The Physical Origin of Long Gas Depletion Times in Galaxies

Vadim A. Semenov1,2 ©, Andrey V. Kravtsov1,2,3 ©, and Nickolay Y. Gnedin1,2,4

1 Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637, USA; semenov@uchicago.edu 2 Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, IL 60637, USA 3 Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637, USA 4 Fermilab Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510-0500, USA

Received 2017 April 19; revised 2017 July 11; accepted 2017 July 16; published 2017 August 18

Abstract

We present a model that explains why galaxies form stars on a timescale significantly longer than the timescales of processes governing the evolution of interstellar gas. We show that gas evolves from a non-star-forming to a star-forming state on a relatively short timescale, and thus the rate of this evolution does not limit the star formation rate (SFR). Instead, the SFR is limited because only a small fraction of star-forming gas is converted into stars before star-forming regions are dispersed by feedback and dynamical processes. Thus, gas cycles into and out of a star-forming state multiple times, which results in a long timescale on which galaxies convert gas into stars. Our model does not rely on the assumption of equilibrium and can be used to interpret trends of depletion times with the properties of observed galaxies and the parameters of star formation and feedback recipes in simulations. In particular, the model explains how feedback self-regulates the SFR in simulations and makes it insensitive to the local star formation efficiency. We illustrate our model using the results of an isolated L* -sized galaxy simulation that reproduces the observed Kennicutt–Schmidt relation for both molecular and atomic gas. Interestingly, the relation for molecular gas is almost linear on kiloparsec scales, although a nonlinear relation is adopted in simulation cells. We discuss how a linear relation emerges from non-self-similar scaling of the gas density PDF with the average gas surface density.

Key words: galaxies: evolution – ISM: kinematics and dynamics – methods: numerical – stars: formation

1. Introduction

One of the widely recognized basic facts about observed galaxies is that they convert gas into stars inefficiently. Star formation rates (SFRs) of galaxies are surprisingly low, given the total mass and density of their interstellar gas and the expected timescales of the processes driving star formation.

As an example, the SFR of the Milky Way (MW) is \( \dot{M}_* \sim 1 \erm{1} M_\odot \erm{yr}^{-1} \), while its gas mass is \( M_g \sim 10^{10} M_\odot \). Thus, the timescale at which the Galaxy would deplete its gas while forming stars at the current rate is \( \tau_{\text{dep}} \equiv M_g/\dot{M}_* \sim 5 \erm{Gyr} \). The depletion timescales of a population of normal star-forming galaxies are comparable and span a range of \( \sim 2 \erm{Gyr} \) (Kennicutt 1989, 1998; Bigiel et al. 2008). The denser molecular phase of the interstellar medium (ISM) is depleted on a similarly long timescale of \( \tau_{\text{dep}, M_\text{H}_2} \equiv M_{\text{H}_2}/\dot{M}_* \sim 1 \erm{Gyr} \) (Kennicutt 1989, 1998; Wong & Blitz 2002; Bigiel et al. 2008; Leroy et al. 2008, 2013).

Compared to the timescales of any dynamical processes that are potentially relevant for star formation, the observed gas depletion times are very long indeed. For example, the orbital period of gas at the solar radius is \( t_{\text{orb}} \sim 200 \erm{Myr} \), and the MW is thus depleting its gas on a timescale of \( \sim 25 \erm{Gyr} \). On average, galaxies deplete their gas on a timescale of \( \sim 10 \erm{Gyr} \) orbital periods (Kennicutt 1998; Wong & Blitz 2002; Leroy et al. 2008; Daddi et al. 2010).

The orbital period, \( t_{\text{orb}} \), is probably the longest of the relevant dynamical timescales one can think of. For example, the turbulent crossing time is usually \( t_{\text{cross}} = h/\sigma \sim 10 \erm{Myr} \), where \( \sigma \gesim 10 \erm{km \, s}^{-1} \) is the velocity dispersion of gas in galactic disks and \( h \sim 100 \erm{pc} \) is the disk scale height in the inner regions of galaxies. The freefall time at the mean or midplane density, \( \rho_0 \), of galaxies spans a similar range: \( t_{\text{ff}, 0} \equiv \sqrt{3\pi/32G\rho_0} \sim 10 \erm{Myr} \). The timescale of molecular cloud collisions is \( t_{\text{cl}} \lesssim 20 \erm{Myr} \) (e.g., Tan 2000). A given gas mass encounters a spiral arm on a timescale of \( t_{\text{arm}} \sim 2\pi/(\Omega(R) - \Omega_p) \), where \( \Omega(R) = V_\text{rot}/R \) is the angular frequency of the gas rotation, \( \Omega_p \) is the pattern speed of the spiral arms, and \( m \) is the number of spiral arms. This timescale is \( t_{\text{arm}} \sim 50 \erm{Myr} \) if we assume \( \Omega_p \sim 20 \erm{km \, s}^{-1} \erm{kpc}^{-1} \) (e.g., Bissantz et al. 2003), \( m \sim 2 \erm{Gyr} \) (e.g., Davis et al. 2015), and \( V_\text{rot} \sim 220 \erm{km \, s}^{-1} \) typically derived for MW-like galaxies. Numerical simulations of gaseous galactic disks show that star-forming molecular clouds may form on even shorter timescales of a few tens of Myr (Dobbs et al. 2012, 2015).

In addition to being slow on global galactic scales, star formation is inefficient even in dense molecular star-forming regions, which convert only \( \lesssim 1\%\sim 10\% \) of gas into stars per local freefall time (Zuckerman & Evans 1974; Zuckerman & Palmer 1974; Krumholz & Tan 2007; Krumholz et al. 2012; Evans et al. 2014; Heyer et al. 2016; Lee et al. 2016). Such low efficiency arises because only \( \sim 0.1\%\sim 10\% \) of the dense gas is self-gravitating and collapsing into stars (Froebrich & Rowles 2010).

However, the inefficiency of star-forming regions alone cannot explain long global depletion times. The local depletion time in observed star-forming regions is \( t_s \sim 40 \erm{Gyr} \) (e.g., Evans et al. 2009, 2014; Heiderman et al. 2010; Lada et al. 2010, 2012; Gutermuth et al. 2011; Schruba et al. 2017). Thus, although the scatter is significant, typical values of \( t_s \) are considerably smaller than the global depletion time of molecular gas, \( \tau_{\text{dep}, M_\text{H}_2} \sim 1 \erm{Gyr} \).

The large scatter in depletion times measured on small scales and the difference between local and global depletion time values indicate that only a fraction of molecular gas is actively forming stars at any given moment. Indeed, the global
depletion time can be expressed as

\[ \tau_{\text{dep}} \equiv \frac{M_g}{M_*} = \frac{\tau_*}{f_{\text{sf}}} \]  

(1)

where \( M_g \) is the total gas mass of the galaxy; \( \tau_* = M_{sf}/M_* = (1/t_{\text{sf}})^{-1} \) is the mass-weighted average over the depletion time distribution in star-forming regions, \( t_* \); and \( f_{\text{sf}} \equiv M_{sf}/M_g \) is the gas-mass fraction in actively star-forming regions. A similar expression can be written for the global depletion time of molecular gas, \( \tau_{\text{dep,H}_2} \), via a corresponding star-forming fraction \( f_{\text{sf,H}_2} \equiv M_{\text{sf,H}_2}/M_{\text{H}_2} \).

Thus, the depletion time measured on larger scales is longer than that in star-forming regions because, as the scale increases, more non-star-forming gas is incorporated in the gas-mass estimate. Likewise, when the depletion time is estimated on larger scales, the scatter in \( \tau_{\text{dep}} \) decreases as we average over the distribution of local \( t_* \). This is indeed observed (Schruba et al. 2010, 2017), although some of the obtained variation may be due to observational effects (Feldmann et al. 2011; Kuijssen & Longmore 2014).

The small scatter in the global depletion time in observed galaxies is manifested in the Kennicutt–Schmidt relation (KSR) between the surface density of gas and the SFR (Kennicutt 1989, 1998; see also Sanduleak 1969; Madore et al. 1974). The relation is particularly tight and close to linear when only molecular hydrogen is used to estimate the surface density of gas (Wong & Blitz 2002; Bigiel et al. 2008; Leroy et al. 2013). Thus, a comprehensive model for the global depletion time must explain not only its value, i.e., the normalization of the KSR, but also both the scatter and the shape of the KSR on different scales.

A number of useful global star formation frameworks and models have been developed over the last three decades (e.g., Wyse & Silk 1989; Silk 1997; Tan 2000; Elmegreen 2002; Krumholz & McKee 2005; Li et al. 2005; Krumholz & Thompson 2007; Saitoh et al. 2008; Krumholz et al. 2009b; Silk & Norman 2009; Ostriker et al. 2010; Ostriker & Shetty 2011; Renaud et al. 2012; Faucher-Giguère et al. 2013; Federrath 2013; Elmegreen 2015; Salim et al. 2015). Many of these models consider the physical processes shaping the form of the KSR while treating its normalization as a flexible constant. Other models also consider the physical origin of the normalization and long depletion timescale.

One class of the latter models associates a long depletion time with the fraction of gas in dense, self-gravitating regions of cold, supersonic molecular clouds with the log-normal gas density probability density function (PDF) (Elmegreen 2002; Krumholz & McKee 2005; Krumholz et al. 2012). Such models, however, assume that all of the molecular gas is in “virialized” star-forming molecular clouds and that the star formation efficiency in these clouds sets the global depletion time. This assumption, which has also been frequently adopted in galaxy simulations (e.g., Robertson & Kravtsov 2008; Gnedin et al. 2009; Christensen et al. 2012; Kuhlen et al. 2012), is at odds with a growing number of observations indicating that the depletion time of star-forming molecular gas is in general considerably shorter than the global depletion time of all molecular gas, \( \tau_{\text{dep,H}_2} \), estimated on \( \gtrsim \) kpc scales. Moreover, models and simulations of star formation in supersonic turbulent clouds show that the local efficiency of star formation is primarily a strong function of the virial parameter of the region, not just its density, temperature, and molecular fraction (e.g., Krumholz & McKee 2005; Padoan et al. 2012, 2017), while the virial parameter can span a wide range of values (Dobbs et al. 2011; Semenov et al. 2016).

Some models derive the KSR and its normalization by assuming that stellar feedback regulates ISM turbulence so as to maintain vertical and/or Toomre (1964) equilibrium within gaseous disks (Ostriker & Shetty 2011; Faucher-Giguère et al. 2013; Hayward & Hopkins 2017). However, it is not clear a priori why equilibrium can generically be expected in galaxies as a whole or in kiloparsec-scale patches and why the SFR does not instead reach values at which gas is driven out in a wind. Moreover, it is still debated whether the turbulence within galactic disks is mainly driven by stellar feedback or gravitational instabilities (e.g., Krumholz & Burkhardt 2016).

Saitoh et al. (2008) argued that the SFR is controlled by the rate at which gas is supplied from the general ISM to the star-forming state, which makes it insensitive to the local efficiency of star formation. However, these authors measured the timescale at which gas is supplied to the star-forming state to be \( \sim 100 \) Myr and did not explain how this timescale relates to the much longer observed depletion times of total gas, \( \tau_{\text{dep}} \sim 2–10 \) Gyr.

