SUPPLEMENTARY for Exploring the departure of Autocorrelation Functions from normality and its consequences in MA(q) modeling.

M. Royer-Carenzi[a,⋆], H. Hassani[b]

[a]Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France

[b]The Research Institute of Energy Management and Planning (RIEMP), University of Tehran, Tehran 19395-4697, Iran.

⋆CONTACT: M. Royer-Carenzi. Email: manuela.royer-carenzi@univ-amu.fr

September 30, 2024

Contents

1 Results for simulated MA(2) processes with $a_2 = \frac{1}{10}$

In our study, we investigate the normality of the sample ACF, and of the SACF . We simulate $N_S = 5\,000$ MA(2) processes with equation

$$Z_t \;=\; \mathcal{E}_t \,+\, \frac{1}{2}\mathcal{E}_{t-1} \,+\, a_2\,\mathcal{E}_{t-2}\,,$$

where $(\mathcal{E}_t)_t$ is either a Gaussian- or an Exponential-WN, with length $n = 500$. In the main paper, in Section 3, we simulate MA(2) processes with coefficient $a_2 = \frac{1}{2}$. Here, we consider $a_2 = \frac{1}{10}$.

1.1 Check for the normality of $\hat{\Xi}(h)$ at a fixed lag h

At a given lag h, using Kolmogorov-Smirnov test, we test for the adequacy of $\hat{\Xi}(h)$ with the Gaussian distribution $\mathcal{N}\left(\rho(h),\, \frac{v_{h,h}}{n}\right)$, and we also merely test for its normality using Shapiro-Wilk. Figure S1 displays the p-values provided by the Kolmogorov-Smirnov test, when applied to either $N_S = 200$ or $N_S = 5\,000$ simulations of MA(2) processes with $a_2 = \frac{1}{10}$, and with length $n = 500$, whereas Figure S2 gives the p-values provided by Shapiro-Wilk. Note that, in the main paper, in Section 3, in Figures 1 and 2, the same tests have also been led on MA(2) processes with $a_2 = \frac{1}{2}$, and the results were very similar.

The results in Figure S1, top line, show that $\hat{\Xi}(h)$ behaves roughly like a Gaussian distribution $\mathcal{N}\left(\rho(h),\, \frac{w_{h,h}}{n}\right)$, as it was expected from Theorem 2.2, even for a lag h much greater than $\sqrt{n}$. But the graphics in the bottom line show that $\hat{\Xi}(h)$ deviates from this specific normal distribution when we lead the normality tests on more simulations. We recall that Theorem 2.2provides an asymptotic result, so that $\hat{\Xi}(h)$ approximately follow the $\mathcal{N}\left(\rho(h),\, \frac{v_{h,h}}{n}\right)$ distribution. Its slight departure from this distribution is more easily detectable with a great number of simulations, and is even more pronounced for Exponential-WN, meaning that $\hat{\Xi}(h)$ distribution may deviate from the $\mathcal{N}\left(\rho(h),\, \frac{v_{h,h}}{n}\right)$ theoretical distribution, with a greater extent.

Nevertheless, if we focus on the Gaussian behavior of the ACF estimators $\hat{\Xi}(h)$, Figure S2 shows that the normality behavior is fairly strong, especially when the underlying WN is Gaussian, since Shapiro nearly never rejects the normality hypothesis, whatever the lag $h < n - 2$, even with a great number of simulations $N_S = 5\,000$. But for an underlying Exponential-WN, the normality property is quickly lost when the number of simulations increases.

1.2 Normality of Sacf

At a given lag H, we test for the normality of the N_S values of $S_{ACF}(H)$. Figure S3 displays the p-values provided by the Shapiro-Wilk test, when applied to either $N_S = 200$ or $N_S = 5\,000$ simulations of MA(2) processes with $a_2 = \frac{1}{10}$, and with length $n = 500$. Note that, in the main paper, in Section 3, in Figure 3 , the same tests have also been led on MA(2) processes with $a_2 = \frac{1}{2}$, and the results were very similar.

