The Maximum of Multifractal Cascades: Exact Distribution and Approximations

By
Daniele Veneziano and Andreas Langousis

Department of Civil and Environmental Engineering, MIT, Cambridge, Mass., USA

Submitted to
Fractals

January, 2005

---

*Corresponding author: Daniele Veneziano, Dept. of Civil and Environmental Engineering, MIT, Cambridge, Mass. 02139. Tel: 617-253-7199, venezian@mit.edu.

Abstract

We study the distribution of the maximum M of multifractal measures using discrete cascade representations. For such discrete cascades, the exact distribution of M can be found numerically. We evaluate the sensitivity of the distribution of M to simplifying approximations, including independence of the measure among the cascade tiles and replacement of the dressing factor by a random variable with the same distribution type as the cascade generator. We also examine how the distribution of M varies with the dimensionality of the support and the multiplicity of the cascade. Of these factors, dependence of the measure among different cascade tiles has the highest effect on the distribution of M. This effect comes mainly from long-range dependence. We use these findings to propose a simple approximation to the distribution of M and give charts to implement the approximation for beta-lognormal cascades.

Keywords: multifractal processes, multiplicative cascades, multifractal extremes.

1. Introduction

Multifractal models have been used to represent stationary scale invariant phenomena in many areas of the applied sciences. In several cases, interest is in the distribution of extreme values, which for example might control the occurrence of hazardous conditions. A growing body of literature deals with how the marginal quantiles of a stationary multifractal measure scale with the resolution and the exceedance probability (Bendjoudi et al., 1997, 1999; Veneziano and Furcolo, 2002; Castro et al., 2004), but the problem of actually calculating the distribution F_M of the maximum M has not been addressed. In order to evaluate hazards (rather than relative hazards), one needs such distribution.

Our main objective is to find F_M using discrete cascade approximations to continuous multifractal processes and study the dependence of F_M on the dimensionality of the support d and the resolution r. Other issues we consider are the relative importance for F_M of short-range and long-range dependence among the cascade tiles and the effect of the cascade multiplicity. The multiplicity, m, is the number of sub-tiles into which a tile is partitioned at each stage of the cascade construction.

Section 2 recalls basic facts about discrete cascades that are pertinent to our analysis. Section 3 derives the exact distribution of M and evaluates its sensitivity to d and m. Section 4 examines the effects on F_M of assuming independence of the measure in different cascade tiles and approximating the dressing factor. A detailed analysis of the relative importance of short-range and long-range dependence is made in Appendix A. Section 5 uses these results to propose a practical approximation to the distribution of M. Implementation of the approximation for beta-lognormal cascades is discussed in detail.

2. Discrete Multifractal Cascades

We start by recalling basic facts about discrete multifractal cascades. Consider a discrete cascade in the unit d-dimensional cube S. Construction of the cascade proceeds as follows (Schertzer and Lovejoy, 1987; Mandelbrot, 1989; Gupta and Waymire, 1993). One starts at level 0 with a single tile $T_{0_1} = S$ and a measure with constant unit density in T_{0_1}. At level $n = 1, 2, \ldots$, each tile at the previous level $n - 1$ is partitioned into $m = m_l{}^d$ cubic tiles where $m_l > 1$ is the integer linear multiplicity of the cascade and m is the volumetric multiplicity. The measure density inside each cascade tile T_{n_i} ($i = 1, \ldots, m^n$) is obtained by multiplying the measure density in the parent tile at level $n - 1$ by a random variable Y_{n_i}. The variables Y_{n_i} are independent copies of a non-negative unit-mean random variable Y, called the generator of the cascade. We call $r = m^n$ the (volumetric) resolution when the cascade construction has reached level n.

A multifractal cascade is characterized by the multiplicity m and the distribution of the generator, F_Y. When using m rather than m_l, the dimensionality d does not matter since a d-dimensional cascade with parameters (m, F_Y) can be mapped into a 1-dimensional cascade with the same parameters (m, F_Y); see Figure 1 for an illustration for $d = 2$ and $m = 4$.

An important quantity in discrete cascades is $\varepsilon_{n,}$ the average measure density in a generic T_n tile. One may distinguish between two such average densities (Schertzer and Lovejoy, 1987): the bare density $\varepsilon_{b,n}$, which is the density when the cascade construction is terminated at level n, and the dressed density $\varepsilon_{d,n}$, which is the average density in T_n for the completely developed cascade. Therefore, $\varepsilon_{d,n}$ includes the effect of fluctuations at sub-tile scales, whereas $\varepsilon_{b,n}$ does not. The bare and dressed densities satisfy

$$\begin{cases} \varepsilon_{b,n} \overset{\mathrm{d}}{=} \prod_{i=1}^{n} Y_i \\ \varepsilon_{d,n} \overset{\mathrm{d}}{=} \varepsilon_{b,n} Z \end{cases} \tag{1}$$

where $Y_1, \ldots, Y_n$ are n independent copies of Y, Z is a random variable independent of $\varepsilon_{b,n}$ called the dressing factor, and $\overset{\mathrm{d}}{=}$ denotes equality in distribution. The factor Z has the same distribution as $\varepsilon_{d,0}$, the dressed measure in S.

Under certain conditions on the moments of Y (see below), Z and $\varepsilon_{d,n}$ have algebraic upper tails of the type

$$\begin{aligned} P[Z > z] &\sim z^{-q^*} \\ P[\varepsilon_{d,n} > \varepsilon] &\sim \varepsilon^{-q^*} \end{aligned} \tag{2}$$

where $\sim$ denotes equality up to a factor that varies slowly with z or ε and $q^* > 1$ is the order above which the moments $E[Z^q]$ and $E[(\varepsilon_{d,n})^q]$ diverge. One can find q^* from the condition $K(q) = \log_m E[Y^{q^*}] = q^* - 1$ (Kahane and Peyriere, 1976). If this equation has no solution greater than 1, then q^* does not exist and Z and $\varepsilon_{d,n}$ do not have algebraic upper tails.

3. The Exact Distribution of Cascade Maxima

Let $M_n = \max_{i=1,\ldots,m^n} (\varepsilon_{d,n_i})$ be the maximum dressed measure density in S at volumetric resolution m^n. The cumulative distribution function F_{M_n} can be found recursively for $n = 0, 1, \ldots$ by noting that $M_0 = Z$ and, for any $n > 0$, M_n is the maximum of m independent variables each with the distribution of YM_{n-1}. Therefore, working with logs,

$$\begin{aligned} F_{\log M_0} &= F_{\log Z} \qquad\qquad , n = 0 \\ F_{\log M_n} &= (F_{\log M_{n-1}} * f_{\log Y})^m \ , n = 1, 2, \ldots \end{aligned} \tag{3}$$

where f_X is the probability density function of X and $F*f$ is the convolution $F * f(s) = \int_{-\infty}^{\infty} F(s-y)\, f(y)\, dy$. The distribution of Z, which is needed for M_0, can itself be obtained through an iterative numerical procedure, as explained in Veneziano and Furcolo (2003).

Often in practice, one is interested in the measure density at some resolution $r = m^n$ with given return period T. This is the value that is exceeded by M_n on average once every T cascade realizations and is given by the upper $(1/T)$-quantile of M_n, $M_{n,T}$. As an example, for a specific "lognormal" cascade with $m = 2$, Figure 2 shows $M_{n,T}$ against $r = m^n$ for different T. The generator of a lognormal cascade with multiplicity m has lognormal distribution with mean value 1 and log-variance $(\sigma_{\ln Y})^2 = 2C \ln(m)$, where $0 < C < 1$ is the so-called co-dimension parameter of the cascade. Then $K(q) = C(q^2 - q)$ and the critical moment order is $q^* = 1/C$. In Figure 2 we have set $C = 0.1$ (hence $q^* = 10$) and $m = 2$. The quantile $M_{n,T}$ is known to have asymptotic power-law dependence on $r = m^n$ and T. Specifically, for T in any finite range and $r \to \infty$, $M_{n,T} \sim r^{-\gamma_1}\ T^{1/q_1}$ and for r in any finite range and $T \to \infty$, $M_{n,T} \sim r\ T^{1/q^*}$. In the first asymptotic expression, the exponent γ_1 is the slope of the tangent to $K(q)$ with K intercept equal to -1 and q_1 is the value of q at the point of tangency (Veneziano and Furcolo, 2002). For the cascade in Figure 2, $\gamma_1 = C(2\sqrt{1/C}-1) \approx 0.532$ and $1/q_1 = \sqrt{C} \approx 0.316$. These asymptotic scaling exponents are shown in Figure 2 and agree well with the numerical results.

It is interesting to examine how for given $K(q)$ the maximum M_r at resolution $r = m^n$ depends on the volumetric multiplicity m. One reason is that discrete cascades are often used as approximations to continuous multifractal processes. If M_r is insensitive to m, then one may expect the extreme of a discrete cascade to approximate well the extreme of the continuous process with the same $K(q)$. Another reason is that m depends on the space dimension d; for

example, $m = 2^d$ in so-called binary cascades. It is of interest to see whether, for given $K(q)$ and r, the maximum M_r depends on d.

The effect of m on M_r may be expressed through the ratio

$$R_{r,\,T|m} = \frac{M_{r,T|m}}{M_{r,T|m=2}} \tag{4}$$

where $M_{r,T|m}$ is the upper $(1/T)$-quantile of M_r for a cascade with multiplicity m. Figures 3a and 3b show the ratio $R_{r,\,T|m}$ for different r and T and $m = 4, 8$. All other parameters are as in Figure 2. The multiplicity m affects $M_{r,T|m}$ in two ways: 1) as m increases, the correlation between tiles at a given distance generally decreases, producing higher values of $M_{r,T|m}$, and 2) as m increases, the distribution of the dressing factor Z tightens around 1, causing a decrease of $M_{r,T|m}$ especially for large T. We have separately evaluated these two effects and found that the first one is negligible. Hence for values of T of practical interest the net effect is that $M_{r,T|m}$ decreases with increasing m. As Figure 3 shows, this effect is largest at low resolutions (because for small r the upper tail of $\varepsilon_{d,n}$ is dominated by the dressing factor Z).

