



HAL
open science

Goodness-of-Fit Tests in Reverberation Chambers: Is Sample Independence Necessary?

Florian Monsef, Ramiro Serra, Andrea Cozza

► **To cite this version:**

Florian Monsef, Ramiro Serra, Andrea Cozza. Goodness-of-Fit Tests in Reverberation Chambers: Is Sample Independence Necessary?. *IEEE Transactions on Electromagnetic Compatibility*, 2015, 57 (6), pp.1748 - 1751. 10.1109/TEMC.2015.2451211 . hal-01241369

HAL Id: hal-01241369

<https://hal.science/hal-01241369>

Submitted on 10 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Goodness-Of-Fit tests in Reverberation Chambers : Is Sample Independence Necessary?

Florian Monsef*, Ramiro Serra†, Andrea Cozza*

*GeePs UMR 8507, Univ Paris-Sud, CentraleSupélec, CNRS, UPMC, Gif-sur-Yvette, France,

E-mail: florian.monsef@geeps.centralesupelec.fr

†Eindhoven University of Technology, Eindhoven, The Netherlands.

Abstract—Goodness-of-fit (GoF) tests are interesting tools to test the overmodedness of the field inside a reverberation chamber (RC). In the definition of GoF tests independent and identically-distributed samples are required. In the present paper we focus on the effect of partial data correlation on GOF tests. We analyze the sensitivity of the most common GOF tests used in electromagnetic compatibility. The results presented herein are based on numerical and experimental data. We show that most of GOF tests are insensitive to correlations up to rather high values depending on the type of GOF test chosen. Implications for practical applications are discussed.

Index Terms—Autocorrelation Coefficient, Data Independence, Random Fields, Reverberation Chambers, Statistical Electromagnetism.

I. INTRODUCTION

Statistical dispersion of field-related quantities in a reverberation chamber (RC) is conditioned by the degree of homogeneity of the field inside the chamber. Ideally, the excitation of a minimum number of modes [1] allows the field to approximatively follow a Gaussian law.

It is common to quantify stirrer efficiency by assessing the number of stirrer steps for which the field can be regarded as reasonably uncorrelated between two contiguous steps [2]. The question of how to estimate this number of positions has been the topic of many papers [3]–[6] and will not be considered here. If the case of the standards are considered, a large number of samples is used in order to obtain a more accurate estimate of the autocorrelation coefficient, from which a decimation factor can then be deduced. This decimation factor allows reducing the number of stirrer positions to the one that ideally corresponds to independent test configurations within the RC, avoiding redundant test results and ultimately saving time.

In order to check the Gaussian fit of the field, goodness-of-fit (GOF) tests can be used as explained in [2]. To the best of our knowledge, it is not clear if the set of data to use must be the decimated one, i.e., the set made up of uncorrelated data, or the complete set of data which provides more samples, even though partially correlated. It is indeed commonly thought that GoF need independent samples to be applied. In practice, samples being correlated, the condition of using independent samples is cut short and one may wonder about the possible use of correlated (and whence dependent) data.

The present work aims to give some insight into this question. To that end, we will recall in Section II the most common

GOF tests used in the context of RCs, while Section III will analyze GOF outcomes according to the autocorrelation model used to generate correlated data. Section III-C will also focus on GOF results but in the case of measured data. Finally, a tentative interpretation of the results will be proposed in Sec. IV.

Our findings show to what extent different GOF tests usually implemented within the EMC community are sensitive to data autocorrelation.

II. GOODNESS-OF-FIT TESTS

In order to determine if a set of samples belong to a given probability distribution, GOF tests can be used as a statistical decision tool. These tests can be performed, e.g., with the real and/or imaginary parts of the electric field in order to check if these follow a Gaussian law. But these tests can also be applied to field modulus samples in order to check whether these follow a Rayleigh distribution. In both cases, a *null hypothesis*, referred to as H_0 , is formulated and tested.

