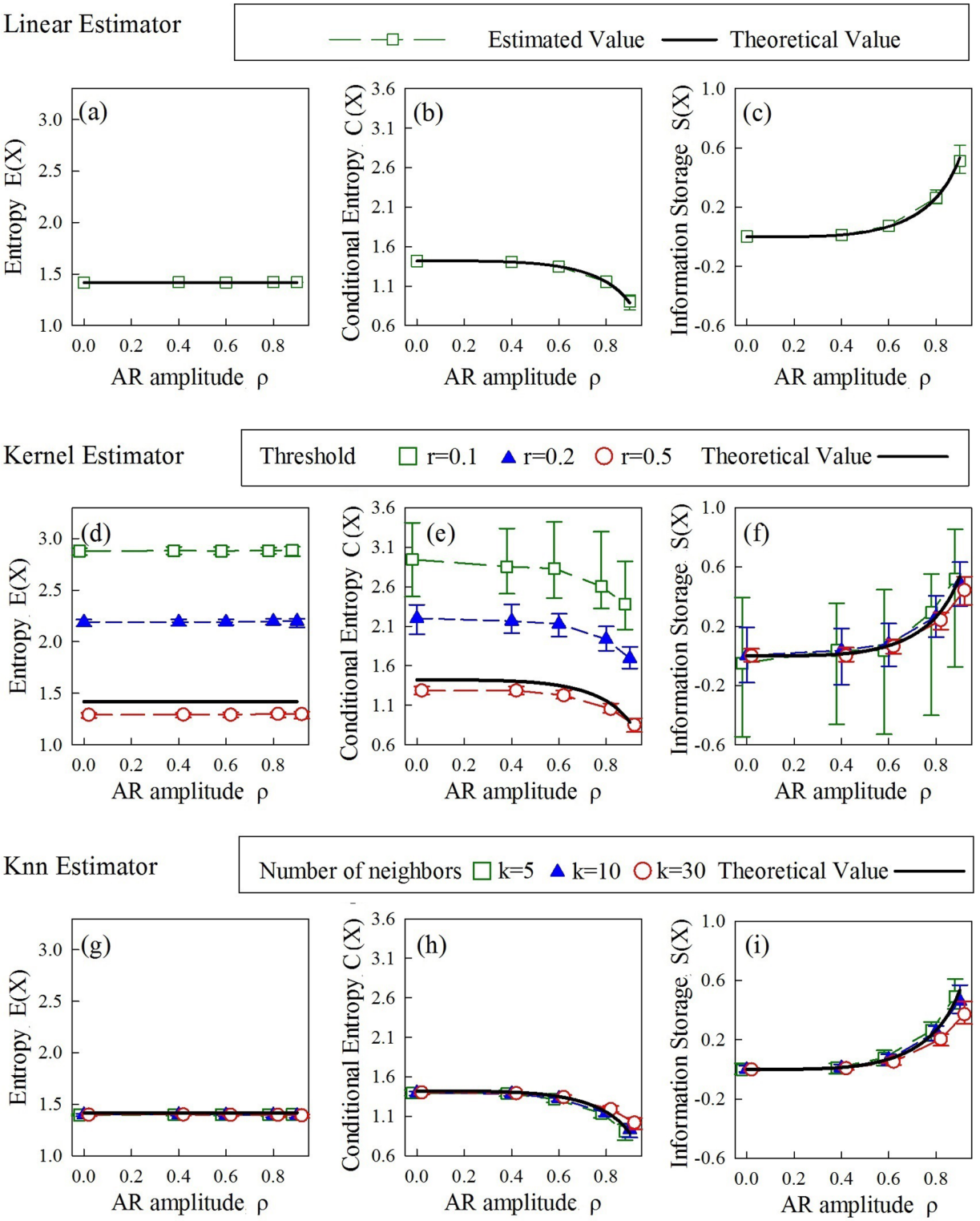

FIG. 9. Dependence of entropy estimators on estimator parameters for stationary AR processes. Plots depict the theoretical values (black solid line) and the estimated distributions (mean and 25%–75% percentiles over 100 realizations, colored symbols and error bars) of entropy (a), (d), (g), conditional entropy (b), (e), (h), and information storage (c), (f), (i) computed using the linear estimator (a)–(c), the kernel estimator (g)–(i) implemented with threshold $r = 0.1$ (green open squares), $r = 0.2$ (blue full triangles), and $r = 0.5$ (red open circles), and the knn estimator (g), (h) implemented with $k = 5$ (green open squares), $k = 10$ (blue full triangles), and $k = 30$ (red open circles) neighbors. Estimates are computed over realizations of length $N = 300$, generated with fixed AR frequency $f = 0.25$ and varying the AR amplitude in the range $\rho \in \{0,0.4,0.6,0.8,0.9\}$. Results: The kernel estimates of entropy and conditional entropy are strongly dependent on the parameter r setting the kernel threshold, whereas the knn estimates are much less sensitive to the parameter k setting the number of neighbors. Since the linear estimator assumes the form of the probability distributions, it has no free parameters.

