A STOCHASTIC THEORY OF THE HIERARCHICAL CLUSTERING
I. HALO MASS FUNCTION

Andrea Lapi1,2,3,4, Luigi Danese1,2

Accepted by ApJ

Abstract

We present a new theory for the hierarchical clustering of dark matter (DM) halos based on stochastic differential equations, that constitutes a change of perspective with respect to existing frameworks (e.g., the excursion set approach); this work is specifically focused on the halo mass function. First, we present a stochastic differential equation that describes fluctuations in the mass growth of DM halos, as driven by a multiplicative white (Gaussian) noise dependent on the spherical collapse threshold and on the power spectrum of DM perturbations. We demonstrate that such a noise yields an average drift of the halo population toward larger masses, that quantitatively renders the standard hierarchical clustering. Then, we solve the Fokker-Planck equation associated to the stochastic dynamics, and obtain the Press & Schechter mass function as a (stationary) solution. Moreover, generalizing our treatment to a mass-dependent collapse threshold, we obtain an exact analytic solution capable of fitting remarkably well the $N$-body mass function over a wide range in mass and redshift. All in all, the new perspective offered by the theory presented here can contribute to better understand the gravitational dynamics leading to the formation, evolution and statistics of DM halos across cosmic times.

Subject headings: Cosmology (343) — Dark matter (353)

1. Introduction

The halo mass function, namely the statistics describing the number of dark matter (DM) halos per unit comoving volume as a function of halo mass and redshift, is a fundamental quantity in astrophysics and cosmology (see textbooks by Mo et al. 2010 and Cimatti et al. 2020). For example, it is a basic ingredient to develop sensible galaxy formation and evolution models (see reviews by Silk & Mamon 2012 and Naab & Ostriker 2017), and it is routinely exploited in cosmological studies relying on the abundance and clustering of collapsed objects and of large-scale structures (see reviews by Frenk & White 2012 and Wechsler & Tinker 2018).

Clearly, the halo mass function can be estimated via high-resolution, large-volume, cosmological $N$-body simulations (see Sheth & Tormen 1999; Jenkins et al. 2001; Warren et al. 2006; Tinker et al. 2008; Crocce et al. 2010; Bhattacharya et al. 2011; Watson et al. 2013). However, given the natural limits on resolution, computational time, and storing capacity, it can be probed only in limited mass and redshift ranges. Moreover, the results of simulations depend somewhat on the algorithm used to identify collapsed halos (e.g., friend-of-friend vs. spherical overdensity), and on specific parameters related to the identification of isolated objects (e.g., the linking length). On the other hand, to estimate the halo mass function from observations is even more challenging (e.g., Castro et al. 2016; Dong et al. 2019; Somnfeld et al. 2019; Li et al. 2020), given the statistical and systematic uncertainties that arise when linking the observable quantities to the halo mass. Therefore, a deep theoretical understanding on how the halo mass function is originated from first principles is of crucial importance.

The modern theoretical framework to address the issue was born with the seminal work by Press & Schechter (1974); these authors were able to compute an analytic expression for the halo mass function by prescribing that a halo would collapse if it resided within a sufficiently overdense region of the initial (Gaussian) perturbation field. Given that the overdensity around a spatial location depends on scale, they recognized that the halo abundance is simply related to the mass fraction in the density field, smoothed on different scales, which is above a critical threshold for collapse. A drawback of this approach is the so called ‘cloud-in-cloud’ problem, i.e., attention must be paid not to double count overdense regions embedded within a larger collapsing perturbation; in other words, one has to consider only the perturbations that overcome the threshold on a given smoothing scale, but not on a larger one.

The problem was solved with the development of the excursion set framework by Bond et al. (1991), which still nowadays constitutes the standard theory. This envisages the overdensity around a given spatial location to execute a random walk when considered as a function of the smoothing scale; if the smoothing is performed with a sharp filter in Fourier space, the walk is Markovian. The collapse threshold here plays the role of a barrier, and the halo mass function is related to the distribution of first crossing, i.e., the probability that a walk crosses the barrier for the first time on a specific scale. In the original theory, the collapse threshold was gauged on the spherical collapse model of DM perturbations (Gunn & Gott 1972), and as such it was assumed to be independent on halo mass; subsequent developments adopted a mass-dependent threshold, inspired from the ellipsoidal
collapse model (see Sheth & Tormen 2002), in order to better reproduce the results of N-body simulations.

The excursion sets approach was then exploited to derive the 'conditional' halo mass function (see Lacey & Cole 1993), describing the mass and redshift distribution of a halo's progenitors, to build up Monte Carlo realizations of the merging process known as merger trees (see Kauffmann & White 1993; Somerville & Kolatt 1999; Cole et al. 2000; Parkinison et al. 2008), and to develop models for the large-scale halo bias (see Mo & White 1996; Sheth & Lemson 1999). Further, more recent refinements include non-Markovian walks (Maggiore & Riotto 2010a; Musso & Sheth 2012), stochastic collapse thresholds (Maggiore & Riotto 2010b; Corasaniti & Achitouv 2011), extension to peaks theory (see Paranjape et al. 2012), descriptions of the void distribution (see Sheth & van de Weygaert 2004; Jennings et al. 2013), and non-standard cosmologies (e.g., von Braun-Bates & Devriendt 2018; Lovell 2020).

Despite this rich literature focused on the theoretical foundations and a number of undoubted successes in practical applications, the excursion set framework is known to hide some pitfalls and drawbacks: no exact analytic expression of the mass function for a general mass-dependent collapse threshold is known, but for very simple shapes (see Zhang & Hui 2006; Lapi et al. 2013); the merging kernel associated to the excursion set theory is not symmetric, and this causes a mathematical inconsistency or at least an ambiguity in the definition of the merger rates (see Benson et al. 2005; Neistein & Dekel 2008; Zhang et al. 2008); even adopting a mass-dependent collapse threshold and other refinements, the excursion set formalism struggles to reproduce the halo mass function, as driven by a multiplicative white (Gaussian) noise dependent on the spherical collapse threshold; in Sect. 2.2 we point out that the solution is stationary when the original equation is written in convenient variables.

In this paper we submit a new theory of the hierarchical clustering and halo mass function based on stochastic differential equations in real space, that constitutes a change of perspective with respect to the excursion set formalism. First, we invent a stochastic differential equation that describes fluctuations in the mass growth of DM halos, as driven by a multiplicative white (Gaussian) noise dependent on the spherical collapse threshold and on the power spectrum of DM perturbations; in this approach it is the mass (or mass variance) in a given region of the Universe to perform a (Markovian) random walk as a function of cosmic time. In Sect. 2 we demonstrate that the noise yields an average drift toward larger masses, that quantitatively renders the standard hierarchical clustering. Then, in Sect. 2.1 we solve the Fokker-Planck equation associated to the stochastic dynamics, and obtain as a solution the Press & Schechter mass function; in Sect. 2.2 we point out that the solution is stationary when the original equation is written in convenient variables.

