of \( P_{21} \) during the dark ages (such as those advocated by [11]) will be useful in this context, even if it only measures fluctuations on rather large scales.

Of course, because dark sector processes still only show up indirectly through their effects on the IGM, this probe is degenerate with other processes that can heat and ionize the IGM. The most likely worry is structure formation itself: x rays, Ly\( \alpha \) photons, and ionizing photons all affect the 21 cm signal [24]. Fortunately, barring any extremely powerful sources, these effects will not become significant until \( z \leq 25 \) in standard structure formation models [66]. Thus the higher redshifts we have focused on are unlikely to be contaminated. Moreover, there is a fundamental difference between heating by structure formation and by dark matter decay. Any halos that do form at these high-redshifts are far out on the tail of the mass function, so the fraction of material inside of galaxies increases nearly exponentially at \( z \geq 10 \). Thus it is difficult to produce any significant, low-level effect at \( z \geq 20 \) without over-ionizing the IGM shortly afterwards (thus violating existing constraints on the CMB optical depth; see, e.g., [67]).

This is in contrast with dark matter decay and annihilation, which are weak functions of time and so produce effects over long time intervals.

More detailed measurements can begin to constrain the decay and annihilation processes themselves by inferring the properties of the products that interact with the IGM. For example, if the decay produces either soft (\( \lesssim 3 \) keV) or extremely hard (\( \gtrsim 10 \) GeV) photons, the energy will be deposited into the IGM nearly instantaneously [1]. In this case the effects will be most obvious at lower redshifts, when there is more time for heating to take place and Compton coupling with the (spatially uniform) CMB is weaker. On the other hand, if the IGM is optically thin to the products (or if the dark matter annihilates), the consequences are more confined to higher redshifts (where the optical depth is larger because of the increased density). More subtly, the interaction processes between the IGM and the decay products determine the fluctuations in the ionizing and heating rates, which in turn affect \( P_{21} \) (see Figs. 3 and 7). In some cases, this could even introduce extra spatial dependence into the power spectrum, although we have not examined such effects here.

Rather than examining specific particle physics models, we used a generic and flexible formulation for decay and annihilation. We refer the interested reader to the discussions in § III A and in Refs. [8,22] for the thermal and ionization histories in specific models. In general, however, we note that among recently popular models, decaying light dark matter (such as axinos and sterile neutrinos) would have the strongest effects. The decay of heavy dark matter could also affect the thermal history, though this depends strongly on the allowed decay channels. Neutralino annihilation, or the annihilation of light dark matter, could also provide an interesting signal.

We have also shown that the heat deposited in the IGM, as well as the excess ionization, can affect the chemistry of the IGM. In particular, the increased temperature and ionization can dramatically increase the rate of \( \text{H}_2 \) formation. While the resulting abundance can be orders of magnitude higher than the standard value (and in some cases comparable to the maximal asymptotic abundance \( x_{\text{H}_2} \sim 10^{-3} \) for gas phase \( \text{H}_2 \) formation), it is not an important coolant at the low densities of the IGM. Thus its direct observable effects are small. However, the increased \( \text{H}_2 \) abundance will strongly simulate early star formation in dense halos, quite possibly violating WMAP constraints on \( \tau_c \). With better modeling, observations of the first stars and \( \tau_c \) may also provide limits on dark matter decay and annihilation.

We must of course acknowledge the tremendous difficulty posed by 21 cm observations at \( z \approx 50 \), the regime in which dark matter decay signatures would be clearest. The principal challenge is the enormous brightness of the Galactic synchrotron foreground, which has a brightness temperature \( T_{\text{sky}} \gtrsim 10^4 \) K at the relevant frequencies of \( \sim 30 \) MHz. It will make measurements of the smoothly varying \( \delta T_b \) extremely difficult; searches for fluctuations will probably be much easier (though still well beyond current capabilities). The largest transverse wavenumber observable by an array distributed in a circle with radius \( R_{\text{max}} \) is \( k_{\perp \text{max}} \approx 0.2 [50/(1 + z)](R_{\text{max}}/2 \text{ km}) \text{ Mpc}^{-1} \). For a crude sensitivity estimate, we consider an array with uniform baseline coverage observing the spherically-averaged signal. Then the error on the power spectrum at wavenumber \( k \) would be [68].
Moving to larger scales can help slightly, but foreground removal will probably compromise measurements at $k \leq 0.01$ Mpc$^{-1}$ (see Sec. 9.3 of Ref. [24] and references therein). Fortunately, more compact array designs improve the sensitivity over a uniform baseline distribution by factors of a few (see Fig. 6 of Ref. [68]). Exploration of the highly-redshifted 21 cm sky is just beginning, and over the next few years we should learn much more about what is possible. When experiments to open up the dark ages do eventually come along, they will provide important constraints on dark matter decay and annihilation—which can plausibly have order unity effects on the 21 cm signals.

ACKNOWLEDGMENTS

S. R. F. thanks the Tapir group at Caltech for their hospitality while much of this work was completed and M. McQuinn for helpful discussions. E. P. is supported by NSF Grant No. AST-0340648 and is also supported by NASA Grant No. NAG5-11489. S. P. O. gratefully acknowledges NSF Grant No. AST-0407084 and NASA Grant No. NNG06GH95G for support.