In this paper, we aim to clarify the origin of the observed long gas depletion timescale in galaxies, taking into account both the inefficiency of star formation in star-forming clouds and the fact that not all of the molecular gas is actively forming stars. To this end, in Section 2, we present a simple conceptual framework that views the ISM as a highly dynamic medium in which gas evolves between non-star-forming and star-forming states. The key aspect of our framework is that it considers global gas depletion as a result of gas evolution that is driven by processes with associated characteristic timescales. This approach is conceptually similar to the framework of Kruĳssen & Longmore (2014), developed to explain and interpret the scatter and possible biases in observational measurements of the KSR on different scales.

We illustrate our framework using a realistic simulation of a galactic disk that reproduces the observed depletion time and the KSR. We describe the simulation in Section 3. In Section 4, we demonstrate that the long global depletion time originates from the rapid cycling of ISM gas between non-star-forming and star-forming states on timescales \( \sim 20–100 \) Myr, in accord with the above estimates and the results of previous galactic disk simulations. On each such cycle, only a small fraction of the gas mass is converted into stars; thus, \( \tau_{\text{dep}} \) is long because a large number of such cycles would be required to deplete all available gas. We analyze the processes driving the rapid gas evolution and use simulations to shed light on why the depletion time of molecular gas is nearly independent of the gas surface density. We discuss our results and summarize our conclusions in Sections 5 and 6.

2. A Model for Gas Depletion Time

The interstellar gas in galaxies is a multiphase, dynamic medium spanning several orders of magnitude in density and temperature. To get a sense of the processes affecting the gas evolution in such a medium, we consider the evolution of individual gas parcels, massless tracers of gas flows in the ISM. One can think of a representative set of ISM atoms as such tracers. At any given time, the local environment of such tracers can be estimated by averaging gas properties on some
scale \( l \) around the position of each tracer. The gas around tracers will expand and contract under the influence of dynamical processes, such as turbulence driven by gravitational instabilities and stellar feedback. Therefore, during evolution over a sufficiently long timescale, the conditions around each gas parcel can evolve between the states of long and short depletion time, \( \tau_* \), once or over many transition cycles. Such cycling of a gas parcel is schematically shown in Figure 1.

The probability density per unit time for a parcel to be converted into a star is given by \( 1/\tau_* \), and one can define the depletion time for a single parcel as the time required for the integrated probability to reach unity. In what follows, for conceptual simplicity, we will adopt a sharp threshold, \( \tau_{*,\text{max}} \), separating non-star-forming, \( \tau_* > \tau_{*,\text{max}} \), and actively star-forming, \( \tau_* < \tau_{*,\text{max}} \), gas states. For a given distribution of \( \tau_* \), this threshold can be chosen in such a way that regions with \( \tau_* < \tau_{*,\text{max}} \) include most of the total star formation.

The duration of a single cycle between the successive stages when the parcel’s environment is in the star-forming state equals the sum of the time spent in the non-star-forming, \( \tau_{\text{nsf}} \), and star-forming, \( \tau_{\text{sf}} \), stages. If we denote the average depletion time of the parcel during the star-forming stage as \( \tau_* \), the parcel will have a probability of \( \tau_{\text{sf}}/\tau_* \) to be converted into a star during the entire cycle. In other words, \( N_c = \tau_* / \tau_{\text{sf}} \) such cycles would be needed for the parcel to be incorporated into a star. Its depletion time can thus be written as

\[
\tau_{\text{dep}} = N_c (\tau_{\text{nsf}} + \tau_{\text{sf}}) = N_c \tau_{\text{nsf}} + \tau_* = \left( \frac{\tau_{\text{nsf}}}{\tau_{\text{sf}}} + 1 \right) \tau_*
\]

Hence, the depletion time of a gas parcel is always longer than \( \tau_* \), and it can be long because star formation during \( \tau_{\text{sf}} \) is inherently inefficient, i.e., \( \tau_* \) is long, and/or because only a small fraction of the evolution cycle is spent in the star-forming state, i.e., \( \tau_{\text{nsf}}/\tau_{\text{sf}} \) is large.

A group of parcels in a patch of the ISM has a distribution of \( \tau_*, \tau_{\text{nsf}}, \tau_{\text{sf}} \), and the collective depletion time of the group is the average \( \tau_{\text{dep}} = (1/\tau_{\text{dep}})^{-1} \) over these distributions. It is clear that if the depletion times of individual parcels are long, \( \tau_{\text{dep}} \) will also be long.

In practice, the \( \tau_{\text{dep}} \) of an ISM patch is estimated from the instantaneous mass of gas, \( M_g \), and young stars formed over a time interval \( \Delta t \), \( \langle M_* \rangle_{\Delta t} \), such that the average SFR is \( \langle M_* \rangle_{\Delta t} = M_* / (\langle <\Delta t \rangle_{\Delta t}) \) and the depletion time is defined as \( \tau_{\text{dep}} = M_g / \langle M_* \rangle_{\Delta t} \). This estimate of \( \tau_{\text{dep}} \) can be related to the dynamics of individual gas parcels within the patch by noting that due to mass conservation, the instantaneous total mass of the gas parcels in the star-forming state, \( M_{sf} \), evolves as \( M_{sf} = M_{sf} - M_* \), where \( F_* \) is the net instantaneous flux of the gas parcels through the star formation threshold and \( M_* \) is the instantaneous SFR of all parcels. After averaging this expression over the time interval \( \Delta t \), we get

\[
\tau_{\text{dep}} = \frac{M_g}{\langle M_* \rangle_{\Delta t}} = \frac{M_g}{\langle F_* \rangle_{\Delta t} - \langle M_* \rangle_{\Delta t}}.
\]

For brevity, in the following derivation, we omit explicit averaging, \( (\ldots)_{\Delta t} \), but assume all fluxes and rates to be averaged over \( \Delta t \).

In general, the average net flux of gas through the star formation threshold can be decomposed into positive and negative contributions,

\[
F_* = F_+ - F_-
\]

which correspond to the supply and removal of star-forming gas illustrated in Figure 1.

The positive flux \( F_+ \) is controlled by a combination of global dynamical processes, e.g., gravitational instabilities, turbulence, cooling, etc., with a significant contribution from stellar feedback. The latter comes in the form of turbulence stirred by interactions of supernova-driven bubbles and by fountain outflows of gas from star-forming regions. The negative flux \( F_- = F_{-\text{fb}} + F_{-\text{d}} \) results from the destruction of star-forming regions both by feedback from young stars formed inside the regions, \( F_{-\text{fb}} \), and by shearing due to large-scale turbulence or differential rotation, expansion of gas behind galactic spiral arms, and other dynamical processes not directly related to star formation inside the regions, \( F_{-\text{d}} \). All fluxes can be parameterized with the characteristic timescales, i.e., \( \tau_{\tau}, \tau_{-\text{fb}}, \) and \( \tau_{-\text{d}} \), on which gas is supplied to and removed from the star-forming state by the corresponding processes,

\[
F_+ \equiv \frac{M_{\text{nsf}}}{\tau_{\tau}} = \frac{M_g}{\tau_{\tau}} \left( 1 - \frac{f_{\text{sf}}}{\tau_{\tau}} \right),
\]

\[
F_- \equiv \frac{M_{\text{sf}}}{\tau_{-\tau}} = \frac{M_g}{\tau_{-\tau}} \left( \frac{1}{\tau_{-\text{fb}}} + \frac{1}{\tau_{-\text{d}}} \right),
\]

where \( f_{\text{sf}} = M_{\text{sf}}/M_g \) is the star-forming mass fraction.

To make the relation between star formation and stellar feedback explicit, we can also parameterize \( F_{-\text{fb}} \) in a way similar to the parameterization of the mass outflow rate of feedback-driven galactic winds,

\[
F_{-\text{fb}} \equiv \frac{\xi}{\tau_{-\tau}} = \frac{M_{\text{fb}}}{\tau_{-\tau}} \frac{\xi}{\tau_{\tau}},
\]

where \( \xi \) is the mass-loading factor and we used the definition of the average depletion time of star-forming gas, \( \tau_{\tau} \equiv M_{\text{sf}}/M_* \). In the context of Equation (6), the mass-loading factor can also be interpreted as a relative rate of gas removal by feedback compared to the rate of star formation, i.e., \( \xi \equiv \tau_*/\tau_{-\tau} \).
An imbalance between the net gas flux into the star-forming state, $F_{sf}$, and the average SFR may result in the evolution of the star-forming mass, which we also parameterize with the characteristic timescale, $\tau_{e, sf}$:

$$|M_d| \equiv \frac{M_d}{\tau_{e, sf}}. \quad (8)$$

The final expression for the global depletion time can be readily derived by substituting Equations (4)–(8) and $f_{sf} = \tau_*/\tau_{dep}$ into Equation (3),

$$\tau_{dep} = \left(1 + \xi + \frac{\tau_*}{\tau_{-d}} \pm \frac{\tau_*}{\tau_{e, sf}}\right) \tau_* + \tau_*, \quad (9)$$

where the sign in front of $\tau_*/\tau_{e, sf}$ reflects the sign of $M_d$.

If we compare the terms in this equation with those in Equation (2) for the depletion time of a single gas parcel, $t_{dep} = N_t \tau_{int} + \tau_*$, their physical meaning becomes clear. The timescale $\tau_{e, sf}$ is analogous to $\tau_{int}$, i.e., the average number of evolution cycles it would take for a single parcel to deplete its gas. Indeed, Equation (5) gives $M_{dep} = F_{sf} \tau_*$, which means that $\tau_*$ is the time over which all of the non-star-forming gas will reach the star-forming state. Thus, $\tau_{e, sf}$ is analogous to the average $t_{int}$ timescale for a collection of parcels. Likewise, the average rate at which the gas mass in the non-star-forming state is decreasing due to star formation, dispersal, and the overall evolution of the gas PDF during $\Delta t$ is given by $M_d/\tau_* + M_d/\tau_{-d, sf}. \quad (10)$

where we used Equation (6) and the definition of the mass-loading factor, $\xi \equiv \tau_{-d}/\tau_{-fb}$. $\quad (10)$

Equation (9) is the key expression of our framework. It states that the global depletion time is the sum of the total time that gas spends in the non-star-forming state over $N_c$ cycles and the total time over which star-forming regions convert this gas into stars, $\tau_*$. $\quad (10)$

This equation elucidates how long $\tau_{dep}$ values can be reconciled with the relatively short local depletion times, $\tau_*$, and even shorter dynamical timescales, $\tau_{-fb}$, discussed in the Introduction. The global depletion time is longer than the depletion time in star-forming regions, $\tau_{int}$, due to the significant fraction of time that gas spends in the non-star-forming state. $\quad (10)$

The global depletion time is longer than the timescale associated with dynamical processes supplying star-forming gas, $\tau_*$, because gas must evolve through the non-star-forming state $N_c$ times, and $N_c$ is large due to either efficient feedback, i.e., large $\xi$, or fast dynamical processes destroying star-forming regions, i.e., short $\tau_{-d, sf}$ (see Equation (10)). $\quad (10)$

When feedback dominates the removal of gas from the star-forming state, the number of cycles becomes $N_c \sim \tau_*/\tau_{-fb}$. $\quad (10)$

This clarifies how feedback can self-regulate star formation, i.e., how $t_{dep}$ can become insensitive to $\tau_*$. Indeed, the timescale $\tau_{-fb}$ is proportional to the rate of energy and momentum injection by feedback, which, in turn, is set by the local rate of star formation, i.e., $\tau_*$. Hence, $\tau_{-fb} \propto \tau_*$, which renders $N_c$ insensitive to $\tau_*$. Thus, when $N_c \tau_* \gg \tau_*$, the depletion time, $\tau_{dep} \approx N_c \tau_*$, will be insensitive to $\tau_*$. $\quad (10)$