In Sect. 3 we introduce a minimal modification of the stochastic equation in terms of a mass-dependent collapse threshold. Using a parametric shape analogous to that adopted in the excursion set framework, we obtain a closed-form analytical solution of the associated Fokker-Planck equation. Remarkably, such a solution has a shape similar to the empirical fitting formula introduced since Sheth & Tormen (1999); moreover, for specific values of the parameters describing the mass dependence of the collapse threshold, our result reproduces remarkably well the N-body mass function over an extended range of masses and redshifts.

As an aside, in Sect. 4 we explore how to generalize our framework when a colored rather than a white noise is considered, so as to enforce a non-Markovian evolution. Adopting for definiteness a multiplicative Ornstein-Uhlenbeck noise and a constant threshold for collapse, we are able to solve exactly the corresponding Fokker-Planck equation and obtain a closed form solution. With respect to the white noise case, the redshift evolution of the mass function is found to be modified somewhat, in a fashion depending on the correlation time characterizing the noise. Finally, in Sect. 5 we summarize our findings and envisage possible outlooks.

Throughout this work, we adopt the standard flat ΛCDM cosmology (Planck Collaboration 2018) with rounded parameter values: matter density ΩM ≈ 0.3, dark energy density ΩΛ ≈ 0.7, baryon density Ωb ≈ 0.05, Hubble constant H0 = 100h km s⁻¹ Mpc⁻¹ with h ≈ 0.7, and mass variance σ8 ≈ 0.8 on a scale of 8 h⁻¹ Mpc. The most relevant expressions are highlighted with a box.

## 2. A STOCHASTIC EQUATION FOR THE HIERARCHICAL CLUSTERING

Our proposal is to capture the essence of the hierarchical clustering via the following nonlinear stochastic differential equation:

\[
\frac{d}{dt} \ln \sigma^{-1} = \frac{\sigma}{\delta_c} \left( \frac{\delta_c}{\sigma} \right)^{1/2} \eta(t),
\]

or equivalently, in terms of mass:

\[
\frac{d}{dt} M = \frac{\sigma^2}{|\delta \sigma / \delta c|} \left( \frac{\delta_c}{\sigma} \right)^{1/2} \eta(t).
\]

Here \(\eta(t)\) is a Gaussian white noise (physical dimension 1/√(time)) with ensemble-average properties \(\langle \eta(t) \rangle = 0\) and \(\langle \eta(t) \eta(t') \rangle = 2 \delta_\delta(t-t')\), where \(\delta_\delta\) is the Dirac delta-function (the factor 2 is only a convention and clearly it could be reabsorbed into the multiplicative term). The quantity \(\delta_c(z) = \delta_c D(0)/D(z)\) is the critical threshold for collapse extrapolated from linear perturbation theory; in a flat Universe (see Mo et al. 2010; Weinberg 2008), one can use the approximations \(\delta_c \simeq \frac{3}{4\pi} (12\pi)^{2/3} [1 + 0.0125 \log_{10} \Omega_M(z)] \approx 1.68\) and \(D(z) \approx \)

5 Throughout the paper we adopt, in line with the majority of the physics community, the Stratonovich convention; this allows to use the rules of ordinary calculus on stochastic variables, at the price of originating a noise-induced drift term in the Fokker-Planck coefficients associated to a given stochastic differential equation (see Appendix A). The alternative Itô convention, mostly used by mathematicians, removes such a noise-induced drift, but requires to develop new rules for the differential calculus of stochastic variables.
A STOCHASTIC THEORY OF HIERARCHICAL CLUSTERING I. 3

stochastic variables that fluctuate over cosmic time
function of \( M \), \( \delta \) the influence of the noise. We stress the change of but both \( \sigma \)
et al. 1986\), standard cold dark matter power spectra (e.g., Bardeen
here \( P \)
\( \sigma \) cutes a random walk as a function of the mass variance
with \( \Omega \)
\( M \)
\( \sigma \)
\( \delta \)

| Fig. 1.— Euler-Maruyama integration of the stochastic differential Eq. (1), yielding the evolution of the mass variance \( \ln \sigma \) (left y–axis) and of the mass \( M \) (right y–axis) as a function of cosmic time \( t \); the initial condition \( M \approx 10^4 M_\odot \) at \( z \approx 10 \) (corresponding to \( t \sim 0.5 \) Gyr) has been adopted. The grey lines are 30 randomly chosen evolutionary tracks, while the red line and shaded area illustrate the median and the quartiles over 3000 realization of the noise; the dashed blue line shows the evolution of the spherical collapse threshold \( \delta_c(t) \).

\[
\frac{5}{2} \frac{\Omega_M(z)}{1+z} \left[ \frac{1}{3} + \frac{209}{70} \Omega_M(z) - \frac{1}{70} \Omega_M^2(z) + \Omega_M^{1/7}(z) \right]^{-1}
\]

with \( \Omega_M(z) \equiv \Omega_M(1+z)^3/\left[\Omega_M + \Omega_M(1+z)^3\right] \). Finally, \( \sigma \) is the mass variance filtered on the mass scale \( M \):

\[
\sigma^2(M) = \frac{1}{(2\pi)^3} \int d^3k \, P(k) \, \hat{W}_M^2(k) ;
\]

here \( P(k) \) is the power spectrum of density fluctuation, and \( \hat{W}_M^2(k) \) is the Fourier transform of a window function whose volume in real space encloses the mass \( M \); for standard cold dark matter power spectra (e.g., Bardeen et al. 1986), \( \sigma(M) \) is an inverse, convex, slowly-varying function of \( M \). Note that in the present theory the relation \( \sigma(M) \) between \( \sigma \) and \( M \) is purely deterministic, but both \( M(t) \) and \( \sigma(t) = \sigma(M(t)) \) are to be considered stochastic variables that fluctuate over cosmic time \( t \) under the influence of the noise. We stress the change of perspective with respect to the standard excursion set formalism: in the latter the overdensity field \( \delta(\sigma) \) executes a random walk as a function of the mass variance \( \sigma \), which plays the role of a pseudo-time variable; here the mass \( M(t) \) or the mass variance \( \sigma(M(t)) \) are themselves stochastic variables, undergoing a Markovian evolution as a function of (real) time \( t \). Note that in the excursion set approach the choice of the filter function in Eq. (3) has a crucial impact, since the random trajectories \( \delta(\sigma) \) are Markovian only when a sharp filter in Fourier space is adopted; in the present theory, assuming a different filter function \( \sigma(t) \), the initial condition

\[
\ln \sigma(t_{i+1}) = \ln \sigma(t_i) + \frac{1}{2} \frac{\sigma^2(t_i)}{\delta_c^2(t_i)} \frac{\left| \delta_c(t_{i+1}) - \delta_c(t_i) \right|}{\delta_c(t_i)} \left| \frac{\sigma(t_i)}{\delta_c(t_i)} \right| \frac{\left| \delta_c(t_{i+1}) - \delta_c(t_i) \right|}{\delta_c(t_i)}^{1/2} w_i ,
\]