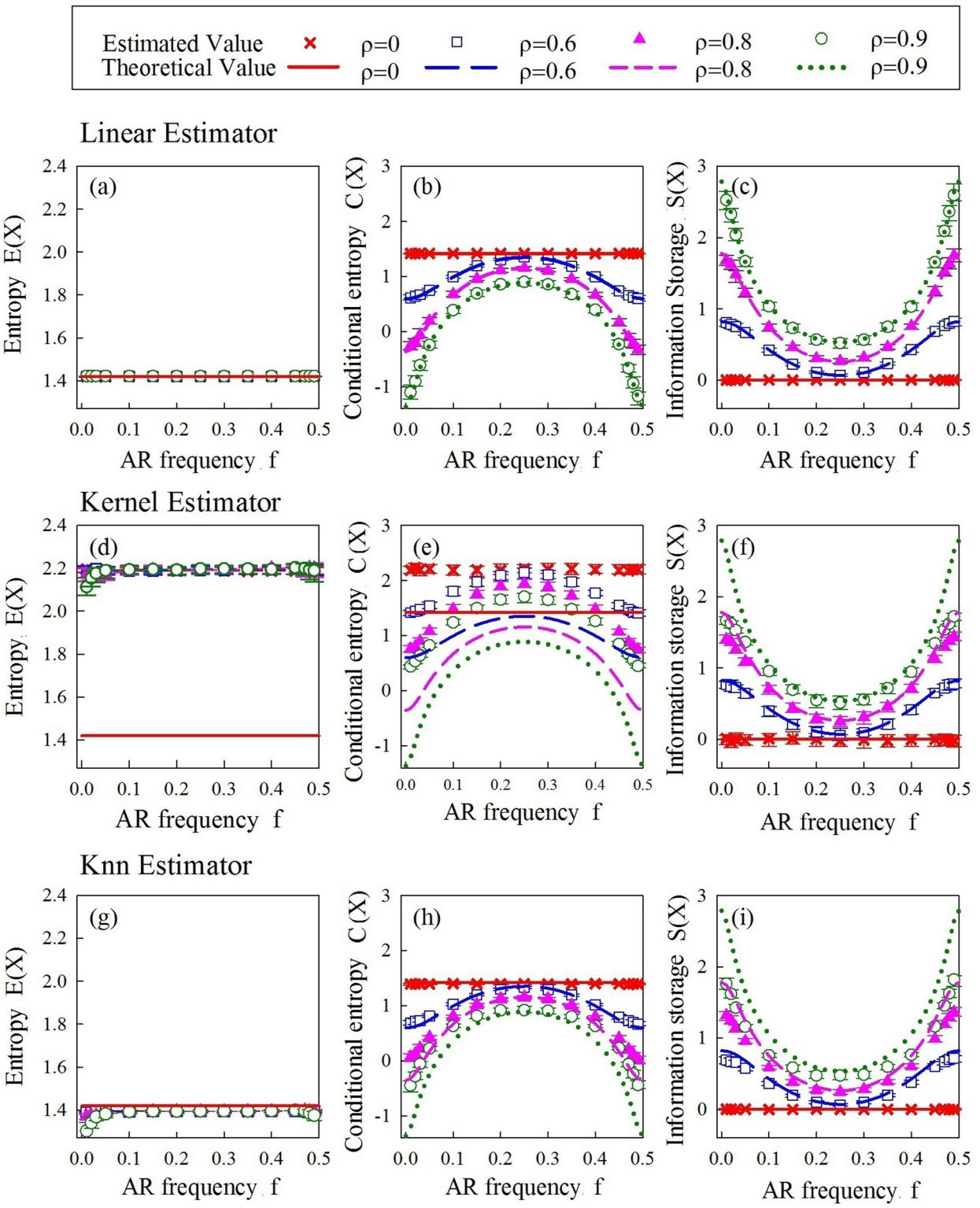

FIG. 10. Dependence of entropy measures and entropy estimators on AR process parameters. Plots depict the theoretical values (lines) and the estimated distributions (mean and 25%–75% percentiles over 100 realizations lasting $N = 300$ samples, colored symbols and error bars) of entropy (a), (d), (g), conditional entropy (b), (e), (h) and information storage (c), (f), (i) computed using the linear estimator (a)–(c), the kernel estimator implemented with threshold $r = 0.2$ (d)–(f), and the knn estimator implemented with $k = 10$ neighbors (g)–(i). Each measure is computed as a function of the AR frequency varying in the range $f \in (0\text{–}0.5)$ for different values of the AR amplitude ($\rho = 0$, red crosses and solid lines; $\rho = 0.6$, blue long-dashed lines and open squares; $\rho = 0.8$, pink short-dashed lines and full triangles; $\rho = 0.9$, green dotted lines and open circles). Results: The linear estimates of entropy measures are close to the theoretical values regardless of the values of the AR amplitude ρ and frequency f. The kernel and knn estimates exhibit a bias that is more evident for high values of the AR amplitude ($\rho \geqslant 0.8$), and for very low or very high values of the AR frequency ($f \leqslant 0.1, f \geqslant 0.4$).

of the simulated data (i.e., stationary Gaussian process), the model-free analysis is complicated by empirical factors, such as the data length, but also by the statistical properties of the underlying process. Specifically, we found that the estimates of conditional entropy and information storage are strongly biased for processes exhibiting very slow or very fast regular oscillations. Moreover, the kernel estimates of these measures, though being extremely popular when implemented in measures like approximate entropy and sample entropy, are highly biased with a bias strongly dependent on the estimation parameter. On the other hand, small bias and low estimation variance can be attained by computing the same measures through the knn method.

B. Performance of entropy estimators and entropy measures for AR processes with nonstationarities

Using the results for the stationary AR process as a benchmark, in this section we study the effect of nonstationary behaviors on the performance of the estimators of entropy, conditional entropy, and information storage. Starting from stationary AR processes, we induce three types of nonstationarity by superimposing sinusoidal trends, adding random spikes and inflating the amplitude of segments of the original time series. For each type of nonstationarity, we first compare the statistical properties and the entropy measures estimated for of individual realizations of AR processes with and without nonstationarity and then perform exhaustive analysis assessing how the estimation performance varies with the severity of the simulated nonstationary behavior.

1. Nonstationarity due to sinusoidal trends

The first type of nonstationarity we consider is the sinusoidal trend. As shown in Figs. 11(a) and 11(b), the presence of a sinusoidal trend changes the probability distribution of X_n, which departs from Gaussianity and becomes bimodal. Trends have also the effect of distorting the temporal relation between X_n and X_{n-2}, making it more evident but changing the sign of their correlation [see Figs. 11(c) and 11(d), where the cloud of points is less dispersed and the fitting line changes its slope]. In this case, the change of the distribution after superimposition of the trend is not reflected by alterations of the estimates of the entropy of the time series, while the higher predictability is reflected by a susbstantial decrease of the conditional entropy and a clear increase of the information storage [Figs. 11(e) and 11(f)]. These effects are evident regardless of the entropy estimator.

The effects described above are confirmed by the analysis of 100 process realizations with and without sinusoidal trends reported in Fig. 12. The analysis performed as a function of the AR amplitude shows that, regardless of the period or the amplitude of the trend, the presence of trends does not have big effects on the estimation of entropy but totally impairs the ability of all entropy estimators in following the variations of the regularity of the AR process. While such an ability was documented in Fig. 8 for the original stationary AR process, here we see that none of the estimators can correctly follow the theoretical behaviors of conditional entropy and information storage as a function of AR amplitude ρ, not even qualitatively [see the difference between the estimated values for signals with trends (colored lines with markers) and the theoretical value for original AR signals (black lines)]. Moreover, for all estimators we find that the estimation bias depends more on the trend amplitude A (represented by red line with cross and pink line with triangle) than on the trend period T (represented by blue line with square and green line with circle). With trend amplitude equal to the variance of the original process ($A = 1$), the conditional entropy is underestimated for $\rho < 0.7$ and overestimated for $\rho > 0.7$, while the opposite happens for the information storage; with trend amplitude $A = 5$ the conditional entropy is systematically underestimated and the information storage is systematically overestimated.

Overall, we find that trends have a big impact on the detection of the dynamical complexity of stochastic processes. In all cases, the negative impact of the presence of trends is documented by the flat response of the entropy measures to variations in the predictability of the underlying original process.

2. Nonstationarity due to spikes

Next, we consider the case in which the stationary AR process is corrupted by spikes with random temporal location and amplitude. Spikes are extremely common in real life signals [18,49,52,53,55,108] and may be manifested as artifacts originating from external conditions or from the intrinsic dynamics of the system. Here we simulate spikes with random temporal location and amplitude. As shown in Figs. 13(a) and 13(b), the presence of spikes concentrates the probability distribution of X_n in a way such that the largest part of the signal variance is due to the spikes, which are outliers of the distribution. As a result of the presence of random outliers, the points of the 2D phase plot of (X_{n-2}, X_n) are concentrated around the origin and the estimation of the temporal relation between X_n and X_{n-2} is strongly biased with respect to the clean case [see Fig. 13(d), where the linear fit follows the outliers rather than the noncorrupted points]. As documented in Figs. 13(e) and 13(f), the more concentrated probability distribution induced by the presence of spikes result in a decrease in the model-free estimate of entropy (kernel and knn estimators), while the linear model-based estimate is unchanged for these normalized time series. Moreover, the errors in the detection of the linear relation between time series samples result in a clear overestimation of conditional entropy and underestimation of information storage by the linear estimator. On the contrary, the kernel and knn estimators are still able to capture the predictability of the time series at least to some extent, as demonstrated by the detection of a significant amount of information storage resulting from the estimation of a decrease in the conditional entropy compared with the entropy. We ascribe the higher robustness to spikes of kernel and knn estimators of conditional entropy and information storage to the fact that these estimators explore locally the state space in the computation of probabilities, thus excluding from the estimate the points corrupted by spikes.

