Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **411,** 2644–2652 (2011) doi:10.1111/j.1365-2966.2010.17871.x

The bias and mass function of dark matter haloes in non-Markovian extension of the excursion set theory

Chung-Pei Ma,[1★] Michele Maggiore,[2] Antonio Riotto[3,4] and Jun Zhang[5]

[1]*Department of Astronomy, University of California, Berkeley, CA 94720, USA*
[2]*Département de Physique Théorique, Université de Genève, CH-1211 Geneva, Switzerland*
[3]*CERN, PH-TH Division, CH-1211, Genève 23, Switzerland*
[4]*INFN, Sezione di Padova, Via Marzolo 8, I-35131 Padua, Italy*
[5]*Texas Cosmology Center, University of Texas, Austin, TX 78712, USA*

Accepted 2010 October 14. Received 2010 October 13; in original form 2010 July 23

ABSTRACT

The excursion set theory based on spherical or ellipsoidal gravitational collapse provides an elegant analytic framework for calculating the mass function and the large-scale bias of dark matter haloes. This theory assumes that the perturbed density field evolves stochastically with the smoothing scale and exhibits Markovian random walks in the presence of a density barrier. Here, we derive an analytic expression for the halo bias in a new theoretical model that incorporates non-Markovian extension of the excursion set theory with a stochastic barrier. This model allows us to handle non-Markovian random walks and to calculate perturbatively these corrections to the standard Markovian predictions for the halo mass function and halo bias. Our model contains only two parameters: κ, which parametrizes the degree of non-Markovianity and whose exact value depends on the shape of the filter function used to smooth the density field, and a, which parametrizes the degree of stochasticity of the barrier. Appropriate choices of κ and a in our new model can lead to a closer match to both the halo mass function and the halo bias in the latest N-body simulations than the standard excursion set theory.

Key words: cosmology: theory.

1 INTRODUCTION

Dark matter haloes typically form at sites of high density peaks. The spatial distribution of dark matter haloes is therefore a biased tracer of the underlying mass distribution. A standard way to quantify this difference between haloes and mass is to use a halo bias parameter $b_{\rm h}$, which can be defined as the ratio of the overdensity of haloes to mass or as the square root of the ratio of the two-point correlation function (or the power spectrum) of haloes to mass.

Like the halo mass function, analytic expressions for the halo bias can be obtained from the excursion set theory (Bond et al. 1991) based on the spherical gravitational collapse model (Cole & Kaiser 1989; Mo & White 1996). In the excursion set theory, the density perturbation evolves stochastically with the smoothing scale, and the problem of computing the probability of halo formation is mapped into the first-passage time problem in the presence of a (constant) barrier. The approach to the clustering evolution is based on a generalization of the peak-background split scheme (Bardeen et al. 1986), which basically consists in splitting the mass perturbations into a fine-grained (peak) component filtered on a scale R and a coarse-grained (background) component filtered on a scale $R_0 \gg R$. The underlying idea is to ascribe the collapse of objects on small scales to the high-frequency modes of the density fields, while the action of large-scale structures of these non-linear condensations is due to a shift of the local background density.

Comparison with N-body simulations finds that the spherical collapse model underpredicts the halo bias for low-mass haloes (Jing 1998; Sheth & Tormen 1999). The discrepancy reaches a factor of $\sim$2 at $M \sim 0.01 M_*$, where M_* is the characteristic non-linear mass scale [defined by $\sigma(M_*) = 1$, where $\sigma^2(M)$ is the variance of the density field in a volume of radius R containing the mass M]. Sheth, Mo & Tormen (2001) obtained an improved formula for the halo bias by using a moving barrier whose scale-dependent shape is motivated by the ellipsoidal gravitational collapse model. Compared to the spherical collapse model, this formula predicts a lower bias at the high-mass end and a higher bias at the low-mass end (see Fig. 1). The resulting bias is shown to be too high at the low-mass end by $\sim$20 per cent compared with simulation results. Further modifications have been introduced that either used the functional form of Sheth & Tormen (1999) or Sheth et al. (2001) with new fitting parameters (e.g. Tinker et al. 2005), or proposed

★E-mail: cpma@berkeley.edu

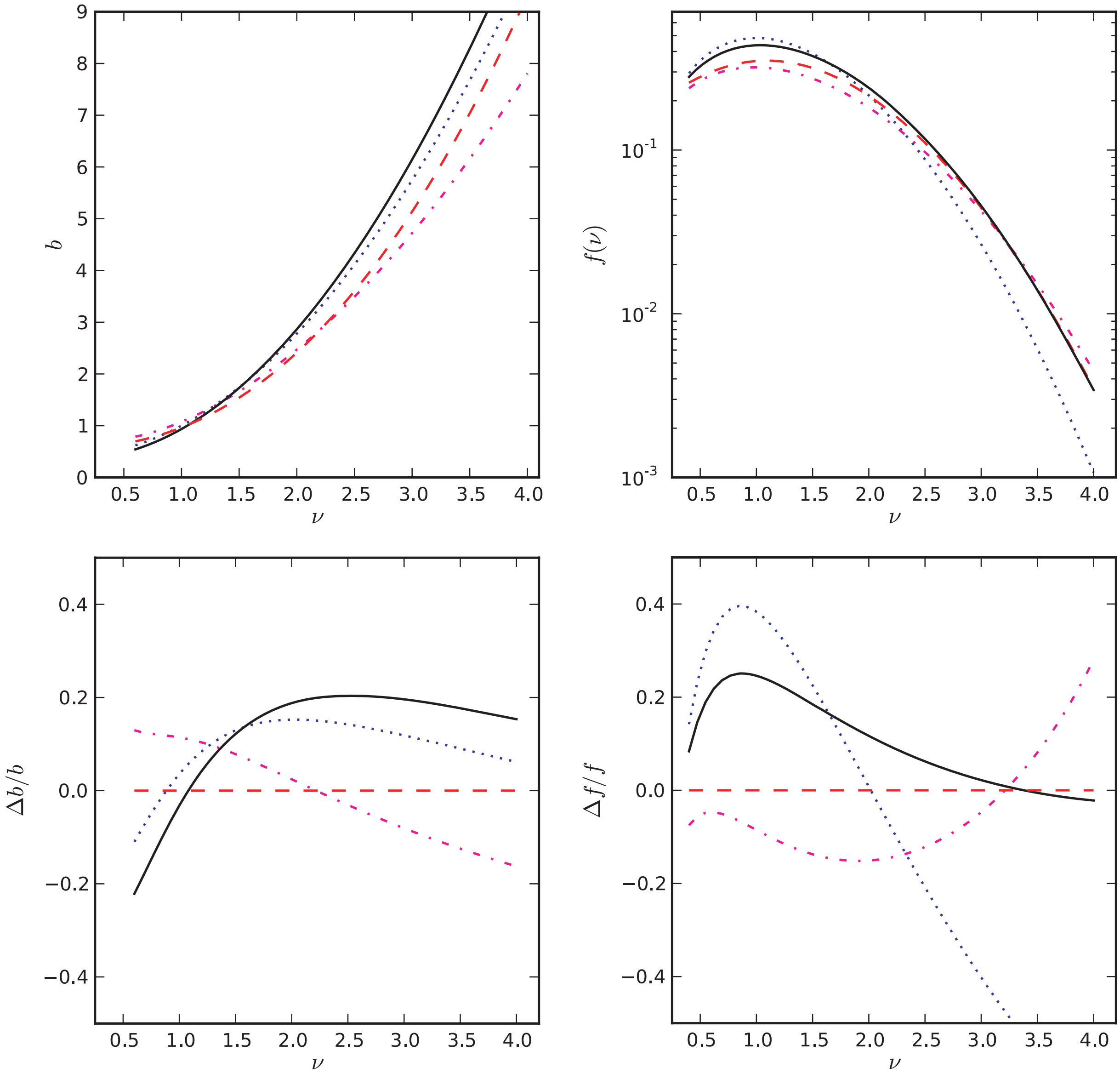

Figure 1. Comparison of the Eulerian halo bias $b_h(\nu)$ (left-hand panels) and the halo mass function $f(\nu)$ (right-hand panels) from various analytic models and simulations: our non-Markovian and stochastic barrier models from equation (27), with $a = 0.818$, $\kappa = 0.23$ (black solid line); the standard Markovian spherical collapse (blue dotted) and ellipsoidal collapse (magenta dot–dashed) models; and the fits to N-body simulation results (red dashed) from Tinker et al. (2008, 2010). The bottom panels show the fractional difference between each of the analytic model prediction and the fit to N-body result. With only two free parameters, our new model is able to match the N-body results to within ∼20 per cent for a wide range of ν.

new fitting forms altogether (e.g. Seljak & Warren 2004; Pillepich, Porciani & Hahn 2010; Tinker et al. 2010).

