6-1-2022

A Self-consistent Model for Brown Dwarf Populations

R. E. Ryan Jr.

P. Thorman

C. Aganze

A. J. Burgasser

S. H. Cohen

See next page for additional authors

Original Publication Information
R. E. Ryan Jr. *et al* "A Self-consistent Model for Brown Dwarf Populations." 2022 *Astrophysical Journal* 932(2): 1-10.

The Astrophysical Journal, 932:96 (10pp), 2022 June 20 https://doi.org/10.3847/1538-4357/ac6de5

A Self-consistent Model for Brown Dwarf Populations

R. E. Ryan, Jr.[1], P. Thorman[2], C. Aganze[3], A. J. Burgasser[3], S. H. Cohen[4], N. P. Hathi[1], B. Holwerda[5], N. Pirzkal[1], and R. A. Windhorst[6]
[1] Space Telescope Science Institute, 3700 San Martin Ave. Baltimore, MD 21218, USA
[2] Department of Physics & Astronomy, Haverford College Haverford, PA 19041, USA
[3] Department of Physics & Astronomy, University of California San Diego, CA 92093, USA
[4] School of Earth & Space Exploration, Arizona State University Tempe, AZ 85287-1404, USA
[5] Department of Physics and Astronomy, University of Louisville Louisville, KY 40292, USA
[6] School of Earth & Space Exploration, Arizona State University Tempe, AZ 85287, USA
Received 2021 May 11; revised 2022 April 20; accepted 2022 May 1; published 2022 June 20

Abstract

We present a self-consistent model of the Milky Way to reproduce the observed distributions (spectral type, absolute *J*-band magnitude, effective temperature) and total velocity dispersion of brown dwarfs. For our model, we adopt parametric forms for the star formation history and initial-mass function, published evolutionary models, and theoretical age–velocity relations. Using standard Markov Chain Monte Carlo methods, we derive a power-law index of the initial-mass function of $\alpha = -0.71 \pm 0.11$, which is an improvement over previous studies. We consider a gamma-function form for the star formation history, though we find that this complex model is only slightly favored over a declining exponential. We find that a velocity variance that linearly increases with age and has an initial value of $\sigma_0 = 9.0^{+11}_{-9.0}$ km s^{-1} best reproduces the total velocity dispersions. Given the similarities to main-sequence stars, this suggests brown dwarfs likely form via similar processes, but we recognize that the sizable uncertainties on σ_0 preclude firm conclusions. To further refine these conclusions, we suggest that wide-field infrared imaging or low-resolution spectroscopic surveys, such as with the Nancy Grace Roman Space Telescope or Euclid, could provide large samples of brown dwarfs with robust spectral types that could probe the thickness of the thin disk. In this way, the number counts and population demographics could probe the same physical processes as with the kinematic measurements, however may provide larger samples and be subject to different selection biases.

Unified Astronomy Thesaurus concepts: Brown dwarfs (185)

1. Introduction

Brown dwarfs represent the substellar extension of the main sequence, and are canonically defined as objects that lack sufficient mass to sustain hydrogen fusion. Consequently, these objects cool throughout their lifetimes, despite other properties (such as mass) remaining roughly constant, which leads to the well-known age–mass degeneracy: a brown dwarf of a given temperature (or spectral type) can have range of ages and masses (e.g., Burrows et al. 1989). This strong coupling between the age and temperature makes brown dwarfs a viable standard clock with which we can examine the history of the Milky Way.

The distribution of observed properties (such as spectral type, effective temperature, absolute magnitude, etc.) are a cornerstone observation of stellar population studies. But since the brown dwarfs evolve considerably, the present-day distributions are sensitive to the initial, fundamental distributions and assumptions on the evolutionary models. Specifically, various authors have explored the relationship between a parameterized initial-mass function (IMF) and observed distributions (such as the luminosity function; see, e.g., Burgasser 2004; Allen et al. 2005). Although this is only an indirect probe of the IMF, it does provide some support for/against particular functional forms. Kirkpatrick et al. (2021) describe how different brown dwarf formation scenarios are manifested in different functional forms of the IMF. In brief, a power-law indicates a scale-free formation, while a skewed normal (or log-normal) distribution implies multiplicative processes and/or multiple intertwined mechanisms.

In addition to the directly observable distribution functions, the kinematics of the brown dwarf population provide a unique glimpse into their formation and evolutionary history. It has long been known that the velocity dispersion of a stellar system will increase with time (e.g., von Hoerner 1959; Mayor 1974; Wielen 1977). Specifically, the thin disk of the Milky Way has been observed to increase in thickness or vertical velocity dispersion throughout its lifetime (sometimes referred to as *disk heating*), and various authors have proposed several driving mechanisms, such as scattering off molecular clouds (Spitzer & Schwarzschild 1951), minor mergers (Quinn et al. 1993), the presence of a central bar (Grand et al. 2016), disk warping (Mackereth et al. 2019), etc. Whatever the causes may be, their cumulative effects (increasing velocity dispersion or disk thickness with population age or spectral type) have been well-studied (e.g., Dehnen & Binney 1998; Aumer & Binney 2009), and the initial velocity dispersions probe the state of the formation. For example, Reipurth & Clarke (2001) suggest that, as stellar systems are forming in their birth clouds, the lowest-mass objects are preferentially ejected via dynamical interactions, which ultimately limits the mass they can accumulate during their formation. Consequently, it is natural to expect the brown dwarf populations to have initial velocity dispersions higher than those of main-sequence stars. Building on prior studies of brown dwarf kinematics (e.g., Schmidt et al. 2007; Zapatero Osorio et al. 2007;

Burgasser et al. 2015), Ryan et al. (2017) suggest that, because brown dwarfs cool over time, their velocity dispersions (or disk vertical scale height) should exhibit strong deviations from the trends set by their slightly more massive, main-sequence cousins. A chief goal of the present work is to expand on the work of Ryan et al. (2017) and use published distribution functions and kinematics of observed brown dwarfs to understand their formation history.

This paper is organized as follows: Section 2 describes our model for the brown dwarf disk population, Section 3 catalogs the published observations that we have adopted, Section 4 details our data modeling, Section 5 discusses our key results, and Section 6 gives our conclusions. We quote all magnitudes in the Vega system.

2. Galaxy Model

We assume that the creation function of brown dwarfs is a separable function of mass (m) and time (t):

$$C(t, m)\, dt\, dm = \psi(t)\, \phi(m)\, dt\, dm, \tag{1}$$

where $\psi(t)$ is the star formation history (SFH) and $\phi(m)$ is the IMF. We adopt parametric forms for each distribution:

$$\phi(m)\, dm = \phi^* \left(\frac{m}{m^*}\right)^{\alpha} \frac{dm}{m^*} \tag{2}$$

$$\psi(t)\, dt = \psi^* \left(\frac{t}{t^*}\right)^{\beta} e^{-t/t^*} \frac{dt}{t^*}, \tag{3}$$

but since the number of brown dwarfs is given as the product of these distributions, we cannot simultaneously constrain ϕ^* and ψ^*. Therefore, we fix ψ^* such that $\int \psi(t)\, dt = 1$, which leaves four free parameters associated with the formation of brown dwarfs: (ϕ^*, α, β, t^*). The parameter m^* is degenerate with the IMF normalization, therefore we fix $m^* = 0.1\ M_\odot$. To avoid confusion, we will define the time since the formation of the Milky Way disk as t and the age of a given population as τ, but these are trivially related as $t + \tau = \tau_{\rm MW}$, where $\tau_{\rm MW}$ is the age of the Milky Way disk. We assume $\tau_{\rm MW} = 10$ Gyr to be consistent with assumptions in previous works (e.g., Burgasser 2004; Allen et al. 2005; Kirkpatrick et al. 2019). In Appendix A and Appendix B, we briefly discuss results for a log-normal IMF and changes to the assumed age of the Milky Way disk, respectively. This form of the SFH in Equation (3) provides a highly flexible family of distributions that can reproduce a constant ($\beta = 0$ and $t^* \gg \tau_{\rm MW}$), a monotonically decreasing ($\beta = 0$ and $t^* > 0$) or increasing ($\beta t^* \geqslant \tau_{\rm MW}$), or something peaked[7] for $0 < (\beta t^*) \leqslant \tau_{\rm MW}$. In Figure 1, we illustrate the range of SFHs this functional form can generate.

