Large scale reduction principle
Joint works with F. Roueff, M.S. Taqqu and C. Tudor

Memory parameter of a time series

- $X = \{X_t\}_{t \in \mathbb{Z}}$: centered stationary time series with unit variance and spectral density f.
- d_X memory parameter of X (Hurvich et al. 1995) if

$$f(\lambda) \underset{\lambda=0}{\sim} |\lambda|^{-2d_X} .$$

- X : long memory process if $0 < d_X < 1/2$, short memory process if $d_X = 0$, negative memory process if $d_X < 0$.
- Extension to the case where $\Delta^K X$ stationary for $K \geq 1$ considering the generalized spectral density of X

$$f(\lambda) = |1 - e^{-\mathrm{i}\lambda}|^{-2K} f_{\Delta^K X}(\lambda) .$$

Examples

- FARIMA model : $\Delta^d X_\ell = \xi_\ell$ with Δ^d fractional differentiation operator of order $d \in (-1/2, 1/2)$ and (ξ_t) iid $\mathcal{N}(0,1)$. Stationary time series with memory parameter $d_X = d$.
- $\{B_H(k)\}_{k\in\mathbb{Z}}$ discretized version of usual FBM $\{B_H(t)\}_{t\in\mathbb{R}}$ with Hurst index $H \in (0,1)$. Memory parameter $d_{B_H} = H + 1/2$.

Main goals

- Estimation of the memory parameter of a non linear time series of the form $G(X)$, X Gaussian time series.
- Statistical properties and asymptotical behavior of the estimator.
- Application to hypothesis testing.

A wavelet based estimator

Wavelet bases

- Compactly supported MRA defined from $\varphi, \psi \in L^2(\mathbb{R})$ compactly supported.
- ψ : function admitting M vanishing moments.
- Wavelet coefficients of $F \in L^2(\mathbb{R})$

$$W_{j,k}^{(F)} = \int_{\mathbb{R}} F(t)\psi_{j,k}(t)\mathrm{d}t, \text{ with } \psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k) .$$

- Wavelet expansion of F in $L^2(\mathbb{R})$: $F = \sum_{(j,k)\in\mathbb{Z}^2} W_{j,k}^{(F)}\psi_{j,k}$.
- Case X time series ? $\mathbf{x}(t) = \sum_\ell X_\ell \varphi(t-\ell)$ and

$$W_{j,k}^{(X)} = \int_{\mathbb{R}} \mathbf{x}(t)\psi_{j,k}(t)\mathrm{d}t = \sum_\ell h_{j,2^jk-\ell} X_\ell = (h_{j,\cdot} \star X)_{2^j,k}$$

with $h_j(m) = \int_{\mathbb{R}} \phi(t+m)\psi_{j,0}(t)\mathrm{d}t$.

A wavelet based estimator

The example of FBM (Wornell et al. 1992, Bardet 2002)

- FBM case $\{B_H(t)\}$ with Hurst index H, variance of wavelet coefficients related to $d_{B_H} = H + 1/2$.
- H–self–similarity

$$\mathbb{E}[|W_{j,k}^{B_H}|^2] = C 2^{2j(H+1/2)} = C 2^{2jd_X}$$

- Gaussian or linear time series X

$$\mathbb{E}[|W_{j,k}^{X}|^2] \sim C(f^*(0), d) 2^{2jd_X} \text{ as } j \to \infty \; .$$

A wavelet based estimator

The estimator of Abry–Veicht (1998)

- $X_1, \ldots, X_N$ sample of the time series X with memory parameter d_X.
- Empirical variance of the wavelet coefficients at scale j

$$\hat{\sigma}_{N,j} = \frac{1}{n} \sum_{k=0}^{n-1} \left(W_{j,k}^{(X)} \right)^2 \ ,$$

with $n \sim N2^{-j}$ number of coefficients available at scale j.

- Expected result $\hat{\sigma}_{N,j} \sim \mathbb{E}[|W_{j,k}^X|^2] \sim C(f^*(0), \psi) 2^{2jd_X}$ as $N, j \to \infty$

A wavelet based estimator

The estimator of Abry–Veicht (1998)

- Wavelet estimator

$$\hat{d}_{N,j}(X) = \sum_{i=0}^{p} w_i \log \hat{\sigma}^2_{N,j+i}$$

with $w_0, \ldots, w_p$ s.t. $\sum_{i=0}^{p} w_i = 0$ and $\sum_{i=0}^{p} i w_i = 1/(2 \log 2)$.