Our goal in this paper is not to improve on the accuracy of the fits to the halo bias, but rather to gain deeper theoretical insight by deriving an analytical expression for the halo bias using a new model. This model modifies the excursion set theory by incorporating non-Markovian random walks in the presence of a stochastic barrier. It is based on a path integral formulation introduced in Maggiore & Riotto (2010a, hereafter MR1) and Maggiore & Riotto (2010b), which provides an analytic framework for calculating perturbatively the non-Markovian corrections to the standard version of the excursion set theory. Mathematically, the non-Markovianity in the theory is related to the choice of the filter function necessary to smooth out the density contrast. As soon as the filter function is different from a step (tophat) function in momentum space, the excursion of the smoothed density contrast is non-Markovian, namely every step depends on the previous ones and the random walk acquires memory. As the computation of the bias parameter b_h amounts to computing the first-crossing rate with a non-trivial initial condition at a large, but not infinite, radius, the non-Markovianity makes the calculation much harder than the Markovian case.

Furthermore, the critical value for collapse in our model is itself assumed to be a stochastic variable, whose scatter reflects a number of complicated aspects of the underlying dynamics. The gravitational collapse of haloes is a complex dynamical phenomenon, and modelling it as spherical, or even as ellipsoidal, is a significant oversimplification. In addition, the very definition of what is a dark matter halo, both in simulations and observationally, is a non-trivial problem. Maggiore & Riotto (2010b) proposed that some of the physical complications inherent to a realistic description of halo formation can be included in the excursion set theory framework, at least at an effective level, by taking into account the fact that the critical value for collapse is itself a stochastic variable.

In Section 2, we review briefly the derivation for the halo mass function in the Markovian excursion set theory (Section 2.1) and the path integral approach used to introduce non-Markovian terms (Section 2.2) and stochastic barriers (Section 2.3) into the theory. Section 3 is devoted to the discussion of halo bias, including a review of the standard derivation in the Markovian case (Section 3.1) and a summary of our new derivation in the non-Markovian model (Sections 3.2 and 3.3). The details of how the halo bias is calculated from the conditional probability for two barrier crossings in the new model are provided in Appendix A. In Section 4, we compare the predictions for the halo bias and mass function in our new model with those from the Markovian model (both spherical and ellipsoidal collapse) and N-body simulations.

2 NON-MARKOVIAN EXTENSION AND STOCHASTIC BARRIER

In this section, we review the main points of the excursion set theory (Bond et al. 1991) and then summarize how to introduce non-Markovian terms in the presence of a stochastic barrier (MR1; Maggiore & Riotto 2010b).

2.1 Brief review of the excursion set theory

The basic variable is the smoothed density contrast:

$$\delta(\boldsymbol{x}, R) = \int d^3x' \, W(|\boldsymbol{x} - \boldsymbol{x}'|, R)\, \delta(\boldsymbol{x}'), \quad (1)$$

where $\delta(\boldsymbol{x}) = \rho(\boldsymbol{x})/\bar{\rho} - 1$ is the density contrast about the mean mass density $\bar{\rho}$ of the universe, $W(|\boldsymbol{x} - \boldsymbol{x}'|, R)$ is the filter function and R is the smoothing scale. We are interested in the evolution of $\delta(\boldsymbol{x}, R)$ with smoothing scale R at a fixed point $\boldsymbol{x}$ in space, so we suppress the argument $\boldsymbol{x}$ from this point on. It is also convenient to use, instead of R, the variance S of the smoothed density field defined by

$$S(R) \equiv \sigma^2(R) = \int \frac{d^3k}{(2\pi)^3}\, P(k)\tilde{W}^2(k, R), \quad (2)$$

where $P(k)$ is the power spectrum of the matter density fluctuations in the cosmological model under consideration and $\tilde{W}$ is the Fourier transform of the filter function W. A smoothing radius $R = \infty$ corresponds to $S = 0$ and, in hierarchical models of structure formation such as the Λ cold dark matter model, S is a monotonically decreasing function of R. We can therefore use S and R interchangeably and denote our basic variable by $\delta(S)$.

If the filter function is taken to be a tophat in momentum space, $\delta(S)$ then satisfies a simple Langevin equation with S playing the role of a 'pseudo-time' (Bond et al. 1991):

$$\frac{\partial \delta(S)}{\partial S} = \eta(S), \quad (3)$$

where $\eta(S)$ represents a stochastic 'pseudo-force' whose two-point correlation statistic obeys a Dirac-delta function:

$$\langle \eta(S_1)\eta(S_2)\rangle = \delta_{\rm D}(S_1 - S_2). \quad (4)$$

It then follows that the function $\Pi(\delta_0; \delta; S)$, which gives the probability density of reaching a value δ at 'time' S starting at a value δ_0 at $S = 0$, satisfies the Fokker–Planck equation:

$$\frac{\partial \Pi}{\partial S} = \frac{1}{2}\frac{\partial^2 \Pi}{\partial \delta^2}. \quad (5)$$

In Bond et al. (1991), this equation was supplemented by the boundary condition

$$\Pi(\delta, S)|_{\delta=\delta_{\rm c}} = 0 \quad (6)$$

to eliminate the trajectories that have reached the critical value $\delta_{\rm c}$ for collapse.

The corresponding solution of the Fokker–Planck equation is

$$\Pi(\delta_0; \delta; S) = \frac{1}{\sqrt{2\pi S}}\left[e^{-(\delta-\delta_0)^2/(2S)} - e^{-(2\delta_{\rm c}-\delta_0-\delta)^2/(2S)}\right]. \quad (7)$$

The probability $\mathcal{F}(S)\,dS$ of first crossing the threshold density $\delta_{\rm c}$ between 'time' S and $S + dS$ is then given by

$$\mathcal{F}(S) = -\int_{-\infty}^{\delta_{\rm c}} d\delta\, \frac{\partial \Pi}{\partial S} = \frac{1}{\sqrt{2\pi}}\frac{\delta_{\rm c}}{S^{3/2}}e^{-\delta_{\rm c}^2/(2S)}, \quad (8)$$

where we have set $\delta_0 = 0$. The number density of virialized objects with mass between M and $M + dM$ is related to the first-crossing probability between S and $S + dS$ by

$$\begin{aligned}\frac{dn(M)}{dM}dM &= \frac{\bar{\rho}}{M}\mathcal{F}(S)\,dS \\ &= \sqrt{\frac{2}{\pi}}\,\frac{\delta_{\rm c}}{\sigma}\,e^{-\delta_{\rm c}^2/(2\sigma^2)}\,\frac{\bar{\rho}}{M^2}\frac{d\ln\sigma^{-1}}{d\ln M}dM \\ &\equiv f(\nu)\,\frac{\bar{\rho}}{M^2}\frac{d\ln\sigma^{-1}}{d\ln M}dM, \end{aligned} \quad (9)$$

where $\sigma = S^{1/2}$ is defined in equation (2) and $\nu \equiv \delta_{\rm c}/\sigma$. This expression reproduces the mass function of Press & Schechter (1974), including the correct overall normalization that had to be adjusted by hand in Press & Schechter (1974). We will use the dimensionless function $f(\nu)$ defined in equation (9) to denote the halo mass function below.