Here, our goal is to model the measured distributions of various observables, and mass is seldom a directly measured quantity. Therefore, we must transform the mass function to a relevant observable. However, because brown dwarfs are constantly cooling, this transformation must be time-dependent. Therefore, we transform from mass to temperature as a function of age:

$$\phi(T, \tau)\, dT = \phi(m) \left| \frac{dT(\tau, m)}{dm} \right| dm, \tag{4}$$

[7] The peak star formation rate will occur at $t_{\rm max} = \beta t^*$.

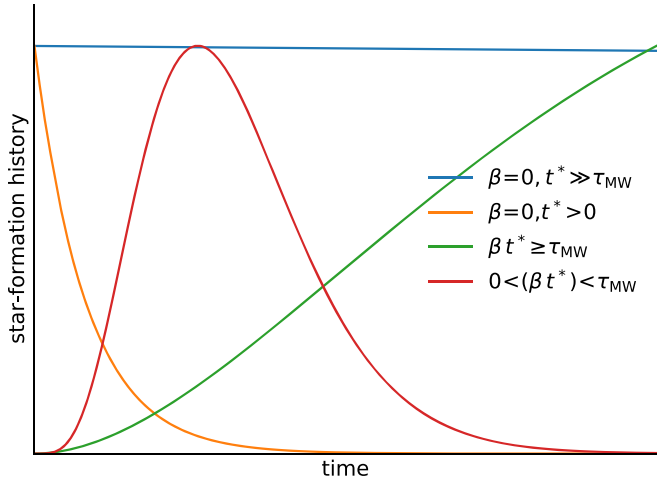

Figure 1. Illustrative forms of our adopted SFH parameterization. Each SFH is normalized to its respective peak, but they can exhibit a range of properties: approximately constant (blue line), monotonically decreasing (orange line), peaked (green line), and monotonically increasing (red line).

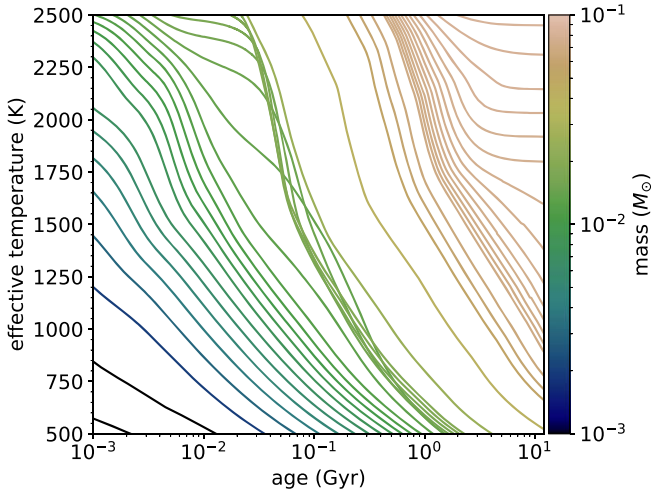

Figure 2. Brown dwarf temperatures from Burrows et al. (1997). Each line represents a different initial mass, as indicated by a color bar.

where $T(\tau, m)$ is the cooling model. We adopt the Burrows et al. (1997) models because they span the broadest range of mass, age, and temperature (see Figure 2), which restrict our creation function to $10^{-3} \leqslant t \leqslant \tau_{\rm MW}$ and $0.0005 \leqslant m \leqslant 0.1\ M_\odot$. We have examined other cooling models (Burrows et al. 2001; Baraffe et al. 2003; Saumon & Marley 2008; Marley & Saumon 2020; Phillips et al. 2020, but see Appendix C for more details), and find results similar to those presented below, but over a narrower range of observed parameters. The Burrows et al. (1997) models assume solar metallicity, therefore that assumption is implicit in our work as well. The temperature distribution is given as a convolution with the SFH:

$$\phi(T)\, dT = \int_0^{\tau_{\rm MW}} \phi(T, \tau)\, \psi(t)\, dt. \tag{5}$$

Finally, the temperature distribution can be transformed to any observable $X = X(T)$ as:

$$\phi(X)\,dX = \phi(T)\left|\frac{dX}{dT}\right|\,dT, \tag{6}$$

which can be directly compared to the observations.

To model the kinematics, we employ the age–velocity dispersion relations (AVR) derived by Wielen (1977):

$$\sigma_x(\tau)^2 = \sigma_{x,0}^2 + C_w\tau \tag{7}$$

$$\sigma_x(\tau)^3 = \sigma_{x,0}^3 + \frac{3}{2}\gamma_x\tau \tag{8}$$

$$\sigma_x(\tau)^3 = \sigma_{x,0}^3 + \frac{3}{2}\frac{\gamma_x}{\omega_x}(e^{\omega_x\tau} - 1), \tag{9}$$

where x can represent any single component of the velocity vector (i.e., $x \in (U, V, W)$) or its magnitude, which we denote as $\sigma^2 = \sigma_U^2 + \sigma_V^2 + \sigma_W^2$. The distinguishing characteristics between Equations (7), (8), and (9) are the assumptions placed on the diffusion coefficient: namely, that it is constant, velocity-dependent, or velocity/time-dependent, respectively. A final remark on Equation (9): Wielen (1977) denote the diffusion timescale as T, but we use an inverse time ω in order to avoid confusion with temperature.

Now we shall establish a model velocity dispersion that is directly comparable to the published measurements. If we consider a single stellar population of disk brown dwarfs created at some time in the past, then it will have a present-day velocity dispersion given by one of the above forms (Equations (7)–(9)). Similarly, subsequent populations will have a lower present-day velocity dispersion (as they will be younger). However, the measured velocity dispersions are not divided by population age, but are effectively averaged over the age of the Milky Way disk. Therefore, we average the above AVRs over the age of the disk age and weight them by the creation function in order to account for the number of varying SFH. Finally, the velocity dispersions are reported for a specified range of spectral type, so we also integrate over the relevant range of effective temperature that we determine from the spline fit in Figure 3. We denote this range of effective temperature as (T_0, T_1), so the model velocity dispersion is given by:

$$\sigma_x^2(T_0, T_1) = \frac{\int_{T_0}^{T_1}\int_0^{\tau_{\rm MW}}\phi(T,\tau)\psi(t)\sigma_x^2(\tau)\,dt\,dT}{\int_{T_0}^{T_1}\int_0^{\tau_{\rm MW}}\phi(T,\tau)\psi(t)\,dt\,dT}, \tag{10}$$

and we recognize that the denominator is the integral over temperature distribution given by Equation (5).