In a nonequilibrium state, in which $M_d > 0 (<0)$ during $\Delta t$, the term $\pm \tau_*/\tau_{e, sf}$ in Equation (9) accounts for the correction of the average rates estimated using the star-forming gas fraction, $f_{sf}$, defined for the instantaneous masses $M_d$ and $M_d$. This correction appears because, when $M_d > 0 (<0)$, the actual average fraction of $\Delta t$ that gas spends in the star-forming state is smaller (higher) than $f_{sf}$ and therefore more (fewer) transition cycles are required for depletion. $\quad (10)$

In a steady state, by contrast, the gas distribution is stationary and the SFR is in equilibrium with the gas fluxes into and out of the star-forming state: $M_d = F_{sf} - M_\odot \approx 0$. In this case, $\tau_{e, sf} \rightarrow \infty$, and the term $\tau_*/\tau_{e, sf}$ can be neglected in Equation (9). In such a steady state, $\tau_{dep} = M_\odot / F_{sf}$ (see Equation (3)), and depletion time is determined by the net rate of gas inflow into the star-forming state, $F_d$. When $F_d$ is small, the depletion time is long. Galaxies as a whole reach the steady state with $M_d \approx 0$ on the shortest of the timescales that control the global depletion time in Equation (9). Thus, globally, such an assumption is justified. However, individual ISM patches may deviate from the steady state, and the $\tau_*/\tau_{e, sf}$ term will be one of the sources of the scatter in depletion times. $\quad (10)$

In the remainder of the paper, we illustrate the framework described above using the results of an isolated galaxy simulation. Although the simulation adopts specific choices for many parameters, including resolution and prescriptions for star formation and feedback, the overall features and implications of our model do not depend on these specific choices. Our framework generically allows one to relate the depletion time on a large scale, e.g., the scale of an entire galaxy, to the star formation and feedback model that operates on a smaller scale, e.g., a resolution scale of a simulation, where the distribution of local depletion times, $\tau_*$, is defined. $\quad (10)$

3. Simulation $\quad (10)$

To illustrate the framework outlined above and elucidate the physical processes that give rise to long global depletion times, we use a simulation of an isolated $\sim L_\odot$-sized galaxy. We carried out the simulation with the adaptive mesh refinement $N$-body and gas dynamics code ART (Kravtsov 1999; Kravtsov et al. 2002; Rudd et al. 2008; Gnedin & Kravtsov 2011) and followed the evolution of an isolated gaseous disk in a live potential of a dark matter halo, stellar bulge, and stellar disk that are modeled with collisionless particles. $\quad (10)$

We adopt the initial conditions that were used in the AGORA code comparison project (Kim et al. 2016) and in the studies of Agertz et al. (2013) and Semenov et al. (2016). Specifically, the isolated disk is initialized inside a dark matter halo with $v_{c,200} = 150$ km s$^{-1}$ and an initial concentration of $c = 10$. The initial disk of old stars has an exponential density profile with a radial scale length of $r_0 \approx 3.4$ kpc and a vertical scale height of $h_\odot = 0.1r_0$ with a total mass of $M_{e, d} \approx 3.4 \times 10^9 M_\odot$. The stellar bulge has an initial mass of $M_{s, b} \approx 4.3 \times 10^9 M_\odot$ that is distributed with a Hernquist density profile with $a = 0.1r_0$ (Hernquist 1990). The initial exponential gaseous disk has the same $r_0$ and $h_\odot$ as the stellar disk; its total mass is $M_d \approx 8.6 \times 10^9 M_\odot$, which corresponds to the disk gas fraction of $f_{gas} \equiv M_d/(M_{e, d} + M_d) \approx 20\%$. $\quad (10)$

Gas evolution is governed by modified hydrodynamical equations that include terms related to cooling and heating, dynamical effects of subgrid turbulence, gas consumption by
star formation, and injection of mass, momentum, and energy by feedback from young stars.

Cooling in the optically thin limit is implemented following the model of Gnedin & Hollon (2012). We assume a fixed metallicity of $Z = Z_\odot$ and constant background heating by interstellar radiation in the Lyman–Werner band with a photodissociation rate of $10^{-10} \, \text{s}^{-1}$ (Stecher & Williams 1967).

To model temperatures in dense self-shielded gas, we assume that extinction is proportional to the local column density of atomic gas, which we approximate as $n L_{24,0}$, where $n$ is the gas number density in a cell and $L_{24,0}$ is the local Jeans length with an applied temperature ceiling of 40 K (model “L1a” in Safranek-Shrader et al. 2017).

In our simulation, we adopt the unresolved turbulence model of Schmidt et al. (2014) that dynamically follows the kinetic energy of gas motions on subgrid scales, $K$, as a separate hydrodynamical field akin to thermal energy. Subgrid turbulence is generated by random velocities on the resolution scale; it exerts pressure and viscous forces on resolved gas motions and decays on the local cell-crossing timescale, as motivated by simulations of both subsonic and supersonic turbulence (e.g., Mac Low et al. 1998). For details about the model implementation in ART, we refer the reader to Semenov et al. (2016).

We adaptively resolve cells where the total gas mass exceeds $\sim 8300 \, M_\odot$ and reach a maximum resolution of $\Delta = 40$ pc. Such a $\Delta$ is sufficient to resolve ISM structure down to densities of $n \sim 100$–$1000 \, \text{cm}^{-3}$, and therefore we do resolve the dynamical processes that are sometimes claimed to limit the star-forming gas supply from the general ISM with an average density of $n \sim 1 \, \text{cm}^{-3}$. At the highest resolution level, we do not apply an artificial pressure floor in cold gas. Thus, the densities of star-forming regions are limited only by the effects of stellar feedback and the effective pressure due to thermal and both subgrid and resolved turbulent motions.

As a star formation prescription, we adopt the local rate $\dot{\rho}_s = \rho / t_s$, where $\rho$ is the gas density in a cell, and the local depletion time $t_s$, is related to the freefall time, $t_{ff} = (3\pi/2G\rho)^{1/2}$, with the efficiency per freefall time, $t_s = t_{ff} / \epsilon_{ff}$.

Numerical and analytical models of star formation in turbulent giant molecular clouds (GMCs) generally predict a strong increase of $\epsilon_{ff}$ with an increasing relative importance of gravity that facilitates star formation, in comparison to turbulence that provides support against gravity (for a review, see Padoan et al. 2014). For example, Padoan et al. (2012) found that the star formation efficiency of a turbulent cloud increases exponentially with a decreasing virial parameter, $\epsilon_{ff} \approx \exp(-\sqrt{\alpha_{vir}/0.53})$, where $\alpha_{vir}$ is defined for a box with a side $\Delta$ as for a uniform sphere of radius $R = \Delta/2$:

$$\alpha_{vir} \equiv \frac{5\sigma_{1D}^2 R}{GM} \approx 9.35 \left(\frac{\sigma_{rot}/10 \, \text{km s}^{-1}}{n/100 \, \text{cm}^{-3}}(\Delta/40 \, \text{pc})^2\right).$$

In our simulations, to apply the Padoan et al. (2012) fit in thermally supported gas, in the definition of $\alpha_{vir}$ we consider contributions of both the sound speed, $c_s$, and the explicitly modeled subgrid turbulent velocities, $\sigma_t \equiv \sqrt{2K/\rho}$:

$$\sigma_{rot} = \sqrt{\sigma_c^2 + c_s^2}.$$  \hspace{1cm} (12)

Also, even though we are able to model $\epsilon_{ff}$ following the Padoan et al. (2012) formula (see, e.g., Semenov et al. 2016; Li et al. 2017), in the simulation used here, we approximate the continuous exponential dependence of $\epsilon_{ff}$ on $\alpha_{vir}$ assuming a constant $\epsilon_{ff} = 1\%$ for $\alpha_{vir} < \alpha_{vir, sf} = 10$ and $\epsilon_{ff} = 0$ elsewhere. This simulation was carried out as part of a simulation suite in which the $\epsilon_{ff}$ value was varied systematically to explore its effect on star formation in galaxies. Using a constant $\epsilon_{ff}$ and a sharp $\alpha_{vir, sf}$ threshold makes the interpretation of simulation results easier, and we will present the results of this simulation suite in a forthcoming study. We explicitly checked that the global depletion times and the KS radii are similar in runs where $\epsilon_{ff}$ follows the Padoan et al. (2012) fit and we approximate this fit with a threshold. A qualitatively similar star formation prescription but with a different choice of parameters was studied by Hopkins et al. (2013a).

We stress that the scenario of gas depletion described in Section 2 remains valid for any choice of star formation prescription, although in Section 5.2 we argue that such a prescription should be chosen carefully, as it is important for the prediction of realistic ISM properties. We note that our adopted threshold value, $\alpha_{vir, sf} = 10$, is consistent with the observed distribution of $\alpha_{vir}$ in molecular clouds with sizes comparable to our resolution of $\Delta = 40$ pc: such clouds tend to have $\alpha_{vir} \lesssim 10$ (Miville-Deschênes et al. 2017). The adopted value of $\epsilon_{ff} = 1\%$ is consistent with the average values deduced for observed actively star-forming clouds.

In our feedback prescription, we inject momentum and energy from the supernovae (SN) type II at a uniform rate between 3 and 43 Myr after the formation of a stellar particle. The total number of SN exploded over this time interval is computed assuming the Chabrier (2003) initial mass function. The thermal energy and radial momentum injected by each supernova are calibrated against simulations of an SN remnant evolution in a nonuniform medium (Martizzi et al. 2015), taking into account ambient gas density and with an additional boost of radial momentum by a fiducial factor of 5. We adopt such a boosting factor to compensate for the numerical loss of injected momentum and to account for the actual physical uncertainties of the radial momentum estimates (e.g., Gentry et al. 2017). In addition to SN type II feedback, during evolution, stellar particles return a fraction of their mass following the prescription of Leitner & Kravtsov (2011).

Figure 2 shows the spatial distribution of gas number density, temperature, and subgrid turbulent velocity in our simulated galaxy at $t = 500$ Myr. The figure highlights the multiphase, dynamic structure of the ISM. Comparison with a simulation with no feedback, which we will present in a companion paper, shows that the structure of the ISM is significantly affected by stellar feedback. Its effect is manifested not only in the ubiquitous regions of hot, turbulent gas but also in the overall flocculent nature of the spiral pattern.

In our analysis, we estimate the molecular gas density in each cell using the model of Krumholz et al. (2008, 2009a); McKee & Krumholz (2010). $\rho_{HI} = \max(0, (1 - 0.75s/(1 + 0.25s))\rho)$, where at solar metallicity $s \approx 1.8/\tau_c$ and $\tau_c = 320/(\Delta/g \, \text{cm}^{-2})$. This model predicts a rapid increase of the molecular fraction at densities $n > 10 \, \text{cm}^{-3}$, indicated by the black contour in Figure 2. The resulting total mass fraction of molecular gas in our simulation is $f_{HI} \equiv M_{HI}/M_{\odot} \approx 20\%$.