4. The Maximum Under Independence Conditions

The exact procedure in equation (3) is tedious to implement due to the repeated convolutions (including convolutions in the numerical evaluation of F_Z; see Veneziano and Furcolo, 2003). In seeking approximations to M_n, we start by considering the distribution of $M_{n,ind}$, the maximum of the measure densities ε_{d,n_i} under the condition that the densities in different tiles are independent. Comparison of $M_{n,ind}$ with M_n will allow us to quantify the effect of dependence on multifractal extremes. In addition, $M_{n,ind}$ will be used in Section 5 to approximate M_n.

First we compare the exact distributions of $M_{n,ind}$ and M_n and then we consider an approximation of $M_{n,ind}$ in which the dressing factor Z is replaced with a random variable of the same type as the cascade generator Y.

4.1 Exact Distribution of $M_{n,ind}$

Under the assumption that the measure densities ε_{d,n_i} ($i = 1, \ldots, m^n$) are independent, one obtains

$$F_{\log M_{n,ind}} = (F_{\log Z} * f_{\log Y_1 + \ldots + \log Y_n})^{m^n} \tag{5}$$

where the variables $Y_1, \ldots, Y_n$ are independent copies of Y and Z is the dressing factor.

For the cascade in Figure 2, Figure 4a compares the distribution of M_n from equation (3) with the distribution of $M_{n,ind}$ from equation (5), for $n = 10$ and 20. Figure 4b and 4c display the same information as Figure 4a, but they show the non-exceedance and exceedance probabilities in log scale, thus focusing respectively on the lower and upper tail of the distribution.

It is clear from Figure 4 that dependence among the cascade tiles has a significant effect on the body and lower tail of the maximum distribution, but not on the extreme upper tail (in Figure 4c, the solid and dashed lines cannot be distinguished). The reason for convergence of the upper tails of $M_{n,ind}$ and M_n is that, as $x \to \infty$, the probability of more than one exceedance event $\varepsilon_{d,n_i} > x$ in S is an infinitesimal of higher order relative to the probability of just one event. Then, for large x,

$$\begin{aligned} P[M_n > x] &= P[\text{at least one exceedance event}] \approx P[\text{one exceedance event}] \\ &\approx E[\text{no. of exceedance events}] = m^n P[\varepsilon_{d,n} > x] \end{aligned} \tag{6}$$

The same reasoning holds for $M_{n,ind}$. Therefore, as $x \to \infty$, both $P[M_n > x]$ and $P[M_{n,ind} > x]$ approach $m^n P[\varepsilon_{d,n} > x]$. Like Z and $\varepsilon_{d,n}$, also M_n and $M_{n,ind}$ have power-law upper tails with exponent $-q^*$ (in Figure 4, $q^* = 10$). As n increases, this power-law behavior appears farther into

the tails of M_n and $M_{n,ind}$; this is why in Figure 4c the algebraic tail is visible for $n = 10$ but not for $n = 20$, although exceedance probabilities as small as 10^{-50} are shown.

It is not immediately obvious whether the difference between M_n and $M_{n,ind}$ is due mainly to small-scale or large-scale dependence (multifractal cascades display dependence at all scales below the size of the region where they are defined; see for example Cates and Deutsch, 1987; O'Neil and Meneveau, 1993; and Marsan *et al*., 1996). Appendix A investigates this issue by using modified cascades in which short-range or long-range dependence is progressively suppressed, while the marginal distribution is kept the same. The qualitative finding is that the distribution of M_n is affected mainly by long-range dependence.

In conclusion, dependence has an effect on the extremes of multifractal cascades. Therefore, when devising approximations to M_n, this effect will need to be considered.

4.2 An Approximation for Z

The distribution of the dressing factor Z, which is needed to calculate the exact distributions of M_n and $M_{n,ind}$, is not known analytically and its numerical evaluation is tedious (Veneziano and Furcolo, 2003). However, for n not very small and away from the extreme upper tail, the dressed density $\varepsilon_{d,n} = \varepsilon_{b,n} Z$ is dominated by the bare density $\varepsilon_{b,n}$ and one may expect small effects on M_n and $M_{n,ind}$ from approximating F_Z. Next we propose a particularly simple approximation.

Suppose that the log-generator $\log Y$ has infinitely divisible distribution (this is always the case for continuous multifractal processes and their approximating discrete cascades) and denote by $Y^{(r)}$ the generator for a volumetric scale-change factor $r > 1$. Notice that r is analogous to the volumetric multiplicity m, but is not necessarily an integer. If $Y^{(r)}$ has characteristic function $h^{(r)}(t)$, then for any $r_1, r_2 > 1$ the characteristic functions of $Y^{(r_1)}$ and $Y^{(r_2)}$ are related as $\left(h^{(r_1)}(t)\right)^{\ln r_1} = \left(h^{(r_2)}(t)\right)^{\ln r_2}$ (Veneziano, 1999). One can use this relationship to obtain the

distribution of $Y^{(r)}$ for any $r \geq 1$ from the distribution of the cascade generator $Y = Y^{(m)}$. In some cases, $Y^{(r)}$ has a simple distribution. For example, for a lognormal multifractal process with co-dimension parameter $0 < C < 1$, $Y^{(r)}$ has lognormal distribution with mean value 1 and log-variance $\mathrm{Var}[\ln Y^{(r)}] = 2C \ln r$.

Here we propose to approximate Z as $Y^{(r_Z)}$ where r_Z is such that Z and $Y^{(r_Z)}$ share some higher moment, in addition to the mean value. The q^{th} moment of Z is finite if $K(q) < q - 1$. Hence, if $K(2) < 1$, one may match the second moments

$$E[Z^2] = \frac{m - 1}{m - m^{K(2)}}$$
$$E\left[\left(Y^{(r_Z)}\right)^2\right] = r_Z{}^{K(2)} \tag{7}$$

The expression for $E[Z^2]$ follows from the consistency condition $Z \overset{\mathrm{d}}{=} \frac{1}{m}\sum_{i=1}^{m} Y_i^{(m)} Z_i$ and the fact that $E\left[\left(Y^{(m)}\right)^2\right] = m^{K(2)}$ (Kahane and Peyriere, 1976; Veneziano and Furcolo, 2003). By equating the moments in equation (7), one obtains

$$r_Z = \left(\frac{m - 1}{m - m^{K(2)}}\right)^{1/K(2)}, \quad K(2) < 1 \tag{8}$$

Figure 5 shows plots of r_Z against m for $K(2) = 0^+$, 0.2, 0.4, 0.6, 0.8, 0.92. Notice that for $K(2) = 1$, $E[Z^2]$ and r_Z diverge. Hence for $K(2)$ close to 1 (rare in practice), other approximations should be used. For small m and $0.2 < K(2) < 0.5$, r_Z is about 2-3.

For $n = 0$, 5, 10 and 20 and a lognormal cascade, Figure 6 compares the distribution of $M_{n,ind}$ using either the exact distribution of Z (solid lines) or the approximation $Y^{(r_Z)}$ (dashed-dotted lines). Other parameters are $m = 2$ and $C = 0.1$ as in Figure 4. For lognormal cascades $K(2) = 2C$; hence in our case $K(2) = 0.2$ and equation (8) gives $r_Z = 2.24$. Since $M_{0,ind} = Z$, the curves for $n =$

0 compare the exact distribution of Z (solid line) with the approximating distribution of $Y^{(r_Z)}$ (dashed-dotted line).

The reciprocal of the exceedance probability plotted in Figure 6 is the return period T (in units of cascade realizations). With this interpretation, Figure 6 shows that, for n = 0, 5, 10 and 20, the proposed approximation of Z produces accurate results for T less than approximately 10^4, 10^{10}, 10^{20} and 10^{30} cascade realizations, respectively. Hence the approximation is accurate, except for the combination of very low resolutions with extremely long return periods.

Replacement of Z with $Y^{(r_Z)}$ gives $\varepsilon_{d,n} \overset{d}{\approx} Y^{(r_Z r)}$. With this replacement, the upper $(1/T)$-quantile of $M_{n,ind}$, $M_{n,ind,T}$, equals the $(1\text{-}1/T)^{1/r}$-quantile of $Y^{(r_Z r)}$, $Y^{(r_Z r)}_{(1-1/T)^{1/r}}$. This approximation of $M_{n,ind,T}$ is used next to approximate the quantiles of M_n.

5. Approximation of M_n for Beta-lognormal Cascades

We obtain an approximation to the upper $(1/T)$-quantile of M_n, $M_{n,T}$, by correcting the quantile $M_{n,ind,T} \approx Y^{(r_Z r)}_{(1-1/T)^{1/r}}$ for dependence; i.e. by using

$$M_{n,T} \approx \gamma(n, T)\ Y^{(r_Z r)}_{(1-1/T)^{1/r}} \tag{9}$$

where $\gamma(n, T)$ is the ratio between the upper $(1/T)$-quantiles of M_n and $M_{n,ind}$,

$$\gamma(n, T) = \frac{M_{n,T}}{M_{n,ind,T}} \tag{10}$$

Next we give the factors $Y^{(r_Z r)}_{(1-1/T)^{1/r}}$ and $\gamma(n, T)$ in equation (9) for the case of beta-lognormal cascades.

Lognormal multifractal measures are among the most widely used stationary scale-invariant models; see for example Schertzer and Lovejoy (1987). Their generator $Y = Y_{ln}$ has lognormal distribution with $K(q) = C_{ln}(q^2 - q)$ where $C_{ln} = \mathrm{Var}[\ln(Y_{ln})]/(2\ln m)$. In some applications, for

example to rainfall, the measure of interest has fractal support (due to the alternation of rainy and dry conditions at different scales). To include this feature, one may use so-called beta-lognormal models, in which the generator Y is the product of two independent random factors: a factor Y_{ln} with lognormal distribution as above, and a discrete factor Y_β which attains value $1/p \geq 1$ with probability p and value 0 with probability $1 - p$ (for an application to rainfall, see for example Over and Gupta, 1996). In this case

$$K(q) = C_\beta (q\text{-}1) + C_{ln} (q^2 - q) \tag{11}$$

where $C_{ln} = \mathrm{Var}[\ln(Y_{ln})]/(2\ln m)$ as for lognormal cascades and $C_\beta = -\log_m p$. Beta-lognormal cascades are non-degenerate if $C_\beta + C_{ln} < 1$.