As a final remark, most authors employ Monte Carlo simulations to transform the creation function (Equation (1)) into relevant observable distributions (e.g., Burgasser 2004; Deacon & Hambly 2006; Day-Jones et al. 2013; Ryan et al. 2017). We have verified that our method is consistent with a Monte Carlo simulation for a model, but is significantly more efficient, which permits a more thorough exploration of the uncertainties and degeneracies.

3. Observations

We seek to fit the above model to the observed properties of brown dwarfs as published in various studies. Below, we describe the relevant attributes of our adopted data.

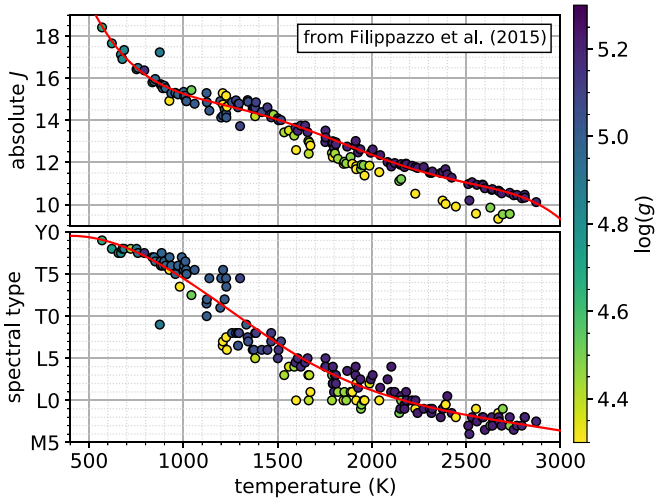

Figure 3. Transformation functions. The data are taken from Filippazzo et al. (2015), the surface gravity $\log(g)$ is encoded in the point color as indicated in the color bar, and the red line represents a cubic-spline fit. We use these splines when transforming from effective temperature to either of these observable (spectral type or absolute J-band magnitude). As described in the accompanying text, we tailor these splines to the high-surface gravity objects, as published samples of field brown dwarfs find very small fractions of low-surface gravity objects.

3.1. Distribution Functions

We consider three observed distributions: effective temperature, spectral type ($\mathcal{S}$), and J-band absolute magnitude (M_J). We use the data from Filippazzo et al. (2015) to establish the transformations from effective temperature to a given observable of the form $X = X(T)$. However, we update the parallaxes (and hence M_J) of two objects (0949−1545 and 1647 + 5632) with higher-precision data from Best et al. (2020) and Kirkpatrick et al. (2021). As seen in Figure 3, these observed relationships depend on the surface gravity (as encoded by the plot symbol color). When studies of field brown dwarfs have sufficient data to distinguish low- and high-surface gravity sources, they find that the samples contain small fractions of low-gravity or young brown dwarfs. For example, Best et al. (2021) find that $5.5 \pm 1.2\%$ and $2.6 \pm 1.6\%$ of their sample of 369 brown dwarfs are young ($\leqslant$200 Myr) and subdwarfs, respectively. Therefore, we tailor our transformations for the high-surface gravity objects, and we show the cubic-spline fits (red lines in Figure 3) to these objects. In Section 5, we will discuss this effect as a systematic difference between our model (tailored to high-surface gravity brown dwarfs) and the observations (containing small proportions of low-gravity objects).

3.2. Kinematic Data

There are many published results for brown dwarf velocity dispersions, which are often quoted in the *UVW*-coordinate system, which is defined such that they increase in the direction of the Galactic center, Galactic rotation, and Galactic north pole, respectively. Additionally, to adopt the Wielen (1977) relations, it is critical to use $|W|$-weighted velocity dispersions (Wielen 1974). Therefore, we recompute the velocity dispersions with a consistent weighting scheme based on Equations (1)–(3) of Wielen (1977) for five publications that present *UVW*-velocities of individual brown dwarfs (Zapatero Osorio et al. 2007; Blake et al. 2010; Schmidt et al. 2010;

Table 1
Observed Data

Observation	Reference	Domain	Number[a]	Count[a]
spectral type	Best et al. (2021)	L0 $\leqslant \mathcal{S} \leqslant$ T8.5	19	369
	Burningham et al. (2013)	T6 $\leqslant \mathcal{S} \leqslant$ Y0	4	76
	Day-Jones et al. (2013)	L4 $\leqslant \mathcal{S} \leqslant$ T4.5	3	63
	Warren et al. (2021)	M9 $\leqslant \mathcal{S} \leqslant$ L3	4	33958
effective temperature	Kirkpatrick et al. (2021)	450 $\leqslant T \leqslant$ 2100	11	525
J-band absolute magnitude	Bardalez Gagliuffi et al. (2019) [b]	10.5 $\leqslant M_J \leqslant$ 14	6	410
	Cruz et al. (2007)	10.5 $\leqslant M_J \leqslant$ 14	7	198
	Reylé et al. (2010)	14 $\leqslant M_J \leqslant$ 17	6	102
velocity dispersions	Blake et al. (2010)	M9 $\leqslant \mathcal{S} \leqslant$ L6	2	59
	Burgasser et al. (2015)	M6 $\leqslant \mathcal{S} \leqslant$ L6	2	85
	Schmidt et al. (2010)	L0 $\leqslant \mathcal{S} \leqslant$ L5	3	484
	Seifahrt et al. (2010)	L0 $\leqslant \mathcal{S} \leqslant$ L8	2	43
	Zapatero Osorio et al. (2007)	M6.5 $\leqslant \mathcal{S} \leqslant$ T8	1	21
totals			70	36393

Notes.
[a] Here, "number" and "count" refer to the number of data points we model and the total number of brown dwarfs in the parent publication, respectively.
[b] As described in Section 3.3, the Bardalez Gagliuffi et al. (2019) are scaled by 1.4 in order to be consistent with the Cruz et al. (2007) results, and as such the number of observations is reduced by one.

Seifahrt et al. 2010; Burgasser et al. 2015). It is worth commenting on a few properties of the underlying samples in these five publications. The authors of each of these previous works have obtained their own respective radial velocities and proper motions, with the exceptions of Blake et al. (2010) and Seifahrt et al. (2010), who employ the same proper motion catalog. Furthermore, these two samples share 29 brown dwarfs (approximately half of each sample separately). Therefore, the bulk of the kinematic data should be considered uncorrelated, with the exception of a fraction of the Blake et al. (2010) and Seifahrt et al. (2010) samples. However, we keep the two data sets in our analysis because they use different radial velocities and still have a many unique objects.

In Figure 4, we show the $|W|$-weighted velocity dispersions normalized to their magnitudes as a function of spectral type. We see that the average ratios are $\langle \sigma_U/\sigma \rangle = 0.77 \pm 0.08$, $\langle \sigma_V/\sigma \rangle = 0.51 \pm 0.15$, and $\langle \sigma_W/\sigma \rangle = 0.36 \pm 0.083$, which is comparable to the original findings of Wielen (1977) of 0.77, 0.49, and 0.41 for a vast range of stellar types (see their Table 1), respectively. Since these ratios are not seen to vary with age (Wielen 1977), we will use the total velocity dispersion (σ) as our kinematic observable in order to minimize the free parameters in the model, though in principle one could consider each component of the velocity vector separately.