2.2 Non-Markovian extension

As already discussed in Bond et al. (1991), a difficulty of the excursion set approach in Section 2.1 is that an unambiguous relation between the smoothing radius R and the mass M of the corresponding collapsed halo only exists when the filter function is a tophat in coordinate space: $M(R) = (4/3)\pi R^3\rho$. For all other filter functions (e.g. tophat in momentum space, Gaussian), it is impossible to associate a well-defined mass $M(R)$ [see also the recent review by Zentner (2007)]. More importantly, $\delta(S)$ obeys a Langevin equation with a Dirac-delta noise as in equations (3) and (4) *only* when the filter function is a tophat in momentum space. Otherwise, the evolution of δ with the smoothing scale becomes non-Markovian, and the distribution function $\Pi(\delta_0; \delta; S)$ of the trajectories no longer obeys the Fokker–Planck equation nor any local generalization of it. In this case, $\Pi(\delta_0; \delta; S)$ obeys a complicated equation that is non-local with respect to the variable S (MR1).

To deal with this problem, MR1 proposed a 'microscopic' approach, in which one computes the probability associated with each trajectory $\delta(S)$ and sums over all relevant trajectories. As with any path integral formulation, it is convenient to discretize the time variable and to take the continuum limit at the end. Therefore, we discretize the interval $[0, S]$ in steps $\Delta S = \epsilon$, so $S_k = k\epsilon$ with $k = 1, \ldots n$, and $S_n \equiv S$, and a trajectory is defined by the collection of values $\{\delta_1, \ldots, \delta_n\}$, such that $\delta(S_k) = \delta_k$. All trajectories start at a value δ_0 at time $S = 0$.

The basic quantity in this approach is the probability density in the space of trajectories, defined as

$$W(\delta_0; \delta_1, \ldots, \delta_n; S_n) \equiv \langle \delta_{\rm D}(\delta(S_1) - \delta_1)\cdots\delta_{\rm D}(\delta(S_n) - \delta_n)\rangle, \quad (10)$$

where δ_D denotes the Dirac-delta function. In terms of W, we define

$$\Pi_\epsilon(\delta_0; \delta_n; S_n) \equiv \int_{-\infty}^{\delta_c} \mathrm{d}\delta_1 \cdots \mathrm{d}\delta_{n-1}\, W(\delta_0; \delta_1, \ldots, \delta_n; S_n), \tag{11}$$

where $S_n = n\epsilon$ and $\Pi_\epsilon(\delta_0; \delta; S)$ is the probability density of arriving at the 'position' δ in a 'time' S, starting from δ_0 at time $S_0 = 0$, through trajectories that never exceeded δ_c. The problem of computing the distribution function of excursion set theory is therefore mapped into the computation of a path integral with a boundary at $\delta = \delta_c$.

The probability density W can be computed in terms of the connected correlators of the theory. When the density field δ is a Gaussian random variable, only the two-point connected function is non-zero, and one finds

$$W(\delta_0; \delta_1, \ldots, \delta_n; S_n) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\lambda_1}{2\pi} \cdots \frac{\mathrm{d}\lambda_n}{2\pi}\, \mathrm{e}^{i\sum_{i=1}^{n}\lambda_i\delta_i - \frac{1}{2}\sum_{i,j=1}^{n}\lambda_i\lambda_j\langle\delta_i\delta_j\rangle_c}, \tag{12}$$

and $\delta_i \equiv \delta(S_i)$. We will restrict the discussion here to the Gaussian case since higher order connected correlators must be included in the non-Gaussian case (Maggiore & Riotto 2010c,d).

First, consider the case of a tophat filter in momentum space, so the evolution of $\delta(S)$ is Markovian and obeys equations (3) and (4). Then one can show that the connected two-point correlator is given by

$$\langle\delta(S_i)\delta(S_j)\rangle_c = \min(S_i, S_j), \tag{13}$$

and the integrals over $\mathrm{d}\lambda_1, \ldots, \mathrm{d}\lambda_n$ in equation (12) can be performed explicitly to give

$$W^{\mathrm{gm}}(\delta_0; \delta_1, \ldots, \delta_n; S_n) = \frac{1}{(2\pi\epsilon)^{n/2}}\, \mathrm{e}^{-\frac{1}{2\epsilon}\sum_{i=0}^{n-1}(\delta_{i+1}-\delta_i)^2} \tag{14}$$

where the superscript 'gm' stands for 'Gaussian and Markovian'. Inserting this expression into equation (11) it can be shown (MR1) that, in the continuum limit, the corresponding distribution function $\Pi^{\mathrm{gm}}_{\epsilon=0}(\delta_0; \delta; S)$ satisfies the Fokker–Planck equation (5) as well as the boundary condition (6), and therefore we recover the standard result (7) of excursion set theory.

The interesting case is to generalize the above computation to filter functions different from the conventional tophat in momentum space. The two-point correlator depends on the filter function. For a Gaussian filter and a tophat filter in coordinate space, for instance, we find

$$\langle\delta(S_i)\delta(S_j)\rangle = \min(S_i, S_j) + \Delta(S_i, S_j), \tag{15}$$

where $\Delta(S_i, S_j) = \Delta(S_j, S_i)$ and, for $S_i \leq S_j$, the function $\Delta(S_i, S_j)$ is well approximated by

$$\Delta(S_i, S_j) \simeq \kappa\, \frac{S_i(S_j - S_i)}{S_j}, \tag{16}$$

with $\kappa \approx 0.35$ for a Gaussian filter and $\kappa \approx 0.44$ for a tophat filter in coordinate space. The parameter κ gives a measure of the non-Markovianity of the stochastic process, and the computation of the distribution function $\Pi(\delta_0; \delta; S)$ can be performed order by order in κ. The technique necessary for evaluating the path integral in equation (11) to first order in κ has been developed in MR1 and will be further discussed below. To first order in the non-Markovian corrections, the resulting first-crossing rate becomes

$$\mathcal{F}(S) = \frac{1-\kappa}{\sqrt{2\pi}}\, \frac{\delta_c}{S^{3/2}}\mathrm{e}^{-\delta_c^2/(2S)} + \frac{\kappa}{2\sqrt{2\pi}}\, \frac{\delta_c}{S^{3/2}}\Gamma\left(0, \frac{\delta_c^2}{2S}\right), \tag{17}$$

where $\Gamma(0, z)$ is the incomplete Gamma function. For $\kappa = 0$, one recovers the Markovian result in equation (8). The halo mass function is then obtained by substituting this expression for $\mathcal{F}$ into equation (9).

2.3 Stochastic barrier

The constant barrier $\delta_c \simeq 1.686$ in the spherical collapse model is a significant oversimplification of the complex dynamics leading to halo formation and growth. Such a model can be improved in various ways. For instance, the excursion set theory results for the mass function have been shown to match more closely those from N-body simulations by considering a moving barrier whose shape is motivated by the ellipsoidal collapse model (Sheth & Tormen 1999, 2002; Sheth et al. 2001; De Simone et al. 2010). The equations are summarized in Table 1. The parameters a, b, c are fixed by fit to N-body simulations, while A is fixed by the normalization condition on the halo mass function. As already remarked in Sheth et al. (2001), these expressions can be obtained from a barrier shape that is virtually identical to the ellipsoidal collapse barrier, except for the factor of a, which is not a consequence of the ellipsoidal collapse model. In fact, the ellipsoidal collapse model reduces to the spherical collapse model in the large mass limit. In this limit the mass function is determined by the slope of the exponential factor, so even in an ellipsoidal collapse model we must have $a = 1$, as in the spherical model. However, numerical simulations show that $a < 1$ and its precise value also depends on the details of the algorithm used for identifying haloes in the simulation, e.g. the link length in a friends-of-friends halo finder or the critical overdensity in a spherical density finder.