where \( w_i \) are random weights extracted from a normal distribution with mean zero and unit variance. In Fig. 1...
we show the resulting evolution of \( \ln \sigma(M(t)) \) as a function of cosmic time, with initial condition \( M \sim 10^4 M_\odot \) or \( \ln \sigma \sim 2.2 \) at \( z \sim 10 \) (reasonable initial conditions do not change significantly the outcome). The grey lines are 30 randomly chosen evolutionary tracks, while the red line and shaded area illustrate the median and the quartiles over 3000 realization of the noise, and the dashed blue line shows the evolution of the spherical collapse threshold \( \delta_c(t) \). It is easily seen that the noise induces a drift of \( \ln \sigma \), making it to decrease, and hence making the mass \( M \) to increase. Remarkably, after a burn-in period needed to erase memory of the initial condition, \( \sigma(M(t)) \) tends to copy the evolution of \( \delta_c(t) \); this means that the noise-induced drift effectively renders the standard hierarchical clustering of the halo population, as expressed by the increase with time of the characteristic mass \( M_c(t) \) set by the condition \( \sigma(M_c(t)) \sim \delta_c(t) \).

### 2.1. Fokker-Planck equation and the Press & Schechter mass function

We now look for the probability density \( \mathcal{P}(M,t) \) for a region to be in a state of mass between \( M \) and \( M + dM \) at time \( t \); this can be found by solving the Fokker-Planck equation associated to Eq. (2), which reads (see Appendix A for details)

\[
\frac{\partial}{\partial t} \mathcal{P}(M,t) = -T^2(t) \frac{\partial}{\partial M} D(M) D'(M) \mathcal{P}(M,t) + T^2(t) \frac{\partial^2}{\partial M^2} \left[ D^2(M) \mathcal{P}(M,t) \right],
\]

where we have defined the two quantities \( D(M) \equiv \sigma^2 / [d\sigma/dM] \) and \( T(t) \equiv |\delta_c^2|^{1/2} / \delta_c^3 \) such that \( M = D(M) T(t) \eta(t) \) factorizes the mass and time dependencies. The Fokker-Planck equation may also be written as a pure continuity equation \( \partial_t \mathcal{P} + \partial_M \mathcal{J} = 0 \) in terms of a probability current \( \mathcal{J}(M,t) = -T^2(t) D(M) \partial_M \mathcal{P}(M,t) \). The natural boundary conditions \( \lim_{M \to \infty} \mathcal{P}(M,t) = 0 \) and \( \mathcal{P}(M,0) = \delta_0(M) \) and the constraint \( \mathcal{P}(M,t) = 0 \) whenever \( M < 0 \) must apply; the latter corresponds to a reflective barrier condition \( \mathcal{J} \mid_{M=0} = D \partial_M (D P) \) if \( M=0 \) at the initial mass \( M = 0 \). Then the probability mass function \( \mathcal{P} \) is normalized as \( \int_0^\infty dM \mathcal{P} = 1 \) and thus it must be related to the halo mass function by

\[
\frac{dN}{dM dV}(M,t) = \frac{\bar{\rho}_M}{M} \mathcal{P}(M,t),
\]

in terms of the average comoving matter density \( \bar{\rho}_M \).

Now to solve the Fokker-Planck equation we employ the transformations:

\[
\begin{align*}
X & \equiv \int \frac{dM}{D(M)} = \frac{1}{\sigma} \\
Y & \equiv \int dt T^2(t) = \frac{1}{2 \delta_c^2} \\
W(X,Y) & \equiv D(M) \mathcal{P}(M,t).
\end{align*}
\]

Then Eq. (5) turns into

\[
\partial_Y W(X,Y) = \partial_X^2 W(X,Y),
\]

which is a standard diffusion equation. In terms of these new variables, the boundary conditions stated below Eq. (4) read \( \lim_{X \to \infty} W = 0 \), \( W(X,0) = \delta_0(X) \) and \( \partial_X W \mid_{X=0} = 0 \). These stem from the following circumstances: (i) for reasonable power spectra \( \sigma \) is a slowly-varying inverse function of \( M \), so that \( X \propto \sigma^{-1} \) tends to zero or infinity as \( M \) does; (ii) the collapse threshold \( \delta_c \propto D^{-1}(t) \) scales inversely with \( t \), so that \( Y \propto \delta_c^{-2} \) tends to zero as \( t \) does; (iii) finally, \( W(X,Y) = D(M) \mathcal{P}(M,t) \) vanishes for large \( X \) since \( \mathcal{P}(M,t) \) is expected to be exponentially suppressed for \( M \to \infty \) so overwhelming any slow (at most powerlaw) divergence of \( D(M) = \sigma^2 / [d\sigma/dM] \).

The solution of this differential problem is standard (it can be easily found via a Fourier transform) and writes

\[
W(X,Y) = \frac{1}{\sqrt{\pi Y}} e^{-X^2/4Y}.
\]

Coming back to the original variables we get

\[
\mathcal{P}(M,t) = \frac{1}{D(M) \sqrt{\pi Y(t)}} e^{-X^2(M)/4Y(t)},
\]

which after Eqs. (6) and (7) yields the Press & Schechter mass function:

\[
N(M,t) = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}_M \delta_c(t)}{M \sigma^2} \frac{d\sigma}{dM} e^{-\delta_c^2(t)/2 \sigma^2}.
\]

Three interesting remarks follow. First, note that the presence of the multiplicative noise in Eqs. (1) and (2) is fundamental in originating a second-order, diffusion-like term in the associated Fokker-Planck equation (see also Appendix A): from the derivation above it is seen that such a term yields the exponential cut-off of the mass function at the high-mass end. Second, one can easily compute the moments

\[
\left\langle \frac{(\delta_c/\sigma)^k}{\sigma} \right\rangle = \int_0^\infty dM \left( \delta_c/\sigma \right)^k \mathcal{P} = \frac{2^{k/2}}{\sqrt{\pi}} \Gamma \left( \frac{1+k}{2} \right).
\]

For scale-free power spectra \( M \propto \sigma^{-6/(n+3)} \) holds in terms of the effective spectral index \( n > -3 \), so that the above implies \( \langle M^k(t) \rangle \propto \delta_c^{-6k/(n+3)} \); this in turn scales as \( \langle M^k(t) \rangle \propto t^{4k/(n+3)} \) in the redshift range where \( \delta_c \propto t^{-2/3} \) applies. Third, it is interesting to note that the Fokker-Planck Eq. (5) highlights that the evolution of the probability density function is driven by the two terms on the r.h.s.: the first represents its noise-induced drift toward larger masses because of hierarchical collapses, and the second describes its diffusive reshaping
at the high-mass end due to the stochasticity in merging/accretion events.

