

Deciphering the archaeological record: Cosmological imprints of nonminimal dark sectors

Keith R. Dienes^{1,2,*} Fei Huang^{3,4,†} Jeff Kost^{5,‡} Shufang Su^{1,§} and Brooks Thomas^{6,||}

¹*Department of Physics, University of Arizona, Tucson, Arizona 85721, USA*

²*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*

³*Department of Physics and Astronomy, University of California, Irvine, California 92697, USA*

⁴*CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics,
Chinese Academy of Sciences, Beijing 100190 China*

⁵*Center for Theoretical Physics of the Universe, Institute for Basic Science, Daejeon 34126 Korea*

⁶*Department of Physics, Lafayette College, Easton, Pennsylvania 18042, USA*



(Received 22 January 2020; accepted 28 April 2020; published 15 June 2020)

Many proposals for physics beyond the Standard Model give rise to a dark sector containing many degrees of freedom. In this work, we explore the cosmological implications of the nontrivial dynamics which may arise within such dark sectors, focusing on decay processes which take place entirely among the dark constituents. First, we demonstrate that such decays can leave dramatic imprints on the resulting dark-matter phase-space distribution. In particular, this distribution need not be thermal—it can even be multimodal, exhibiting a nontrivial pattern of peaks and troughs as a function of momentum. We then proceed to show how these features can induce modifications to the matter power spectrum. Finally, we assess the extent to which one can approach the archaeological “inverse” problem of deciphering the properties of an underlying dark sector from the matter power spectrum. Indeed, one of the main results of this paper is a remarkably simple conjectured analytic expression which permits the reconstruction of many of the important features of the dark-matter phase-space distribution directly from the matter power spectrum. Our results therefore provide an interesting toolbox of methods for learning about, and potentially constraining, the features of nonminimal dark sectors and their dynamics in the early universe.

DOI: [10.1103/PhysRevD.101.123511](https://doi.org/10.1103/PhysRevD.101.123511)

I. INTRODUCTION

Dark matter remains one of the great mysteries of modern physics. We know that it exists, and through its gravitational interactions we know how much of it there has been at different epochs during the evolution of the universe. We also know that it is approximately pressureless (with equation of state $w \approx 0$), and that it is relatively cold. However, our knowledge of the dark sector is extremely limited. We do not know how many species of particles comprise the dark sector, nor do we know their masses or spins or whether they are fundamental or composite. We likewise do not know whether the dark matter interacts nongravitationally with the visible sector, nor do we know much about these interactions if they do occur. We do not even know if the dark-sector constituents interact nongravitationally with each other. As a result, the dark sector remains one of the most compelling enigmas facing physics today.

We likewise do not understand the *origins* of the dark sector. We have no idea what production mechanism originally populates this sector, or whether this mechanism is even thermal. Likewise, we do not know what kinds of nontrivial dynamics might be involved in establishing the dark matter that we observe today.

Tackling these problems would not be so urgent if the dark sector were not so important. However, the total energy density of the dark sector is approximately five times that of the visible sector. As a result, it is primarily the dark physics which drives the evolution of the universe through much of cosmological history. Likewise, dark matter seeds structure formation and thereby gives rise to the structure-filled universe that we observe today.

This then leads to two critical questions:

- (i) What imprints might nontrivial early-universe dark-sector dynamics leave in present-day observables?
- (ii) To what extent can we *decipher* this archaeological record, exploiting information about the present-day universe in order to learn about or constrain the properties of the dark sector?

These are clearly very broad questions, and in this paper we shall attempt to take a step towards addressing these

*dienes@email.arizona.edu

†huangf4@uci.edu

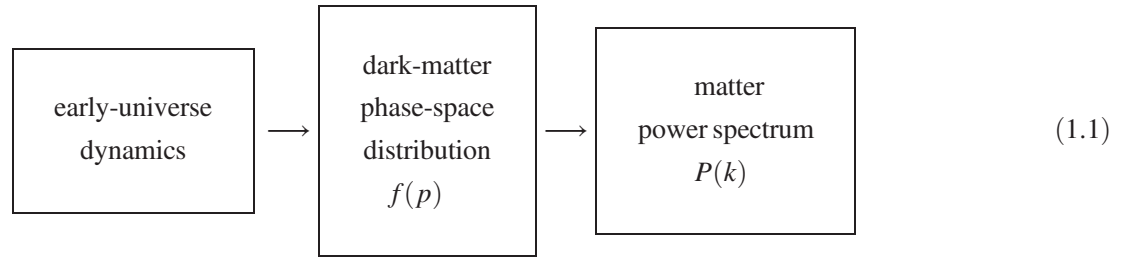
‡jeffkost@ibs.re.kr

§shufang@email.arizona.edu

||thomasbd@lafayette.edu

questions. In particular, we shall concentrate on the linear matter power spectrum $P(k)$, which describes the spatial distribution of matter. As such, this quantity carries an immense amount of information regarding how structure

was formed in the early universe, as sketched below: the matter power spectrum $P(k)$ depends on the net dark-matter phase-space distribution $f(p)$, and this in turn is highly sensitive to the early-universe dynamics we wish to constrain.



Clearly, a given early-universe dynamics leads to a specific $f(p)$ and then to a specific $P(k)$. However, this process is not invertible; in fact, the mappings sketched above may not even be one-to-one. Nevertheless, we can ask: To what extent can we find *signatures* or *patterns* in $f(p)$ and $P(k)$ which might give us at least *partial* information about the early-universe dynamics that produced the dark matter?

Answering this question is the primary goal of this paper. Note that we see this exercise as having two primary motivations beyond those outlined above. First, it is only the matter power spectrum $P(k)$ which is ultimately observable; by contrast, $f(p)$ and the early-universe dynamics which produces it are not observationally accessible. Thus, learning how to approach this “inverse” problem—even in a rough, approximate way—will ultimately become increasingly urgent as further observational data is accumulated. But perhaps even more critically, it is possible that the dark matter interacts with the visible sector only gravitationally. This would be unfortunate, as presumed nongravitational interactions between the dark and visible sectors are the underpinnings of all collider-based, direct-detection, and indirect-detection dark-matter search experiments. Thus, if the dark sector interacts with the visible sector too weakly, it may ultimately only be through studies of quantities such as the matter power spectrum that we will ever learn about the dark sector and its early-universe dynamics.

This paper is organized in two parts. The first part, consisting of Secs. II and III, is primarily concerned with explorations of the two connections sketched in Eq. (1.1) above, with Sec. II devoted to explorations of how we might uncover aspects of the early-universe dynamics by studying the dark-matter phase-space distribution $f(p)$, and Sec. III devoted to explorations of how we might uncover aspects of $f(p)$ given a particular matter power spectrum $P(k)$. As discussed above, our goals are merely to observe and interpret certain patterns and signatures. Although a complete inverse map is not possible, we shall nevertheless demonstrate there are many ways in which we can “invert” certain aspects of the mappings in Eq. (1.1), and indeed in Sec. III we shall conjecture a remarkably

simple closed-form expression which will enable us to reconstruct many of the salient features of the underlying dark-matter phase-space distribution $f(p)$, given a particular matter power spectrum $P(k)$.

By contrast, the second part of this paper, consisting of Sec. IV, presents the detailed analysis of an explicit example model which illustrates all of our main points. In particular, we shall begin with an explicit Lagrangian describing a hypothetical nonminimal dark sector at early times, and we shall then demonstrate that the dynamics implied by this Lagrangian indeed leaves the predicted imprints in $f(p)$ and $P(k)$. As such, this model will permit us to perform a complete “end-to-end” analysis of the connections sketched in Eq. (1.1). We shall also take the opportunity to utilize our conjectured relation from Sec. III in order to test our ability to reconstruct many features of $f(p)$ from $P(k)$, and thereby demonstrate that our conjecture is indeed remarkably accurate for this purpose. Finally, we conclude in Sec. V with a discussion of our main results and possible future research directions.

This paper also contains four appendixes. Appendix A discusses certain aspects of the Boltzmann equations which underlie the physics of this paper, and also presents the specific Boltzmann equations that are used in our analysis of the model in Sec. IV. In Appendix B we then use these Boltzmann equations in order to provide a full numerical example of certain results quoted in Sec. II. By contrast, Appendix C contains a short derivation of the adiabatic sound speed associated with dark matter of a given momentum, as background for a technical point to be discussed in Sec. III. Finally, Appendix D provides details concerning the time evolution of the dark sector in our model in Sec. IV—details which are likely to have relevance for the behavior of nonminimal dark sectors more broadly, even beyond the specific model studied in this paper.

There are, of course, many different theoretical possibilities for early-universe cosmology, each giving rise to different dynamical patterns and different resulting behaviors for $f(p)$ and $P(k)$. For this reason, we stress that it is not our goal in this paper to advocate for a particular model

of the early universe, or even to attempt a complete survey of all logical possibilities. Rather, our goal is to develop tools which can be exploited quite generally—not only to recognize and interpret the observational signatures of various dark sectors, but also to ultimately constrain their properties. Of course, we shall find that the particular dark-matter phase-space distributions $f(p)$ which interest us the most are those which exhibit nonstandard features such as multimodality, with many identifiable peaks and troughs. As we shall demonstrate, such distributions emerge quite naturally from *nonminimal* dark sectors—sectors which transcend the typical WIMP paradigm. As such, many of our results will be particularly useful in such cases. Our results, however, will be completely general, and will hold even for more minimal theories of the early universe.

II. PACKETS TO PACKETS, DUST TO DUST: FROM EARLY-UNIVERSE DYNAMICS TO DARK-MATTER PHASE-SPACE DISTRIBUTIONS

In this section we shall discuss a variety of issues pertaining to the first connection sketched in Eq. (1.1), and the degree to which this connection might be inverted.

A. The cosmological conveyor belt

Once dark matter has been produced in the early universe, its properties can be described through its phase-space distribution $f(\vec{x}, \vec{p}, t)$. The assumption of spatial homogeneity to first order simplifies this quantity to $f(\vec{p}, t)$. Given $f(\vec{p}, t)$, we can calculate the corresponding number density $n(t)$, energy density $\rho(t)$, and pressure $P(t)$ via

$$\begin{aligned} n(t) &\equiv g_{\text{int}} \int \frac{d^3 \vec{p}}{(2\pi)^3} f(\vec{p}, t), \\ \rho(t) &\equiv g_{\text{int}} \int \frac{d^3 \vec{p}}{(2\pi)^3} E(p) f(\vec{p}, t), \\ P(t) &\equiv g_{\text{int}} \int \frac{d^3 \vec{p}}{(2\pi)^3} \frac{p^2}{3E(p)} f(\vec{p}, t), \end{aligned} \quad (2.1)$$

where $E(p) \equiv \sqrt{p^2 + m^2}$ with $p \equiv |\vec{p}|$ and where g_{int} is the number of internal degrees of freedom. We can even calculate the corresponding equation of state $w(t) \equiv P(t)/\rho(t)$. The dark-matter distribution $f(\vec{p}, t)$ is therefore a central quantity in understanding the cosmological properties of the dark sector, telling us whether the dark matter is cold or hot, thermal or nonthermal, and so forth. Indeed, for nonrelativistic (“cold”) dark matter we have $E(p) \sim m \gg p$ and therefore $w \approx 0$, while for relativistic (“hot”) dark matter we have $E(p) \sim p \gg m$ and therefore $w \approx 1/3$. Likewise, if the dark matter is thermal (i.e., produced while in thermal equilibrium with a heat bath), we expect $f(\vec{p}, t)$ to follow a Bose-Einstein or Fermi-Dirac distribution, while all other $f(\vec{p}, t)$ are necessarily non-thermal. Of course, we are assuming here that the dark

matter consists of only a single species. If the dark matter consists of multiple species $i = 0, 1, \dots, N$, the dark sector would be described through a separate phase-space distribution $f_i(\vec{p}, t)$ for each species.

In this section, we shall begin our study by exploring the manner in which early-universe dynamics affects $f(\vec{p}, t)$. This is the first connection sketched in Eq. (1.1). Towards this end, we shall start by studying how the distribution function $f(\vec{p}, t)$ evolves with time.

In the standard Friedmann-Robertson-Walker (FRW) cosmology, the physical distance $x(t)$ between two otherwise stationary points at time t as the universe expands is related to that at time t' by $x(t) = x(t')a(t)/a(t')$, where $a(t)$ is the scale factor. In order to maintain the Poisson bracket $\{x, p\} = 1$, the momentum p of a given free particle therefore has to evolve inversely with respect to the coordinate x , implying $p(t) = p(t')a(t')/a(t)$, or $d \log p / dt = -H(t)$, where $H \equiv \dot{a}/a$ is the Hubble parameter. This describes the gravitational redshift of momentum due to the expansion of the universe. From this we see that time evolution corresponds to additive p -independent shifts in the value of $\log p$, which makes $\log p$ a particularly convenient variable for studying the time evolution of any dark-matter phase-space distribution. Assuming isotropy of the phase-space distribution, quantities such as the physical number density in Eq. (2.1) can therefore be rewritten as

$$\begin{aligned} n(t) &= \frac{g_{\text{int}}}{2\pi^2} \int_0^\infty dp p^2 f(p, t) \\ &= \frac{g_{\text{int}}}{2\pi^2} \int_{-\infty}^\infty d \log p p^3 f(p, t), \end{aligned} \quad (2.2)$$

whereupon we see that the *comoving* number density $N(t) \equiv na^3$ is given as

$$N(t) = \frac{g_{\text{int}}}{2\pi^2} \int_{-\infty}^\infty d \log p p g(p, t) \equiv \frac{g_{\text{int}}}{2\pi^2} \mathcal{N}(t), \quad (2.3)$$

where we have defined

$$g(p, t) \equiv (ap)^3 f(p, t) \quad (2.4)$$

and where $\mathcal{N}(t)$ is the total area under the curve of $g(p, t)$ plotted versus $\log p$.

Particle interactions can of course change the value of the comoving number density $N(t)$. However, barring such interactions, this quantity should remain time-independent. Likewise, we have already seen that $\log p$ merely accrues additive p -independent shifts under time evolution, which implies that the measure $d \log p$ in Eq. (2.3) is also unaffected. Given the relation in Eq. (2.3), this implies that the value of $g(p, t)$ should also be invariant under time evolution, in the sense that

$$g(p(t'), t') = g(p(t), t) \quad (2.5)$$

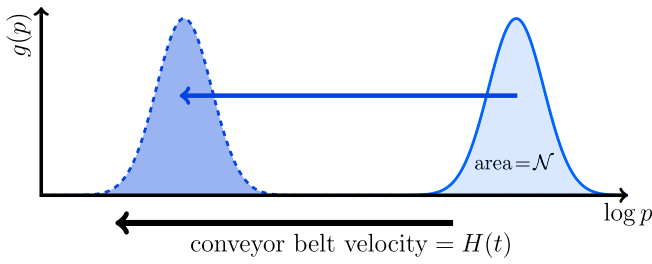


FIG. 1. The cosmological conveyor belt. A given dark-matter phase-space distribution $g(p) \equiv (ap)^3 f(p)$ is rigidly carried towards smaller values of $\log p$ with velocity $H(t)$ as the universe expands. The area \mathcal{N} under the curve is proportional to the corresponding (fixed) comoving number density $N(t) \equiv N$, as given in Eq. (2.3).

for any times t, t' . This implies that under time evolution (and barring any further dark-matter creation or decay), the curve for $g(p)$ merely slides towards smaller values of $\log p$ without distortion, as if carried along on a “conveyor belt” moving with velocity $H(t)$. Indeed, $\mathcal{N}(t)$ —the area under this curve—will also be invariant. For these reasons, we shall often concentrate on $g(p)$ rather than $f(p)$ in this paper when discussing the dark-matter phase-space distribution. This situation is sketched in Fig. 1.

In this paper, we shall refer to this as the *cosmological conveyor belt*. This conveyor belt will ultimately play an important role in our thinking. Of course, we have already seen that it provides a useful way of conceptualizing the

time evolution of the dark-matter phase-space distribution $g(p)$. However, as we shall now discuss, it also allows us to understand how different phase-space distributions $g(p)$ may arise as the result of physical processes that occur during cosmological evolution.

As an example, let us consider the case of a minimal dark sector consisting of a single dark-matter species subject to a specific dark-matter production mechanism. Accordingly, this dark-matter species is produced with a corresponding phase-space distribution $g(p)$ which we may regard as coming into existence on the cosmological conveyor belt at the time of production, with the particular shape of $g(p)$ depending on the details of the production mechanism. This distribution then simply redshifts towards smaller values of $\log p$ for all subsequent times prior to the decay of the dark matter, if any.

However, things can be very different for a nonminimal dark sector containing an entire *ensemble* of particle species instead of a single dark-matter component. In such cases, the phenomenology of the dark sector is not determined by the properties of any individual constituent alone, but is instead determined collectively across all components. However, for such nonminimal dark sectors, it is possible that the dark-matter production may be more complicated, with different “deposits” onto the cosmological conveyor belt occurring at different moments in cosmological history.

For example, let us consider the situation sketched in Fig. 2. In this scenario, an initial dark-matter constituent

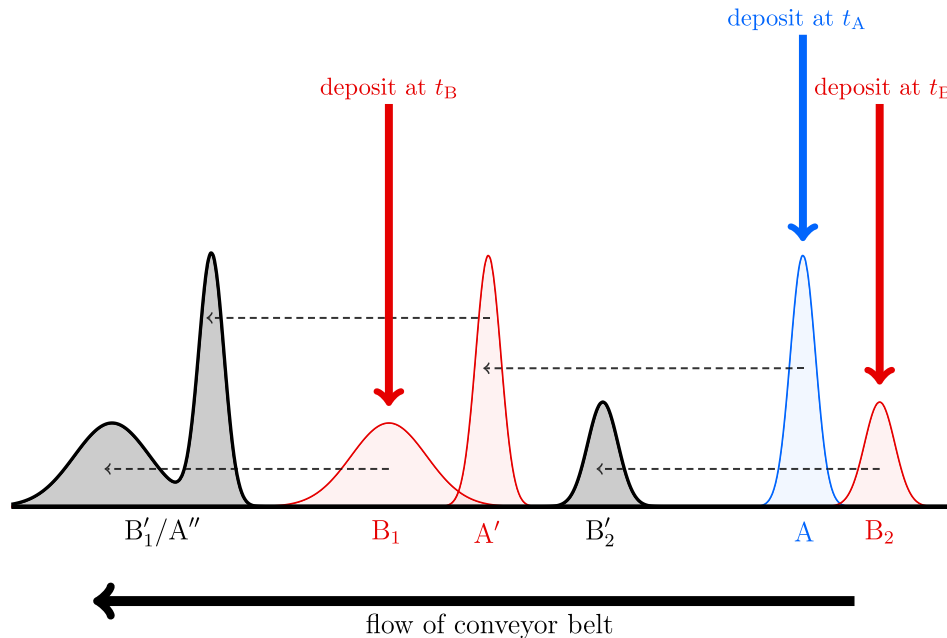


FIG. 2. Sketch of the cosmological conveyor belt at three different times: an early time t_A (blue) at which packet A is deposited, a later time t_B (red) at which packets B_1 and B_2 are deposited, and a final time (black). During the time interval between t_A and t_B , packet A has redshifted to position A' , where it overlaps with the newly deposited packet B_1 . After time t_B , both the combined packet B_1/A' and the packet B_2 continue to redshift into their final positions B_1'/A'' and B_2' . Thus, the resulting dark-matter phase-space distribution consists of two disjoint contributions (black): a warmer unimodal distribution B_2' and a colder multimodal distribution B_1'/A'' . However, in the absence of further information, knowledge of this final distribution allows us to resurrect only certain aspects of this cosmological history.

A is produced at time t_A (blue), with a corresponding phase-space distribution (sketched as packet A in Fig. 2) deposited onto the conveyor belt at that time. However, at a later time t_B (red), two further contributions are produced, corresponding to a colder packet B_1 and a warmer packet B_2 , both of which are also deposited onto the conveyor belt with appropriate momenta. Of course, during the time interval between t_A and t_B , our original packet A has redshifted into a new location A' , where it partially overlaps with B_1 . These then superpose to form a nontrivial combined packet B_1/A' . At all later times beyond t_B (such as that sketched in black in Fig. 2), the resulting phase-space distribution consists of two disjoint contributions: a warmer distribution (labeled B'_2) which is unimodal and a colder distribution (labeled B'_1/A'') which exhibits a complex, bimodal shape. Thus, we see that the final resulting dark-matter distribution can be fairly complicated and even multimodal. However, without further information, it would be possible to resurrect only certain aspects of this cosmological history.

We may quantify this sort of analysis as follows. In general, for any dark sector, the final dark-matter distribution $g(p)$ is realized as the accumulation of all previous deposits occurring at all previous times during cosmological history. Towards this end, let $\Delta(p, t)$ denote the profile of the dark-matter deposit rate at time t . Then, at any time t , we have

$$g(p, t) = \int^t dt' \Delta\left(p \frac{a(t)}{a(t')}, t'\right). \quad (2.6)$$

Of course, if the deposits occur at discrete times t_i , then

$$\Delta(p, t) = \sum_i \Delta_i(p) \delta(t - t_i). \quad (2.7)$$

Inserting this into Eq. (2.6) then yields

$$g(p, t) = \sum_i \Delta_i\left(p \frac{a(t)}{a(t_i)}\right), \quad (2.8)$$

demonstrating that $g(p, t)$ reflects a particular cosmological history. Indeed, the situation sketched in Fig. 2 is nothing but the special case with

$$\Delta(p, t) = \Delta_A \delta(t - t_A) + (\Delta_{B_1} + \Delta_{B_2}) \delta(t - t_B). \quad (2.9)$$

Of course, the important archaeological question is to determine the extent to which knowledge of $g(p, t)$ at a given time t can be used to resurrect the cosmological history. The mathematical answer is provided by Eq. (2.6) or Eq. (2.8): for each value of p , we can constrain only the corresponding sum or integral in these equations such that this sum or integral has the value $g(p, t)$. Note that this is an independent constraint for each value of p . However, for

any given time t , it is clearly impossible to resurrect a unique historical deposit profile $\Delta(p, t')$ for all times $t' < t$ given only the information contained within $g(p, t)$. To see this, let us write Eq. (2.6) in the suggestive form

$$g(p, t) = \int_{-\infty}^{\infty} dt' \int_{-\infty}^{\infty} dp' \Delta(p', t') \delta\left(p' - p \frac{a(t)}{a(t')}\right) \Theta(t - t') \quad (2.10)$$

and treat $\Delta(p', t')$ as the “source” which produces the “response” $g(p, t)$. Indeed, the factors in the integrand of this expression (particularly the argument of the Dirac δ -function in conjunction with the Heaviside Θ -function) make clear that $g(p)$ in some sense tallies the separate contributions along what we may call the “backwards FRW momentum light cone,” in the same way as $\delta(x - ct)$ would have indicated a spatial light cone in flat space. Moreover, given the form of Eq. (2.10), we can identify

$$G(p, p', t, t') \equiv \delta\left(p' - p \frac{a(t)}{a(t')}\right) \Theta(t - t') \quad (2.11)$$

as a Green’s function connecting the source to the response. We can thereby convert Eq. (2.10) into the differential form

$$\frac{\partial g}{\partial t} - Hp \frac{\partial g}{\partial p} = \Delta(p, t), \quad (2.12)$$

where we have recognized $\partial/\partial t - Hp\partial/\partial p$ as the differential operator corresponding to the Green’s function in Eq. (2.11). Indeed, the result in Eq. (2.12) is nothing but the Boltzmann equation describing the evolution of the phase-space distribution $g(p, t)$ in the presence of the source term $\Delta(p, t)$, here functioning as an effective collision term. *This demonstrates that our conveyor-belt/deposit picture is ultimately a physical representation of the integral form of the Boltzmann equation.*

However, the result in Eq. (2.12) also illustrates mathematically why we cannot solve for the complete cosmological history of deposits $\Delta(p, t')$ for all times $t' < t$. We could certainly use this relation to solve for $\Delta(p, t')$ if we had information about $g(p, t')$ at all such times. However, we presumably only have access to $g(p, t)$ at a fixed late time t , long after the final deposits onto the conveyor belt have occurred. This then prevents us from resurrecting the desired profile history $\Delta(p, t')$.

We have already seen in Fig. 2 that multimodality of the phase-space distribution suggests that separate deposits occurred at different moments in cosmological history. However, this immediately begs the question as to whether such a pattern of deposits can arise naturally. In particular, it is important to determine what kinds of nonminimal dark sectors can give rise to such deposit patterns. However, as we shall now demonstrate, such deposit patterns can easily

emerge if our nonminimal dark sector contains an ensemble of states with different masses, lifetimes, and cosmological abundances. In such cases, it is the *intra-ensemble decays* (i.e., decays from heavier to lighter dark-sector components) which will naturally give rise to such deposit patterns.

B. From parents to daughters: Decays and their effects on dark-matter phase-space distributions

As an initial step towards explaining how this occurs, let us first study the fundamental process in which a parent packet decays, producing a daughter packet. As we shall see, this process is surprisingly subtle within an expanding universe because of the relativistic effects which emerge if the momenta involved are large compared to the mass of the parent. We shall therefore proceed to discuss this process in several steps.

In general, a parent packet describes the momentum distribution of a collection of parent particles. Such particles ultimately decay at different times and with different momenta. However, we can group these parent particles into subsets depending on when they decay and the momenta they have when they decay, so that each (p, t) subset consists of the particles which decay at a given common time t with a given common momentum p . The decays of the particles within each (p, t) subset then collectively produce a deposit onto the daughter conveyor belt, and integrating these deposits with those from all other relevant (p, t) subsets (while accounting for appropriate redshifting effects) then yields the final daughter packet $g(p)$. It is in this way that we shall organize our discussion of the decay process through which a daughter packet emerges through the decay of a parent packet. Of course, the idea that we can view our daughter packet as the (p, t) -integral of a deposit profile $\Delta(p, t)$ is already familiar from Eq. (2.6). The only difference here is that we are reorganizing our deposits in terms of the momenta of the decaying parents rather than the momenta of the resulting daughters—a reorganization which proves particularly useful when the dark matter is produced through decays.

We shall therefore begin our discussion of the decay process by studying the properties of the daughter deposit that emerges from a single (p, t) subset of parent particles—i.e., from a population of parent particles which share a common momentum and decay at a common time. Although we shall ultimately refrain from specifying a particular decay process, for simplicity we shall initially assume that each parent particle X undergoes a two-body decay of the form $X \rightarrow YY$ with parent and daughter masses m_X and m_Y respectively, where $m_Y < m_X/2$. In this case, the energy of each daughter in the rest frame of the parent is simply $m_X/2$, and likewise the magnitude of the momentum of each daughter in the rest frame of the parent is given by $\sqrt{(m_X/2)^2 - m_Y^2}$. Indeed, this is the nonzero momentum that is imparted to the daughters in the

cosmological background frame (i.e., the “lab frame”) when the parent is at rest. However, if the parent has momentum p_X in the lab frame at the time of its decay, then the energy of each daughter in the lab frame is given by

$$E_Y = \frac{1}{2} \sqrt{m_X^2 + p_X^2} + \frac{p_X}{2} \sqrt{1 - 4m_Y^2/m_X^2} \cos \theta, \quad (2.13)$$

where θ is the angle between the daughter momentum in the parent rest frame and the parent momentum. Given that the angle θ is unfixed by the kinematics of the decay (implying that all values of $\cos \theta$ arise with equal probability), these daughter energies E_Y will therefore vary uniformly within a total range of magnitude

$$\Delta E_Y = E_Y^{(\max)} - E_Y^{(\min)} = p_X \sqrt{1 - 4m_Y^2/m_X^2}. \quad (2.14)$$

Likewise, for any daughter with energy E_Y within this range, the corresponding daughter momentum in the lab frame is simply given by $p_Y = \sqrt{E_Y^2 - m_Y^2}$. Thus, the existence of a range of possible daughter energies ΔE_Y implies the existence of a corresponding range of daughter momenta Δp_Y , all of which are populated through the decays of the parent particles in the given subset. Indeed, we observe that daughter energy range ΔE_Y —and indeed the corresponding daughter momentum range Δp_Y —both increase as functions of p_X .

In this context, we remark that this kinematic analysis actually allows us to draw an even stronger conclusion if $m_Y \approx 0$: our deposits will be rectangular in shape even in p_Y -space (and not only in E_Y -space), and will actually be logarithmically *centered* around a p_X -independent “anchor” momentum

$$p_Y^{\text{anchor}} \approx E_Y^{\text{anchor}} \equiv \sqrt{E_Y^{(\min)} E_Y^{(\max)}} \approx \frac{m_X}{2}. \quad (2.15)$$

Every deposit, regardless of p_X , will then include this anchor momentum, even though $\langle p_Y \rangle$ and Δp_Y will continue to grow with p_X .

Our main interest, of course, is not in the kinematic details of this specific decay process. Rather, we are interested in certain features which are exemplified above but which are generic across many different decay processes. These include processes such as $X \rightarrow YZ$ where the daughter masses m_Y and m_Z do not have a large hierarchy between them. These also include three-body decays such as $X \rightarrow YZW\dots$ where the number of daughter particles remains $\mathcal{O}(1)$, where the daughter masses do not have large hierarchies between them, and where we assume “typical” decays within the Dalitz plot which are characterized by generic $\mathcal{O}(1)$ angles between the decay products, so that no hierarchies emerge amongst the momenta of the daughters when measured in the rest frame of the decaying parent.

In all such cases, although the precise shapes of these deposits depend on the detailed kinematics associated with the specific decay process, certain general conclusions about the corresponding daughter deposits from each (p, t) parent subset can nevertheless be drawn. For example, because the sum of the masses of the daughters is always less than the mass of the parent, each daughter is produced with a nonzero momentum in the rest frame of the parent. The magnitude of this momentum is uniquely determined for two-body decays, and is also determined to within an order of magnitude for the “typical” three-body decays discussed above. Likewise, because the direction of the daughter momentum in the parent rest frame is uncorrelated with the direction of the momentum of the parent in the lab frame, the magnitude of the momentum of each daughter is broadened by the boost of the parent into a *range* of momentum magnitudes as measured in the lab frame. Thus, the corresponding daughter deposit will stretch across a range of momenta whose width generally grows as a function of the momentum of the parent at the time of decay. Indeed, such considerations are generic, and should hold within the classes of decay processes discussed above.

Many other properties of the possible daughter deposits can be similarly deduced from general considerations such as these. As a result, given the properties of a particular daughter deposit, it is often possible to “reconstruct” certain generic properties of the parent particles within the corresponding (p, t) subset as well as certain generic properties of the corresponding decay process. Our results are shown in Table I, where the relevant properties of possible daughter deposits are listed on the left and the corresponding parent and decay properties are listed on the right.

In Table I, column headings are defined as follows. For the daughter deposit, “rel ($\langle p \rangle$)” at the top of the first column indicates whether the daughter momenta within the deposit—as represented by the average momentum $\langle p \rangle$ —are relativistic. Likewise, the second and third columns indicate whether the width Δp of the daughter deposit is large, order-one, or small (“wide,” “ $\mathcal{O}(1)$,” or “narrow,” respectively) when compared with either the daughter mass m or the mean deposit momentum $\langle p \rangle$. For the parent subset, “rel at decay” indicates whether the parent momentum is relativistic or nonrelativistic at the time of decay, with “rel_~” and “rel_≫” further specifying situations with $p_P \sim m_P$ and $p_P \gg m_P$ respectively, where m_P and p_P are the corresponding parent mass and momentum. By contrast, the final two columns of the table indicate the degree to which the decay process is near marginality. In rough terms, this refers to the degree to which the decay process endows the daughters with additional kinetic energy beyond that which is directly inherited from the parent. More precisely, “absolute marginality” and “relative marginality” are respectively assessed in terms of the fractions p_D^{rest}/m_P and p_D^{rest}/p_P , where p_D^{rest} is the magnitude of the

daughter momentum in the rest frame of the parent. In the case of relative marginality, decays for which p_D^{rest}/p_P is much less than 1, much greater than 1, or $\mathcal{O}(1)$ are respectively considered “near” marginal, “far” from marginal, or $\mathcal{O}(1)$ marginal. By contrast, in the case of absolute marginality, there is a maximum value of p_D^{rest}/m_P which can ever kinematically arise. This maximum value depends on the details of the particular decay process, and corresponds to the limit in which all of the daughter masses vanish. (For example, this maximum value is 1/2 for decays of the form $X \rightarrow YY$.) We then consider the corresponding decay to be respectively “near” absolute marginality, “far” from absolute marginality, or exhibiting $\mathcal{O}(1)$ absolute marginality depending on whether p_D^{rest}/m_P is respectively much less than 1, very close to its maximum value, or somewhere in between.

Note that this table is designed to indicate only those fundamental trends that emerge in limiting hierarchical cases involving only rough orders of magnitude. In other words, in this table we only consider cases in which we can clearly identify our daughter particles as either very nonrelativistic ($p \ll m$), moderately relativistic ($p \sim m$), or ultrarelativistic ($p \gg m$), with an implied hierarchy of momenta between these three situations. Likewise we only consider cases in which our daughter widths Δp are either much less than, or approximately the same size as, or much bigger than m or $\langle p \rangle$, with hierarchically large separations between these different cases. It is because we are only considering such limiting cases with large hierarchies that we do not provide any further information of a more specific nature within this table. Likewise, in constructing this table, we have made certain assumptions which are consistent with this general purpose. For example, as already stated above, these results assume that each decay produces $\mathcal{O}(1)$ daughters whose masses are all of the same overall scale $\mathcal{O}(m_D)$. These results likewise assume the order-of-magnitude estimate that $m_P \sim \mathcal{O}(m_D)$ (respectively, $m_P \gg m_D$) for decays which are near (respectively, far from) absolute marginality.

Despite these assumptions and restrictions, this table contains a wealth of information about the possible daughter deposits that can emerge from the simultaneous decays of parents with a given momentum. As an example, within the limits discussed above, we learn that there is only one way to produce a nonrelativistic daughter deposit for which the width Δp is extremely small compared to both the daughter mass as well as the mean deposit momentum: we must begin with nonrelativistic parents experiencing a decay which is near absolute marginality and yet far from relative marginality (Case A). Indeed, for a two-body decay of the form $X \rightarrow YY$, this would correspond to a situation in which $\frac{1}{2}\sqrt{1 - 4m_Y^2/m_X^2}$ is much smaller than 1 but much larger than p_X/m_X . Conversely, we see that highly *relativistic* daughter deposits with $\Delta p \ll m$ and $\Delta p \ll \langle p \rangle$ can only emerge from extremely nonrelativistic parents experiencing decays which are *far*

TABLE I. The various daughter deposits that can arise from the simultaneous decays of a population of parent particles of fixed momentum. Given the properties of the daughter deposit, we can therefore reconstruct the extent to which the parents were relativistic and the extent to which the corresponding decay process was near marginality. In some cases this reconstruction is unique, while in other cases several possibilities exist. This table nevertheless exhibits a rich pattern of correlations between daughter deposits, parents, and associated decay properties. The quantities at the top of each column of this table are discussed in the text, along with associated definitions and underlying assumptions.

Case	Daughter deposit			Parent	Decay		
	rel? $\langle p \rangle$	width $\Delta p/m$	relative width $\Delta p/\langle p \rangle$		rel at decay?	near absolute marginality?	near relative marginality?
A	$p \ll m$	narrow	narrow	nonrel	near	far	
B			$\mathcal{O}(1)$			$\mathcal{O}(1)$	
C			narrow			far	
D	$p \sim m$	$\mathcal{O}(1)$	narrow	rel $_{\sim}$	near	near	
E			$\mathcal{O}(1)$			$\mathcal{O}(1)$	
F	$p \gg m$	narrow	narrow	nonrel	far	far ($p_{\text{parent}} \ll m_{\text{daughter}}$)	
G		$\mathcal{O}(1)$				far ($p_{\text{parent}} \sim m_{\text{daughter}}$)	
H		wide		$\mathcal{O}(1)$	rel $_{\gg}$	near	near
I					nonrel	far	far ($p_{\text{parent}} \gg m_{\text{daughter}}$)
J	wide	$\mathcal{O}(1)$	rel $_{\sim}$	far	$\mathcal{O}(1)$		
K			rel $_{\gg}$	$\mathcal{O}(1)$ or far	near		

from both absolute and relative marginality, so long as the parent momentum is significantly smaller than the daughter mass (Case F). Similarly, we learn from this table that there are only two ways of producing a highly relativistic daughter deposit for which $\Delta p \sim m$ but $\Delta p \ll \langle p \rangle$: the parents must either be nonrelativistic when experiencing a decay which is far from both absolute and relative marginality (Case G) or highly relativistic when experiencing a decay which is near both absolute and relative marginality (Case H). Indeed, in the former case, we can further conclude that the parent momentum must be of the same order as the daughter mass—otherwise our decay would no longer produce daughter deposits for which $\Delta p \sim m$. The results in this table thus provide a useful guide towards archaeological reconstruction, at least at the level of the individual daughter deposits that emerge from (p, t) parent subsets.