A physical understanding of the parameter a is given by a second independent improvement of the spherical collapse model, the diffusing barrier model proposed in Maggiore & Riotto (2010b). These authors suggested that at least some of the physical complications inherent to a realistic description of halo formation, which involves a mixture of smooth accretion, violent encounters and fragmentations, can be included in the excursion set theory framework by assuming that the critical value for collapse is itself a stochastic variable, whose scatter reflects a number of complicated aspects of the underlying dynamics. In the simple example of a barrier performing a random walk with diffusion coefficient D_B around the spherical collapse barrier, one indeed finds a mass function in which δ_c is effectively replaced by $a^{1/2}\delta_c$, with $a = 1/(1 + D_B)$, while at the same time κ is replaced by $a\kappa$ (see Table 1).

3 HALO BIAS

We now apply the technique in Section 2 to the computation of the halo bias, including the non-Markovian corrections with stochastic barriers. We sketch here the main steps of the computations, leaving the details to Appendix A.

3.1 Conditional probability: the Markovian case

To compute the bias, we need the probability of forming a halo of mass M, corresponding to a smoothing radius R, under the condition that the smoothed density contrast on a much larger scale R_m has a specified value $\delta_m = \delta(R_m)$. We use $\mathcal{F}(S_n|\delta_m, S_m)$ to denote the conditional first-crossing rate. This is the rate at which trajectories first cross the barrier at $\delta = \delta_c$ at time S_n, under the condition that they passed through the point $\delta = \delta_m$ at an earlier time S_m. We also use the notation $\mathcal{F}(S_n|0) \equiv \mathcal{F}(S_n|\delta_m = 0, S_m = 0)$, so $\mathcal{F}(S_n|0)$ is

Table 1. Summary of mass function and halo bias predicted by various analytic models.

Model	Mass function $f(\nu)$	Halo bias $b_{\rm h}(\nu)$	Parameters
Spherical collapse	$\sqrt{\frac{2}{\pi}}\,\nu\exp\left(-\frac{\nu^2}{2}\right)$	$1+\frac{\nu^2-1}{\delta_{\rm c}}$	$\delta_{\rm c}=1.686$
Ellipsoidal collapse	$A\sqrt{\frac{2}{\pi}}\sqrt{a}\nu\exp\left(-\frac{a\nu^2}{2}\right)\left[1+(a\nu^2)^q\right]$	$1+\frac{1}{\sqrt{a}\delta_{\rm c}}\left[\sqrt{a}(a\nu^2)+\sqrt{a}b(a\nu^2)^{1-c}\right.$	$A=0.322$, $q=-0.3$
		$\left.-\frac{(a\nu^2)^c}{(a\nu^2)^c+b(1-c)(1-c/2)}\right]$	$a=0.707$, $b=0.5$, $c=0.6$
Non-Markovian	$\sqrt{\frac{2}{\pi}}\left[(1-\kappa)\nu\exp\left(-\frac{\nu^2}{2}\right)+\kappa\frac{\nu}{2}\Gamma\left(0,\frac{\nu^2}{2}\right)\right]$	$1+\frac{1}{\delta_{\rm c}\left[1-\kappa+\frac{\kappa}{2}e^{\nu^2/2}\Gamma(0,\nu^2/2)\right]}\left\{\nu^2-1\right.$	$\kappa=0$ for tophat-k filter
		$\left.+\frac{\kappa}{2}\left[2-\exp\left(\frac{\nu^2}{2}\right)\Gamma\left(0,\frac{\nu^2}{2}\right)\right]\right\}$	$\kappa=0.35$ for Gaussian
			$\kappa=0.44$ for tophat-x
Non-Markovian	$\kappa\to a\kappa,\quad \nu\to\sqrt{a}\nu$	$\kappa\to a\kappa,\quad \nu\to\sqrt{a}\nu,\quad \delta_{\rm c}\to\sqrt{a}\delta_{\rm c}$	$a=\frac{1}{1+D_{\rm B}}$
+ Stochastic barrier			$D_{\rm B}=$ diffusion coefficient

the first-crossing rate when the density approaches the cosmic mean value on very large scales.

The halo overdensity in Lagrangian space is given by (Kaiser 1984; Efstathiou et al. 1988; Cole & Kaiser 1989; Mo & White 1996; see also Zentner 2007 for a review)

$$1+\delta^{\rm L}_{\rm halo}=\frac{\mathcal{F}(S_n|\delta_m,S_m)}{\mathcal{F}(S_n|0)}\,. \tag{18}$$

In a sufficiently large region, we have $S_m \ll S_n \equiv S$ and $\delta_m \ll \delta_{\rm c}$. Then, using the first-crossing rate of excursion set theory and retaining only the term linear in δ_m, we obtain

$$\delta^{\rm L}_{\rm halo}=\frac{\nu^2-1}{\delta_{\rm c}}\,\delta_m\,, \tag{19}$$

where $\nu=\delta_{\rm c}/\sigma$. After mapping to Eulerian space, one finds $\delta_{\rm halo}\approx 1+\delta^{\rm L}_{\rm halo}$ in the limit of small overdensity $\delta_m\simeq\delta$ (Mo & White 1996), and

$$b_{\rm h}(\nu)=1+\frac{\nu^2-1}{\delta_{\rm c}}\,. \tag{20}$$

3.2 Non-Markovian corrections

We now use the path integral formalism discussed in Section 2.2 to compute the non-Markovian corrections to the halo bias. The relevant quantity for our purposes is the conditional probability

$$P(\delta_n,S_n|\delta_m,S_m)\equiv \\ =\frac{\int_{-\infty}^{\delta_c}{\rm d}\delta_1\cdots\widehat{{\rm d}\delta_m}\cdots{\rm d}\delta_{n-1}W(\delta_0=0;\delta_1,\ldots,\delta_n;S_n)}{\int_{-\infty}^{\delta_c}{\rm d}\delta_1\cdots{\rm d}\delta_{m-1}W(\delta_0=0;\delta_1,\ldots,\delta_m;S_m)}\,, \tag{21}$$

where the hat over ${\rm d}\delta_m$ means that ${\rm d}\delta_m$ must be omitted from the list of integration variables. The numerator is a sum over all trajectories that start from $\delta_0=0$ at $S=0$, have a given fixed value δ_m at S_m, and a value δ_n at S_n, while all other points of the trajectory, $\delta_1,\ldots,\delta_{m-1},\delta_{m+1},\ldots,\delta_{n-1}$ are integrated from $-\infty$ to $\delta_{\rm c}$. The denominator gives the appropriate normalization to the conditional probability.

Similarly to equation (8), the conditional first-crossing rate $\mathcal{F}(S_n|\delta_m,S_m)$ is obtained from the conditional probability $P(\delta_n, S_n|\delta_m,S_m)$ using

$$\mathcal{F}(S_n|\delta_m,S_m)=-\int_{-\infty}^{\delta_c}{\rm d}\delta_n\,\frac{\partial P(\delta_n,S_n|\delta_m,S_m)}{\partial S_n}\,. \tag{22}$$

In the Gaussian and Markovian cases, the probability $W^{\rm gm}$ satisfies

$$W^{\rm gm}(\delta_0;\delta_1,\ldots,\delta_n;S_n)=W^{\rm gm}(\delta_0;\delta_1,\ldots,\delta_m;S_m)\\ \times W^{\rm gm}(\delta_m;\delta_{m+1},\ldots,\delta_n;S_n-S_m)\,, \tag{23}$$

and $P(\delta_n,S_n|\delta_m,S_m)$ in equation (21) becomes identical to the probability of arriving in δ_n at time S_n, starting from δ_m at time S_m, which is given by equation (7) (with δ_m identified with δ_0 at $S=S_n-S_m$), and we therefore recover the excursion set theory result.