Figure 14 reports the results of the systematic analysis of the effects of spikes, performed studying the behavior of the entropy measures as a function of the AR amplitude of the uncorrupted AR process at varying the frequency of occurrence and the amplitude of the spikes. We found that

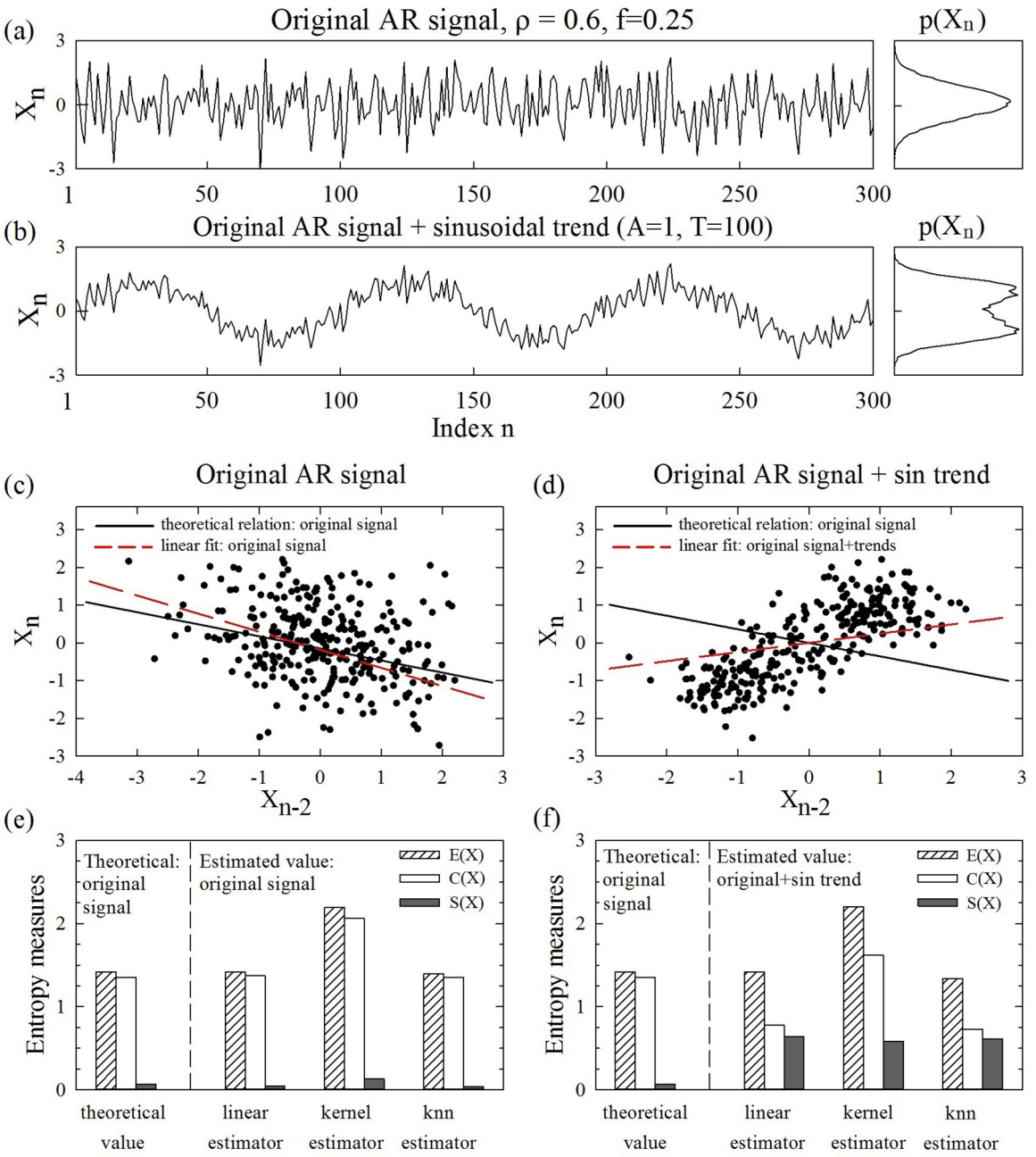

FIG. 11. Performance comparison for AR signals with and without sinusoidal trends: alteration of signal properties and entropy estimates. (a) and (b) show exemplary realizations of a stationary AR process with amplitude $\rho = 0.6$ and frequency $f = 0.25$ before and after superposition of a sinusoidal trend with amplitude $A = 1$ and period $T = 100$; the corresponding probability distributions are shown on the right. Signals are normalized to zero mean and unit variance. (c), (d) 2D phase plots of (X_n, X_{n-2}) derived from the time series in (a), (b). The generating equation of this AR process with $\rho = 0.6$ and $f = 0.25$ is $X_n = -0.36X_{n-2} + U_n$, which yields the theoretical temporal relation between X_n and X_{n-2} shown by the solid black line; the estimated temporal relation obtained through linear least-squares fit of the two clouds of points is shown by the red dashed lines. (e), (f) Entropy (shaded bars), conditional entropy (white bars) and information storage (gray bars) expressed as theoretical values computed for the stationary AR process without trends and estimated values computed for the time series in (a), (b). Estimations are performed using the linear estimator, the kernel estimator with threshold $r = 0.2$, and the knn estimator with $k = 10$ neighbors. Results: The presence of a sinusoidal trend superimposed to a realization of the AR process alters the probability distribution of X_n and distorts the temporal relation between X_n and X_{n-2}. This results in a significant decrease of the conditional entropy and in a significant increase of the information storage for all estimators.

the linear estimates of conditional entropy and information storage are highly affected by spikes, which blunt the capability of the measures to respond to changes in the AR amplitude [Figs. 14(b) and 14(c), except for the case of low spike amplitude ($A = 1$) and percentage ($P = 5\%$) in which a certain performance is preserved]. On the other hand, spikes were found to be less problematic for the kernel and knn estimates of the entropy measures. Figures 14(d)–14(i)