The results in this table are also significantly influenced by the relativistic effects connected with the boosting associated with highly relativistic parents. For example, were it not for such relativistic effects, highly relativistic parents experiencing decays which are far from absolute marginality but close to relative marginality (such as would arise for two-body decays of the form $X \rightarrow YY$ when $p_X \gg m_X/2$ and $m_Y \ll m_X$) would have yielded daughter deposits with $\Delta p \gg m$ but $\Delta p \ll \langle p \rangle$. However, it is the relativistic effects associated with this boosting that actually broaden the width Δp of the resulting daughter deposits and force $\Delta p \sim \langle p \rangle$ (as indicated within Case K). Thus, even though the dark matter may be nonrelativistic today as the result of the gravitational redshifting that has occurred since the

conclusion of the decay process, such relativistic effects may nevertheless be forever imprinted within the deposits that ultimately comprise the daughter packets.

Of course, as we transition across different parent subsets within any physical decay process, the parent and daughter masses are fixed—indeed, only the parent momentum at the time of decay varies. Thus the absolute marginality of the decay process is fixed for all parent subsets in any given decay process, and varying the parent momentum induces transitions between only certain cases within Table I. For example, if the parent and daughter masses correspond to a decay process close to absolute marginality, then we can only transition from Case A to Case B to Case D to Case H as the parent momentum increases. In such cases we find that the corresponding daughter deposits likewise shift from nonrelativistic to relativistic (as expected for a decay process near absolute marginality), but that $\Delta p \ll m$ until the parent momentum becomes highly relativistic. Interestingly, however, we see from Table I that $\Delta p \ll \langle p \rangle$ except near a “resonance” that occurs when p_D^{rest} (the daughter momentum in the rest frame of the parent) becomes approximately equal to the parent momentum. At that point, as illustrated in Case B, we find $\Delta p \approx \langle p \rangle$.

It is also possible to understand certain general aspects of the *shapes* of these deposits. As long as the decay process produces a unique magnitude for the momentum of the daughter in the rest frame of the parent, the boost due to the momentum of the parent will cause the energy of the daughter to have the general form $E_D = E_1 + E_2 \cos \theta$. Indeed, we have already seen an example of this for the

$X \rightarrow YY$ decay process in Eq. (2.13). Isotropy implies that all values of $\cos \theta$ are equally likely to occur, and thus the resulting deposit has a *flat* profile when plotted in E -space. In other words, when plotted in E -space the resulting deposit is “brick”-shaped. We will discuss this further below. However, changing variables to p -space then converts this flat profile into one that rises as a function of p —an effect which is negligible for highly relativistic deposits (for which $E \approx p$) but otherwise sizable. This effect is even more dramatic when the deposit is plotted versus $\log p$.

Having discussed the individual daughter deposits that emerge from each (p, t) subset of parent particles, we now must combine these deposits in order to construct the final, total daughter packet. Indeed, this is the only way in which we can properly discuss the manner in which the decays of *all* of the parent particles within a complete parent packet produce a complete daughter packet. However, in order to combine these deposits correctly, there are a number of additional effects that must be taken into account:

- (i) First, the different parent particles within the parent packet do not all decay at the same time. The decay process is a probabilistic exponential one, with a survival probability scaling with time as $e^{-t/\tau}$ where $\tau = 1/\Gamma$ is the proper decay lifetime and Γ the decay width. Thus we can never determine precisely when a given parent particle will decay.
- (ii) Second, as the result of time-dilation effects, the different parent particles do not even share a common effective lifetime. Indeed, parent particles with larger momenta within the parent packet will have longer effective lifetimes, thereby delaying their decays relative to the decays of those particles with smaller momenta.
- (iii) Finally, we must account for the continual pull of gravitational redshifting which persists throughout this entire process. This affects the parent particles, as their momenta continually redshift until the moment of decay. However, this also affects the daughter deposits, as each deposit also immediately begins to redshift while waiting for subsequent deposits to appear. The final daughter packet is then determined only after each deposit has arrived.

All of these effects combine to render the transition from parent packet to daughter packet a fairly complicated affair.

In order to determine the extent to which the features described in Table I for the individual daughter deposits might eventually survive for the full daughter packets—and also to determine what additional features might accumulate for these daughter packets as the result of the effects itemized above—we shall proceed in several steps. First, it will prove instructive to analyze the decay process while incorporating the second and third features above, but disregarding the first. Then, we shall consider the opposite situation in which we incorporate the first and third features but disregard the second. Each of these

approaches will provide useful, complementary information concerning the final daughter packets that emerge. Finally, we shall analyze the full decays, taking all three features into account simultaneously.

1. Instantaneous-decay approximation with time-dilation and redshifting effects

We begin, therefore, by considering the decay process from parent packet to daughter packet under the so-called instantaneous-decay approximation in which each parent particle is assumed to decay precisely at its lifetime $\tau \equiv 1/\Gamma$, as measured in its rest frame. We will also assume that the parent species has mass m and is produced at some time $t = t_0$ with a simple, unimodal phase-space distribution $g_P(p)$. The resulting decay process is sketched in Fig. 3.

Because the parent packet $g_P(p)$ stretches across a nonzero range of momenta, each momentum slice of the parent packet will experience a different time-dilation factor. This then increases the effective lifetime of each slice prior to decay, causing slices with larger momenta to decay later than those with smaller momenta. However, during the time interval prior to decay, each momentum slice also experiences a cosmological redshift. This redshift decreases the momentum of each slice and thereby partially mitigates the time-dilation effect. Indeed, combining both effects, we see that each momentum slice with initial momentum p_0 at $t = t_0$ decays at

$$t_{\text{decay}}(p_0) = t_0 + \int_{t_0}^{t_0+\tau} dt' \frac{\sqrt{p_0^2(t_0/t')^{4/[3(1+w)]} + m^2}}{m}, \quad (2.16)$$

where we have assumed an FRW cosmology with the Hubble factor scaling as $H \sim 2/[3(1+w)t]$ where w is the equation-of-state parameter. Likewise, the momentum of this slice at the time of its decay is given by

$$p_{\text{decay}}(p_0) = p_0 \left[\frac{t_0}{t_{\text{decay}}(p_0)} \right]^{2/[3(1+w)]}, \quad (2.17)$$

where $t_{\text{decay}}(p_0)$ is given in Eq. (2.16). Indeed, these results are completely general within the instantaneous-decay approximation.

As we move upwards within the parent packet towards momentum slices with increasing values of p_0 , the time-dilation factor increases—but this also provides a longer time interval during which cosmological redshifting occurs. It is therefore important to understand the extent to which this extra redshifting might compensate for the greater original momentum of the momentum slice. Indeed, at first glance it might even seem that momentum slices with greater initial momenta p_0 could ultimately be redshifted to *smaller* final momenta p_{decay} when they actually decay,

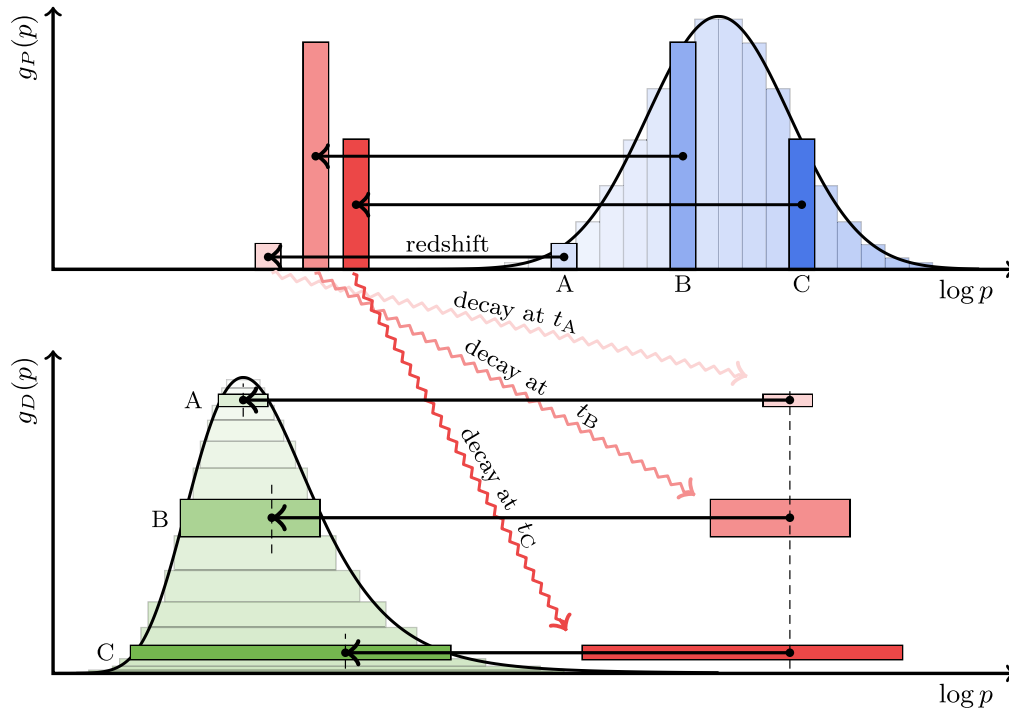


FIG. 3. The process through which a parent packet $g_P(p)$ (blue) decays into a daughter packet $g_D(p)$ (green) in an expanding FRW cosmology, with intermediate steps indicated (red/pink). Each momentum slice of the parent packet experiences a different time-dilation factor which not only extends the lifetimes of the parent particles associated with that slice but also increases the total accumulated cosmological redshift experienced by that momentum slice prior to decay. The sequential decays of these redshifted momentum slices then make sequential contributions to the daughter packet, with each contribution extending over an increasingly broad range of momenta and experiencing its own cosmological redshift immediately upon production until further contributions arrive. This process ends with the decay of the last momentum slice of the parent, thereby completely depleting the parent packet (blue) and yielding the total daughter packet (green). Note that this illustration assumes the instantaneous-decay approximation in which the decay of each parent particle is treated as occurring promptly once its lifetime $\tau = 1/\Gamma$ is reached.

with $p_{\text{decay}}(p'_0) < p_{\text{decay}}(p_0)$ even when $p'_0 > p_0$. However, it is straightforward to verify that $dp_{\text{decay}}(p_0)/dp_0 > 0$ for all p_0 , ensuring that $p_{\text{decay}}(p'_0) > p_{\text{decay}}(p_0)$ for all $p'_0 > p_0$. Thus, momentum slices with greater initial momenta p_0 continue to have greater momenta when they each decay. However, it is also straightforward to verify that $dp_{\text{decay}}(p_0)/dp_0 < 1$ for all p_0 . Thus, any two momentum slices whose original momenta p_0 differ by an amount Δp_0 will have decay momenta p_{decay} differing by an amount $\Delta p_{\text{decay}} < \Delta p_0$. Unless $p_0 \gg p_{\text{decay}}$, this further implies that $\Delta \log p_{\text{decay}} < \Delta \log p_0$. This “momentum compression” is illustrated along the top portion of Fig. 3, where the relative horizontal spacings between the original (blue) momentum slices labeled A, B, and C are larger than the relative horizontal spacings between the corresponding redshifted (red/pink) momentum slices which are sketched at the (different) times of their decays.

The decay of each momentum slice of the parent then yields a contribution to the emerging phase-space distribution $g_D(p)$ of the daughters. Indeed, these contributions are nothing but the deposits already discussed in Table I. These individual deposits contributions are sketched as the

red/pink “bricks” within the lower portion of Fig. 3. As already noted above, the precise shapes of these deposits onto the daughter conveyor belt depend on the detailed kinematics associated with the decay process, and thus our representation of these contributions in Fig. 3 as rectangular bricks whose height does not vary with momentum should at this stage be understood merely as an approximate abstraction adopted for visual simplicity (albeit an abstraction which becomes increasingly accurate for relativistic daughter packets). However, a number of features associated with these bricks are rather important and are incorporated into this sketch. First, we observe that the total area associated with each brick is proportional to the total area of the original decaying momentum slice of the parent, with the proportionality constant signifying the number of daughters (in this case, two) produced through the decay of each parent. Thus, the vertical normalization of the lower portion of this figure need not be assumed to be the same as that of the upper portion. Second, because our original parent packet $g_P(p)$ has a profile which first rises and then falls as a function of momentum, the total areas of the bricks that are deposited first grow and then shrink as

the decay process proceeds. Third, in making this sketch we have assumed that the mass of the parent greatly exceeds the combined masses of the daughters. It is for this reason that we have sketched our bricks as having momenta which exceed p_{decay} . Indeed, as indicated in Table I, these brick momenta will *greatly* exceed p_{decay} if the parent momentum slice is sufficiently nonrelativistic at the time of its decay, and if the decay process itself is relatively far from absolute marginality. In such cases, the bulk of the brick momentum comes from the energy liberated during the decay process itself. Finally, as discussed above, the width of each deposited brick grows as a function of p_{decay} (and therefore grows as a function of p_0). Likewise, if the daughter is sufficiently relativistic as well, each deposited brick will be “anchored” around a common anchor momentum p_{anchor} . These features are also illustrated for the red/pink bricks shown in Fig. 3, with the black vertical dashed line indicating p_{anchor} .

The final stage of the evolution from parent to daughter packet once again involves cosmological redshifting. Because the narrowest bricks are deposited first, they begin redshifting towards smaller momenta before the subsequent, wider bricks have even been deposited. Thus, as shown in Fig. 3, the narrowest bricks experience the largest redshifts—an effect which causes the upper portion of the emerging daughter packet $g_D(p)$ to experience an effective “pull” to the left relative to the lower portion when plotted versus momentum. This effect will be discussed further below.

Ultimately, the final shape of the daughter packet $g_D(t)$ is then given by the sum of these different bricks at the time when the final (bottom) brick is deposited. Indeed, we shall take the deposit of the final daughter brick (or equivalently the decay of the highest parent momentum slice) as marking the completion of the decay process. The daughter packet at this moment is sketched in green in Fig. 3. It is important to realize that because the bottom brick within the green daughter packet has not experienced any redshift at the completion time, it will still be anchored around the dashed black vertical anchor line. In other words, in such cases, the base of the final green daughter packet technically extends all the way to (and well beyond) the anchor line, even if this behavior is not directly visible in Fig. 3 because the bricks below brick *C* have increasingly small heights. Of course, after the decay completion time, the shape of the daughter green packet is fully determined; the packet then redshifts rigidly along the momentum conveyor belt towards smaller momenta.

We have already noted that relativistic time-dilation effects cause the narrowest daughter bricks to experience larger redshifts than those experienced by the wider daughter bricks. This relativistic effect therefore provides a leftward contribution to the overall “tilt” (or skewness) of the daughter packet. In general, the skewness S of a given

$g(p)$ packet can be quantified as the third moment of the $g(p)$ distribution in $\log p$ -space:

$$S \equiv \frac{1}{\sigma^3} \langle (\log p - \langle \log p \rangle)^3 \rangle, \quad (2.18)$$

where

$$\sigma^2 \equiv \langle (\log p - \langle \log p \rangle)^2 \rangle \quad (2.19)$$

and where all averaging is done with respect to the $g(p)$ distribution. Thus, any effect which provides a leftward contribution to the overall tilt of the $g(p)$ packet is one which tends to *increase* the skewness S of the packet. Indeed, this effect is particularly pronounced when the parent particles are highly relativistic when produced. By contrast, when the parent is nonrelativistic at production, there are no significant time-dilation effects within the decay process. Thus the upper portion of the daughter packet no longer experiences extra redshifting relative to the lower portion, and the corresponding extra leftward contribution to the overall tilt of the daughter packet is eliminated.

This is important because there is already a strong tendency towards *rightward* tilting (negative skewness) for $g(p)$ packets simply because such packets are plotted versus $\log p$. Of course, we have already seen that each of our individual deposits has the shape of a flat “brick” when plotted versus the daughter energy E , and thus has no intrinsic skewness. However, the change of variables from E to p already introduces a rightward tilt (one which disappears in the relativistic limit), and the further change from p to $\log p$ introduces an additional rightward tilt. Thus, even a distribution which is completely symmetric in E -space around a mean value $\langle E \rangle$ will nevertheless exhibit a pronounced rightward tilt (or negative skewness) when plotted versus $\log p$. Rightward tilts thus tend to be the “norm” for most daughter $g(p)$ packets. In fact, as shown in Fig. 4 for comparison purposes, even “standard” distributions such as a thermal distribution exhibit a pronounced negative skewness. Because of this strong tendency towards rightward tilting, a parent packet will often have to be exceedingly relativistic before the relativistic leftward-tilt contribution due to time dilation is able to overcome the pre-existing contributions towards rightward tilting and produce a daughter packet exhibiting an overall leftward tilt, with $S > 0$.

Looking over the entire process shown in Fig. 3, we see that each *vertical* momentum slice in the original parent packet $g_P(p)$ ultimately gives rise to a *horizontal* slice in the daughter packet $g_D(p)$. Thus, while the resulting shape of the daughter packet is ultimately sensitive to many kinematic details associated with the decay process, certain general conclusions can be drawn. Each of these can therefore play an “archaeological” role in helping us to

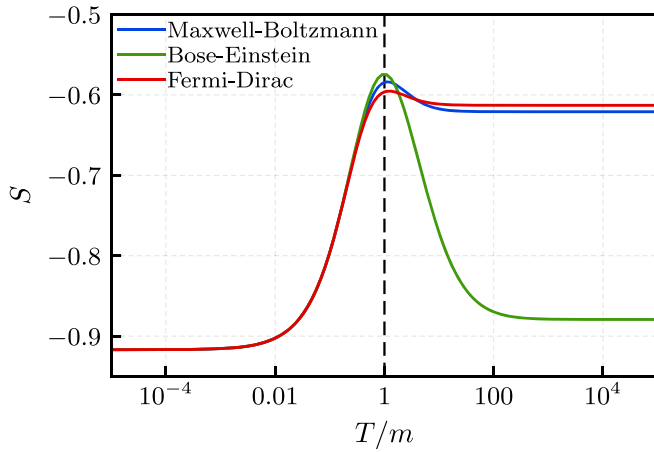


FIG. 4. Many dark-matter phase-space distributions tend to have pronounced rightward tilts (negative skewness S) when plotted versus $\log p$. Here we plot the skewness S of a thermal distribution as a function of the temperature T for a species of mass m obeying Maxwell-Boltzmann, Bose-Einstein, or Fermi-Dirac statistics.

determine the properties of the parent, given only the shape of the daughter.

- (i) First, for example, we see that the *widths* of the daughter packet are correlated with the redshifted momenta of the individual slices of the parent at the times when they decay. Indeed, the maximum width along the *base* of the daughter packet corresponds to the *greatest* momentum within the parent packet, while the narrowest width (which governs the shape of the *peak* of the daughter packet, i.e., its “cuspi-ness”) corresponds to the *smallest* momentum within the parent packet. In technical terms, these two features together characterize the “kurtosis” of the daughter packet.
- (ii) Second, in a similar fashion, the ascending and descending *slopes* of the daughter packet are correlated with the vertical heights of the corresponding deposited bricks. However, the height of each brick is correlated with its *area*, and this area is in turn correlated with the area of the originally decaying parent momentum slice and thus with the corresponding value of $g_p(p)$.
- (iii) Finally, as discussed above, the relative tilt of the daughter packet can provide information about the relativistic nature of the parent at the time when it was initially produced:

$$\begin{aligned} &\text{relativistic parent} \\ &\Leftrightarrow \left\{ \begin{array}{l} \text{leftward contribution} \\ \text{to daughter tilt.} \end{array} \right. \quad (2.20) \end{aligned}$$

Indeed, such relativistic effects tend to increase the skewness of the daughter packet, thereby either

providing the packet with an outright leftward tilt or decreasing the extent to which the packet would otherwise tilt to the right.

Using such pieces of information, many gross features of the daughter packet can be exploited in order to learn about the parent—even independently of the detailed kinematics of the particular decay process. Indeed, one could even proceed beyond skewness and kurtosis to consider higher moments of the daughter phase-space distribution. Of course, the results outlined above have been obtained under the instantaneous-decay approximation. However, we nevertheless expect these features to continue to hold in a rough, average sense even for a full-fledged exponential decay.

Within the instantaneous-decay approximation, it is easy to estimate the conditions under which the relativistic leftward contribution to the tilt of the daughter packet will be significant in size. As discussed above, this contribution arises when the parent is relativistic, and results from the fact that the different momentum slices of the parent packet decay at different times. This in turn implies that the upper portion of the daughter packet experiences a redshift relative to the lower portion, thereby pulling the daughter packet to the left and increasing its skewness. The fact that the parent is relativistic implies that $\langle p \rangle_p / m_p \gg 1$, where $\langle p \rangle_p$ denotes a central momentum value of the parent packet and where m_p is the parent mass. Likewise, we can roughly approximate the time interval between the earliest and latest decays of the parent momentum slices, or equivalently between the deposit times of the narrowest and widest daughter bricks, as $\Delta t \approx (\Delta p)_p / (m_p \Gamma)$, where $(\Delta p)_p$ is the width of the parent packet in momentum space. However, during this interval, the top of the daughter packet redshifts in $(\log p)$ -space by an approximate distance $H \Delta t$, where H represents an average value of the Hubble parameter during this interval. Relative to the rest of the daughter packet, this redshift will induce a significant contribution towards leftward tilt if $H \Delta t \gtrsim \Delta(\log p)_D$ where $\Delta(\log p)_D \approx \log(1 + (\Delta p)_D / \langle p \rangle_D)$ is the width of the daughter packet in $(\log p)$ -space and where $(\Delta p)_D$ and $\langle p \rangle_D$ are respectively the width and mean value of the daughter packet in momentum space. Combining these expressions, we thus find that the resulting leftward contribution to the daughter tilt will be significant if

$$\frac{(\Delta p)_p}{m_p} \gtrsim \frac{\Gamma}{H} \log \left(1 + \frac{(\Delta p)_D}{\langle p \rangle_D} \right). \quad (2.21)$$

For the sake of later comparison, it is also instructive to understand the limit of the decay process sketched in Fig. 3 when the parent packet is *nonrelativistic* at the time it is produced. In such cases, the entire parent packet redshifts a uniform amount before decaying, with the momentum slices maintaining their relative distances along the $\log p$ axis. Likewise, under the instantaneous-decay

approximation, the momentum slices all decay at the same time, simultaneously depositing daughter bricks of exceedingly narrow widths directly on top of each other. (More precisely, their horizontal displacements, if any, are relatively small compared with their total widths.) Thus, within the instantaneous-decay approximation, we expect that nonrelativistic parent packets will ultimately give rise to extremely sharp (δ -function-like) daughter packets. Moreover, we expect that this will remain true regardless of the shape of the parent packet. Of course, at this stage it remains to determine how these two observations fare when we go beyond the instantaneous-decay approximation. This will be discussed below.

2. Exponential decay in the nonrelativistic limit: A universal dark-matter distribution

Thus far, we have analyzed the decay process from parent packet to daughter packet while taking into account the effects of time dilation as well as cosmological redshifting. However, it is also instructive to consider the complementary situation in which we disregard time-dilation effects and instead properly treat the decay of each parent particle as a probabilistic process following an exponential profile $e^{-t/\tau}$ where $\tau \equiv 1/\Gamma$ is the proper lifetime of the species. Indeed, such a treatment describing the decay from parent packet to daughter packet will actually be exact in the limit that the parent packet is nonrelativistic.

Under these assumptions, it proves possible to proceed completely analytically. We assume the existence of a parent packet of total number density $N_P(t_0)$ at some initial time t_0 . Ignoring the time-dilation effects associated with the different particles within the parent packet is tantamount to assuming that the parent particles are extremely nonrelativistic. Thus all of the parent particles have the same effective lifetime, so that the actual *shape* of the parent packet is irrelevant. Given that the parent particles all share the same effective lifetime τ , at any future time t beyond t_0 we have

$$N_P(t) = N_P(t_0) \exp[-\Gamma(t - t_0)] \Theta(t - t_0), \quad (2.22)$$

or equivalently

$$\frac{dN_P(t)}{dt} = -\Gamma N_P(t). \quad (2.23)$$

We have explicitly inserted the Heaviside Θ -function into Eq. (2.22) simply as a way of enforcing our condition that $t \geq t_0$. Although we shall never seek to understand what might have happened at earlier times, this Θ -function will soon play an important role.

For concreteness, let us assume that our parent decays into two identical daughters. (Other cases pose no special difficulties and would proceed analogously.) The daughter number density therefore grows with time as

$$\frac{dN_D(t)}{dt} = +2\Gamma N_P(t). \quad (2.24)$$

Indeed, this growth in the daughter number density is realized as the result of a continuous stream of deposits onto the daughter conveyor belt. In the nonrelativistic limit, these deposits are infinitely narrow and have fixed momentum p_* , where p_* is merely a property of the decay kinematics, depending on the parent and daughter masses alone. Indeed, p_* is nothing but the momentum of the daughters in the rest frame of the parent (which is of course equivalent to the cosmological background lab frame for extremely nonrelativistic parents), and thus serves as the fixed deposit location in momentum-space throughout the decay process. However, although the deposits at any time t have momentum p_* , each deposit redshifts while subsequent deposits continue to arrive. We shall let t_f denote the “final” time at which our decay process has essentially concluded. We then find that any deposit which occurs at time t will redshift to a final momentum given by

$$p(t) = p_* \left[\frac{a(t)}{a(t_f)} \right] = p_* \left(\frac{t}{t_f} \right)^{\kappa/3}, \quad (2.25)$$

where $a(t) \sim t^{\kappa/3}$ is the FRW scale factor, with $\kappa = 3/2$ for a radiation-dominated universe and $\kappa = 2$ for a matter-dominated universe. In writing Eq. (2.25) we are of course assuming that the entire decay process occurs within a time interval during which the value of κ is constant. Note that this final momentum $p(t)$ of the deposit depends on the time t at which this deposit occurs, and is thus a function of t . Using Eq. (2.25), we can therefore change variables from t to $p(t)$, reflecting the fact that the cosmological redshifting has induced the spread in decay *times* to become a spread in final daughter *momenta*. Indeed, from Eq. (2.25) we find that $d \log p(t)/dt = H(t)$, whereupon Eq. (2.24) becomes

$$\begin{aligned} \frac{dN_D(t)}{d \log p} &= \frac{2\Gamma}{H(t)} N_P(t) \\ &= \frac{2\Gamma}{H(t)} N_P(t_0) \exp[-\Gamma(t - t_0)] \Theta(t - t_0). \end{aligned} \quad (2.26)$$

However, from Eq. (2.3) we see that

$$\frac{dN_D(t)}{d \log p} = \frac{g_{\text{int}}}{2\pi^2} g(p) \quad (2.27)$$

where $g(p)$ is the phase-space distribution of the daughter particles. Using the relations $H(t) = H(t_f)(p_*/p)^{3/\kappa}$ and $t/t_f = (p/p_*)^{3/\kappa}$ we therefore obtain our final result for the daughter phase-space distribution at time t_f :

$$g(p) = \frac{4\pi^2 N_P(t_0)}{g_{\text{int}}} \frac{\Gamma}{H(t_f)} \left(\frac{p}{p_*}\right)^{3/\kappa} \times \exp\left\{-\Gamma\left[t_f\left(\frac{p}{p_*}\right)^{3/\kappa} - t_0\right]\right\} \Theta(p - p_{\text{min}}), \quad (2.28)$$

where $p_{\text{min}} \equiv p_*(t_0/t_f)^{\kappa/3}$. Within Eq. (2.28), the Heaviside Θ -function now simply indicates that there exists a minimum momentum p_{min} reached by the daughter packet. This is the momentum to which the very first daughter deposit is redshifted when the decay process concludes. Indeed, there are no daughter particles with $p < p_{\text{min}}$ because no parent particles decay prior to t_0 .

As the above derivation demonstrates, our result in Eq. (2.28) is a *universal* functional form—a universal phase-space distribution to which *any* phase-space distribution resulting from a decay must tend in the limit that the parent is extremely nonrelativistic. Indeed, the only assumption that entered into this derivation is the assumption that the magnitudes of the daughter momenta are all equal to a unique value p_* . This assumption is true for two-body decays and is also largely valid for higher-body decays that are “typical” in the sense described earlier. Consequently, the only features within this functional form that are sensitive to the particular decay process are the value of p_* which enters into Eq. (2.28)—thereby merely serving as a reference momentum—and an overall numerical factor describing the daughter multiplicity. Moreover, in the nonrelativistic limit, this result is wholly independent of the shape of the parent packet. Although this result has appeared previously in a variety of specific contexts (see, e.g., Ref. [1]), our derivation of this result is completely general and proceeds rather simply and directly from our conveyor-belt picture. This universal form for $g(p)$ in Eq. (2.28) is plotted in Fig. 5 for different values of Γ .

One notable feature of this functional form for $g(p)$ is that it is not continuous at $p = p_{\text{min}}$. Indeed, $g(p) = 0$ for all $p < p_{\text{min}}$, while

$$\lim_{(p \rightarrow p_{\text{min}})^+} g(p) = \frac{4\pi^2 N_P(t_0)}{g_{\text{int}}} \frac{\Gamma}{H(t_0)}. \quad (2.29)$$

At first glance, this might seem disturbing, as we are used to phase-space packets which rise smoothly from zero at small momenta and fall smoothly to zero at high momenta. However, we must recall that $g(p)$ is a *distribution function*—specifically, a particle-number $\log p$ -space *density*. Just as with a Dirac δ -function, there is nothing inconsistent about discontinuous densities because the physical quantities of interest (such as the actual numbers of physical particles, or their total energies, pressures, etc.) are *integrals* over these densities, as in Eq. (2.1). Indeed, the discontinuity in $g(p)$ at $p = p_{\text{min}}$ ultimately stems from the discontinuity in the slope dN_P/dt at the initial time t_0 —a feature which is endemic to *all* decay processes which are

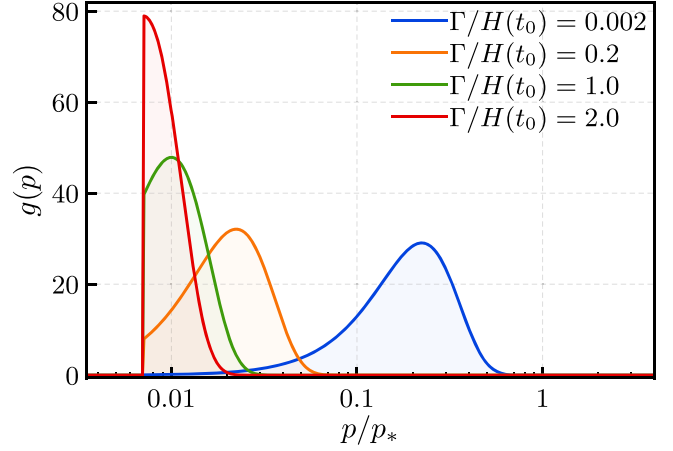


FIG. 5. The universal form $g(p)$ for all decay daughters in the limit that the parent momentum is extremely nonrelativistic. This universal form is given in Eq. (2.28) and plotted at a fixed final time t_f for different values of Γ , with all other quantities held fixed. For this plot we have chosen the reference values $\kappa = 3/2$ and $N_P(t_0) = g_{\text{int}} = 1$. Note, in particular, that we have held t_0 fixed and taken $t_f = 20\Gamma_{\text{min}}^{-1}$, where Γ_{min} is the smallest of the decay widths for which the corresponding $g(p)$ distributions are shown in this figure. Thus t_f represents the time at which the decay process concludes for the case with the smallest Γ . For all other cases in this figure with larger decay widths, we retain this value of t_f : thus the corresponding daughter packets are already fully produced prior to t_f and then redshift rigidly until $t = t_f$. Moreover, in the formal $\Gamma/H(t_0) \gg 1$ limit, the decay timescale is so rapid compared with the redshifting timescale that all decay deposits occur essentially simultaneously at $p = p_*$ and then redshift together to $p = p_{\text{min}}$. Thus, in this limit, we find $g(p) \rightarrow \delta(p - p_{\text{min}})$.

modeled through exponentials as in Eq. (2.22). Despite this discontinuity, the actual *number* of daughter particles $N_D(t)$ rises continuously from zero at $t = t_0$, as required.

It is interesting to consider our expression in Eq. (2.28) in the formal $\Gamma/H(t_0) \rightarrow \infty$ limit. In this limit, the decay timescale is so rapid compared with the redshifting timescale that all decay deposits essentially occur at $t = t_0$ and then redshift together from p_* at $t = t_0$ to p_{min} at $t = t_f$. It then follows that $g(p) \rightarrow \delta(p - p_*)$ as $\Gamma/H(t_0) \rightarrow \infty$. Thus in such cases the Dirac δ -function even emerges as a limiting case for $g(p)$. However, we hasten to point out that numerous other effects enter into the game when $\Gamma/H(t_0) \gg 1$. For example, as $\Gamma/H(t_0)$ grows large, we can no longer ignore the *inverse-decay* process in which daughter particles recombine to form the parent. The presence of both processes tends to thermalize the parent and daughter packets to each other. Likewise, as $\Gamma/H(t_0)$ grows large, the resulting phase-space distribution $f(p)$ no longer satisfies the condition $f(p) \ll 1$ for all momenta. Quantum statistical effects will therefore also become important.

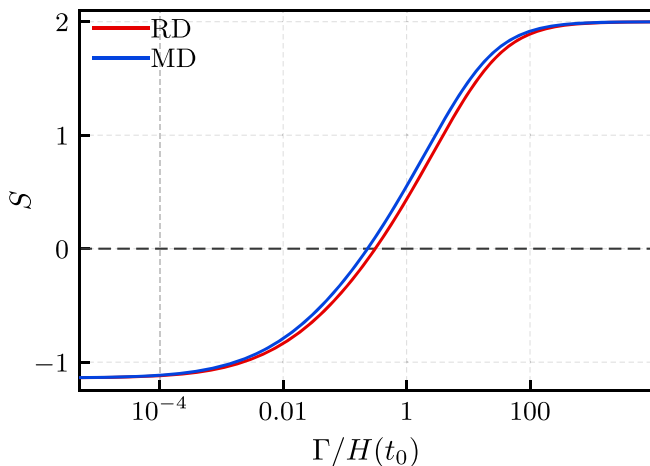


FIG. 6. The skewness S of the universal functional form for $g(p)$ in Eq. (2.28), plotted as a function of $\Gamma/H(t_0)$ for $\kappa = 3/2$ and $\kappa = 2$ [corresponding to radiation-domination (RD) and matter-domination (MD) respectively]. We see that this skewness is negative (signifying rightward tilt) for $\Gamma/H(t_0) \ll 1$, but then rises to positive values (leftward tilt) as $\Gamma/H(t_0) \rightarrow \infty$, ultimately asymptoting at $S = 2$.

Another notable feature associated with this universal asymptotic form for $g(p)$ in Eq. (2.28) is its skewness S . This skewness S is plotted as a function of $\Gamma/H(t_0)$ in Fig. 6. For $\Gamma/H(t_0) \ll 1$, we see that $g(p)$ is highly rightward-tilted, with $S < 0$. However, as $\Gamma/H(t_0)$ increases, we see that S also begins to increase, with $g(p)$ actually becoming leftward-tilted for $\Gamma/H(t_0) \gtrsim 1$. This increasing tendency towards leftward tilting is largely due to the Heaviside cutoff in Eq. (2.28), as this cutoff plays an increasingly dominant role in eliminating the low-momentum tail of the $g(p)$ distribution. Ultimately, for $\Gamma/H(t_0) \gg 1$, we see that $S \rightarrow 2$. Given that this is the limit in which $g(p) \rightarrow \delta(p - p_{\min})$, it may seem strange at first glance that S does not return to zero in this limit (even if this limit is unphysical because of the effects described above). However, strictly speaking, the Dirac δ -function lacks a width and therefore does not have a well-defined skewness; like the longitude of the North Pole, this quantity depends on the precise route through which the limit is approached.