To obtain the quantile $Y^{(r_Z r)}_{(1\text{-}1/T)^{1/r}}$, we observe that for beta-lognormal cascades

$$Y^{(r)} = \begin{cases} 0\,, & \text{with probability } 1\text{-}r^{-C_\beta} \\ r^{C_\beta} \exp\left(-C_{ln} \ln r + \Phi\sqrt{2C_{ln} \ln r}\right), & \text{with probability } r^{-C_\beta} \end{cases} \tag{12}$$

where Φ is the standard normal variable. In obtaining equation (12), we have used the equality $p^{\log_m r} = r^{-C_\beta}$. It follows from equation (12) that

$$Y^{(r_Z r)}_{(1\text{-}1/T)^{1/r}} = \begin{cases} 0\,, & \text{if } 1\text{-}r^{-C_\beta} \geq (1\text{-}1/T)^{1/r} \\ (r_Z r)^{C_\beta} \exp\left(-C_{ln} \ln(r_Z r) + \sqrt{2C_{ln} \ln(r_Z r)}\, \Phi_{(1\text{-}(1\text{-}1/T)^{1/r})\, r^{C_\beta}}\right), & \text{otherwise} \end{cases} \tag{13}$$

where Φ_p is the value exceeded by Φ with probability p.

Now we turn to the quantile ratio $\gamma(n, T)$ in equation (10). For beta-lognormal cascades this ratio can be evaluated numerically, using equations (3) and (5). Figure 7 shows results for volumetric multiplicities m = 2, 4, 8 and two sets of co-dimension parameters, (C_β, C_{ln}) = (0, 0.1) and (0.2, 0.1). The cascade levels considered in Figure 7 vary with m so that the same range of resolution is covered in each panel. In principle, γ depends on the level n, the return period T, the

multiplicity m, and the co-dimension parameters (C_β, C_{ln}). However, as Figure 7 shows, γ is insensitive to C_β (because the intermittency of the process, which is controlled by C_β, has little effect on the dependence among the cascade tiles). Therefore, one may evaluate $\gamma(n, T)$ under $C_\beta = 0$, i.e. for a lognormal cascade with $Y = Y_{ln}$.

For bare lognormal cascades, $\gamma(n, T)$ has a scaling relationship with C_{ln}. A derivation of this relationship for the broader class of log-stable cascades is given in Appendix B. In the case of lognormal cascades (index of stability $\alpha = 2$), equations (B.8) and (B.10) in Appendix B give $\gamma_{b,C_{ln_1}}(n, T) = \left(\gamma_{b,C_{ln_2}}(n, T)\right)^{\sqrt{C_{ln_1}/C_{ln_2}}}$, where the subscript b indicates that this relationship holds for bare cascades. However, the same relation is satisfied in good approximation also for dressed cascades, i.e.

$$\gamma_{C_{ln_1}}(n, T) \approx \left(\gamma_{C_{ln_2}}(n, T)\right)^{\sqrt{C_{ln_1}/C_{ln_2}}} \tag{14}$$

For example, Figure 8 compares the exact dressed ratios $\gamma_{C_{ln}=0.05}(n, T)$ and $\gamma_{C_{ln}=0.2}(n, T)$ with approximations based on equation (14) using $C_{ln_1} = 0.05$ and 0.2 and $C_{ln_2} = 0.1$. The approximation is clearly very accurate. This is why, in Figure 7, we have considered only one value of C_{ln} ($C_{ln} = 0.1$). The factor $\gamma(n, T)$ for any other value of C_{ln} can be obtained through equation (14).

Figure 7 further shows that the dependence of γ on m (for fixed volumetric resolution $r = m^n$) is weak but not negligible. This means that γ depends somewhat on the space dimension d, if for example one considers binary cascades with volumetric multiplicity $m = 2^d$.

The issue of what value of m to use in practice is not trivial. If interest is in the maximum of a discrete cascade, then of course m should be the volumetric multiplicity of the cascade (e.g. $m = 2^d$ for "binary" cascades in d-dimensional space). However, if interest is in the maximum of

a continuous process of which the discrete cascade is an approximation, then m is not uniquely defined. Notice that one can construct approximating cascades with any integer multiplicity $m \geq 2$ in spaces of any dimension. It suffices to specify a scale-invariant tile-splitting rule in which at the generic level each tile is partitioned into m sub-tiles. Two rules with $m = 2$ in two-dimensional space are shown in Figure 9 (the fact that, in scheme (b), the initial tile is a square not a triangle is not a major violation of scale invariance). Hence, when interest is in approximating the maximum of a continuous process, one may use values of m as small as 2, irrespective of d.

Finally, we comment on the dependence of γ on n and T. Figure 7 shows that, for any given (m, C_{ln}, C_β), γ has maximum value 1 when $n = 0$ or $T \to \infty$. The reason why $\gamma \to 1$ as $T \to \infty$ is that, as was noted in Section 4.1, the extreme upper tails of M_n and $M_{n,ind}$ are practically identical. However, the lower quantiles are significantly larger for $M_{n,ind}$ than for M_n. This is why γ decreases as T decreases. A rough explanation for the effect of n on γ is that, as n increases, the number of dependent cascade tiles increases, inducing a larger difference between M_n and $M_{n,ind}$.

To summarize, we suggest the following approximation to $M_{n,T}$, the upper $(1/T)$-quantile of the maximum of a beta-lognormal cascade:

1. Given (C_{ln}, C_β) and the volumetric resolution m (for example $m = 2^d$ where d is the space dimension or $m = 2$ irrespective of d), calculate $K(2) = C_\beta + 2C_{ln}$ and find r_Z from equation (8);
2. For the volumetric resolution r and return period T of interest, find the quantile $Y^{(r_Z r)}_{(1-1/T)^{1/r}}$ from equation (13);
3. Use the scaling relation in equation (14) and the chart in Figure 7 for the appropriate m and (C_{ln}, C_β) = (0.1, 0) to find the correction factor for dependence, $\gamma_{C_{ln}}(n, T)$;

4. Approximate $M_{n,T}$ as $M_{n,T} \approx \gamma(n, T)\ Y^{(r_Z r)}_{(1-1/T)^{1/r}}$.

6. Conclusions

Discrete multiplicative cascades are often used in place of continuous multifractal processes as they are more amenable to theoretical and numerical analysis. Here we have used them to approximate continuous multifractal extremes.

The construction of multifractal cascades leads to a recursive numerical procedure for the distribution of the maximum M_n at volumetric resolution $r = m^n$, where m is the volumetric multiplicity of the cascade and n is the cascade level. A frequently used approximation to M_n, $M_{n,ind}$, ignores dependence among the cascade tiles. The distributions of M_n and $M_{n,ind}$ have almost identical extreme upper tails but differ significantly in the body and lower-tail regions. These differences are due mainly to long-range dependence among the cascade tiles.

The recursive procedure to calculate the exact distribution of M_n, given by equation (3), involves repeated convolution operations and is numerically tedious. Although simpler, calculation of the distribution of $M_{n,ind}$ through equation (5) also requires repeated convolutions (to obtain the distribution of the dressing factor Z).

In developing a practical approximation to $M_{n,ind}$, we have replaced Z with $Y^{(r_Z)}$, a random variable whose distribution is of the same type as that of the cascade generator Y and whose first two moments match the first two moments of Z. Then, to approximate M_n, we have included the effect of dependence among the cascade tiles through the ratio $\gamma(n, T)$ between the upper $(1/T)$-quantiles of M_n and $M_{n,ind}$. For beta-lognormal cascades, $\gamma(n, T)$ depends on the volumetric multiplicity m and the co-dimension parameters C_{ln} and C_β, in addition to the cascade level n and the return period T. We have found however that the effect of C_β is negligible and that $\gamma(n, T)$ has an analytical scaling relation with C_{ln}; see equation (14). Hence we have provided

charts of $\gamma(n, T)$ for $m = 2^d$ with $d = 1$, 2 and 3, $C_\beta = 0$, and just one reference value of C_{ln} ($C_{ln} = 0.1$); see Figure 7. At least for beta-lognormal cascades, the proposed approximation produces accurate estimates of the distribution of multifractal extremes.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant No. EAR-0228835, and in part by the Alexander S. Onassis Public Benefit Foundation under Scholarship No. F-ZA 054/2004-2005.

Appendix A: The Influence of Short- and Long-range Dependence on M_n

We examine whether the difference between the exact maximum of a multifractal cascade, M_n, and the maximum under the condition of independence, $M_{n,ind}$, is due mainly to short-range or long-range dependence. We do so by considering modified cascades in which short-range or long-range dependence is progressively suppressed, while the marginal distribution of the cascade is kept the same.

Suppose one wants to ignore dependence at large scales, say between cascade levels 0 and n_0. Then one should start the cascade construction at level n_0 and independently assign measure densities $\varepsilon_{d,n_0,i}$ to all the tiles at that level. The variables $\varepsilon_{d,n_0,i}$ have the distribution of Y_{n_0}, the product of n_0 independent variables, each with the distribution of the generator Y. The rest of the cascade construction remains the same. To obtain the distribution of M_n under these partial-dependence conditions, one follows the procedure in equation (3) to level $n - n_0$ and then calculates $F_{\log M_n}$ as

$$F_{\log M_n} = (F_{\log M_{n-n_0}} * f_{\log Y_{n_0}})^{m^n} \tag{A.1}$$

Figure 10 shows results for $n = 20$ when excluding large-scale dependence over $n_0 = 0$, 5, 10, 15 and 20 cascade steps. All other parameters are the same as in Figure 4. Notice that the limiting cases $n_0 = 0$ and $n_0 = 20$ correspond to the distribution of M_n and $M_{n,ind}$, respectively. Figure 10 shows that large-scale dependence is influential on the distribution of the maximum, in particular in the lower tail region.

Consider now the case when small-scale dependence is ignored, say between cascade levels $n - n_0$ and n (between volumetric resolutions m^{n-n_0} and m^n). In this case the original cascade construction proceeds to level $n - n_0$. Then, in a single step, each tile at level $n - n_0$ is partitioned into m^{n_0} tiles at resolution m^n. The measure density in each level-n tile is obtained by multiplying the density in the parent tile at level $n - n_0$ by an independent realization of the modified dressing factor $Z_{n_0} \overset{d}{=} Z\,Y_{(n_0)}$, where $Y_{(n_0)}$ is the product of n_0 independent variables each with the distribution of Y and Z is the dressing factor.