- Gaussian/linear case : $\hat{\sigma}_{N,j}$ and $\widehat{d}_{N,j}$ both satisfying a CLT (Moulines–Roueff–Taqqu (2007), Roueff–Taqqu (2009)). This means that under mild assumptions

$$(N2^{-j})^{1/2}(\hat{d}_{N,j}(X) - d_X)$$

admits a Gaussian limit U_1 which can be given explicitly.

A wavelet based estimator

Beyond linear case ?

- Non linear case ? Estimation of the memory parameter using Fourier–based estimator (Dalla et al, 2006).
- Asymptotic behavior of the Abry–Veicht estimator known in the Rosenblatt case (Bardet–Tudor, 2010) using stochastic calculus.
- Extension to the general non linear case using the Abry–Veicht estimator.

Statistical properties of the wavelet–based estimator

Preliminary results

- X Gaussian centered stationary time series with memory parameter d_X, $Y = G(X)$ with G non linear function.
- Memory parameter of Y ?
- Depends on d_X and on the Hermite expansion of G

$$G = \sum_q c_q H_q \ ,$$

where $\sum_q c_q^2/q! < +\infty$, H_q q–th Hermite polynomial.

- Hermite rank of G $q_0 = \min\{q, c_q \neq 0\}$.
- Memory parameter of Y : $\delta(q_0) = d_X q_0 - (q_0 - 1)/2$ (Dalla et al. 2006).

Statistical properties of the wavelet–based estimator

A consistency result

We apply wavelet–based estimation to $Y = G(X)$

Theorem (Clausel et al., 2015) General case $Y = G(X)$

Let (j_N) increasing sequence s.t. $\lim_{N\to\infty} N2^{-j_N} = \infty$. Suppose that $M \geq K + \delta(q_0)$. Then, as $N \to \infty$,

$$\hat{d}_{N,j_N}(Y) \overset{(P)}{\to} d_Y = \delta(q_0) \,.$$

Asymptotical properties of the wavelet–based estimator

Some questions

- Consistency not sufficient in view of statistical applications as hypothesis testing.
- Convergence rate and asymptotical behavior of the estimator ?
- Reduction principle true for the wavelet coefficients (Clausel et al., 2012). For $Y = G(X)$ with $G = c_{q_0} H_{q_0}/q_0! + \cdots$

$$W_{j,k}^{(Y)} \approx W_{j,k}^{(c_{q_0} H_{q_0}(X)/q_0!)}$$

- Does $\hat{d}_{N,j}(Y)$ satisfy the reduction principle ? If such the case, for $Y = G(X)$ with

$$G = c_{q_0} H_{q_0}/q_0! + \cdots$$

  then $\hat{d}_{N,j}(Y)$ behaves as $\hat{d}_{N,j}(c_{q_0} H_{q_0}(X)/q_0!)$ as $N, j \to \infty$
- Behavior of $\hat{d}_{N,j}(c_{q_0} H_{q_0}(X)/q_0!)$?

Asymptotical properties of the wavelet–based estimator

The special case $G = c_{q_0} H_{q_0}(X)/q_0!$, $q_0 \geq 2$

Theorem (Clausel et al., 2014) Case $Y = H_{q_0}(X)/q_0!$, $q_0 \geq 2$

Assume $M \geq K + \delta(q_0)$. As $N \to \infty$, if $j = j(N)$ is such that $j \to \infty$ and $N2^{-j} \to \infty$, then

$$\hat{d}_{N,j}(Y) = d_Y + (N2^{-j})^{2d_X-1} O_P(1) + O\left(2^{-\tilde{\beta} j}\right) .$$

where $\tilde{\beta}$ is related to the smoothness at 0 of $f^*(\lambda) = |\lambda|^{2d_X} f(\lambda)$. Moreover the O_P-term converges in distribution to a Rosenblatt variable U_2.