3.3. Data Summary

In Table 1, we list all the data we model, their relevant citations, the domains over which the measurements are made, and the numbers of individual measurements. However, there is an obvious discrepancy in the J-band luminosity functions between Cruz et al. (2007) and Bardalez Gagliuffi et al. (2019), which has been previously noted (Kirkpatrick et al. 2021, who refer to a private communication with Bardalez Gagliuffi that indicates their measurements are high due to pessimistic completeness corrections). Since the results from Bardalez Gagliuffi et al. (2019) are derived from considerably more data, we do not want to retain them in our analysis. Therefore, we derive a multiplicative correction as the uncertainty-weighted average ratio of the Bardalez Gagliuffi et al. (2019) to Cruz et al. (2007) luminosity functions, which we find as $r = 1.4 \pm 0.3$. In the subsequent analyses, we renormalize the Bardalez Gagliuffi et al. (2019) results by this factor and reduce the degrees of freedom by one. With that, we are left with a total of 77 measurements or bins in a given observable, which were derived from many hundreds to thousands of individual brown dwarfs.

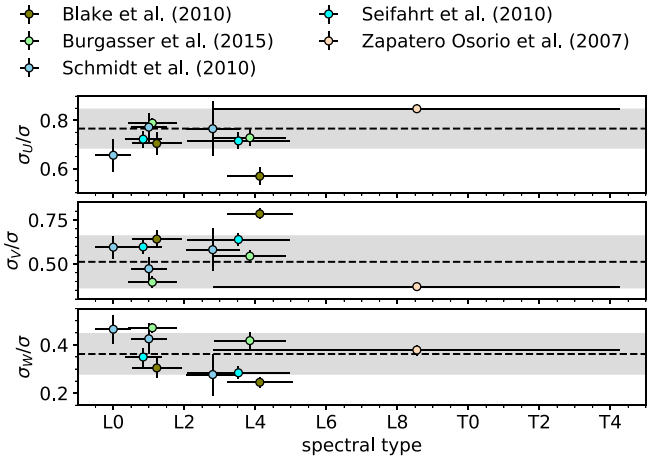

Figure 4. Published $|W|$-weighted velocity dispersions. The symbol colors are described in the above legend, and the dashed line and gray bar indicate the uncertainty-weighted average of each ratio and one standard deviation range. Only the vertical component σ_W/σ shows any evidence for a strong type dependence, which was predicted by Ryan et al. (2017).

4. MCMC Marginalization

Our primary goal is to identify to what extent a single, self-consistent model for the brown dwarf creation function and diffusion law can simultaneously reproduce the observed distribution functions and kinematics. While simple "greedy" algorithms may quickly identify the optimal solution, they generally cannot produce contours for assessing degeneracies in the parameter space. Therefore, we embed our model within a Markov Chain Monte Carlo (MCMC) simulation to draw

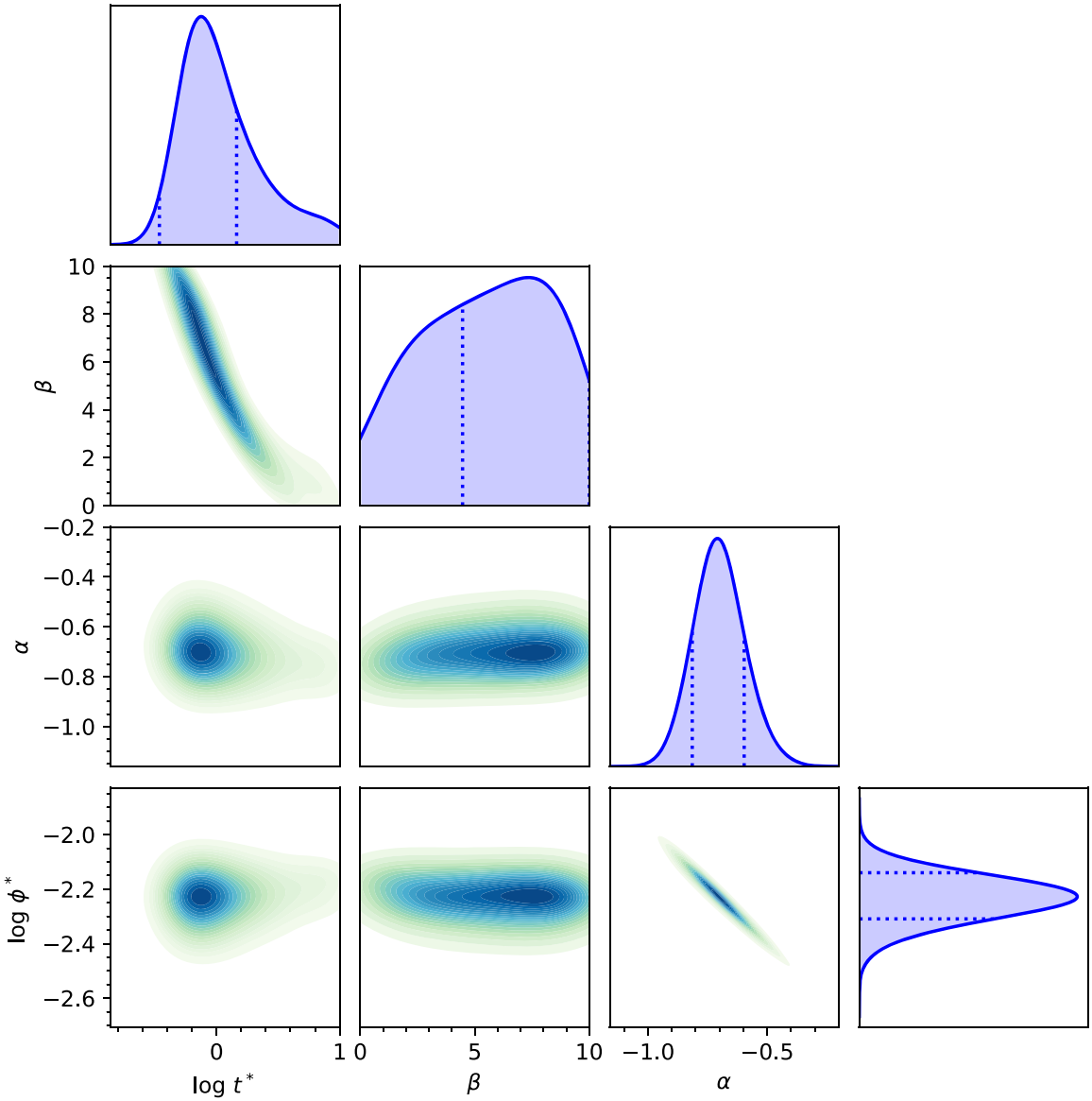

Figure 5. Corner plot for simultaneous fitting of observed distributions. The dotted lines in the marginal distributions indicate the 68.3% confidence range around their respective peaks.

random deviates from the *posterior* distribution, which is given by Bayes' theorem:

$$p(\theta|\mathcal{D}) \propto p(\theta)\,\mathcal{L}(\mathcal{D}|\theta), \tag{11}$$

where θ is a vector of the unknown parameters of the model (discussed in more detail below), $\mathcal{D}$ is the vector of measured data (as described above), $p(\theta)$ is the prior probability of this model vector, and $\mathcal{L}(\mathcal{D}|\theta)$ is the likelihood of getting the data vector given the model. Given the quoted uncertainties of the data, we take the likelihood as a Gaussian, $-2\ln\mathcal{L} = \chi^2$, and χ^2 has the usual definition. We adopt simple top-hat priors in either linear or logarithmic space as stated in Table 2. We use the `emcee` implementation (Foreman-Mackey et al. 2013) of the affine-invariant sampler (Goodman & Weare 2010) for all MCMC simulations.