There are several existing GOF tests, and, with the same set of data, the outcome can differ from one GOF test to another. This is due to the fact that each GOF is based on a different metric. The question raised herein is not to know which GOF is the more adequate to RC applications but rather to use the most popular ones and to analyze their sensitivity to correlated data. Although not exhaustive, the most commonly encountered ones are the following:

- the χ^2 test [7],
- The Lilliefors (L) test [8],
- the Kolmogorov-Smirnov (KS) test [9],
- the Anderson-Darling (AD) test [5], [10].

The χ^2 test determines whether there is a significant difference between the theoretical frequencies of occurrence and the observed frequencies. It tells if the inevitable difference between the expected and observed values is due to stochastic dispersion in the sample, or if it there is a significant underlying difference. The Lilliefors and KS tests both estimate the maximum distance between the empirical cumulative distribution function (*cdf*) and the expected one. The distance is compared to a critical value, often referred to as D_{KS} , in order to decide whether the sample data belong to the expected distribution. The Lilliefors test uses sample moments unlike KS for which the moments of the expected distribution need to be specified *a priori*. The AD test integrates a weighted

difference between the sample *cdf* and the expected one. The weighted function is such that the test focuses more on the tails of the distribution than in other tests.

In what follows, these GOF tests will be performed on synthetic data generated numerically and on data obtained experimentally.

III. SENSITIVITY OF GOF ON SYNTHESIZED DATA

In this section we will present an approach where the first lag correlation coefficient is controlled in order to analyze if the outcomes of GOF tests are somewhat related to the degree of correlation.

A. Correlation Models

The idea is to generate correlated data to emulate, on average, what one could obtain by using a stirrer in practice. We are interested in generating two different correlation models. This allows us to see if the sensitivity is only linked to the lag-1 correlation, referred to as ρ_l , or if the correlation model itself has an impact on GOF outcomes.

We consider a set x_i whose samples follow a standardized normal law. For such a case, let us recall that the lag-1 coefficient is such that $\rho_l = E[x_i x_{i-1}]$, where $E[\cdot]$ is the ensemble average operator.

The first model is a first-order autoregressive (AR(1)) [11] model based on the following recursive equation:

$$\begin{aligned} x_1 &= \eta_1, \\ x_i &= \rho x_{i-1} + \eta_i \sqrt{1 - \rho^2}, \end{aligned} \quad (1)$$

where the random variable η is standardized, i.e., $\eta \sim \mathcal{N}(0, 1)$. For such an AR(1) model, we report in Fig. 1 the autocorrelation function obtained (by using 5000 samples) for $\rho = 0.3$ and $\rho = 0.7$, respectively. It is clear that ρ_l corresponds to ρ .

The second model considered herein is an AR(2) model for the correlation that may exist between a field value obtained at a given position and those obtained at the two preceding positions. The recursive equation reads in that case,

$$x_i = \rho x_{i-1} + \rho' x_{i-2} + \gamma \eta_i. \quad (2)$$

Note that (2) can be regarded as a dynamic system that may be unstable¹ if $\{\rho, \rho'\}$ are not properly chosen. As in (1), the coefficient weighting η_i , here γ , is set to ensure x_i to be standardized such that $\gamma = \sqrt{1 - (\rho + \rho')^2}$. This condition also provides a restriction on the possible values of ρ and ρ' since $(\rho + \rho')^2$ must be less than 1. The (somewhat arbitrary) choice made to fulfill the constraints on ρ and ρ' is such that $\rho' = 0.25 \rho$. Although not intuitive, for an AR(2) model, ρ_l is a function of ρ and ρ' such that $\rho_l = \rho/(1 - \rho')$, i.e., is larger than for an AR(1) model for a given ρ and $\rho' \neq 0$. In what follows, the number of data per set will be referred to as N .