2.2. Stationarity

An alternative derivation, that will be useful for the generalization in the next Section, is the following. We again start from Eq. (11) and change variable from $σ$ to $ν = \delta_c(t)/σ(M)$:

$$\frac{d}{dt} ν = -\frac{δ_c}{δ_e} ν + \frac{δ_e}{δ_c} η(t);$$  \hspace{1cm} (13)

thus now the variable $ν$ is seen to undergo a stochastic Ornstein-Uhlenbeck process. The corresponding Fokker-Planck equation reads

$$\partial_t P(ν, t) = \frac{δ_c}{δ_e} \partial_ν \left[ ν P(ν, t) + \partial_ν P(ν, t) \right].$$  \hspace{1cm} (14)

We set the boundary conditions $\lim_{ν → ∞} P = 0$ and $J_{|ν = 0} = -|\delta_c/δ_e| [ν P + \partial_ν P]|_{ν = 0} = 0$, implying $\int_0^∞ dν P = 1$; the former is the natural boundary condition for large values of $M$, which correspond to large $ν ∝ σ^{-1}$; the latter is the no-current boundary condition at the $ν = 0$ point enforced by the constraint $P(ν, t) = 0$ for $ν < 0$ being $ν$ positively defined.

Having incorporated the time dependent quantity $δ_e(t)$ into the new variable $ν$, under the ergodic hypothesis one expects that the relevant solution in terms of this variable should be stationary (e.g., Paul & Baschnagl 2013), i.e., it should satisfy $δ_t P(ν, t) = 0$. From Eq. (14), $P(ν, t) = P(ν)$ is determined by

$$ν P + \frac{d}{dν} P = 0,$$  \hspace{1cm} (15)

where the constant on the r.h.s. must be zero to satisfy the no-current boundary condition $J_{|ν = 0} = 0$. The solution to this simple ordinary differential equation with normalization $\int_0^∞ dν P(ν) = 1$ is $P = √{2/π} e^{-ν^2/2}$. The mass function writes as

$$N(M, t) = \frac{ρ_m}{M} \frac{dν}{dM} P(ν);$$  \hspace{1cm} (16)

given that $|dν/dM| = (δ_e/σ^2)|dσ/dM|$, this is indeed easily recognized to be again the Press & Schechter mass function.

One may wonder whether the general time-dependent solution $P(ν, t'|ν', t')$ of Eq. (14) for a generic initial condition $P(ν, t = t'|ν', t') = δ_D(ν - ν')$ converges to the stationary state, and over which timescale. To this purpose, we note that defining a new time variable $τ ∝ -ln δ_e(t)$ brings Eq. (14) into a form that can be easily solved via a Fourier transform; the fundamental solutions (a general result for Gaussian and Markovian variables known as Doob’s theorem) read $P ∝ e^{-(ν ± ν')^2/(1 - e^{-2τ})}/√{1 - e^{-2τ}}$. Taking into account the no-current boundary condition, and reverting to the original time variable one obtains

$$P(ν, t|ν', t') = \frac{1}{√{2π (1 - δ^2/δ')^2}} \left\{ \exp \left[ -\frac{1}{2} \frac{(ν - ν' δ/δ')^2}{1 - δ^2/δ'} \right] + \exp \left[ -\frac{1}{2} \frac{(ν + ν' δ/δ')^2}{1 - δ^2/δ'} \right] \right\},$$  \hspace{1cm} (17)

with $δ ≡ δ_e(t)$ and $δ' ≡ δ_e(t')$. Plainly, away from any initial condition, for $δ$ substantially lower than $δ'$, this converges to the stationary solution $P(ν, t)$ derived above. Such transitional states could be possibly related to deviation of the mass function from the self-similar shape Eq. (16) in terms of the variable $ν$. Note, in passing, that the transition probability Eq. (17) cannot be directly related to the halo conditional mass function of the extended Press & Schechter theory; we anticipate that to derive the latter a modification of Eq. (14) is needed, but the issue demands an extended analysis that will be presented in a forthcoming paper.

3. MASS-DEPENDENT THRESHOLD: THE $N$-BODY MASS FUNCTION

It is well known that the halo mass function derived from $N$-body simulations deviates substantially from the Press & Schechter shape (see reference in Sect. 1 for details); in particular, the former is flatter than the latter both at the high and at the low mass end, and evolves more slowly toward high redshift (see Fig. 2). This mismatch is usually cured by introducing a mass-dependent threshold for collapse:

$$δ_e(σ, t) = σ \sqrt{\bar{q} ν} \left[ 1 + \frac{β}{(\sqrt{q} ν)^2 γ} \right] \equiv σ B(ν) = δ_e B(ν)/ν,$$  \hspace{1cm} (18)

where $q$, $β$, and $γ$ are parameters to be set by fitting the $N$-body outcomes. Such a modified threshold is generally ascribed, though a bit naively, to the fact that perturbations undergo an ellipsoidal rather than a spherical collapse (Sheth & Tormen 2002; see also discussion by Mo et al. 2010). Note, however, that the above shape is quite general in describing a variety of phenomena that can influence the collapse, like tidal torques and angular momentum, cosmological constant, dynamical friction (see Del Popolo 2017 and references therein). In the excursion set framework, the parameters $q ≈ 0.707$, $β ≈ 0.47$ and $γ ≈ 0.615$ are required to fit the $N$-body mass function; however, there is some degeneracy, in that for example a square-root barrier with $q ≈ 0.55$, $β ≈ 0.5$ and $γ ≈ 0.5$ fits simulations equally well. We stress that values of $γ > 1/2$ are of some concern in the excursion set framework, because they imply that some walks $δ(σ)$ do not cross the barrier at all, an occurrence thought to represent fragmentation. Notice that no general analytical expression exists for the excursion set mass function, apart
from very particular barrier shapes (e.g., the constant, linear or square-root barriers; see for example Mahmood et al. 2006; Giocoli et al. 2007; Lapi et al. 2013) and in general one must recur to numerical solutions (see Zhang & Hui 2006).