Subgrid turbulent velocities, $\sigma_t$, range from $\lesssim 3 \, \text{km s}^{-1}$ in the diffuse ISM between the spiral arms to $\sim 30$–$300 \, \text{km s}^{-1}$ in hot SN bubbles. In this simulation, supernovae do not explicitly inject turbulent energy, and high $\sigma_t$ in hot bubbles is generated...
by the subgrid turbulence model. In the cold dense gas, turbulent velocities are supersonic and vary significantly, \( \sigma_t \approx 5-15 \, \text{km} \, \text{s}^{-1} \). Strong subgrid turbulence in cold gas results in high values of \( \alpha_{\text{vir}} \) and, according to our star formation criterion, \( \alpha_{\text{vir}} < \alpha_{\text{vir,sf}} = 10 \), only \( \approx 40\% \) of all molecular gas mass is star-forming at any given moment. Such star-forming regions are shown in the left panel of Figure 2 with green contours.

The total SFR of our model galaxy is \( M_* \approx 1-2 \, M_\odot \, \text{yr}^{-1} \), which translates to global depletion times for the total and molecular gas of \( \tau_{\text{dep}} \approx 4-8 \) and \( \tau_{\text{dep,H}_2} \approx 1-2 \) Gyr. The values of the depletion times and molecular fraction are in the ballpark of the typical values observed in nearby spiral galaxies (e.g., Wong & Blitz 2002; Bigiel et al. 2008; Leroy et al. 2013).

Figure 3 compares the KSR between the surface density of the SFR and the total (left panel) and molecular (right panel) gas in our simulation and in observations. To match the typical spatial scales on which this relation is usually measured, we smooth 2D maps of \( \Sigma_* \), \( \Sigma_{\text{H}_1,\text{H}_2} \), and \( \Sigma_{\text{H}_2} \) obtained at 500 Myr with a Gaussian filter with a width of 1 kpc and plot the median, 16th, and 84th percentiles of the resulting pixel distribution (thick and thin dark lines on top of colored distributions). The distributions are colored according to the average galactic radius, \( R \), of pixels in a given bin. To match the averaging timescales of the star formation indicators, we measure the surface density of stars that are younger than 30 Myr and define \( \Sigma_* \) as \( \Sigma_* (<30 \, \text{Myr})/30 \, \text{Myr} \). The thin red line at high \( \Sigma \) indicates the slope adopted in our star formation prescription, \( \rho_* \propto \rho^{1.5} \). Thin dotted lines correspond to the constant depletion times of 0.1, 1, and 10 Gyr (from top to bottom). We compare our results to the observed relations in nearby spiral galaxies (Bigiel et al. 2008, 2010; Leroy et al. 2013) and in the Milky Way (Misiriotis et al. 2006). The violet contour in the right panel shows the range of \( \tau_{\text{dep,H}_2} \approx 0.5-2 \) Gyr estimated using the radial profiles of \( \Sigma_* \) and \( \Sigma_{\text{H}_2} \) for the Milky Way from Figure 7 in Kennicutt & Evans (2012).

Figure 2. Midplane slices of gas number density, \( n \), temperature, \( T \), and subgrid turbulent velocity, \( \sigma_t = \sqrt{2 K/\rho} \), after 500 Myr of evolution. To make comparison easier, the black contours in all panels correspond to \( n = 10 \, \text{cm}^{-3} \), above which the molecular mass fraction rapidly increases at solar metallicity. Green contours in the left panel indicate gas that satisfies our star formation criterion, \( \alpha_{\text{vir}} < \alpha_{\text{vir,sf}} = 10 \).
For a consistent comparison with the observed KSR for H\textsubscript{1} + H\textsubscript{2} gas, in our simulation, we defined neutral hydrogen as all nonmolecular gas denser than n\textsubscript{H,SSS}, given by Equation (13) in Rahmati et al. (2013). This threshold corresponds to the gas self-shielded from the far-ultraviolet (FUV) background with the adopted photoionization rate $\Gamma = 10^{-16}$ s\textsuperscript{-1}. We also excluded all neutral hydrogen that is colder than 1000 K, assuming that it constitutes the optically thick cold neutral medium (CNM) not included in the observed measurements of $\Sigma_{\text{H}1}$. Our temperature threshold is somewhat higher than the CNM temperature estimated in real galaxies ($\lesssim$300 K; e.g., Wolriffe et al. 2003) because, in our simulation, we do not resolve the transition between warm and cold neutral gas phases, which results in intermediate gas temperatures on the resolution scale. The particular value of the temperature threshold was chosen to select $\sim$40% of the neutral hydrogen mass, which is close to the CNM mass fractions estimated in the Milky Way and nearby galaxies (e.g., Heiles & Troland 2003; Braun 2012; Pineda et al. 2013; Sofue 2017).

In the analyses presented below, we consider the processes in the ISM during the time interval between 400 and 600 Myr, which is short enough to neglect the effect of global gas consumption on the ISM dynamics and long enough to accumulate sufficient statistics of rare events and make the considered distributions representative and well-sampled.

To study the detailed dynamics of individual gas parcels, we use gas tracer particles that are passively advected with the local flow velocity interpolated to the positions of the particles with the cloud-in-cell scheme. We populate the disk with 10\textsuperscript{5} tracer particles uniformly initialized within $R < 8$ kpc after 300 Myr of disk evolution when the transients related to the initial off-equilibrium state had dissipated away. After initialization, we wait for 100 Myr to let the tracers equilibrate with the gas density distribution. At that point, the distributions of tracer densities in radial annuli resemble the gas density PDF in computational cells, and the distribution of tracer properties thus can be considered to be a good approximation of the mass-weighted PDF of gas properties.

We average the distribution of tracers to construct statistics, such as their PDF and fluxes in the n–$\sigma_{\text{tot}}$ phase diagram, between 400 and 600 Myr with 1 Myr steps. We checked that at every moment between 400 and 600 Myr, phase distributions of gas and tracer particles resembled their averaged versions, which means that the galaxy remained in approximate equilibrium over the considered period of time.

We focus on the evolution of gas in the phase plane of density, n, and total velocity dispersion, $\sigma_{\text{tot}}$, because the position of a gas parcel in this plane determines its internal consumption time, $\tau_s$, according to our star formation prescription. To quantify gas motions in the n–$\sigma_{\text{tot}}$ plane, we measure the derivatives $d \log n/ dt$ and $d \log \sigma_{\text{tot}}/ dt$ probed by each tracer every 1 Myr. To estimate the average local flow rates of gas, we accumulate fluxes corresponding to these derivatives and normalize them by the local density of tracers in the n–$\sigma_{\text{tot}}$ plane.

To characterize actual fluxes that supply and remove star-forming gas, in addition to the total flux of tracers, we separately track the fluxes of tracers with decreasing or increasing $\alpha_{\text{vir}}$. We quantify the magnitudes of these fluxes with the characteristic evolution timescale, on which $\alpha_{\text{vir}}$ changes by an order of magnitude at a given rate,

$$\tau_{\alpha,>0} \equiv \left( \frac{d \log_{10} \alpha_{\text{vir}}}{dt} \right)_{>0}^{-1},$$

(13)

where we average the derivative of $\log_{10} \alpha_{\text{vir}}$, taking into account only tracers with decreasing (increasing) $\alpha_{\text{vir}}$ to compute $\tau_{\alpha,>0}$ ($\tau_{\alpha,<0}$).

4. Results

4.1. The Origin of Long Global Depletion Times

As we discussed at the end of Section 2, in steady state, when $\langle M_f \rangle_{\Delta t} \approx 0$ on the timescale $\Delta t$ over which the SFR is estimated, long global depletion times, $\tau_{\text{dep}} \equiv \dot{M}_g/M_* = M_g/F_{\text{sd}}$, originate from a small net flux of gas into the star-forming state, $F_{\text{sd}}$. In principle, $F_{\text{sd}}$ could be small if the rate at which gas evolves toward the star-forming state was set by a slow “bottleneck” process. However, as Figure 4 shows, in simulations with efficient feedback, gas rapidly transitions between the star-forming and non-star-forming states, and a small $F_{\text{sd}}$ results from a near-cancelation of large opposite fluxes into and out of the star-forming state.

In this figure, we plot the distribution of gas tracer particles within the disk in the plane of gas number density, n, and $\sigma_{\text{tot}} = \sqrt{\sigma_\|^2 + \sigma_\perp^2}$, that can be viewed as an effective temperature including both thermal and turbulent gas motions on subgrid scales. The gas distribution spans a wide range of densities, $\sigma_{\text{tot}}$, and temperatures and has two distinct peaks. The peak at low densities, $n \sim 1$ cm\textsuperscript{-3}, corresponds to diffuse, warm, subsonic ($n \lesssim c_s$) gas at temperature $T \sim 10^4$ K. The gas in the second peak at $n > 10$ cm\textsuperscript{-3}, by contrast, is cold ($T \lesssim 100$ K) and supersonic ($n > c_s$).

According to our star formation prescription, the star-forming gas has $\alpha_{\text{vir}} < \alpha_{\text{vir, sf}} = 10$. Such gas in Figure 4 resides below the thick green dotted line. The net mass flux of gas in the n–$\sigma_{\text{tot}}$ plane is visualized by the arrows in the top panel of Figure 4, where the length of the arrows is equal to the distance tracers would traverse in 5 Myr for a given flux. The arrows are rather small throughout most of the phase space occupied by tracers and are particularly small near the thick green dotted line. This means that the net evolution of gas in the n–$\sigma_{\text{tot}}$ plane is slow and the net flux through the star formation threshold, $F_{\text{sd}}$, is small. This small net flux results in the long global depletion timescales exhibited by our simulated galaxy, $\tau_{\text{dep}} \sim 5$ Gyr and $\tau_{\text{dep,H}_2} \sim 1$ Gyr (see Section 3 and Figure 3).

However, the middle and bottom panels of Figure 4 show that the small net $F_{\text{sd}}$ results from the near-cancelation of two opposite fluxes. These panels show the fluxes of only those tracers in which $\alpha_{\text{vir}}$ is decreasing, $F_-$, or increasing, $F_+$, and these fluxes are significantly stronger than the net flux in the top panel. A typical tracer evolves toward and away from the star-forming state on a timescale of order $\tau_{\alpha,>0} \sim 5–30$ Myr, consistent with the estimates of the timescales of relevant processes in Section 1. Thus, the rate of gas supply from the diffuse, warm ISM to the star-forming state cannot be the factor limiting the global SFR, as envisioned by Saitoh et al. (2008). Instead, gas generally evolves from the diffuse to the star-forming state on a timescale of tens of Myr, much shorter than the global depletion time. The latter is long because gas rapidly leaves the star-forming state at the rate that nearly cancels the rate at which gas is reaching this state. In
of individual tracer particles. The particular shapes of such trajectories vary depending on local conditions and specific physical processes that govern gas evolution.

In our simulation, gas evolution between diffuse, warm, subsonic and dense, cold, supersonic ISM phases is governed by large-scale disk instabilities and the turbulent flows generated by them. The evolution of gas in the dense phase is determined predominantly by stellar feedback that disperses star-forming regions, drives large-scale ISM turbulence, and launches fountain-like outflows.

In the following subsections, we consider these processes using three illustrative tracer trajectories integrated over 200 Myr and shown in Figure 5. We chose these particular tracers because their evolution over the considered period of time is governed predominantly by the same processes over several consequential cycles of compression and expansion.

4.2.1. Compression and Expansion Due to Spiral Arms

The blue line in the top left panel of Figure 5 shows an example of the evolution followed by a tracer that swings between the subsonic and supersonic phases during cycles of compression and expansion as it enters and exits the spiral arms. The trajectory of this tracer in the $n$-$\sigma_{\text{tot}}$ plane is shown with the same color in the right panel.