B. Results

By considering a AR(1) model and by using (1), we generated data used as input to the GOF tests. The simulation

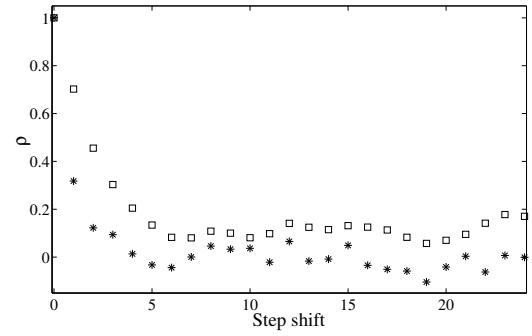


Fig. 1. Autocorrelation function obtained by using the model given by (1) for $\rho = 0.3$ (asterisks) and $\rho = 0.7$ (squares), respectively.

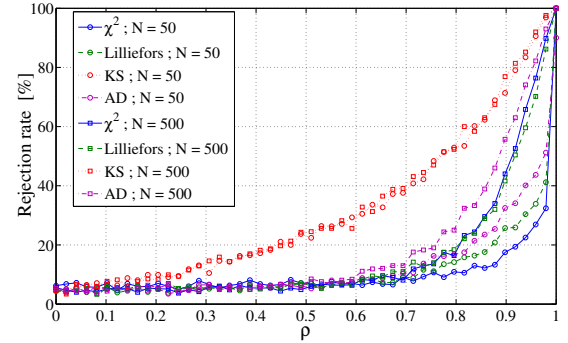


Fig. 2. Rejection rates ($\alpha = 5\%$), obtained with an AR(1) model, for different samples sizes and GOF tests as a function of data autocorrelation.

study is carried out by generating sets of x_i with $N = 50$ and $N = 500$ and for 50 values of ρ uniformly distributed over $[0; 1]$. To reduce statistical uncertainties, each realization was repeated a thousand times.

Fig. 2 summarizes the percentage of rejected null hypotheses of normality for different GOF tests under study and different sample numbers, for a significance level $\alpha = 5\%$ (p-value). We stress the fact that these rejection rates are solely coming from data correlation, since the tested data sets come from a normal parent distribution, as described by the process in (1).

It can be seen that the KS test shows a higher sensitivity to data autocorrelation than other GOF tests. It seems also more robust against different sample sizes. From these results, it appears that KS test could be suggested as a way of identifying residual autocorrelation on data, even for low values of ρ .

In order to highlight the case of the KS test sensitivity, critical values are plotted in Fig. 3 for a varying number of independent realizations² and different correlation levels. We can see that D_{KS} values are shifted upwards as the correlation rises. For $\rho < 0.5$, we notice that D_{KS} is almost insensitive to the degree of correlation, i.e., is almost similar to the *iid* case.

Another interesting way to tackle this analysis is to wonder which correlation level corresponds to an average p-value of 5%, the average being performed over sets of p-values taken from tests repeated a thousand times. Fig. 4 shows the results

¹if the poles modulus is larger than 1.

²that correspond to stirrer states ideally independent in practice.

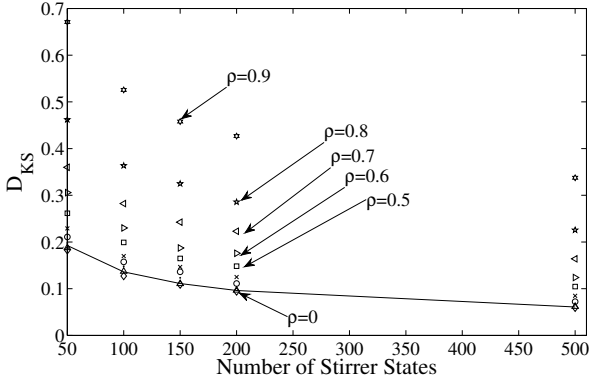


Fig. 3. The KS threshold obtained for different correlation levels as a function of the number of stirrer states. The threshold obtained for the classic (uncorrelated) case (solid line) is shown as well as the thresholds obtained for values of ρ reported in the figure and running from 0 to 0.9.