In the framework presented here, we aim to show that a minimal modification of the basic Eq. (1), which incorporates a mass-dependent collapse threshold with shape analogous to the above Eq. (18), will yield a mass function in excellent agreement with N-body simulations. Moreover, we will provide an analytic expression valid for any triples of parameters $q$, $\beta$, and $\gamma < 1/2$, that in the limit $\nu \gg 1$ (i.e., large masses and/or early times) matches the empirical formula introduced since Sheth & Tormen (1999); this will establish a direct connection between the parameters describing the barrier and the shape of the halo mass function (which is absent in the excursion set approach). We start from Eq. (1) by replacing, in the multiplicative noise term, the standard spherical collapse threshold $\delta_c(\sigma,t)$ with the mass-dependent $\delta_c(\sigma,t)$ given above; thus now the ratio $\sigma/\delta_c(\sigma,t)$ modulates the noise toward enforcing collapse. We also retain the term $|\delta_c/\delta_c|^{1/2} \approx |D(t)/D(t)|^{1/2}$ involving the characteristic timescale for the linear growth of perturbations (see discussion in Sect. 2). When formulated in terms of $\nu$, Eq. (1) modified in such a way writes as

$$\frac{d \nu}{dt} = -\nu \left[ \frac{\nu}{B(\nu)} \left( \frac{\nu}{B(\nu)} \right) \right]^{1/2} \eta(t),$$

where $\delta_c(t)$ is the standard threshold for spherical collapse. The corresponding Fokker-Planck equation reads:

$$\partial_t \mathcal{P}(\nu,t) = \left[ \frac{\nu}{B(\nu)} \right] \left( \frac{\nu}{B(\nu)} \right) \eta(t),$$

with boundary conditions $\lim_{\nu \rightarrow \infty} \mathcal{P} = 0$ and $\mathcal{J}_{\nu=0} = -[\delta_c/\delta_c] [\nu \mathcal{P} + (\nu/B) \partial_\nu \mathcal{P} ]|_{\nu=0} = 0$, implying the constraint $\int_0^\infty d\nu \mathcal{P} = 1$.

In analogy to the procedure followed in Sect. 2.2, we look for stationary solutions $\mathcal{P}(\nu,t) = \mathcal{P}(\nu)$ with $\partial_t \mathcal{P} = 0$; one obtains the equation

$$\frac{\nu}{B(\nu)} \frac{d}{d\nu} \left[ \frac{\nu}{B(\nu)} \mathcal{P} \right] = -\nu \mathcal{P},$$

where the integration constant must be null to satisfy the no-current boundary condition $\mathcal{J}_{\nu=0} = 0$. The above equation can be easily solved by multiplying both sides by $1/B$ and recognizing that it becomes separable for the function $(\nu/B) \mathcal{P}$; we find

$$\mathcal{P}(\nu) = A \frac{B(\nu)}{\nu} \exp \left[ -\int \frac{B^2(\nu)}{\nu} d\nu \right],$$

where the normalization constant $A$ is determined by the condition $\int_0^\infty d\nu \mathcal{P}(\nu) = 1$. We stress that this result holds for any mass-dependent collapse threshold that can be expressed as $\delta_c(\sigma,t) \equiv \delta_c(t) B(\nu)/\nu$ in terms of the scaled variable $\nu$.

Specializing now to the shape of $B(\nu)$ from Eq. (18), we note that for $\nu \rightarrow 0$ the behavior $\nu \mathcal{P}(\nu) \propto B(\nu) \propto \nu^{1-2\gamma}$ applies; to satisfy the normalization constraint $\int_0^\infty d\nu \mathcal{P} = 1$ one must require $\gamma < 1/2$. Performing explicitly the integration, we get the closed form expression

$$\mathcal{P}(\nu) = A \sqrt{\frac{\beta}{(\sqrt{\gamma})^2}} \exp \left\{ -\frac{\nu^2}{2} \left[ \frac{1}{1-\gamma (\sqrt{\nu})^2} + \frac{\beta^2}{1-2\gamma (\sqrt{\nu})^2} \right] \right\}. $$

**TABLE 1**

### Solution Parameters for a Mass-Dependent Threshold

|          | $q$     | $\beta$ | $\gamma$ |
|----------|---------|---------|----------|
| ST99     | 0.62    | 0.16    | 0.37     |
| Bh+11    | 0.69    | 0.09    | 0.42     |
| Wa+13    | 0.69    | 0.12    | 0.37     |

*Note: Values of the parameters $q$, $\beta$ and $\gamma$ are set by fitting the Fokker-Planck solution Eqs. (23) and (24) to the halo mass function from the N-body simulations by Sheth & Tormen (1999; ST99), Bhattacharya et al. (2011; Bh+11) and Watson et al. (2013; Wa+13), see Sect. 2 for details.*
The resulting mass function just writes

\[ N(M, t) = \frac{\bar{\rho}}{M \sigma} \left| \frac{d\sigma}{dM} \right| \nu \mathcal{P}(\nu) ; \]  

incidentally, note that the multiplicity function \( f(\ln \sigma) \) used in some literature works is just \( f(\ln \sigma) = \nu \mathcal{P}(\nu) \).

We stress that the above is an exact expression, valid for any triple of values \( q, \beta, \gamma < 1/2 \); the Press & Schechter function is recovered for \( q = 1 \) and \( \beta = 0 \). Remarkably, the asymptotic behavior for \( \nu \gg 1 \), corresponding to large masses and/or early cosmic times, is seen to produce a shape akin to the empirical fit of \( N \)-body simulations adopted since Sheth & Tormen (1999); however, for finite \( \nu \) the terms in the exponential are important and must be taken into account.

We now use a Levenberg-Marquardt least-squares minimization routine to fit the multiplicity function \( \nu \mathcal{P}(\nu) \) to the simulation outcomes for FoF halos by Sheth & Tormen (1999), Bhattacharyya et al. (2011) and Watson et al. (2013), sampled from their (somewhat different) fitting formulas for \( -0.6 \leq \ln(\sigma(M) D(z))^{-1} \leq 1.3 \) in mass bins of 0.25 dex. For Sheth & Tormen (1999), we find best fit parameters \( q \approx 0.62, \beta \approx 0.16 \) and \( \gamma \approx 0.37 \), yielding a value of the normalization constant \( \mathcal{A} \approx 0.63 \). For Bhattacharyya et al. (2011), we obtain \( q \approx 0.69, \beta \approx 0.09 \) and \( \gamma \approx 0.42 \), yielding \( \mathcal{A} \approx 0.64 \). For Watson et al. (2013), we get \( q \approx 0.69, \beta \approx 0.12 \) and \( \gamma \approx 0.37 \), yielding \( \mathcal{A} \approx 0.66 \). These triples of values, reported for convenience in Table 1, are consistent within the uncertainties in the simulation results and in the fitting procedure.

In Fig. 2 we compare the mass function \( N(M, t) \) from Eqs. (23) and (24) to the \( N \)-body results by Bhattacharyya et al. (2011), finding an excellent agreement over a wide range of masses \( M \approx 10^8 - 10^{16} M_\odot \) and redshifts \( z \approx 0 - 10 \). In the same Figure we also plot for reference the Press & Schechter mass function; moreover, we show the mass function computed from the excursion set approach (following the numerical algorithm by Zhang & Hui 2006) with the barrier shape of Eq. (20) and the standardly adopted parameters \( q \approx 0.707, \beta \approx 0.47 \) and \( \gamma \approx 0.615 \) (see above).