To find the distribution of M_n ignoring short-range dependence over n_0 cascade levels, one first finds the distribution of $\log Z_{n_0}$, $F_{\log Z_{n_0}}$. Raising this distribution to the m^{n_0} power gives the distribution of the maximum of $\log Z_{n_0}$ among the n-level tiles within a single $(n - n_0)$-level tile. One uses this distribution in place of $F_{\log Z}$ in the initialization step of equation (3) and then proceeds with (3) for $n - n_0$ steps. Figure 11 shows results when ignoring dependence between cascade levels $n - n_0$ and n, for $n_0 = 0$, 5, 10, 15 and 20. All other parameters are as in Figure 10. Again, the limiting cases $n_0 = 0$ and $n_0 = 20$ correspond to the exact distribution of M_n and the distribution of the completely independent approximation $M_{n,ind}$, respectively. One can see that removing dependence at the smallest scales (for example, for $n_0 = 5$, between volumetric resolutions 2^{15} and 2^{20}) has minimal effect on the distribution of the maximum.

Together, Figures 10 and 11 indicate that the body of the distribution of M_n is sensitive to long-range dependence, whereas short-range dependence is largely unimportant (especially when long-range dependence is present).

Appendix B: Dependence of Bare Log-stable Cascades and Their Maxima on the Co-dimension Coefficient

Let $\varepsilon_{b;n,\alpha,C}$ be the bare density at level n of a log-stable cascade with index of stability $0 < \alpha \leq 2$, volumetric multiplicity m and co-dimension parameter C, inside the unit cube S. Also let $M_{b,C}(\alpha, n)$ be the maximum of $\varepsilon_{b;n,\alpha,C}$ in S and $M_{b,C,ind}(\alpha, n)$ be the maximum in S under the condition that $\varepsilon_{b;n,\alpha,C}$ has independent values in different tiles T_{n_i}. Finally, let $M_{b,C}(\alpha, n, T)$ and $M_{b,C,ind}(\alpha, n, T)$ be the upper $(1/T)$-quantiles of $M_{b,C}(\alpha, n)$ and $M_{b,C,ind}(\alpha, n)$, respectively, and denote by $\gamma_{b,C}(\alpha, n, T)$ the quantile ratio

$$\gamma_{b,C}(\alpha, n, T) = \frac{M_{b,C}(\alpha, n)}{M_{b,C,ind}(\alpha, n)} \tag{B.1}$$

We are interested in how the measure density $\varepsilon_{b;n,\alpha,C}$, its maximum $M_{b,C}(\alpha, n)$ and the quantile ratio $\gamma_{b,C}(\alpha, n, T)$ depend on C, when all other parameters are kept fixed.

The log-generator of the cascade, $\log_m Y_{\alpha,C}$, has stable distribution $S_\alpha(\sigma, \beta = -1, \mu)$ with maximum negative skewness $\beta = -1$. The parameters σ and μ are such that

$$\log_m E[(Y_{\alpha,C})^q] = \begin{cases} \dfrac{C}{\alpha\text{-}1}(q^\alpha - q), & \text{for } \alpha \neq 1 \\ C\, q \ln(q), & \text{for } \alpha = 1 \end{cases} \tag{B.2}$$

and are found as follows from α, C and m.

Let $X \sim S_\alpha(\sigma, \beta = -1, \mu)$. Then for any a and any $b > 0$,

$$a + bX \sim \begin{cases} S_\alpha(b\sigma, \beta = \text{-}1, b\mu + a)\,, & \text{for } \alpha \neq 1 \\ S_{\alpha=1}\left(b\sigma, \beta = \text{-}1, b\mu + a + \frac{2}{\pi}\sigma b \ln(b)\right), & \text{for } \alpha = 1 \end{cases} \tag{B.3}$$

and for any $q > 0$,

$$E[e^{qX}] = \begin{cases} exp\left(q\mu - \frac{\sigma^\alpha}{cos\left(\frac{\pi\alpha}{2}\right)}q^\alpha\right), & \text{for } \alpha \neq 1 \\ exp\left(q\mu + \frac{2\sigma}{\pi}q\,\ln q\right), & \text{for } \alpha = 1 \end{cases} \tag{B.4}$$

(see for example Samorodnitsky and Taqqu, 1994, pp. 11 and 15). It follows from equation (B.3) that

$$\ln(Y_{\alpha,C}) \sim \begin{cases} S_\alpha(\ln(m)\sigma, \beta = \text{-}1, \ln(m)\mu)\,, & \text{for } \alpha \neq 1 \\ S_{\alpha=1}\left(\ln(m)\sigma, \beta = \text{-}1, \ln(m)\left(\mu + \frac{2}{\pi}\sigma\ln(\ln m)\right)\right), & \text{for } \alpha = 1 \end{cases} \tag{B.5}$$

and from equation (B.4),

$$\log_m E[(Y_{\alpha,C})^q] = \begin{cases} q\mu - \frac{(\ln m)^{\alpha\text{-}1}\,\sigma^\alpha}{cos\left(\frac{\pi\alpha}{2}\right)}q^\alpha, & \text{for } \alpha \neq 1 \\ q\mu + \frac{2\sigma}{\pi}q\,\ln[q\,\ln(m)]\,, & \text{for } \alpha = 1 \end{cases} \tag{B.6}$$

Using equations (B.2) and (B.6) one obtains,

$$\begin{aligned} &\mu = \frac{C}{1\text{-}\alpha}\,, \quad \sigma = C^{1/\alpha}\,[\ln(m)]^{(1\text{-}\alpha)/\alpha}\left(\frac{cos\left(\frac{\pi\alpha}{2}\right)}{1\text{-}\alpha}\right)^{1/\alpha}, && \text{for } \alpha \neq 1 \\ &\mu = \text{-}C\,\ln(\ln m)\,, \quad \sigma = \frac{\pi C}{2}\,, && \text{for } \alpha = 1 \end{aligned} \tag{B.7}$$

Having determined the distribution of the generator $Y_{\alpha,C}$ of log-stable cascades, we now show that a power transformation relates the generators for different C, i.e. that for any given $0 < \alpha \leq 2$ and $C_1, C_2 > 0$, constants a and $b > 0$ exist such that $Y_{\alpha,C_1} \overset{d}{=} m^a (Y_{\alpha,C_2})^b$ or equivalently

$\log_m Y_{\alpha,C_1} \overset{d}{=} a + b \log_m Y_{\alpha,C_2}$. Using equations (B.3) and (B.7), one obtains that the latter equality is satisfied for

$$\begin{aligned} a &= \frac{C_1^{1/\alpha}(C_1^{1/\alpha'} - C_2^{1/\alpha'})}{1-\alpha}, \quad b = \left(\frac{C_1}{C_2}\right)^{1/\alpha}, \quad \text{for } \alpha \neq 1 \\ a &= -C_1 \ln\left(\frac{C_1}{C_2}\right), \quad b = \left(\frac{C_1}{C_2}\right), \quad \text{for } \alpha = 1 \end{aligned} \tag{B.8}$$

where $\frac{1}{\alpha} + \frac{1}{\alpha'} = 1$. Since a power transformation of the generator induces the same power transformation of the bare cascades, we conclude that, for a and b in equation (B.8),

$$\varepsilon_{b;n,\alpha,C_1} \overset{d}{=} m^a \, (\varepsilon_{b;n,\alpha,C_2})^b \tag{B.9}$$

It follows from equation (B.9) that the extremes of log-stable bare cascades scale with the co-dimension parameter C as

$$\begin{aligned} M_{b,C_1}(\alpha, n) &\overset{d}{=} m^a \left(M_{b,C_2}(\alpha, n)\right)^b \\ M_{b,C_1}(\alpha, n, T) &= m^a \left(M_{b,C_2}(\alpha, n, T)\right)^b \\ M_{b,C_1,ind}(\alpha, n, T) &= m^a \left(M_{b,C_2,ind}(\alpha, n, T)\right)^b \\ \gamma_{b,C_1}(\alpha, n, T) &= \left(\gamma_{b,C_2}(\alpha, n, T)\right)^b \end{aligned} \tag{B.10}$$

where m is the volumetric multiplicity and the exponents a and b are given by equation (B.8). Notice that, for any given C_1/C_2 , $b = (C_1/C_2)^{1/\alpha}$ is closer to 1 for larger α. This means that the effect of dependence on the cascade extremes, which is expressed by the ratio γ, is minimum for $\alpha = 2$ (for lognormal cascades). An intuitive explanation is that, among the class of log-stable cascades, the lognormal cascades have the strongest positive singularities and hence have less dependent positive extreme values.

Another observation on equation (B.10) is that, while scaling of the maximum values involves the volumetric multiplicity m, the scaling of γ does not and therefore is the same for cascades in spaces of any dimension d.

Since averaging and power transformations do not commute, dressed measures and their maxima do not strictly satisfy the scaling relations in equations (B.9) and (B.10). However, for n not very small and T not very large, these equations provide accurate approximations also for dressed cascades; see numerical comparison in the main text.

References

Benjoudi, H, P. Hubert, D. Schertzer, and S. Lovejoy (1997), "Interpretation multifractale des courbes intensite'-duree-frequance des precipitations," *Geosciences de Surface (Hydrologie)/Surface Geosciences (Hydrology),* **325**: 323-336.

Benjoudi, H, P. Hubert, D. Schertzer, and S. Lovejoy (1999), "Multifractal Explanation of Rainfall Intensity-Duration-Frequency Curves," *Proceedings*, EGS, 24th General Assembly, The Hague, The Netherlands.

Castro, J., A. Cârsteanu, and C. Flores (2004), "Intensity-duration-area-frequency Functions for Precipitation in a Multifractal Framework," *Physica A*, **338**: 206-210.

Cates, M. E. and J. M. Deutsch (1987), "Spatial Correlations in Multifractals," *Phys. Rev. A,* **35**(11): 4907-4910.

Gupta, V.K. and Waymire, E. (1993), "A Statistical Analysis of Mesoscale Rainfall as a Random Cascade," *J. Appl. Meteorol.*, **32**(2): 251-267.