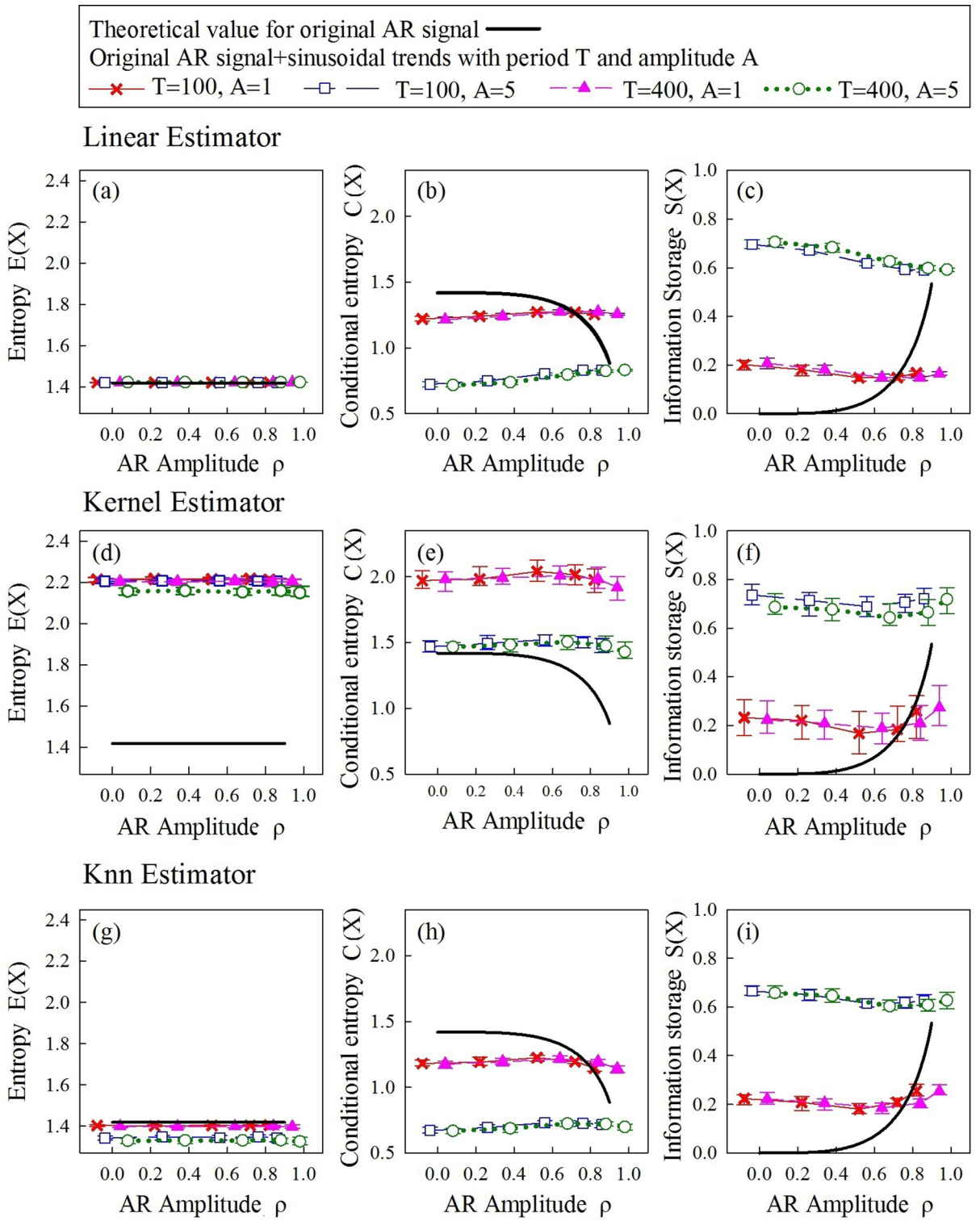

FIG. 12. Effects of nonstationarity due to sinusoidal trends on the estimation of entropy measures. Plots depict the behavior of entropy (a), (d), (g), conditional entropy (b), (e), (h) and information storage (c), (f), (i) computed as a function of the amplitude ρ of an AR process with fixed frequency $f = 0.25$, expressed as theoretical values computed for the original process without trends (black solid lines), and estimated distributions (mean and 25%–75% percentiles) computed over 100 realizations of $N = 300$ samples of the process, each corrupted with an additive sinusoidal trend of period T and amplitude A and normalized to zero mean and unit variance (colored symbols and error bars: $T = 100, A = 1$, red crosses and solid lines; $T = 100, A = 5$, blue open squares and long-dashed lines; $T = 400, A = 1$, pink full triangles and short-dashed lines; $T = 400, A = 5$, green open circles and dotted lines). Estimates are performed using the linear estimator (a)–(c), the kernel estimator implemented with threshold $r = 0.2$ (d)–(f), and the knn estimator implemented with $k = 10$ neighbors (g)–(i). Results: The presence of trends impairs the ability of all estimators to quantify the changes of conditional entropy and information storage induced by variations in the AR amplitude ρ. Moreover, trends induce an estimation bias bias proportional to the trend amplitude A.

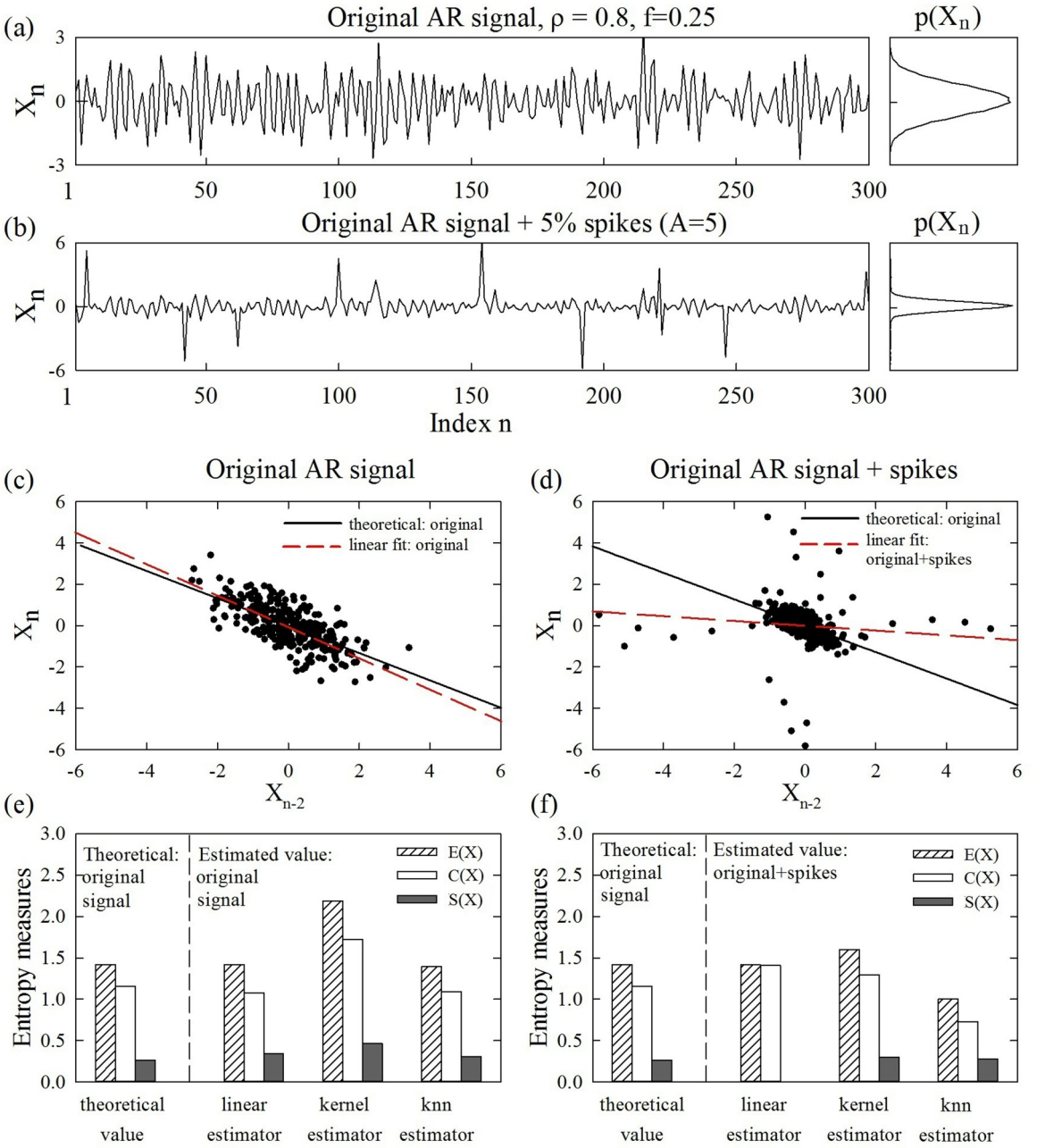

FIG. 13. Performance comparison for AR signals with and without random spikes: alteration of signal properties and entropy estimates. (a), (b) Exemplary realizations of a stationary AR process with amplitude $\rho = 0.8$ and frequency $f = 0.25$ before and after superposition of spikes with amplitude $A = 5$ to 5% of the time series points; the corresponding probability distributions are shown on the right. Signals are normalized to zero mean and unit variance. (c), (d) 2D phase plots of (X_n, X_{n-2}) derived from the time series in (a), (b). The generating equation of this AR process with $\rho = 0.8$ and $f = 0.25$ is $X_n = -0.64X_{n-2} + U_n$, which yields the theoretical temporal relation between X_n and X_{n-2} shown by the solid black line; the estimated temporal relation obtained through linear least-squares fit of the two clouds of points is shown by the red dashed lines. (e)–(f) Entropy (shaded bars), conditional entropy (white bars), and information storage (gray bars) expressed as theoretical values computed for the stationary AR process without spikes and estimated values computed for the time series in (a), (b). Estimations are performed using the linear estimator, the kernel estimator with threshold $r = 0.2$, and the knn estimator with $k = 10$ neighbors. Results: The presence of spikes superimposed to a realization of the AR process concentrates the probability distribution of the process and adds random outliers, which blurs the detection of the temporal relation between between X_n and X_{n-2}. This results in the inability of the linear estimator to detect the information storage in the process, while the kernel and knn estimators are less affected.

display that, apart from a negative bias in the estimation of entropy and conditional entropy, the estimated values of the entropy measures could correctly follow the variations in their theoretical values for the original process induced by changing the AR amplitude. The kernel and knn estimates of information storage exhibit a lower bias and a higher variance than the corresponding estimates of conditional entropy. The dependence of the estimation biases on the spike percentage P