Asymptotical properties of the wavelet–based estimator

Hint of the proof (1)

- Harmonizable representation of X

$$X_\ell = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda\ell} f^{1/2}(\lambda) \mathrm{d}\widehat{W}(\lambda) ,$$

- $Y_\ell = c_{q_0} H_{q_0}(X_\ell)/q_0!$ multiple stochastic integral of order q_0.
- $W_{j,k}^{(Y)}$ linear in Y : multiple stochastic integral of order q_0.
- Centered empirical variance ? Need to estimate

$$\frac{1}{n}\sum_{k=0}^{n-1}[W_{j,k}^{(Y)}]^2 - \mathbb{E}\left[\frac{1}{n}\sum_{k=0}^{n-1}[W_{j,k}^{(Y)}]^2\right]$$

Asymptotical properties of the wavelet–based estimator

Hint of the proof (2)

- Product formula for multiple stochastic integrals applied to the multiple stochastic integral $W_{j,k}^{(Y)}$
  $\Rightarrow$ decomposition into Wiener chaos of the centered empirical variance
  $$\hat{\sigma}_{N,j} - \mathbb{E}[\hat{\sigma}_{N,j}] = \sum_{q=1}^{q_0} I_{N,j}^{(2q)}$$
  with I_{2q} multiple integrals of order $2q$.
- Dominating term $I_{N,j}^{(2)}$: Rosenblatt variable whose asymptotic variance is known
  $\Rightarrow$ asymptotical behavior of the empirical variance and the estimator of the memory parameter using the delta method.

Asymptotical properties of the wavelet–based estimator

A reduction principle ?

- End of the story ?
  - We know the asymptotic limit of the estimator if $G = H_{q_0}$ in the two cases $q_0 = 1$ (Gaussian) and $q_0 \geq 2$ (Rosenblatt).
  - If $G = c_{q_0} H_{q_0}/q_0! + \cdots$ and reduction principle true $d_{N,j}(Y) \approx d_{N,j}(c_{q_0} H_{q_0}(X)/q_0!)$....
- Unfortunately not so simple (Abry et al. 2011) ! !
- The reduction principle may not hold....

Asymptotical properties of the wavelet–based estimator

A counterexample for the reduction principle

- Case $Y = G(X)$ with $G = H_{q_0} + H_{q_0+1}$, $q_0 \geq 2$.
- $W_{j,k}^{(Y)} = W_{j,k}^{(q_0)} + W_{j,k}^{(q_0+1)}$ with $W_{j,k}^{(q)}$ in the q–th Wiener chaos for $q = q_0, q_0 + 1$.
- Product formula

$$[W_{j,k}^{(Y)}]^2 = [W_{j,k}^{(q_0)}]^2 + [W_{j,k}^{(q_0+1)}]^2 + 2W_{j,k}^{(q_0)} W_{j,k}^{(q_0+1)}$$

- If reduction principle true

$$\frac{1}{n}\sum_{k=0}^{n-1}[W_{j,k}^{(Y)}]^2 \approx \frac{1}{n}\sum_{k=0}^{n-1}[W_{j,k}^{(q_0)}]^2$$

Statistical properties of the wavelet–based estimator

A counterexample for the reduction principle

- If $N \ll 2^{2j}$, the sum

$$1/n\left(\sum_{k=0}^{n-1}[W_{j,k}^{(q_0)}]^2\right)$$

is dominating in the empirical variance and the reduction principle holds.

- Unfortunately, if $2^{2j} \ll N$, the sum

$$1/n\left(\sum_{k=0}^{n-1} W_{j,k}^{(q_0)} W_{j,k}^{(q_0+1)}\right)$$

is dominating and the reduction principle does not hold ! !

- Extension to the case $G = H_{q_0} + H_{q_{\ell_0}} + H_{q_{\ell_0}+1}$. The reduction principle holds or not depending whether $N \ll 2^{j(\nu+1)}$ or $2^{j(\nu+1)} \ll N$ with $\nu = 2q_{\ell_0} + 1 - 2q_0$.

Statistical properties of the wavelet–based estimator

A counterexample for the reduction principle

- General case (Clausel et al. 2013) . Maybe complicated : limit may be Gaussian, Rosenblatt or Hermite and the rate of convergence of the estimator can be different !
- Depends on the value of ν such that $N \sim 2^{j(\nu+1)}$ with respect to some critical indices depending on the whole function G....
- Sufficient conditions for the reduction principle to hold ?