Even though the discontinuity in Eq. (2.29) is a true physical effect—one which becomes increasingly prominent for large $\Gamma/H(t_0)$ —the existence of this discontinuity does signal a certain internal stress within our original assumption of a truly exponential decay beginning promptly at $t = t_0$. As discussed above, this discontinuity ultimately stems from our decision to model our decay process through an exponential as in Eq. (2.22), thereby intrinsically introducing a discontinuity in the slope dN_p/dt at the initial time t_0 . However, the validity of an exponential-decay process implicitly rests upon the assumption that the *parent* came into existence sharply at $t = t_0$. This, of course, cannot be strictly true, since the parent itself must be produced through

some process involving its own internal timescale. There is thus a tension between our exponential-decay assumption and the timescale associated with the production of the parent—a tension which becomes increasingly severe as $\Gamma/H(t_0)$ becomes increasingly large and as the size of the discontinuity in Eq. (2.29) grows. Indeed, for sufficiently large $\Gamma/H(t_0)$, the decays of the parent will begin even before the parent itself is fully produced. Such effects can therefore be expected to soften this discontinuity.

When performing our analyses in the rest of this paper, it will often be convenient to assume the existence of an initial time t_0 at which we may presume the validity of initial conditions involving the prior existence of certain parent particles. In such cases, we shall therefore either incorporate the existence of such discontinuities into our analysis or restrict ourselves to regions of parameter space for which

$$\Gamma/H(t_0) \ll 1, \quad (2.30)$$

where Γ is the decay width of the parent. In such regions of parameter space, the resulting discontinuities in the resulting daughter packet will be relatively small and the daughter packets will have essentially smooth tails at both low and high momenta. However, none of the main results of this paper will depend on this assumption. Thus our results will not depend on the existence or removal of these discontinuities.

3. Full numerical analysis, and the interplay between decay kinematics and cosmological expansion

Finally, we marshal our forces and examine the process of a parent packet decaying to a daughter packet including all relativistic, redshifting, and exponential-decay effects. It is most efficient to analyze this case numerically, time-evolving our system according to the full Boltzmann equations, and in this way we have extracted general results for our final daughter *packets* along the lines of the general results we previously obtained for the individual daughter *deposits* in Table I. For concreteness, in performing this analysis we have restricted our attention to the case of two-body decays with identical daughters, and we have likewise assumed a relatively narrow parent packet, with $\Delta p_p \ll m_p$. Other than these restrictions, we have examined the full range of possibilities in this class, focusing on the fundamental trends that emerge in limiting hierarchical cases involving only rough orders of magnitude, as in Table I. Our results are shown in Table II, where the column headings have the same definitions as in Table I. Note that an explicit numerical example which may help to further elucidate these general results appears in Appendix B.

This table describing the daughter packets exhibits many of the same sorts of correlations that we already saw in Table I for the individual deposits. For example, under the assumptions inherent in this table, we see that there is only

TABLE II. The full daughter phase-space distributions $g(p)$ that can arise from the decays of a relatively narrow parent distribution, assuming a rigorous exponential decay with all relativistic time-dilation and redshifting effects included. For concreteness in this analysis we have assumed a two-body decay with identical daughters. Given the properties of the daughter distribution, we can therefore reconstruct the extent to which the parent population was relativistic and the extent to which the corresponding decay process was near marginality. In some cases this reconstruction is unique, while in other cases several possibilities exist. The quantities at the top of each column of this table are the same as for Table I.

Case	Daughter distribution			Parent	Decay		
	rel? $\langle p \rangle$	width $\Delta p/m$	relative width $\Delta p/\langle p \rangle$		rel at decay?	near absolute marginality?	near relative marginality?
A	$p \ll m$	narrow	$\mathcal{O}(1)$	nonrel	near	far	
B						$\mathcal{O}(1)$	
C						$\mathcal{O}(1)$	
D	$p \sim m$	narrow	narrow	rel $_{\sim}$	near	near	
E						$\mathcal{O}(1)$	
F	$p \gg m$	wide	$\mathcal{O}(1)$	nonrel	far	far ($p_{\text{parent}} \ll m_{\text{daughter}}$)	
G						far ($p_{\text{parent}} \sim m_{\text{daughter}}$)	
H		$\mathcal{O}(1)$	narrow	rel $_{\gg}$	near	near	
I		wide	$\mathcal{O}(1)$	nonrel	far	far ($p_{\text{parent}} \gg m_{\text{daughter}}$)	
J						rel $_{\sim}$	far
K	rel $_{\gg}$					$\mathcal{O}(1)$ or far	near

one way to achieve an ultrarelativistic daughter packet with $\Delta p \sim m$ and $\Delta p \ll \langle p \rangle$: our parent must also have been ultrarelativistic and experienced decays which are near both absolute and relative marginality (Case H). Likewise, we see that there are only two ways of producing a relativistic daughter packet for which $\langle p \rangle \sim \Delta p \sim m$: we can have either a nonrelativistic parent packet experiencing a decay which is far from relative marginality [but exhibiting $\mathcal{O}(1)$ absolute marginality] (Case C), or a relativistic parent packet experiencing $\mathcal{O}(1)$ absolute and relative marginality (Case E). Finally, and perhaps most interestingly, we see that under the assumptions inherent in this table, it is not possible to realize an ultrarelativistic daughter packet with $\Delta p \ll m$.

It is instructive to compare the results of Table II for the final daughter packets with those of Table I for the individual deposits. In Cases B, D, E, H, J, and K, we see that the widths Δp of the final daughter packets are not significantly different from those of the individual daughter deposits. By contrast, in Cases A, C, F, G, and I, we see that our final daughter packets are significantly broader than the individual deposits from which they were constructed. Indeed, comparing Tables I and II, we see that the most significant broadening of all occurs for Case F, in which “narrow” individual daughter deposits with $\Delta p \ll m$ actually combine to produce a “wide” daughter packet with $\Delta p \gg m$.

At first glance, comparing Tables I and II, we see that broadening feature appears to be perfectly correlated with the degree to which the daughter momentum in the rest

frame of the parent exceeds the parent momentum at the time of decay, or equivalently the degree to which the decay process is far from relative marginality: Cases A, C, F, G, and I correspond to cases which are all far from relative marginality, while Cases B, D, E, H, J, and K correspond to cases which either are near relative marginality or experience $\mathcal{O}(1)$ relative marginality. Indeed, as we might expect, Case F is actually the farthest from relative marginality. It is also natural to speculate that for Cases B, D, E, H, J, and K, broadening fails to occur because the daughter deposits all essentially accrue *on top of each other* throughout the decay process. By contrast, for Cases A, C, F, G, and I, it is natural to speculate that each daughter deposit tends to be horizontally displaced relative to the previous one, so that they land *side by side*.

While these expectations are mostly correct, the full story is actually more subtle. It turns out that the manner in which successive decay deposits are stacked throughout the decay process is ultimately a function of not only the relative marginality p_D^{rest}/p_P of the decay process, but also the degree to which the parent is relativistic at the time of its decay. Moreover, the resulting deposit stacking patterns are actually more complex than mere “on top of each other” or “side by side.”

In Table III, we describe the actual stacking patterns that result for a two-body decay of a relatively narrow parent packet of mass m_P into two identical daughters of mass m_D . For ease of visualization, we describe these stacking patterns *within the comoving frame* (i.e., as if we are riding

TABLE III. Stacking dark-matter deposits: the interplay between decay kinematics and cosmological expansion. This table describes how successive dark-matter deposits are stacked in order to build a complete daughter phase-space distribution, as seen in the comoving frame and assuming the two-body decay of a parent of mass m_P with a very narrow phase-space distribution into two identical daughters of mass m_D . Here p_P is the momentum of the parent at the decay time $\tau = 1/\Gamma$. As the decay process unfolds, the manner in which successive decay deposits are stacked is ultimately a function of not only the degree to which the parent is relativistic at the time of its decay but also the relative marginality p_D^{rest}/p_P of the decay process itself. Note that we can equivalently express the relationships between p_D^{rest}/p_P and m_D/m_P along the left-most column as relationships between p_D^{rest}/m_D and p_P/m_P —i.e., as relationships between the degree to which the daughter is relativistic within the rest frame of the parent and the degree to which the parent is relativistic within the background frame. The information in this table explains how the data in Table I ultimately leads to the data in Table II, and outlines the cases in which cosmological expansion has a particularly significant impact on the resulting dark-matter phase-space distribution.

Decay relative marginality p_D^{rest}/p_P	Parent relativistic at decay?		
	$p_P \ll m_P$	$p_P \sim m_P$	$p_P \gg m_P$
$\frac{p_D^{\text{rest}}}{p_P} \ll \frac{m_D}{m_P}$	<i>Weak broadening:</i> Deposits are initially very narrow, with p_{min} decreasing and p_{max} increasing slightly throughout decay process. Resulting daughter packet is narrow with sharp cusp, reflecting small phase space for decay.	<i>Weak broadening:</i> Deposits initially land on top of each other, then transition to behavior in which p_{min} decreases while p_{max} increases. (Case D)	<i>No broadening:</i> Deposits land on top of each other throughout decay process. Resulting daughter packet has sharp left and right edges. (Cases H,K1)
$\frac{p_D^{\text{rest}}}{p_P} \sim \frac{m_D}{m_P}$	<i>Weak broadening:</i> Deposits initially land on top of each other, then transition to landing side by side. (Case B)	<i>Weak broadening:</i> Deposits initially land on top of each other, then transition to landing side by side. (Case E)	<i>Weak broadening:</i> Deposits initially land on top of each other, then transition to behavior in which p_{max} remains constant while p_{min} drops to zero and then begins to grow. Decay ends before p_{min} reaches p_{max} . Resulting daughter packet has sharp right edge. (Case K2)
$\frac{p_D^{\text{rest}}}{p_P} \gg \frac{m_D}{m_P}$	<i>Strong broadening:</i> Deposits land side by side throughout the decay process. (Cases A,C,F,G,I)	<i>Weak broadening:</i> Deposits initially land with p_{min} increasing and p_{max} remaining constant, but then transition to landing side by side. (Case J)	<i>No broadening:</i> Deposits land on top of each other throughout decay process. Resulting daughter packet has sharp left and right edges.

along our cosmological momentum-space conveyor belt). As the decay process unfolds, the momentum of the parent continually redshifts; we shall nevertheless identify a “typical” momentum p_P as the central momentum of the (narrow) parent packet at $t = 1/\Gamma$ where Γ is the decay width. In general, the momentum of each deposit stretches between minimum and maximum values p_{min} and p_{max} , and our stacking patterns are determined by examining how these two quantities independently evolve with time in the comoving frame. For the purposes of Table III, we strictly define deposits “landing on top of each other” as the situation in which both p_{min} and p_{max} remain constant in the comoving frame (implying that the deposits not only stack directly on top of each other but also share a common width). By contrast, we define deposits “landing side by side” as the situation in which both p_{min} and p_{max} increase

with time in the comoving frame, while again holding the deposit width $p_{\text{max}} - p_{\text{min}}$ fixed. For ease of comparison we have also indicated which of the cases within Table II corresponds to which stacking pattern, with K1 and K2 indicating to the two possibilities for Case K within Table II with absolute marginalities which are either $\mathcal{O}(1)$ or far, respectively.

As evident from Table III, several different stacking patterns are possible. These stacking patterns then allow us to assess the degree to which significant broadening might occur when the individual daughter deposits are combined to form the final daughter packet. As anticipated, Cases A, C, F, G, and I all experience “side-by-side” stacking. As such, these cases experience the strongest degree of broadening—a broadening which is sufficiently large that it can be seen in comparing the results of Tables I and II.

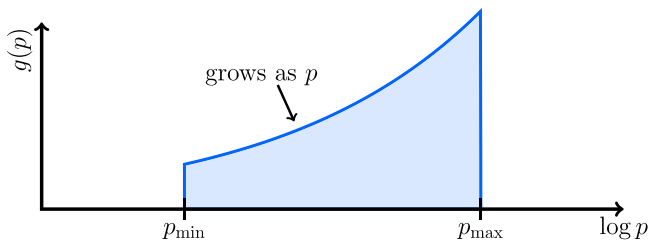


FIG. 7. The universal form for the dark-matter phase-space distribution $g(p)$ corresponding to Cases H and K1, under the assumptions inherent in Table III. This distribution exhibits sharp left and right edges at p_{\min} and p_{\max} , respectively, and exhibits growth scaling linearly with p for $p_{\min} \leq p \leq p_{\max}$.

Likewise, for Cases H and K1 the deposits are stacked “on top of each other.” Thus in these cases there is literally no broadening in transitioning from the width of the individual deposits to the width of the resulting packet. Finally, however, in Cases B, D, E, J, and K2, we see that broadening does occur. However, it is evident from the stacking patterns described in Table III that this broadening is much weaker than that experienced for true side-by-side stacking. This broadening is therefore too small to appear in the comparison between Tables I and II.

Ultimately, the results in Table III describe a subtle interplay between the *kinematics* of the decay process and the *cosmological expansion* of the universe in which the decay is embedded. In some cases, we see that the cosmological expansion has little overall effect on the shape of the resulting daughter packet—indeed, in such cases a packet of similar shape would have emerged even if the universe were not expanding. Moreover, as indicated in Table III, these packets tend to have relatively sharp edges (ultimately stemming from our assumption of a very narrow parent packet). In other cases, by contrast, we see that the cosmological expansion plays a significant role in shaping the resulting daughter packet. In such cases the resulting packets tend not to have such sharp edges, even when the parent packet is extremely narrow.

Finally, when combined with previous observations, the results in Table III even occasionally allow us to predict the full functional form for the resulting daughter packet. For example, under the assumption of a two-body decay into identical daughters, we have already seen that our deposits have flat profiles when plotted as functions of the daughter energy E . In other words, when plotted in E -space, each deposit is “brick”-shaped. However, we now see from Table III that Cases H and K1 have the unique property that all of the deposits land directly on top of each other while sharing the same comoving edge momenta p_{\min} and p_{\max} throughout the decay process. Furthermore, we see from Table II that the daughters for Cases H and K1 are both highly relativistic, which means that a flat profile in E -space is equivalently a flat profile in p -space. Such a profile therefore grows linearly with p when plotted versus $\log p$. This results in the daughter packet shown Fig. 7.

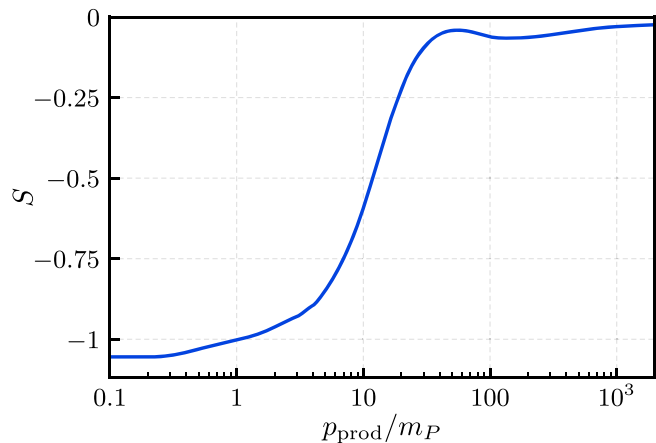


FIG. 8. The skewness S of the daughter packet resulting from the decay of a (skewless) log-Gaussian parent packet, plotted as a function of the central momentum p_{prod} of the parent at production. In all cases the width of the parent packet is held constant, stretching across two orders of magnitude in momentum, and we have assumed a two-body decay with identical daughters for which $m_D/m_P \approx 0.4$. We have also taken $\Gamma/H(t_0) = 10^{-3}$ and $\kappa = 3/2$. As the parent becomes increasingly relativistic, we see that the skewness of the daughter increases dramatically from the universal asymptotic nonrelativistic limit $S \approx -1.06$ (in agreement with the results in Fig. 6) towards $S \approx 0$. This confirms that our expectations from Eq. (2.20) survive a full Boltzmann analysis.

We can also determine the extent to which our previous expectation from Eq. (2.20) concerning the tilt of the daughter packet actually survives a full Boltzmann analysis, with all relativistic, redshifting, and exponential-decay effects included. In Fig. 8 we plot the skewness S of the daughter packet that emerges from the decay of a log-Gaussian parent packet (i.e., a parent packet which has a Gaussian shape when plotted on a logarithmic axis) as a function of the average parent momentum at the time the parent is produced, holding the width of the parent packet fixed and taking the reference value $\Gamma/H(t_0) = 10^{-3}$. We see that this skewness indeed shows a strong dependence on the average momentum of the parent at production, rising from $S \approx -1.06$ in the nonrelativistic limit [in agreement with Fig. 6 for this value of $\Gamma/H(t_0)$] towards $S \approx 0$ as the parent becomes increasingly relativistic at production. Thus we see that our observations regarding skewness survive a full Boltzmann analysis. Indeed, the maximum value of S in the ultra-relativistic limit depends on the skewness of the original parent packet.

These observations concerning skewness provide another tool which can help us determine the properties of the parent from the properties of the daughter. For example, we have already seen in Table II that a relativistic daughter packet which is narrow, with $\Delta p \ll m$ as well as $\Delta p \ll \langle p \rangle$, could in general be the result of either a relativistic parent experiencing a near-marginal decay

(Case E) or a nonrelativistic parent experiencing a decay which is farther from marginality (Case C). The skewness of the daughter packet may therefore be a useful tool in helping to distinguish between these two cases.

It is interesting that the curve in Fig. 8 is nonmonotonic. Indeed, even though the skewness S generally increases as the parent becomes increasingly relativistic at production, we see that there is a region $p_{\text{prod}}/m_P \sim \mathcal{O}(50\text{--}500)$ for the parent momentum at production within which the skewness S experiences a slight dip. Given the parameters chosen for this plot, this is the region in which the average parent momentum p_P at decay is approximately equal to the momentum p_D^{rest} of the daughters in the rest frame of the parent. Indeed, in the language of Tables I and II, this is the region in which our decay process transitions from being *near* relative marginality to being *far* therefrom. When the decay process is near relative marginality, the width Δp of the daughter packet is largely set by p_D^{rest} while its average momentum $\langle p \rangle$ is set by p_P . By contrast, when the decay process is far from relative marginality, the reverse is true. Thus the region with $p_D^{\text{rest}}/p_{\text{decay}} \approx \mathcal{O}(1)$ [which for this figure translates to the region $p_{\text{prod}}/m_P \sim \mathcal{O}(50\text{--}500)$] marks the transition between these two behaviors, with the slight dip in the skewness emerging as a transitional effect.

C. Nonminimal dark sectors: Overlapping decay chains and multimodal dark-matter distributions

Having understood the individual parent-to-daughter decay process, let us now examine how intra-ensemble decays within a nonminimal, multicomponent dark sector can yield a multimodal dark-matter phase-space distribution. To frame our discussion, let us consider the three-state system illustrated in Fig. 9. Here the three states are labeled $\ell = 0, 1, 2$ in order of increasing mass, and we shall assume that only the heaviest state is initially populated. For simplicity, we shall assume that this state corresponds to a single nonrelativistic unimodal packet (sketched in blue in Fig. 9)—indeed, this packet can even be thermal. This packet then redshifts until it decays. For simplicity, let us assume that each decay is a two-body decay, and in this case we shall assume that the $\ell = 2$ state undergoes a decay of the form $2 \rightarrow 1 + 0$ in order to produce two nonrelativistic daughters (green), one with $\ell = 1$ and the other with $\ell = 0$. Each of these new daughter packets then redshifts until the $\ell = 1$ daughter undergoes its own decay into two kinematically identical nonrelativistic $\ell = 0$ granddaughters (orange). As shown in Fig. 9, these granddaughters happen to have a non-negligible overlap with the redshifted $\ell = 0$ daughter (dashed green). They therefore superpose, and the resulting distributions continue to redshift, ultimately producing a final $g(p)$ distribution (black).

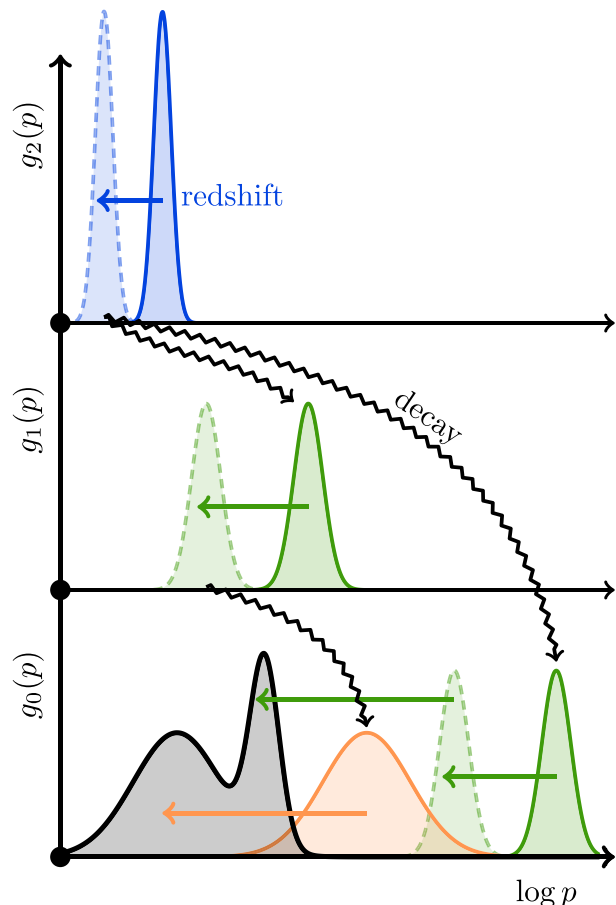


FIG. 9. One possible scenario leading to a nontrivial phase-space distribution $g(p)$ at late times through a sequence of successive “deposits” of the sort sketched in Fig. 2. An excited state is initially created with a simple unimodal (and possibly even thermal) distribution at level $\ell = 2$ (blue) and redshifts towards smaller momenta before undergoing a two-body decay $2 \rightarrow 1 + 0$. Each of these daughter distributions then redshifts until the $\ell = 1$ daughter undergoes its own decay $1 \rightarrow 0 + 0$. The resulting distributions combine and continue to redshift, ultimately producing a final phase-space distribution (black). Thus, even though we began with a simple unimodal distribution, a complex bimodal distribution eventually emerges as the result of a superposition of results from two competing decay chains.

Note that each decay, whether from parent to daughter or from daughter to granddaughter, exhibits the characteristics discussed above. In particular, because all of our packets are assumed nonrelativistic, each decay produces offspring packets whose widths in momentum space are larger than those of the packet from the preceding generation. These offspring packets also have higher momenta than those of the corresponding parent packet, as discussed above. Each of the offspring packets nevertheless has the same total area as the parent packet from which it emerged, since each daughter is descended from a single parent. Indeed, all of the other daughter properties discussed in Sec. II B also continue to apply.

As anticipated, this system furnishes us with an explicit example in which the deposits onto the $\ell = 0$ conveyor belt exactly mirror those of packets A and B_1 in Fig. 2. It is for this reason that both processes produce the same sort of bimodal distribution. Thus, we see from this example that the A/B_1 portion of the deposit pattern in Fig. 2 can indeed emerge naturally—these deposits can simply arise as the end-products of separate decay chains from a single heavy source. (Likewise, the B_2 portion of Fig. 2 can also emerge through an additional independent decay chain.) We therefore conclude that a complex, multimodal dark-matter distribution function can easily be produced, even when our original parent state is unimodal (and potentially even thermal).

We may also study the deposit rate functions $\Delta(p, t)$ arising from such scenarios. To do this, we shall work backwards and begin by focusing on those states whose decays directly produce ground-state daughter particles. In particular, we begin by considering the set of parent packets labeled by ℓ which decay at times t_ℓ and produce daughter packets directly on the $\ell = 0$ dark-matter conveyor belt. Equation (2.7) then becomes

$$\Delta(p, t) = \sum_{\ell} \Delta_{\ell}(p) \delta(t - t_{\ell}), \quad (2.31)$$

where $\Delta_{\ell}(p)$ is the profile of the dark-matter phase-space contribution coming from the decay of parent ℓ . Moreover, we can further write

$$\Delta_{\ell}(p) = \int dp' g_{\ell}(p') P_{\ell \rightarrow 0}(p', p), \quad (2.32)$$

where $g_{\ell}(p')$ is the phase-space distribution of the parent ℓ and where $P_{\ell \rightarrow 0}$ denotes the probability per unit momentum that a parent ℓ with momentum p' produces an $\ell = 0$ daughter with momentum p , normalized by the multiplicity of the decay. Indeed, it is within these probability functions that much of the kinematic information embedded in Table II resides. Substituting Eqs. (2.31) and (2.32) into Eq. (2.6) then yields

$$g(p) = \sum_{\ell} \int dp' g_{\ell}(p') P_{\ell \rightarrow 0}\left(p', p \frac{a(t)}{a(t_{\ell})}\right). \quad (2.33)$$

This result is complete as is. However, we could push this further by recognizing that the parents ℓ are most likely to be descendants of grandparents k , which in turn are likely to be descendants of great-grandparents j , and so forth. We can therefore write the parent distribution $g_{\ell}(p')$ in Eq. (2.33) in terms of deposits from grandparent decays, and recursively iterate this process back to a primordial initial state a . We then obtain

$$\begin{aligned} g(p) &= \sum_{\{a,b,\dots,k,\ell\}} \int dp_{\ell} \int dp_k \dots \int dp_b \int dp_a g_a(p_a) \\ &\times P_{a \rightarrow b}\left(p_a, p_b \frac{a(t_b)}{a(t_a)}\right) P_{b \rightarrow c}\left(p_b, p_c \frac{a(t_c)}{a(t_b)}\right) \dots \\ &\times P_{k \rightarrow \ell}\left(p_k, p_{\ell} \frac{a(t_{\ell})}{a(t_k)}\right) P_{\ell \rightarrow 0}\left(p_{\ell}, p \frac{a(t)}{a(t_{\ell})}\right). \end{aligned} \quad (2.34)$$

Here the summation over $\{a, b, \dots, k, \ell\}$ represents the summation over all possible *decay chains* that start from the primordial state a ; note, in this context, that not all decay chains are of equal length, as this depends on how many intermediate steps are taken during the decay process. Likewise, each t_i ($i = a, b, \dots, \ell$) denotes the time at which the state i decays, where we are implicitly assuming that each state decays instantaneously once its lifetime τ_i is reached. Thus we are implicitly taking $t_i \equiv t_0^{(a)} + \tau_a + \tau_b + \dots + \tau_i$, where $t_0^{(a)}$ is the time at which the primordial ancestor a is produced.

Thus far, we have studied the manner in which a final phase-space distribution $g(p)$ is constructed from the results of individual decay chains, and we have seen that multimodality can emerge naturally through the superposition of the results from separate chains. However, even the analyses we have performed miss certain features. For example, we have been assuming in our discussions until this point that all decays happen instantaneously at $t = 1/\Gamma$, where Γ is the associated decay width. In reality, however, decays occur continuously both before and after $1/\Gamma$. In a similar vein, we also implicitly assumed that each momentum slice of a given parent is created at the same time and therefore feels the same “clock.” However, this also is not necessarily true. For example, we assumed that the $\ell = 1$ state in Fig. 9 is produced, redshifts, and then decays. However, because of the continuous nature of the decay process, some parts of the $\ell = 1$ packet might still remain to be created while other parts of the packet might have already begun to decay. This is a continuous process in which the $\ell = 1$ state is essentially an intermediate resonance.

For accurate results, features such as these must also be taken into account. It is then natural to ask whether effects such as these might “wash out” the features (such as multimodality) that we have been discussing—thereby restoring a traditional packet shape—or whether such features survive.

To study this, we shall perform a full numerical analysis of the Boltzmann equations corresponding to a physical system involving multiple independent decay chains. (The Boltzmann equations for a general multistate system are presented in Appendix A.) In particular, we shall begin with the same setup as in Fig. 9, namely a three-state system which begins with only the heaviest state populated. For concreteness we shall even imagine that this state has a nonrelativistic thermal distribution $g_2 \sim p^3 \exp(-p^2/2m_2T)$ where $T = m_0/20$ is chosen as a

reference value. We shall also take the masses in the ratio $m_2/7 = m_1/3 = m_0$, and let Γ_{ij}^ℓ denote the width for the decay $\ell \rightarrow i + j$. For simplicity, we shall restrict our attention to symmetric decays with $i = j$.

Once the highest level is populated, a decay cascade is initiated and proceeds until the total resulting abundance eventually collects in the ground state. However, the ground-state phase-space distribution that results depends on the specific choice of decay widths Γ_{ij}^ℓ , as each choice corresponds to a different decay pattern. In Fig. 10 we plot the resulting phase-space distributions of the ground state for five different combinations of values for $\{\Gamma_{00}^2, \Gamma_{11}^2, \Gamma_{00}^1\}$, with each value expressed in Fig. 10 as a fraction of the Hubble parameter H at the time when the phase space distribution $g_2(p)$ is initially established. In each case, we show this final distribution at the moment when the intra-ensemble decay processes has concluded (defined as the moment at which this distribution has accumulated 99.5% of its final asymptotic expected abundance).

Each of the cases shown in Fig. 10 has a direct physical interpretation.

- (i) $\{\Gamma_{00}^2, \Gamma_{11}^2, \Gamma_{00}^1\}/H(t_0) = \{10^{-3}, 0, 0\}$ (blue): In this case, the heaviest state decays directly into the ground state. This is therefore effectively a two-component system. Because the rest mass of the daughters is significantly less than the rest mass of the parent, the daughter packets emerge with considerable average momentum and are in fact relativistic.
- (ii) $\{\Gamma_{00}^2, \Gamma_{11}^2, \Gamma_{00}^1\}/H(t_0) = \{0, 10^{-3}, 10^{-5}\}$ (orange): In this case, the direct decay into the ground state is forbidden. Instead, in order to reach the ground state, the decays must go through two steps, $2 \rightarrow 1 + 1$ followed by $1 \rightarrow 0 + 0$. However, as compared with the first step, the second step has a much smaller decay rate. As a result, the intermediate state $\ell = 1$ is able to fully develop and carry significant abundance before it begins decaying to the ground state.
- (iii) $\{\Gamma_{00}^2, \Gamma_{11}^2, \Gamma_{00}^1\}/H(t_0) = \{0, 10^{-3}, 10^{-3}\}$ (purple): This case is similar to the previous case, but the decay rate from $\ell = 1$ to $\ell = 0$ is the same as that from $\ell = 2$ to $\ell = 1$. Therefore, the intermediate state begins decaying *while* is still forming, thereby putting an effective upper limit on the abundance this state can carry.
- (iv) $\{\Gamma_{00}^2, \Gamma_{11}^2, \Gamma_{00}^1\}/H(t_0) = \{0, 10^{-3}, 10^{-2}\}$ (jade): This case is similar to the previous two cases, but with a much faster second step. Therefore, as soon as any portion of the intermediate state begins to form, it almost immediately decays into the ground state. As a result this state never accumulates appreciable abundance.
- (v) $\{\Gamma_{00}^2, \Gamma_{11}^2, \Gamma_{00}^1\}/H(t_0) = \{10^{-3}, 10^{-3}, 10^{-5}\}$ (red): Here the $\ell = 2$ state has an equal probability (branching ratio) to decay into the $\ell = 1$ state or

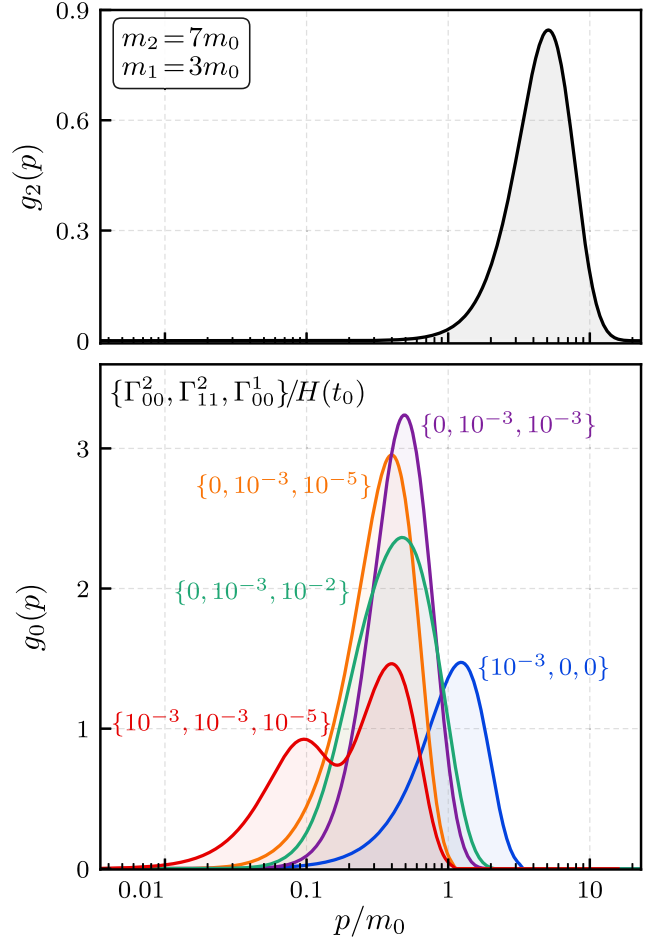


FIG. 10. Decays within a three-state system with mass ratios 7:3:1. We begin with an initial thermal distribution (top panel, black curve) and for each of five different combinations of decay widths $\{\Gamma_{00}^2, \Gamma_{11}^2, \Gamma_{00}^1\}$, we calculate the resulting phase-space distribution of the ground state (bottom panel, colored curves) at the moment when all decays are complete. This final time can therefore be different for each curve. Note that our chosen decay widths are expressed as fractions of the Hubble parameter $H(t_0)$ at the time t_0 when the initial distribution is established. We have also taken $\kappa = 3/2$. Within these plots the vertical axes are normalized such that the parent packet has unit area.

the $\ell = 0$ state, with the former then eventually decaying into the $\ell = 0$ state. Thus, unlike each of the previous cases, this case exhibits two independent decay chains. Indeed, this case represents a superposition of the blue and orange cases listed above, and one may easily verify that the (red) curve for this case is nothing but a weighted combination of the orange and (redshifted) blue curves.

We see, then, that even though the initial phase-space distribution of the parent state is the same for all cases (and is even chosen to be thermal), the resulting phase-space distributions that emerge for the ground state differ rather substantially from each other. For example, they differ in their total areas (comoving number densities),

with the parent distribution and our five different ground-state distributions having total areas in the ratios 1:2:4:4:4:3. This is precisely in accordance with our expectations, given that each decay produces two daughters and that the final case involves two independent decay chains, each occurring with 50% probability. However, these curves also differ in their overall shapes, yielding results which are either unimodal or bimodal, the latter occurring in the single case that exhibits two independent decay chains. Indeed, the case shown in red in Fig. 10 is essentially the same as that sketched in Fig. 9, except that in Fig. 10 so much time elapses between the two deposits onto the ground state that the first deposit has had time to redshift *beyond* the eventual location of the second deposit. The fact that the red curve in Fig. 10 is nothing but a weighted sum of the orange and (redshifted) blue curves ultimately rests on the approximate linearity of the underlying Boltzmann equations, a feature which allows us to treat each decay chain independently and then simply add the results. It is in this final addition process that the multimodality emerges, and thus we see that multimodality is a robust phenomenon, even when a full numerical Boltzmann analysis is performed. Moreover, each “mode” of the resulting distribution will individually exhibit the properties already described in Sec. II B for individual daughters.

Note that the Boltzmann equations are generally nonlinear. However, these equations are approximately linear under certain assumptions. One necessary assumption is that the phase-space distribution $f_i(p, t)$ for each species be relatively small, i.e., $f_i(p, t) \ll 1$, so that any Bose enhancement (or Pauli blocking in the case of a fermionic species) can be safely neglected. However, the Boltzmann equations will also be nonlinear if there exist any significant processes amongst the different species for which the initial state includes two or more species. For example, if our system includes two-body decay processes of the form $\phi_\ell \rightarrow \phi_i \phi_j$, then our system necessarily also includes the inverse-decay processes $\phi_i \phi_j \rightarrow \phi_\ell$ as well as scattering processes of the form $\phi_i \phi_j \rightarrow \phi_\ell \rightarrow \phi_i \phi_j$. Both kinds of processes will render the Boltzmann equations nonlinear, and consequently have the power to weaken or even potentially eliminate the multimodality (or even the non-thermality) that would have otherwise arisen from independent decay chains. However, as we shall see in Sec. IV, it is often the case that the amplitudes for two-body decays can be significant while the amplitudes for the corresponding inverse-decay and scattering processes remain small. In such cases, we expect our multimodality to be robust.