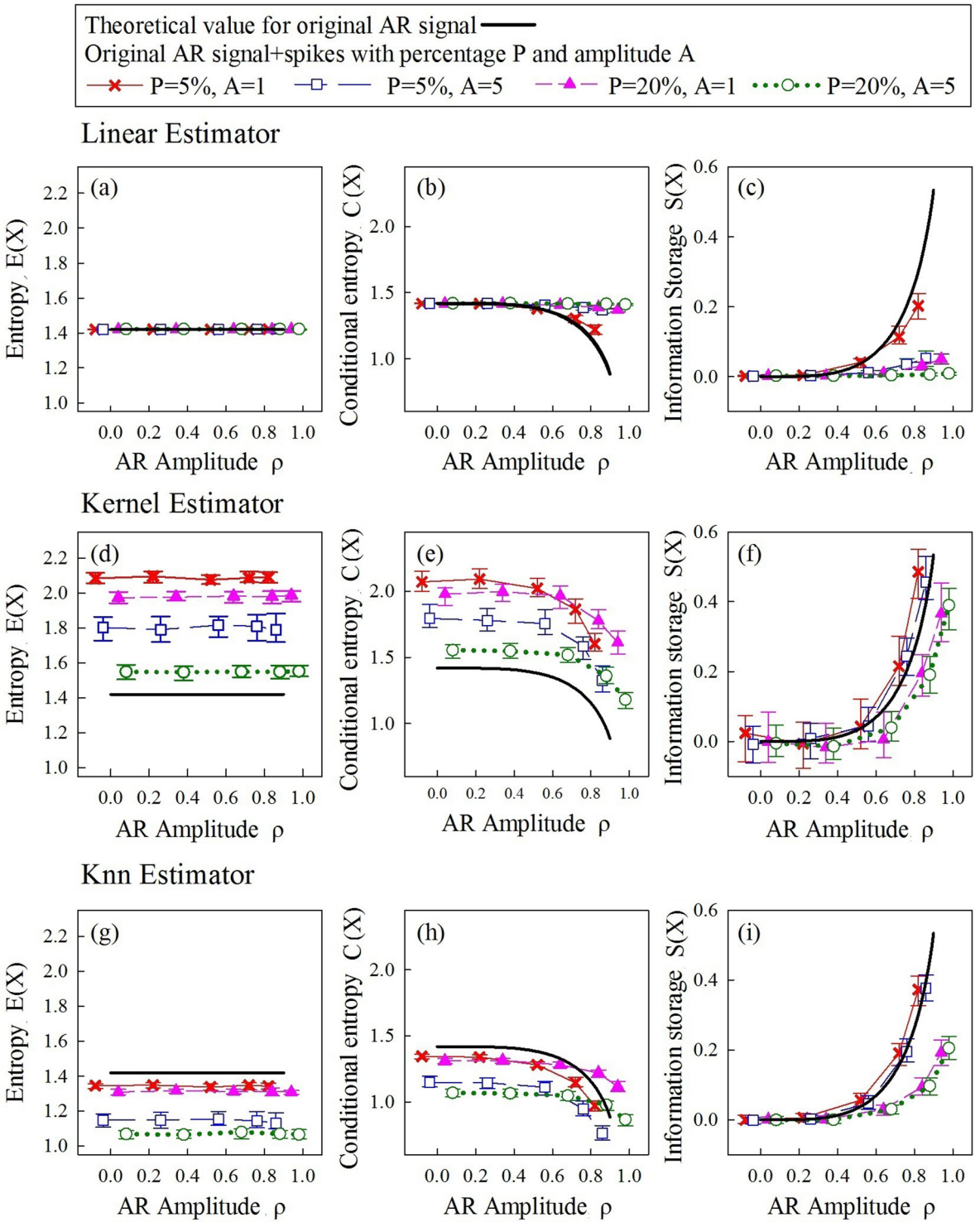

FIG. 14. Effects of nonstationarity due to random spikes on the estimation of entropy measures. Plots depict the behavior of entropy (a), (d), (g), conditional entropy (b), (e), (h), and information storage (c), (f), (i) computed as a function of the amplitude ρ of an AR process with fixed frequency $f = 0.25$, expressed as theoretical values computed for the original process without spikes (black solid lines), and estimated distributions (mean and 25%–75% percentiles) computed over 100 realizations of $N = 300$ samples of the process, each corrupted with additive random spikes of amplitude A occurring with probability P and normalized to zero mean and unit variance (colored symbols and error bars: $P = 5\%, A = 1$, red crosses and solid lines; $P = 5\%, A = 5$, blue open squares and long-dashed lines; $P = 20\%, A = 1$, pink full triangles and short-dashed lines; $P = 20\%, A = 5$, green open circles and dotted lines). Estimates are performed using the linear estimator (a)–(c), the kernel estimator implemented with threshold $r = 0.2$ (d)–(f), and the knn estimator implemented with $k = 10$ neighbors (g)–(i). Results: The presence of spikes partially impairs the ability to quantify the changes of conditional entropy and information storage induced by variations in the AR amplitude ρ; the impairment is more evident for the linear estimator and for high percentages of spikes. Moreover, spikes induce an estimation bias proportional to both the amplitude and the percentage of spikes.

and the spike amplitude A varies across estimators: the linear estimator fails when $P > 5\%$ or $A > 1$, the kernel estimator is equally affected by P and A and the knn estimator is more affected by P than A.

Overall, these results indicate that spikes have a deleterious impact on the model-based estimation of the measures of dynamical complexity. Since spikes are commonly encountered in a large variety of practical settings, we conclude that cautions should be used in adopting linear approaches to the computation of conditional entropy and information storage in the presence of these artifacts. On the contrary, model-free estimates are less affected by spikes and, in the presence of a moderate amount and amplitude of spikes, they are still sensitive to variations in the dynamical complexity of the clean time series.

3. Nonstationarity due to local changes in the signal variance

As a third nonstationary behavior, we consider the alteration in the amplitude of segments of the original AR process. As shown in Figs. 15(a) and 15(b), the local alteration of the signal variance has the effect of concentrating the probability distribution of X_n in a similar way than for the case of spikes. Similarly, the 2D phase plot of (X_{n-2}, X_n) exhibits a percentage of outliers that surround the cloud of points representing the non-corrupted portions of the original time series [Fig. 15(d)]. However, since this type of nonstationarity does not destroy the temporal relation between the time series samples, the linear fit in the 2D phase plot of (X_{n-2}, X_n) is still quite accurate. As a result, the linear estimation of conditional entropy and information storage is a bit degraded, but not fully impaired as in the case of random spikes [see Figs. 15(e) and 15(f)]. In this individual realization, the kernel and knn estimators provide slightly better performances in terms of estimation of conditional entropy and information storage. Note that, as in the case of random spikes, the concentration of the probability distribution is reflected by lower values of the entropy estimated using the kernel and knn estimators, while the linear estimates are again unaffected by the shape of the distribution.

Figure 16 reports the results of the complete analysis whereby the estimation of entropy measures is performed as a function of the AR amplitude for different values of the percentage and maximal amplitude of the segments of high variance imposed in the AR process. The main effect of the presence of segments of high variance is the introduction of a negative bias in the model-free estimates of entropy and conditional entropy, as well as of a positive bias in the model-free estimates of information storage [Figs. 16(d)–16(i)]; the bias in the information storage is higher for the kernel estimates than for the knn estimates. The linear estimates of the entropy measures are less affected by this bias [Figs. 16(a)–16(c)]. In spite of the bias we found that, in all conditions of local alteration of the signal variance, the values of the entropy measures computed using all estimators could follow the changes in their theoretical value imposed by varying the AR amplitude ρ.

These results suggest that the presence of nonstationarity due to segments of high variance is not as detrimental as other types of nonstationary behaviors, as it introduces a bias in the entropy measures but does not preclude the capability of these measures to detect changes in dynamical complexity induced by alterations of the predictable structure of the observed process.