Due to strong compression, gas entering a spiral arm rapidly cools and loses the thermal support that initially dominates in the subsonic ISM. At the same time, initially low subgrid turbulent velocities of the subsonic ISM, $\sigma_s \lesssim 3 \kms$, rapidly grow due to compressional heating (see the detailed discussion in Section 3.1 of Semenov et al. 2016 and also Robertson & Goldreich 2012). At $n \sim 10 \cm^{-3}$, when subgrid turbulent velocities become comparable to the thermal speed, gas detaches from the lower envelope of the distribution shown in Figure 4 and enters the supersonic ISM phase. Similarly, when gas leaves a spiral arm, it expands, and subgrid turbulent velocities decrease. Eventually, under the influence of expansion and interstellar FUV heating, gas returns to the subsonic ISM phase with $n \sim 1 \cm^{-3}$ and $T \sim 10^4 \K$.

The actual transition of gas between the subsonic and supersonic phases is fast, as it is controlled by a strong compression rate in the spiral arms and short cooling times at $n > 1 \cm^{-3}$. Hence, the rate at which diffuse gas is promoted into the dense phase is mostly determined by the time that gas waits between subsequent passages of the spiral arms,

$$\tau_{\text{arm}} \sim \frac{2\pi R}{mV_{\text{gas}}} \sim 80 \Myr \frac{(R/8 \kpc)}{(m/6)(V_{\text{gas}}/100 \kms)},$$

where $V_{\text{gas}} \equiv v_{\text{gas}}-v_{\text{pat}}$ is the speed of gas relative to the spiral wave pattern, and we set $m = 6$, as our simulated galaxy develops six spiral arms.

The typical time that gas spends inside a spiral arm before expansion contributes to the dynamical rate of gas removal from the star-forming state, $\tau_{\text{...}}$. This timescale depends on the spiral arm width, gas velocity, and angle at which gas flows inside the arm. Depending on local conditions, this timescale can be as long as a few tens of Myr.

In Figure 6, we plot distributions of $n$ and $\sigma_{\text{tot}}$ separately for tracers residing at different galactic radii and therefore experiencing different ISM conditions. The distribution in the outer disk (top panel) is shaped predominantly by the compression and expansion due to the spiral arms. Specifically,
most of the gas mass in the outer disk resides in the diffuse subsonic phase and forms a peak at \( n \sim 1 \text{ cm}^{-3} \) and \( T \sim 10^4 \text{ K} \). The tail extending along the lower envelope of the distribution toward the dense supersonic phase corresponds to the gas currently being compressed in the spiral arms. As the figure also shows, the compression of diffuse gas in the spiral arms is only relevant at large radii, whereas closer to the disk center, less gas remains in the diffuse phase, and this process becomes much less important.

4.2.2. SN-induced Shocks and ISM Turbulence

We find that the evolution of dense, supersonic gas in the \( n-\sigma_{\text{tot}} \) phase is driven by feedback from SNe in a star-forming region. Injection of momentum by SNe in an ISM at these radii has a complex structure (see Figure 2) reflecting the chaotic nature of the gas.

Supernova feedback also affects some of the gas by accelerating it in the direction perpendicular to the disk plane. Such gas expands in fountain-like outflows but eventually cycles back to the ISM under the influence of the disk potential. Interactions of such outflows with the halo gas adjacent to the disk result in an increase of small-scale turbulent velocities that quickly dissipate when the gas falls back onto the disk.

An example of a tracer trajectory during expansion and subsequent recycling of a fountain outflow is shown with the red lines in Figure 5. This particular tracer was ejected and recycled twice, at \( \sim 20-80 \) and \( \sim 130-150 \) Myr. In each event, after its star-forming region was dispersed by feedback, this tracer acquired a moderate vertical velocity of \( v_z \sim 50 \text{ km s}^{-1} \) and elevated as high as \( \sim 400 \text{ pc} \) above the disk plane, i.e., a few scale heights, before falling back onto the disk. At the highest elevation point, the gas in these outflows expands only to the densities comparable to those of the diffuse subsonic ISM phase, \( n \sim 0.2-2 \text{ cm}^{-3} \), and its virial parameter reaches the values of \( \alpha_{\text{vir}} \sim 10^{-3}-10^{-4} \) due to the strong turbulence generated by the interaction of the expanding outflow with the surrounding gas. Outflows launched by feedback from regions of more vigorous star formation reach even lower \( n \) and higher \( \alpha_{\text{vir}} \).

The timescale of the fountain cycle can be estimated as a dynamical time in the gravitational field of a massive infinite sheet of constant surface density \( \Sigma_{\text{tot}} \), corresponding to the local total surface density of the disk,

\[
\tau_{\text{grav}} \sim \frac{v_z}{\pi G \Sigma_{\text{tot}}} \sim 20 \text{ Myr} \left( \frac{v_z}{50 \text{ km s}^{-1}} \right) \left( \Sigma_{\text{tot}}/200 \ M_\odot \text{ pc}^{-2} \right),
\]

where \( v_z \) is the initial vertical velocity of gas in the outflow and \( \Sigma_{\text{tot}} = \Sigma_s + \Sigma_g \sim 200 \ M_\odot \text{ pc}^{-2} \) is the typical total surface density of gas and stars in our simulated galaxies.
along the upper envelope of the distribution shown in the top panel. Gas in such outflows at large radii usually returns to the diffuse, warm, subsonic ISM between the spiral arms. Closer to the center, outflows are ubiquitous, and, after falling back, their gas directly rejoins the tumultuous large-scale turbulent motions of dense, supersonic gas.

### 4.3. Duration and Number of Evolution Cycles

Typical tracer trajectories considered in the previous section explicitly confirm that during the evolution, gas parcels perform many fast cycles and rapidly explore a significant portion of the PDF extent, frequently switching between non-star-forming and star-forming states. As we discussed in Section 2, the distribution of the time that the gas parcels spend on each cycle in these states, \( t_{\text{nsf}} \) and \( t_{\text{sf}} \), determines the global depletion time of the galaxy.

In Figure 7, we plot the distribution of \( t_{\text{nsf}} \) and \( t_{\text{sf}} \) directly measured from the trajectories of all tracers as the time between consequential crossings of the star formation threshold. The results in the previous section indicate that the mix of processes governing gas evolution may change with the galactic radius, \( R \), and surface density, \( \Sigma_g \). Thus, to explore possible trends, we plot the distributions of timescales as a function of \( R \) and \( \Sigma_g \).

The distribution of \( t_{\text{nsf}} \) shown in the top panel indicates that the majority of tracers spend less than 100 Myr in the non-star-forming stage of evolution during each cycle. At higher average surface densities closer to the disk center, this time is even shorter, \( t_{\text{nsf}} \lesssim 50 \) Myr, with a very low median value (thick red line in the top panel of Figure 7).

Colors in the top panel show the average fraction of time that gas tracers spend in the dense phase, \( n > 10 \text{ cm}^{-3} \), over the non-star-forming stage of evolution. A blue color at small radii implies that gas preferentially stays in the dense, molecular phase even when it does not form stars. This is also evident from the bottom panel of Figure 6, which shows that only a small fraction of gas expands to \( n < 10 \text{ cm}^{-3} \) and does so as a part of fountain outflows.

At larger radii, the relatively slow rate of star-forming gas replenishment via compression in the spiral arms becomes important, and the median \( t_{\text{nsf}} \) increases to \( \sim 80 \) Myr. Gas governed by this process spends significant time in the diffuse subsonic ISM, and such tracer trajectories occupy the areas of the longest \( t_{\text{nsf}} \) at \( R > 4 \) kpc (green color in Figure 7). However, as indicated by the blue color, at such radii many tracers still perform short cycles with \( t_{\text{nsf}} < 50 \) Myr without leaving the dense phase.

The increase of the \( t_{\text{nsf}} \) median value is consistent with the scaling proportional to the freefall time at the mean or midplane density at a given radius, \( t_{\text{ff,0}} \propto \rho_0^{1/2}/(\Sigma_0/2h_0)^{1/2} \), shown by the dotted line in Figure 7 (see also Saitoh et al. 2008). Such scaling is sometimes adopted in analytical models of galactic star formation to define the timescale on which star-forming regions are created (e.g., Krumholz et al. 2012; Elmegreen 2015). As we discussed above, \( t_{\text{nsf}} \) in our simulations is set by both stellar feedback that drives turbulence and dynamical processes within the ISM. The scaling of the median \( t_{\text{nsf}} \) with density indicates that gravity and the associated timescale plays at least some role in setting the time that gas spends in the non-star-forming state. For example, the fall of the gas driven out in a fountain outflow back to the disk will occur on a timescale of order \( \sim t_{\text{nsf}} \).

![Figure 6. Distributions and average fluxes of tracers residing at different galactic radii, \( R \) (shown in the top left corner of each panel). Notation follows that of the top panel of Figure 4. Comparison with the right panel of Figure 5 hints that the distribution of tracers on the disk outskirts (top panel) is predominantly shaped by gas compression and expansion due to the spiral arms, while close to the disk center (bottom panel), the distribution is shaped by feedback-driven turbulence and outflows. Whirl-like patterns of velocities in the cold, supersonic phase indicate that the distribution of dense gas at all radii is affected by star formation feedback (see text for details).](image-url)
The Astrophysical Journal, 845:133 (17pp), 2017 August 20

Figure 7. Distribution of time that tracers spend in non-star-forming ($t_{\text{nsf}}$, top panel) and star-forming ($t_{\text{sf}}$, bottom panel) states between successive crossings of the star formation threshold. Distributions of $t_{\text{nsf}}$ and $t_{\text{sf}}$ are shown as a function of galactic radius, $R$. The top axis in each panel also indicates the average surface density of gas at a given radius, $\Sigma_g = \exp(-R/t_\text{sf})$, where $t_\text{sf}$ is the initial scale radius of the disk. Contours indicate 68% and 95% of trajectories. The thick red lines show the median timescales at every radius. The dotted line in the top panel corresponds to the freefall time at the average midplane density at a given radius. Colors in the top panel show the average fraction of time spent in the dense phase with $\Sigma_g > 10 \text{ cm}^{-2}$; the dotted line in the bottom panel shows the average number of passages through the star-forming state required for complete depletion, $N_c$ (Equation (17)). To increase the statistics for long cycles, we follow tracers between 400 and 1000 Myr of the disk evolution. For presentation purposes, we smooth the resulting distributions, preserving their main features.

The bottom panel of Figure 7 shows the distribution of time spent by tracers in the star-forming state on each cycle. This timescale is close to the typical “lifetime” of star-forming regions and is quite short: $t_{\text{sf}} \lesssim 20$ Myr, or 2–4 freefall times at the typical densities of star-forming regions. The fact that $t_{\text{sf}}$ is, on average, significantly shorter than $t_{\text{nsf}}$ is consistent with the small mass fraction of star-forming gas.

As we discussed in Section 2, the average time that a gas parcel spends in the star-forming state on a single cycle determines the total number of such cycles required for complete depletion as $N_c = t_\text{sf}/t_{\text{nsf}}$. For every tracer on each passage through the star-forming state, we estimate this number as the inverse fraction of mass depleted during the passage,

$$N_c^{-1} = \int \frac{dt}{t_\text{sf}} = \int \frac{dt}{t_{\text{nsf}}},$$

where the integral is accumulated for each tracer particle while it is in the star-forming state between subsequent crossings of the star formation threshold. The resulting distribution of $N_c$ is shown by the colors in the bottom panel of the figure.