Kahane, J.-P., and J. Peyriere (1976), "Sur certaines martingales de Benoit Mandelbrot," *Adv. Math.,* **22**: 131-145.

Mandelbrot, B. B. (1989), "Multifractal Measures, Especially for the Geophysicist," *Pure Appl. Geophys.,* **131**(1/2): 5-42.

Marsan, D., D. Schertzer, and S. Lovejoy (1996), "Causal Space-time Multifractal Processes: Predictability and Forecasting of Rain Fields," *J. Geophys. Res.,* **101**(D21): 26,333-26,346.

O'Neil, J. and C. Meneveau (1993), "Spatial Correlations in Turbulence: Predictions from the Multifractal Formalism and Comparison with Experiments." *Phys. Fluid A,* **5**(1): 158-172.

Over, T. M. and V. K. Gupta (1996), "A Space-time Theory of Mesoscale Rainfall Using Random Cascades, " *J. Geophys. Res.* **101**: 26,319-26,331.

Samorodnitsky, G. and M. S. Taqqu (1994), *Stable Non-Gaussian Random Processes,* Chapman & Hall, New York.

Schertzer, D. and S. Lovejoy (1987), "Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling of Multiplicative Processes," *J. Geophys. Res.*, **92**: 9693-9714.

Veneziano, D. (1999), "Basic Properties and Characterization of Stochastically Self-Similar Processes in R^D," *Fractals*, **7**(1): 59-78.

Veneziano, D. and P. Furcolo (2002), "Multifractality of Rainfall and Intensity-duration-frequency Curves," *Wat. Resour. Res.*, **38**(12): 1306-1317.

Veneziano, D. and P. Furcolo (2003), "Marginal Distribution of Stationary Multifractal Measures and Their Haar Wavelet Coefficients," *Fractals,* **11**(3): 253-270.

Figure Captions

Figure 1: One-dimensional representation of a two-dimensional cascade with multiplicity $m = 4$.

Figure 2: Lognormal cascade with $m = 2$ and $C_1 = 0.1$. $(1/T)$-quantile of $M_{n,}$, $M_{n,T}$ for different resolutions $r = 2^n$ and return periods T (T is expressed in units of cascade realizations).

Figure 3: Lognormal cascade with $C = 0.1$. Ratio $R_{r,\,T|m} =$ in equation (4) for (*a*) $m = 4$ and (*b*) $m = 8$.

Figure 4: Distribution of the maximum M_n (solid lines) and its independent approximation (dashed lines) for a cascade with multiplicity $m = 2$ at resolutions 2^{10} and 2^{20}. The cascade generator has a lognormal distribution with co-dimension coefficient $C_1 = 0.1$: (*a*) Body of the distribution, (*b*) lower tail, (*c*) upper tail.

Figure 5: Dependence of the resolution r_Z in equation (8) on $K(2)$ the volumetric multiplicity m.

Figure 6: Distribution of the maximum $M_{n,ind}$ using the exact distribution of Z and a second-moment matching approximation. The cascade is identical to that of Figure 1 and is developed to levels $n = 0, 5, 10, 20$.

Figure 7: Ratio $\gamma(n, T)$ in equation (10) as a function of the return period T and the level n. Lognormal and beta-lognormal cascades with co-dimension parameter $C_{ln} = 0.1$ and $C_\beta = 0, 0.2$ and volumetric multiplicity $m = 2, 4, 8$

Figure 8: Comparison between the exact dressed ratios $\gamma(n, T)$ and $\gamma(n, T)$ with approximations based on equation (14) using $C = 0.05$ and 0.2, $C = 0.1$, and volumetric multiplicity $m = 2$.

Figure 9: Two nested binary partitions of a square with volumetric multiplicity $m = 2$.

Figure 10: Exclusion of long-range dependence over n_0 cascade levels. Distribution of the maximum M_n of a binary cascade at resolution $r = 2^{20}$ for $n_0 = 0, 5, 10, 15, 20$. The

cascade generator has lognormal distribution with co-dimension coefficient $C_1 = 0.1$. (*a*) Body of the distribution, (*b*) lower tail.

Figure 11: Exclusion of short-range dependence over n_0 cascade levels. Distribution of the maximum M_n of a binary cascade at resolution $r = 2^{20}$ for $n_0 = 0, 5, 10, 15, 20$. The cascade generator has lognormal distribution with co-dimension coefficient $C_1 = 0.1$. (*a*) Body of the distribution, (*b*) lower tail.

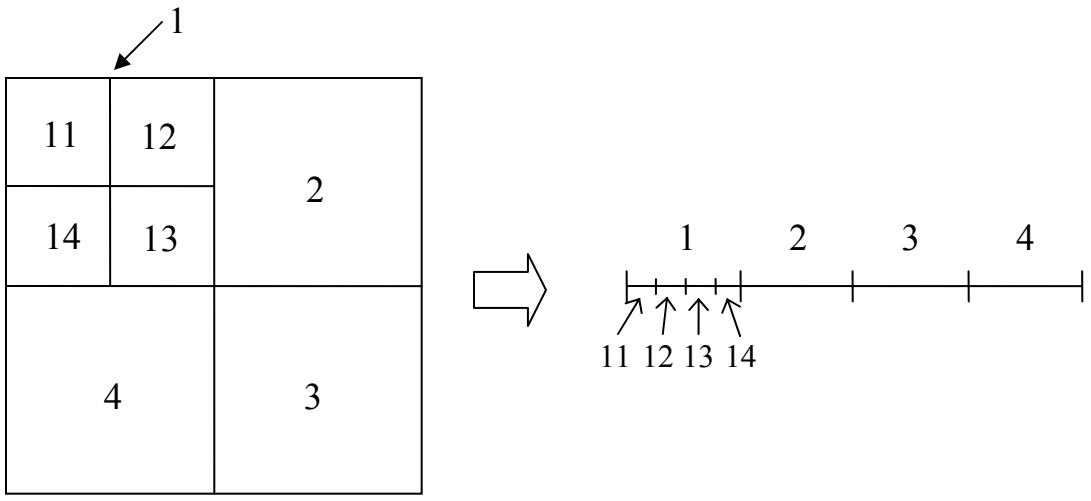


Figure 1: One-dimensional representation of a two-dimensional cascade with multiplicity $m = 4$.

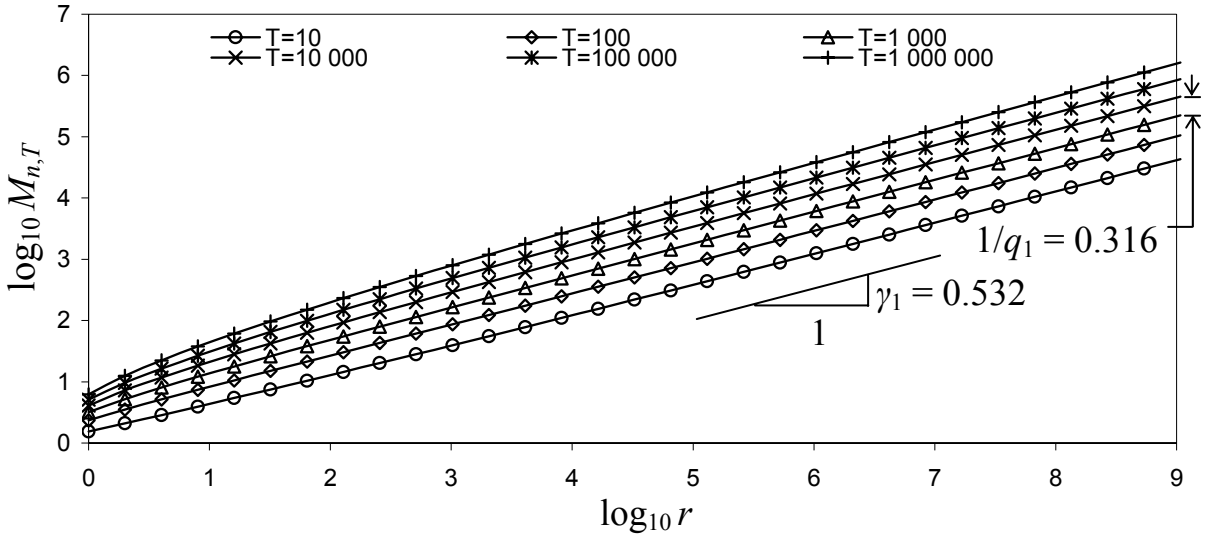


Figure 2: Lognormal cascade with $m = 2$ and $C_1 = 0.1$. $(1/T)$-quantile of $M_{n,}$, $M_{n,T}$ for different resolutions $r = 2^n$ and return periods T (T is expressed in units of cascade realizations).