C. Performance of entropy estimators and entropy measures for fractionally integrated white noise processes

In this section, we investigate the theoretical behavior of the entropy measures, as well as the performance of all entropy estimators in computing these measures, for processes with power-law long-range correlations. After setting the properties of fractionally integrated white noise processes as described in the methods (Sec. II C 3), the theoretical values of the entropy measures are computed as a function of the differencing parameter d, which controls the sign and the strength of long range correlations using the derivations described in the Appendix. We then compare these theoretical values with the distribution of the estimated values, in order to evaluate comparatively the efficacy of the various entropy estimators.

Results of this analysis are reported in Fig. 17. First, we find that both the theoretical and the estimated values of conditional entropy decrease, and the values of information storage increase, with the strength of long-range correlations modulated by the differencing parameter d. Additionally, the asymmetric behaviors of conditional entropy and information storage in response to positive or negative variations of the differencing parameter d [Figs. 17(b), 17(c), 17(e), 17(f), 17(h), and 17(i)] document that entropy measures are more sensitive to positive long-range correlations than to anti-correlations of the same strength. These results mirror the fact that signals with positive correlation are often associated with longer memory than signals with a negative correlation of the same strength.

Moreover we investigate the dependence of the entropy estimates on the time series length N, finding that not only the variance, but also the bias of of the estimates of conditional entropy and information storage, decrease for longer time series; this behavior is particularly evident for positive long-range correlations [Figs. 17(b), 17(c), 17(e), 17(f), 17(h), and 17(i)]. Similar discrepancies between numerical and mean estimated values of complexity measures were observed in [12] for $1/f$ noise time series, indicating that stationarity is an important prerequisite for the analysis of short time series, and trend-like behaviors may impair entropy estimation. A potential explanation for this finding lies in the the fact that, since signals with stronger positive correlation exhibit more trendlike behaviors [see Fig. 6(a) for a representative example], longer time series are needed to capture the similarity in the patterns that determines accurate estimates of conditional entropy and information storage. These results hold for all estimators, and also confirm the higher variance of the kernel estimator, compared to the linear and nearest-neighbor estimators, as is previously observed for the simulations of pure AR processes.

The findings reported in this section document that, in addition to traditionally used analytical tools for the quantification of long-range correlations such as the detrended fluctuation analysis (DFA), also the entropy measures studied in this work, which are commonly used to assess short-range dependencies, are able to quantify the the degree of long-range dependency

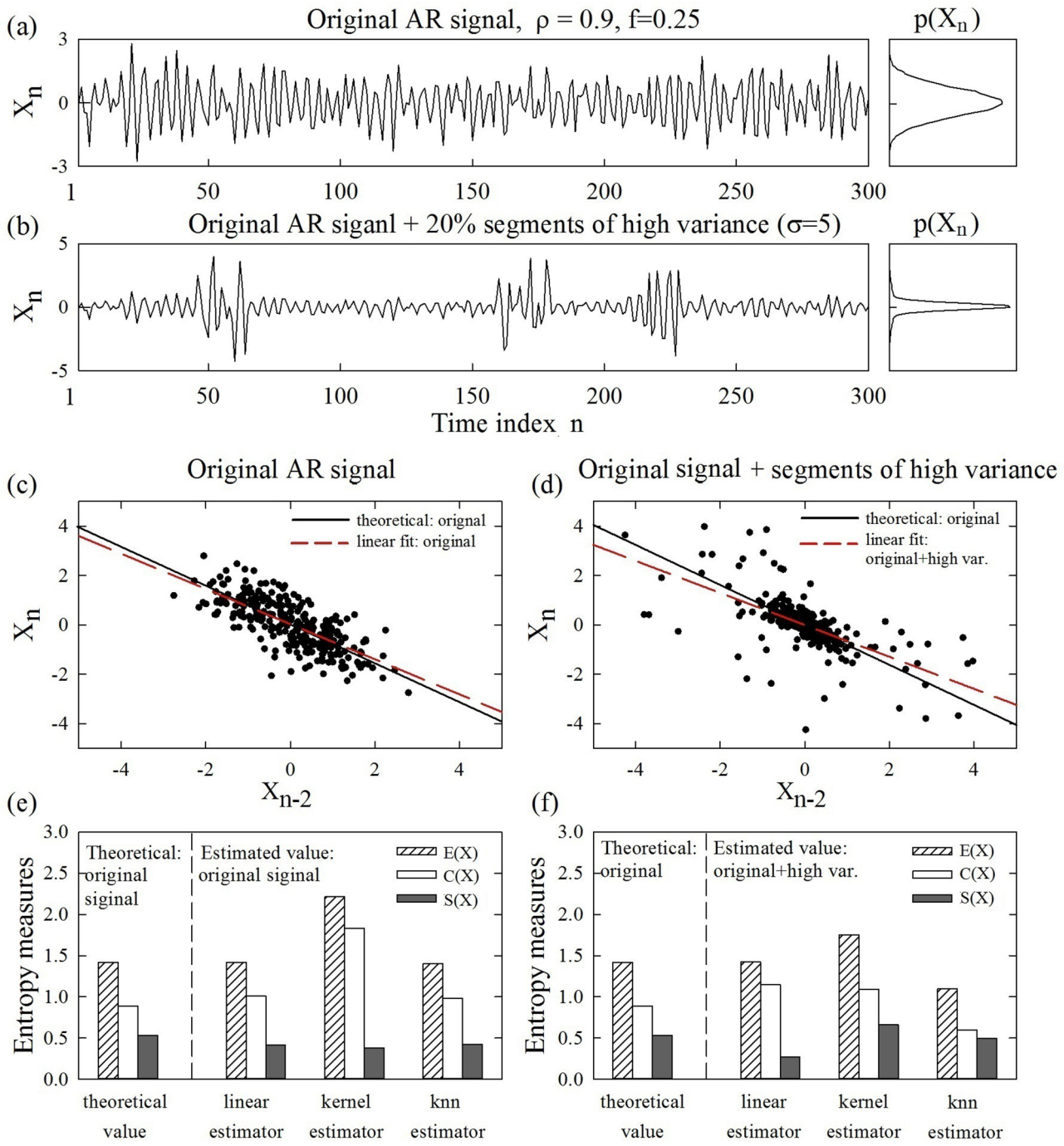

FIG. 15. Performance comparison for AR signals with and without segments of high variance: alteration of signal properties and entropy estimates in the presence of local changes in signal variance. (a), (b) Exemplary realizations of a stationary AR process with amplitude $\rho = 0.9$ and frequency $f = 0.25$ before and after inflating random segments by an amplification factor $\sigma = 5$ (a total of $P = 20\%$ of the time series points are inflated); the corresponding probability distributions are shown on the right. Signals are normalized to zero mean and unit variance. (c), (d) 2D phase plots of (X_n, X_{n-2}) derived from the time series in (a), (b). The generating equation of this AR process with $\rho = 0.9$ and $f = 0.25$ is $X_n = -0.81X_{n-2} + U_n$, which yields the theoretical temporal relation between X_n and X_{n-2} shown by the solid black line; the estimated temporal relation obtained through linear least-squares fit of the two clouds of points is shown by the red dashed lines. (e), (f) Entropy (shaded bars), conditional entropy (white bars), and information storage (gray bars) expressed as theoretical values computed for the stationary AR process without changes in variance and estimated values computed for the time series in (a), (b). Estimations are performed using the linear estimator, the kernel estimator with threshold $r = 0.2$, and the knn estimator with $k = 10$ neighbors. Results: The presence of segments with higher variance concentrates the probability distribution of the process and disperses a portion of the points without distorting their temporal relation. This results in a mild decrease of the conditional entropy and increase of the information storage for the kernel and knn estimators, while opposite changes are appreciated for the linear estimator.