In agreement with our model, typical $N_c \sim 50$ and the lifetimes of gas in the star-forming state, $t_{\text{nsf}} \sim 10–20$ Myr, are consistent with the range of the star-forming gas depletion times, $\tau_s \sim N_c t_{\text{nsf}} \sim 300–500$ Myr, obtained in our simulation (see the next subsection). In addition, assuming $t_{\text{nsf}} \sim 50–100$ Myr, Equation (2) for the typical depletion time of a gas parcel gives a value of $\sim 2–5$ Gyr, which is consistent with the actual global depletion time obtained in our simulation (see Section 3).

4.4. Emergence of Linear $\Sigma_{\text{H}_2}$–$\Sigma_{\text{H}_2}$ Relation

One of the most intriguing results of our simulation is the emergence of the linear $\Sigma_{\text{H}_2}$–$\Sigma_{\text{H}_2}$ relation consistent with observations (see the right panel of Figure 3), even though on small scales the SFR is assumed to scale nonlinearly with the gas density: $\dot{\rho}_\text{sf} = \rho_\text{H}_2 \rho \propto \rho^{1.5}$. This result provides a counterexample to the arguments that in simulations, the slope of the KSR on large and small scales should be the same (Schaye & Dalla Vecchia 2008; Gnedin et al. 2014).

The observed linearity of the molecular KSR is often explained by the “counting argument,” which posits that all molecular gas is in star-forming clouds of approximately the same depletion time. In this case, the surface density of the SFR will vary linearly with the surface density of molecular gas, as long as the geometric covering fraction of star-forming clouds is less than unity.

However, if not all of the molecular gas is star-forming, as is the case in our simulations and is likely the case in observed galaxies, the explanation for the linearity must be more nuanced. Analogously to Equation (1), the depletion time of molecular hydrogen is given by

$$\tau_{\text{dep}, \text{H}_2} = \frac{\Sigma_{\text{H}_2}}{\Sigma_s} = \frac{\tau_s}{f_{\text{nsf}, \text{H}_2}},$$

where $f_{\text{nsf}, \text{H}_2} = \Sigma_{\text{sf}}/\Sigma_{\text{H}_2}$. The linearity of the molecular KSR, or, equivalently, the independence of the $\tau_{\text{dep}, \text{H}_2}$ of $\Sigma_{\text{H}_2}$, must therefore arise from a lack of dependence of $\tau_s$ and $f_{\text{nsf}, \text{H}_2}$ on $\Sigma_{\text{H}_2}$ or a cancellation of any such dependence in their ratio.

The depletion time of star-forming gas is given by the average over the inverse local depletion time distribution in star-forming regions, $\tau_s = \Sigma_{\text{sf}}/\Sigma_s = \langle 1/t_{\text{nsf}} \rangle^{1/2}$. Thus, if the distribution of $t_{\text{nsf}}$ is independent of $\Sigma_{\text{H}_2}$, so is $\tau_s$, regardless of the actual shape of the distribution. In our simulation, $t_{\text{nsf}} = t_{\text{ff}}/c_{\text{ff}}$, where $c_{\text{ff}} = \text{const}$, and thus the distribution of $t_{\text{nsf}} \propto n^{-3/2}$ is set by the density distribution of star-forming gas.

The PDFs of the star-forming gas density in radial annuli with widths of 500 pc are shown in Figure 8. The figure shows that $\Sigma_{\text{H}_2}$ spans two orders of magnitude, while the average star-forming gas density and the shape of the PDF do vary, but the variation is quite mild. Thus, $\tau_s$ depends on $\Sigma_{\text{H}_2}$, but weakly, as shown by the thin line in Figure 9.

This dependence is weak for two reasons. First, the average density of star-forming gas in all of the ISM environments in our simulation is limited by the stellar feedback (at high densities) and FUV background (at $n < 20 \text{ cm}^{-3}$) to a nearly constant range of $n \sim 10–10^3 \text{ cm}^{-3}$, comparable to the typical densities of observed star-forming regions. Second, as in the...
is also obtained in simulations of the diffuse interstellar medium (ISM) that imply that gas parcels must undergo ~5 100 pc transverse velocities during gas dynamics before reaching star-forming regions. As a result, longer evolution.

The trends of the shown star-forming gas distributions between 400 and 500 Myr are similar to those seen in Figure 9, which indicates that the key expression of our model (Equation 9) was derived by applying the mass conservation law to all star-forming gas in a galaxy or given ISM patch. Mass conservation can also be invoked to build a model for the formation, evolution, and destruction of individual GMCs (e.g., Feldmann & Gnedin 2011; Zamora-Avilés et al. 2012; Lee et al. 2016). The short, ~10–100 Myr, timescales of the physical processes driving the evolution of gas in the ISM (see the top panel of Figure 7) indicate that the ISM is vigorously “boiling” when considered on the global depletion timescale. During this vigorous evolution, gas cycles between non-star-forming and star-forming stages and spends only $t_{df} \sim 5–15$ Myr in the star-forming stage on each cycle (see the bottom panel of Figure 7), which is consistent with the short lifetimes derived for observed GMCs (e.g., Kawamura et al. 2009; Murray 2011; Schruba et al. 2017).

Observational estimates of the integral star formation efficiency during a star-forming stage, defined for a given star-forming region containing a mass of $m_e$ and a mass of formed young stars of $m_*$ as $\epsilon_{int} = m_*/(m_e + m_*)$, give $\epsilon_{int} \sim 1%–20\%$ (e.g., Evans et al. 2009; Lada et al. 2010). This fraction is even smaller in less-efficient clouds (e.g., Rebollo et al. 2015; Lee et al. 2016; Vutisalchavakul et al. 2016). A similar range of $\epsilon_{int}$ is also obtained in simulations of star cluster formation (Gavagnin et al. 2017), models of star formation in GMCs (e.g., Zamora-Avilés & Vázquez-Semadeni 2014), and cosmological simulations of a Milky Way–sized galaxy that resolve the growth of globular clusters and self-consistently capture its termination by stellar feedback (Li et al. 2017).

Such values of $\epsilon_{int}$ imply that gas parcels must undergo $N_c \sim \epsilon_{int}^{-1} \sim 5–100$ cycles transitioning from the non-star-forming to star-forming state before they convert their gas into stars. This number of cycles is also consistent with the typical depletion times of star-forming gas, $\tau_c \equiv (1/t_{df})_{int}^{-1}$, and lifetime, $t_{df}$, derived for observed star-forming regions, $N_c = \tau_c/t_{df}$. Specifically, depletion times of gas in observed star-forming regions are estimated to be $t_{df} \sim 5–15$ Myr quoted above, these $\tau_c$ give $N_c \sim (50–500)/(5–15) \sim 3–100$. Such $N_c$ are in the ballpark of the $N_c$ range that we estimate for the gas in our simulations (see the bottom panel of Figure 7), although we note that the number of cycles in our simulation can be somewhat overestimated due to the artificially sharp threshold in the definition of star-forming gas.

Note that specific values of $t_{df}$, $t_{sf}$, and $N_c$ depend on the scale, $l$, on which the small-scale distribution of $t_c$ is defined. Clearly, if we consider the evolution of gas parcels on the scale of protostellar cores, $\sim 0.01$ pc, the star-forming stage of evolution will correspond to the formation of one or a handful of stars, which will consume most of the gas in a single event. The gas parcels on this scale will spend a long time in the non-star-forming stage and will consume their gas in one or a few cycles, $N_c = 1 + \xi$, where $\xi$ corresponds to the ratio of the protostellar core mass returned to the ISM to the mass of the formed star.

The key expression of our model (Equation 9) was derived by applying the mass conservation law to all star-forming gas in a galaxy or given ISM patch. Mass conservation can also be invoked to build a model for the formation, evolution, and destruction of individual GMCs (e.g., Feldmann & Gnedin 2011; Zamora-Avilés et al. 2012; Lee et al. 2016). Our model is based on mass conservation alone, and thus the overall physical explanation for long gas depletion times presented in Sections 2 and 4 does not require the assumption of dynamical equilibrium. Indeed, when a gas parcel undergoes...
some inherently nonequilibrium process, such as compression in a spiral arm, the parcel’s depletion time will still be given by 
\[ t_{\text{dep}} = (\tau_e / f_d)(t_{\text{nsf}} + t_{\text{sf}}) = \tau_e (t_{\text{nsf}} / f_d + 1) \]
and therefore \( t_{\text{dep}} \) will be long if \( t_{\text{nsf}} \gg t_d \) and/or \( \tau_e \) is long. The absence of a long-term equilibrium requirement is an essential difference of our framework from the models for the KSR that rely on the assumption of self-regulation to the vertical or Toomre (1964) equilibrium state (e.g., Ostriker & Shetty 2011; Faucher-Giguère et al. 2013).

In fact, our model explicitly accounts for the deviations from the equilibrium state in which \( M_g \approx 0 \). Such deviations, along with fluctuations of other quantities that enter Equation (9), can be important sources of the depletion time scatter. These deviations can be substantial for individual ISM patches, which generally will not be in equilibrium, even if a galaxy as a whole is. Additional scatter can arise due to observational tracers sampling different stages of gas evolution incompletely (Kuijssen & Longmore 2014).

We note, however, that unlike the models of Ostriker & Shetty (2011) and Faucher-Giguère et al. (2013), our conceptual framework cannot quantitatively predict the depletion time by itself. It only elucidates how the depletion time is related to the timescales of the processes driving gas evolution. The variables through which this relation is parameterized can be either calibrated in simulations, as is done in this study, or derived in analytical models (e.g., Zamora-Avilés et al. 2012). Nevertheless, as we discuss in the following subsections, our framework is very useful for interpreting and explaining a number of puzzling facts about star formation in both observed and simulated galaxies.

5.1. Implications for Observations

Rapid cycling of gas between non-star-forming and star-forming states explains the large discrepancy between long global depletion times of \( \gtrsim 1 \) Gyr and short, \( \sim 10–100 \) Myr, timescales associated with the dynamical processes in the ISM. Only a small fraction of gas is converted into stars during each cycle, and therefore gas would have to go through a large number of cycles to be depleted.

Our model also naturally explains the difference between observed local depletion times of (mostly molecular) gas in star-forming regions, \( t_e \sim 50–500 \) Myr, and global depletion times of both total gas, \( \tau_{\text{dep}} \sim 2–10 \) Gyr, and molecular gas, \( \tau_{\text{dep,H}_2} \sim 1–3 \) Gyr. The global depletion times, \( \tau_{\text{dep}} \) and \( \tau_{\text{dep,H}_2} \), are longer than the average gas depletion time in star-forming regions, \( \tau_e \), due to the significant fraction of time that gas spends in the non-star-forming state (see Equation (2)). This implies that only a fraction of total and molecular gas is forming stars at any given moment. For example, the observed values of \( t_e \) and \( \tau_{\text{dep,H}_2} \) indicate that only \( f_{\text{H}_2,H}\equiv M_{\text{H}_2}/M_H \approx \tau_e/\tau_{\text{dep,H}_2} \approx 5\%–50\% \) of molecular gas is forming stars. The range of \( f_{\text{H}_2,H} \) in our simulation is consistent with this estimate (see the dotted line in Figure 9), with the non-star-forming molecular gas state corresponding to strongly turbulent cold gas. Diffuse, non-star-forming molecular gas is also observed in the Milky Way, but its mass fraction (\( \sim 25\% \)) from Roman-Duval et al. 2016) is a factor of \( \sim 2–4 \) too low to explain the discrepancy between global and local depletion times of molecular gas. Thus, a substantial fraction of non-star-forming molecular gas must be missed in such observations.