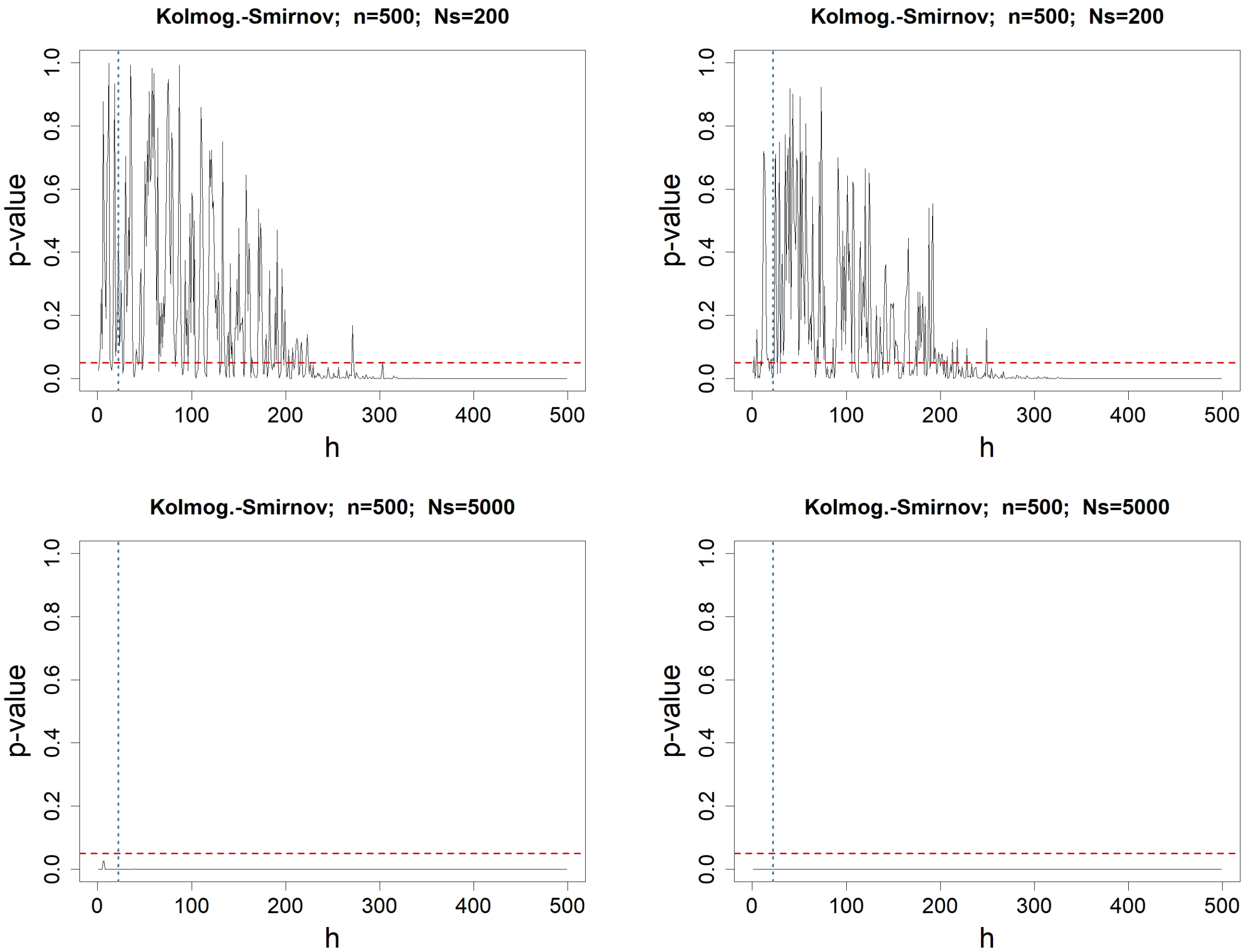

Figure S1: P-values when testing for the adequacy of the N_S values of $\hat{\rho}(h)$ with $\mathcal{N}\left(\rho(h),\, \frac{v_{h,h}}{n}\right)$, for any fixed lag h varying from 1 to $n-1$. The involved normality test is Kolmogorov-Smirnov's. The left column concerns MA(2) driven by a Gaussian WN whereas the right one deals with Exponential WN process. The length of the simulated WN process is $n = 500$. In the upper figures, the number of simulated MA(2) processes is $N_S = 200$, whereas it is $N_S = 5\,000$ in the bottom. The red-dotted horizontal line represents 5%, while the blue-dotted vertical line represents $h = \sqrt{n}$.