4. COLORED NOISE AND NON-MARKOVIAN WALKS

In Nature, white noise is never found to be perfectly realized, but rather constitutes an idealization of the stochastic driving force affecting a physical phenomenon. Thus one may wonder whether the previous treatment can be extended to a colored instead of a white noise; this will correspondingly enforce a non-Markovian evolution of the system. To have a grasp on the impact of colored noise and non-Markovianity on the mass function, we consider a multiplicative stochastic Ornstein-Uhlenbeck process; for simplicity we adopt a mass-independent collapse threshold, so that the endpoint of this computation should be compared with the Press & Schechter mass function. Specifically, we modify our Eq. (2) into the two-dimensional stochastic system

\[
\begin{align*}
\dot{M} &= T(t) D(M) \Gamma(t) \\
\dot{\Gamma} &= -\omega \Gamma(t) + \zeta \eta(t)
\end{align*}
\]  

where, besides already defined quantities, \( \Gamma(t) \) is an Ornstein-Uhlenbeck noise with average \( \langle \Gamma(t) \rangle = 0 \) and a nontrivial correlation between different times

\[
\langle \Gamma(t) \Gamma(t') \rangle = \frac{\zeta^2}{\omega} \left[ e^{-\omega |t-t'|} - e^{-\omega(t+t')} \right] ,
\]

controlled by the parameters \( \omega \) and \( \zeta \) (both have physical dimension of 1/time); \( \zeta \) represents the degree of volatility, i.e. the sensitivity of the system to random changes, while \( \omega \) is the dissipation rate at which the system tends to reverse toward its zero mean. This is perhaps the simplest generalization of the white noise case, since \( \Gamma(t) \) is still Markovian, but of course \( M(t) \) is not. In the limit \( \zeta = \omega \to \infty \) one recovers a pure white-noise, since \( \langle \Gamma(t) \Gamma(t') \rangle \approx \lim_{\omega \to \infty} \omega e^{-\omega |t-t'|} = 2 \delta_D(t-t') \).

The Fokker-Planck equation regulating the dynamics of the probability density function \( \mathcal{P}(M, \Gamma, t) \) for the above system is

\[
\frac{\partial}{\partial t} \mathcal{P}(M, \Gamma, t) = -T(t) \Gamma \frac{\partial}{\partial M} [D(M) \mathcal{P}(M, \Gamma, t)] + \omega \frac{\partial}{\partial \Gamma} [\Gamma \mathcal{P}(M, \Gamma, t)] + \zeta^2 \frac{\partial^2}{\partial \Gamma^2} \mathcal{P}(M, \Gamma, t) ;
\]

this is as a continuity equation \( \partial \mathcal{P} + \nabla \cdot \mathcal{J} = 0 \) in terms
of the vectorial differential operator $\nabla = [\partial_M, \partial_t]$ and probability current $J = [J_M, J_t] = [T \nabla \mathcal{P}, -\omega \Gamma \mathcal{P} - \zeta \partial_t \mathcal{P}]$. In particular, we are interested in the marginalized $\mathcal{P}(M, t) = \int d\Gamma \mathcal{P}(M, \Gamma, t)$ with boundary conditions $\lim_{M \to \infty} \mathcal{P}(M, t) = 0$, $\mathcal{P}(M, 0) = \delta_D(M)$ and $\left(\int d\Gamma \mathcal{P}\right)_{M=0} = 0$; the latter expresses in the two-dimensional space $(M, \Gamma)$ the requirement of a zero current on the $M = 0$ line, i.e. $(\int d\Gamma J_M)|_{M=0} = 0$. In Appendix B we show that the Fokker-Planck equation is solved by

$$\mathcal{P}(M, t) = \frac{1}{D \sqrt{\pi Y}} e^{-X^2/4Y}, \quad (28)$$

in terms of:

$$\begin{align*}
X(M) &= \int \frac{dM}{D(M)} = \frac{1}{\sigma(M)} \\
Y(t) &= \frac{\zeta^2}{\omega} \int dt \mathcal{T}(\tau) e^{-\omega \tau} \int dr \mathcal{T}(r') [e^{\omega r'} - e^{-\omega r'}],
\end{align*}$$

This is the equivalent for colored noise of Eq. (29), which is recovered in the limit $\zeta = \omega \to \infty$ as $Y(t) \to \int dt \mathcal{T}^2(t) = 1/2 \delta^2$. The corresponding expression for the mass function is written as:

$$N(M, t) = \frac{1}{\sqrt{\pi Y(t)}} \frac{\bar{\rho}_M}{M \sigma^2} \frac{dM}{dM} e^{-1/4 \sigma^2 Y(t)}; \quad (30)$$

this is similar to the Press & Schechter shape, but for a modified redshift evolution encoded in $Y(t)$. The quantity $(2Y)^{-1/2}$, that in the white-noise limit is just $\delta_c$, can be regarded as a modified collapse threshold; in Fig. 3 we show how its evolution and absolute value differ from $\delta_c(t)$, depending on the volatility $\zeta$ and mean-reversal rate $\omega$ characterizing the noise. We conclude that only values of $\zeta \sim \omega \gtrsim$ several Gyr$^{-1}$ are required not to move far away from the Press & Schechter mass function, and hence from simulations.

5. SUMMARY AND OUTLOOK

In this paper we have submitted a new theory of the hierarchical clustering based on stochastic differential equations in real space, that constitutes a change of perspective with respect to the excursion set formalism; this work is specifically focused on the halo mass function.

First, we have invented a stochastic differential equation that describes fluctuations in the mass growth of DM halos, as driven by a multiplicative white (Gaussian) noise dependent on the spherical collapse threshold and on the power spectrum of DM perturbations. By numerically integrating such a stochastic differential equation, in Sect. 2 we have demonstrated that the noise yields an average drift of the halo population toward larger masses, that quantitatively renders the standard hierarchical clustering (see Fig. 1). Then, in Sect. 2.1.1 we have solved the Fokker-Planck equation associated to the stochastic dynamics, and obtained as a solution the Press & Schechter mass function; in Sect. 2.2 we have pointed out that the solution is stationary when the original equation is written in convenient variables.

Then in Sect. 3 we have introduced a minimal modification of the stochastic equation in terms of a mass-dependent collapse threshold. Using a parametric shape analogous to that adopted in the excursion set framework, we have obtained a closed-form analytical solution of the associated Fokker-Planck equation. Remarkably, such a solution has a limiting shape for large masses/early times similar to the empirical fitting formula introduced since Sheth & Tormen (1999); in fact, for specific values of the parameters describing the mass dependence of the collapse threshold, our result reproduces extremely well the $N$-body mass function over a wide range of masses and redshifts (see Fig. 2).

As an aside issue, in Sect. 4 we have generalized our stochastic approach to a colored, instead of a white, noise; in particular, we have investigated the modification to the Press & Schechter mass function when the stochastic dynamics is ruled by a multiplicative Ornstein-Uhlenbeck noise with finite volatility and mean-reversal rate. We have exactly solved the related Fokker-Planck equation, finding that the mass function has shape analogous to the Press & Schechter one when expressed in terms of a modified, effective collapse threshold; the latter may substantially differ from the standard $\delta_c(t)$ in absolute value and time evolution, depending on the correlation parameters of the noise (see Fig. 3). We conclude that values of such parameters larger than several Gyr$^{-1}$ are required not to move far away from the Press & Schechter mass function, and hence from simulations.