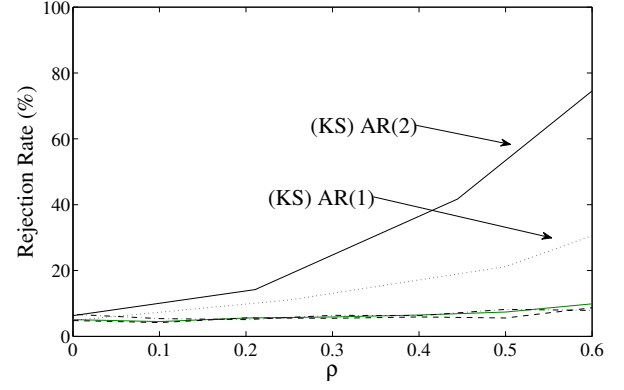


Fig. 5. Rejection rates obtained with an AR(2) model for $N = 500$. χ^2 (dashed line) AD (dashed-dotted line) and Lilliefors (Green) do not seem to be sensitive to correlation, unlike the KS test, for which the AR(1) result has been reported for comparison.

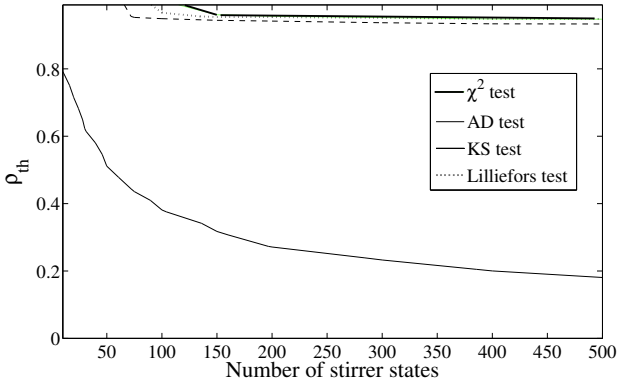


Fig. 4. Correlation threshold leading to an average empirical p-value of 5% as a function of the number of stirrer states.

obtained for the different GOF tests (see legend). We observe that the correlation level for non-KS tests is quite high and almost independent of the number of stirrer states.

Now, let us analyze the rejection rate obtained when using an AR(2) model. Fig. 5 shows the corresponding results. It can be seen that non-KS tests (dashed line, dashed-dotted line and green line) are still insensitive to correlation, whereas the KS rejection rate increases faster (solid line) than what was obtained with an AR(1) model (dotted line). In a way, the AR(2) provides a higher average correlation level² which could explain the difference of slopes in the rejection-rate curves obtained for the two AR models considered here.

C. Experimental Setup and Measurements

We will check now if the sensitivity of the GOF tests studied herein is also observed when experimental data are used.

Measurements were performed in the RC facility of the Eindhoven University of Technology. The chamber is $4 \times 5 \times 2.85$ (h) m with a lowest useable frequency of 200 MHz, according to the criteria in [2]. A signal generator, a RF power

² as the for the lag-1 case, for a given ρ , the lag-2 coefficient is larger for the case of an AR(2) model than for an AR(1) model; it equals $\rho' + \rho^2 / (1 - \rho')$ for the former and only ρ^2 for the latter.

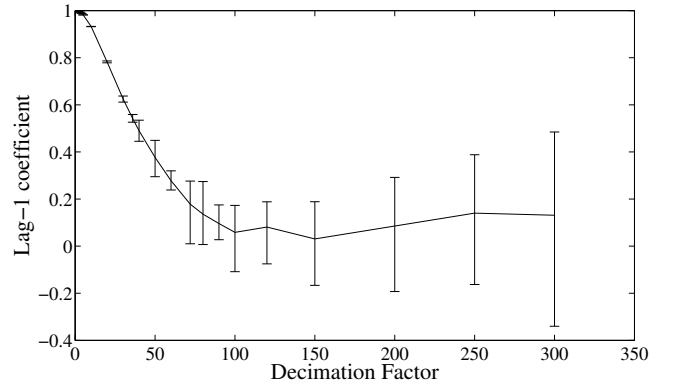


Fig. 6. Average lag-1 autocorrelation coefficient of measured data inside an RC at 1 GHz, for different levels of sample decimation (different sample size). Vertical bars stand for 95% confidence intervals.