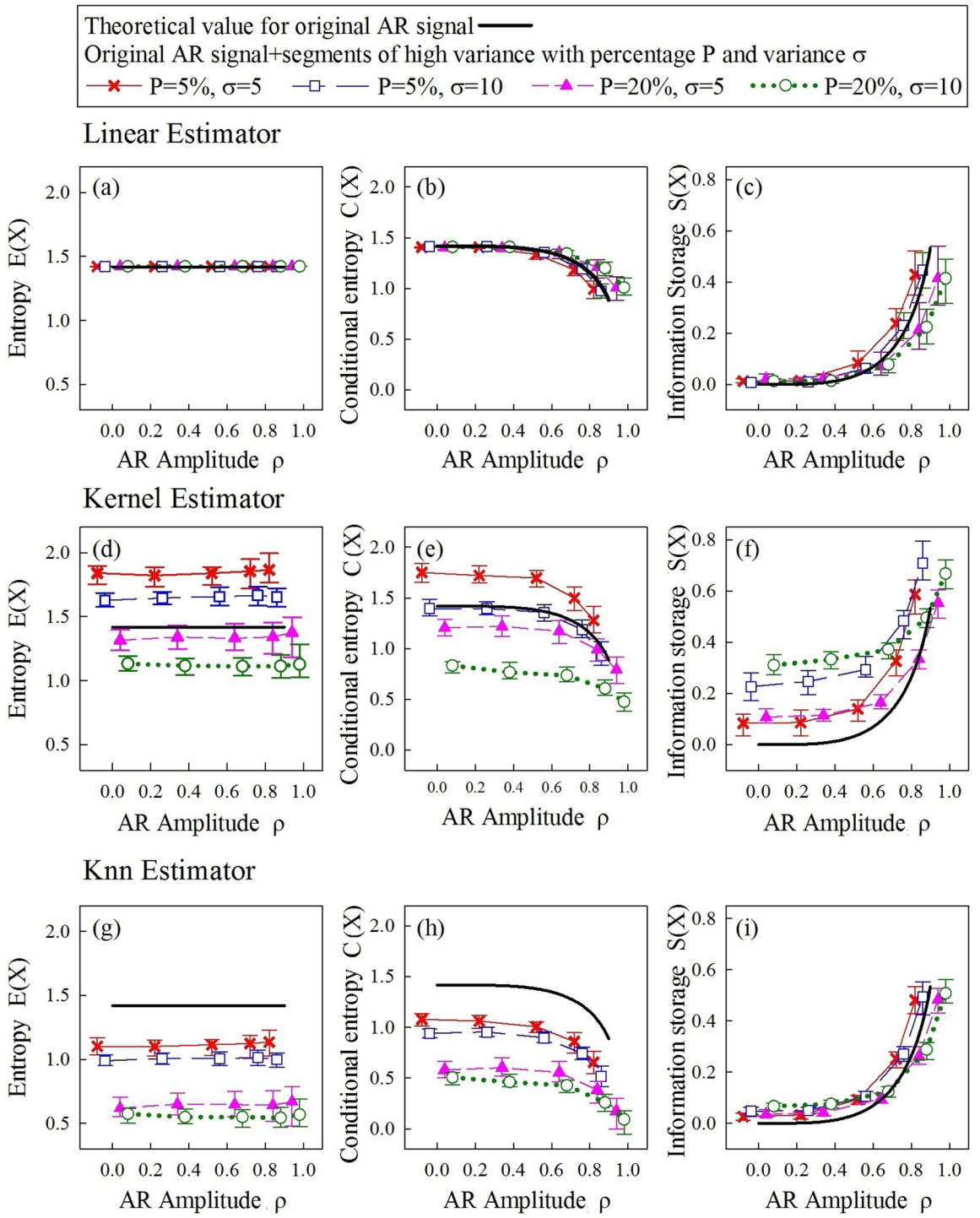

FIG. 16. Effects of nonstationarity due to local changes in the signal variance on the estimation of entropy measures. Plots depict the behavior of entropy (a), (d), (g), conditional entropy (b), (e), (h), and information storage (c), (f), (i) computed as a function of the amplitude ρ of an AR process with fixed frequency $f = 0.25$, expressed as theoretical values computed for the original process without changes in variance (black solid lines), and estimated distributions (mean and 25%–75% percentiles) computed over 100 realizations of $N = 300$ samples of the process, each corrupted by randomly distributed inflated segments and normalized to zero mean and unit variance. Each inflated segment lasts 20 points and is generated by magnifying original data points by a factor of σ, with the percentage of inflated points to the total signal length being $P\%$ (colored symbols and error bars: $P = 5\%, \sigma = 5$, red crosses and solid lines; $P = 5\%, \sigma = 10$, blue open squares and long-dashed lines; $P = 20\%, \sigma = 5$, pink full triangles and short-dashed lines; $P = 20\%, \sigma = 10$, green open circles and dotted lines). Estimates are performed using the linear estimator (a)–(c), the kernel estimator implemented with threshold $r = 0.2$ (d)–(f), and the knn estimator implemented with $k = 10$ neighbors (g)–(i). Results: The presence of segments with high variance does not impair significantly the ability to quantify the changes of conditional entropy and information storage induced by variations in the AR amplitude ρ. However, local changes in the signal variance induce an estimation bias proportional to the percentage of inflated points and—to a lower extent—to the amplitude of inflation.