In the following subsections, we will fit various subsets of the observed data (i.e., the distributions in spectral type, effective temperature, and absolute *J*-band magnitude as well as total velocity dispersion). We treat each observable equally and do not apply any relative weighting when computing the *posterior*. Therefore, the *posterior* used in the MCMC simulation(s) is the simple product of the *posteriors* derived for each distribution

Table 2
Model Parameters

Name	Units	Range
IMF normalization	$\log\ \#/\mathrm{pc}^3$	$-4 \leqslant \log\phi^* \leqslant 0$
IMF power-law index	⋯	$-2 \leqslant \alpha \leqslant 0$
SFH power-law index	⋯	$0 \leqslant \beta \leqslant 8$
SFH *e*-folding timescale	log Gyr	$-1 \leqslant \log t^* \leqslant 3$
initial velocity dispersion	$\mathrm{km\ s}^{-1}$	$0 \leqslant \sigma_0$
constant diffusion coefficient	$(\mathrm{km\ s})^{-2}/\mathrm{Gyr}$	$0 \leqslant C$
velocity-dependent diffusion coefficient	$(\mathrm{km\ s})^{-3}/\mathrm{Gyr}$	$0 \leqslant \gamma$
velocity-dependent *e*-folding timescale	$\log\ \mathrm{Gyr}^{-1}$	$-2 \leqslant \log\omega \leqslant 2$

Table 3
Best-Fit Model Parameters[a]

Description	$\log\phi^*$	α	β	$\log t^*$	σ_0	C	γ	$\log\omega$	χ^2/ν[b]	BIC
photometric distributions	$-2.22^{+0.09}_{-0.09}$	$-0.71^{+0.11}_{-0.11}$	$7.34^{+2.66}_{-2.87}$	$-0.12^{+0.28}_{-0.34}$	⋯	⋯	⋯	⋯	258/63	274
constant diffusion	$-2.23^{+0.13}_{-0.13}$	$-0.72^{+0.13}_{-0.14}$	$2.56^{+2.93}_{-2.56}$	$0.17^{+0.52}_{-0.59}$	$9.01^{+12.16}_{-9.01}$	671^{+224}_{-217}	⋯	⋯	371/71	397
velocity-dependent	$-2.21^{+0.09}_{-0.09}$	$-0.74^{+0.11}_{-0.11}$	$3.12^{+2.63}_{-3.12}$	$-0.00^{+0.32}_{-0.47}$	$11.2^{+11.1}_{-11.2}$	⋯	$2.36^{+0.57}_{-0.59}$	⋯	407/71	433
velocity/time-dependent	$-2.21^{+0.09}_{-0.09}$	$-0.75^{+0.11}_{-0.10}$	$2.51^{+2.08}_{-2.51}$	$0.20^{+0.38}_{-0.44}$	$6.56^{+7.32}_{-6.56}$	⋯	$0.15^{+0.27}_{-0.15}$	$-0.04^{+0.26}_{-0.24}$	1389/70	1419

Notes.
[a] The units of each parameter are given in Table 2.
[b] The χ^2 values and the degrees of freedom for the peak in the marginalized *posteriors*.

separately. Ultimately, this culminates in simultaneously fitting all of the observations with a single model.

4.1. Distribution Functions

MCMC methods are known to have difficulties converging in high-dimensional parameter spaces. Therefore, we begin by only considering the observed distributions, and seed the chains with the solution from a bounded optimization with the Levenburg–Marquardt method (Newville et al. 2014, hereafter referred to as *LM results*) in order to expedite convergence of the chains. For all MCMC simulations, we use 20 walkers, burn 1000 steps, draw 3000 samples from the *posterior*, and thin the chains by a factor of 10:1, which results in 6000 random deviates. In Figure 5, we show the kernel-density estimation of the *posterior* for the observed distributions described in Section 3.1. We record the peak positions of the marginalized *posterior* distributions and their 68.3% upper and lower confidence intervals in Table 3. We find two notable results from this preparatory simulation: the power-law index (β) of the SFH is largely unconstrained, and the power-law index of the IMF is well-localized, $\alpha = -0.71 \pm 0.11$. In Figure 6, we show the fit of our model to the observed distributions.

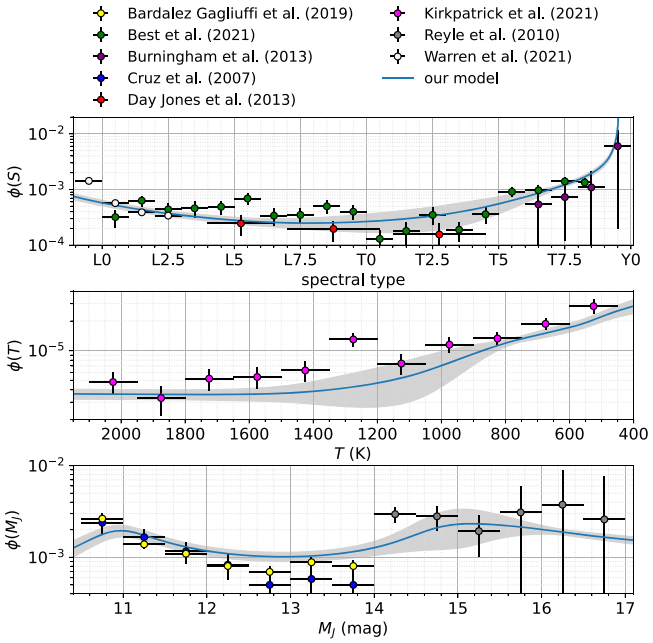

Figure 6. Best-fit model to observed distributions. The blue line represents the model for the peak in the marginalized *posteriors* from the MCMC fit (see Figure 5) for $\theta = -2.25, -0.67, 7.64, -0.14$) (see Table 3). The light gray region indicates the 1σ uncertainty derived from the covariance from the MCMC chains. The legend in the upper margin details the authorship of the data points by color.

4.2. Kinematic Data

Next, we expand the parameter space to include the total velocity dispersions, but start these MCMC simulations on the peaks of the marginal distributions from the observed-distribution-only analysis (see Figure 5). We consider each of the AVR models (Equations (7), (8), and (9)) in turn, and present their best-fit model parameters in Table 3. To distinguish between the different model AVRs, each with their own degrees of freedom, we consider the Bayesian information criterion (BIC) associated with the goodness of fit and degrees of freedom: $\mathrm{BIC} = k \ln N - 2 \ln \mathcal{L}$, where k and N represent the number of parameters and the number of measurements in the model, respectively. We find that the constant-diffusion model (see Figure 7) has the lowest BIC of the kinematic models, and so it becomes the prime focus of our interpretations. However, in all cases, we find that the constraints on the power-law index of the SFH become much tighter with $\beta \lesssim 4$, which is formally consistent at the 1σ level with an exponentially declining model. For this family of declining SFHs, we find the exponential timescale is $t^* \sim 1$ Gyr; however we remain cautious of overinterpreting this result, given the sizeable uncertainties on the SFH. Similarly, the initial velocity dispersion shows a sizeable uncertainty, which is a direct consequence of the inconsistencies in published velocity dispersions (see upper right panel of Figure 8). Nevertheless, the kinematic data have a clear role to play in further refining the model.

We assessed the convergence for every parameter of every MCMC simulation by considering the autocorrelation timescale for the parameter. We find that the autocorrelation function tends to a constant after ∼200 steps, but this quick convergence is not surprising given that each chain was started at the maximum of the *posterior* identified by the LM-fit. Furthermore, in all cases, the MCMC procedure does not find that the position of the peak in the *posterior* is considerably different than that determined by the LM-fit. To restate the above, the MCMC simulations are important for assessing degeneracies in the model, which improves our overall understanding of the uncertainties on the individual parameters. Based on these arguments, we believe that the MCMC chains have converged and robust conclusions based on our quoted uncertainties can be made.