The next-order development of this work will concern the computation of the conditional mass function, i.e., the mass function of a halo’s progenitors. This investigation will naturally extend to merger rates, formation time distributions, and large-scale halo bias. A more detailed comparison of our results with the outcomes of $N$-body simulations, that includes the specificity of both the numerical experiments as well as of the theory, will be welcome. Other future applications could involve a re-examination of the two-phase mass growth of DM halos, the halo specific angular momentum distribution, the void mass function, and halo statistics in non-standard cosmological frameworks. We very much hope that the new perspective offered by the theory presented here will contribute to a better understanding of the gravitational dynamics leading to the formation and evolution of DM halos and hosted baryonic structures across cosmic times.

We thank our referee for a constructive report, and for the insightful comments and helpful suggestions. We acknowledge Carlo Baccigalupi, Alessandro Bressan, and Giovanni Bussi for enlightening discussions and critical reading. This work has been partially supported by PRIN MIUR 2017 prot. 20173ML3WW 002, ‘Opening the ALMA window on the cosmic evolution of gas, stars and supermassive black holes’. A.L. has taken advantage of the MIUR grant ‘Finanziamento annuale individuale attività base di ricerca’ and of the EU H2020-MSCA-ITN-2019 Project 860744 ‘BiD4BEST: Big Data applications for black hole Evolution STudies’. 
In terms of the conditional (transition) probability $P(x, t + \tau|\tau, x')$ between the times $t$ and $t + \tau$; in other words, for a Markovian system the transition probability depends only on the value at the next earlier time. We rewrite the integrand as $P(x, t + \tau|\tau, x') P(x', t) = P(x + \Delta, t + \tau|x - \Delta, t) P(x - \Delta, t)$ in terms of $\Delta \equiv x - x'$ and then perform a Taylor expansion in $\Delta$ to obtain

$$P(x, t + \tau|\tau, x') P(x', t) \approx \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \Delta^n \partial_x^n [P(x + \Delta, t + \tau|x, t) P(x, t)].$$

(A2)

Now we insert this expression in Eq. (A1) and perform the integration after changing variable from $x'$ to $\Delta$: noting that in the $n = 0$ term $\int d\Delta P(x + \Delta, t + \tau|x, t) = 1$ holds since the conditional probability is normalized, we get

$$P(x, t + \tau) - P(x, t) \approx \sum_{n=1}^{\infty} \frac{1}{n!} \partial_x^n M_n(x; t; \tau),$$

(A3)

where we have defined the conditional moments

$$M_n(x; t; \tau) \equiv \langle |\xi(t + \tau) - \xi(t)|^n \rangle_{\xi(t) = x} = \int d\Delta \Delta^n P(x + \Delta, t + \tau|x, t).$$

(A4)

Now we Taylor-expand the moments with respect to $\tau$ as follows

$$M_n(x; t; \tau)/n! \approx D_n(x, t) \tau + \vartheta(\tau^2);$$

(A5)

note that terms of order $\tau^0$ cannot be present since $P(x + \Delta, t|x, t) = \delta_D(\Delta)$ by definition and in Eq. (A4) all the conditional moments for $n \geq 1$ vanish. For future reference the coefficients $D_n$ are defined as

$$D_n(x, t) \equiv \lim_{\tau \to 0} \frac{1}{\tau} \frac{M_n(x; t; \tau)}{n!} = \lim_{\tau \to 0} \frac{1}{\tau} \langle |\xi(t + \tau) - \xi(t)|^n \rangle_{\xi(t) = x}.$$

(A6)

All in all, we obtain the so called Kramers-Moyal expansion in terms of the partial differential equation

$$\partial_t P(x, t) \approx \sum_{n=1}^{\infty} (-\partial_x)^n D_n(x, t);$$

(A7)

this ends the first step in the derivation.

We now compute explicitly the coefficients $D_n$ when the variable $\xi(t)$ satisfies a stochastic differential equation

$$\dot{\xi} = h(\xi, t) + g(\xi, t) \eta(t),$$

(A8)

with initial condition $\xi(t) = x$. Here $\eta(t)$ is a white (Gaussian) noise with ensemble-average properties $\langle \eta(t) \rangle = 0$ and $\langle \eta(t) \eta(t') \rangle = 2 \delta_D(t - t')$; the coefficient 2 in this last expression is only a convenient arbitrary choice, since it can be reabsorbed into the multiplicative function $g$ without loss of generality. We start by transforming the differential into an integral stochastic equation

$$\xi(t + \tau) - x = \int_t^{t+\tau} dt' [h(\xi(t'), t') + g(\xi(t'), t') \eta(t')].$$

(A9)

We expand near $x$ the functions $h(\xi(t'), t') \approx h(x, t') + \partial_x h(x, t') (\xi(t') - x)$ and $g(\xi(t'), t') \approx g(x, t') +$
Then we iterate for \(\xi(t') - x\) in the integrand to get

\[
\xi(t + \tau) - x \approx \int_t^{t+\tau} dt' h(x, t') + \int_t^{t+\tau} dt' \partial_x h(x, t') (\xi(t') - x) + \ldots
\]  

(A10)

Now taking the ensemble average and using the properties of the white noise yields

\[
\langle \xi(t + \tau) - x \rangle = \int_t^{t+\tau} dt' h(x, t') + \int_t^{t+\tau} dt' \int_t^{t'} dt'' h(x, t'') \partial_x h(x, t') + \ldots
\]  

(A11)

where in the last term we have used that \(\int_t^{t'} dt'' 2 \delta_{D}(t'' - t') g(x, t'') = g(x, t')\) since the Dirac-\(\delta\) operates on an extremal of the integration. Dividing by \(\tau\) and taking the limit for \(\tau \to 0\) one immediately recognizes the coefficient \(D_1 = h(x, t) + g(x, t) \partial_x g(x, t)\). For higher-order coefficients notice that terms containing the noise are proportional to \(\tau^n\) where \(n\) is the number of integrals involved, and vanish for small \(\tau\); actually only one of such terms, containing two integrals and two noises contributes and yields \(D_2 = \lim_{\tau \to 0} (1/2\tau) \int_t^{t+\tau} dt' \int_t^{t'} dt'' g(x, t') g(x, t'') 2 \delta_{D}(t' - t'') = g^2(x, t)\), while \(D_n = 0\) for any \(n \geq 3\). This ends the second step of the derivation.