amplifier and a log-periodic antenna (80 MHz - 1 GHz) were used in order to inject electromagnetic energy into the RC at a fixed frequency of 1 GHz. The three cartesian components of the electric field magnitude were measured by an electric field probe located inside the working volume for a total of 3600 stirrer positions (one field sample every 0.1°). As is well known, ideally each field component is expected to follow a Rayleigh distribution. By taking such small stirrer step angles, we intend to generate highly correlated data that are also reasonably Rayleigh distributed as well.

The procedure is to start with the complete set of 3600 samples, estimate ρ_l and apply the different GOF tests. Subsequently, we will decimate the set, in order to get a lower ρ_l (as well as a lower number of samples) and again apply the different GOF tests. We perform this process with increasing decimation until reaching a set of 12 samples. Fig. 6 shows the estimated ρ_l by using the three field components as a function of the sample size. As expected, the degree of correlation lowers as the decimation factor increases.

Three different GOF tests were applied to the increasingly decimated data: the χ^2 , the KS, and the AD tests. The Lilliefors test could be applied to test against a Rayleigh

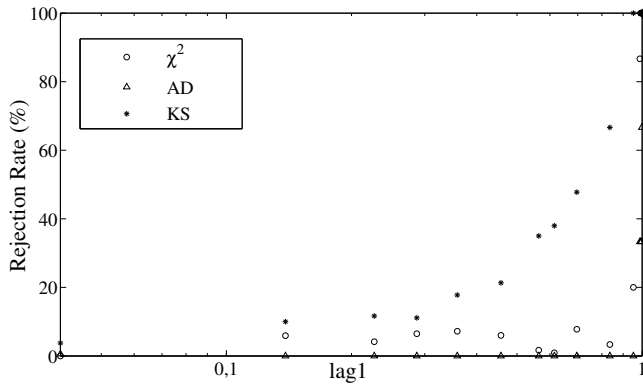


Fig. 7. Rejection rates ($\alpha = 5\%$) for different GOF tests (see legend) as a function of lag-1 autocorrelation for measured data.

distribution but tables needed for that specific case are not available. Fig. 7 highlights the rejection rates as a function of lag-1 correlation for the measured (and decimated) sets of data. Similar conclusions as from the simulation study can be drawn: GOF tests other than the KS appear insensitive to data autocorrelation, even for relatively high values of ρ_l , whereas KS tests seem to be as sensitive as what was observed with numerical results.

IV. DISCUSSION

To understand the pertinence of the results presented in the previous sections, it is worth recalling that the main effect of correlation is to reduce the range value spanned by the successive data samples, by introducing a direct dependence, roughly said, a sort of drag or inertia.

This has an impact on the number of bins that need to be set for the χ^2 test. As long as the correlation allows the data to span a sufficiently wide range of values to balance the frequencies in the different classes according to Cochran's criterion, the χ^2 test remains operational. When the correlation becomes too strong, the values spread collapses to a narrow value interval that makes possible only the definition of a single bin. And then the test fails.

The same idea can be transposed to the AD test where the empirical *cdf* fails to be representative of the whole real distribution; only a small portion of the *cdf* becomes accessible and the comparison to a reference distribution becomes no longer possible.