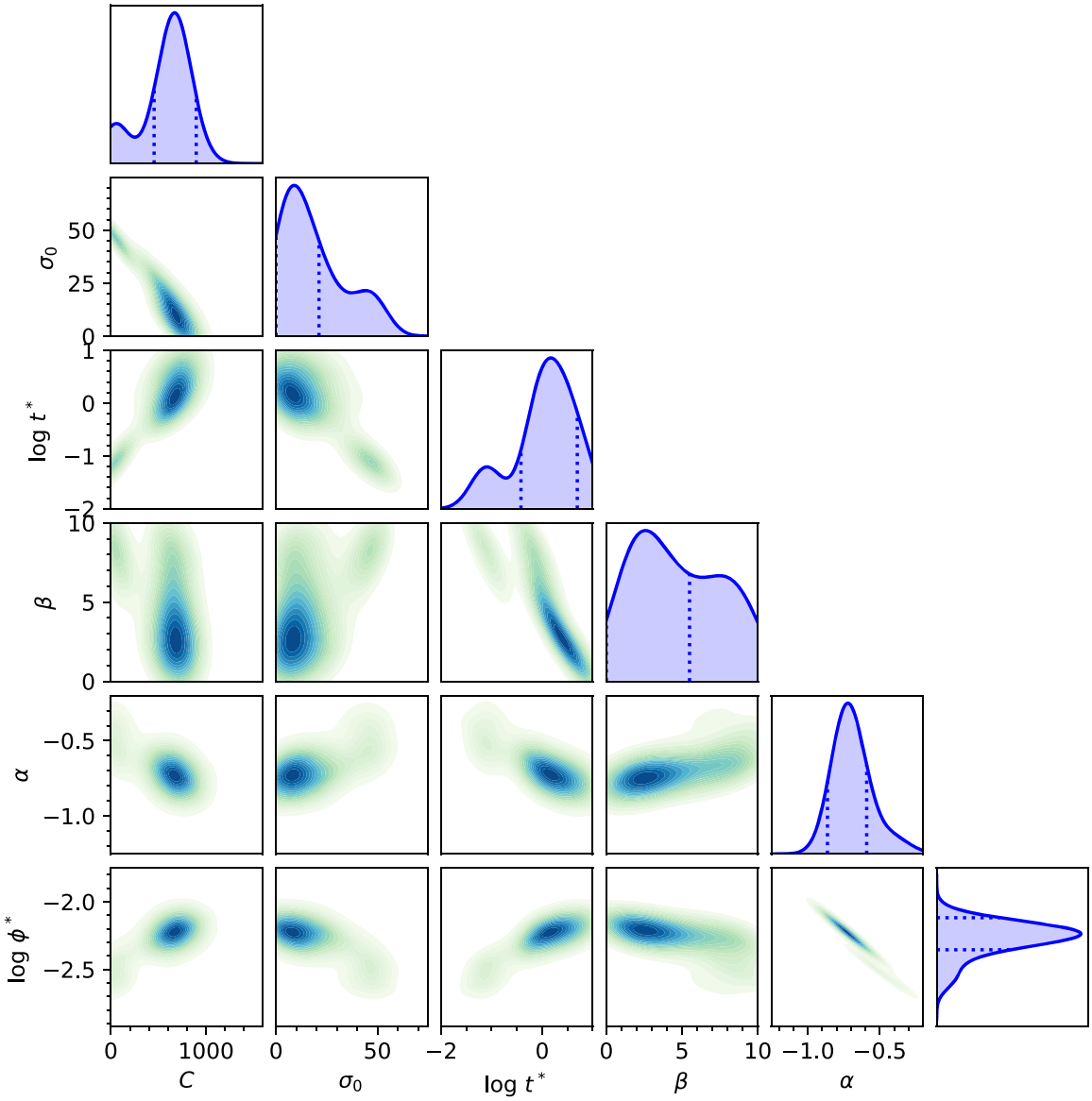

Figure 7. Corner plot for simultaneous fit of observed distributions and total velocity dispersions with a constant diffusion term: Equation (7). The plot symbols have the same meaning as in Figure 5.

5. Discussion

The observed distributions alone are capable of placing tight constraints on the power-law index of the IMF of $\alpha = -0.71 \pm 0.11$, which has been examined before. For example, our approach is very similar to the Bayesian model developed by Allen et al. (2005) to examine the *J*-band luminosity functions from Burgasser (2002), Cruz et al. (2007). Although they considered a larger suite of IMF forms, including a two-component power law and a truncated power law, we find considerably tighter constraints on the power-law index (for a single-component power law): -0.3 ± 0.6 versus -0.67 ± 0.11. Additionally, Kirkpatrick et al. (2021) analyze their observed temperature distribution from 525 brown dwarfs, and find a power-law index of $\alpha = -0.6 \pm 0.1$. Our improvement with respect to these previous studies is almost certainly related to the substantially larger data set that we have explored[8] or a more robust statistical treatment of the uncertainties (Kirkpatrick et al. 2021 consider only a coarse sampling in α and identify the range where their χ^2 values show no significant difference).

Here, we consider the sources of systematic uncertainty that we have tacitly ignored, the foremost of which concerns the aggregation of published results. This is most evident in our estimated rescaling of a factor of 1.4 for the Bardalez Gagliuffi et al. (2019) based on the reported pessimistic completeness corrections (Kirkpatrick et al. 2021, described in Table 1). Second, we have explicitly focused on brown dwarfs by calibrating the transformations to observables around sources with high-surface gravity $\log g \gtrsim 5$, as most published results contain few low-gravity objects. But this discrepancy between the observations with few low-gravity objects and our simulations calibrated to only high-gravity objects may explain some differences in our findings. The largest deviations between our derived cubic splines (red lines in Figure 3) occur for $M_J \lesssim 12$ mag or $T \gtrsim 1800$ K. In general, these portions of our model space (Figure 6 or Figure 8) do not show any

[8] The data analyzed by Allen et al. (2005) and Kirkpatrick et al. (2021) are included in our study.

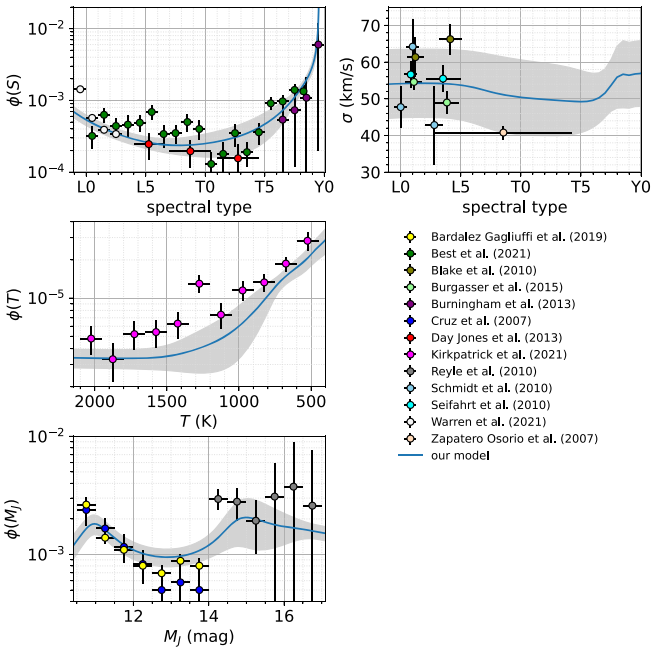

Figure 8. Best-fit model to the combined observed distributions and kinematic data. The plot symbols have the same meaning as in Figure 6, but the model here represents the constant AVR from Equation (7) with $\theta = (-2.23, -0.72, 2.06, 0.33, 8.60, 651)$.

significant differences between the data and the model. In fact, the regions of the largest differences occur at $1000 \lesssim T \lesssim 1400$ K, where the spline fits seem representative of the underlying population. A final systematic comes from uncertainties in the calibration data of Filippazzo et al. (2015), where there is a clear scatter around the cubic splines (the red line in Figure 3). Based on the rms of the residuals to those interpolants, we estimate that the systematic errors are $\sigma_J \sim 0.4$ mag and $\sigma_S \sim 1.5$. In Appendix D, we give a rough estimate of the consequences these systematic terms have for our derived parameters, and we leave the full propagation of these uncertainties to future work.