We have computed the non-Markovian corrections to this result for the case of Gaussian fluctuations and a tophat filter in coordinate space, i.e. with the two-point function given in equations (15) and (16). The computation is quite involved, and we leave the details to Appendix A. Taking finally $S_m=0$ and developing to first order in $\delta_m\equiv\delta_0$, which is the case relevant to the computation of the bias, for the conditional first-crossing rate we find

$$\mathcal{F}(S|\delta_0,S_0=0)=\frac{\delta_{\rm c}}{\sqrt{2\pi}\,S^{3/2}}{\rm e}^{-\delta_c^2/(2S)}\\ \times\left\{\left(1-\kappa+\frac{\kappa}{2}{\rm e}^{\nu^2/2}\Gamma(0,\nu^2/2)\right)\right.\\ \left.+\frac{\delta_0}{\delta_{\rm c}}\left[(\nu^2-1)+\frac{\kappa}{2}\left(2-{\rm e}^{\nu^2/2}\Gamma(0,\nu^2/2)\right)\right]\right\}. \tag{24}$$

From this, we obtain the Lagrangian halo bias $b^{\rm L}_{\rm h}$ and the Eulerian halo bias $b_{\rm h}$:

$$b_{\rm h}(\nu)=1+b^{\rm L}_{\rm h}\\ =1+\frac{1}{\delta_{\rm c}}\,\frac{1}{1-\kappa+\frac{\kappa}{2}{\rm e}^{\nu^2/2}\Gamma(0,\nu^2/2)}\\ \times\left\{(\nu^2-1)+\frac{\kappa}{2}\left[2-{\rm e}^{\nu^2/2}\Gamma(0,\nu^2/2)\right]\right\}. \tag{25}$$

For $\kappa=0$, we recover the usual Markovian result in equation (20). For $\nu\gg1$, corresponding to large masses, ${\rm e}^{\nu^2/2}\Gamma(0,\nu^2/2)\to 2/\nu^2$, and the above expression simplifies to

$$b_{\rm h}(\nu)\to 1+\frac{\nu^2-1}{\delta_{\rm c}}\left(\frac{1+\frac{\kappa}{\nu^2}}{1-\kappa+\frac{\kappa}{\nu^2}}\right)\\ \approx\frac{1}{1-\kappa}\,\frac{\nu^2}{\delta_{\rm c}}\,,\qquad \text{for } \nu\gg1\,. \tag{26}$$

3.3 Adding a stochastic barrier

In the presence of the stochastic barrier described in Section 2.3, we can easily modify the halo bias in equation (25) using the substitution $\delta_c \to a^{1/2}\delta_c$ and $\kappa \to a\kappa$, where the parameter a is related to the diffusion coefficient of the barrier. The Eulerian halo bias finally reads as

$$b_h(\nu) = 1 + \frac{1}{\sqrt{a}\delta_c} \frac{1}{1 - a\kappa + \frac{a\kappa}{2} e^{a\nu^2/2}\Gamma(0, a\nu^2/2)} \times \left\{ (a\nu^2 - 1) + \frac{a\kappa}{2}\left[2 - e^{a\nu^2/2}\Gamma(0, a\nu^2/2)\right] \right\}. \quad (27)$$

We note that equation (27) raises the halo bias in the large ν (i.e. large halo mass) region compared to the bias in the ellipsoidal collapse model (Sheth et al. 2001), and in fact get closer to the spherical result. For $\nu \gg 1$, using again the asymptotic expression of the incomplete Gamma function and keeping only the leading term $\sim\nu^2$, our result reads as

$$b_h(\nu) \simeq \frac{a^{1/2}}{(1 - a\kappa)} \frac{\nu^2}{\delta_c}, \qquad \text{for } \nu \gg 1, \quad (28)$$

which differs from the asymptotic spherical collapse result by an overall factor of $a^{1/2}/(1 - a\kappa)$.

4 COMPARISONS

We now compare the predictions for the Eulerian halo bias $b_h(\nu)$ from our non-Markovian and stochastic barrier model with those from the standard excursion set theory as well as *N*-body simulations. We present the results for the halo mass function in parallel since as we have shown in Sections 2 and 3, an analytic theory for halo formation provides simultaneous predictions for the mass function and bias.

Our model contains two parameters: (1) κ, which parametrizes the degree of non-Markovianity and its exact value depends on the filter function used to smooth the density, e.g. $\kappa = 0,\ 0.35, 0.45$ for tophat in momentum-space, Gaussian and tophat in coordinate filters, respectively, and (2) a, which parametrizes the stochasticity of the diffusing barrier with the diffusion coefficient D_B, where $a = 1/(1 + D_B)$. There is no a priori reason to favour one filter to another nor is the choice of filters limited to the three functional forms given above. Furthermore, we recall that in MR1 the scaling $\kappa \to a\kappa$ is obtained under the simplified assumption that the barrier makes a simple Brownian motion around the spherical collapse barrier; for more complicated stochastic motions of the barrier (and also for fluctuations around the ellipsoidal barrier), the rescaling of κ might be different. For these reasons, we prefer in this work to treat both a and κ as free parameters and use simulations to calibrate their values.

For the simulations, we choose to compare with the fits to the latest *N*-body simulations (Tinker et al. 2008, 2010). These papers provide detailed discussions about the comparison within the different *N*-body results and numerical issues such as the dependence of the results on simulation resolution and halo definitions and finders.

Fig. 1 shows the results for the bias (left-hand panels) and mass function (right-hand panels) as a function of halo mass (as parametrized by $\nu = \delta_c/\sigma$). The bottom panels show the fractional difference between each model prediction and the fit to *N*-body results. For the spherical and ellipsoidal models, we use the standard parameters listed in Table 1. For our model, we plot the predictions using $a = 0.818$ and $\kappa = 0.23$, which provide a good match (within $\sim$20 per cent) to both the bias and the mass function from *N*-body, and in particular to the mass function at high mass.

We note that since κ is related to the two-point correlation function of the density field (see equation 91 of MR1), it can in principle depend on the cosmological model. As discussed in footnote 10 of MR1, however, the dependence of κ on the cosmological parameters is extremely weak. Both κ and a can therefore be treated as universal parameters whose values can be calibrated with simulations.

5 CONCLUSIONS

We derived an analytic expression (equation 27 with $a = 1$) for the halo bias in the non-Markovian extension of the excursion set theory. This new model is based on a path integral formulation introduced in MR1, which provides an analytic framework for handling the non-Markovian nature of the random walk and for calculating perturbatively the non-Markovian corrections to the standard version of the excursion set theory. The degree of non-Markovianity in our theory is parametrized by a single variable, κ, whose exact value depends on the shape of the filter function used to smooth the density field, e.g. $\kappa = 0$ for a tophat filter in momentum space, $\kappa \approx 0.35$ for a Gaussian filter and $\kappa \approx 0.44$ for a tophat filter in coordinate space.

As already discussed in Bond et al. (1991), Robertson et al. (2009) and MR1, changing the filter function in the spherical collapse model from a tophat in momentum space to a tophat in coordinate space does not help alleviate the discrepancy in the mass function between the Press–Schechter model and *N*-body simulations. In another word, had we plotted the corresponding curves in Fig. 1 using $a = 1$ (i.e. a constant barrier height as in the spherical collapse model) and $\kappa = 0.44$ (for a tophat filter in coordinate space), the bias would be too high by up to $\sim$80 per cent at large ν compared to the *N*-body result and the mass function would be too low by up to $\sim$80 per cent at large mass (see also fig. 9 of MR1). Using a Gaussian filter reduces κ by only $\sim$20 per cent and has only a minor effect. Additional modifications to the theory beyond including non-Markovian corrections must therefore be introduced to match the *N*-body results.