III. NO LONGER VOID AND WITHOUT FORM: FROM DARK-MATTER PHASE-SPACE DISTRIBUTIONS TO MATTER POWER SPECTRA

We now turn to the second connection indicated in Eq. (1.1), namely that between the dark-matter phase-space

distribution $f(p)$ —or equivalently $g(p)$ —and the matter power spectrum $P(k)$. As in Sec. II, our goal is to develop intuition for how different $g(p)$ distributions affect the ultimate shape of $P(k)$. In this way we can potentially begin to address the “inverse” question of reconstructing certain rough characteristics of $g(p)$ given only the information associated with $P(k)$. Indeed, we shall ultimately present a closed-form conjecture which will enable us to “resurrect” many features of $g(p)$, given only the information in $P(k)$. This is important because it is ultimately only through quantities such as $P(k)$ that the dark-matter phase-space distribution $f(p)$ makes contact with observational data.

A. Nonminimal dark sectors and the matter power spectrum

The basic idea underpinning the connection between $f(p)$ and $P(k)$ is that dark matter helps to promote the growth of structure in the early universe. According to the inflationary paradigm, inflation gives rise to a spectrum of primordial curvature perturbations in the early universe. These perturbations serve as seeds for cosmological structure, allowing overdense and underdense regions to develop across a spectrum of physical scales. However, the manner in which this structure evolves is sensitive to the equation of state of the different species present. For example, in a universe composed entirely of cold dark matter (CDM), the primordial perturbations grow linearly with the scale factor, generating significant structure on even the smallest scales. By contrast, if some of the dark matter is not purely cold (i.e., if the dark matter has some nonzero pressure due a nonzero momentum), then dark-matter particles with a sufficient speed may escape the developing gravitational wells, thereby smoothing out the structure over the corresponding scales. As a result, one finds that only perturbations of a certain minimum size are able to remain stable and/or grow with time.

All of this information is encoded within the matter power spectrum $P(k)$. To define this quantity, we consider the spatial perturbations of the energy density $\rho(\vec{x}, t)$ relative to the zeroth order, unperturbed, spatially homogeneous energy density $\bar{\rho}(t)$:

$$\delta(\vec{x}, t) \equiv \frac{\delta\rho(\vec{x}, t)}{\bar{\rho}(t)} = \frac{\rho(\vec{x}, t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \quad (3.1)$$

where

$$\rho(\vec{x}, t) \equiv \frac{1}{(2\pi)^3} \sum_i g_i \int d^3p E_i(p) f_i(\vec{x}, \vec{p}, t) \quad (3.2)$$

and where

$$\bar{\rho}(t) \equiv \frac{1}{V} \int_V d^3x \rho(\vec{x}, t) \quad (3.3)$$

for a suitably large fiducial volume V . Note that in Eq. (3.2) we are explicitly summing over all species that contribute

to the total energy density of the universe, including the dark matter. Our main interest is then in the two-point correlation function $\xi(\vec{x}, t) \equiv \langle \delta(\vec{x} + \vec{y}, t) \delta(\vec{y}, t) \rangle$, which, under the assumptions of homogeneity and isotropy, can be Fourier-decomposed in the form

$$\xi(\vec{x}, t) \equiv \frac{1}{(2\pi)^3} \int d^3k P(k, t) e^{-i\vec{k}\cdot\vec{x}}, \quad (3.4)$$

where $k \equiv |\vec{k}|$ and where $P(k, t)$ is the time-dependent matter power spectrum. Since $\xi(\vec{x}, t)$ depends only on $r \equiv |\vec{x}|$, we can invert Eq. (3.4) and perform the angular integration to obtain

$$P(k, t) = 4\pi \int_0^\infty dr r^2 \frac{\sin kr}{kr} \xi(r, t). \quad (3.5)$$

This quantity therefore describes the degree to which the two-point correlation function of energy-density fluctuations shows power at the scale k , and for convenience we shall henceforth define $P(k) \equiv P(k, t_{\text{now}})$. Moreover, in this work we shall mostly be interested in the so-called *transfer function* $T(k)$, defined as

$$T(k) \equiv \sqrt{\frac{P(k)}{P_{\text{CDM}}(k)}}, \quad (3.6)$$

where $P_{\text{CDM}}(k)$ is the matter power spectrum that would emerge in the CDM limit in which all of the dark matter is assumed cold. Thus, for any $f(p)$, the transfer function $T(k)$ indicates the degree to which the corresponding matter power spectrum $P(k)$ deviates from our CDM-based expectations as a function of the scale k . Conveniently, $T(k)$ is independent of the particular choice of normalizations for $P(k)$, many of which exist in the literature.

As a dominant contributor to this process, dark matter plays a leading role in structure formation. In order to determine which scales k within the matter power spectrum $P(k)$ might be affected by dark matter of a given momentum p , we can perform a straightforward horizon calculation. Given that dark matter of momentum p has a velocity $v = p/E = p/\sqrt{p^2 + m^2}$, the corresponding horizon size d_{hor} in a static flat universe would simply be vt , whereupon we could identify a corresponding wave number $k_{\text{hor}} \sim 1/d_{\text{hor}}$. However, in an expanding universe, the momentum is continually redshifting towards lower values. Thus we must actually integrate over the cosmological history between the time of dark-matter production and today. This then yields the result

$$\begin{aligned} k_{\text{hor}}(p) &\equiv \xi \left[\int_{t_{\text{prod}}}^{t_{\text{now}}} \frac{p/a(t)}{\sqrt{p^2/a(t)^2 + m^2} a(t)} dt \right]^{-1} \\ &= \xi \left[\int_{a_{\text{prod}}}^1 \frac{da}{Ha^2} \frac{p}{\sqrt{p^2 + m^2 a^2}} \right]^{-1}, \end{aligned} \quad (3.7)$$

where the quantity p in this expression signifies the dark-matter momentum *today*. We emphasize that this expression for k_{hor} is merely an order-of-magnitude estimate, one which holds only up to an multiplicative factor of order $\mathcal{O}(1)$. Indeed, while the horizon length within the square brackets in Eq. (3.7) is a precisely defined quantity, there are many different ways of extracting a corresponding wave number, depending on the particular conventions adopted and on the particular system under study. The quantity ξ in our definition of $k_{\text{hor}}(p)$ in Eq. (3.7) is therefore inserted in order to represent this $\mathcal{O}(1)$ factor. Note that the integral in Eq. (3.7) generally receives contributions from the radiation-dominated era as well as later matter-dominated era, and these scale differently with time. For this reason it is convenient to treat these separately, leading to the approximate total result

$$k_{\text{hor}}(p) \approx \xi \left[\frac{p}{m(a^2 H)_{\text{MRE}}} \left\{ 2 + \tanh^{-1} \left(\frac{m}{\sqrt{m^2 + p_{\text{MRE}}^2}} \right) - \tanh^{-1} \left(\frac{m}{\sqrt{m^2 + p_{\text{prod}}^2}} \right) \right\} \right]^{-1}, \quad (3.8)$$

where ‘‘MRE’’ indicates the time of matter-radiation equality and where p_X indicates the value of the momentum redshifted back to time t_X .

Any dark matter with momentum p today has the potential to suppress structure at all scales $k \geq k_{\text{hor}}(p)$. *In other words, for any p , the quantity $k_{\text{hor}}(p)$ defines the minimum value of k for which the matter power spectrum $P(k)$ could be affected.* However, our interest is not merely in situations involving a single momentum slice—we are interested in understanding situations involving an entire nontrivial distribution $f(p)$. It is here that subtleties arise.

One standard approach that is commonly employed in the literature is to average over the different momenta within the distribution, essentially defining a free-streaming horizon scale k_{FSH} in terms of an average packet velocity $\langle v(t) \rangle$ via a relation of the form

$$k_{\text{FSH}} \sim \left[\int_{t_{\text{prod}}}^{t_{\text{now}}} \frac{dt}{a(t)} \langle v(t) \rangle \right]^{-1}. \quad (3.9)$$

However, while such an analysis is robust for most thermal or unimodal momentum distributions in which averaged quantities faithfully represent the full distribution, this is not generally the case for nonthermal and/or multimodal distributions [2]. Indeed, there are implicit assumptions buried within the definition in Eq. (3.9) that may cause it to fail for more general momentum distributions of the sort we have discussed in Sec. II. For example, if we have a bimodal $f(p)$, it is not necessarily even the case that the average velocity $\langle v(t) \rangle$ represents the speed of *any* particle that populates the distribution.

In order to understand these complications at a more fundamental level, one can consider the Jeans equation which governs the time evolution of energy-density perturbations on different length scales. For dark matter with a given sound speed c_s , the form of this equation typically allows one to read off a characteristic Jeans wave number k_{Jeans} which separates gravitationally stable (i.e., growing) and unstable (i.e., nongrowing) modes. This wave number thereby signals the onset of suppression for the matter power spectrum $P(k)$. Essentially k_{Jeans} indicates the critical perturbation size at which the tendency towards gravitational collapse is precisely balanced against the increased dark-matter pressure resisting collapse.

However, this analysis assumes that the pressure perturbations propagate through the dark matter with a single well-defined sound speed c_s , much as we would expect for dark matter of a single species. For dark-matter particles of mass m with a single momentum p , a well-defined sound speed c_s always exists. However, it is easy to see that c_s depends on p . For example, we know that $c_s = 0$ for $p = 0$ (signifying the case of infinitely cold dark matter), while $c_s = 1/\sqrt{3}$ for $p \gg m$ (signifying the case of highly relativistic matter). Indeed, an explicit expression for c_s as a function of p is derived in Appendix C. Thus, dark matter with a nontrivial distribution $f(p)$ does not have a single well-defined sound speed. From the perspective of the Jeans equation, such dark matter therefore behaves as if it were composed of different “species” with different sound speeds, one for each momentum slice.

If the dark-matter phase-space distribution $f(p)$ is highly peaked around a central averaged momentum $\langle p \rangle$, it may of course be possible to work with a single averaged sound speed $\langle c_s \rangle$ and thereby extract a single Jeans scale, in the same spirit as Eq. (3.9). However, for more complicated distributions, this need no longer be the case. In general, the presence of multiple “species” is problematic in the presence of gravity because gravity feels all species simultaneously, thereby causing the evolution of the density perturbations for each species to be affected by the density perturbations for all of the other species. In other words, the existence of dark matter across a nontrivial momentum distribution $f(p)$ leads to important cross-correlations that eliminate the existence of critical Jeans scales that cleanly separate growing from nongrowing (suppressed) perturbation modes. This is true even if the dark matter exhibits only two distinct momenta.

In addition to these difficulties, we are not merely interested in the range of k -values for which the matter power spectrum $P(k)$ might be affected—we also wish to evaluate $P(k)$ itself. Though relatively straightforward, the calculation of $P(k)$ is also highly nontrivial. One begins with certain initial conditions for the primordial perturbations arising from inflation. One then propagates these forward in time through the linearized Einstein equations, assuming a dark-matter component with a given phase-

space distribution $f(p, t)$. This also requires taking into account the evolution of the background cosmology. Evolving to the present time and explicitly calculating the two-point correlation function, one thereby obtains the matter power spectrum $P(k)$ corresponding to a given $f(p, t)$ at $t = t_{\text{now}}$.

In this work, we shall perform these calculations numerically using the CLASS software package [3–6]. However, for all the reasons discussed above, we do not expect any simple relationship to exist between a given $f(p)$ distribution and the corresponding matter power spectrum $P(k)$. That said, we nevertheless wish to develop a rough, phenomenological way of understanding how the phase-space distribution $f(p)$ affects the matter power spectrum $P(k)$. Even more ambitiously, we wish to have a way of “inverting” this mapping from $f(p)$ to $P(k)$, so that we can use our knowledge of $P(k)$ in order to “resurrect” the most important features of $f(p)$.

B. Defining a k -space dark-matter profile $\tilde{g}(k)$

Our approach to this problem is somewhat unorthodox. First, we shall consider momentum slices through our dark-matter distribution packet $g(p)$, relating each slice of momentum p to a corresponding value $k_{\text{hor}}(p)$. Indeed, we shall do this without any assumptions concerning the overall shape of $g(p)$. Normally, as indicated above, k_{hor} would be interpreted as defining the *minimum* value of k within $P(k)$ which could potentially be affected by dark matter in that slice. However, we shall instead take the defining relation for $k_{\text{hor}}(p)$ in Eq. (3.7) as *defining a mapping* between the p -variable of $g(p)$ and the k -variable of $P(k)$. In other words, we shall identify k_{hor} with k and thereby consider $g(p)$ as having a corresponding profile in k -space:

$$\tilde{g}(k) \equiv g(k_{\text{hor}}^{-1}(k)) |\mathcal{J}(k)|. \quad (3.10)$$

Here k_{hor}^{-1} denotes the inverse function which relates k back to p [i.e., the inverse of the mapping in Eq. (3.7)], while $\mathcal{J}(k) \equiv d \log p / d \log k$ is the Jacobian for the change of variables from $\log p$ to $\log k$ [again as defined through the $k_{\text{hor}}(p)$ function in Eq. (3.7)]. It then follows that

$$\mathcal{N}(t) = \int d \log p g(p) = \int d \log k \tilde{g}(k). \quad (3.11)$$

Thus, just as the p -profile $g(p)$ describes the dark-matter distribution in p -space, the k -profile $\tilde{g}(k)$ describes the dark-matter distribution in k -space. Moreover, because $\tilde{g}(k)$ lives in the same space as $P(k)$, these two functions can even be plotted together along the same axis. *We shall therefore approach the question of understanding the relationship between $g(p)$ and $P(k)$ by instead seeking to understand the relationship between $\tilde{g}(k)$ and $P(k)$.*

Strictly speaking, our identification of k_{hor} with k is potentially subject to the same sorts of $\mathcal{O}(1)$ ambiguities which entered into our original definition of k_{hor} , as represented by the $\mathcal{O}(1)$ prefactor ξ in Eq. (3.7). However, for our purposes in this paper, any further $\mathcal{O}(1)$ factor involved in relating k_{hor} to k can be absorbed into an effective change in the value of ξ . Thus, for simplicity, we shall consider ξ to incorporate all such $\mathcal{O}(1)$ effects. We shall therefore identify k_{hor} directly with k , as stated above, and proceed to investigate the relationship between $\tilde{g}(k)$ and $P(k)$.

C. Connecting $\tilde{g}(k)$ to $P(k)$: General phenomenological observations

In order to develop an intuition regarding this relationship, we shall examine the response of $P(k)$ to a series of idealized forms for $g(p)$ and $\tilde{g}(k)$. For concreteness and simplicity, each of these idealized forms for $g(p)$ will be taken to consist of one or more log-normal distributions of the form

$$g(p) = \frac{A}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2} \left[\log\left(\frac{p}{\langle p \rangle}\right) + \frac{1}{2}\sigma^2\right]^2\right\}, \quad (3.12)$$

where σ and $\langle p \rangle$ are respectively the width and average momentum of the distribution and where A is an overall normalization (or dark-matter abundance). These distribution functions are essentially Gaussians in $(\log p)$ -space. Indeed, with this form, we can easily verify that $\int g(p) d \log p = A$ and $\int p g(p) d \log p = \langle p \rangle$, as claimed. Thus either σ or $\langle p \rangle$ may be altered independently. However, it is important to note that $\langle \log p \rangle \equiv \int \log p g(p) d \log p = \log \langle p \rangle - \sigma^2/2$. In other words, this Gaussian is centered (with maximum height) at $\log(p) - \sigma^2/2$ in $(\log p)$ -space, and thus any change in the value of σ will cause the location of the maximum height to shift even though $\langle p \rangle$ is kept fixed.

Let us consider the case in which $g(p)$ consists of a single Gaussian peak in $(\log p)$ -space, and let us study the effects of changing the overall dark-matter abundance A associated with this peak, keeping $\langle p \rangle$ and σ fixed. Such distributions are shown in the upper panel of Fig. 11, where we have chosen values of A such that the total abundance associated with this peak in each case is given by $\Omega = r\Omega_{\text{DM}}$ where Ω_{DM} is the total dark-matter abundance and where the ratio r ranges over the set of values specified within Fig. 11. For each value of $r < 1$, we are of course implicitly assuming the existence of an additional infinitely cold dark-matter component (not shown) with abundance $\Omega_{\text{cold}} = (1-r)\Omega_{\text{DM}}$ such that the total dark-matter abundance remains fixed at Ω_{DM} . For each value of r , the corresponding k -space distributions $\tilde{g}(k)$ and transfer functions $T^2(k) \equiv P(k)/P_{\text{CDM}}(k)$ are shown together in the lower panel of Fig. 11. For concreteness in defining $\tilde{g}(k)$, we have taken $\xi = 5/3$ within Eq. (3.7), as this value tends to horizontally align the peaks $\tilde{g}(k)$ with the onset of

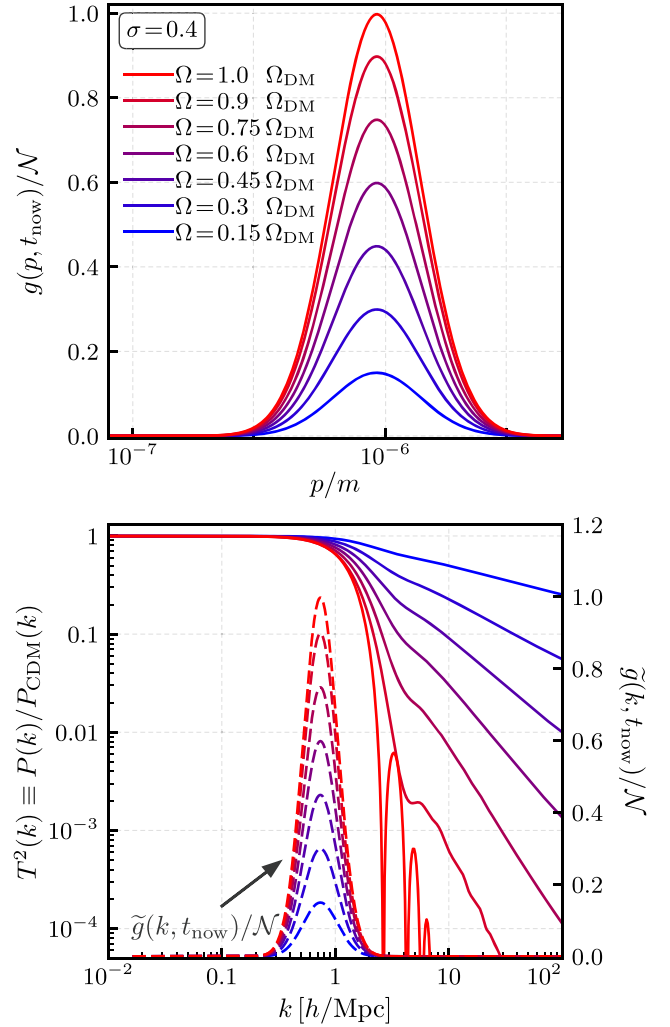


FIG. 11. Dark-matter distributions $g(p)$ which differ only in their overall normalizations (i.e., their associated abundances Ω), and their corresponding transfer functions $T^2(k) \equiv P(k)/P_{\text{CDM}}(k)$. Upper panel: A variety of dark-matter distributions $g(p)$ which share a common width σ and average momentum $\langle p \rangle$. In each case we imagine the existence of an additional cold dark-matter component (not shown) such that the total dark-matter abundance is held fixed at Ω_{DM} . Lower panel: The corresponding k -space dark-matter distributions $\tilde{g}(k)$, calculated for $\xi = 5/3$, and the corresponding transfer functions $T^2(k)$, calculated via the CLASS software package [3–6]. We see that the greater the total dark-matter abundance associated with $g(p)$, the stronger the suppression for $T^2(k)$ and the steeper the slope $d \log T^2(k)/d \log k$ at large values of k .

the suppression of structure within $T^2(k)$, as shown in Fig. 11. Indeed, we will find that taking $\xi \approx 5/3$ has a similar alignment effect for every case to be considered in this paper, not only here but also for the highly nontrivial cases to be examined in Sec. IV. We shall therefore adopt $\xi = 5/3$ as a universal reference value for the rest of this paper.

For values of k below the location of $\tilde{g}(k)$, we see that the transfer function remains close to 1, indicating that the

development of structure at these scales is insensitive to the fact that a fraction r of the dark matter is no longer being treated as infinitely cold. In other words, for the large length scales corresponding to such values of k , even the dark matter carrying the nonzero momenta within $g(p)$ is effectively cold. However, as k increases and approaches the scales associated with nonzero $\tilde{g}(k)$, our transfer function begins to peel away from unity and fall rather dramatically. This signifies the onset of the suppression of structure at these scales, as anticipated. In the case with the greatest suppression, we see from Fig. 11 that the curve for $T^2(k)$ actually begins to exhibit sharp wiggles as a function of k . These are ultimately the effects of dark-matter acoustic oscillations; such effects are generally subleading, appearing within the transfer function only when the formation of structure is significantly suppressed, and they shall not be our focus in this paper. We shall therefore disregard such wiggles in our subsequent discussion. However, we also note the degree of suppression is greater when $\tilde{g}(k)$ carries greater abundance. This immediately tells us, as expected, that there is a correlation between the abundance associated with $\tilde{g}(k)$ and the degree of structure suppression induced at larger values of k . Likewise, we also note from Fig. 11 that as k passes beyond the range with nonzero $\tilde{g}(k)$, the evolution of our transfer function $T^2(k)$ on this log-log plot in each case ultimately seems to develop a negative slope which remains essentially constant for the values of k shown. This slope also appears to be correlated with the abundance associated with $\tilde{g}(k)$, with steeper (more negative) slopes corresponding to larger abundances. Thus, at this stage, we conclude that a greater abundance for $\tilde{g}(k)$ appears to correspond not only to a stronger *suppression* for $T^2(k)$ at larger values of k , but also to a steeper *slope* for $T^2(k)$.

In order to disentangle these two effects, we now consider the case of a single peak where we now vary the *width* σ of the peak, holding A and $\langle p \rangle$ constant. This situation is shown in Fig. 12. First, we confirm from the upper panel of Fig. 12 that holding $\langle p \rangle$ fixed and increasing σ indeed causes the maximum of the peak to shift towards smaller values of $\log p$, as discussed above. Second, as we progress from smaller to larger values of k within the lower panel of Fig. 12, we observe that increasing the width σ of the dark-matter distribution induces a more gradual suppression of $T^2(k)$ as a function of k , ultimately resulting in less net suppression at large values of k .

Even more importantly, however, we also observe that increasing the width of $\tilde{g}(k)$ appears to have no effect on the logarithmic slope of the transfer function $d \log T^2 / d \log k$ at large k beyond $\tilde{g}(k)$. Indeed, the logarithmic slope of the transfer function appears to be fixed, even though the overall suppression clearly varies with the width of $g(p)$. Of course, even though we are varying the width of $g(p)$ in Fig. 12, we are holding the overall *abundance* associated with $g(p)$ fixed. This then

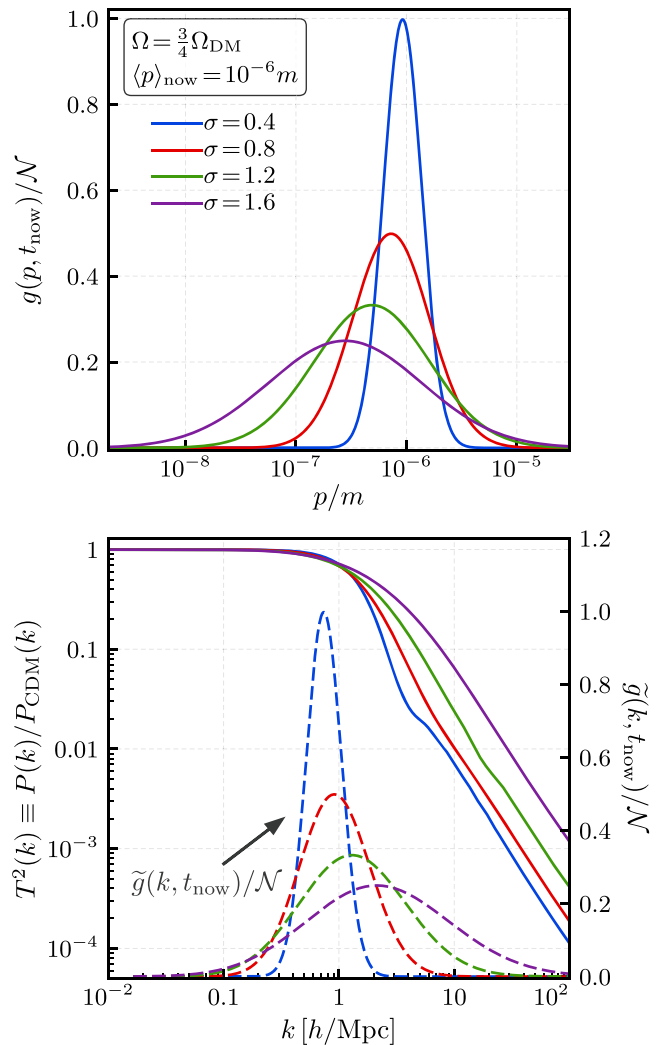


FIG. 12. Similar to Fig. 11, but for dark-matter distributions in which only the width σ is varied while the total abundance and average momentum $\langle p \rangle$ are held fixed. We see that varying the width affects the net suppression of the transfer function, *but does not affect its eventual slope* $d \log T^2 / d \log k$ at large k beyond $\tilde{g}(k)$. As discussed in the text, this suggests that the accumulated abundance correlates not with the net suppression of the transfer function, but with its logarithmic slope.

suggests that the total dark-matter abundance associated with $g(p)$ correlates with the eventual logarithmic slope of the transfer function rather than with the net amount by which $T^2(k)$ is suppressed.

This behavior actually holds throughout the range of k -values plotted in Fig. 12, and not merely at large k . Indeed, as we sweep a reference value k^* from left to right in k -space and pass through the $\tilde{g}(k)$ distribution, we accumulate an increasing abundance of dark matter for which $k < k^*$. Likewise, as we sweep from left to right in k -space, we see that the logarithmic slope of the transfer function at $k = k^*$ simultaneously becomes increasingly steep. Indeed, it is only for the values of k shown in Fig. 12

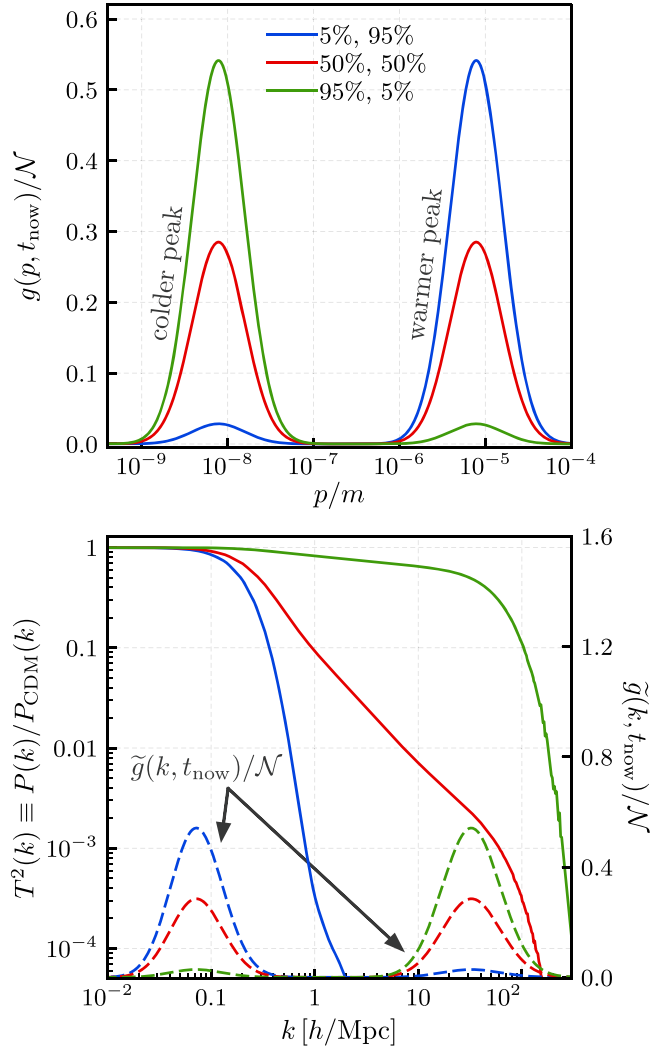


FIG. 13. Similar to Figs. 11 and 12, but for dark-matter distributions $g(p)$ consisting of two log-Gaussian peaks whose widths and average momenta are held fixed but amongst which the apportionment of the total dark-matter abundance Ω_{DM} is varied. As we sweep from smaller to larger values of k in each case, we see that $d \log T^2 / d \log k$ becomes increasingly steep as our accumulated abundance increases—precisely as anticipated from the results of Fig. 12.

beyond $\tilde{g}(k)$ that we stop accumulating abundance and the logarithmic slope of the transfer function becomes effectively constant. This observation thus allows us to refine our previous conclusions from Fig. 11 and suggests that *the accumulated abundance correlates not with the net suppression of the transfer function, but with its logarithmic slope* $d \log T^2 / d \log k$.

To see if this behavior survives for more complex $g(p)$ distributions, in Fig. 13 we consider the case of two disjoint peaks, one corresponding to colder dark matter with smaller $\langle p \rangle$ and the other to warmer dark matter with larger $\langle p \rangle$. (Note that we refer to these peaks as colder and warmer even though they need not be thermal.) In each case

we hold the width and average momentum $\langle p \rangle$ of each peak fixed, but we vary the manner in which the total dark-matter abundance Ω_{DM} is apportioned between them. We nevertheless find that the above behavior persists even in such cases. Indeed, as we sweep from smaller to larger values of k in the regions *within* the peaks, we see that the logarithmic slope of the transfer function becomes increasingly steep as our accumulated abundance increases—indeed, as the relative abundance of the warmer peak increases, the slope of the transfer function between the peaks also increases. However, as we sweep from smaller to larger values of k in the region *between* the peaks, we are no longer accumulating abundance. Likewise, the logarithmic slope of the transfer function in this region is approximately constant. Thus, we continue to find that accumulated abundance remains correlated with the logarithmic slope $d \log T^2 / d \log k$ of the transfer function.

D. From $P(k)$ back to $f(p)$: The hot fraction function $F(k)$ and a reconstruction conjecture

We could, of course, continue to examine further cases with more nontrivial $g(p)$ distributions. However, at this stage, we are actually able to formulate a conjecture which will incorporate all of our observations and which we believe holds quite generally. Moreover, as we shall see, this conjecture will not only allow us to correlate features of the dark-matter distribution $g(p)$ with those of the transfer function $T^2(k)$, but also allow us to invert the process and resurrect the salient features of $g(p)$ directly from $T^2(k)$.

In order to phrase our conjecture mathematically, we first note that at any value of k , the total abundance of dark matter accumulated from even smaller k -values is given by

$$F(k) \equiv \frac{\int_{-\infty}^{\log k} \tilde{g}(k') d \log k'}{\int_{-\infty}^{+\infty} \tilde{g}(k') d \log k'}. \quad (3.13)$$

Note that this can be equivalently written in terms of $g(p)$ rather than $\tilde{g}(k)$ as

$$F(k) \equiv \frac{\int_{\log p(k)}^{\infty} g(p') d \log p'}{\int_{-\infty}^{\infty} g(p') d \log p'} \quad (3.14)$$

where the $p(k)$ function within the lower limit of the integral in the numerator is the inverse of the $k_{\text{hor}}(p)$ function in Eq. (3.7). Thus, for any value of k , we see that $F(k)$ may be interpreted physically as that fraction of the dark matter which may effectively be considered as free-streaming (or “hot”) relative to the corresponding value of $p = k_{\text{hor}}^{-1}(k)$. We shall therefore refer to $F(k)$ as the *hot fraction function*. Note that the denominators in Eqs. (3.13) and (3.14) are nothing but the overall normalizations \mathcal{N} associated with the phase-space distribution, as defined in Eq. (2.3).

Our conjecture, then, is that there is a direct approximate relationship between the hot fraction function $F(k)$ and the logarithmic slope $d \log T^2 / d \log k$ of the transfer function. We shall write this in terms of an as-yet unknown function η :

$$F(k) \approx \eta \left(\left| \frac{d \log T^2}{d \log k} \right| \right). \quad (3.15)$$

Equivalently, taking the $(\log k)$ -derivative of both sides, we conjecture that

$$\frac{\tilde{g}(k)}{\mathcal{N}} \approx \eta' \left(\left| \frac{d \log T^2}{d \log k} \right| \right) \left| \frac{d^2 \log T^2}{(d \log k)^2} \right|, \quad (3.16)$$

where the prime on the η -function indicates a derivative with respect to its argument. Thus, if the first derivative of $\log T^2$ with respect to $\log k$ is related to the hot fraction function $F(k)$, then it is a combination of the first and second derivatives that is related to the phase-space distribution $\tilde{g}(k)$.

This conjecture is quite significant, correlating the hot fraction function $F(k)$ not with the transfer function $T^2(k)$ but with its slope. However, we can even push this one step further. In Fig. 11, we observed that the slopes $d \log T^2 / d \log k$ that emerge at the large values of k shown in this plot are correlated with the abundances Ω associated with the distributions $\tilde{g}(k)$ shown. However, within these large- k regions beyond $\tilde{g}(k)$, the accumulated hot fraction functions $F(k)$ are nothing but the fractions $r \equiv \Omega / \Omega_{\text{DM}}$ that relate these abundances Ω to the total dark-matter abundance Ω_{DM} . Thus, given these ratios and the corresponding logarithmic slopes within Fig. 11, we can investigate whether there might exist a simple empirical relation between these two quantities, and indeed we find that the relation

$$\left| \frac{d \log T^2}{d \log k} \right| \approx [F(k)]^2 + \frac{3}{2} F(k) \quad (3.17)$$

holds to rather high precision. This then fixes our unknown function η in Eq. (3.15), whereupon Eq. (3.16) takes the form

$$\frac{\tilde{g}(k)}{\mathcal{N}} \approx \frac{1}{2} \left(\frac{9}{16} + \left| \frac{d \log T^2}{d \log k} \right| \right)^{-1/2} \left| \frac{d^2 \log T^2}{(d \log k)^2} \right|. \quad (3.18)$$

This, then, is the final form of our conjecture. Indeed, if accurate, this relationship would allow us to “resurrect” $\tilde{g}(k)$ from the transfer function $T^2(k)$ and thereby deduce the dark-matter distribution $g(p)$ that produced it.

In Sec. IV, starting from an explicit Lagrangian describing a complex nonminimal dark sector, we shall explicitly test the observations we have developed thus far in this paper, including our reconstruction conjecture in

Eq. (3.18). As we shall find, in each case our conjecture (3.18) is indeed remarkably successful in reproducing the salient features of the dark-matter distribution, even when this distribution highly nontrivial and/or multimodal. Our conjecture can therefore serve as a powerful tool in the archaeology toolbox, enabling a reconstruction of many features of the early universe starting from a present-day observable such as the matter power spectrum.

Two important caveats must be borne in mind regarding our conjecture in Eq. (3.18). First, we emphasize that our conjecture is not meant to be a precise mathematical statement. Indeed, given the rather complicated nature of the Einstein evolution equations which connect $g(p)$ to $T^2(k)$, we do not expect a relation of the simple form in Eq. (3.18) to provide a precise archaeological inverse (except perhaps under some limiting approximations and simplifications). Rather, this conjecture is intended merely as an approximate practical guide—a way of reproducing the rough characteristics of $g(p)$ given a particular transfer function $T^2(k)$.