There are several noteworthy portions of parameter space where our model fails to accurately reproduce the observed data, and they seem to be related to the L/T transition, which may implicate the evolutionary models. First, our model predicts a factor of ∼10 too few brown dwarfs at $T \sim 1300$ K, compared to the findings of Kirkpatrick et al. (2021, see middle panels of Figures 6 and 8). Second, there appears to be a discontinuity in the observed J-band luminosity function around $J \approx 14$ mag that is not completely reflected in the model (see lower panel in Figures 6 and 8). But, as a general rule, our model tends to *underpredict* the observations, which may not be surprising, since the infrared color–magnitude diagram for field brown dwarfs shows a flattening around MKO $J \sim 14$ mag (e.g., Dupuy & Liu 2012), and theoretical isochrones often struggle to reproduce this sharp feature (for example, see Figure 11 of Phillips et al. 2020). However, we note that the discontinuity in the J-band luminosity function conspicuously occurs between three data sets (Cruz et al. 2007 and Bardalez Gagliuffi et al. 2019 with Reylé et al. 2010), which may hint at some systematic in the observations (such as subtle differences in the magnitude system or otherwise unaddressed completeness corrections). Finally, we note that there is considerable disagreement in the kinematic data, whether compared to our model or other published results. These differences are likely a property of the observations and not our model, as the kinematic samples are rarely unbiased and/or complete.

We find that the constant-diffusion AVR model best reproduces the observed distribution and kinematic data with $\Delta\mathrm{BIC} \gtrsim 10$ (compared to other AVR relations), with brown dwarfs requiring an initial velocity dispersion of $\sigma_0 \sim 9.0^{+11}_{-9.0}$ km s^{-1}. Luhman (2012) reviews several formation scenarios for brown dwarfs, specifically highlighting the contrasting predictions for the initial velocity dispersion. To establish a baseline for comparison, we determine the initial (total) velocity dispersion from the kinematic data from LAMOST-Gaia (Yu & Liu 2018). Since the brown dwarfs are almost certainly associated with the thin disk, we only consider the chemically defined thin disk from Yu & Liu (2018) (i.e., $|z| < 270$ pc and $\tau \leqslant 7$ Gyr). We fit Equation (7) for the constant-diffusion AVR model to their total velocity dispersions for all stars of all metallicities, and find $\sigma_{0,MS} = 13.4 \pm 3.4$ km s^{-1} and $C_{\rm MS} = 390 \pm 25$ (km s)$^{-2}$/Gyr, which suggests that brown dwarfs are likely formed with initial velocity dispersions comparable to those of main-sequence stars. However, we remain cautious of overinterpreting these results, given the considerable uncertainty on our inferred initial velocity dispersions for brown dwarfs, which is likely related to the inherent difficulties in obtaining complete and unbiased kinematic measurements. Instead, we reflect on these results as highlighting the degree to which a single model single model can (or cannot) reproduce a range of observations, whether distribution functions or kinematics. Future studies of disk brown dwarfs will be able to build on this approach and extend similar findings on young stellar systems (e.g., White & Basri 2003; Joergens 2006; Luhman et al. 2006; Parker et al. 2011) to the Galactic scale.

While the importance of kinematic data cannot be overstated, these measurements are challenging to obtain free of sample biases (for example, see Hsu et al. 2021, for an illustration of the difficulties). Therefore, Ryan et al. (2017) suggest using star counts to estimate the thickness of the thin disk ($z_{\rm scl}$) that can be linked to the vertical velocity dispersion ($\sigma_{\rm W}$) for a self-gravitating disk: $\sigma_W^2 \propto 4\pi G z_{scl}$, where the constant of proportionality is of order unity and is based on the disk shape parameter n (van der Kruit 1988). In this way, if wide-field surveys can provide robust spectral types from high-fidelity photometry to distinguish brown dwarfs from other astronomical sources, then they can probe a similar Galactic-scale physics with different biases or sample completeness. To this end, the large surveys with the Nancy Grace Roman Space Telescope (near-infrared imaging and slitless spectroscopy), Euclid (near-infrared imaging and slitless spectroscopy), and/or the Rubin Observatory (optical imaging) may be able to catalog brown dwarfs $\gtrsim$1 kpc (depending on spectral type).

6. Conclusions

We find a power-law IMF and a constant-diffusion AVR with an initial velocity dispersion of $\sigma_0 \sim 9.0^{+11}_{-9.0}$ km s^{-1}, which is perhaps smaller or comparable to that of main-sequence stars (of 13.4 ± 3.4 km s^{-1}; see Yu & Liu 2018). Taken together, these results suggest that brown dwarfs and main-sequence stars form via very similar processes, and the Reipurth & Clarke (2001) model with enhanced velocity dispersions for brown dwarfs is generally supported. Future work may utilize large-area surveys to map the thickness of the

thin disk of the Milky Way, which probes similar physics as the kinematics, but subject to different selection biases.

We would like to thank S. Schmidt and J. Lunine for helpful discussions and advice. We are extremely grateful to the anonymous referee for a careful and thorough reading and critique of this paper.

Software: `emcee` (Foreman-Mackey et al. 2013), `lmfit` (Newville et al. 2014).

Appendix A
A Log-normal IMF

Above, we exclusively considered a single-component power-law IMF. Here, we briefly describe the results for a log-normal IMF of the form:

$$\phi(m)\,dm = \frac{\phi^*}{m \ln s \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln m - \ln m^*}{\ln s}\right)^2\right] dm. \tag{A1}$$

This model has three parameters: (ϕ^*,$\ln m^*$,$\ln s$), therefore we use the ΔBIC to distinguish between the two functional forms. We find that, for all of the AVR models, the $\Delta\mathrm{BIC} = \mathrm{BIC}_{\log-\mathrm{normal}} - \mathrm{BIC}_{\mathrm{power-law}} \gtrsim 800$, which strongly disfavors the log-normal model.

Appendix B
The Effect of the Assumed the Milky Way Disk Age

We do not attempt to marginalize over the age of the Milky Way disk, as there are far better data sets to establish this parameter. However, we considered two alternate values, $\tau_{\mathrm{MW}} = 8$ Gyr and 12 Gyr, to estimate the sensitivity of our adopted value (10 Gyr) on the results. If we briefly consider the observed distribution functions and the LM results, then we find only a marginal change in our observed parameters (i.e., the first row of Table 3). For example, the changes in the goodness of fit with respect to the $\tau_{\mathrm{MW}} = 10$ Gyr assumption are $\Delta\chi^2 = -0.03$ and $+0.92$, respectively. While this is far from a thorough statistical exploration, it does not indicate a strong dependence on the assumed disk age.