In Figure S3, we observe that the sum of sample ACF departs from normality for almost all the lags H, except maybe for the first lags H when the underlying white noise is Gaussian, and/or when the number of simulations remains low. Of course, Kolmogorov-Smirnov test confirmed the departure of $S_{ACF}(H)$ from the theoretical distribution $\mathcal{N}\left(\sum_{h=1}^{H} \rho(h),\, \frac{w_{H,H}}{n}\right)$ at any lag H, whatever the nature of the underlying WN (Gaussian or Exponential) and even for a low number of simulations (results not shown). The departure of $S_{ACF}(H)$ from $\mathcal{N}\left(\sum_{h=1}^{H} \rho(h),\, \frac{w_{H,H}}{n}\right)$ can be explained by the previous finding that $\hat{\Xi}(h)$ do not perfectly

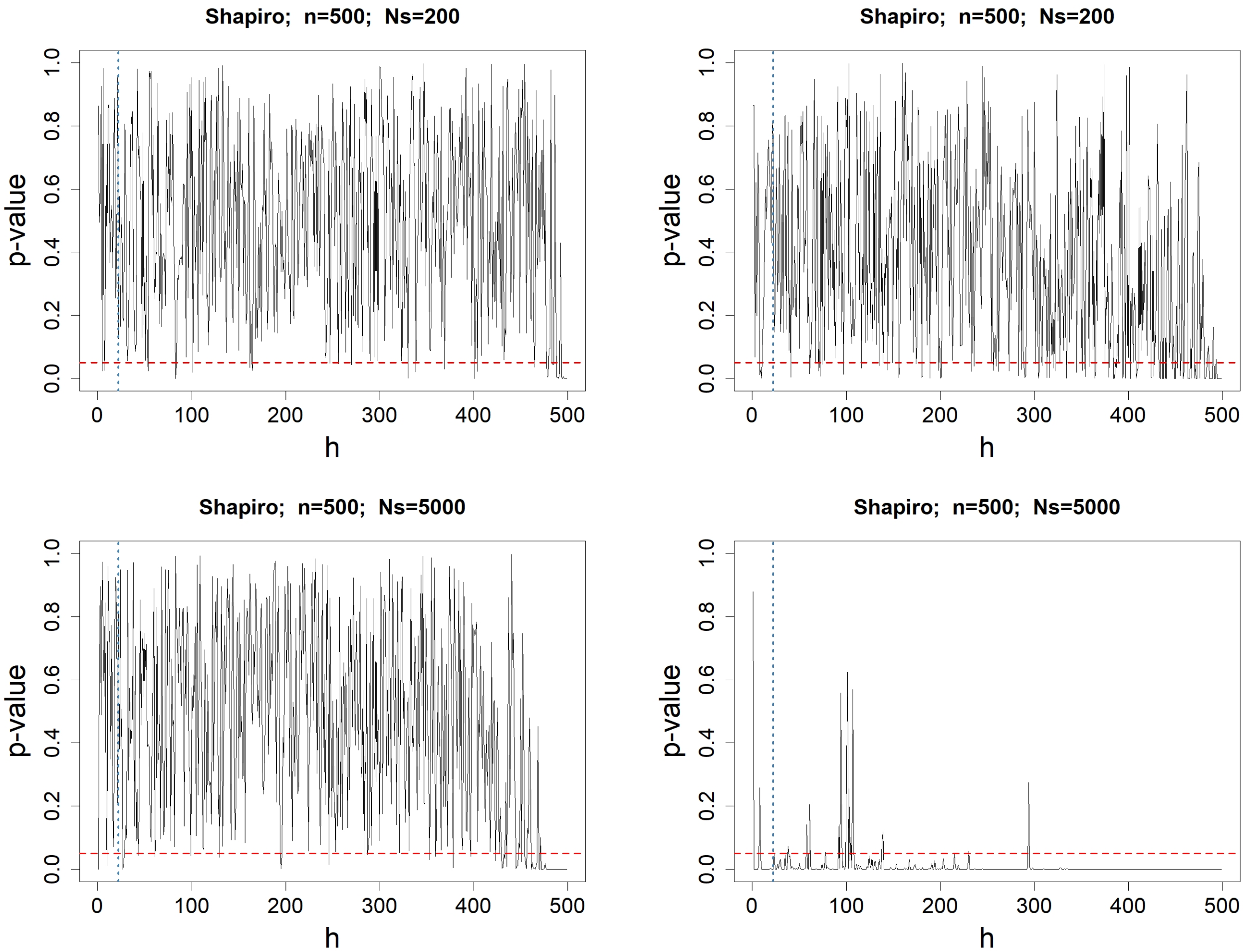

Figure S2: P-values when testing for the normality of the N_S values of $\hat{\rho}(h)$, for any fixed lag h varying from 1 to $n-1$. The involved normality test is Shapiro-Wilk's. The left column concerns MA(2) driven by a Gaussian WN whereas the right one deals with Exponential WN process. The length of the simulated WN process is $n = 500$. In the upper figures, the number of simulated MA(2) processes is $N_S = 200$, whereas it is $N_S = 5\,000$ in the bottom. The red-dotted horizontal line represents 5%, while the blue-dotted vertical line represents $h = \sqrt{n}$.

converge to a Gaussian distribution with $\mu = \rho(h)$ and $\sigma^2 = \frac{v_{h,h}}{n}$.