We have explored one such modification by allowing the barrier height itself to be a stochastic variable (Section 2.3 and Maggiore & Riotto 2010b). This new ingredient introduces a second parameter a in our theory, as summarized in Table 1. As the solid black curves in Fig. 1 illustrate, an appropriate choice of these two parameters for the non-Markovian correction and stochastic barrier (e.g. $\kappa = 0.23$ and $a = 0.818$) produces a good match to *N*-body results for *both* the halo mass function and bias, with fractional deviations being $\sim$20 per cent or less. In comparison, the ellipsoidal collapse model contains four fitting parameters (a, b, c and q; see Table 1) and does a comparable job at matching *N*-body simulations (dot–dashed magenta curves in Fig. 1).

Further improvement to the model presented in this paper can be obtained by computing the bias through the excursion set theory starting from the ellipsoidal model and including the effects of non-Markovianity. A step towards this computation has been taken recently in De Simone et al. (2010), where the first-crossing rate has been computed using the path integral method for a generic barrier. As the next step, one could envisage to combine the diffusing barrier model with the ellipsoidal model, i.e. consider a barrier that fluctuates around an average value given by the ellipsoidal collapse model, rather than around the constant value provided by the spherical collapse model as done in this paper. We expect this combined model to be able to provide an even closer match to *N*-body results than the $\sim$20 per cent accuracy achieved by either model alone.

ACKNOWLEDGMENTS

Support for C-PM is provided in part by the Miller Institute for Basic Research in Science, University of California Berkeley. The work of MM is supported by the Fonds National Suisse. We thank Roman Scoccimarro for a useful comment.

REFERENCES

Bardeen J. M., Bond J. R., Kaiser N., Szalay A., 1986, ApJ, 304, 15
Bond J., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Cole S., Kaiser N., 1989, MNRAS, 237, 1127
De Simone A., Maggiore M., Riotto A., MNRAS, submitted, preprint (arXiv:1007.1903 [astro-ph.CO])
Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715
Jing Y. P., 1998, ApJ, 503, L9
Kaiser N., 1984, ApJ, 284, L9
Maggiore M., Riotto A., 2010a, ApJ, 711, 907 (MR1)
Maggiore M., Riotto A., 2010b, ApJ, 717, 515
Maggiore M., Riotto A., 2010c, ApJ, 717, 526
Maggiore M., Riotto A., 2010d, MNRAS, 405, 1244
Mo H., White S., 1996, MNRAS, 282, 347
Pillepich A., Porciani C., Hahn O., 2010, MNRAS, 402, 191
Press W. H., Schechter P., 1974, ApJ, 187, 425
Robertson B., Kravtsov A., Tinker J., Zentner A., 2009, ApJ, 696, 636
Seljak U., Warren M., 2004, MNRAS, 355, 129
Sheth R., Tormen G., 1999, MNRAS, 308, 119
Sheth R., Tormen G., 2002, MNRAS, 329, 61
Sheth R., Mo H., Tormen G., 2001, MNRAS, 323, 1
Tinker J., Weinberg D., Zheng Z., Zehavi I., 2005, ApJ, 631, 41
Tinker J. et al., 2008, ApJ, 688, 709
Tinker J., Robertson B., Kravtsov A., Klypin A., Warren M., Yepes G., Gottlober S., 2010, ApJ, 724, 878
Zentner A., 2007, Int. J. Modern Phys. D, 16, 763

APPENDIX A: DETAILS OF THE COMPUTATION

To perform the computation, we use the technique discussed in detail in MR1. We first consider the numerator in equation (21). We start from equation (12), with the two-point function $\langle\delta_i\delta_j\rangle_c$ given in equations (15) and (16), and we expand to first order in κ (recall that Δ_{ij} is proportional to κ). This gives W in terms of $W^{\rm gm}$:

$$\begin{aligned} W(\delta_0;\ldots,\delta_n;S_n) &= \int \mathcal{D}\lambda\ {\rm e}^{i\sum_{i=1}^n \lambda_i\delta_i - \frac{1}{2}\sum_{i,j=1}^n \lambda_i\lambda_j(\min(S_i,S_j)+\Delta_{ij})} \\ &\simeq W^{\rm gm}(\delta_0;\ldots,\delta_n;S_n) \\ &+\frac{1}{2}\sum_{i,j=1}^n \Delta_{ij}\partial_i\partial_j W^{\rm gm}(\delta_0;\ldots,\delta_n;S_n)\,, \end{aligned} \tag{A1}$$

where $\Delta_{ij} \equiv \Delta(S_i, S_j)$, $\partial_i \equiv \partial/\partial\delta_i$ and we have used the identity

$$\lambda_k {\rm e}^{i\sum_{j=1}^n \lambda_j\delta_j} = -{\rm i}\partial_k {\rm e}^{i\sum_{j=1}^n \lambda_j\delta_j} \tag{A2}$$

to transform the factor $-\Delta_{ij}\lambda_i\lambda_j$ coming from the expansion of the exponential into $\Delta_{ij}\partial_i\partial_j$. It is convenient to split the sum into various pieces:

$$\begin{aligned} \frac{1}{2}\sum_{i,j=1}^n \Delta_{ij}\partial_i\partial_j &= \frac{1}{2}\sum_{i,j=1}^{m-1} \Delta_{ij}\partial_i\partial_j + \sum_{i=1}^{m-1}\Delta_{im}\partial_i\partial_m \\ &+\frac{1}{2}\sum_{i,j=m+1}^{n-1}\Delta_{ij}\partial_i\partial_j + \sum_{i=m+1}^{n-1}\Delta_{in}\partial_i\partial_n \\ &+\sum_{i=1}^{m-1}\Delta_{in}\partial_i\partial_n + \Delta_{mn}\partial_m\partial_n \\ &+\sum_{i=1}^{m-1}\sum_{j=m+1}^{n-1}\Delta_{ij}\partial_i\partial_j + \sum_{j=m+1}^{n-1}\Delta_{jm}\partial_j\partial_m\,. \end{aligned} \tag{A3}$$

First, consider the contribution from the first line of this expression. Using the factorization property (23) of $W^{\rm gm}$, its contribution to the numerator in equation (21) can be written as

$$\begin{aligned} &\int_{-\infty}^{\delta_c} {\rm d}\delta_1\cdots{\rm d}\delta_{m-1}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1} \\ &\left[\frac{1}{2}\sum_{i,j=1}^{m-1}\Delta_{ij}\partial_i\partial_j + \sum_{i=1}^{m-1}\Delta_{im}\partial_i\partial_m\right] \\ &\quad\times W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m)W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n-S_m) \\ &= \int_{-\infty}^{\delta_c}{\rm d}\delta_1\cdots{\rm d}\delta_{m-1}\left[\frac{1}{2}\sum_{i,j=1}^{m-1}\Delta_{ij}\partial_i\partial_j + \sum_{i=1}^{m-1}\Delta_{im}\partial_i\partial_m\right] \\ &\quad\times W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m) \\ &\quad\times\int_{-\infty}^{\delta_c}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1}W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n-S_m) \\ &\quad+\int_{-\infty}^{\delta_c}{\rm d}\delta_1\cdots{\rm d}\delta_{m-1}\sum_{i=1}^{m-1}\Delta_{im}\partial_i W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m) \\ &\quad\times\int_{-\infty}^{\delta_c}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1}\partial_m W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n-S_m)\,. \end{aligned} \tag{A4}$$

The first term is easily dealt by observing that

$$\begin{aligned} &\int_{-\infty}^{\delta_c}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1}W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n-S_m) \\ &= \Pi^{\rm gm}(\delta_m;\delta_n;S_n-S_m)\,. \end{aligned} \tag{A5}$$