But second—and perhaps more importantly—we stress that our conjecture implicitly assumes/requires that $\log T^2$ has a negative-semidefinite second derivative with respect to $\log k$, so that it either has a constant slope or is concave-down when plotted versus $\log k$. Generally, this tends to occur in situations in which our dark-matter momentum distributions—no matter how complex in shape—are relatively “clustered” in k -space. By contrast, if there are widely separated clusters in the dark-matter phase-space distribution, the transfer function can cross an inflection point and become concave-up, potentially even leading to a plateau [7]. Indeed, this is precisely what would happen if we were to plot, e.g., the transfer functions in Fig. 11 out to even larger values of k . In such cases, our conjecture is expected to hold only within each cluster individually (just as it indeed holds for the “clustered” region of Fig. 11 shown). As we shall see in Sec. IV, this restriction to clusters is not severe, and still allows us to resurrect the salient features of the dark-matter phase-space distribution $g(p)$ for a wide variety of dynamical histories.

Finally, our conjecture in Eq. (3.18) also allows us to understand the limitations of certain proposed functional forms for $T^2(k)$ which have appeared in the literature. For example, in Ref. [8] it was shown that the functional form [9–11]

$$T(k) = [1 + (\alpha k)^\beta]^\gamma \quad \text{with} \quad \alpha, \beta > 0, \gamma < 0 \quad (3.19)$$

is remarkably successful in fitting the results from a relatively large number of underlying models of early-universe dynamics. (Note that more complex functional forms are also discussed in Ref. [8].) However, using our conjecture in Eq. (3.18), we can reconstruct the corresponding $\tilde{g}(k)$, obtaining

$$\tilde{g}(k) = \frac{Ax}{(1+x)^{3/2}\sqrt{9+Bx}}, \quad (3.20)$$

where $x \equiv (\alpha k)^\beta$, $A \equiv 4\beta^2|\gamma| > 0$, and $B \equiv 9 + 32\beta|\gamma| > 9$. As expected, this functional form corresponds to a localized packet, with $\tilde{g}(k) \sim Ak/3$ as $k \rightarrow 0$ and $\tilde{g}(k) \sim A/(\sqrt{B}k)$ as $k \rightarrow \infty$. Moreover, it is also straightforward to verify that this packet is always unimodal, with a unique maximum located at

$$x_{\max} = \frac{B-9 + \sqrt{(B-9)^2 + 144B}}{4B}. \quad (3.21)$$

Thus we learn that the generic functional form in Eq. (3.19), although quite flexible, is ultimately limited to dark-matter phase-space distributions which are unimodal. Indeed, as we have seen (and as we shall explicitly verify in Sec. IV), this is only a small part of what is possible.

IV. DECIPHERING THE ARCHAEOLOGICAL RECORD: AN ILLUSTRATIVE END-TO-END EXAMPLE

In Sec. II, we examined the manner in which decays within the dark sector can give rise to a nontrivial and often multimodal phase-space distribution $f(p)$ for the lightest particle within this sector, i.e., for the particle which constitutes the dark matter at the present time. Likewise, in Sec. III, we studied the relationship between a given phase-space distribution $f(p)$ and the corresponding linear matter power spectrum $P(k)$, ultimately proposing a conjecture in Eq. (3.18) which would allow us to approximately reconstruct $f(p)$ from knowledge of $P(k)$. However, our analysis thus far has been primarily conceptual and limited to examples involving only a small number of dark-sector states. Likewise, we have only examined each step within Eq. (1.1) independently, without performing a complete end-to-end analysis.

In this section, we shall examine how these ideas play out within a more complex framework—one involving a relatively large number of dark-sector states and decay pathways. In this way, we seek to determine the extent to which our general observations from Secs. II and III are robust, remaining more or less intact as the dark sector grows in complexity, even under a full numerical Boltzmann analysis. We also wish to see how the different steps of the process within Eq. (1.1) actually connect to each other, starting with a particular model of early-universe dynamics and ending with a predicted matter power spectrum $P(k)$.

Towards this end, in this section we shall construct a phenomenologically rich illustrative model within which our calculations will take place. In particular, we shall begin with an explicit Lagrangian describing a relatively large number of dark-sector states. From this we shall then proceed to study the various decay amplitudes, the patterns of possible allowed decay chains, the resulting dark-matter

phase-space distributions $f(p)$, and the corresponding matter power spectra $P(k)$. This will ultimately enable us to determine the extent to which the phenomena discussed in Secs. II and III are realized within a relatively complex dark sector. This will also afford us the opportunity of explicitly testing our reconstruction conjecture in Eq. (3.18). We emphasize, however, that the illustrative model we shall consider is not meant to be a UV-complete description of an actual fully realized dark sector. Rather, our purpose here is merely to establish a framework in which to study the rich set of possible phenomenologies associated with intra-ensemble decays and to verify that the basic expectations we have discussed in Secs. II and III are indeed realized, including the potential for archaeological reconstruction.

A. The model

Our model consist of an ensemble of $N + 1$ real scalar fields ϕ_ℓ , where the index $\ell = 0, 1, \dots, N$ labels these fields in order of increasing mass. We take these scalars to be singlets under the Standard-Model (SM) gauge group and assume that they couple only negligibly to the fields of the visible sector. Likewise, we shall assume that the behavior of these fields is governed by the Lagrangian

$$\mathcal{L} = \sum_{\ell=0}^N \left(\frac{1}{2} \partial_\mu \phi_\ell \partial^\mu \phi_\ell - \frac{1}{2} m_\ell^2 \phi_\ell^2 \right) + \mathcal{L}_{\text{int}}, \quad (4.1)$$

where m_ℓ denotes the mass of ϕ_ℓ and where the interaction Lagrangian \mathcal{L}_{int} consists of terms involving the fields ϕ_ℓ alone. For simplicity, we shall consider a form for \mathcal{L}_{int} in which two-body intra-ensemble decays of the form $\phi_\ell \rightarrow \phi_i \phi_j$ dominate the decay width of each unstable ϕ_ℓ . In particular, we consider an interaction Lagrangian which includes trilinear terms of the form

$$\mathcal{L}_{\text{int}} \ni - \sum_{\ell=0}^N \sum_{i=0}^{\ell} \sum_{j=0}^i c_{\ell ij} \phi_\ell \phi_i \phi_j, \quad (4.2)$$

where the $c_{\ell ij}$ are coupling coefficients with dimensions of mass. We emphasize that in addition to the trilinear terms in Eq. (4.2), \mathcal{L}_{int} will in general include additional quartic terms involving the ϕ_ℓ . Such terms play an important role in stabilizing the scalar potential; moreover, they may also play a role in establishing the primordial abundances of these fields at early times. However, due to phase-space considerations, they typically play only a subleading role in the decay phenomenology of the ϕ_ℓ at times well after those abundances have been established. In what follows, we shall therefore assume that the coupling coefficients associated with such terms are sufficiently small that these terms have negligible impact on particle dynamics within the dark sector at late times, and focus on the consequences of the trilinear terms in Eq. (4.2).

The different masses m_ℓ and couplings $c_{\ell ij}$ associated with our individual ensemble constituents will not be considered to be independent parameters of our model. Rather, both the masses and couplings will be assumed to scale across the ensemble according to a set of scaling relations involving only a small number of free parameters. In particular, we shall assume that the m_ℓ scale across the ensemble according to the scaling relation

$$m_\ell = m_0 + \ell^\delta \Delta m, \quad (4.3)$$

where m_0 denotes the mass of the lightest ensemble constituent, where Δm is a parameter with dimensions of mass, and where δ is a dimensionless scaling exponent which controls how the density of states scales across the ensemble.

A variety of scaling relations are in principle possible for the coupling coefficients $c_{\ell ij}$. Given that our primary aim in this paper is to investigate the extent to which information about the decay history of the dark sector can be gleaned from the phase-space distribution of a single, stable particle species which plays the role of the dark matter today, we shall adopt a scaling relation for the $c_{\ell ij}$ which is amenable to such a study. First, we shall assume that the coupling structure of our model is such that only the lightest ensemble constituent ϕ_0 is stable and that a contribution to the decay width of each constituent ϕ_ℓ with $\ell > 0$ arises from the interaction Lagrangian in Eq. (4.2). Moreover, we also wish to examine how the interplay between the coupling structure of the dark sector, the decay kinematics within the ensemble, and the cosmological background ultimately gives rise to the phase-space distribution $g_0(p, t_{\text{now}})$ for this lightest ensemble constituent at present time.

There are two salient kinematic quantities which are of particular interest in characterizing two-body decays of the form $\phi_\ell \rightarrow \phi_i \phi_j$. The first of these quantities is the total energy released in the decay—i.e., the difference between the mass of the decaying particle and the sum of the masses of the two daughter particles. The second is the difference in mass between the two daughter particles. Motivated by these considerations, we shall therefore adopt a parametrization for the $c_{\ell ij}$ in which

$$c_{\ell ij} = \mu R_{\ell ij} \left(\frac{m_\ell - m_i - m_j}{\Delta m} \right)^r \left(1 + \frac{|m_i - m_j|}{\Delta m} \right)^{-s} \times \Theta(m_\ell - m_i - m_j). \quad (4.4)$$

Here μ is a parameter with dimensions of mass which sets the overall scale of the couplings, while $\Theta(m_\ell - m_i - m_j)$ is a Heaviside theta function, r and s are dimensionless free parameters, and the combinatoric factor

$$R_{\ell ij} \equiv \begin{cases} 6 & \text{all indices different} \\ 3 & \text{only two indices equal} \\ 1 & \text{all indices equal} \end{cases} \quad (4.5)$$

is defined such that

$$\sum_{\ell=0}^N \sum_{i=0}^{\ell} \sum_{j=0}^i R_{\ell ij} \phi_\ell \phi_i \phi_j = \sum_{m,n,p=0}^N \phi_m \phi_n \phi_p. \quad (4.6)$$

Note that even though $i \geq j$ in Eq. (4.2), the absolute-value signs in Eq. (4.4) ensure that $c_{\ell ij} = c_{\ell ji}$. This property will be useful later.

The parameters r and s appearing in Eq. (4.4) have a straightforward interpretation and will be critical for our analysis. The parameter r governs the manner in which $c_{\ell ij}$ scales with the overall kinetic energy released during the decay process $\phi_\ell \rightarrow \phi_i \phi_j$. Taking $r > 0$ establishes a preference for highly exothermic decays involving a large conversion of mass energy into kinetic energy—in other words, decays from heavy parents directly into relatively light daughters which therefore behave more like radiation. By contrast, taking $r < 0$ establishes a preference for minimally exothermic decays in which relatively little kinetic energy is released and the daughter particles behave more like matter. Likewise, the parameter s governs the extent to which the daughter particles are close in mass to each other. Taking $s > 0$ establishes a preference for decays in which there is a high degree of symmetry between the masses of the daughter particles. By contrast, taking $s < 0$ disfavors such decays relative to those in which such a symmetry between the daughters is significantly broken. Thus, while positive r favors decays producing radiation, positive s favors decays whose daughters are symmetric.

Examining the phenomenology that results from different values of r and s therefore allows us to survey many different kinds of decay chains and their corresponding dark-matter phase-space distributions. Of course, a fully realistic model of the dark sector is unlikely to exhibit a coupling structure of the specific form in Eq. (4.4). However, as we shall see, this structure is capable of realizing many if not most of the different decay phenomenologies that could emerge within a fully realistic multi-component dark sector, and our goal in this section is to study the implications of these different decay phenomenologies rather than their specific realization within a UV-complete theory. Adopting the coupling structure in Eq. (4.4) will therefore be sufficient for our purposes.

Given the number of free parameters which govern our model and the variety of possible initial conditions for the $g_\ell(p, t_I)$, it will prove useful for us to adopt a few benchmark assumptions as we proceed. In particular, for concreteness, we shall focus on the case of an ensemble with $N = 9$ —i.e., an ensemble comprising ten constituent particles with ten distinct masses. Such a value of N is sufficiently large that a highly nontrivial pattern of intra-ensemble decays can arise, yet sufficiently small that the evolution of the Boltzmann system is not computationally onerous. We shall also fix $\delta = 1$, $\Delta m = 2m_0$, and $\mu = m_0/10$ in what follows. Given

these benchmark assumptions, there remains one free parameter m_0 on which the mass spectrum of our model depends, as well as two free parameters r and s which govern the coupling structure of the model.

For simplicity, we shall consider the case in which only the most massive constituent ϕ_9 in the ensemble is initially populated at t_I , while the energy density in all other ensemble constituents is negligible. For sake of generality, we shall remain largely agnostic about the precise nature of the mechanism which establishes the initial energy density $\rho_9(t_I)$ for this field. Rather, we shall simply assume that this occurs during a production-time window Δt_P and effectively ceases acting by some cosmological time $t_P \leq t_I$. Note that in order to ensure that the change in $\rho_9(t)$ due to decays is not significant during the production-time window, it is sufficient to require that

$$\begin{cases} t_I - t_{P_\ell} \ll \tau_9 \\ \Delta t_{P_\ell} \ll \tau_9, \end{cases} \quad (4.7)$$

where $\tau_9 \equiv 1/\Gamma_9$ is the lifetime of ϕ_9 . Indeed, the conditions in Eq. (4.7) are equivalent to the condition already mentioned in Eq. (2.30) which ensures effectively continuous daughter packets, even for extremely nonrelativistic parents. We emphasize that a production mechanism for which $\tau_9 \leq t_I$ is not in conflict with these conditions, provided that Δt_P is small and provided that the time t_I at which we begin our numerical Boltzmann evolution is sufficiently close to t_P .

Depending on the nature of the production mechanism through which $\rho_9(t_I)$ is established, the initial phase-space distribution $g_9(p, t_I)$ for ϕ_9 can take a variety of forms. In what follows, we shall assume that the initial population of ϕ_9 is sufficiently cold that the detailed shape of $g_9(p, t_I)$ has essentially no impact on the subsequent evolution of the phase-space distributions of all particles produced through the decay chains initiated by ϕ_9 decay. Indeed, in the regime in which $g_9(p, t)$ effectively only receives support at very small p , the phase-space packet for each daughter particle produced directly by the decay process $\phi_9 \rightarrow \phi_i \phi_j$ will be sharply peaked around $p \approx \frac{1}{2}(m_0 - m_i - m_j)$. Thus, within this regime, the profiles of the phase-space packets generated for the two daughter particles ϕ_i and ϕ_j are not particularly sensitive to the shape of $g_9(p, t_I)$, nor are the profiles of the packets for the particles produced via subsequent decays. Indeed, we have verified numerically that varying the shape of $g_9(p, t_I)$ while holding the initial number density $n_9(t_I)$ fixed does not significantly impact our results for either $g_0(p, t_{\text{now}})$ or $P(k)$, provided that the condition $f_9(p, t) \ll 1$ is satisfied for all $t \geq t_I$.

Given these assumptions, only two initial conditions for our example scenario remain to be specified. These are the initial time t_I itself and the overall normalization of $g_9(p, t_I)$ at this initial time. The system then evolves dynamically according the Boltzmann equations outlined

in Appendix A. However, as briefly noted in Sec. II, the Boltzmann equations are greatly simplified and possess several attractive properties if they can be approximated as linear. We can make such an approximation if the phase-space distributions are sufficiently small (i.e., if $f_\ell(p, t) \ll 1$, so that Pauli-blocking and Bose-enhancement effects are negligible) and if scattering and inverse-decay processes amongst the scalars are also negligible. As discussed in Appendix A, these conditions are in turn both satisfied as long as the normalization of $g_9(p, t_I)$ remains sufficiently small. Thus, the Boltzmann evolution in our ensemble is effectively linear, and rescaling the initial normalization of $g_9(p, t_I)$ within this limit simply has the effect of rescaling the other phase-space distributions by the same constant factor. Furthermore, we have shown that our system of Boltzmann equations is also invariant under a particular set of transformations of the dimensionful parameters in our model. Thus, for the sake of generality, in what follows we shall refrain from specifying particular values for these quantities until we examine the consequences of our model for the matter power spectrum and thereby make contact with observational constraints. Of course, due to phenomenological considerations, we shall ultimately require that these parameters be chosen such that the energy density carried by each ϕ_n with $n > 1$ is negligible by the time of big-bang nucleosynthesis (BBN).

B. Partial widths and decay chains

Given the interaction Lagrangian in Eq. (4.2) and the mass spectrum and coupling structure in Eqs. (4.3) and (4.4), it is straightforward to determine the partial decay widths Γ_{ij}^ℓ for all kinematically allowed intra-ensemble decay processes of the form $\phi_\ell \rightarrow \phi_i + \phi_j$ arising in our model. The partial widths are given by

$$\Gamma_{ij}^\ell = \frac{1}{1 + \delta_{ij}} \frac{|\vec{p}_{\text{CM}}|}{8\pi m_\ell^2} |c_{\ell ij}|^2, \quad (4.8)$$

where the momentum of either daughter particle in the rest frame of ϕ_ℓ is given by

$$|\vec{p}_{\text{CM}}| = \frac{\sqrt{[m_\ell^2 - (m_i + m_j)^2][m_\ell^2 - (m_i - m_j)^2]}}{2m_\ell}. \quad (4.9)$$

Note that $\Gamma_{ij}^\ell = \Gamma_{ji}^\ell$. Since no additional decay processes with non-negligible partial widths exist for any of the ensemble constituents, the total width Γ_ℓ of ϕ_ℓ is merely

$$\Gamma_\ell \equiv \sum_{i=0}^{\ell-1} \sum_{j=0}^i \Gamma_{ij}^\ell, \quad (4.10)$$

where the summation is over all configurations of daughters (i, j) with $j \leq i$ such that $m_i + m_j \leq m_\ell$.

In Fig. 14, we show how the partial widths Γ_{ij}^9 for the decays $\phi_9 \rightarrow \phi_i\phi_j$ which initiate these decay chains depend on the parameters r and s introduced in Eq. (4.4). Each individual panel of the figure corresponds to a particular choice of the parameters r and s , with r increasing from top to bottom and s increasing from left to right. Within a given panel, the color of a particular square in the grid indicates the value of Γ_{ij}^9 for the corresponding assignment of r and s , normalized to the partial width Γ_{00}^9 obtained for the parameter assignments $r = s = 0$.

It is evident from Fig. 14 that a variety of scaling behaviors for the Γ_{ij}^9 as a function of the indices i and j can arise as a result of the interplay between the parameters r and s . For example, for $r < 0$ and $s < 0$ (top left panel), decays to daughters ϕ_i and ϕ_j with vastly different values of i and j are preferred, as are decays with small mass gaps $m_9 - m_i - m_j$. As a result, the partial widths are largest for final states such as $\phi_9 \rightarrow \phi_8\phi_0$ and $\phi_9 \rightarrow \phi_7\phi_1$. By contrast, for $r > 0$ and $s < 0$ (bottom left panel), the same preference for daughters with vastly different values of i and j persists, but is now combined with a preference for highly exothermic decays. This leads to a situation in which the largest partial widths are those to final states such as $\phi_9 \rightarrow \phi_5\phi_1$ and $\phi_9 \rightarrow \phi_4\phi_0$. For $r < 0$ and $s > 0$ (top right panel), a strong preference for symmetry between the daughter-particle masses, combined with a preference against strongly exothermic decays, leads to a situation in which the single decay channel $\phi_9 \rightarrow \phi_4\phi_4$ dominates. For the case in which $r = s = 0$, the Γ_{ij}^9 for all possible decay channels are quite similar. However, we emphasize that they are not precisely identical, due to the dependence of $|\vec{p}_{\text{CM}}|$ in Eq. (4.9). The scaling behaviors for the Γ_{ij}^9 displayed in the remaining panels of the figure can be also understood based on these same considerations.

We also see from Fig. 14 that the values of Γ_{ij}^9 for the processes which dominate the total width Γ_9 of ϕ_9 are significantly larger for certain combinations of r and s than others. In particular, we see that Γ_9 is largest when r and s are positive and negative, respectively, and smallest when r and s are respectively negative and positive. Thus, we see that Γ_9 itself varies significantly as a function of r and s .

We now proceed to examine the decay chains which ultimately determine the structure of $g_0(p, t_{\text{now}})$. Indeed, while an examination of the partial widths Γ_{ij}^9 provides some insight into how the decay phenomenology of our model depends on the choice of our model parameters r and s , we must consider the full set of possible decay pathways through which the energy density initially stored in ϕ_9 is ultimately transferred to ϕ_0 in order to characterize how the coupling structure of our model impacts the present-day phase-space distribution $g_0(p, t_{\text{now}})$ of the lightest ensemble constituent ϕ_0 .

Because we are considering two-body decays of the form $\phi_\ell \rightarrow \phi_i\phi_j$ (each with a different branching fraction), and

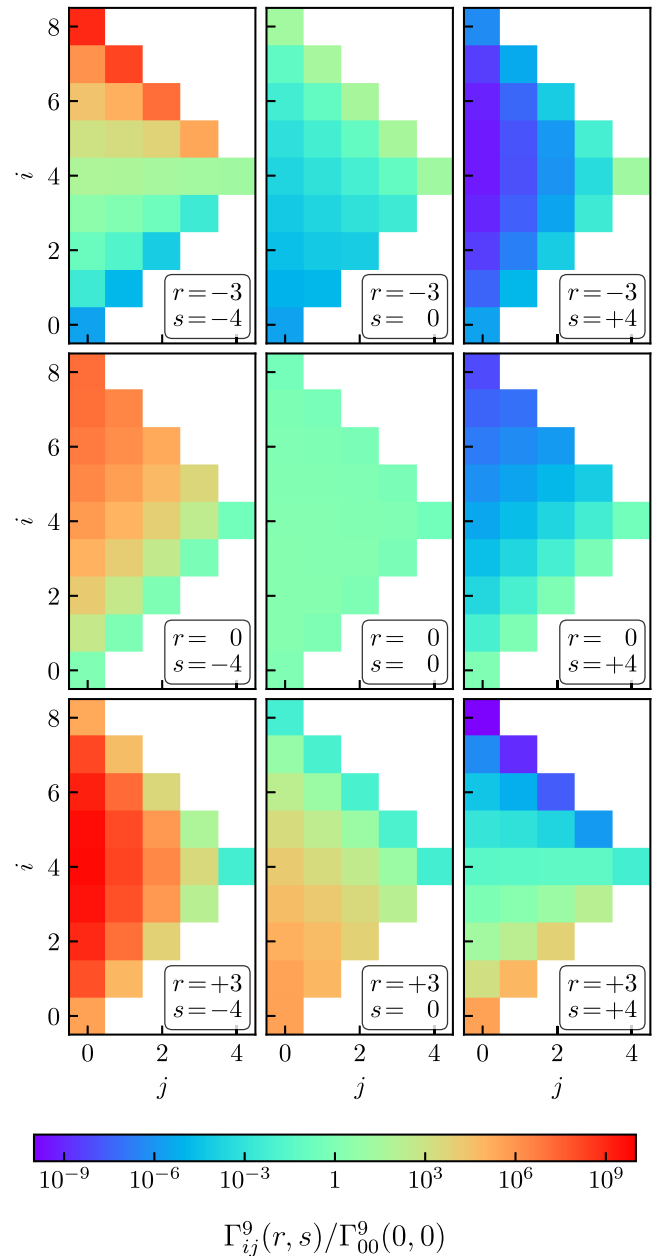


FIG. 14. Partial widths for the various decay processes of the form $\phi_9 \rightarrow \phi_i\phi_j$ with $i \geq j$ through which the heaviest constituent ϕ_9 in our particle ensemble decays. Each panel of this figure corresponds to a particular choice of r and s , with r increasing from top to bottom and s increasing from left to right. Within a given panel, the color of a particular square in the grid indicates the value of Γ_{ij}^9 for the corresponding assignment of r and s , normalized to the partial width Γ_{00}^9 obtained for the parameter assignments $r = s = 0$.

because each of these daughters then produces two granddaughters (whose identities are determined according to another set of branching fractions), the collection of particles produced from a single ancestor proliferates quickly through each generation, as does the full set of potential decay chains and their associated net branching

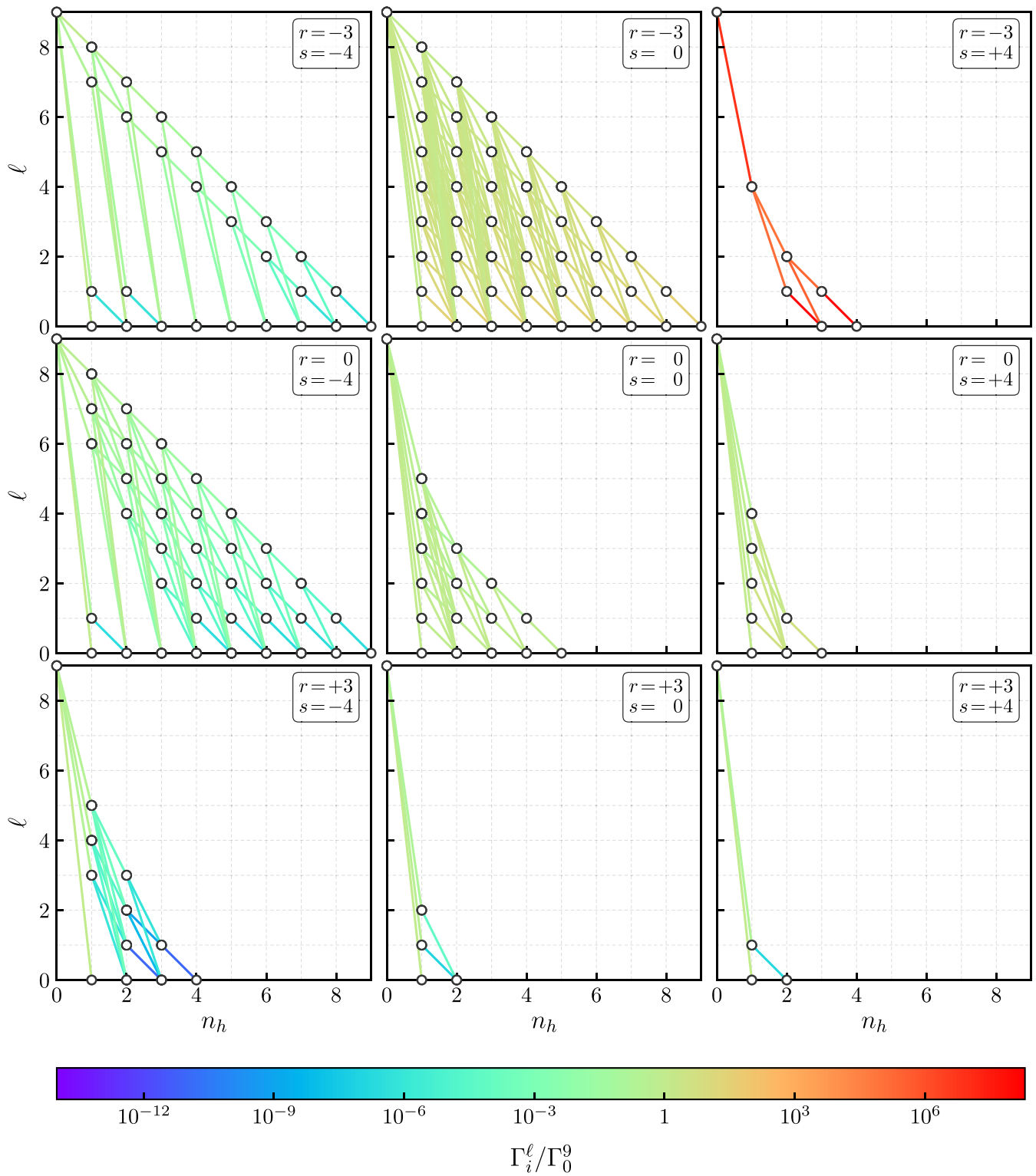


FIG. 15. The dominant reduced decay chains that arise within our model, starting from the heaviest field ϕ_9 . The different panels of this figure correspond to the same choices of r and s as in Fig. 14. The vertical axis in each panel indicates the index ℓ associated with a particular particle produced along the decay chain, while the horizontal axis indicates the number n_h of individual steps or “hops” that have occurred along the reduced decay chain in order to produce that particle. Each line segment from parameters (ℓ_1, n_h) to parameters $(\ell_2, n_h + 1)$ thus corresponds to the production of a daughter particle ϕ_{ℓ_2} directly from the decay of a parent particle ϕ_{ℓ_1} , with an associated production rate $\Gamma_{\ell_2}^{\ell_1}$. For visual simplicity, we have indicated only those “dominant” parent-to-daughter segments for which $\Gamma_{\ell_2}^{\ell_1}$ exceeds 5% of the total production rate $2\Gamma_{\ell_1}$ stemming from ϕ_{ℓ_1} , with the corresponding colors indicating the value of $\Gamma_{\ell_2}^{\ell_1}$ normalized to Γ_0^9 in each panel.

fractions. In this paper we have nevertheless performed this analysis exactly, and all of the results that we shall show in this paper have traced all of the possible decay chains in this way. However, in order to interpret our results physically and illustrate them graphically, it will suffice to simplify our discussion somewhat by considering what we shall call *reduced* decay chains. Essentially, for any given parent ϕ_ℓ , we can ask which of all possible daughters ϕ_i with masses $m_i \leq m_\ell$ is most likely to be produced *regardless of the identity of the other sibling*. Indeed, the total rate for producing a given daughter ϕ_i from a parent ϕ_ℓ is given by

$$\Gamma_i^\ell \equiv \sum_{j=0}^N \mathcal{N}_{ij} \Gamma_{ij}^\ell, \quad (4.11)$$

where $\mathcal{N}_{ij} = 1 + \delta_{ij}$ is the multiplicity of ϕ_i in the final state of any individual process $\phi_\ell \rightarrow \phi_i \phi_j$. We shall then construct our reduced decay chains by stitching together sequences of such $\ell \rightarrow i$ parent-to-daughter transitions.

In Fig. 15, we show the dominant reduced decay chains which arise within the context of our model, given our choice of initial conditions. Each line segment shown in the figure corresponds to the production of a daughter particle ϕ_i directly from the decay of a parent particle ϕ_ℓ , while the color of the segment provides a measure of the rate Γ_i^ℓ at which daughters ϕ_i are produced directly from the decays of ϕ_ℓ . Thus the segments whose colors tend towards red in each panel indicate a more rapid production of daughters than those in the same panel whose colors tend towards blue. In order to maintain visual clarity, for each parent ϕ_ℓ we have indicated only those segments whose daughters ϕ_i have the highest net probabilities for being produced. These are defined as those processes for which the corresponding production rate Γ_i^ℓ exceeds 5% of the total particle-production rate stemming from ϕ_ℓ decay—i.e., processes for which $\Gamma_i^\ell / (2\Gamma_\ell) \geq 0.05$.

It is straightforward to interpret the results shown in Fig. 15 in terms of deposits onto the cosmological conveyor belt associated with the lowest state ϕ_0 , as discussed in Sec. II. Of course, each decay chain shown within the different panels of Fig. 15 ultimately terminates when ϕ_0 is produced. This then is the time of the deposit onto the ϕ_0 conveyor belt. However, it is evident from the different colors associated with these decay chains that they can potentially proceed with very different overall rates. In general, we may associate a characteristic timescale associated with a given reduced decay chain by considering the sum of the inverses of the aggregate production rates Γ_i^ℓ for all of the decay processes occurring within that decay chain. Thus, a rough estimate of this timescale can be obtained from the timescale associated with the slowest individual decay step within the chain. In Fig. 15, these are segments whose colors tend towards the blue end of the color spectrum rather than the red.

In the four panels of the figure for which $r \leq 0$ and $s \geq 0$ —i.e., the four panels constituting the upper right corner of the figure—the timescales for all decay chains which contribute significantly to the production of ϕ_0 are quite similar. In other words, each of these decay chains in these panels tends to make its deposit onto the g_0 conveyor belt at roughly the same time. As discussed in Sec. II, we therefore expect the corresponding $g_0(p, t_{\text{now}})$ to be essentially unimodal for such combinations of r and s . By contrast, we observe that in the remaining figures in the plot—i.e., those for which $r > 0$ and/or $s < 0$ —many of the decay chains which contribute significantly to the production of ϕ_0 have vastly different timescales, as indicated in Fig. 15 through their significantly different colors. Consequently, in such cases we expect that a nontrivial, multimodal phase-space distribution will be generated for ϕ_0 .

Our understanding of the conveyor-belt dynamics from Sec. II also enables us to form comparative expectations regarding $g_0(p, t_{\text{now}})$ across the different panels. For example, we have already seen that in Fig. 14 that the decay widths involved for $r < 0$ and $s > 0$ are significantly smaller than those involved for $r > 0$ and $s < 0$. Thus, even though all of the decay chains within the first case arise might make their deposits at approximately the same time as each other, this time of deposit is much later than the time of the first deposits that occur in the second case. Thus we expect the unimodal dark-matter phase-space distribution $g_0(p, t_{\text{now}})$ in the first case to have less time to redshift than the contributions from the first mode of the phase-space distribution in the second case, implying that $g_0(p, t_{\text{now}})$ should be situated at higher momenta than the lowest-momentum peak within the distribution $g_0(p, t_{\text{now}})$ in the second case.

C. Dark-matter phase-space distributions and matter power spectra

In order to determine the extent to which these expectations are realized within our model, we have calculated the actual dark-matter phase-space distributions $g_0(p, t_{\text{now}})$ that result in each case at the present time, long after all of the decays within each decay chain have terminated. We stress that this calculation has been performed through a full Boltzmann analysis, as outlined in Appendix A, with all source and decay terms included. Indeed, as the decay process unfolds towards the ground state, there is considerable variety in the intermediate stages through which our dark sector passes. This nontrivial time evolution of the dark sector is outlined in Appendix D. However, in this section, our main interest concerns the final results that emerge after all of the decays are concluded.