Statistical properties of the wavelet–based estimator

Sufficient conditions for the reduction principle

Theorem

Assume $M \geq K + \delta(q_0)$. There exists some critical index ν_c which can be defined explicitly and depends only on G, d_X, such that if $N \ll 2^{j(\nu_c+1)}$, the reduction principle holds as $j, N \to \infty$.

In some cases this critical index is very simple.

- G even : $\nu_c = \infty$.
- $q_0 \geq 2$ and there is two LRD terms in the Hermite expansion of G

$$\nu_c = 1 + 2(q_{\ell_0} - 2q_0)$$

Statistical properties of the wavelet–based estimator

Application to hypothesis testing

- Definition of a statistical test procedure which applies to a general G.
- Let d_0^* : possible value for the true unknown memory parameter d_Y of Y.
- Hypotheses

$$H_0 : d_Y = d_0^* \quad \text{against} \quad H_1 : d_Y \in (0, \bar{K} + 1/2) \setminus \{d_0^*\}.$$

Statistical properties of the wavelet–based estimator

Application to hypothesis testing

- $\alpha \in (0, 1)$ be a significance level.
- Statistical test

$$\delta_s = \begin{cases} 1 & \text{if } |\hat{d}_0 - d_0^*| > s_N(\alpha), \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

where $s_N(\alpha)$ is the $(1 - \alpha/2)$ quantile of $U_1/(N2^{-j})^{1/2}$ or $U_1/(N2^{-j})^{1/2}$ depending on the Hermite rank of G.

Statistical properties of the wavelet–based estimator

Application to hypothesis testing

The constant ζ is a constant depending on the behavior of $f(\lambda)|\lambda|^{2d}$ at $\lambda = 0$.

Theorem

Let $j = (j_N)$ s.t. $N2^{-j} \to \infty$ holds, $M \geq K + \delta(q_0)$. Suppose moreover that, as $N \to \infty$,

$$N2^{-j} \ll 2^{j\nu_c^*},$$

and that

$$2^{-\zeta j} \ll u_N^{-1},$$

with $u_N = (N2^{-j})^{1/2}$ if $q_0 = 1$, $u_N = (N2^{-j})^{1-2d_X}$ otherwise. Then, the test δ_s is a consistent test with asymptotic significance level α.

Numerical experiments

X_t Gaussian ARFIMA(0,d,0). Tests with two models

- Model 1 : $Y_t = H_1(X_t) + 1/(2\sqrt{3})H_3(X_t)$: $q_0 = 1$, $d_0 = d$ and $\nu_c = (1-2d)/(2d-1/2)$
- Model 2 : $Y_t = 1/\sqrt{2}H_2(X_t) + 1/(2\sqrt{3})H_3(X_t)$, $q_0 = 2$, $d_0 = 2d - 1/2$ and $\nu_c = 1$

Monte–Carlo simulations involving 1000 independent draws of samples.

Numerical experiments

Performances of the regression estimator

	bias	std	MSE
d=0.3	-0.0338	0.0402	0.0028
d=0.325	-0.0878	0.1112	0.0201
d=0.35	-0.0368	0.0425	0.0032
d=0.375	-0.0363	0.0619	0.0051
d=0.4	-0.0513	0.1289	0.0192

Table: Model 1, $N = 2^{15}$.

Numerical experiments

Performances of the regression estimator

	d_0	bias	std	MSE
d=0.35	0.2	-0.0302	0.0811	0.0075
d=0.375	0.25	-0.0722	0.1026	0.0157
d=0.4	0.3	-0.0462	0.0891	0.0101
d=0.425	0.35	-0.0455	0.0831	0.0090
d=0.45	0.4	-0.0409	0.0856	0.0090

TABLE: Model 2, $N = 2^{15}$.

Numerical experiments

Finite sample performances of the test

d_0^*	0.3	0.325	0.35	0.375	0.4
$\alpha = 0.01$	0.0730	0.0700	0.0930	0.0730	0.0590
$\alpha = 0.05$	0.1780	0.1840	0.1830	0.1590	0.1530
$\alpha = 0.1$	0.2630	0.2460	0.2620	0.2390	0.2150

TABLE: Rejection rates under H_0 for different values of d_0^* and α for Model 1 with $N = 2^{15}$.