As pointed out by Kuijssen & Longmore (2014), a model considering different evolutionary stages and corresponding chemical phases of the ISM gas can be used to interpret the dependence of the depletion time and its scatter on the averaging scale (see, e.g., Schruba et al. 2017). This dependence in observed galaxies can also be used as a stringent test of the star formation and feedback implementation in galaxy formation models.

Our model for gas depletion time provides a natural framework for predicting and interpreting trends with galaxy properties and redshift. For instance, we show that the duration of cycles decreases with increasing surface density as \( t_{\text{nsf}} \propto \Sigma_g^{-0.5} \), which is accompanied by a milder but nevertheless nonnegligible decrease in \( \tau_e \) (see Section 4.4 and the thin line in Figure 9). This means that the observed decrease in global depletion times in high-redshift and starburst galaxies (e.g., Kennicutt 1998; Bouché et al. 2007; Genzel et al. 2010; Tacconi et al. 2017) can be explained by shorter dynamical timescales, \( t_{\text{nsf}} \), and star-forming gas depletion times, \( \tau_e \), associated with high-density environments. In addition, the nonequilibrium state of starburst galaxies may result in short \( \tau_{\text{dep}} \) due to the contribution of the \( \tau_e/t_{\text{nsf}} \) term in Equation (9).

Our framework also predicts the dependence of depletion times and KSR shape on metallicity. Gas must be shielded by a certain column density in order to become cold and molecular. This column density has a corresponding number density at which such a transition occurs, as can be seen in the phase diagrams in Figure 6 that show the sharp change from the warm, transonic phase (yellow) to the cold, supersonic phase (blue) at \( n \sim 10^3 \text{cm}^{-3} \). At lower metallicities, both the characteristic number density and column density of the transition increase, leading to the decrease of \( f_{\text{H}_2,H} \) and, \( \tau_e \). Thus, the overall gas depletion time, \( \tau_{\text{dep}} \), increases. The mechanism of \( f_{\text{H}_2,H} \) regulation by feedback-driven turbulence discussed in Section 4.4, however, should operate regardless of the gas metallicity. Thus, we expect the KSR for molecular gas to remain linear at low metallicity. Although we also expect its normalization to be approximately independent of metallicity, a weak dependence is possible due to the changing properties of star-forming regions with metallicity.

The higher characteristic density of the transition at lower metallicity also results in the shift of the turnover in the KSR for total gas to higher surface densities. This shift is qualitatively similar to that predicted by the Gnedin & Kravtsov (2011) models, where star formation is tied to molecular gas. We have confirmed this explicitly by resimulating our model galaxy at a lower metallicity and will present these results in a forthcoming paper.

5.2. Implications for Galaxy Formation Simulations

Recent studies show that in simulations with strong feedback, the galaxy-scale SFR—and hence the global depletion time—is insensitive to the local star formation efficiency (e.g., Dobbs et al. 2011; Agerz et al. 2013; Hopkins et al. 2013a, 2017; Agerz & Kravtsov 2015; Benincasa et al. 2016; Orr et al. 2017). This behavior is thought to be due to “self-regulation” of star formation by feedback (e.g., Dobbs et al. 2011).

Our framework naturally explains the physical mechanism of such self-regulation. Equation (9) accounts for feedback via the term proportional to the mass-loading factor, \( \xi \). When feedback is sufficiently strong so that this term dominates, the depletion
time is given by $\tau_{\text{dep}} \approx \xi \tau_s$, where $\tau_s$ is the timescale of the star-forming gas supply from the ISM. Physically, $\tau_s$ is not related to the local star formation efficiency, while $\xi$ is the ratio of the star-forming gas depletion time, $\tau_s = (\epsilon_f/\dot{M}_\mathrm{f})_\mathrm{d}$, and the timescale with which feedback disperses star-forming gas, $\tau_{\text{sf}, \text{fb}}$ (see Section 2). Timescales $\tau_s$ and $\tau_{\text{sf}, \text{fb}}$ are proportional to each other, because the rate of gas removal from star-forming regions is proportional to the rate of energy and momentum injection by feedback, which, in turn, is proportional to the SFR: $M_s/\tau_{\text{sf}, \text{fb}} = \dot{M}_\mathrm{S} = \dot{M}_\mathrm{S}/\tau_s$. Thus, in this regime, the global SFR, $M_s = M_s/\tau_{\text{dep}} \approx M_s/\tau_s$, does not depend on the local depletion time and thus on local star formation efficiency. At the same time, $M_s$ is inversely proportional to the strength of feedback, $\xi$, as observed in simulations with self-regulation (e.g., Hopkins et al. 2017; Orr et al. 2017).

Note that for a given strength of feedback, the regime of self-regulation exists only for a certain range of $\epsilon_f$ and $\tau_s$. For sufficiently low $\epsilon_f$ (long $\tau_s$), the second term in the sum in Equation (9), i.e., $\tau_s$, will dominate, and the global depletion time will scale with the local star formation efficiency as $\tau_{\text{dep}} \sim \tau_s \propto \epsilon_f^{-1}$. Indeed, we have checked that when the feedback parameters and star formation threshold in our simulations are fixed, the global depletion time is insensitive to the variation of local efficiency for $\epsilon_f > 1\%$ but scales as $\tau_{\text{dep}} \propto \epsilon_f^{-1}$ when $\epsilon_f < 1\%$.

The critical value of $\epsilon_f$, above which $\tau_{\text{dep}}$ becomes insensitive to $\epsilon_f$, depends on the relative contribution of the $\tau_s$ term in the sum of Equation (9) and thus depends on the factors that control the first term in the sum, in particular the feedback strength, $\xi$. Dependence of the critical $\epsilon_f$ on feedback strength explains the results of Agertz et al. (2013) and Agertz & Kravtsov (2015), who found that the normalization of the KSR, i.e., $\tau_{\text{dep}}$, scales with $\epsilon_f$ in simulations without feedback but is similar for $\epsilon_f = 1\%$ and $10\%$ when feedback is strong.

The weak sensitivity of the SFR to the value of $\epsilon_f$ in the regime of strong feedback does not mean that this value is not important. For example, the mass fraction of star-forming gas in this regime scales inversely with $\epsilon_f$: $f_{\text{sf}} \equiv M_{\text{sf}}/M_g = \tau_s/\tau_{\text{dep}} \propto \epsilon_f^{-1}$. Thus, an incorrect $\epsilon_f$ will result in an incorrect mass fraction and an incorrect spatial distribution of star-forming gas. A similar conclusion was drawn by Hopkins et al. (2013b), who suggested that the amount of dense gas probed by HCN decreases with increasing star formation efficiency and feedback strength.

The mass fraction of star-forming gas and its distribution is also a nontrivial function of the feedback strength and the specific mix of processes that define it (e.g., Hopkins et al. 2013b; Butler et al. 2017). Stronger feedback shortens the time that gas spends in the star-forming state, $t_{\text{sf}}$, and, therefore, decreases the mass fraction of star-forming gas: $f_{\text{sf}} \equiv M_{\text{sf}}/M_g \sim t_{\text{sf}}/t_{\text{inf}} + t_{\text{dep}}$. Besides, strong feedback shapes the overall gas PDF and may increase timescales $t_{\text{inf}}$, which also results in a decrease of $f_{\text{sf}}$. This happens, for example, when a gaseous disk is stabilized by pressure corresponding to slowly dissipating energy directly injected by feedback (e.g., Agertz et al. 2013; Benincasa et al. 2016).

Finally, $t_{\text{inf}}$, $t_{\text{sf}}$, and $f_{\text{sf}}$ depend on the overall adopted definition of star-forming gas. Usually, the boundary between star-forming and non-star-forming gas is established by a set of thresholds in various gas properties, such as gas density, molecular fraction, virial parameter $\alpha_{\text{vir}}$, etc. Our results indicate that these thresholds should be carefully chosen, because if thresholds grossly misidentify the star-forming gas, the values of $\epsilon_f$ and feedback strength may need to compensate for the wrong choice to get reasonable $t_{\text{sf}}$ and $f_{\text{sf}}$ values. It is not yet clear whether such compensation is possible in general, but, in any case, wrong thresholds may drive the $\epsilon_f$ and feedback parameters to incorrect values.

In addition, the choice of thresholds defining star-forming gas indirectly affects the efficiency of feedback. When more gas is designated as star-forming for a given fixed SFR, the feedback mass-loading factor, $\xi$, will have to be larger; i.e., feedback has to be stronger. The specific choice of thresholds matters as well. For example, when star-forming gas is defined using a density threshold, $n_{\text{sf}}$, feedback has to drive star-forming gas to lower densities, $n < n_{\text{sf}}$, in order to shut down star formation. If star-forming gas is defined using a threshold in the virial parameter instead, $\alpha_{\text{vir}}$, can be increased to values larger than the threshold very quickly by injecting thermal or turbulent energy without changing the gas density significantly.

Thus, the framework presented in this paper implies that to get the value of $f_{\text{sf}}$ and depletion time, $\tau_{\text{dep}} = \tau_{\text{sf}}/t_{\text{inf}}$, and the distribution of star-forming gas correctly, the overall definition of star-forming gas, its star formation efficiency, and the strength of stellar feedback are all important and should be modeled carefully. Ideally, the modeling choices should be based on solid physical ground or subgrid models calibrated on higher-resolution simulations. For example, recent studies show that local $\epsilon_f$ can be modeled using the results of high-resolution simulations of turbulent star-forming gas (Padoan et al. 2012, 2017; Federrath 2015) and a subgrid model for turbulence calibrated on turbulence simulations (e.g., Braun et al. 2014; Braun & Schmidt 2015; Semenov et al. 2016; Li et al. 2017).

Interestingly, the results presented in Section 4.4 indicate that feedback can break the direct connection between the small- and large-scale relations between gas density and SFR. Specifically, our simulations assume $\beta_\mathrm{S} \propto \rho_{\text{iso}}^{1.5}$ on the scales of individual cells. However, as can be seen in Figures 3 and 9, on kiloparsec scales, the relation between the surface density of molecular gas and SFR is almost linear. This result provides a counterexample to the arguments that in simulations, the slope of the KSR on kiloparsec scales simply reflects the assumed density dependence of the local SFR on small scales (Schaye & Dalla Vecchia 2008; Gnedin et al. 2014). Our analysis shows that the linear relation arises because only a fraction of cold, dense gas is forming stars and its density PDF does not scale self-similarly with the large-scale surface density of molecular gas, $\Sigma_{\text{H}_2}$. Variation in covering fraction allows $\Sigma_{\text{H}_2}$ to vary significantly, while the average density of star-forming gas and the corresponding depletion time vary only weakly. In addition, self-regulation by feedback ties the fraction of molecular gas that is forming stars and the depletion time of star-forming gas, so that their trends with $\Sigma_{\text{H}_2}$ are similar. These trends cancel out, resulting in an almost constant $\tau_{\text{dep}, \text{H}_2}$ and linear $\Sigma_{\alpha_{\text{vir}}^{-1}}$ relation.

The emergence of the large-scale KSR via nontrivial, nonlinear effects of the star formation--feedback loop motivates efforts to model such processes in high-resolution simulations rather than tuning the star formation prescription to produce a particular large-scale relation. Our model also sheds some light on the importance of the so-called “early feedback”—a collective name for the energy and
momentum injection by young stars during the first 3 Myr of the life of a stellar population, before the first supernovae explode. The time lag between the formation of stellar particles and the onset of feedback is important when $\tau_s$ is comparable to the lag. In this case, a substantial fraction of the gas mass can be converted into stars in the first 3 Myr before supernova feedback can limit star formation. When $\tau_s$ is long, the global depletion time becomes less sensitive to the presence or absence of stellar feedback during the first 3 Myr of the star-forming stage.