Combining this with the contribution coming from the zeroth-order term $W^{\rm gm}(\delta_0;\ldots,\delta_n;S_n)$ in equation (A1) and using again the factorization property (23) of $W^{\rm gm}$, we therefore get

$$\begin{aligned} &\Pi^{\rm gm}(\delta_m;\delta_n;S_n-S_m)\int_{-\infty}^{\delta_c}{\rm d}\delta_1\cdots{\rm d}\delta_{m-1} \\ &\quad\times\left[1+\frac{1}{2}\sum_{i,j=1}^{m-1}\Delta_{ij}\partial_i\partial_j + \sum_{i=1}^{m-1}\Delta_{im}\partial_i\partial_m\right] \\ &\quad\times W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m) \\ &+\int_{-\infty}^{\delta_c}{\rm d}\delta_1\cdots{\rm d}\delta_{m-1}\sum_{i=1}^{m-1}\Delta_{im}\partial_i W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m) \\ &\quad\times\int_{-\infty}^{\delta_c}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1}\partial_m W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n-S_m)\,. \end{aligned} \tag{A6}$$

We now observe that the terms in brackets give just the expansion to $O(\kappa)$ of the denominator in equation (21). Therefore, to $O(\kappa)$, we can write

$$P(\delta_n, S_n|\delta_m, S_m) = \Pi^{\rm gm}(\delta_m;\delta_n;S_n - S_m) + P^{\rm non\text{-}mark}(\delta_n, S_n|\delta_m, S_m)\,, \tag{A7}$$

where

$$P^{\rm non\text{-}mark}(\delta_n, S_n|\delta_m, S_m) = \frac{N_a + N_b + N_c + N_d}{\Pi^{\rm gm}(\delta_0;\delta_m;S_m)}\,, \tag{A8}$$

and $N_a, \ldots, N_d$ are defined by

$$N_a = \int_{-\infty}^{\delta_{\rm c}} {\rm d}\delta_1\cdots{\rm d}\delta_{m-1}\sum_{i=1}^{m-1}\Delta_{im}\partial_i W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m) \times\int_{-\infty}^{\delta_{\rm c}}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1}\partial_m W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n - S_m)\,, \tag{A9}$$

$$N_b = \int_{-\infty}^{\delta_{\rm c}} {\rm d}\delta_1\cdots{\rm d}\delta_{m-1}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1} \left[\frac{1}{2}\sum_{i,j=m+1}^{n-1}\Delta_{ij}\partial_i\partial_j + \sum_{i=m+1}^{n-1}\Delta_{in}\partial_i\partial_n\right] \times W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m)W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n - S_m)\,, \tag{A10}$$

$$N_c = \int_{-\infty}^{\delta_{\rm c}} {\rm d}\delta_1\cdots{\rm d}\delta_{m-1}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1} \left[\sum_{i=1}^{m-1}\Delta_{in}\partial_i\partial_n + \Delta_{mn}\partial_m\partial_n\right] \times W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m)W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n - S_m)\,, \tag{A11}$$

$$N_d = \int_{-\infty}^{\delta_{\rm c}} {\rm d}\delta_1\cdots{\rm d}\delta_{m-1}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1} \left[\sum_{i=1}^{m-1}\sum_{j=m+1}^{n-1}\Delta_{ij}\partial_i\partial_j + \sum_{j=m+1}^{n-1}\Delta_{jm}\partial_j\partial_m\right] \times W^{\rm gm}(\delta_0;\ldots,\delta_m;S_m)W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n - S_m)\,. \tag{A12}$$

The contribution N_a comes from equation (A4), while N_b, N_c and N_d come from the second, third and fourth lines in equation (A3), respectively. Observe that in the denominator in equation (A8) we could replace $\Pi(\delta_0;\delta_m;S_m)$ by $\Pi^{\rm gm}(\delta_0;\delta_m;S_m)$, since the numerator is proportional to Δ_{ij} and therefore is already $\mathcal{O}(\kappa)$.

The contributions $N_a, \ldots, N_d$ can be computed using the techniques developed in MR1. The term N_a is immediately obtained using equations (105) and (110) of MR1 and is given by

$$N_a = \kappa\,\frac{\delta_{\rm c}(\delta_{\rm c} - \delta_m)}{S_m}{\rm Erfc}\left(\frac{2\delta_{\rm c} - \delta_m}{\sqrt{2S_m}}\right) \times\partial_m\Pi^{\rm gm}(\delta_m;\delta_n;S_n - S_m)\,, \tag{A13}$$

where Erfc is the complementary error function. The term N_b is given by

$$N_b = \Pi^{\rm gm}(\delta_0;\delta_m;S_m) \times[\Pi^{b1}(\delta_m, S_m;\delta_n, S_n) + \Pi^{b2}(\delta_m, S_m;\delta_n, S_n)]\,, \tag{A14}$$

where

$$\Pi^{b1}(\delta_m, S_m;\delta_n, S_n) \equiv \int_{-\infty}^{\delta_{\rm c}}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1} \times\sum_{i=m+1}^{n-1}\Delta_{in}\partial_i\partial_n W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n - S_m) \tag{A15}$$

and

$$\Pi^{b2}(\delta_m, S_m;\delta_n, S_n) \equiv \int_{-\infty}^{\delta_{\rm c}}{\rm d}\delta_{m+1}\cdots{\rm d}\delta_{n-1} \times\frac{1}{2}\sum_{i,j=m+1}^{n-1}\Delta_{ij}\partial_i\partial_j W^{\rm gm}(\delta_m;\ldots,\delta_n;S_n - S_m)\,. \tag{A16}$$

The computation of Π^{b1} and Π^{b2} is quite similar to the computation of the terms called $\Pi^{\rm mem}$ and $\Pi^{\rm mem\text{-}mem}$ in MR1, and in the continuum limit $\epsilon \to 0$ we get

$$\begin{aligned}\Pi^{b1}(\delta_m, S_m;\delta_n, S_n) &= \partial_n\lim_{\epsilon\to 0}\frac{1}{\epsilon}\int_{S_m}^{S_n}{\rm d}S_i \times\Delta(S_i, S_n)\Pi_\epsilon^{\rm gm}(\delta_m;\delta_{\rm c};S_i - S_m)\Pi_\epsilon^{\rm gm}(\delta_{\rm c};\delta_n;S_n - S_i)\\ &= \frac{\kappa}{\pi}(\delta_{\rm c} - \delta_m)\partial_n\left\{(\delta_{\rm c} - \delta_n)\int_{S_m}^{S_n}{\rm d}S_i \times\frac{S_i}{S_n(S_i - S_m)^{3/2}(S_n - S_i)^{1/2}} \times\exp\left[-\frac{(\delta_{\rm c} - \delta_m)^2}{2(S_i - S_m)} - \frac{(\delta_{\rm c} - \delta_n)^2}{2(S_n - S_i)}\right]\right\}\end{aligned} \tag{A17}$$

and

$$\begin{aligned}\Pi^{b2}(\delta_m, S_m;\delta_n, S_n) &= \lim_{\epsilon\to 0}\frac{1}{\epsilon^2}\int_{S_m}^{S_n}{\rm d}S_i\int_{S_i}^{S_n}{\rm d}S_j \times\Delta(S_i, S_j)\Pi_\epsilon^{\rm gm}(\delta_m;\delta_{\rm c};S_i - S_m) \times\Pi_\epsilon^{\rm gm}(\delta_{\rm c};\delta_{\rm c};S_j - S_i)\Pi_\epsilon^{\rm gm}(\delta_{\rm c};\delta_n;S_n - S_j)\\ &= \frac{\kappa}{\pi\sqrt{2\pi}}(\delta_{\rm c} - \delta_m)(\delta_{\rm c} - \delta_n) \times\int_{S_m}^{S_n}{\rm d}S_i\,\frac{S_i}{(S_i - S_m)^{3/2}}{\rm e}^{-(\delta_{\rm c}-\delta_m)^2/[2(S_i - S_m)]} \times\int_{S_i}^{S_n}{\rm d}S_j\,\frac{{\rm e}^{-(\delta_{\rm c}-\delta_n)^2/[2(S_n - S_j)]}}{S_j(S_j - S_i)^{1/2}(S_n - S_j)^{3/2}}\,.\end{aligned} \tag{A18}$$

This can be rewritten as a total derivative with respect to δ_n, as

$$\Pi^{b2}(\delta_m, S_m;\delta_n, S_n) = \frac{\kappa}{\pi\sqrt{2\pi}}(\delta_{\rm c} - \delta_m)\partial_n \times\int_{S_m}^{S_n}{\rm d}S_i\,\frac{S_i}{(S_i - S_m)^{3/2}}{\rm e}^{-(\delta_{\rm c}-\delta_m)^2/[2(S_i - S_m)]} \times\int_{S_i}^{S_n}{\rm d}S_j\,\frac{{\rm e}^{-(\delta_{\rm c}-\delta_n)^2/[2(S_n - S_j)]}}{S_j(S_j - S_i)^{1/2}(S_n - S_j)^{1/2}}\,. \tag{A19}$$

The fact that both Π^{b1} and Π^{b2} can be written as a derivative with respect to δ_n simplifies considerably the computation of the flux $\mathcal{F}(S)$, since we can integrate $\partial_n \equiv \partial/\partial\delta_n$ by parts, and then we only need to evaluate the integrals in equations (A17) and (A19) in $\delta_n = \delta_{\rm c}$, which can be done analytically, as discussed in MR1.