Numerical experiments

Finite sample performances of the test

In case of Model 2, asymptotic limit may be difficult to deal with !
$\Rightarrow$ bootstrap–like strategy

1. Pick m sub-samples of the original time series of 2^{N-L} consecutive observations, randomly with replacement.
2. For $\ell = 1, \dots, m$, compute an estimator $\hat{d}_0(\ell)$ based on the ℓth sub–sample (with the same j and weights w_i as for $\hat{d}_0$).
3. Compute the empirical variance $\hat{v}_L$ of the sample $\hat{d}_0(\ell)$, $\ell = 1, \dots, m$ obtained in previous step and set the empirical variance of the full sample estimate $\hat{d}_0$ to $\hat{v} = 2^{-L(1-2d_0^*)}\hat{v}_L$.

Numerical experiments

Finite sample performances of the test

d_0^*	0.15	0.2	0.25	0.3	0.35	0.4
$\alpha = 0.01$	0.4210	0.3230	0.3860	0.4470	0.4250	0.4250
$\alpha = 0.05$	0.5110	0.4280	0.4770	0.4840	0.4580	0.4450
$\alpha = 0.1$	0.5760	0.4970	0.5290	0.5430	0.4980	0.4820

TABLE: Rejection rates under H_0 for different values of d_0^* and different levels α for Model 2 ($N = 2^{15}$).

Numerical experiments

Finite sample performances of the test

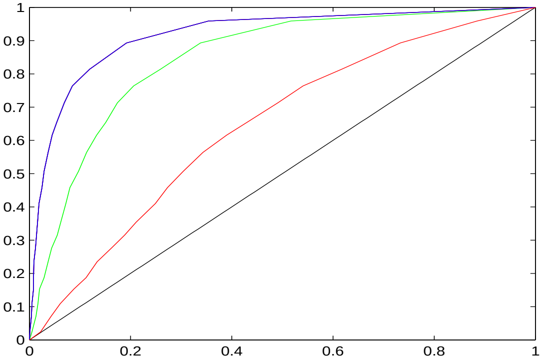

FIGURE: ROC curves for Model 1 and $d_0^* = 0.4$ for three data sets : $d_0 = 0.3$ (blue top curve), $d_0 = 0.325$ (green middle curve), $d_0 = 0.35$ (red bottom curve). $N = 2^{15}$.

Numerical experiments

Finite sample performances of the test

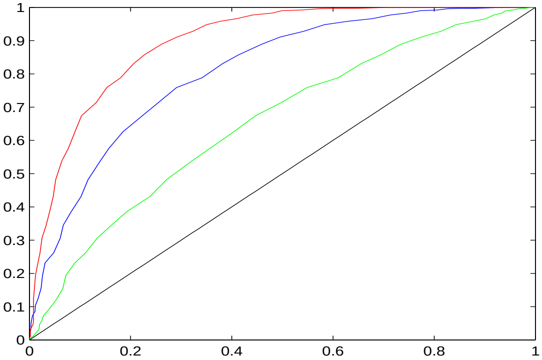

FIGURE: ROC curves for Model 2 and $d_0^* = 0.3$ for three data sets : $d_0 = 0.15$ (red top curve), $d_0 = 0.2$ (blue middle curve), $d_0 = 0.25$ (green bottom curve). $N = 2^{15}$.

Bibliography

- G.W.Wornell, A.V.Oppenheim, Estimation of fractal signals from noisy measurements using wavelets. IEEE SP. (1992).
- C.M.Hurvich, B.K. Ray, Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes. JTSA (1995).
- P. Abry, D. Veitch, Wavelet analysis of long-range-dependent traffic. IEEE TIT (1998).
- J.M.Bardet, Statistical study of the wavelet analysis of fractional Brownian motion. IEEE TIT (2002).
- V.Dalla, L.Giraitis, J. Hidalgo, Consistent estimation of the memory parameter for nonlinear time series. JTSA (2006).

Bibliography

- E. Moulines, F. Roueff, M. S. Taqqu. On the spectral density of the wavelet coefficients of long memory time series with application to the log–regression estimation of the memory parameter. JTSA (2007).
- F. Roueff, M. S. Taqqu. Central limit theorems for arrays of decimated linear processes. SPA (2009).
- J.M. Bardet, C. A. Tudor. A wavelet analysis of the Rosenblatt process : chaos expansion and estimation of the self-similarity parameter. SPA (2010)
- P.Abry, H.Helgason, V. Pipiras. Wavelet–based analysis of non-Gaussian long–range dependent processes and estimation of the Hurst parameter. Lith. Math. J. (2011).