Our results for the dark-matter phase-space distributions $g_0(p, t_{\text{now}})$ are shown in Fig. 16. Once again, the different panels of Fig. 16 correspond to the same choices of r and s as in Figs. 14 and 15. In order to facilitate comparison

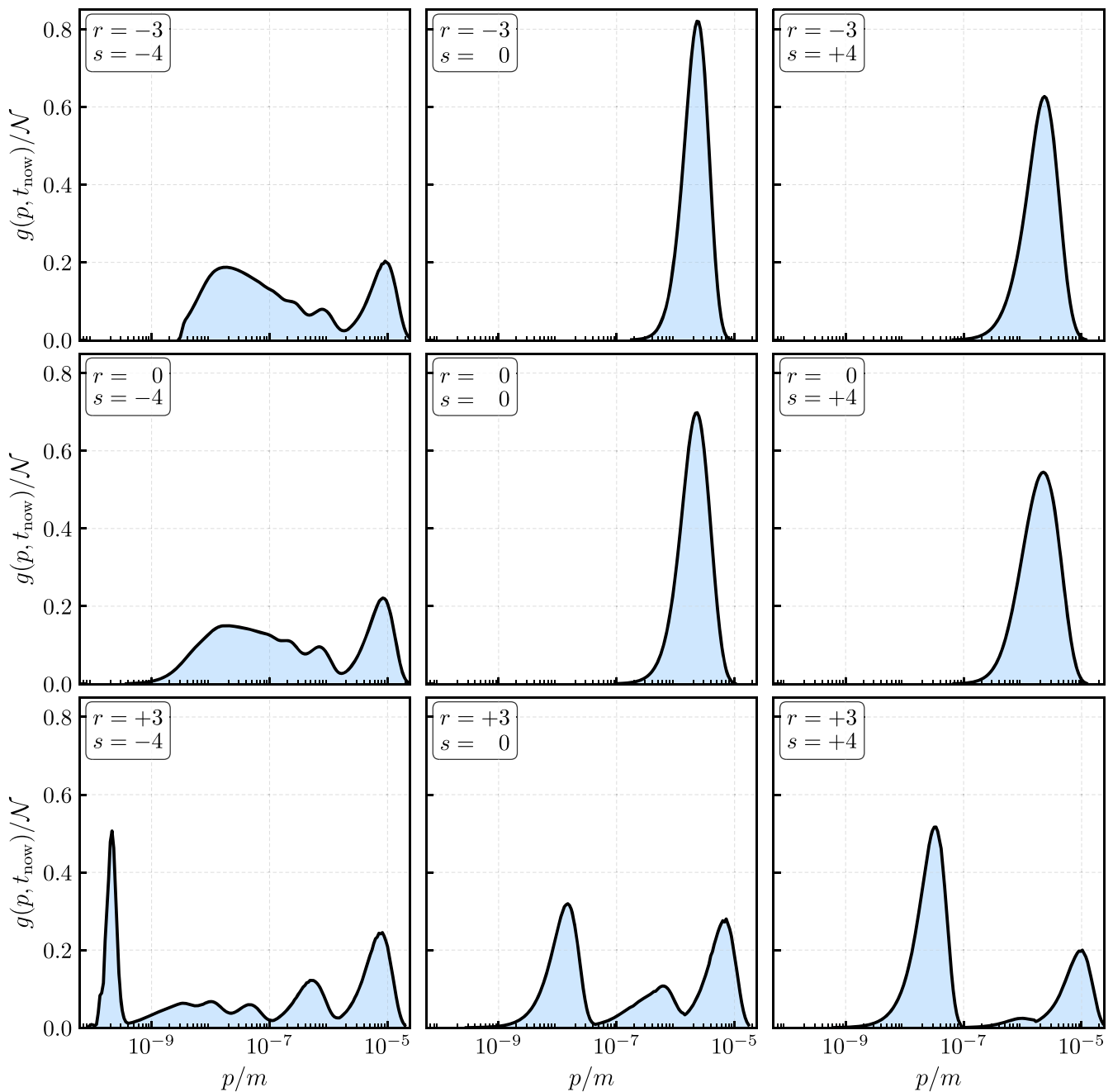


FIG. 16. The phase-space distributions $g_0(p, t_{\text{now}})$ in our model for the lightest ensemble constituent ϕ_0 at the present time. The different panels correspond to the same choices of r and s as in Figs. 14 and 15. The initial conditions in each case have been chosen such that these distributions share not only a common overall normalization but also a common value for the ratio $\langle p \rangle_{\text{now}}/m_0$, where $\langle p \rangle_{\text{now}}$ is the average momentum of the distribution at present time.

between the results obtained for different values of r and s , we have chosen the initial conditions in each case such that all of our final distributions $g_0(p, t_{\text{now}})$ share not only a common overall normalization but also a common value for the ratio $\langle p \rangle_{\text{now}}/m_0$, where $\langle p \rangle_{\text{now}}$ is the average present-day momentum of the distribution. The overall normalization for $g_0(p, t_{\text{now}})$ is chosen to correspond to the observed present-day dark-matter energy density

$\rho_{\text{DM}} \approx 10^{-5} h^2 \text{ GeV cm}^{-3}$. Likewise, for all of the plots in Fig. 16 we have chosen the value

$$\frac{\langle p \rangle_{\text{now}}}{m_0} = \frac{T_{\text{CMB}}}{100 \text{ eV}} \approx 2.3 \times 10^{-6}, \quad (4.12)$$

where $T_{\text{CMB}} \approx 0.23 \times 10^{-3} \text{ eV}$ is the present-day temperature of the cosmic microwave background (CMB)

radiation. This value is motivated as follows. In general, we are interested in situations in which a significant portion of the total dark matter exhibits speeds lying within the rough window $v \approx p/m_0 \sim \mathcal{O}(10^{-8} - 10^{-5})$. Indeed, if a significant fraction of the dark matter has speeds above this

window, the resulting dark matter will be too “hot” to accord with observation. On the other hand, the free-streaming of dark matter with speeds below this window will eventually have effects on $P(k)$ which are evident only at large values of k which are beyond the range at which

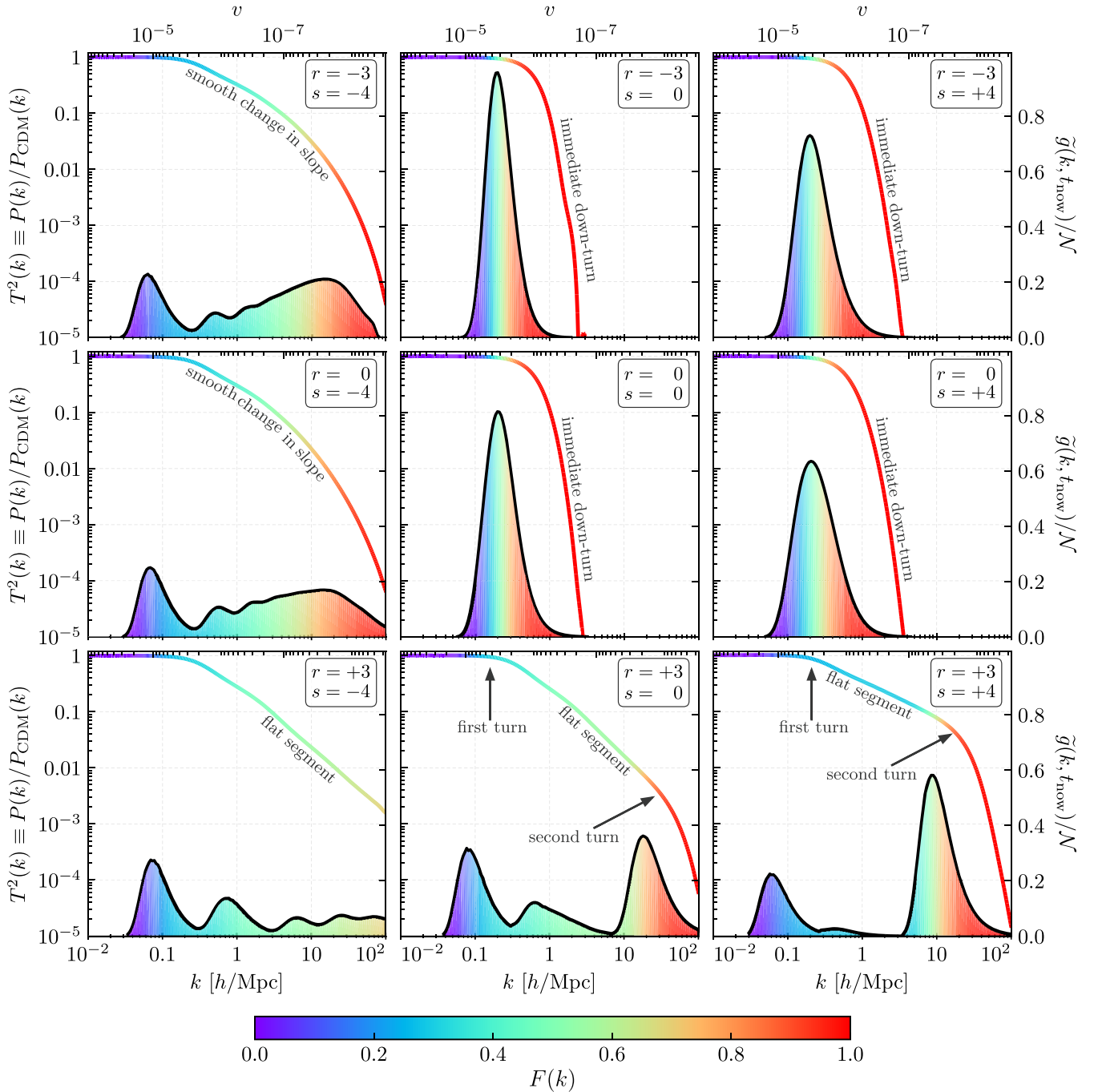


FIG. 17. The k -space distributions $\tilde{g}(k)$ defined in Eq. (3.10) and squared transfer functions $T^2(k) \equiv P(k)/P_{\text{CDM}}(k)$ corresponding to each of the phase-space distributions $g_0(p, t_{\text{now}})$ shown in Fig. 14. As we sweep from smaller to larger values of k within the k -space distributions $\tilde{g}(k)$ (outlined in solid black curves), we have shaded the area inside the $\tilde{g}(k)$ distributions according to the accumulated hot fraction function $F(k)$ defined in Eq. (3.13). These same colors are then used for plotting the corresponding transfer functions $T^2(k)$ as functions of k . The tick marks along the bottom, top, left, and right axes of each panel correspond respectively to the wave number k , the corresponding velocity v , the transfer function $T^2(k)$, and the k -space distribution $\tilde{g}(k)$.

they can reliably be probed observationally. Motivated by these considerations, we have therefore chosen to anchor the overall scale of the present-day momentum $\langle p \rangle$ for our $g_0(p, t_{\text{now}})$ distribution as in Eq. (4.12) such that the speed of a typical dark-matter particle lies within our window of interest.

In this connection, we observe that the Boltzmann equations which give rise to $g_0(p, t)$ are invariant under the rescalings specified in Eq. (A12). Likewise, our eventual power spectrum $P(k)$ depends only on the velocities $v \approx p/m_0$ of ϕ_0 rather than on p and m_0 independently. This result of course assumes that the conditions $f_i(p, t) \ll 1$ are likewise satisfied for all of the ϕ_i distributions in our ensemble during all times $t \geq t_I$ during which our Boltzmann evolution takes place. In particular this must be true for $f_9(p, t_I)$ —corresponding to the heaviest state the ensemble—prior to the decay process. We have verified numerically that even for m_0 as small as 20 keV, which corresponds to a present-day average momentum $\langle p \rangle_{\text{now}} = 200T_{\text{CMB}}$, these consistency criteria are satisfied. Indeed, the required normalization for $g_9(p, t_I)$ is such that $f_i(p, t) \ll 1$ for all ϕ_i in the ensemble at all $t \geq t_I$ for all combinations of r and s considered in Fig. 16.

We see from the results in Fig. 16 that all of our expectations from Sec. II are indeed borne out. For example, as anticipated, the phase-space distributions $g_0(p, t_{\text{now}})$ obtained for $r \leq 0$ and $s \leq 0$ are essentially unimodal, while the distributions which arise either for $r > 0$ or for $s < 0$ exhibit a complicated, multimodal structure. Likewise, we see that single peaks within the unimodal distributions sit at higher momenta than the lowest-momentum peaks within the multimodal distributions, again in line with our expectations above.

These results also demonstrate one of our main themes of this paper, namely that fairly complex and even multimodal dark-matter phase-space distributions can easily arise, even when the full details of the Boltzmann evolution are incorporated into the analysis. Indeed, the existence of many independent levels and the proliferation of overlapping decay chains do not wash out the nontrivial structures for $g_0(p)$ that we anticipated in Sec. II. Instead, they only serve to enhance these structures.

We now turn to examine the k -space distributions $\tilde{g}(k)$ and squared transfer functions $T^2(k) \equiv P(k)/P_{\text{CDM}}(k)$ corresponding to each of the phase-space distributions $g_0(p, t_{\text{now}})$ shown in Fig. 16. Our results are shown in Fig. 17. For each phase-space distribution $g_0(p, t_{\text{now}})$, the corresponding k -space distribution $\tilde{g}(k)$ is defined in Eq. (3.10). Likewise, we have calculated each corresponding transfer function $T^2(k)$ using the CLASS [3–6] software package. As we sweep from smaller to larger values of k within the k -space distributions $\tilde{g}(k)$ in Fig. 17, we have shaded the area inside the $\tilde{g}(k)$ distributions according to the accumulated hot fraction function $F(k)$ defined in Eq. (3.13). These same colors are then used for

plotting the corresponding transfer functions $T^2(k)$ as functions of k .

The results in Fig. 17 are once again in complete accordance with our expectations from Sec. III. In particular, we see that the logarithmic slope $d \log T^2(k)/d \log k$ of the transfer function indeed appears to correlate with the hot fraction function $F(k)$. Indeed, this slope holds steady in regions where the hot fraction function is relatively constant (with relatively little change of color), while this slope changes more rapidly in regions where the hot fraction function (and thus the corresponding color) is also changing rapidly. This is precisely the behavior anticipated in Sec. III. However, we now see that this behavior survives robustly even for dark-matter phase-space distributions which are fairly complicated, potentially even exhibiting many different peaks and troughs as functions of k .

D. An explicit test of our reconstruction conjecture

Finally, given the transfer functions in Fig. 17, we can now perform a test of our conjectured relation in Eq. (3.18). In particular, we can now determine the extent to which this relation allows us to reconstruct the dark-matter phase-space distribution $\tilde{g}(k)$ directly from the transfer function $T^2(k)$. Our results are shown in Fig. 18. In each panel of this figure, the black curve shows the transfer function $T^2(k)$ from Fig. 17, while the blue outline shows the original underlying k -space dark-matter phase-space distribution $\tilde{g}(k)$ from which it was derived, also from Fig. 17. By contrast, the pink shaded regions show the reconstructed phase-space distributions $\tilde{g}(k)$ that follow directly from the transfer function $T^2(k)$ via our conjectured “inverse” relationship in Eq. (3.18).

Although the reconstructed phase-space distributions do not match the original phase-space distributions exactly, they nevertheless do faithfully capture the most salient features of these distributions. Indeed, this is true not only for the unimodal distributions that emerge for $r \leq 0$ and $s \leq 0$, but also for the multimodal distributions that emerge otherwise. Likewise, this remains true even in cases for which our distributions are relatively extended, with no sharp peaks at all, such as those which emerge for $s < 0$. We thus conclude that our conjectured relationship holds remarkably well across a variety of possible dark-matter distribution shapes (thermal, nonthermal, unimodal, multimodal, and so forth). *A priori*, the relationship between the phase-space distribution $g(p)$ and the resulting power spectrum $P(k)$ is complex, involving rather nontrivial initial conditions incorporating primordial perturbations of the inflaton and other fields which are then propagated forward in time using the linearized Einstein equations in the presence of these noncold dark-matter species. Given this situation, we find this degree of agreement for what is essentially a relatively simple analytic “inversion” formula to be rather stunning.

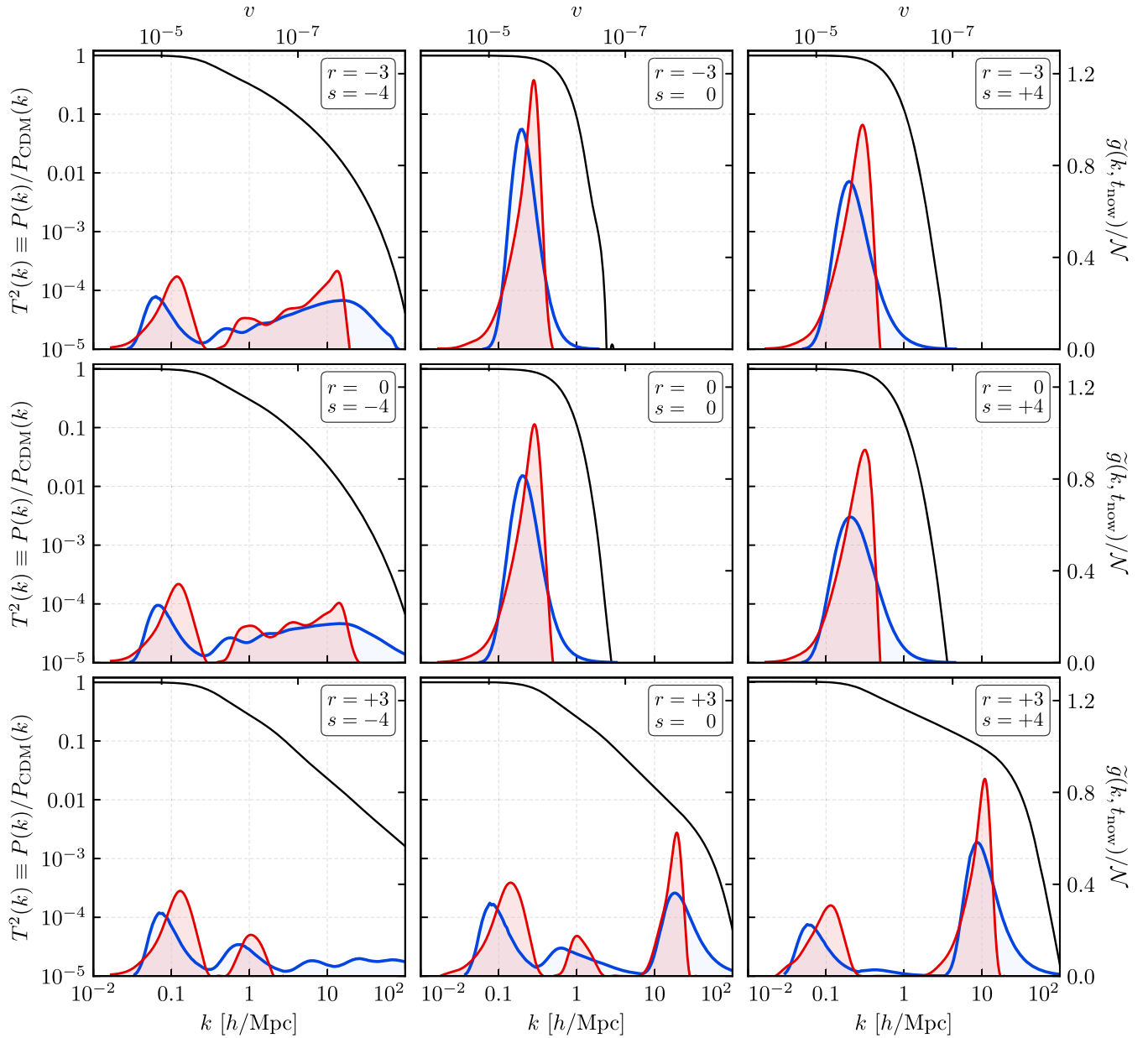


FIG. 18. An explicit test of our reconstruction conjecture in Eq. (3.18). In each panel, we show three curves: the original dark-matter phase-space distribution (blue), the corresponding transfer function (black) to which it gives rise, and the phase-space distribution (red/pink) which has been “reconstructed” from the transfer function via Eq. (3.18). We see that in all cases—including those with unimodal, bimodal, and even trimodal distributions, with peaks of assorted heights and widths—our conjecture reproduces many of the critical features of the original phase-space distribution with remarkable accuracy.

The results of this section thus illustrate that many features of the early-universe dynamics associated with the nonminimal dark sector of our illustrative model are indeed imprinted on the resulting matter power spectrum, and that an “archaeological” study of the matter power spectrum is indeed capable of reconstructing many aspects of this cosmological history. Indeed, we have seen that many features of the dark-matter phase-space distribution can be reconstructed from the transfer function via our conjecture in Eq. (3.18). Likewise, the specific pattern of

peaks and valleys within this reconstructed phase-space distribution reflects the integrated deposit history onto the dark-matter conveyor belt, and through the particular properties of the peaks involved we may even use results from Sec. II—such as those in Table II—in order to elucidate many aspects of the particular decay phenomenology involved. Thus, while no archaeological study can reconstruct every aspect of the cosmological past, we have seen there is much that we can indeed learn through these sorts of analyses.

V. CONCLUSIONS AND DISCUSSION

Throughout this paper, our goal has been to develop a toolbox of methods for learning about, and potentially constraining, the features of nonminimal dark sectors and their dynamics in the early universe. By necessity this quest has been fundamentally “archaeological”—we seek to exploit present-day data in order to learn about the past. In this paper our approach to this question has been centered around two somewhat independent linkages, as first outlined in Eq. (1.1): the first connects early-universe dynamics (such as that associated with a nonminimal dark sector) to a resulting dark-matter phase-space distribution $f(p)$, and the second connects this phase-space distribution to a corresponding matter-power spectrum $P(k)$. In principle, this latter quantity is observable—an assertion which we shall discuss further below—and thus our overarching goal has been to examine the extent to which we can “invert” the flow chart in Eq. (1.1) and thereby constrain not only the corresponding dark-matter distribution $f(p)$ but also the possible early-universe dynamics from which this $f(p)$ distribution might have arisen—all while starting from a given matter power spectrum $P(k)$.

Needless to say, a complete inversion is not possible, and we have seen numerous examples in this paper in which several different cosmological histories give rise to the same (or similar) present-day data. Yet, there have also been instances in which we can draw fairly robust conclusions regarding the inverse map with only a handful of reasonable assumptions. Collectively, such results may be of critical importance if the dark sector turns out to interact with the visible sector exceedingly weakly, or perhaps even only gravitationally. Studies focusing on inverting such quantities as the matter power spectrum may then be the only ways of ever learning about the dark sector and its early-universe dynamics.

In this paper, we have examined each of the linkages within Eq. (1.1) from a number of different directions, always with an eye towards assessing the extent to which the mappings they represent might be inverted. Along the way, we have generated what we believe to be a number of interesting and potentially useful results.

In Sec. II B, we began by studying the decay process from a given parent phase-space distribution (or “packet”) to a corresponding daughter packet. We found that this process is highly complex, as illustrated in Fig. 3, involving a mixture of effects due to time dilation, cosmological redshifting, and the exponential nature of the decay process itself. Under certain assumptions that are discussed in the text, we believe that our chief results in this direction include the following:

- (i) First, under certain assumptions, we found that the width and average momentum of a given daughter packet allow us to greatly constrain the properties of the parent from which it arose as well as the marginality of the corresponding decay process.

These results are shown in Table II. Indeed, we found that in some cases this reconstruction is unique, while in other cases several possibilities exist.

- (ii) Second, in some cases we were able to obtain *universal functional forms* which describe the resulting daughter packets. For example, in the limit that the parent is highly nonrelativistic, we found that the daughter packet always approaches the universal functional form shown in Eq. (2.28) and illustrated in Fig. 5. By contrast, for certain highly relativistic decays, we found that the daughter packet will instead have the very different universal functional form shown in Fig. 7.
- (iii) Third, by comparing the results of Tables I and II, we were able to explore the interplay between decay kinematics and cosmological expansion. This resulted in our observations, as outlined in Table III, concerning the manner in which individual decay deposits onto the daughter conveyor belt accrue as time evolves. These general observations allowed us to determine the extent to which cosmological expansion causes the daughter packets to be wider than would have been predicted on the basis of decay kinematics alone. In some cases, we found that cosmological expansion has little overall effect on the shape of the resulting daughter packet—indeed, in such cases a packet of similar shape would have emerged even if the universe had not been expanding. As indicated in Table III, these packets tend to have relatively sharp edges. In other cases, by contrast, we found that cosmological expansion plays a significant role in shaping the resulting daughter packet. In these cases the resulting packets tend not to have such sharp edges.
- (iv) Finally, we demonstrated that the overall tilt or *skewness* of the daughter packet also carries information concerning the parent from which the daughter emerged. For example, as anticipated in Eq. (2.20) and verified explicitly in Fig. 8, we found that there is a close relationship between the skewness of the daughter packet and the degree to which the parent was relativistic at the time it was produced. In Eq. (2.21) we even determined the conditions under which such relativistic skewness effects will be significant. Moreover, in a similar way, we were also able to demonstrate an entirely different skewness relationship: one between the skewness of the daughter packet and the *decay rate* Γ experienced by the parent [as expressed in units of the Hubble parameter $H(t_0)$ at the time t_0 when the parent was established]. These results are illustrated in Fig. 6. Together, these results imply that the skewness of the daughter packet can also serve as a useful tool in reconstructing the properties of the parent.

Next, in Sec. II C, we enlarged our discussion beyond the case of a single parent to examine the full set of decays that

can occur within a nonminimal dark sector containing many different constituents. Within this context, our primary observations are as follows:

- (v) We explicitly demonstrated that such a nonminimal dark sector can leave dramatic imprints on the phase-space distribution associated with the dark-sector ground state (presumed to be the dark matter that survives to the present time). We found that the resulting dark-matter phase-space distribution *need no longer be thermal*—in fact, it may even be *multimodal*, exhibiting a nontrivial pattern of peaks and troughs as a function of momentum. Such results therefore carry us beyond the sorts of thermal phase-space distributions that are traditionally assumed for dark-matter particles. In fact, we found that multimodal phase-space distributions emerge quite generically as the result of overlapping decay chains, even if we start from a single ancestor with a simple, unimodal phase-space distribution. This remains true even if our ancestor was produced thermally. *Thus we see that our final dark-matter phase-space distribution can be nonthermal (and potentially even multimodal)—even if the dark sector was initially populated through thermal means.*

In Sec. III, we turned to examine the second linkage in Eq. (1.1)—that connecting the dark-matter phase-space distribution $g(p)$ to the matter power spectrum $P(k)$. Given our results from Sec. II C, we allowed our analysis to remain completely general and we did not assume any particular form for the dark-matter phase-space distribution. In this context, we regard our main results to be the following:

- (vi) First, given an arbitrary phase-space distribution $g(p)$, we introduced a corresponding new quantity, a “dual” k -space distribution function $\tilde{g}(k)$. We view this as an important conceptual step since the p and k variables are not dual to each other in any “Fourier” sense. Rather, we utilized the defining relation for the horizon wave number $k_{\text{hor}}(p)$ in Eq. (3.7) in order to define a mapping between the p -variable of $g(p)$ and the k -variable of $P(k)$, implicitly identifying k_{hor} with k itself. We stress that this treatment is completely unorthodox, since $k_{\text{hor}}(p)$ is technically the *minimum* value of k for which $P(k)$ could potentially be affected by dark matter of momentum p . As such, there is therefore no direct relation between k_{hor} and the k -variable of $P(k)$. Nevertheless, by choosing to identify these two variables with each other, it becomes possible to consider the dark-matter phase-space distribution $g(p)$ in the same k -space as the matter power spectrum $P(k)$. This ultimately proved to be a critical step in reconstructing the former from the latter.
- (vii) Second, prompted by our phenomenological observations in Sec. III C, we also defined a so-called “hot fraction function” $F(k)$. Once again, this is a fairly nontrivial step because it too rests again on our

unorthodox mapping between the p -variable of $g(p)$ and the k -variable of $P(k)$. Indeed, the hot fraction function $F(k)$ is the total accumulated dark-matter abundance from all momenta p greater than $k_{\text{hor}}^{-1}(k)$.

- (viii) Finally, given these two new quantities $\tilde{g}(k)$ and $F(k)$, we were able to put forth what may prove to be one of the most important results of this paper: our “reconstruction” conjecture in Eq. (3.18). In this context, it is important to emphasize that we regard this conjecture as having two fairly independent components. The first component of our conjecture is the assertion that the hot fraction function $F(k)$ is directly correlated *not* with the transfer function $T^2(k) \equiv P(k)/P_{\text{CDM}}(k)$ itself, but rather with its logarithmic slope $d \log T^2 / d \log k$. This assertion is written explicitly in Eq. (3.15) in terms of an unknown correlation function η , and thus stands independently of any particular correlation function η . Given this, the second component of our conjecture asserts a particular empirical form for this correlation between $F(k)$ and the logarithmic slope $d \log T^2 / d \log k$. This correlation is given in Eq. (3.17). This result then implicitly furnishes us with an explicit form for the η -function described above, and thereby leads directly to our final conjecture in Eq. (3.18).

One remarkable feature of our conjecture is that it relates $P(k)$ to $\tilde{g}(k)$ *point by point* in k -space. Indeed, as long as we know the transfer function $T^2(k)$ and its derivatives at a specific value of k , our conjecture allows us to reconstruct $\tilde{g}(k)$ at that value of k . In this sense our reconstruction is *local*, mapping each portion of the matter power spectrum curve to a corresponding portion of the dark-matter phase-space distribution. This is a very useful feature because each portion of the $\tilde{g}(k)$ curve directly maps back to a corresponding portion of the original $g(p)$ phase-space distribution, and this in turn maps back to a sum of deposits along a particular “backwards FRW momentum lightcone” [here borrowing the language below Eq. (2.10)]. Thus we can trace our inverse map point by point along these curves, and thereby potentially correlate specific features in the matter power spectrum with specific deposit profiles stemming from early-universe dynamics.

In Sec. IV, we then proceeded to test all of these ideas within the context of an actual illustrative model. We began by specifying a particular Lagrangian governing early-universe dynamics and proceeded all the way to a calculation of the corresponding matter power-spectrum. As expected, this model was able to give rise to dark-matter phase-space distributions of great variety, including some which are unimodal, some which are bimodal or trimodal, and even some which have fairly nontrivial patterns of peaks and troughs—all arising as a consequence of the large number of dark-sector states and a corresponding multiplicity of decay pathways involving different characteristic timescales. These distributions are shown in Fig. 16.

Within this model we were also able to perform a rigorous test of our reconstruction conjecture in Eq. (3.18). In all of the cases explored, we obtained results for the reconstructed distributions $\tilde{g}(k)$ which, though not exact, successfully captured all of their salient phenomenological features. This is illustrated explicitly in Fig. 18. Likewise, in Appendix D, we traced the time development of this model *during* the intra-ensemble decay process and demonstrated how the total energy density and equation of state of the dark sector evolved nontrivially as the decay process unfolded. All of these results confirmed our main conclusions as highlighted above.

Finally, last but not least, we would be remiss not to mention one further unique feature of our work:

- (ix) Within this paper, we developed a very intuitive picture for viewing the cosmological time development of a given dark-matter phase-space distribution $g(p)$ in terms of a cosmological “conveyor belt” and deposits onto it. Of course, as we discussed, this picture is nothing but a physical representation of the integral form of the standard Boltzmann equations, and as such it is not necessarily “new.” Nevertheless, throughout this paper, we repeatedly found this physical picture to be a particularly useful one, especially for nonminimal dark sectors and the decays that can arise amongst their constituents. Indeed, this intuitive picture of deposits onto a cosmological conveyor belt has served as the backbone of much of the discussion in this paper.

Needless to say, this paper has spanned considerable territory. As such, a number of comments are in order, concerning not only the work we have done but also possible directions for future investigation.

First and foremost, our ability to perform the kind of archaeological reconstruction we have described in this paper is predicated on our ability to measure $P(k)$ at large k . Understanding the extent to which we are currently able to probe the matter power spectrum is therefore critical for our work. It is also important to assess the future prospects for probing $P(k)$ at even higher k .

Measurements of the linear matter power spectrum based on data obtained at low redshifts are currently reliable for $k \lesssim 0.05\text{--}0.1 \text{ Mpc}^{-1}$. However, observations of the Lyman- α forest at higher redshifts up to $z \sim 5$ can provide additional information about the linear matter power spectrum at wave numbers up to $k \sim 1 \text{ Mpc}^{-1}$. At higher z , the density of neutral hydrogen is so large that the hydrogen emission spectrum becomes difficult to measure. However, observations of the 21-cm spectral line of neutral hydrogen at redshifts up to $z \sim 30$ could potentially yield information about the linear matter power spectrum at much higher k . There are of course certain practical considerations which would render such a measurement challenging. For example, any instrument capable of attaining the requisite sensitivity would need to have an enormous collection

area, and ionospheric effects on radio signals with frequencies below 100 MHz would likely render a terrestrial telescope unsuitable for this purpose. Such a measurement is nevertheless possible in principle.

Another strategy for probing the dynamics of a nonminimal dark sector during the early universe would be to perform a reconstruction of the dark-matter phase-space distribution similar to the one we have discussed in this paper, but based on the *nonlinear* matter power spectrum. For example, information about the clustering of matter on very small scales can be obtained when light emitted by quasars and other distant astrophysical objects is gravitationally lensed by small foreground objects with masses down to $\sim 10^6 M_\odot$, producing arcs and other similar observable features.

The challenges involved in extracting meaningful information concerning the underlying dark sector from the nonlinear matter power spectrum are not only observational, but also theoretical and computational. A detailed analysis of how structure evolves in the nonlinear regime generally requires computationally expensive N -body simulations such as GADGET-2 [12]. That said, less computationally intensive tools and approximation methods exist which can provide insight into how dark-matter velocity distributions affect the power spectrum on small scales. Approaches along these lines include the application of fitting procedures such as HALOFIT [13] and approximations concerning the collapse of dark-matter halos [14]. While these approaches have been applied to warm-dark-matter scenarios [15], any results obtained from the highly nonthermal and multimodal distributions we have discussed here would likely be considerably different.

Despite the complications involved in extending our analysis to the nonlinear regime, there are compelling motivations for studies along these lines. It has been suggested [16] that single-component dark-matter scenarios with a thermal-like distribution are not capable of successfully addressing small-scale structure anomalies such as the “too-big-to-fail” problem [17] while at the same time satisfying Lyman- α constraints. This incompatibility is ultimately due to the steepness of the suppression in the power spectra associated with thermal-like distributions. Likewise, standard warm-dark-matter models have been proposed as a way to reproduce the observed density profiles of dark-matter halos, but suffer from the so-called “Catch-22” problem [18]. By contrast, the highly nonthermal phase-space distributions we have studied here may be able to address these issues with more success. Indeed, the linear power spectra which arise from such distributions often do not fall as steeply with k , primarily because $\tilde{g}(k)$ can span a broader range of scales. Moreover, even in the simplest such scenarios—e.g., dark matter produced through the decay of a single unstable species—constraints are easier to satisfy and structure on small scales differs appreciably from the predictions of thermal dark-matter models [1,19].

We emphasize that small-scale structure in the theories we have studied here is also likely to differ in other ways from that expected in theories of thermal warm dark matter. For example, in thermal warm-dark-matter scenarios, a minimum mass $M_{\min} \propto k_{\text{FSH}}^{-3}$ (ultimately dictated by the free-streaming horizon k_{FSH}) exists below which dark-matter halos do not form. Such a cutoff in the associated halo-mass function is corroborated by N -body simulations [20]. By contrast, as we have seen, the free-streaming horizon is generally not a reliable indicator of the dark-matter phase-space distribution for nonminimal dark sectors.

As our conjecture makes clear, some features of nonminimal dark sectors may be more easily reconstructed than others. As noted above, our conjecture relates each point in the matter power spectrum to a corresponding point in the dark-matter phase-space distribution. Thus, our conjecture gives us information about the dark-matter phase-space distribution only for the specific momentum scales which correspond, via Eq. (3.7), to the wave number scales at which we can observe the matter power spectrum. If, as discussed above, one can eventually extend the linear regime towards even higher values of k , our conjecture may then allow us to reconstruct portions of the dark-matter phase-space distribution at even lower velocity. Further details might also be accessible depending on the precision with which such observations can be made and perhaps even the extent to which our theoretical conjecture might be refined. This provides additional motivation for extending the reach of the linear regime.

There are many possible dark-matter scenarios for which our reconstruction techniques could be useful. Of course, in this paper we have demonstrated the utility of these techniques within the context of an explicit model presented in Sec. IV. This model also demonstrated how competing decay chains in particular yield highly non-thermal dark-matter momentum distributions. As already noted, this model was chosen for its illustrative power and was not meant to describe a UV-complete description of the dark sector. However, models with many of the same qualitative features emerge naturally within a number of UV-complete scenarios for new physics.

For example, such features generically arise in theories involving extra spacetime dimensions. Indeed, in scenarios where the SM is localized on a brane within the higher-dimensional bulk, the Kaluza-Klein (KK) excitations of bulk fields are necessarily neutral under the SM gauge group. From a four-dimensional perspective, such excitations therefore manifest themselves as towers of “dark” particles. The self-interactions of bulk fields can therefore give rise to intra-ensemble decays among the corresponding KK modes, as can interactions between bulk fields and fields on the brane [21–23]. In either case, the decay properties of the KK modes are determined both by the couplings involved and by the geometry of the extra dimensions. For example, in scenarios involving

warped extra dimensions, the pattern of decays—and therefore the shape of the resulting dark-matter phase-space distributions—depends sensitively on the degree of warping [24].

Clockwork scenarios [25–27] also involve large numbers of dark particles. The structure of the dark sector in clockwork models effectively consists of a single light weakly coupled state, along with a compressed mass spectrum of heavier states. In a manner similar to dimensional deconstruction [28], these models can be constructed as discretizations of extra-dimensional theories which are similar to the linear-dilaton model [27,29–32], an approximate holographic dual to little string theory [33]. The interactions among the fields which drive the clockwork mechanism can naturally involve intra-ensemble decays among these states. The particular spectrum of masses and couplings in clockwork models implies a pattern of intra-ensemble decays that would differ from those in the other scenarios described above.

There also exist other frameworks in which the dark sector generally comprises large numbers of dark particles. For example, such dark particles may emerge as the “hadrons” associated with the confining phase of a strongly coupled dark sector [24,34]. Likewise, the dark particles may also arise naturally as the gauge-neutral bulk states of type I string theories [34]. Such frameworks are likely to have very different phenomenologies from those discussed above, since the density of dark states in such theories grows *exponentially* with mass and the states themselves lie along linear Regge trajectories. It would be interesting to explore the phenomenology of the intra-ensemble decays and resulting dark-matter phase-space distributions that might emerge in such scenarios.

There also exist many possible generalizations and extensions of our work. One aspect of our work that may be broadly extended is the production mechanism for our dark sector. To a large extent, the production mechanism is dependent on how our ensemble couples to external fields. In this work, we have generally assumed that our ensemble has only gravitational couplings to the visible sector. Additionally, within our example model in Sec. IV, we assumed that our ensemble has “top-heavy” initial conditions—i.e., conditions in which the heavier states acquire larger initial energy densities.