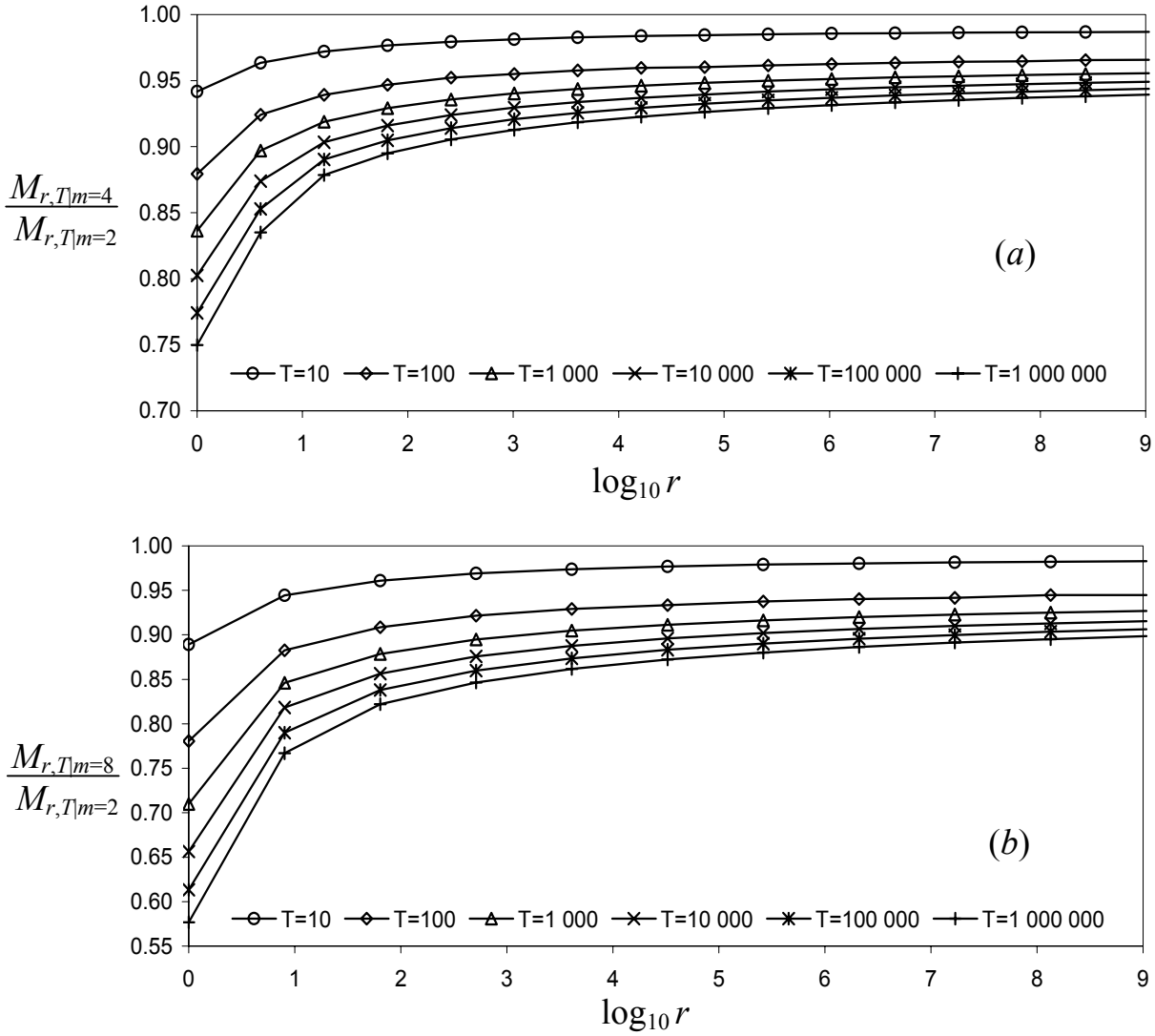


Figure 3: Lognormal cascade with $C = 0.1$. Ratio $R_{r,\,T|m} = \frac{M_{r,T|m}}{M_{r,T|m=2}}$ in equation (4) for (*a*) $m = 4$ and (*b*) $m = 8$.

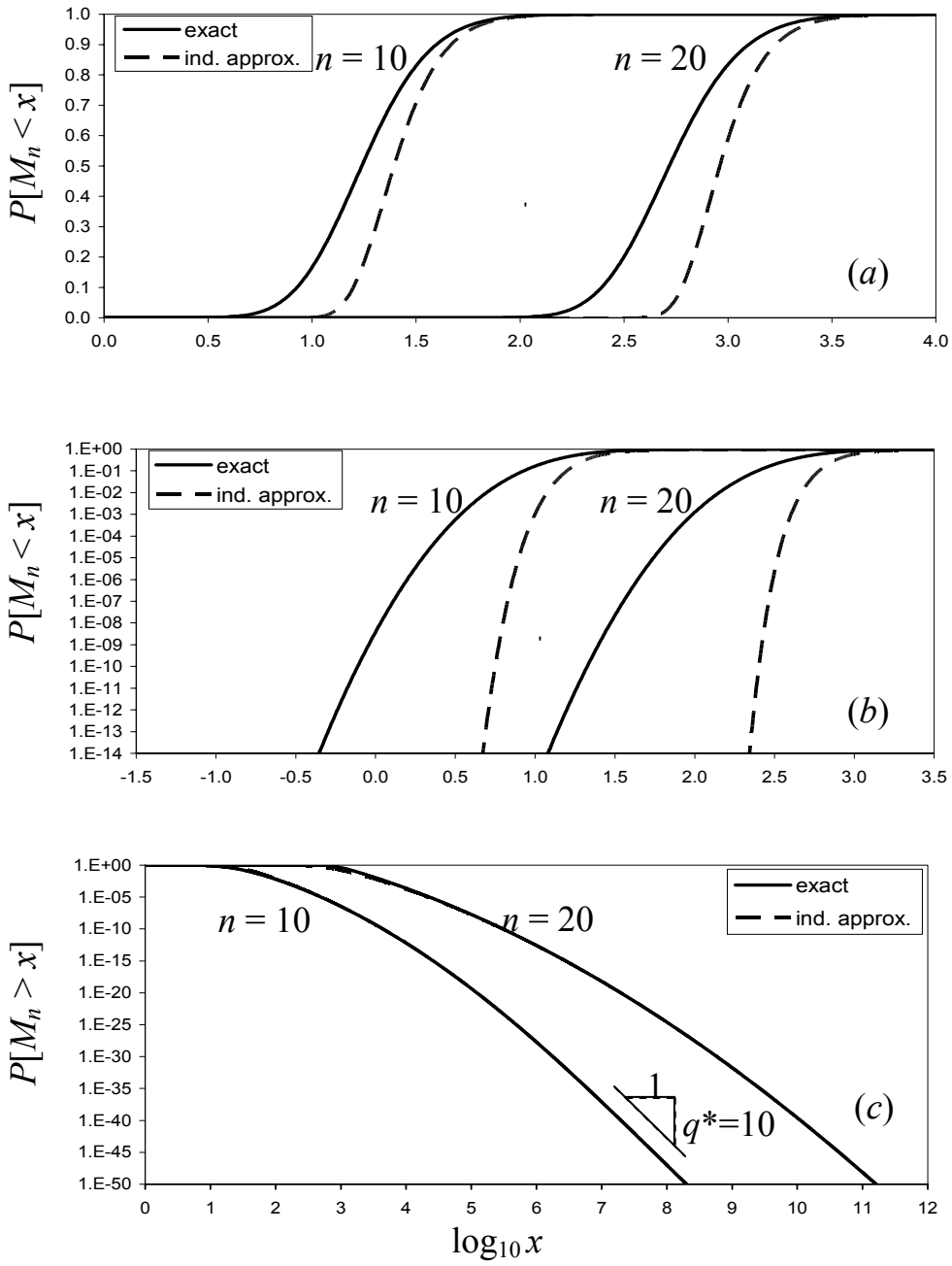


Figure 4: Distribution of the maximum M_n (solid lines) and its independent approximation (dashed lines) for a cascade with multiplicity $m = 2$ at resolutions 2^{10} and 2^{20}. The cascade generator has a lognormal distribution with co-dimension coefficient $C_1 = 0.1$: (*a*) Body of the distribution, (*b*) lower tail, (*c*) upper tail.

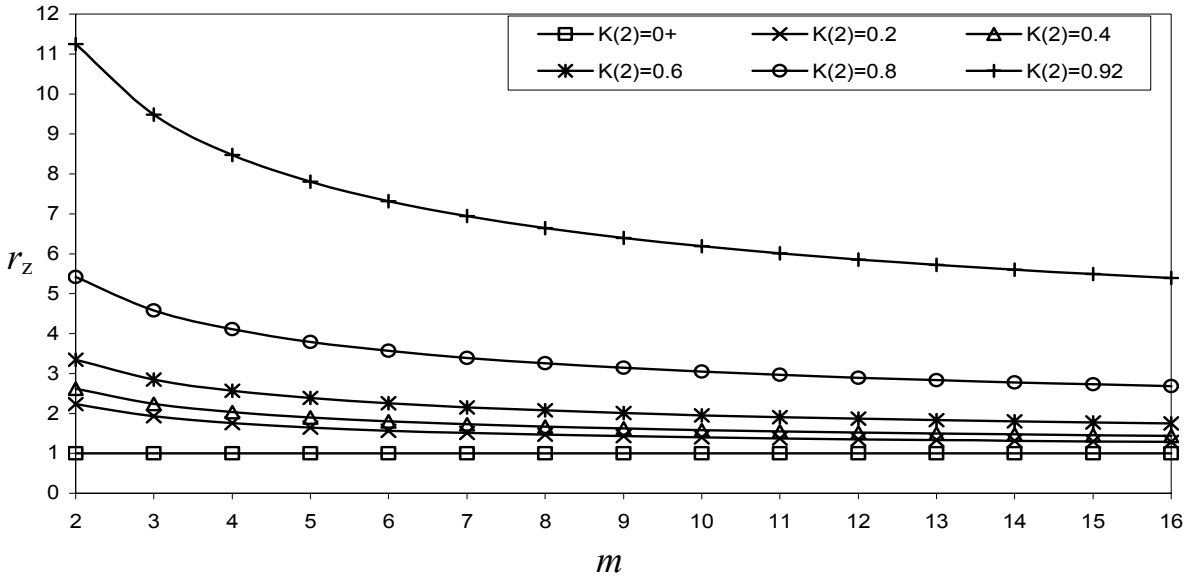


Figure 5: Dependence of the resolution r_Z in equation (8) on $K(2)$ the volumetric multiplicity m.

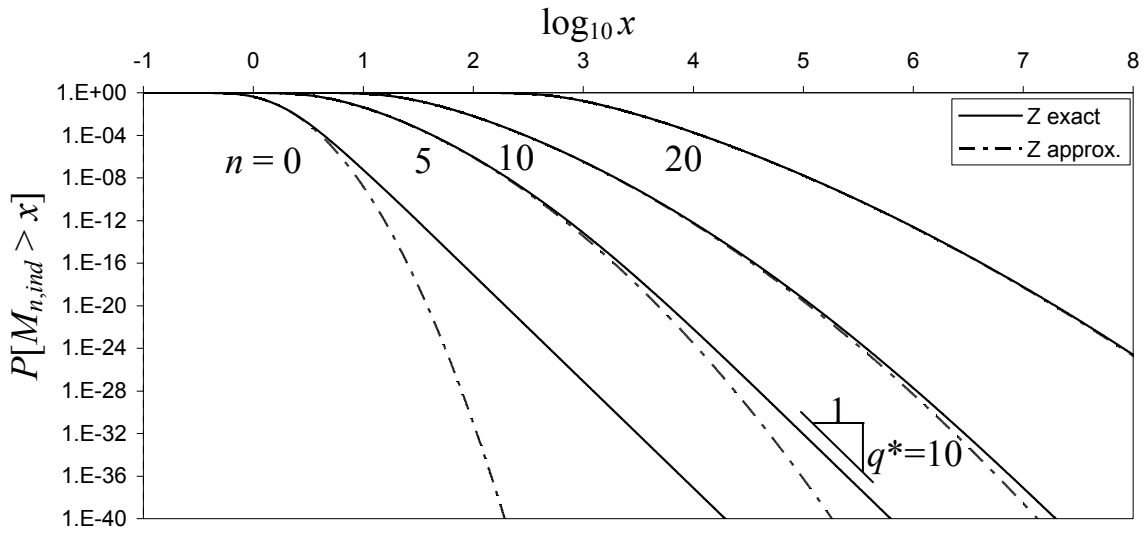


Figure 6: Distribution of the maximum $M_{n,ind}$ using the exact distribution of Z and a second-moment matching approximation. The cascade is identical to that of Figure 1 and is developed to levels n = 0, 5, 10, 20.