The term N_c is a total derivative with respect to ∂_n of a quantity that vanishes in $\delta_n = \delta_{\rm c}$ so, when inserted into equation (22), it gives a vanishing contribution to the first-crossing rate. The most complicated term is N_d. Using the techniques developed in MR1, a rather long computation gives

$$N_d = \frac{\kappa}{\pi}\partial_n\left\{\delta_c(\delta_c-\delta_m)\mathrm{Erfc}\left(\frac{2\delta_c-\delta_m}{\sqrt{2S_m}}\right)\partial_m I\right. \quad (A20)$$

$$\left.+\Pi^{gm}(\delta_0;\delta_m;S_m)\tilde{N}_d\right\}, \quad (A21)$$

where

$$\tilde{N}_d = -\delta_m(\delta_c-\delta_m)I(\delta_m,\delta_n)+S_m(\delta_c-\delta_m)\partial_m I(\delta_m,\delta_n) -S_m I(\delta_m,\delta_n), \quad (A22)$$

and

$$I(\delta_m,\delta_n)\equiv\int_{S_m}^{S_n}\mathrm{d}S_j\frac{1}{S_j(S_j-S_m)^{1/2}(S_n-S_j)^{1/2}} \times\exp\left\{-\frac{(\delta_c-\delta_m)^2}{2(S_j-S_m)}-\frac{(\delta_c-\delta_n)^2}{2(S_n-S_j)}\right\}. \quad (A23)$$

Using equations (A8) and (A14), equation (A7) can be rewritten as

$$P(\delta_n,S_n|\delta_m,S_m)=\Pi^{gm}+\Pi^{b1}+\Pi^{b2}+\frac{N_a+N_c+N_d}{\Pi^{gm}(\delta_0;\delta_m;S_m)}. \quad (A24)$$

We can now compute the contribution to the flux from the various terms. The term Π^{gm} gives the zeroth-order term:

$$\mathcal{F}^{gm}(S_n|\delta_m,S_m)=-\frac{\partial}{\partial S_n}\int_{-\infty}^{\delta_c}\mathrm{d}\delta_n\,\Pi^{gm}(\delta_m;\delta_n;S_n-S_m) =\frac{1}{\sqrt{2\pi}}\frac{\delta_c-\delta_m}{(S_n-S_m)^{3/2}}\mathrm{e}^{-(\delta_c-\delta_m)^2/[2(S_n-S_m)]}. \quad (A25)$$

The contribution of Π^{b1} to the flux is zero since it is the derivative with respect to ∂_n of a quantity that vanishes in $\delta_n=\delta_c$, and the same holds for N_c. The contribution of Π^{b2} is

$$\mathcal{F}^{b2}(S_n|\delta_m,S_m) \quad (A26)$$

$$=-\frac{\partial}{\partial S_n}\int_{-\infty}^{\delta_c}\mathrm{d}\delta_n\,\Pi^{b2}(\delta_m,S_m;\delta_n;S_n)$$

$$=-\frac{\kappa}{\pi\sqrt{2\pi}}(\delta_c-\delta_m)\frac{\partial}{\partial S_n} \times\int_{S_m}^{S_n}\mathrm{d}S_i\,\frac{S_i}{(S_i-S_m)^{3/2}}\mathrm{e}^{-(\delta_c-\delta_m)^2/[2(S_i-S_m)]} \times\int_{S_i}^{S_n}\mathrm{d}S_j\frac{1}{S_j(S_j-S_i)^{1/2}(S_n-S_j)^{1/2}}. \quad (A27)$$

The inner integral is elementary:

$$\int_{S_i}^{S_n}\mathrm{d}S_j\frac{1}{S_j(S_j-S_i)^{1/2}(S_n-S_j)^{1/2}}=\frac{\pi}{(S_iS_n)^{1/2}}, \quad (A28)$$

and we end up with

$$\mathcal{F}^{b2}(S_n|\delta_m,S_m)=-\frac{\partial}{\partial S_n}\left[\frac{\kappa(\delta_c-\delta_m)}{\sqrt{2\pi S_n}} \times\int_{S_m}^{S_n}\mathrm{d}S_i\,\frac{S_i^{1/2}}{(S_i-S_m)^{3/2}}\mathrm{e}^{-(\delta_c-\delta_m)^2/[2(S_i-S_m)]}\right], \quad (A29)$$

which generalized equation (118) of MR1 to $\delta_m\neq0$ and $S_m\neq0$. For S_m generic, the integral cannot be performed analytically. However, for computing the bias we are actually interested in the limit $S_m\to0$ with δ_m generic, and we see that in this case this contribution reduces to that computed in MR1, with the replacement $\delta_c\to\delta_c-\delta_m$.

The remaining contributions can be computed similarly. For the term N_d, again, rather than computing explicitly the derivative $\partial_n=\partial/\partial\delta_n$ in equation (A20), it is convenient to insert this expression directly into the first-crossing rate (22) and use the fact that it is a total derivative with respect to ∂_n to perform the integral over $\mathrm{d}\delta_n$. So, in the end, we only need

$$I(\delta_m,\delta_n=\delta_c)=\frac{\pi}{(S_mS_n)^{1/2}}\mathrm{e}^{+(\delta_c-\delta_m)^2/(2S_m)} \times\mathrm{Erfc}\left[(\delta_c-\delta_m)\sqrt{\frac{S_n}{2S_m(S_n-S_m)}}\right]. \quad (A30)$$

We can now put together all the terms and take the limit $S_m\to0$ (with $\delta_c-\delta_m>0$). In this limit the Erfc function in equation (A30) reduces to an exponential, so

$$I(\delta_m,\delta_n=\delta_c)\simeq\frac{\sqrt{2\pi(S_n-S_m)}}{(\delta_c-\delta_n)S_n}\mathrm{e}^{-(\delta_c-\delta_n)^2/[2(S_n-S_m)]}. \quad (A31)$$

Denoting $\delta_m=\delta_0$ in this limit, we finally get

$$\mathcal{F}(S|\delta_0,S_m=0)=\frac{1-\kappa}{\sqrt{2\pi}}\frac{\delta_c-\delta_0}{S^{3/2}}\mathrm{e}^{-(\delta_c-\delta_0)^2/(2S)} +\frac{\kappa}{2\sqrt{2\pi}}\frac{\delta_c-\delta_0}{S^{3/2}}\Gamma\left(0,\frac{(\delta_c-\delta_0)^2}{2S}\right) -\frac{\kappa}{\sqrt{2\pi}}\frac{\delta_0}{S^{3/2}}\left[1-\frac{(\delta_c-\delta_0)^2}{S}\right]\mathrm{e}^{-(\delta_c-\delta_0)^2/(2S)}. \quad (A32)$$

Expanding this result to first order in δ_0, we obtain equation (24).

This paper has been typeset from a TEX/LATEX file prepared by the author.