Although the latter assumption was not critical for our main results, there exist many production mechanisms that yield such initial conditions. Perhaps the most obvious of these is gravitational production. For example, dark particles can be generated due to the changing spacetime metric at the end of inflation [35]. Indeed, this process is often invoked as a potential abundance-generation mechanism for ultraheavy dark-matter candidates [36,37]. Furthermore, as shown in Refs. [38–45], oscillations of background quantities induced by inflaton dynamics can also contribute to particle production in the early universe.

This also constitutes an abundance-generation mechanism for dark matter. In general, depending on the details of the inflationary model and how the dark-sector particles interact with one another, gravitational production can make substantial contributions to the cosmological abundances of particles with a potentially broad range of mass scales. Moreover, in many cases the energy density bestowed upon a given dark-sector species in such scenarios generally increases as a function of its mass. Thus, we see that gravitational production often furnishes a “top-heavy” distribution of energy densities. Indeed, gravitational production is particularly appealing within the context of this paper because it is not predicated on the presence of any nongravitational interactions between the fields of the dark and visible sectors.

Beyond gravitational production, our dark sector can also be populated through interactions with the inflaton field. For example, the initial population could simply be produced through the perturbative decay of the inflaton as part of the reheating process. As with gravitational production, this production mechanism can also yield a “top-heavy” distribution of energy densities, provided that the branching ratios for inflaton decays are larger for the more massive states in our ensemble. Furthermore, the distribution of dark-sector energy densities can also be sensitive to the shape of the inflaton potential during reheating, particularly if the curvature of this potential changes significantly over the range of field values through which the inflaton oscillates [46].

Interactions with the inflaton can also populate our ensemble through nonperturbative processes. In particular, once the inflaton begins oscillating coherently, parametric instabilities develop which lead to the explosive production of bosonic dark-sector particles within certain momentum bands [47–50]. The resulting dark-sector energy densities and momentum distributions depend on a variety of factors, including their couplings to the inflaton and the shape of the inflaton potential. Interactions among the ensemble constituents can have an impact on the outcome as well, since these affect the modulation of the effective masses of the dark-sector fields, a possibility we shall discuss in more detail below. Additionally, we note that the process of thermalization is generally complicated in these scenarios—with phases of nonlinear dynamics, turbulence, etc.—and these could play an important role in establishing the relative initial abundances of the different dark-sector states.

Additional production mechanisms are also available if the ensemble couples to other external states. For example, production can occur through the freeze-in mechanism [51–54], which is of particular interest in this regard because the couplings involved need not be large. The freeze-in production of our ensemble can occur through interactions with either visible-sector particles, or particles in some other separate dark sector. Similarly, depending on the strength of our intra-ensemble couplings, the freeze-in

and freeze-out mechanisms can also occur through intra-ensemble interactions. In particular, these processes could play an important role in the thermalization of our phase-space distributions at early times.

Another possible extension of our work relates to visible-sector couplings. In this paper we have assumed for simplicity that our dark sector is essentially decoupled from the visible sector. We therefore assumed that the decays of our dark-sector states do not produce SM particles. While this is certainly a viable possibility, it would also be interesting to examine the consequences of relaxing this assumption. For example, as discussed above, the presence of nongravitational interactions between these sectors opens up new possible dark-matter production mechanisms. Moreover, such interactions generically give rise to scattering processes involving both the dark and visible sectors. These processes, which bring the dark-sector states towards kinetic equilibrium, could potentially distort the phase-space distributions of these states and thereby wash out the associated imprints in the matter power spectrum. However, the presence of these interactions also opens up the possibility that the dark sector could be probed through other, complementary means. For example, these interactions can introduce dark-matter decay channels involving SM particles in the final state. Not only do such decay channels transfer energy from the dark to the visible sector, but they can also have other observational consequences. Thus, it might be possible to simultaneously observe both a signal of dark-sector dynamics within the matter-power spectrum and a complementary signal of interactions between the dark and visible sectors. It would be interesting to investigate how large these couplings could be without erasing information which would otherwise have been imprinted within the matter power spectrum.

Another assumption we have made that can potentially be relaxed concerns the lifetimes of our dark-sector states. In particular, we have focused on the regime in which the lifetimes of all unstable dark-sector states are sufficiently short that all of their energy density has been transferred to the lightest species by the beginning of the BBN epoch. We have made this assumption since the decays of unstable particles at subsequent times—even decays solely to other, lighter states within the dark sector—are constrained by their impact of the cosmic expansion rate and its effect on various observables [55–58]. These observables include the spatial distribution of CMB anisotropies, baryon-acoustic-oscillation data, and the relationship between the redshifts and luminosity distances of Type Ia supernovae. Moreover, if these particles can decay into visible-sector states, even more stringent bounds apply. To allow for dark-matter decays during or after BBN, one must ensure that observational limits on the production of such states are satisfied. The corresponding bounds on the lifetime and abundance of a single decaying particle species were derived in Refs. [59–70]. Likewise, the constraints on decays to electromagnetically interacting final states for an entire

ensemble of decaying particles have been formulated in a model-independent way in Ref. [71].

Of course, if one or more of these unstable dark-sector states is extremely long-lived, with $\tau_\ell \gtrsim t_{\text{now}}$, these states contribute to the present-day dark-matter abundance. Such dark-matter scenarios then fall within the purview of the Dynamical Dark Matter framework [21–23]. The phase-space distributions associated with all of these dark-matter components will generically have nonthermal profiles as a consequence of intra-ensemble decays. As a result, any such component which retains a non-negligible energy density into the matter-dominated epoch contributes to the growth (or suppression) of matter perturbations, rendering the process of structure formation far more complicated. Furthermore, these dark-matter components presumably continue to decay throughout the epoch during which these perturbations become nonlinear, thereby altering the dynamics of dark-matter halo formation and modifying the resulting spatial mass distribution of the halos.

Another assumption that may be relaxed concerns the effects of Bose enhancement or Pauli blocking. Throughout this paper, for simplicity we focused exclusively on the $f_\ell(p, t) \ll 1$ regime for each of our dark-sector species. However, there are several reasons to extend our analysis by relaxing this assumption. For example, in scenarios in which the dark-sector states are scalar fields—e.g., moduli or axion-like particles—these fields can acquire vacuum expectation values (VEVs) which are misaligned from the minimum of the scalar potential. Indeed, such a misaligned VEV is associated with a zero-momentum condensate, severely violating the condition $f_\ell(p, t) \ll 1$ [72]. A complete description of the scalar ensemble then necessarily involves this condensate, in addition to any contributions to $f_\ell(p, t)$ from other production mechanisms.

There are several considerations that can impact the dynamics of the ensemble when the criterion $f_\ell(p, t) \ll 1$ does not hold and quantum effects come into play. One of these is that the contributions to the collision operator $C[f]$ in Eq. (A2) associated with scattering, decay, and inverse-decay processes all involve factors of $(1 + f_\ell)$ —factors which cannot be approximated as unity in the presence of a condensate. The impact of these effects on the phenomenology of scalar fields has been investigated, for example, in the context of asymmetric reheating [73,74]. Other considerations associated with the VEVs of dark-sector fields can alter the phenomenology in significant ways. For example, within the context of the illustrative model presented in Sec. IV, the presence of a set of VEVs $\langle \phi_\ell \rangle$ for the ϕ_ℓ implies that each dark-sector species acquires a field-dependent contribution to its mass as a consequence of the trilinear coupling in Eq. (4.2). In particular, the effective mass of each such species in this case is given by $(m_\ell^2)_{\text{eff}} = m_\ell^2 + 2 \sum_{i=0}^N c_{i\ell} \langle \phi_i \rangle$.

While this is interesting in and of itself, even more interesting is the possibility that the $\langle \phi_\ell \rangle$ —and therefore

the effective mass of each dark-sector species—could be time-dependent. This occurs naturally if any of the $\langle \phi_\ell \rangle$ are displaced from the minimum of the scalar potential at early times. Then, once the Hubble parameter falls to $H \sim m_\ell$, the corresponding field VEV will begin to oscillate coherently. As a result, parametric instabilities can develop—instabilities which lead to an enhanced production of scalars within particular momentum bands. Alternatively, a time dependence for the $(m_\ell^2)_{\text{eff}}$ can also arise directly as a consequence of dynamical processes which do not involve modulation of the VEVs of other dark-sector fields—e.g., from a cosmological phase transition. Even considering only the effect on the VEVs, multiple fields receiving such dynamical mass contributions exhibit an array of possible behaviors [75–77]. Furthermore, when the $(m_\ell^2)_{\text{eff}}$ evolve in this way, the pattern of intra-ensemble decays could be drastically altered, thereby modifying the ultimate shape of the resulting dark-matter phase-space distributions.

In this paper we have shown that multimodality of the dark-matter phase-space distribution can emerge quite generically within a multicomponent dark sector when there are overlapping decay chains with different decay rates. However, strictly speaking, one does not require independent decay chains in order to produce multimodality. Indeed, even a single decay pathway can produce a multimodal distribution for the daughter if the parent itself experiences production while the decay is proceeding. For example, a grandparent might decay and thereby replenish the parent after the parent has already decayed. Alternatively, the parent might be in thermal contact with an external source (such as might occur if the parent is experiencing thermal freezeout), and thus its abundance might be continually replenished as it decays. In fact, a model exhibiting the latter phenomenon already exists within the context of sterile-neutrino dark matter [2,78].

In general, as we have discussed in Sec. II, a multimodal dark-matter phase-space distribution will emerge whenever there are widely separated dark-matter “deposits” onto the cosmological conveyor belt. Of course, such deposits need not all be the results of decays from more massive states—any sequence of production episodes separated in time and/or momentum can realize the same end result. An example of such a phenomenon can be found in Ref. [79].

One of the most important results of our paper is our reconstruction conjecture in Eq. (3.18). As emphasized above, we regard this conjecture as having two distinct components: the first is the assertion that the hot fraction function $F(k)$ is connected to the *slope* of the transfer function $d \log T^2 / d \log k$, as indicated in Eq. (3.15), and the second is the assertion of a particular function η which describes this connection, as indicated implicitly through Eq. (3.17). Indeed, these two assertions together yield our final conjecture in Eq. (3.18). Although our conjecture is remarkably successful in reproducing the salient features of

$\tilde{g}(k)$, we regard this conjecture as at best purely empirical. It is therefore possible that one or both aspects of this conjecture might be further refined. For example, it is possible that the hot fraction function $F(k)$ might also carry a weak dependence on other (higher) derivatives of the transfer function, or on the value of the transfer function itself. Likewise, even with the assumption given in Eq. (3.15), it is possible that the η -function might carry higher-order corrections beyond those in Eq. (3.18). Although it is not possible to rigorously invert the mathematical procedure through which a given dark-matter phase-space distribution $g(p)$ produces a corresponding matter power spectrum $P(k)$, it may be possible to trace through such a calculation algebraically to leading order in order to learn which features of $g(p)$ might dominate the resulting $P(k)$, and vice versa. In this way, one might hope to eventually derive our conjecture analytically, along with possible correction terms.

Finally, it is interesting to consider how machine-learning techniques could be applied to the archaeological inverse problem of deciphering the properties of an underlying dark sector from the matter power spectrum. Indeed, there has recently been considerable interest in how machine-learning techniques, such as the implementation of neural networks, can be applied to various aspects of early-universe cosmology. For example, emulators trained on Einstein-Boltzmann solvers have been used to generate estimates for observables such as the linear matter power spectrum and the CMB directly from either standard cosmological parameters [80–84] or the parameters associated with specific models [85]. Neural networks have also been used to eliminate computational bottlenecks involving the most time-intensive or least-parallelizable steps in the calculations performed by these solvers [86]. There are several ways in which machine learning might likewise be applied to the work in this paper. For example, one could potentially employ these techniques in order to learn solutions to the Boltzmann evolution of our dark-matter phase-space distributions or the resulting cosmological perturbations.

As is abundantly clear from this discussion, the work we have presented here represents but a first foray in the general direction of the archaeological reconstruction of the dark sector based on the matter power spectrum. As such, many avenues remain open for future research. Some constitute potential refinements or generalizations of the work we have presented here, while others extend our results in a number of new directions. As further observational data accumulates concerning the properties of the matter power spectrum, many different ideas along these lines will be needed in order to exploit this data in pursuit of our overall archaeological goals. We therefore hope that our results can play a significant role in this endeavor.

ACKNOWLEDGMENTS

We would like to thank K. Boddy, D. Curtin, P. Draper, A. Erickcek, L. Leskovec, S. McDermott, C. Miller, Y. Park, M. Peloso, E. Rozo, G. Shiu, T. Tenkanen, K. Zurek, and especially J. Shelton and S. White for discussions. The research activities of K. R. D., F. H., and S. S. were supported in part by the Department of Energy under Grant DE-FG02-13ER41976 / DE-SC0009913, while those of J. K. were supported in part by the IBS under project code IBS-R018-D1 and those of B. T. were supported in part by the National Science Foundation under Grant PHY-1720430. The research activities of K. R. D. were also supported in part by the National Science Foundation through its employee IR/D program. This work was completed in part at the Aspen Center for Physics, which is supported by the National Science Foundation under Grant PHY-1607611. K. R. D. and B. T. would also like to acknowledge the hospitality of the Institute for Basic Science in Daejeon, South Korea, as well as the hospitality of the Institut Pascal at the Université Paris-Saclay, which is supported by the IPHT, the P2I and SPU research departments, and the P2IO Laboratory of Excellence (Programs ANR-11-IDEX-0003-01 Paris-Saclay and ANR-10-LABX-0038). The opinions and conclusions expressed herein are those of the authors, and do not represent any funding agencies.

APPENDIX A: BOLTZMANN EQUATIONS

In this appendix, we describe the Boltzmann equations which govern the time evolution of the phase-space distributions $f_\ell(p_\ell, t)$ for the dark-sector fields ϕ_ℓ in the model introduced in Sec. IV. In general, the Boltzmann equation which governs the evolution of $f_\ell(p_\ell, t)$ for a given dark-sector species ϕ_ℓ may be written in the form

$$\frac{\partial f_\ell(p_\ell, t)}{\partial t} = H(t)p_\ell \frac{\partial f_\ell(p_\ell, t)}{\partial p_\ell} + \frac{C[f]}{E_\ell}, \quad (\text{A1})$$

where $H(t)$ is the Hubble parameter, where $E_\ell = \sqrt{p_\ell^2 + m_\ell^2}$ is the energy of a particle of this species with momentum p_ℓ , and where $C[f]$ is the collision operator. The collision operator for our model can be written as a sum of three terms

$$C[f] = C_D^{(-)}[f] + C_D^{(+)}[f] + C_S[f]. \quad (\text{A2})$$

The first of these terms represents the contribution from decay processes of the form $\phi_\ell \rightarrow \phi_i\phi_j$, which serve as a sink for ϕ_ℓ , along with the corresponding inverse-decay processes. The second represents the contribution from decay processes of the form $\phi_i \rightarrow \phi_j\phi_\ell$, which serve as a source for ϕ_ℓ , along with the corresponding inverse-decay processes. The third term represents the contribution from $2 \rightarrow 2$ scattering processes which involve one or more particles of species ϕ_ℓ in the initial state, the final state, or both.

The first term on the right side of Eq. (A2) is given by

$$C_D^{(-)}[f] = -\sum_{i,j} \int d\Pi_i d\Pi_j \left\{ |\mathcal{M}_{ij}^{\ell}|^2 \right. \\ \times (2\pi)^4 \delta^4(p_\ell - p_i - p_j) \\ \left. \times \left[f_\ell(1+f_i)(1+f_j) - f_i f_j(1+f_\ell) \right] \right\}, \quad (\text{A3})$$

where the phase-space measure $d\Pi_i$ for the real scalar ϕ_i is given by

$$d\Pi_i = \frac{1}{(2\pi)^3} \frac{d^3 p_i}{2E_i}. \quad (\text{A4})$$

The squared matrix element $|\mathcal{M}_{ij}^{\ell}|^2$ for the decay process $\phi_\ell \rightarrow \phi_i \phi_j$ in our model is simply

$$|\mathcal{M}_{ij}^{\ell}|^2 = \frac{1}{\mathcal{N}_{ij}} c_{\ell ij}^2, \quad (\text{A5})$$

where $\mathcal{N}_{ij} = 1 + \delta_{ij}$ is the multiplicity of ϕ_i in the final state of the decay process. We note that $|\mathcal{M}_{ij}^{\ell}|^2$ is defined here such that it incorporates the symmetry factor which arises for combinations of the indices i and j for which multiple identical particles appear in the final state of the decay process (or in the initial state of the inverse-decay process). Likewise, the second term on the right side of Eq. (A2) is given by

$$C_D^{(+)}[f] = -\sum_{i,j} \mathcal{N}_{\ell j} \int d\Pi_i d\Pi_j \left\{ |\mathcal{M}_{ij}^i|^2 \right. \\ \times (2\pi)^4 \delta^4(p_i - p_j - p_\ell) \\ \left. \times \left[f_i(1+f_j)(1+f_\ell) - f_j f_\ell(1+f_i) \right] \right\}. \quad (\text{A6})$$

The squared matrix element in this case is

$$|\mathcal{M}_{ij}^i|^2 = \frac{1}{\mathcal{N}_{\ell j}} c_{i\ell j}^2, \quad (\text{A7})$$

where once again the squared matrix element has been defined such that it incorporates the relevant symmetry factor. Finally, the third term on the right side of Eq. (A2) is given by

$$C_S[f] = -\sum_{i,j,k} \mathcal{N}_{\ell i} \int d\Pi_i d\Pi_j d\Pi_k \left\{ |\mathcal{M}_{jk}^{\ell i}|^2 \right. \\ \times (2\pi)^4 \delta^4(p_\ell + p_i - p_j - p_k) \\ \left. \times \left[f_\ell f_i(1+f_j)(1+f_k) - f_j f_k(1+f_i)(1+f_\ell) \right] \right\}. \quad (\text{A8})$$

Provided that the coupling coefficients associated with any quartic terms in the interaction Lagrangian for the ϕ_ℓ are sufficiently small that they can be safely neglected, the squared matrix element for the scattering process $\phi_\ell \phi_i \rightarrow \phi_j \phi_k$ is

$$|\mathcal{M}_{jk}^{\ell i}|^2 = \frac{1}{\mathcal{N}_{\ell i} \mathcal{N}_{jk}} \left| \sum_n \left[\frac{c_{\ell i n} c_{n j k}}{[(p_\ell + p_i)^2 - m_n^2 + im_n \Gamma_n]} \right. \right. \\ \left. \left. + \frac{c_{\ell j n} c_{n i k}}{(p_\ell - p_j)^2 - m_n^2 + im_n \Gamma_n} \right. \right. \\ \left. \left. + \frac{c_{\ell k n} c_{n i j}}{(p_\ell - p_k)^2 - m_n^2 + im_n \Gamma_n} \right] \right|^2, \quad (\text{A9})$$

where once again the squared matrix element has been defined such that it incorporates the relevant symmetry factor.

The expressions in Eqs. (A3), (A6), and (A8) are completely general and applicable across the full parameter space of our model. However, within our region of interest within this parameter space, two conditions are satisfied which enable us to simplify these equations considerably. The first of these conditions, which we explicitly enforce throughout our numerical study, is that $f_\ell(p_\ell, t) \ll 1$ for all dark-sector species ϕ_ℓ at all times $t \geq t_I$. Within this regime, all Bose-enhancement factors can be neglected in Eqs. (A3), (A6), and (A8). These individual contributions to the collision operator therefore reduce to

$$C_D^{(-)}[f] \approx -\sum_{i,j} \int \frac{d^3 p_i d^3 p_j}{4(2\pi)^2 E_i E_j} [|\mathcal{M}_{ij}^{\ell}|^2 \\ \times \delta^4(p_\ell - p_i - p_j)(f_\ell - f_i f_j)] \\ C_D^{(+)}[f] \approx -\sum_{i,j} \mathcal{N}_{j\ell} \int \frac{d^3 p_i d^3 p_j}{4(2\pi)^2 E_i E_j} [|\mathcal{M}_{j\ell}^i|^2 \\ \times \delta^4(p_i - p_j - p_\ell)(f_i - f_j f_\ell)] \\ C_S[f] \approx -\sum_{i,j,k} \mathcal{N}_{i\ell} \int \frac{d^3 p_i d^3 p_j d^3 p_k}{8(2\pi)^5 E_i E_j E_k} [|\mathcal{M}_{jk}^{\ell i}|^2 \\ \times \delta^4(p_\ell + p_i - p_j - p_k)(f_\ell f_i - f_j f_k)]. \quad (\text{A10})$$

The second condition which is satisfied within our parameter-space region of interest is that the overall scale of the couplings among the dark-sector fields ϕ_ℓ —a scale set by the value of the parameter μ —be sufficiently small that the terms in $C[f]$ associated with scattering and inverse-decay processes have a negligible effect on the evolution of the phase-space distributions $f_\ell(p_\ell, t)$ for these fields. In the regime in which these terms can be neglected, the expressions in Eq. (A10) reduce to

$$\begin{aligned}
C_D^{(-)}[f] &\approx -\sum_{i,j} \int \frac{d^3 p_i d^3 p_j}{4(2\pi)^2 E_i E_j} [|\mathcal{M}_{ij}^\ell|^2 \\
&\quad \times \delta^4(p_\ell - p_i - p_j) f_\ell], \\
C_D^{(+)}[f] &\approx -\sum_{i,j} \mathcal{N}_{j\ell} \int \frac{d^3 p_i d^3 p_j}{4(2\pi)^2 E_i E_j} [|\mathcal{M}_{j\ell}^i|^2 \\
&\quad \times \delta^4(p_i - p_j - p_\ell) f_i], \\
C_S[f] &\approx 0.
\end{aligned} \tag{A11}$$

Indeed, we have verified numerically that for the parameter choices specified in Sec. IV, the interaction rates for these processes are all much smaller than $H(t)$ for all $t \geq t_I$.

One useful property of the expressions in Eq. (A11) is that they are linear in the phase-space densities $f_\ell(p_\ell, t)$. It therefore follows that the Boltzmann equations for the ϕ_ℓ are linear in the $f_\ell(p_\ell, t)$ within our parameter-space region of interest as well. This in turn implies that the effect of rescaling the initial normalization $f_9(p_9, t_I)$ by an overall constant factor is simply to rescale the normalizations of all of the $f_\ell(p_\ell, t)$ at all subsequent times by the same factor—provided, of course, that the condition $f_\ell(p_\ell, t) \ll 1$ continues to be satisfied for all ϕ_ℓ at all times $t \geq t_I$.

Another useful property of the Boltzmann equations which arises within our parameter-space region of interest is an invariance under a certain class of transformations involving an arbitrary dimensionless scaling parameter α . In particular, for any $\alpha > 0$, it can be shown that the Boltzmann equations are invariant under the algebraic replacements

$$\begin{cases} m_0 \rightarrow \alpha m_0 \\ \Delta m \rightarrow \alpha \Delta m \\ \mu \rightarrow \alpha \mu \\ f_\ell(p_\ell, t) \rightarrow f_\ell(p_\ell/\alpha, \alpha t) \\ t_I \rightarrow t_I/\alpha \end{cases} \tag{A12}$$

for all of the ϕ_ℓ . Physically speaking, the transformation listed in Eq. (A12) for the phase-space distribution $f_\ell(p_\ell, t)$ represents a uniform shift such that the occupation density in phase space at momentum p_ℓ becomes the occupation density at momentum αp_ℓ .

The invariance of the Boltzmann equations under the transformations in Eq. (A12) has important implications for the phenomenology of our model. In particular, the phase-space distribution $f_\ell(p_\ell, t)$ for any of our dark-sector fields ϕ_ℓ , expressed as a function of the dimensionless ratio p_ℓ/m_0 , is identical to the phase-space distribution obtained for any other choice of model parameters for which the values of the dimensionless quantities $\Delta m/m_0$, μ/m_0 , and $m_0 t_I$ are the same.

APPENDIX B: DECAY FROM PARENT TO DAUGHTER: AN EXPLICIT NUMERICAL EXAMPLE

In Table II, we described certain general properties of the daughter packets which result from relatively narrow parent packets undergoing two-body decays into identical daughters. In this appendix, we provide an explicit set of numerical examples which may further elucidate these general results.

These examples are shown in Table IV. In particular, the data in this table was generated through a full numerical Boltzmann analysis including all relativistic, redshifting, and exponential-decay effects. For this analysis, we have assumed $\Gamma/H(t_0) = 10^{-3}$ where t_0 is the time at which the parent packet is produced, and we have taken $\kappa = 3/2$. Within this table, m_P and m_D are respectively the parent and daughter masses; $\langle p_{\text{prod}} \rangle$ is the average momentum of the parent packet at production; and $\langle p_{\text{decay}} \rangle$ is the average momentum of the parent packet when the proper time elapsed since production reaches $\tau \equiv \Gamma^{-1}$. Likewise p_D^{rest} is the momentum of the daughter in the rest frame of the parent, a quantity which depends on the masses alone and which is given in this case by $p_D^{\text{rest}} = \sqrt{(m_P/2)^2 - m_D^2}$. Each different case within Table IV can be specified by choosing values of m_P , m_D , and either p_{prod} or p_{decay} (we shall select p_{decay} for this purpose); the remaining entries in this table are then calculated accordingly.

Note that we have chosen the different cases within Table IV in such a way as to correspond to the different cases in Table II. These different cases are therefore listed in Table IV using the same labeling scheme as in Table II, with K1 and K2 in Table IV corresponding to the two possibilities for Case K within Table II with absolute marginalities which are either $\mathcal{O}(1)$ or far, respectively. In this connection, recall that the *absolute* marginality in Table II is determined by the ratio p_D^{rest}/m_P , while the *relative* marginality in Table II is determined by the ratio $p_D^{\text{rest}}/p_{\text{decay}}$. The different cases shown in Table IV therefore correspond to situations in which $p_D^{\text{rest}} \ll p_{\text{decay}}$, $p_D^{\text{rest}} \sim p_{\text{decay}}$, or $p_D^{\text{rest}} \gg p_{\text{decay}}$. Likewise, for each choice, we have then considered only those choices for which $p_{\text{decay}} \ll m_P$, $p_{\text{decay}} \sim m_P$, or $p_{\text{decay}} \gg m_P$. As a result, Cases A, B, D, and H are near absolute marginality (with $p_D^{\text{rest}}/m_P \ll 1$), while Cases F, G, I, J, and K2 are far from absolute marginality (with p_D^{rest}/m_P approaching its maximum kinematically allowed value, which in this case is $1/2$). Likewise, Cases D, H, K1, and K2 are near relative marginality (with $p_D^{\text{rest}} \ll p_{\text{decay}}$), while Cases A, C, F, G, and I are far from relative marginality (with $p_D^{\text{rest}} \gg p_{\text{decay}}$). Indeed, Case F is the farthest from relative marginality, as already anticipated below Table II.

It is clear from Table IV that our numerical results conform quite well to our general expectations in Table II. In making this assessment, we note that the relative

TABLE IV. An explicit numerical example of the decay of a parent packet into a daughter packet, shown for a variety of cases corresponding to the cases in Table II. In obtaining this data we have assumed a parent of mass m_P undergoing a two-body decay into identical daughters of mass m_D , and we have taken $\Gamma/H(t_0) = 10^{-3}$ and $\kappa = 3/2$, where t_0 is the time at which the parent is originally produced. The data in this table was generated through a full numerical Boltzmann analysis including all relativistic, redshifting, and exponential-decay effects. The different cases in this table are labeled according to the same labeling scheme as in Table II, with K1 and K2 corresponding to the two possibilities for Case K within Table II with absolute marginalities which are either $\mathcal{O}(1)$ or far, respectively. In all cases we find that the results of this explicit example agree with the general properties outlined in Table II.

Case	Parent		Decay		Daughter distribution		
	$\frac{\langle P_{\text{prod}} \rangle}{m_P}$	$\frac{\langle P_{\text{decay}} \rangle}{m_P}$	$\frac{m_D}{m_P}$	$\frac{\langle P_D^{\text{rest}} \rangle}{m_P}$	$\frac{\langle P_D \rangle}{m_D}$	$\frac{\Delta p_D}{m_D}$	$\frac{\Delta p_D}{\langle P_D \rangle}$
A	4.6×10^{-5}	10^{-6}			6.4×10^{-4}	3.3×10^{-4}	0.52
B	4.5×10^{-2}	10^{-3}			7.2×10^{-4}	3.3×10^{-4}	0.46
D	53	1	$\frac{1}{2} - 10^{-6}$	10^{-3}	0.38	6.7×10^{-4}	1.8×10^{-3}
H	$1.4 \times 10^{+6}$	10^{+3}			405	0.55	1.4×10^{-3}
C	4.6×10^{-2}	10^{-3}			0.19	0.1	0.52
E	12	1/4	0.4333	1/4	0.22	0.1	0.45
K1	$1.4 \times 10^{+6}$	10^{+3}			472	136	0.29
F	4.5×10^{-5}	10^{-6}			$1.6 \times 10^{+3}$	832	0.52
G	4.7×10^{-3}	10^{-4}			$1.6 \times 10^{+3}$	831	0.52
I	0.47	10^{-2}	10^{-4}	1/2	$1.5 \times 10^{+3}$	803	0.52
J	25	1/2			$2.0 \times 10^{+3}$	848	0.43
K2	$1.5 \times 10^{+6}$	10^{+3}			$2.1 \times 10^{+6}$	$1.2 \times 10^{+6}$	0.58

hierarchies between our input parameters in Table IV have usually been chosen to be approximately 10^3 or larger. We therefore use this same hierarchical scale when deciding, for example, whether a given daughter width Δp_D is to be considered much smaller than, of the same order of magnitude as, or much larger than m_D or $\langle p_D \rangle$. The agreement between these two tables thus provides further confirmation of the results in Table II, with Table IV serving as an explicit example of the physics underlying Table II.

APPENDIX C: ADIABATIC SOUND SPEED

As background for a technical point discussed in Sec. III, in this appendix we provide a short derivation of the adiabatic sound speed c_s associated with dark matter of a given momentum p .

In general, c_s describes the response of the pressure P due to a change in the energy density ρ . Any pressure perturbation δP can be written in terms of an entropy perturbation δs and an energy-density perturbation $\delta \rho$ as

$$\begin{aligned} \delta P &= \frac{\partial P}{\partial s} \Big|_{\rho} \delta s + \frac{\partial P}{\partial \rho} \Big|_s \delta \rho \\ &= \delta P_{\text{nad}} + c_s^2 \delta \rho, \end{aligned} \quad (\text{C1})$$

where P_{nad} is a nonadiabatic contribution which will not concern us and where

$$c_s^2 \equiv \frac{\partial P}{\partial \rho} \Big|_s. \quad (\text{C2})$$

As such, the adiabatic sound speed depends only on background quantities, and these do not vary in space but only in time. We can therefore reexpress c_s in the form

$$c_s^2 = \dot{\bar{P}} / \dot{\bar{\rho}} \quad (\text{C3})$$

where $\bar{\rho}$ and \bar{P} are average background values.

Some dark sectors have dark matter exhibiting relatively simple phase-space distributions $f(p)$. In such cases, as discussed in Sec. III, one might consider calculating c_s by averaging across all momenta in order to consider the time variations of a momentum-averaged pressure and a momentum-averaged energy density. This would then hopefully provide a characteristic value of the associated sound speed. In this paper, by contrast, we are interested in situations in which the dark-matter phase-space distribution is relatively complex and potentially even multimodal. In such cases, a momentum average might then fail to capture all of the relevant information.

For this reason, we shall proceed by viewing each momentum slice through the $f(p)$ distribution as its own effective ‘‘species’’ having its own sound speed $c_s(p)$. Indeed, as we shall demonstrate, the sound speed varies nontrivially with p , thereby justifying this approach. In order to calculate $c_s(p)$, let us assume that n_p is the number density of the dark-matter particles with momentum p . We then have

$$\rho_p = \bar{\rho}_p = n_p E_p, \quad P_p = \bar{P}_p = n_p \frac{p^2}{3E_p}, \quad (\text{C4})$$

where $E_p \equiv \sqrt{p^2 + m^2}$. Given that $n_p \sim a^{-3}$ and $p \sim a^{-1}$, we find that $dE_p/da = -p^2/(E_p a)$, whereupon it follows that

$$\frac{d\rho_p}{da} = -\frac{\rho_p}{a} \left(3 + \frac{p^2}{E_p^2} \right), \quad \frac{dP_p}{da} = -\frac{P_p}{a} \left(5 - \frac{p^2}{E_p^2} \right). \quad (\text{C5})$$

We therefore find that the sound speed is given by

$$c_s^2(p) = \frac{dP_p/da}{d\rho_p/da} = \frac{1}{3} \frac{p^2}{E_p^2} \left(\frac{5 - p^2/E_p^2}{3 + p^2/E_p^2} \right). \quad (\text{C6})$$

From this result we easily verify that $c_s \rightarrow 0$ for cold (nonrelativistic) matter with $p/E_p \rightarrow 0$, while $c_s \rightarrow 1/\sqrt{3}$ for highly relativistic matter (or radiation) with $p/E_p \rightarrow 1$. This accords with our usual expectation that $c_s = \sqrt{w}$ for matter with equation-of-state parameter w .

APPENDIX D: TIME EVOLUTION OF THE DARK SECTOR

In Sec. IV, we introduced an illustrative model involving a collection of decaying fields ϕ_ℓ and focused on the properties of the resulting present-day phase-space distribution $g_0(p, t_{\text{now}})$ associated with the stable, lightest constituent ϕ_0 . We also analyzed the impact of this phase-space distribution on the matter power spectrum $P(k)$. While the properties of $g_0(p, t_{\text{now}})$ are an important focus of this paper, it is nevertheless also interesting to examine how the properties of the dark sector evolve *while* the individual constituent fields within the dark sector are actively evolving and decaying. These properties include not only the phase-space distributions and energy densities of the individual ϕ_ℓ , but also the aggregate equation of state for the dark sector as a whole.

These quantities are defined and calculated as follows. Given the particular decay patterns that arise in our model, solving the associated Boltzmann equations provides us with the phase-space distribution $f_\ell(p, t)$ for each species ϕ_ℓ as the decays proceed. Using the expressions in Eq. (2.1) it is then straightforward to calculate the corresponding number densities $n_\ell(t)$, energy densities $\rho_\ell(t)$, and pressures $P_\ell(t)$ which are also functions of time as the decays proceed. Likewise, the total number density $n_{\text{tot}}(t)$, total energy density $\rho_{\text{tot}}(t)$, and total pressure $P_{\text{tot}}(t)$ of the dark sector as a whole are simply sums of these individual contributions:

$$X_{\text{tot}}(t) \equiv \sum_{\ell=0}^N X_\ell(t) \quad \text{for } X = n, \rho, P. \quad (\text{D1})$$

We shall also define the individual species equation-of-state parameters $w_\ell(t) \equiv P_\ell(t)/\rho_\ell(t)$ as well as an aggregate equation-of-state parameter $w_{\text{tot}}(t) \equiv P_{\text{tot}}(t)/\rho_{\text{tot}}(t)$ for the dark sector as a whole. Note that these quantities are time-dependent as a result of the decays which occur within the dark sector. It is also important to note that unlike the aggregate quantities appearing in Eq. (D1), $w_{\text{tot}}(t)$ is actually a *weighted* sum of the individual $w_\ell(t)$ contributions:

$$w_{\text{tot}}(t) \equiv \frac{P_{\text{tot}}(t)}{\rho_{\text{tot}}(t)} = \sum_{\ell=0}^N \left[\frac{\rho_\ell(t)}{\rho_{\text{tot}}(t)} \right] w_\ell(t), \quad (\text{D2})$$

where the energy-density ratios $\rho_\ell(t)/\rho_{\text{tot}}(t)$ serve as the appropriate weighting factors. Finally, we observe that $w_{\text{tot}}(t)$ may equivalently be expressed as [21]

$$w_{\text{tot}}(t) = - \left(\frac{1}{3H} \frac{d \log \rho_{\text{tot}}}{dt} + 1 \right). \quad (\text{D3})$$

In general, a relation of this sort which directly connects $w(t)$ to $\rho(t)$ follows from the fundamental definition $w(t) \equiv P(t)/\rho(t)$ under the additional assumption that the only changes to the total energy density $\rho(t)$ are those due to Hubble expansion, so that $\delta\rho = P\delta V$ where $V \sim a^3$. In other words, the expression in Eq. (D3) assumes that our dark-matter ensemble couples negligibly to any other nongravitational sector. However, the total energy density of our dark sector as a whole is unaffected by decays that occur purely within the dark sector (such as we are considering in our model). We therefore obtain Eq. (D3) for w_{tot} . By contrast, a similar equation would not hold for each w_ℓ separately during any time period in which ϕ_ℓ particles are being created or lost through decays.