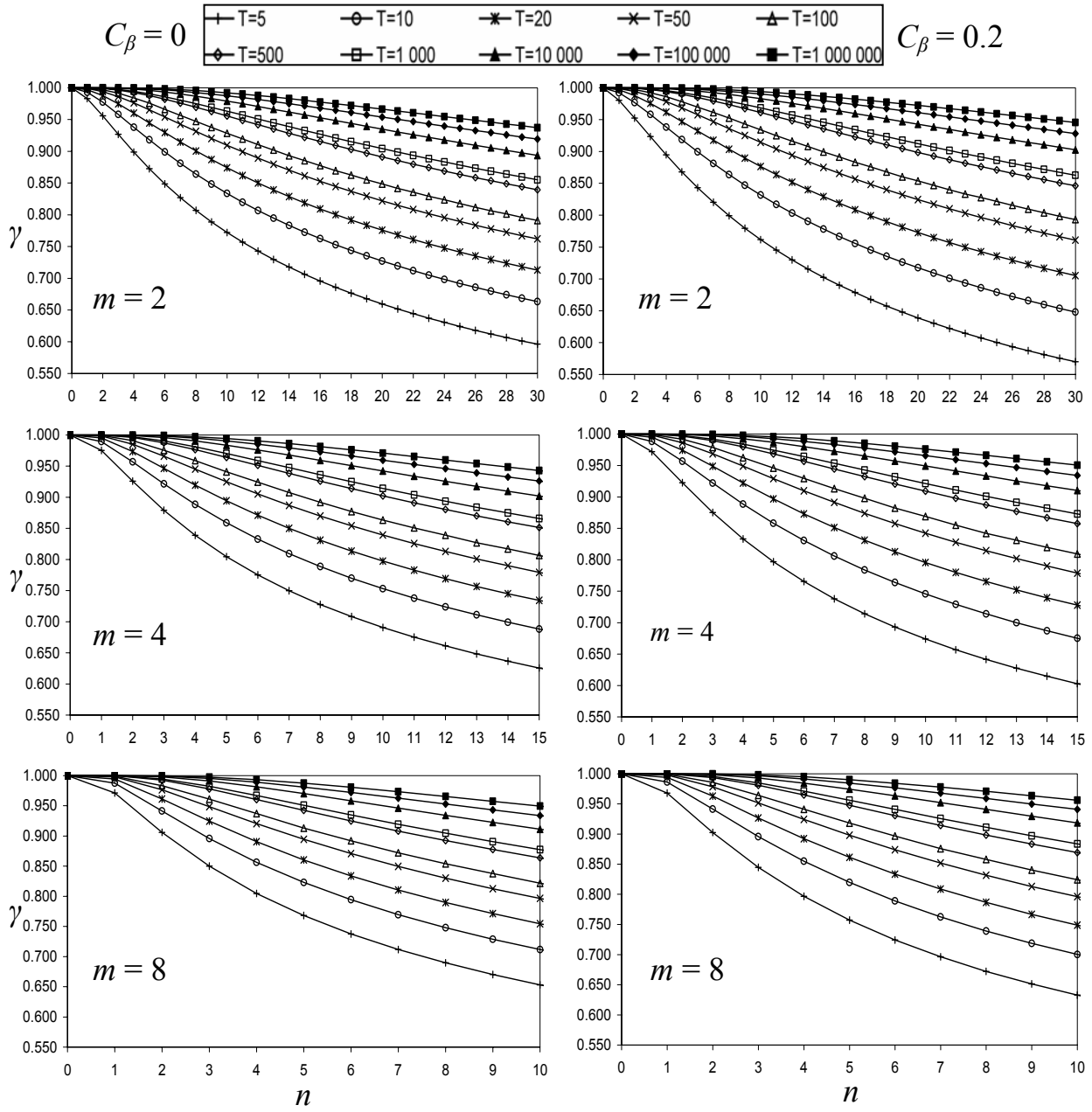


Figure 7: Ratio $\gamma(n, T)$ in equation (10) as a function of the return period T and the level n. Lognormal and beta-lognormal cascades with co-dimension parameter $C_{ln} = 0.1$ and $C_\beta = 0$, 0.2 and volumetric multiplicity $m = 2$, 4, 8.

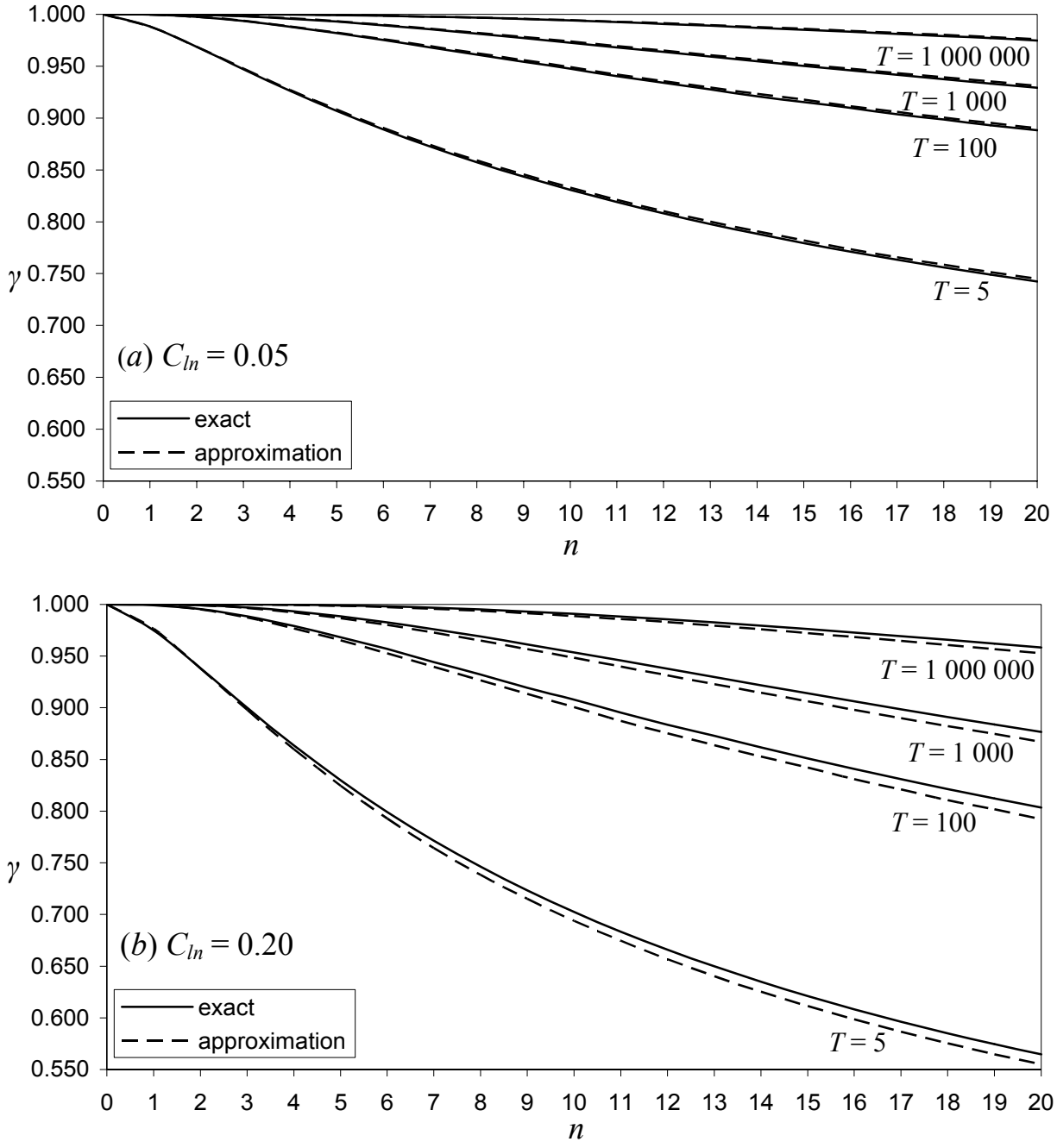
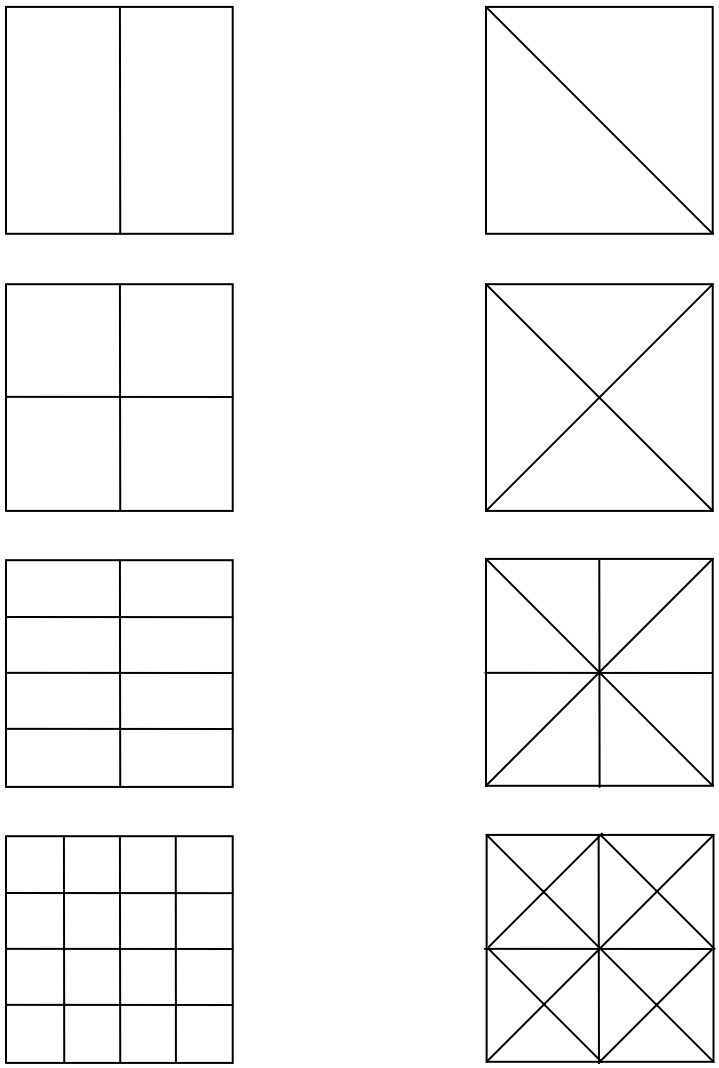