Appendix C
Consideration of Alternative Evolutionary Models

In the main body of this work, we focused on the Burrows et al. (1997) models because they provide a dense and broad sampling of points in the mass–age plane and extend to modest temperatures. In this way, we can probe the largest range of the observed distribution functions. However, here we consider five additional evolutionary models (Burrows et al. 2001; Baraffe et al. 2003; Saumon & Marley 2008; Marley & Saumon 2020; Phillips et al. 2020). While there is broad qualitative agreement in the general characteristics of the cooling model $T(\tau, m)$, the subtle differences may lead to unexpected changes. For simplicity, we only consider the distribution functions (i.e., we ignore kinematics) and work with the LM results (i.e., we skip the MCMC simulations, as their primary role is to characterize the multivariate uncertainty distribution). We restrict our ranges to the temperature, mass, and age regions where all models have dense sampling, and we find largely consistent results: $-1.0 \lesssim \alpha \lesssim -0.7$, with higher uncertainties $\delta\alpha \sim 0.2$. As noted in the text, this increase in uncertainty is likely due to reducing the number of observed data points.

Appendix D
On the Role of Uncertainty in the Transformation Functions

As described above, the scatter around the spline models to the transformation functions in Figure 3 will introduce a systematic uncertainty on our derived model parameters. Here, we provide a rough estimate of those uncertainties, and defer to future work a more complete treatment that marginalizes over the spline-model parameters by including that step in the MCMC simulation(s). As noted, the respective standard deviations of the residuals around the spline model are $\sigma_J \sim 0.4$ mag and $\sigma_S \sim 1.5$. We add normal random variables with zero mean and these standard deviations to the observations in the lower and upper panels of Figure 6, respectively. We repeat the LM-fitting procedure and repeat for 10^3 to establish a distribution of output parameters. Based on these Monte Carlo simulations, we find the typical uncertainty on the derived model parameters to be $\sigma_\theta \sim (0.3, 0.3, 0.0, 0.4)$, where $\theta = (\log \phi^*, \alpha, \beta, \log t^*)$.

ORCID iDs

R. E. Ryan, Jr. https://orcid.org/0000-0003-0894-1588
C. Aganze https://orcid.org/0000-0003-2094-9128
A. J. Burgasser https://orcid.org/0000-0002-6523-9536
S. H. Cohen https://orcid.org/0000-0003-3329-1337
N. P. Hathi https://orcid.org/0000-0001-6145-5090
B. Holwerda https://orcid.org/0000-0002-4884-6756
N. Pirzkal https://orcid.org/0000-0003-3382-5941
R. A. Windhorst https://orcid.org/0000-0001-8156-6281

References

Allen, P. R., Koerner, D. W., Reid, I. N., & Trilling, D. E. 2005, ApJ, 625, 385
Aumer, M., & Binney, J. J. 2009, MNRAS, 397, 1286
Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, A&A, 402, 701
Bardalez Gagliuffi, D. C., et al. 2019, ApJ, 883, 205
Best, W. M. J., Liu, M. C., Magnier, E. A., & Dupuy, T. J. 2020, AJ, 159, 257
Best, W. M. J., Liu, M. C., Magnier, E. A., & Dupuy, T. J. 2021, AJ, 161, 42
Blake, C. H., Charbonneau, D., & White, R. J. 2010, ApJ, 723, 684
Burgasser, A. J. 2002, PhD thesis, California Institute of Technology
Burgasser, A. J. 2004, ApJS, 155, 191
Burgasser, A. J., Logsdon, S. E., Gagné, J., et al. 2015, ApJS, 220, 18
Burningham, B., Cardoso, C. V., Smith, L., et al. 2013, MNRAS, 433, 457
Burrows, A., Hubbard, W. B., & Lunine, J. I. 1989, ApJ, 345, 939
Burrows, A., Hubbard, W. B., Lunine, J. I., & Liebert, J. 2001, RvMP, 37, 719
Burrows, A., Marley, M., Hubbard, W. B., et al. 1997, ApJ, 491, 856
Cruz, K. L., Reid, I. N., Kirkpatrick, J. D., et al. 2007, AJ, 133, 439
Day-Jones, A. C., Morocco, F., Pinfield, D. J., et al. 2013, MNRAS, 430, 1171
Deacon, N. R., & Hambly, N. C. 2006, MNRAS, 298, 387
Dehnen, W., & Binney, J. J. 1998, MNRAS, 298, 387
Dupuy, T. J., & Liu, M. C. 2012, ApJS, 201, 19
Filippazzo, J. C., Rice, E. L., Faherty, J., et al. 2015, ApJ, 810, 158
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Goodman, J., & Weare, J. 2010, Comm. App. Math. Comp. Sci., 5, 65
Grand, R. J. J., Springl, V., Gómez, F. A., et al. 2016, MNRAS, 459, 199
Hsu, C.-C., Burgasser, A. J., Theissen, C. A., et al. 2021, ApJS, 257, 45
Joergens, V. 2006, A&A, 448, 655
Kirkpatrick, J. D., Gelino, C. R., Faherty, J. K., et al. 2019, ApJS, 240, 19
Kirkpatrick, J. D., Martin, E. C., mart, R. L., et al. 2021, ApJS, 253, 7
Luhman, K. L. 2012, ARA&A, 50, 65

Luhman, K. L., Whitney, B. A., Meade, M. R., et al. 2006, ApJ, 647, 1180
Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019, MNRAS, 489, 176
Marley, M., & Saumon, D. 2020, Sonora 2018: Cloud-free, solar C/O substellar evolution and photometry, Zenodo, v1.0, doi:10.5281/zenodo.3733843
Mayor, M. 1974, A&A, 32, 321
Newville, M., Otten, R., Nelson, R., et al. 2014, lmfit/lmfit-py v0.9.11, Zenodo, doi:10.5281/zenodo.11813
Parker, R. J., Bouvier, J., Goodwin, S. P., et al. 2011, MNRAS, 412, 2489
Phillips, M. W., Tremblin, P., Baraffe, I., et al. 2020, A&A, 637, 38
Quinn, P. J., Hernquist, L., & Fullagar, D. P. 1993, ApJ, 403, 74
Reipurth, B., & Clarke, C. 2001, AJ, 122, 432
Reylé, C., Delorme, P., Willott, C. J., et al. 2010, A&A, 522, 112
Ryan, R. E., Jr., Thorman, P. A., Schmidt, S J., et al. 2017, ApJ, 847, 53
Saumon, D., & Marley, M. S. 2008, ApJ, 689, 1327
Schmidt, S. J., Cruz, K. L., Bongiorno, B., Liebert, J., & Reid, J. 2007, AJ, 133, 2258
Schmidt, S. J., West, A. A., Hawley, S. L., & Pineda, J. S. 2010, AJ, 139, 1808
Seifahrt, A., Reiners, A., Almaghrbi, K. A. M., & Basri, G. 2010, A&A, 512, 37
Spitzer, L., Jr., & Schwarzschild, M. 1951, ApJ, 114, 385
van der Kruit, P. C. 1988, A&A, 192, 117
von Hoerner, S. 1959, LIACo, 9, 136
Warren, S., Ahmed, S., & Laithwaite, R. 2021, OJAp, 4, 4
White, R. J., & Basri, G. 2003, ApJ, 582, 1109
Wielen, R. 1974, HiA, 3, 395
Wielen, R. 1977, A&A, 60, 263
Yu, J., & Liu, C. 2018, MNRAS, 475, 1093
Zapatero Osorio, M. R., Martín, E. L., Béjar, V. J. S., et al. 2007, ApJ, 666, 1205