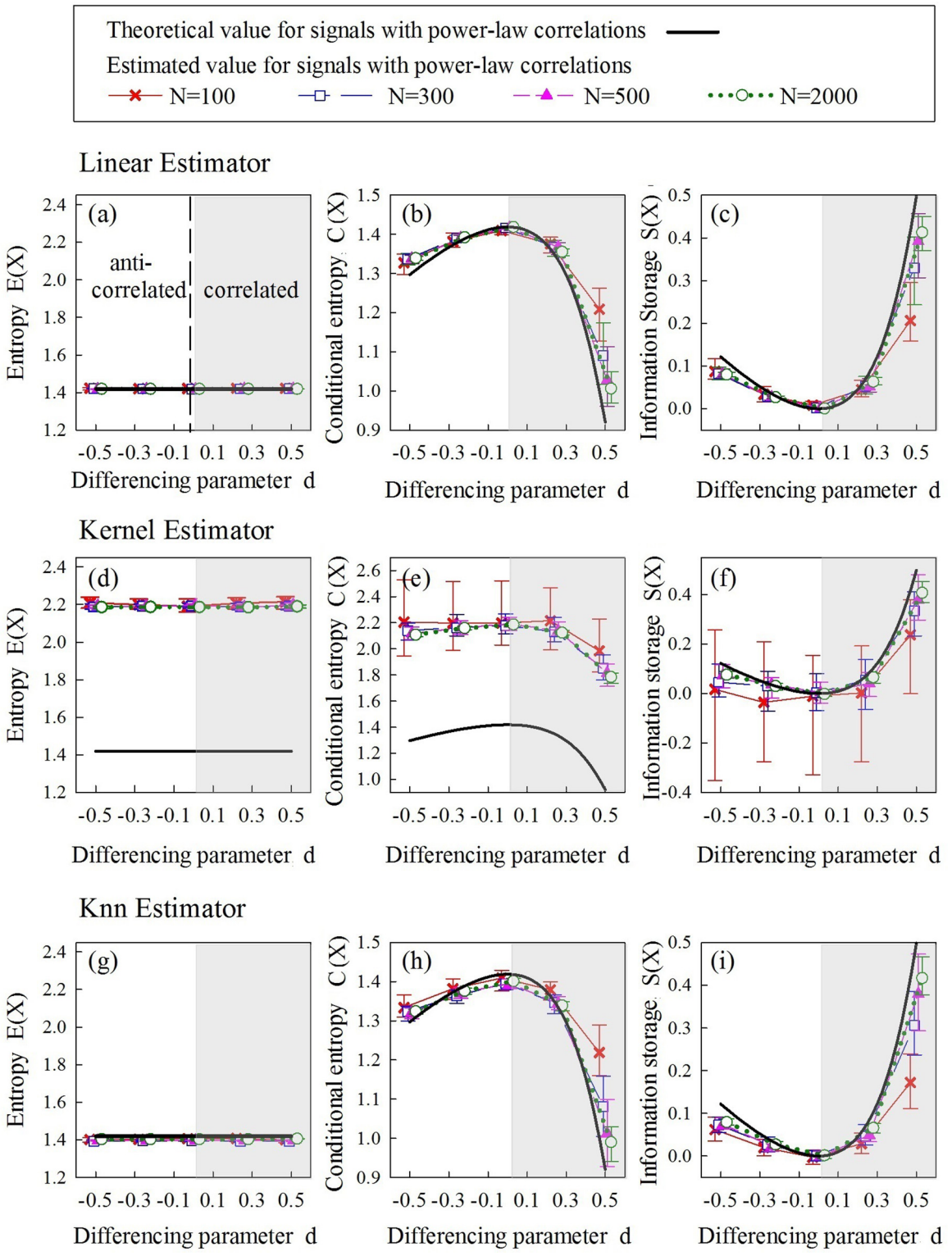

FIG. 17. Performance of entropy estimators and entropy measures for fractionally integrated processes with long-range power-law correlations. Plots depict the theoretical values (black solid lines) and the estimated distributions (mean and 25%–75% percentiles over 100 realizations of length $N = 100$ (red crosses and solid lines), $N = 300$ (blue open squares and long-dashed lines), $N = 500$ (pink solid triangles with short-dashed lines), and $N = 2000$ (green open circles with dotted lines)) of entropy (a), (d), (g), conditional entropy (b), (e), (h), and information storage (c), (f), (i) computed as a function of the differencing parameter d. Estimates are obtained using the linear estimator (a)–(c), the kernel estimator implemented with threshold $r = 0.2$ (d)–(f), and the knn estimator implemented with $k = 10$ neighbors (g)–(i). Results: Signals with positive correlations present lower conditional entropy and higher information storage than signals with anticorrelation of the same strength. All estimators reflect the changes in entropy measures with the type and strength of power-law correlation, with a higher accuracy for longer time series length, but exhibit a bias which increases with the correlation strength.

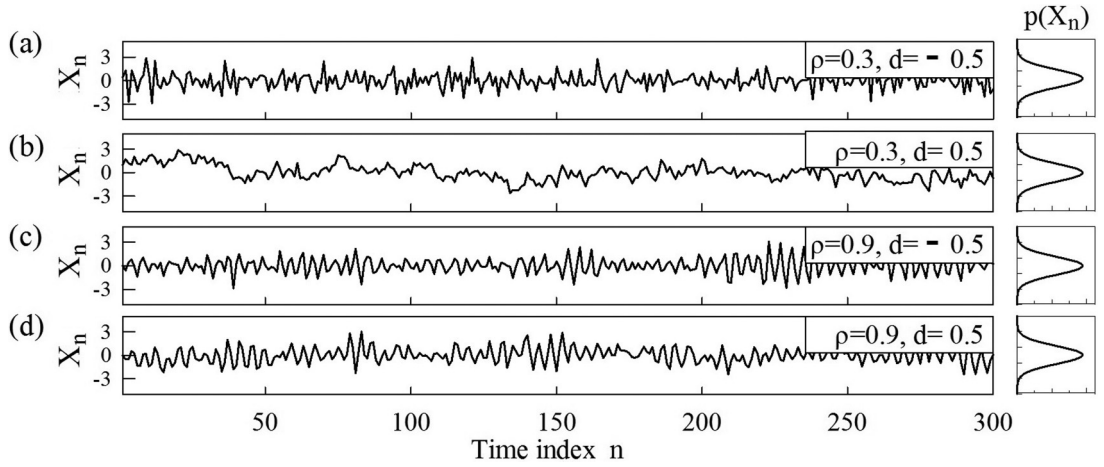

FIG. 18. Exemplary realizations of fractionally integrated autoregressive processes for varying AR amplitude ρ and differencing (correlation) parameter d. Plots depict signals with (a) weak autoregression ($\rho = 0.3$) and strong anti-correlation ($d = -0.5$), (b) weak autoregression ($\rho = 0.3$) and strong positive correlation ($d = 0.5$), (c) strong autoregression ($\rho = 0.9$) and strong anticorrelation ($d = -0.5$), (d) strong autoregression ($\rho = 0.9$) and strong positive correlation ($d = 0.5$). Slow trends in the signal are present for strong positive long-range correlation, and less evident for anticorrelated signals or in the presence of a strong AR component.

of the present of a process on its past values. However, the accuracy of the estimates is highly dependent on the time series length, indicating that—contrary to what happens for the estimation of short-range AR dependencies—very long realizations would be needed to yield accurate estimation of conditional entropy and information storage in the presence of strong positive long-range correlations.

D. Performance of entropy estimators and entropy measures for fractionally integrated autoregressive processes

The results reported in the previous sections describe the capability of entropy measures and entropy estimators to reflect changes in the temporal structure of both pure AR processes producing stochastic oscillations and pure fractionally integrated white noise processes exhibiting power-law long-range correlations. Here we extend the analysis by investigating whether and how the theoretical properties of the entropy measures and the performance of the entropy estimators change when the analyzed processes display both short-term AR dependencies and power-law long-range correlations. Representative examples of these processes are reported in Fig. 18, suggesting that their dynamical structure is altered in a different way depending on the strength of the stochastic oscillation and the sign of the long-range correlations. Specifically, we see that long-range correlations of the same strength ($|d| = 0.5$) determine different structure in the signals depending on their sign when the AR amplitude is low [$\rho = 0.3$, Figs. 18(a) and 18(b)], while they do not affect substantially or differently the dynamical structure when the AR amplitude is high [$\rho = 0.9$, Figs. 18(c) and 18(d)].

Results of the analysis performed at varying the AR amplitude ρ for different values of the differencing parameter d are reported in Fig. 19. First, the analysis confirms that, for these Gaussian normalized time series, the expected values of entropy are not dependent on the parameters ρ and d, and the estimates, apart from the bias of the kernel method known also before, are accurate for all approaches Figs. 19(a), 19(d), and 19(g). The theoretical values of conditional entropy and information storage deviate from their behavior for pure AR processes (pink short-dashed curves) in a way depending on the sign of long-range correlations: for anticorrelated processes ($d < 0$, red solid and blue long-dashed curves) the trend is similar to the case $d = 0$ apart from a shift of the curves toward lower conditional entropy and higher information storage; for positive long-range correlated processes (green dotted and gray dash-dot curves) conditional entropy and information storage are no more increasing monotonically with ρ, showing a nontrivial dependency especially for high values of the differencing parameter. As seen in Figs. 19(b), 19(c), 19(e), 19(f), 19(h), and 19(i), these theoretical trends were followed by the estimated values with a performance comparable to that observed for the various estimators applied to pure AR or fractionally integrated processes (i.e., with the strong bias typical of the kernel estimator and with a slightly better performance of the linear estimator compared with the knn estimator). The main difference is that in this case of combined AR and fractionally integrated processes all estimators (even the linear) produced biased estimates of the entropy measures. The bias was positive for conditional entropy estimates and negative for information storage estimates, increased with the differencing parameter d, and was more marked for positive d than for negative d.

Thus, the combined presence of stochastic oscillations arising from short-term interactions and of power-law long-range correlations, which is a very common situation of real-world time series, complicates both the theoretical behavior and the practical estimation of entropy measures. The interpretation of the values taken by these measures, as well as their accurate estimation, become problematic in the presence of very regular stochastic oscillations, when the increase of the conditional entropy (and the decrease of the information storage) with the strength of long-range correlations is more subtle, or in the presence of strong positive correlations, when it may happen that the conditional entropy does not decrease (and