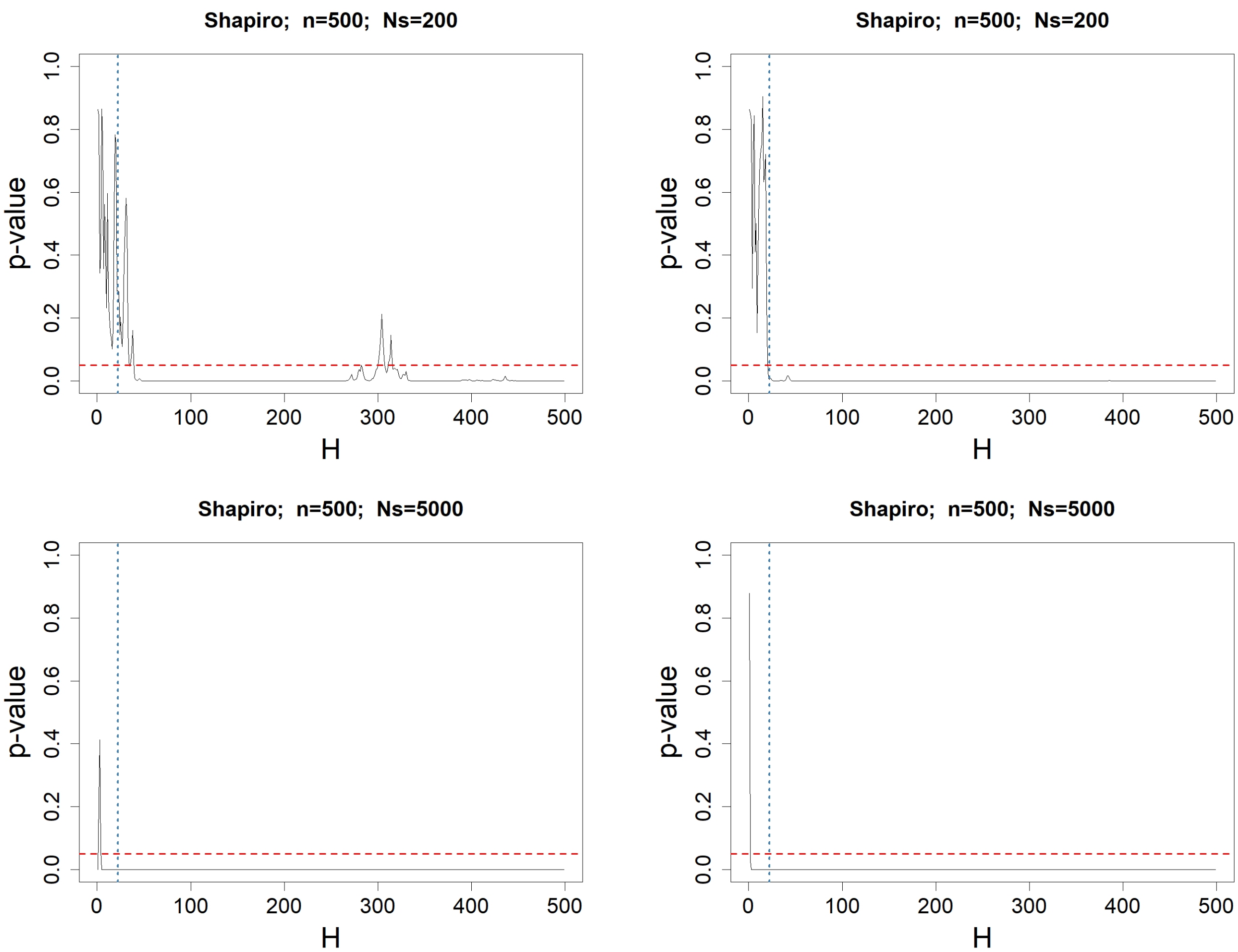

Figure S3: P-values when testing for the normality of the N_S values of $S_{acf}(H)$, for any fixed lag H varying from 1 to $n-1$. The involved normality test is Shapiro-Wilk's. The left column concerns MA(2) driven by a Gaussian WN whereas the right one deals with Exponential WN process. The length of the simulated WN process is $n = 500$. In the upper figures, the number of simulated MA(2) processes is $N_S = 200$, whereas it is $N_S = 5\,000$ in the bottom. The red-dotted horizontal line represents 5%, while the blue-dotted vertical line represents $H = \sqrt{n}$.