6. Summary and Conclusions

We present a simple and intuitive physical model that elucidates why gas depletion times in galaxies are long compared to the timescales of the processes driving the evolution of the ISM. We show that the depletion time is long, not because some bottleneck in the formation of star-forming regions imposes a long evolutionary timescale, but because only a small fraction of the gas mass is converted into stars during a single star-forming stage in the evolution of a gas parcel. This fraction is small due to the short duration of the star-forming stage, as dynamical processes and stellar feedback efficiently disperse star-forming regions, and the low intrinsic star formation efficiency of dense molecular gas. A gas parcel thus must go through many cycles transitioning between non-star-forming and star-forming states before it becomes converted into stars. Hence, even though the duration of each cycle can be short, the global depletion time is long because the number of cycles is large.

Furthermore, the difference between the global and local depletion times of molecular gas in our model arises because not all of the molecular gas is actively forming stars. Non-star-forming molecular gas appears naturally if local star formation efficiency is a strong function of the virial parameter of a region, while the molecular fraction of gas is set by its ability to shield against FUV radiation and is a function of mainly gas number density and metallicity.

We illustrate our model using the results of an isolated $L_\odot$-sized disk galaxy simulation that reproduces the observed KSR for both molecular and atomic gas. We discuss the predictions of our model for the dependence of the global depletion time on properties of observed galaxies and on the parameters of star formation and feedback recipes in galaxy simulations. In particular, our model explains the weak sensitivity of the global KSR to the assumed local star formation efficiency reported in several recent numerical studies and the physics of this “self-regulation” (see Sections 2 and 5.2).

Additional results and conclusions can be summarized as follows.

1. Analysis of our simulation shows that the properties of gas parcels in the ISM evolve on timescales of $\sim 10$–100 Myr under the influence of compression by the spiral arms, ISM turbulence, and SN-driven shocks. The relative importance of these processes varies with galactocentric radius, $R$, and average surface density. At $R \gtrsim 5$ kpc, the evolution from a warm, diffuse state to a dense, cold phase is driven mainly by compression in the spiral arms, while SN-driven shocks and large-scale ISM turbulence dominate at smaller radii.

2. During an evolutionary cycle, gas spends most of the time in the non-star-forming state, $t_{\text{nsf}} > t_{\text{sf}}$, whereas the time spent in the star-forming state, $t_{\text{sf}}$, is limited by stellar feedback and dynamical processes to $t_{\text{sf}} \sim 5$–15 Myr. We find that the median $t_{\text{nsf}}$ varies with gas surface density as $t_{\text{nsf}} \propto \Sigma^{0.5}$.

3. On the resolution scale of our simulation, 40 pc, the typical range of densities in star-forming regions is limited to $n \sim 10^2$–$10^3$ cm$^{-3}$ by the interstellar FUV background and stellar feedback. We find that the resulting depletion times of star-forming gas $\tau_s \sim 300$–500 Myr are consistent with the depletion times estimated for observed GMCs on these scales and exhibit only weak trends with the surface densities of total and molecular gas.

4. The distributions of depletion times and lifetimes of star-forming regions in our simulations imply that a typical gas parcel has to undergo 5–100 cycles transitioning between non-star-forming and star-forming states before converting its mass into stars.

5. On kiloparsec scales, our simulation produces a nearly linear relation between the surface density of H$_2$ and surface density of the SFR, i.e., the molecular KSR, even though a nonlinear local relation, $\dot{\rho}_s \propto \rho^{1.5}$, is adopted for star-forming gas in simulation cells. We show that the linear relation emerges due to stellar feedback, which shapes the gas density PDF (making it non-self-similar) and establishes a correlation between the local depletion time of the star-forming gas and the fraction of molecular gas that is in the star-forming state.

The model for the global depletion timescale presented in this paper is a generic framework that can be applied not only to galaxies as a whole but also to individual ISM patches with sizes ranging from ~kiloparsec to a typical size of star-forming regions, ~10 pc. It can also be used to predict and interpret trends of gas depletion time with the ISM properties and redshift.

As a final comment, we note that in the context of galaxy evolution over cosmological timescales, the actual gas depletion time is often considered to be unimportant. For example, galaxies from the star-forming sequence are predicted to form stars at the rates regulated by gas accretion and gas loss in winds, because gas depletion times in such galaxies are short compared to other relevant timescales (e.g., Bouché et al. 2010; Davé et al. 2012; Lilly et al. 2013). Note, however, that dwarf galaxies and galaxies from the green valley consume gas on extremely long timescales of $\gtrsim 5$–10 Gyr, and, therefore, their depletion times do affect their evolution. Moreover, at $z \gtrsim 6$, when the age of the universe is $\lesssim 1$ Gyr, the Gyr-long gas depletion times become comparable to the cosmological evolution timescale and will therefore play an important role in controlling the SFR during the early stages of galaxy evolution (e.g., Dekel & Mandelker 2014; Peng & Maiolino 2014). The framework for modeling gas depletion time presented in this paper thus opens a way to refine theoretical models of galaxy formation in this regime, which is particularly important in the upcoming era of the James Webb Space Telescope.

We would like to thank the referee, Mark Krumholz, for a constructive report and many detailed comments that have improved the paper. We thank Cameron Liang, Claude-André...
References

Agertz, O., & Kravtsov, A. V. 2015, ApJ, 804, 18
Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, ApJ, 770, 25
Benincasa, S. M., Wadsley, J., Couchman, H. M. P., & Keller, B. W. 2016, MNRAS, 462, 3653
Bigiel, F., Leroy, A., Walter, F., et al. 2008, AJ, 136, 2846
Bigiel, F., Leroy, A., Walter, F., et al. 2010, AJ, 140, 1194
Bisantzi, N., Englmaier, P., & Gerhard, O. 2003, MNRAS, 340, 949
Bouché, N., Cresci, G., Davies, R., et al. 2007, ApJ, 671, 303
Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001
Braun, H., & Schmidt, W. 2015, MNRAS, 454, 1545
Braun, H., Schmidt, W., Niemeyer, J. C., & Almgren, A. S. 2014, MNRAS, 442, 3407
Braun, R. 2012, ApJ, 749, 87
Butler, M. J., Tan, J. C., Teyssier, R., et al. 2017, ApJ, 841, 82
Chabrier, G. 2003, PASP, 115, 763
Christensen, C., Quinn, T., Governato, F., et al. 2012, MNRAS, 425, 3058
Daddi, E., Elbaz, D., Walter, F., et al. 2010, ApJ, 714, L118
Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98
Davis, B. L., Kenne, G., Kenne, J., et al. 2015, ApJ, 802, L13
Dekel, A., & Mandelker, N. 2014, MNRAS, 444, 2071
Dobbs, C. L., Burkert, A., & Pringle, J. E. 2011, MNRAS, 413, 2935
Dobbs, C. L., Pringle, J. E., & Burkert, A. 2012, MNRAS, 425, 2157
Dobbs, C. L., Pringle, J. E., & Duarte-Cabral, A. 2015, MNRAS, 446, 3608
Elmegreen, B. G. 2002, ApJ, 577, 206
Elmegreen, B. G. 2015, ApJL, 814, L30
Evans, N. J., II, Dunham, M. M., Jørgensen, J. K., et al. 2009, ApJS, 181, 321
Evans, N. J., II, Heiderman, A., Allen, L. E., Huard, T., & Heyer, M. 2010, ApJ, 723, 1019
Heiles, C., & Troland, T. H. 2003, ApJ, 586, 1067
Hernquist, L. 1990, ApJ, 356, 359
Hernquist, L. 1990, ApJ, 356, 359
Heyer, M., Gnedin, N. Y., & Kravtsov, A. V. 2014, ApJ, 787, L7
Hill, G. H., & Kravtsov, A. V. 2009, ApJ, 697, 55
Gettmuth, R. A., Paper, I. L., & Fujimoto, Y. 2014, ApJ, 787, L7
Gnedin, N. Y., Tissera, E., & Kravtsov, A. V. 2009, ApJ, 697, 55
Gutermuth, R. A., Megeath, S. T., et al. 2011, ApJ, 739, 84
Hayward, C. C., & Hopkins, P. F. 2017, MNRAS, 465, 1682
Heiderman, A., Evans, N. J., II, Allen, L. E., Huard, T., & Heyer, M. 2010, ApJ, 723, 1019
Heiles, C., & Troland, T. H. 2003, ApJ, 586, 1067
Hernquist, L. 1990, ApJ, 356, 359
Hill, G. H., & Kravtsov, A. V. 2009, ApJ, 697, 55
Gutermuth, R. A., Megeath, S. T., et al. 2011, ApJ, 739, 84
Hayward, C. C., & Hopkins, P. F. 2017, MNRAS, 465, 1682

Orcid iDs

Vadim A. Semenov https://orcid.org/0000-0002-6648-7136
Andrey V. Kravtsov https://orcid.org/0000-0003-4307-634X

The Astrophysical Journal, 845:133 (17pp), 2017 August 20

Semenov, Kravtsov, & Gnedin

Faucher-Giguère, Robert Feldmann, Romain Teyssier, Simon Lilly, and Marcella Carlofor useful discussions. We are also grateful to Oscar Ageritz, Camille Avestruz, Benedikt Diemer, Oleg Gnedin, Hui Li, Philip Mansfield, Enrique Vázquez-Semadeni, and Tony Wong, whose valuable comments and suggestions helped to improve our paper. We also thank Adam Leroy for sharing the data from Leroy et al. (2013) that we used in Figure 3. This work was supported by NASA ATP grant NNH12ZDA001NF, NSF grant AST-1412107, and the Kavli Institute for Cosmological Physics at the University of Chicago through grant PHY-1125897 and an endowment from the Kavli Foundation and its founder, Fred Kavli. The simulation and analyses presented in this paper have been carried out using the Midway cluster at the University of Chicago Research Computing Center, which we acknowledge for support.

The references are cited in the text.
Schruba, A., Leroy, A. K., Walter, F., Sandstrom, K., & Rosolowsky, E. 2010, 
ApJ, 722, 1699
Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2016, ApJ, 826, 200
Silk, J. 1997, ApJ, 481, 703
Silk, J., & Norman, C. 2009, ApJ, 700, 262
Sofue, Y. 2017, MNRAS, 468, 4030
Stecher, T. P., & Williams, D. A. 1967, ApJL, 149, L29
Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2017, ApJ, submitted
(arXiv:1702.01140)
Tan, J. C. 2000, ApJ, 536, 173

Toomre, A. 1964, ApJ, 139, 1217
Vutisalchavakul, N., Evans, N. J., II, & Heyer, M. 2016, ApJ, 831, 73
Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, 
ApJ, 587, 278
Wong, T., & Blitz, L. 2002, ApJ, 569, 157
Wyse, R. F. G., & Silk, J. 1989, ApJ, 339, 700
Zamora-Avilés, M., & Vázquez-Semadeni, E. 2014, ApJ, 793, 84
Zamora-Avilés, M., Vázquez-Semadeni, E., & Colin, P. 2012, ApJ, 751, 77
Zuckerman, B., & Evans, N. J., II 1974, ApJL, 192, L149
Zuckerman, B., & Palmer, P. 1974, ARA&A, 12, 279