Concerning the KS test, the outcome is conditioned by Kolmogorov statistics. This statistics are sometimes called the Brownian bridge that assumes a random Brownian motion. Variations are due to trajectories, assumed to be independent. This independence is not the same concept required for independent data samples, as repeatedly used throughout the present work. However, a causality seems to exist. The correlation has an impact on the KS critical value, but to the best of our knowledge, there is no simple relation between the proper D_{KS} value and ρ_l . From the results presented in the previous section, the typical critical value $1.36/\sqrt{N}$ (solid line in Fig. 3) seems too low when correlation intervenes, such that H_0 is often rejected.

Finally, the Lilliefors test seems to be more robust by using sample moments than the KS test, but it is not easy to be more specific on the reasons that make it more robust.

V. CONCLUSIONS

Uncorrelated data were shown not to be necessary when performing χ^2 , Anderson-Darling, or Lilliefors GOF tests. It seems that these tests fail when a correlation of about 90% is reached. However, the Kolmogorov-Smirnov test was shown to be very sensitive to correlation. These results were obtained by using synthetic data and experimental data. The rejection rate obtained highlighted the same clear trends.

As stated in standards, GoF evaluates the mismatch between the measured and the theoretical (ideal) distribution functions of field or power density. As highlighted in the present work, an RC user can perform some of these GoF tests without caring about the possible correlation that may exist between samples. In other terms, if the test is well chosen, there is no need to decimate the data to have an insight on the overmodedness of the chamber.

As a perspective, the KS GOF test could be thought as a way to detect correlation in scenarios where the data distribution is known in advance.

REFERENCES

- [1] F. Monsef, "Why a Reverberation Chamber Works at Low Modal Overlap," *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 6, pp. 1314–1317, 2012.
- [2] *Reverberation chamber test methods*, International Electrotechnical Commission (IEC), Std. 61 000-4-21, 2011.
- [3] F. Moglie and V. Primiani, "Analysis of the independent positions of reverberation chamber stirrers as a function of their operating conditions," *IEEE Transactions on Electromagnetic Compatibility*, vol. 53, pp. 288–295, 2011.
- [4] S. Pfennig and H. Krauthauser, "A general method for determining the number of independent stirrer positions in reverberation chambers," in *Electromagnetic Compatibility (EMC EUROPE), 2012 International Symposium on*. IEEE, 2012, pp. 1–6.
- [5] C. Lemoine, P. Besnier, and M. Drissi, "Estimating the effective sample size to select independent measurements in a reverberation chamber," *IEEE Transactions on Electromagnetic Compatibility*, vol. 50, no. 2, pp. 227–236, 2008.
- [6] O. Lundén and M. Backstrom, "Stirrer Efficiency in FOA Reverberation Chambers. Evaluation of Correlation Coefficients and chi-squared tests," in *IEEE International Symposium on Electromagnetic Compatibility*, vol. 1. IEEE, 2000, pp. 11–16.
- [7] J. Kostas and B. Boverie, "Statistical Model for a Mode-Stirred Chamber," *IEEE Transactions on Electromagnetic Compatibility*, vol. 33, no. 4, pp. 366–370, 1991.
- [8] F. Monsef and A. Cozza, "Goodness-of-fit tests in radiated susceptibility tests," in *Workshop on Aerospace EMC, 2012 Proceedings ESA*. IEEE, 2012, pp. 1–5.
- [9] P. Corona, G. Ferrara, and M. Migliaccio, "Reverberating chambers as sources of stochastic electromagnetic fields," *IEEE Transactions on Electromagnetic Compatibility*, vol. 38, no. 3, pp. 348–356, 1996.
- [10] V. Mariani Primiani and F. Moglie, "Numerical simulation of reverberation chamber parameters affecting the received power statistics," *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 3, pp. 522–532, 2012.
- [11] C. Lemoine, P. Besnier, and M. Drissi, "Using autoregressive models to estimate the number of independent samples available with stirrer rotation in reverberation chamber," in *Electromagnetic Compatibility, 2007. EMC Zurich 2007. 18th International Zurich Symposium on*. IEEE, 2007, pp. 469–472.