We begin our analysis of the time evolution of the dark sector in our model by examining how the energy densities $\rho_\ell(t)$ evolve in time. For reasons to be discussed below, we shall consider the time dependence of the quantities $a^3\rho_\ell$ rather than ρ_ℓ alone. Note that there are generally three different effects that can cause a given $a^3\rho_\ell$ to vary with time. First, this quantity can increase if ϕ_ℓ particles are created via the decays of heavier ensemble constituents. Second, this quantity can decrease if the ϕ_ℓ particles themselves decay into lighter ensemble constituents. Finally, however, $a^3\rho_\ell$ can also scale nontrivially with Hubble expansion. In the absence of ϕ_ℓ production or decay, $a^3\rho_\ell$ would be constant only if the ϕ_ℓ particles were all nonrelativistic. Indeed, in this case we could even interpret $a^3\rho_\ell$ as a *comoving* energy density. By contrast, if the ϕ_ℓ particles carry a significant momentum, this momentum will experience a gravitational redshift towards smaller values, and this will also cause $a^3\rho_\ell$ to drop as a function of time. For example, if ϕ_ℓ is highly relativistic

(e.g., radiation), we know that ϕ_ℓ scales as a^{-4} rather than a^{-3} . Thus $a^3\phi_\ell$ will scale as a^{-1} , even if no ϕ_ℓ particles are being created or lost through decay.

This last effect also applies to the total energy density $a^3\rho_{\text{tot}}$. Because each decay process within the dark sector necessarily conserves energy, $a^3\rho_{\text{tot}}$ is unaffected by decays within the dark sector. However, Hubble expansion continues its inexorable degradation of the overall kinetic energy associated the dark sector. As a result, if our dark sector consists of highly relativistic matter, its energy density $a^3\rho_{\text{tot}}$ will initially scale as a^{-1} . By contrast, at late times, $a^3\rho_{\text{tot}}$ will become effectively constant as our dark sector becomes increasingly cold. Thus, a significant variation of $a^3\rho_{\text{tot}}$ with time indicates that a significant fraction of the total dark-sector energy density is still carried by relativistic particles.

In Fig. 19 we plot the time evolution of $a^3\rho_\ell$ for each species ϕ_ℓ in our model, where the different panels in this figure correspond to the same choices for the parameters r and s as in Figs. 14 through 18. Note that the ten solid curves (ranging from red to blue) shown within each panel correspond to the individual $a^3\rho_\ell$ for $\ell = 0, 1, \dots, 9$, respectively, with each normalized to the total $a^3\rho_{\text{tot}}$ evaluated at the initial time $t = t_I$. Likewise, the black dashed curve shows the evolution of the total energy density $a^3\rho_{\text{tot}}$, similarly normalized to its initial value.

It is straightforward to understand the curves shown in these panels. Let us first consider the regime in which $r \leq 0$ and $s \geq 0$ —i.e., the four panels in the upper right corner of Fig. 19. As indicated in Fig. 15, the characteristic timescales associated with all relevant decay chains are quite similar in this regime. As a result, even if the intermediate states in these decay chains acquire significant energy densities—as indeed occurs in most of these figure panels—these individual energies ρ_ℓ are ultimately transferred to ϕ_0 in each case on essentially the same timescale. Moreover, within this regime, decays in which a significant amount of kinetic energy is transferred to the daughter particles are not preferred—and for $r < 0$ these decays are actively suppressed. As a result, the fraction of the initial energy density $\rho_9(t_I)$ which is eventually dissipated by the redshifting of daughter particles is small compared to the fraction which is converted into the mass energy of ϕ_0 particles. This implies that the overall decrease in $a^3\rho_{\text{tot}}$ is comparatively small within this regime, and markedly smaller for $r = -3$ than for $r = 0$.

By contrast, in the $s < 0$ and $r \leq 0$ regime, the preference for decays which are highly asymmetric but yet not highly exothermic leads to a decay-chain structure involving multiple subsequent steps of the form $\phi_\ell \rightarrow \phi_i\phi_j$, each of which yields one very light particle ϕ_i and another particle ϕ_j with j only slightly below ℓ . In this regime, the fraction of the initial mass energy of ϕ_9 particles converted to kinetic energy along any given decay chain is likewise

small—and likewise markedly smaller for $r = -3$ than for $r = 0$. However, the timescales on which the individual intermediate ϕ_ℓ acquire significant energy densities are also quite different in this regime. In particular, successively lighter states acquire a significant ρ_ℓ at successively later times due to an injection of energy density from the decays of the states just above them. The fact that a significant population of ϕ_0 particles is generated at each step along these decay chains gives rise to the complex multimodal phase distributions which appear in the corresponding panels of Fig. 16.

Finally, in the $r > 0$ regime, the preference for decays which are highly exothermic leads to a decay-chain structure in which ϕ_9 particles decay directly to final states comprising much lighter ϕ_ℓ . Since a considerable fraction of the initial mass energy of ϕ_9 particles is converted into kinetic energy in the process, the decrease in $a^3\rho_{\text{tot}}$ is comparatively large within this regime. Moreover, the timescales on which ϕ_0 particles are produced from the direct decays of ϕ_9 are significantly shorter than the timescales associated with the decays of other light ϕ_ℓ particles which might also be produced from the decays of ϕ_9 down to ϕ_0 . Thus, as we see in the panels along the bottom row of Fig. 19, the timescales on which the energy densities ρ_ℓ associated with these other light species are transferred to ϕ_0 are quite long. The fact that one population of ϕ_0 particles is generated from the direct decays of ϕ_9 while another population of ϕ_0 particles is produced from the decays of these lighter longer-lived states is ultimately responsible for the multimodal structure of the phase-space distributions appearing in the corresponding panels of Fig. 16.

A similar analysis can also be performed for our time-dependent equation-of-state parameters $w_\ell(t)$ and $w_{\text{tot}}(t)$, as defined in and above Eqs. (D2) and (D3). In Fig. 20, we display $w_\ell(t)$ for $\ell = 0, 1, \dots, 9$ as functions of the ratio t/t_I , with the same color and panel configurations as in Fig. 19. We also display the aggregate equation-of-state parameter $w_{\text{tot}}(t)$ for the dark sector as a whole (dashed black curve).

Once again, just as with Fig. 19, the behavior of each individual equation-of-state parameter $w_\ell(t)$ ultimately reflects the populating and depopulating of the ϕ_ℓ state as the decay process unfolds. Moreover, the relation in Eq. (D3) implies that we may view $w_{\text{tot}}(t)$ at any given time as measure of how rapidly $\rho_{\text{tot}}(t)$ is changing with time. Indeed, we observe that when $w_{\text{tot}}(t)$ differs significantly from zero in any given panel of Fig. 20, the slope of the black dashed curve in the corresponding panel of Fig. 19 also differs significantly from zero. Combinations of r and s for which there exist multiple decay pathways producing ϕ_0 particles on significantly different characteristic timescales therefore typically give rise to $w_{\text{tot}}(t)$ curves with multiple peaks. Thus, if the $w_{\text{tot}}(t)$ curve exhibits such nonmonotonicities as a function of time, there is a good

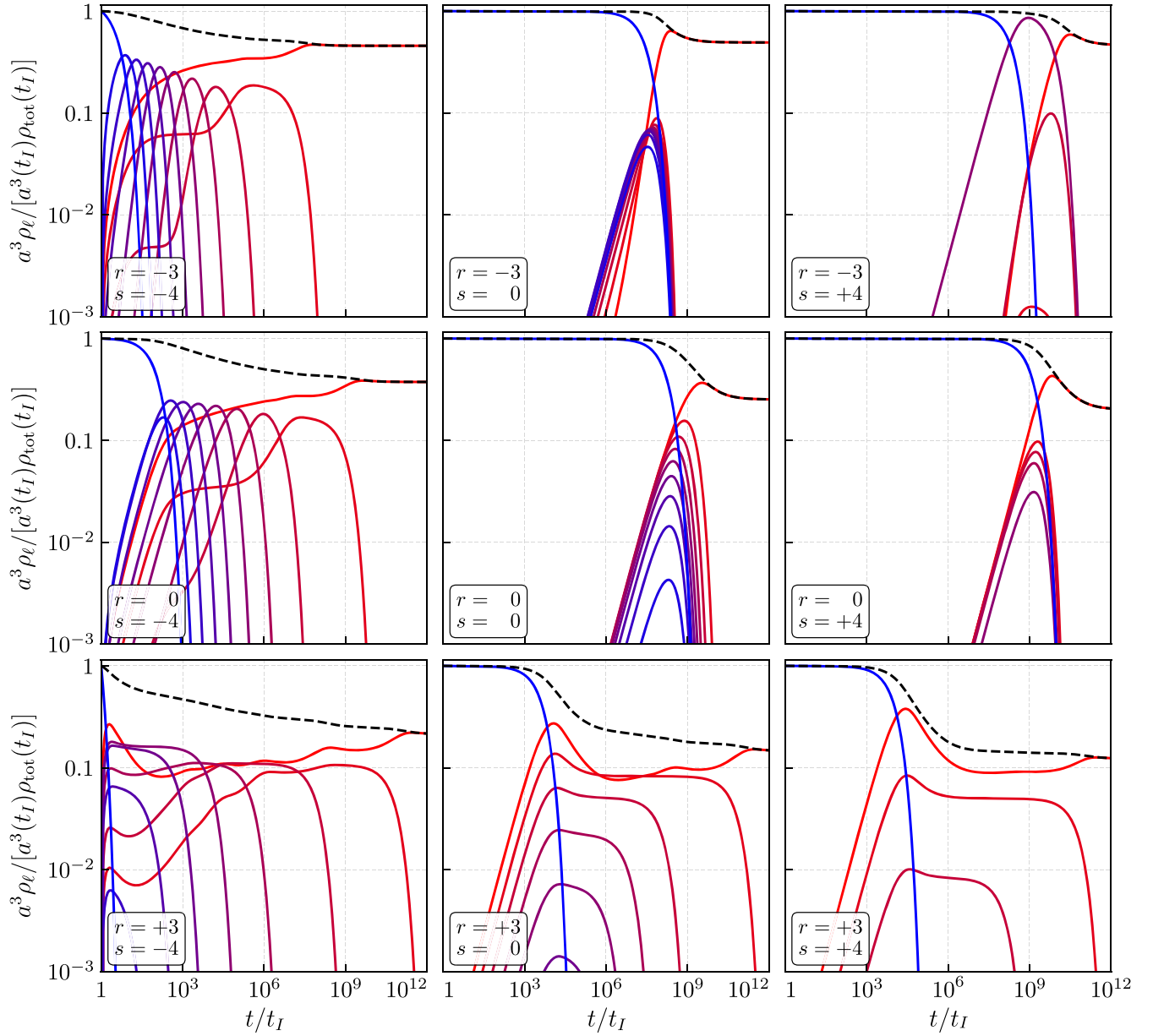


FIG. 19. The individual constituent energy densities $a^3\rho_\ell$ for $\ell = 0, \dots, 9$ (solid colored curves ranging from red to blue, respectively), each normalized to the total energy density $a^3\rho_{\text{tot}}$ evaluated at the initial time $t = t_I$ and plotted as a function of the dimensionless ratio t/t_I . The total energy density $a^3\rho_{\text{tot}}$, similarly normalized to its initial value, is also plotted (black dashed curve).

chance that the corresponding $g_0(p, t_{\text{now}})$ will turn out to be multimodal as a function of p .

Finally, in Fig. 21, we illustrate how the phase-space distribution $g_0(p, t)$ of the lightest state in the dark sector evolves in time. The curves shown in each panel represent “snapshots” of $g_0(p, t)$ at five different values of t . For each combination of r and s , these snapshot times are determined by dividing the time interval between t_I and the time t_f at which the comoving number density of ϕ_0 reaches 99.9% of its late-time asymptotic value into five time intervals which are evenly spaced on a logarithmic scale. The red curve in each panel

corresponds to the earliest snapshot, while the blue curve corresponds to the final snapshot when essentially the entire dark-sector abundance has settled into the ground state. However, we emphasize that since $\tau_9 \gg t_I$ for many of the combinations of r and s shown—in particular, for those combinations with $r \leq 0$ and $s \geq 0$ —a non-negligible population of ϕ_0 particles is not generated until late times. Thus, in the corresponding panels of Fig. 21, the only curve which deviates considerably from $g_0(p, t) \approx 0$ is the one corresponding to $t = t_f$.

We have seen in Sec. IV that the present-day phase-space distribution for ϕ_0 is unimodal for $r \leq 0$ and $s \geq 0$.

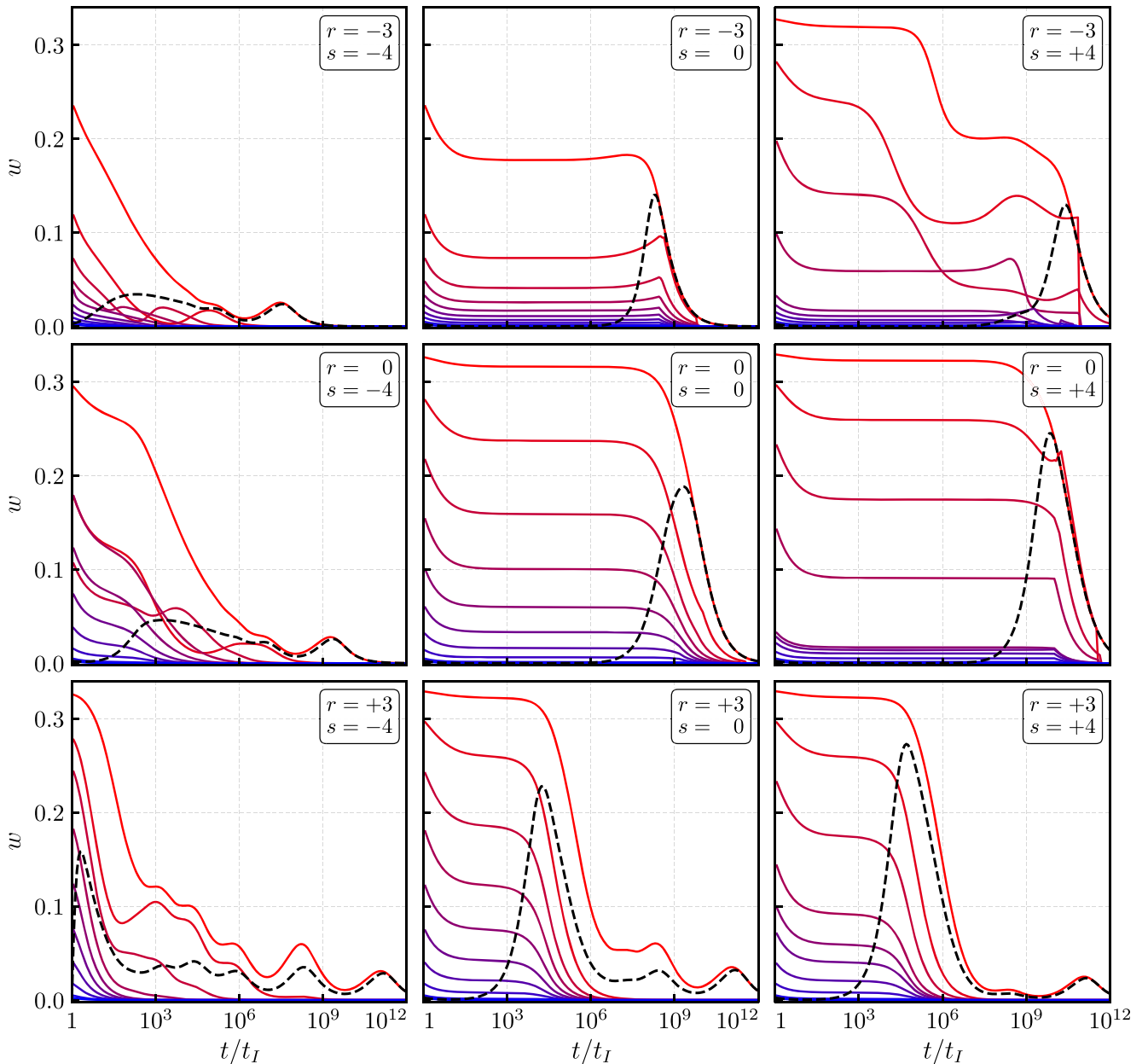


FIG. 20. The equation-of-state parameters $w_\ell(t)$ for each dark-sector species ϕ_ℓ , plotted as functions of the dimensionless ratio t/t_I for $\ell = 0, \dots, 9$ (red to blue, respectively). By contrast, the black dashed curve represents the aggregate equation-of-state parameter $w_{\text{tot}}(t)$ of the dark sector as a whole.

The results shown in the corresponding panels of Fig. 21 provide some additional insight into how such a phase-space distribution develops. In particular, since all relevant decay chains through which ϕ_9 ultimately decays to the ground state have similar characteristic timescales, the deposits from these different decay chains arrive on the $g_0(p, t)$ “conveyor belt” at roughly the same time. Prior to this time—which for the particular parameter choices adopted in these panels is quite late in comparison with t_I —we have $g_0(p, t) \approx 0$. Thus the deposits from all of these decay chains all arrive together at this late time,

without a significant intervening time interval over which substantial redshifting can occur. This then generates a single, narrow peak for $g_0(p, t)$.

By contrast, when $r > 0$ and/or $s > 0$, we have seen in Sec. IV that the present-day phase-space distribution for ϕ_0 is highly nontrivial and multimodal. The results shown in the corresponding panels of Fig. 21 illustrate how this multimodality arises as a consequence of sequential deposits from different decay pathways with different characteristic timescales. Sizable deposits to $g_0(p, t)$ arrive at early times and experience significant redshifts during the time

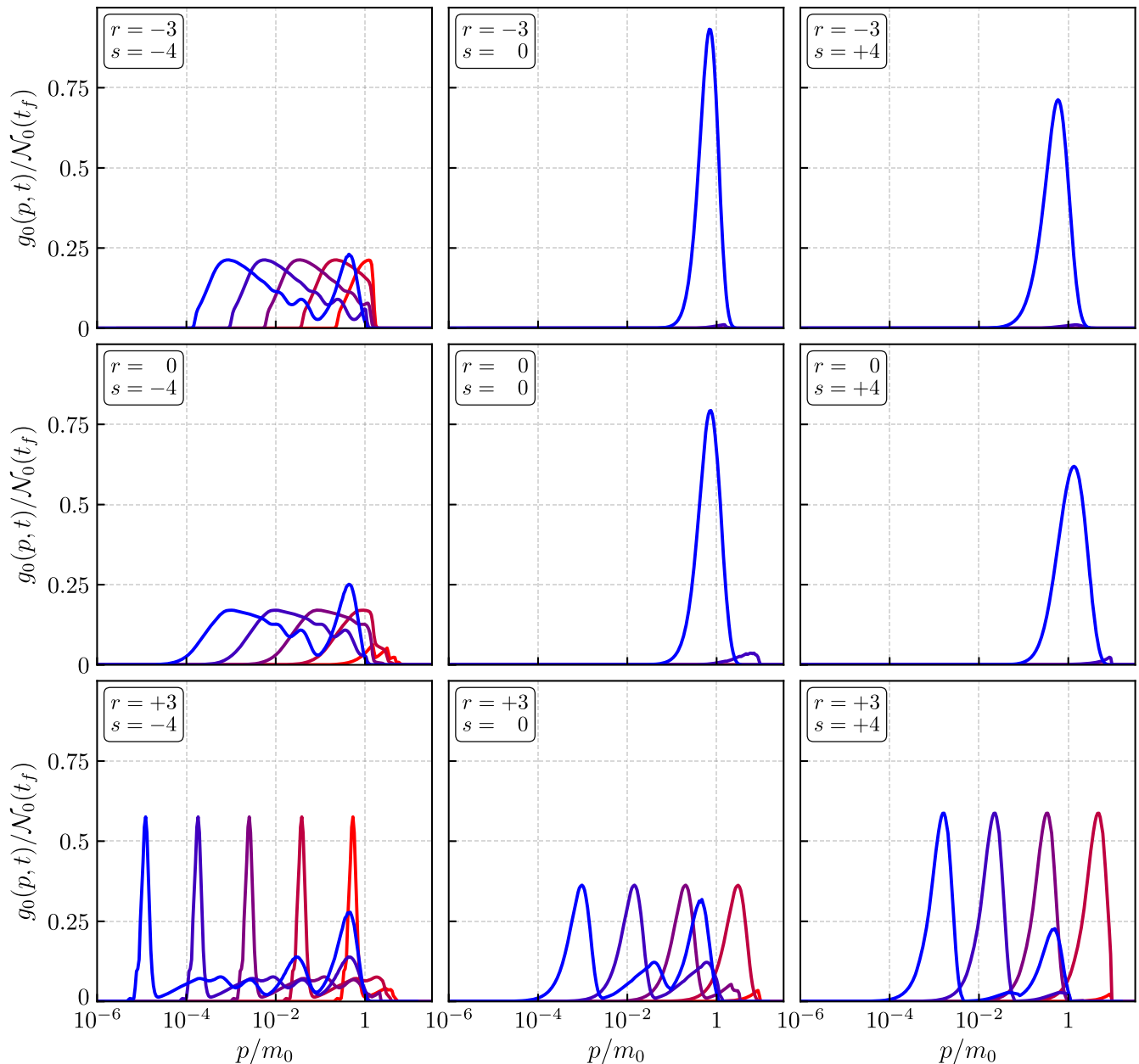


FIG. 21. The evolution of the ground-state phase-space distribution $g_0(p, t)$ as the decay process unfolds. For each value of r and s in a given panel, we show $g_0(p, t)/\mathcal{N}_0$ at five different “snapshot” times (ranging from red to blue) which are evenly spaced on a logarithmic scale between the initial time t_I and the final time t_f at which $g_0(p, t)$ reaches 99.9% of its late-time asymptotic value. Note that for $r \leq 0$ and $s \geq 0$, there are no deposits into the ground state until relatively late times. Consequently fewer phase-space distributions are shown.

interval between t_I and t_f , while additional deposits arrive at subsequent times. This is consistent with our general expectations from Secs. II and IV.

As an example, let us consider the sequence of snapshots that emerges for the case with $(r, s) = (3, 4)$, as indicated in the lower right panel of Fig. 21. At early times, we essentially have only a single large peak (red curve) which then simply rides along the momentum conveyor belt towards smaller momenta at later times. Indeed, by the time of our final snapshot, this large peak has been

transported to values of p/m_0 in the approximate range $10^{-4} \lesssim p/m_0 \lesssim 10^{-2}$ (large bright blue peak). However, by this time, we see from Fig. 21 that another much smaller deposit onto the conveyor belt has been made, corresponding to the small bright blue peak shown within the approximate range $10^{-1} \lesssim p/m_0 \lesssim 1$. Together, these deposits result in the final bimodal phase-space distribution shown in the lower right panel of Fig. 16. Indeed, this sequence of deposits onto the conveyor belt is also consistent with the results shown in the lower right

panel of Fig. 15, in which two dominant reduced decay chains make deposits onto the conveyor belt at different times. However, the snapshots in Fig. 21 now allow us to confirm that the smaller of the two peaks within the final bimodal distribution was deposited later than than the larger peak. Thus the dark-sector ϕ_0 particles which populate the larger peak within the ϕ_0 phase-space distribution were produced earlier than those which populate the smaller peak.

Interestingly, if we look even more closely at the lower right panel of Fig. 21, we observe that there is actually a very small bump (bright red) within the approximate range $1 \lesssim p/m_0 \lesssim 10$ which was deposited even before the larger red peak. (Indeed, the decay chain corresponding to this deposit is too subdominant to appear in Fig. 15.) After redshifting, this eventually contributes to the very small bump shown within the approximate range $10^{-7} \lesssim p/m \lesssim 10^{-6}$ in the lower right panel of Fig. 16. We thus learn that the ϕ_0 particles within this small subdominant bump were actually the first to be produced from amongst the entire phase-space distribution.

We conclude, then, that the time evolution of our dark sector can be highly nontrivial. Although we have focused in this paper on the rich consequences of the late-time phase-space distribution $g_0(p, t)$ for $t \geq t_f$ and its consequences for the matter power spectrum, we now see that the internal dynamics within the dark sector can also exhibit its own richness. For example, we have seen in Fig. 20 that the equation of state of the dark sector can have a nontrivial time evolution at early times—a time dependence which may also leave imprints in the cosmological evolution. Likewise, although we have assumed in this paper that all of the dark-sector decays have concluded long before the present cosmological era, there do exist other interesting cosmological frameworks (such as that associated with Dynamical Dark Matter [21,22]) in which such “intra-ensemble” decays within the dark sector are occurring even at the present time. With only minor changes in the appropriate timescales, the analysis we have performed here should therefore be relevant for such other scenarios as well.

-
- [1] M. Kaplinghat, *Phys. Rev. D* **72**, 063510 (2005).
- [2] J. König, A. Merle, and M. Totzauer, *J. Cosmol. Astropart. Phys.* **11** (2016) 038.
- [3] J. Lesgourgues, [arXiv:1104.2932](https://arxiv.org/abs/1104.2932).
- [4] D. Blas, J. Lesgourgues, and T. Tram, *J. Cosmol. Astropart. Phys.* **07** (2011) 034.
- [5] J. Lesgourgues, [arXiv:1104.2934](https://arxiv.org/abs/1104.2934).
- [6] J. Lesgourgues and T. Tram, *J. Cosmol. Astropart. Phys.* **09** (2011) 032.
- [7] A. Boyarsky, J. Lesgourgues, O. Ruchayskiy, and M. Viel, *J. Cosmol. Astropart. Phys.* **05** (2009) 012.
- [8] R. Murgia, A. Merle, M. Viel, M. Totzauer, and A. Schneider, *J. Cosmol. Astropart. Phys.* **11** (2017) 046.
- [9] R. Barkana, Z. Haiman, and J. P. Ostriker, *Astrophys. J.* **558**, 482 (2001).
- [10] M. Viel, K. Markovič, M. Baldi, and J. Weller, *Mon. Not. R. Astron. Soc.* **421**, 50 (2012).
- [11] C. Destri, H. J. de Vega, and N. G. Sanchez, *Phys. Rev. D* **88**, 083512 (2013).
- [12] V. Springel, *Mon. Not. R. Astron. Soc.* **364**, 1105 (2005).
- [13] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchmann (Virgo Consortium), *Mon. Not. R. Astron. Soc.* **341**, 1311 (2003).
- [14] A. Cooray and R. K. Sheth, *Phys. Rep.* **372**, 1 (2002).
- [15] R. E. Smith and K. Markovic, *Phys. Rev. D* **84**, 063507 (2011).
- [16] A. Schneider, D. Anderhalden, A. Maccio, and J. Diemand, *Mon. Not. R. Astron. Soc.* **441**, L6 (2014).
- [17] M. Boylan-Kolchin, J. S. Bullock, and M. Kaplinghat, *Mon. Not. R. Astron. Soc.* **422**, 1203 (2012).
- [18] A. V. Maccio, S. Paduroiu, D. Anderhalden, A. Schneider, and B. Moore, *Mon. Not. R. Astron. Soc.* **424**, 1105 (2012).
- [19] M.-Y. Wang, A. H. G. Peter, L. E. Strigari, A. R. Zentner, B. Arant, S. Garrison-Kimmel, and M. Rocha, *Mon. Not. R. Astron. Soc.* **445**, 614 (2014).
- [20] R. E. Angulo, O. Hahn, and T. Abel, *Mon. Not. R. Astron. Soc.* **434**, 3337 (2013).
- [21] K. R. Dienes and B. Thomas, *Phys. Rev. D* **85**, 083523 (2012).
- [22] K. R. Dienes and B. Thomas, *Phys. Rev. D* **85**, 083524 (2012).
- [23] K. R. Dienes and B. Thomas, *Phys. Rev. D* **86**, 055013 (2012).
- [24] Y. Buyukdag, K. R. Dienes, T. Gherghetta, and B. Thomas, *Phys. Rev. D* **101**, 075054 (2020).
- [25] K. Choi and S. H. Im, *J. High Energy Phys.* **01** (2016) 149.
- [26] D. E. Kaplan and R. Rattazzi, *Phys. Rev. D* **93**, 085007 (2016).
- [27] G. F. Giudice and M. McCullough, *J. High Energy Phys.* **02** (2017) 036.
- [28] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, *Phys. Rev. Lett.* **86**, 4757 (2001).
- [29] N. Craig, I. Garcia Garcia, and D. Sutherland, *J. High Energy Phys.* **10** (2017) 018.
- [30] G. F. Giudice and M. McCullough, [arXiv:1705.10162](https://arxiv.org/abs/1705.10162).
- [31] K. Choi, S. H. Im, and C. S. Shin, *J. High Energy Phys.* **07** (2018) 113.
- [32] G. F. Giudice, Y. Kats, M. McCullough, R. Torre, and A. Urbano, *J. High Energy Phys.* **06** (2018) 009.
- [33] I. Antoniadis, A. Arvanitaki, S. Dimopoulos, and A. Givon, *Phys. Rev. Lett.* **108**, 081602 (2012).

- [34] K. R. Dienes, F. Huang, S. Su, and B. Thomas, *Phys. Rev. D* **95**, 043526 (2017).
- [35] L. H. Ford, *Phys. Rev. D* **35**, 2955 (1987).
- [36] D. J. H. Chung, E. W. Kolb, and A. Riotto, *Phys. Rev. D* **59**, 023501 (1998).
- [37] D. J. H. Chung, P. Crotty, E. W. Kolb, and A. Riotto, *Phys. Rev. D* **64**, 043503 (2001).
- [38] B. A. Bassett and S. Liberati, *Phys. Rev. D* **58**, 021302 (1998); **60**, 049902(E) (1999).
- [39] Y. Ema, R. Jinno, K. Mukaida, and K. Nakayama, *J. Cosmol. Astropart. Phys.* **05** (2015) 038.
- [40] T. Markkanen and S. Nurmi, *J. Cosmol. Astropart. Phys.* **02** (2017) 008.
- [41] Y. Ema, R. Jinno, K. Mukaida, and K. Nakayama, *Phys. Rev. D* **94**, 063517 (2016).
- [42] Y. Ema, K. Nakayama, and Y. Tang, *J. High Energy Phys.* **09** (2018) 135.
- [43] M. Fairbairn, K. Kainulainen, T. Markkanen, and S. Nurmi, *J. Cosmol. Astropart. Phys.* **04** (2019) 005.
- [44] T. Markkanen, A. Rajantie, and T. Tenkanen, *Phys. Rev. D* **98**, 123532 (2018).
- [45] Y. Ema, K. Nakayama, and Y. Tang, *J. High Energy Phys.* **07** (2019) 060.
- [46] K. R. Dienes, L. Heurtier, J. Kost, and B. Thomas (to be published).
- [47] J. H. Traschen and R. H. Brandenberger, *Phys. Rev. D* **42**, 2491 (1990).
- [48] L. Kofman, A. D. Linde, and A. A. Starobinsky, *Phys. Rev. Lett.* **73**, 3195 (1994).
- [49] Y. Shtanov, J. H. Traschen, and R. H. Brandenberger, *Phys. Rev. D* **51**, 5438 (1995).
- [50] L. Kofman, A. D. Linde, and A. A. Starobinsky, *Phys. Rev. D* **56**, 3258 (1997).
- [51] J. McDonald, *Phys. Rev. Lett.* **88**, 091304 (2002).
- [52] A. Kusenko, *Phys. Rev. Lett.* **97**, 241301 (2006).
- [53] K. Petraki and A. Kusenko, *Phys. Rev. D* **77**, 065014 (2008).
- [54] L. J. Hall, K. Jedamzik, J. March-Russell, and S. M. West, *J. High Energy Phys.* **03** (2010) 080.
- [55] Y. Gong and X. Chen, *Phys. Rev. D* **77**, 103511 (2008).
- [56] G. Blackadder and S. M. Koushiappas, *Phys. Rev. D* **90**, 103527 (2014).
- [57] G. Blackadder and S. M. Koushiappas, *Phys. Rev. D* **93**, 023510 (2016).
- [58] A. Desai, K. R. Dienes, and B. Thomas, *Phys. Rev. D* **101**, 035031 (2020).
- [59] W. Hu and J. Silk, *Phys. Rev. D* **48**, 485 (1993).
- [60] W. Hu and J. Silk, *Phys. Rev. Lett.* **70**, 2661 (1993).
- [61] M. Kawasaki and T. Moroi, *Astrophys. J.* **452**, 506 (1995).
- [62] S. Sarkar, *Rep. Prog. Phys.* **59**, 1493 (1996).
- [63] J. A. Adams, S. Sarkar, and D. W. Sciama, *Mon. Not. R. Astron. Soc.* **301**, 210 (1998).
- [64] R. H. Cyburt, J. R. Ellis, B. D. Fields, and K. A. Olive, *Phys. Rev. D* **67**, 103521 (2003).
- [65] X.-L. Chen and M. Kamionkowski, *Phys. Rev. D* **70**, 043502 (2004).
- [66] T. R. Slatyer, N. Padmanabhan, and D. P. Finkbeiner, *Phys. Rev. D* **80**, 043526 (2009).
- [67] R. H. Cyburt, J. Ellis, B. D. Fields, F. Luo, K. A. Olive, and V. C. Spanos, *J. Cosmol. Astropart. Phys.* **10** (2009) 021.
- [68] D. P. Finkbeiner, S. Galli, T. Lin, and T. R. Slatyer, *Phys. Rev. D* **85**, 043522 (2012).
- [69] T. R. Slatyer, *Phys. Rev. D* **87**, 123513 (2013).
- [70] M. Kawasaki, K. Kohri, T. Moroi, and Y. Takaesu, *Phys. Rev. D* **97**, 023502 (2018).
- [71] K. R. Dienes, J. Kumar, P. Stengel, and B. Thomas, *Phys. Rev. D* **99**, 043513 (2019).
- [72] E. W. Kolb and M. S. Turner, *Front. Phys.* **69**, 1 (1990).
- [73] P. Adshead, Y. Cui, and J. Shelton, *J. High Energy Phys.* **06** (2016) 016.
- [74] E. Hardy and J. Unwin, *J. High Energy Phys.* **09** (2017) 113.
- [75] K. R. Dienes, J. Kost, and B. Thomas, *Phys. Rev. D* **93**, 043540 (2016).
- [76] K. R. Dienes, J. Kost, and B. Thomas, *Phys. Rev. D* **95**, 123539 (2017).
- [77] K. R. Dienes, J. Kost, and B. Thomas, *Phys. Rev. D* **100**, 083516 (2019).
- [78] A. Merle and M. Totzauer, *J. Cosmol. Astropart. Phys.* **06** (2015) 011.
- [79] M. R. Lovell, S. Bose, A. Boyarsky, S. Cole, C. S. Frenk, V. Gonzalez-Perez, R. Kennedy, O. Ruchayskiy, and A. Smith, *Mon. Not. R. Astron. Soc.* **461**, 60 (2016).
- [80] W. A. Fendt and B. D. Wandelt, *Astrophys. J.* **654**, 2 (2007).
- [81] W. A. Fendt and B. D. Wandelt, [arXiv:0712.0194](https://arxiv.org/abs/0712.0194).
- [82] T. Auld, M. Bridges, M. P. Hobson, and S. F. Gull, *Mon. Not. R. Astron. Soc.* **376**, L11 (2007).
- [83] T. Auld, M. Bridges, and M. P. Hobson, *Mon. Not. R. Astron. Soc.* **387**, 1575 (2008).
- [84] A. Manrique-Yus and E. Sellentin, *Mon. Not. R. Astron. Soc.* **491**, 2655 (2020).
- [85] K. J. Bae, R. Jinno, A. Kamada, and K. Yanagi, *J. Cosmol. Astropart. Phys.* **03** (2020) 042.
- [86] J. Albers, C. Fidler, J. Lesgourgues, N. Schneberg, and J. Torrado, *J. Cosmol. Astropart. Phys.* **09** (2019) 028.