Figure 8: Comparison between the exact dressed ratios $\gamma_{C_{ln}=0.05}(n, T)$ and $\gamma_{C_{ln}=0.2}(n, T)$ with approximations based on equation (14) using $C_{ln_1} = 0.05$ and 0.2, $C_{ln_2} = 0.1$, and volumetric multiplicity $m = 2$.



(a) Rectangular cells (b) Triangular cells

Figure 9: Two nested binary partitions of a square with volumetric multiplicity $m = 2$.

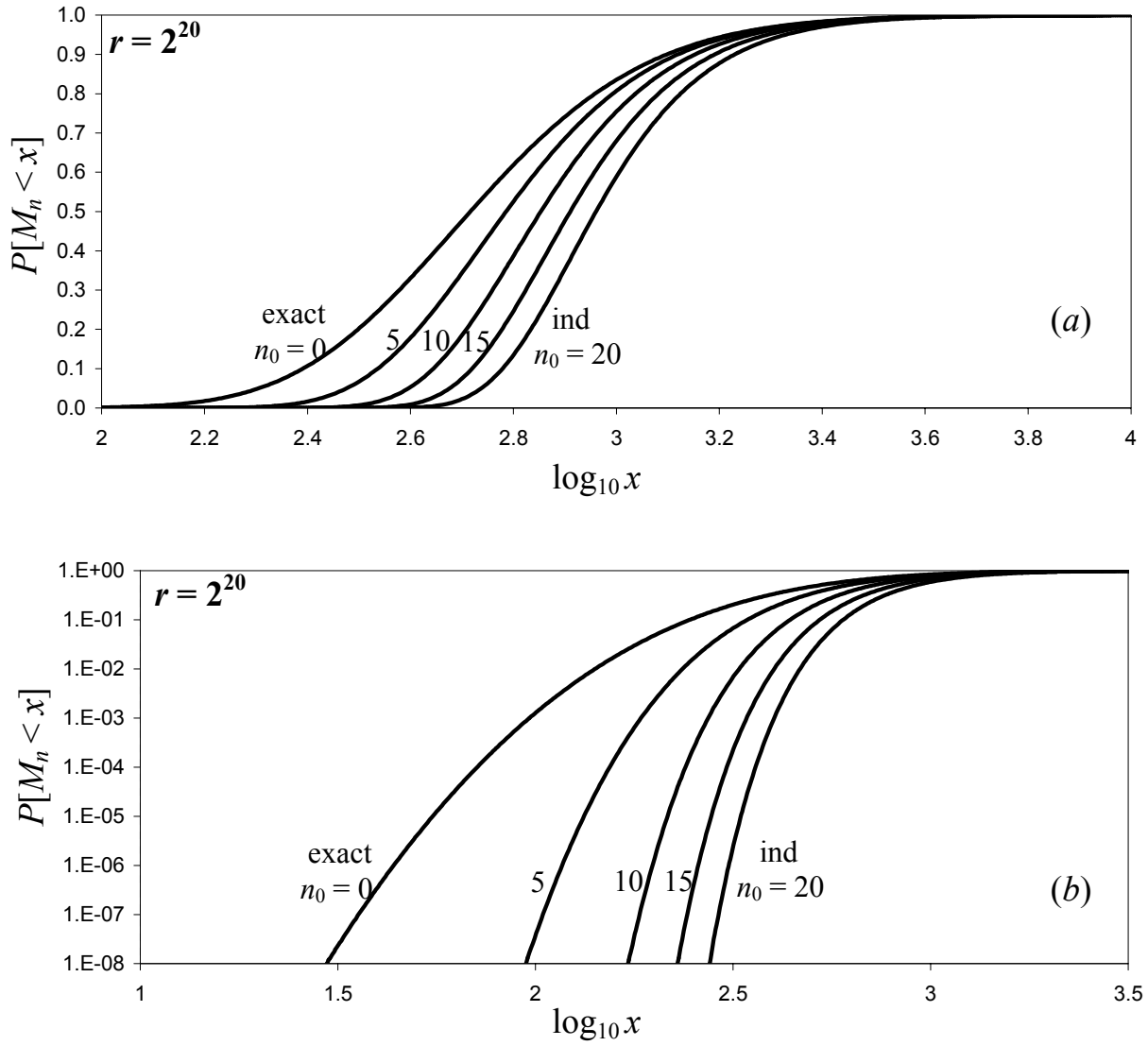


Figure 10: Exclusion of long-range dependence over n_0 cascade levels. Distribution of the maximum M_n of a binary cascade at resolution $r = 2^{20}$ for $n_0 = 0$, 5, 10, 15, 20. The cascade generator has lognormal distribution with co-dimension coefficient $C_1 = 0.1$. (*a*) Body of the distribution, (*b*) lower tail.

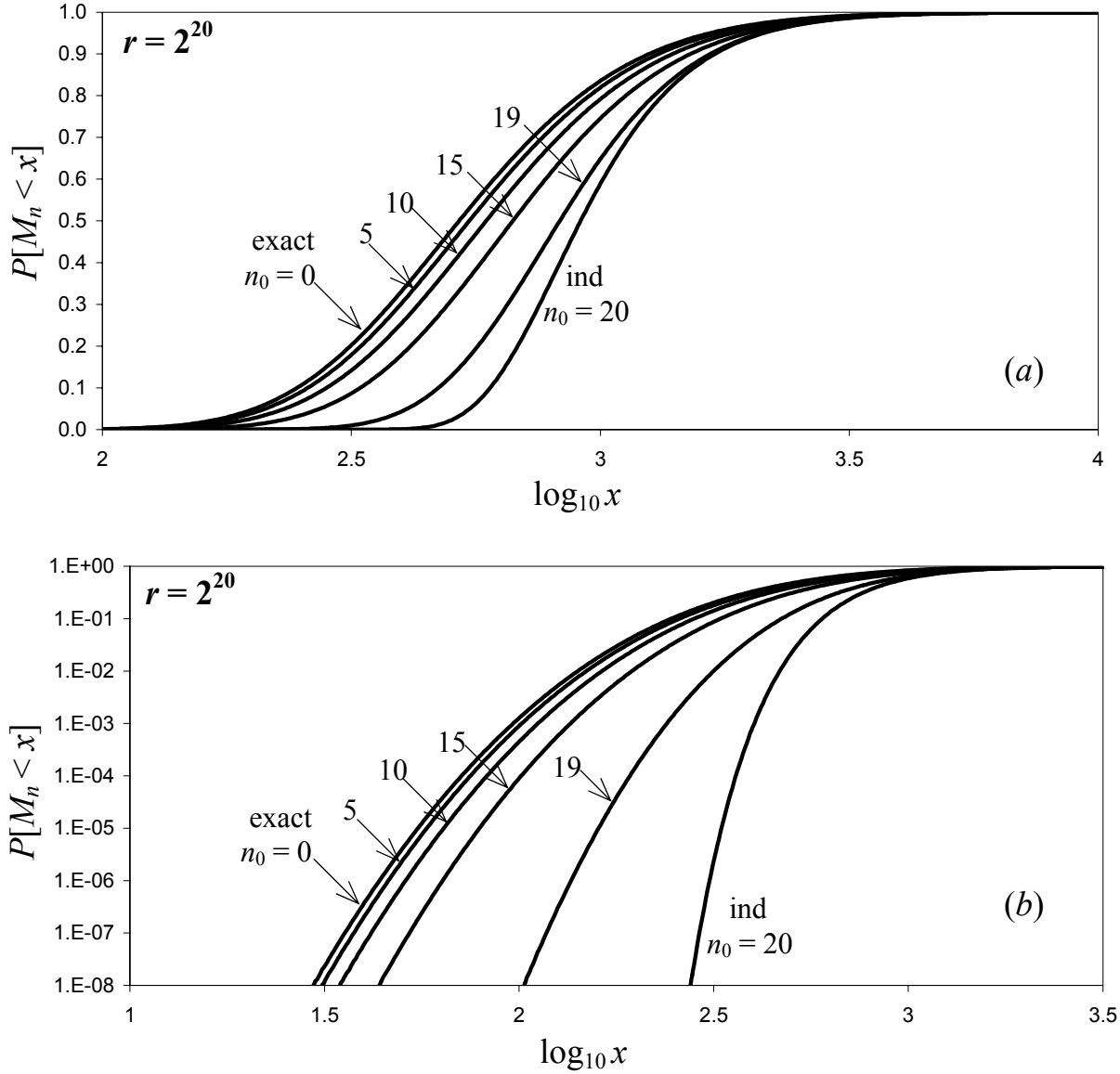


Figure 11: Exclusion of short-range dependence over n_0 cascade levels. Distribution of the maximum M_n of a binary cascade at resolution $r = 2^{20}$ for $n_0 = 0, 5, 10, 15, 20$. The cascade generator has lognormal distribution with co-dimension coefficient $C_1 = 0.1$. (*a*) Body of the distribution, (*b*) lower tail.