Putting together the coefficients just derived in the Kramers-Moyal expansion of Eq. (A7), one finds the Fokker-Planck equation corresponding to the original stochastic equation:

\[
\partial_t \mathcal{P}(x, t) = -\partial_x [D_1(x, t) \mathcal{P}(x, t)] + \partial_x^2 [D_2(x, t) \mathcal{P}(x, t)]
\]

(A12)

The quantity \(g \partial_x g\) appearing in the coefficient \(D_1\) is a noise-induced drift; this stems from the fact that as \(\eta(t)\) fluctuates, also the random variable \(\xi(t)\) and so the function \(g(\xi(t), t)\) varies and therefore \((g(\xi(t), t) \eta(t))\) is not null even if \(\langle \eta(t) \rangle = 0\) is. Finally, simple algebra shows that the Fokker-Planck equation may be written as a source-free continuity equation:

\[
\partial_t \mathcal{P}(x, t) + \partial_x \mathcal{J}(x, t) = 0
\]

(A13)

in terms of a probability current

\[
\mathcal{J}(x, t) \equiv h(x, t) \mathcal{P}(x, t) - g(x, t) \partial_x [g(x, t) \mathcal{P}(x, t)]
\]

(A14)

\[\mathcal{J}(x, t) \equiv h(x, t) \mathcal{P}(x, t) - g(x, t) \partial_x [g(x, t) \mathcal{P}(x, t)]\]  

(A15)

B. SOLUTION OF THE FOKKER-PLANCK EQUATION FOR COLORED NOISE

In this Appendix we show how to solve the Fokker-Planck equation derived in Sect. 4:

\[
\partial_t \mathcal{P}(M, \Gamma, t) = -T(t) \Gamma \partial_M [D \mathcal{P}(M, \Gamma, t)] + \omega \partial_T [\Gamma \mathcal{P}(M, \Gamma, t)] + \zeta^2 \partial_M^2 \mathcal{P}(M, \Gamma, t).
\]

(B1)
for the marginalized $\mathcal{P}(M, t) = \int d\Gamma \mathcal{P}(M, \Gamma, t)$ with boundary conditions $\lim_{M \to \infty} \mathcal{P} = 0$, $\mathcal{P}(M, 0) = \delta_D(M)$ and $(\int d\Gamma \mathcal{P})|_{M=0} = 0$.

As a first step, we introduce a new variable $X \equiv \int dM/D(M)$ in place of $M$, and redefine the probability density as $\mathcal{W}(X, \Gamma, t) = D(M) \mathcal{P}(M, \Gamma, t)$; then the above equation turns into

$$\partial_t \mathcal{W} = -T(t) \Gamma \partial_X \mathcal{W} + \omega \partial_T [\mathcal{W} \Gamma] + \zeta^2 \partial_T^2 \mathcal{W}. \quad (B2)$$

We now perform a two-dimensional Fourier transform

$$\mathcal{W}(X, \Gamma, t) \propto \int dk_X dk_\Gamma \tilde{\mathcal{W}}(k_X, k_\Gamma, t) e^{i(k_X X + k_\Gamma \Gamma)}, \quad (B3)$$

and obtain the following equation for the Fourier modes

$$\partial_t \tilde{\mathcal{W}} = T(t) k_X \partial_{k_X} \tilde{\mathcal{W}} - \omega k_\Gamma \partial_{k_\Gamma} \tilde{\mathcal{W}} - \zeta^2 k_\Gamma^2 \tilde{\mathcal{W}}. \quad (B4)$$

Given the boundary conditions, it is convenient to look for solutions with shape

$$\tilde{\mathcal{W}}(k_X, k_\Gamma, t) \propto e^{-k_X^2 \Sigma_{XX}/2 - k_\Gamma^2 \Sigma_{\Gamma\Gamma}/2 - k_X k_\Gamma \Sigma_{X\Gamma}} \quad (B5)$$

where $\Sigma_{XX}(t)$, $\Sigma_{X\Gamma}(t)$, $\Sigma_{\Gamma\Gamma}(t)$ are only functions of time. Inserting this ansantz into the previous equation yields the following ordinary differential equations

$$\begin{align*}
\dot{\Sigma}_{\Gamma\Gamma} &= -2\omega \Sigma_{\Gamma\Gamma} + 2\zeta^2 \\
\dot{\Sigma}_{X\Gamma} &= T \Sigma_{\Gamma\Gamma} - \omega \Sigma_{X\Gamma} \\
\dot{\Sigma}_{XX} &= 2T \Sigma_{X\Gamma}.
\end{align*} \quad (B6)$$

These can be straightforwardly solved as

$$\begin{align*}
\Sigma_{\Gamma\Gamma}(t) &= \frac{\zeta^2}{\omega} (1 - e^{-2\omega t}) \\
\Sigma_{X\Gamma}(t) &= \int_0^t d\tau T(\tau) e^{-\omega(t-\tau)} \Sigma_{\Gamma\Gamma}(\tau) \\
\Sigma_{XX}(t) &= 2 \int_0^t d\tau T(\tau) \Sigma_{X\Gamma}(\tau).
\end{align*} \quad (B7)$$

Inverting the Fourier transform in Eq. (B3), one finds out the solution

$$\mathcal{W}(X, \Gamma, t) \propto \frac{1}{\sqrt{||\Sigma||}} \exp \left\{ -\frac{X^2}{2 \Sigma_{XX}} - \frac{[\Sigma_{XX} \Gamma - \Sigma_{X\Gamma} X]^2}{2 \Sigma_{XX} ||\Sigma||} \right\}, \quad (B8)$$

where $||\Sigma|| = \Sigma_{XX} \Sigma_{\Gamma\Gamma} - \Sigma_{X\Gamma}^2$ is the determinant of the $2 \times 2$ symmetric matrix constructed with the $\Sigma$s; this can be easily checked to satisfy the desired boundary conditions, that in terms of the variable $X$ and function $\mathcal{W}$ read $\lim_{X \to \infty} \mathcal{W} = 0$, $\mathcal{W}(X, 0) = \delta_D(X)$, and $(\int d\Gamma \mathcal{W})|_{X=0} = 0$.

Finally, marginalizing over $\Gamma$ and coming back to the original variables, one obtains

$$\mathcal{P}(M, t) = \frac{1}{D \sqrt{\pi} Y} e^{-X^2/4Y}, \quad (B9)$$

in terms of the quantities:

$$\begin{align*}
X(M) &= \int \frac{dM}{D(M)} \\
Y(t) &= \frac{\Sigma_{XX}}{2} = \frac{\zeta^2}{\omega} \int_0^t \! d\tau T(\tau) e^{-\omega \tau} \int_\tau^T \! d\tau' T(\tau') [e^{\omega \tau'} - e^{-\omega \tau'}]
\end{align*} \quad (B10)$$

note that the correct white-noise limit is recovered for $\zeta = \omega \to \infty$ since $Y(t) \to \int dt T^2(t) = 1/2 \delta_D^2(t)$, and that actually $(2Y)^{-1/2}$ constitute an effective collapse threshold, dependent on the parameters of the colored noise.

REFERENCES

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15

Benson, A. J., Kamionkowski, M., & Hassani, S. H. 2005, MNRAS, 357, 847