Bibliography

- M. Clausel, F. Roueff, M.S. Taqqu, C. Tudor, Large scale behavior of wavelet coefficients of non–linear subordinated processes with long memory. ACHA (2012).
- M. Clausel, F. Roueff, M.S. Taqqu, C. Tudor, High order chaotic limits of wavelet scalograms under long–range dependence. ALEA (2013).
- M. Clausel, F. Roueff, M.S. Taqqu, C. Tudor, Wavelet estimation of the long memory parameter for Hermite polynomial of Gaussian processes. ESAIM P.S.(2014)
- M. Clausel, F. Roueff, M.S. Taqqu, Large scale reduction principle and application to hypothesis testing. EJS (2015)

Critical exponent

The critical exponent is

$$
\nu_c = \begin{cases}
\infty, \ \text{if } \mathcal{L} = \{0\} \text{ or if } q_0 = 1,\ d \leq 1/4 \text{ and } I_0 = \emptyset, \\[2ex]
\frac{d+1/2-2\delta_+(q_{\ell_0})}{d}, \ \text{if } q_0 = 1,\ d \leq 1/4 \text{ and } I_0 \neq \emptyset, \\[2ex]
\frac{1-2\delta_+(q_1-1)}{2d-1/2}, \ \text{if } q_0 = 1,\ d > 1/4,\ 1 \in \mathcal{L} \text{ and } J_d = \emptyset, \\[2ex]
\min\left(\frac{1-2\delta_+(q_1-1)}{2d-1/2}, \frac{2d+1/2-2\delta_+(q_{\ell_r})-\delta(r+1)}{\delta(r+1)} : r \in \mathcal{I}_r\right), \\
\qquad \text{if } q_0 = 1,\ d > 1/4 \text{ and } J_d \neq \emptyset, \\[2ex]
\infty, \ \text{if } q_0 \geq 2 \text{ and } I_0 = \emptyset, \\[2ex]
1 + \frac{4(\delta(q_0)-\delta_+(q_{\ell_0}))}{1-2d}, \ \text{if } q_0 \geq 2 \text{ and } I_0 \neq \emptyset.
\end{cases}
$$

Numerical experiments

Finite sample performances of the test

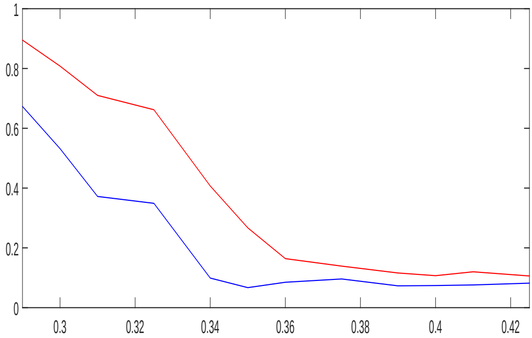

FIGURE: Rejection rates as a function of d_0 for two data sets : $(X_t)_{1\leq t\leq N}$ (blue bottom curve), Model 1 (red top curve), $d_0^* = 0.4$, $N = 2^{15}$.

Numerical experiments

Finite sample performances of the test

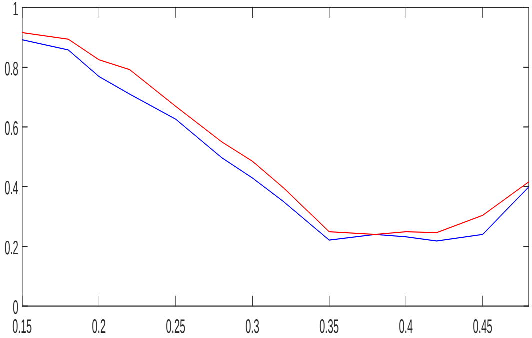

FIGURE: Rejection rates as a function of d_0 for two data sets : $(H_2(X_t))_{1\leq t\leq N}$ (blue bottom curve), Model 2 (red top curve), $d_0^* = 0.3$, $N = 2^{15}$.