Medical Image Analysis 40 (2017) 44–59

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Statistical characterization of noise for spatial standardization of CT scans: Enabling comparison with multiple kernels and doses

Gonzalo Vegas-Sánchez-Ferrero[a,b,*], Maria J. Ledesma-Carbayo[b], George R. Washko[a], Raúl San José Estépar[a]

[a] Applied Chest Imaging Laboratory (ACIL), Brigham and Women's Hospital, Harvard Medical School, 1249, Boylston St., Boston, MA 02115 USA
[b] Biomedical Image Technologies Laboratory (BIT), ETSI Telecomunicacion, Universidad Politecnica de Madrid, and CIBER-BBN, Madrid, Spain

ARTICLE INFO

Article history:
Received 2 November 2016
Revised 16 March 2017
Accepted 3 June 2017
Available online 7 June 2017

Keywords:
Computerized tomography
Non-stationary noise
Statistical characterization

ABSTRACT

Computerized tomography (CT) is a widely adopted modality for analyzing directly or indirectly functional, biological and morphological processes by means of the image characteristics. However, the potential utilization of the information obtained from CT images is often limited when considering the analysis of quantitative information involving different devices, acquisition protocols or reconstruction algorithms. Although CT scanners are calibrated as a part of the imaging workflow, the calibration is circumscribed to global reference values and does not circumvent problems that are inherent to the imaging modality. One of them is the lack of noise stationarity, which makes quantitative biomarkers extracted from the images less robust and stable. Some methodologies have been proposed for the assessment of non-stationary noise in reconstructed CT scans. However, those methods focused on the non-stationarity only due to the reconstruction geometry and are mainly based on the propagation of the variance of noise throughout the whole reconstruction process. Additionally, the philosophy followed in the state-of-the-art methods is based on the reduction of noise, but not in the standardization of it. This means that, even if the noise is reduced, the statistics of the signal remain non-stationary, which is insufficient to enable comparisons between different acquisitions with different statistical characteristics. In this work, we propose a statistical characterization of noise in reconstructed CT scans that leads to a versatile statistical model that effectively characterizes different doses, reconstruction kernels, and devices. The statistical model is generalized to deal with the partial volume effect via a localized mixture model that also describes the non-stationarity of noise. Finally, we propose a stabilization scheme to achieve stationary variance. The validation of the proposed methodology was performed with a physical phantom and clinical CT scans acquired with different configurations (kernels, doses, algorithms including iterative reconstruction). The results confirmed its suitability to enable comparisons with different doses, and acquisition protocols.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Quantitative imaging (QI) is the process of reducing functional, biological and morphological processes to a measurable quantity by means of medical imaging. The uses of QI are even greater in the light of a new healthcare delivery system that becomes more personalized and tries to tailor therapies to the underlying pathophysiology.

QI includes the development, standardization, optimization, and application of structural, functional, or molecular imaging acquisition protocols, data analyses, display methods, and reporting structures, as well as the validation of QI results against relevant biological and clinical data. This way, QI contributes to the radiological interpretation by assessing the degree of a given condition (Buckler et al., 2011; Abramson et al., 2015). QI has been adopted in clinical studies and trials to obtain more sensitive and precise endpoints. The advancement in techniques to automatically interpret and quantify medical images have been recognized by regulatory agencies that have now proposed guidelines for the qualification of image-based biomarkers to be used as valid endpoints in clinical trials (e.g. the Quantitative Imaging Biomarkers Alliance (QIBA) at www.rsna.org/qiba). The utility of quantitative imaging is somehow hampered by the lack of standardization among vendors due to the nuisances of the acquisition and reconstruction processes such as signal-to-noise ratio, spatial resolution,

* Corresponding author.
E-mail address: gvegas@bwh.havard.edu (G. Vegas-Sánchez-Ferrero).

http://dx.doi.org/10.1016/j.media.2017.06.001
1361-8415/© 2017 Elsevier B.V. All rights reserved.

slice thickness, image reconstruction algorithms among others (Mulshine et al., 2015).

Computerized tomography (CT) is recognized as a very suitable and widely adopted modality for quantitative imaging due to its high contrast and physical interpretability of the acquired signal. Uses of the quantitative imaging in CT (qCT) are the assessment of tumor size and texture (Aerts et al., 2014), calcifications (Agatston et al., 1990), emphysema (Müller et al., 1988), stenosis (Boogers et al., 2010) to name a few. In all cases, there is a reliance on the statistics of the CT signal to derive a valid quantity that captures the pathological process.

Although CT scanners are calibrated as part of the imaging workflow, the calibration is circumscribed to global reference values of air and water (Millner et al., 1978). This fact jointly to other inherent factors of the acquisition makes the CT signal more variable than desired (e.g. photon starvation, partial volume effect, beam hardening) (Hsieh, 2003). These effects are particularly important to create a quantitative metric that is consistent among vendors and free of confounding factors due to changes in patient weight and size to fulfill requirements of accuracy and precision (Uppot et al., 2007).

Among all those issues, CT noise is an important factor that has been carefully studied during the last decades at the detector level as part of the transmission process (Whiting, 2002; Whiting et al., 2006). The non-monochromatic nature of the X-ray signal, the amount of total X-ray energy defined by tube current coupled with the effects of the reconstruction and the interaction between X-ray and matter within the scanning field of view make the noise characterization in the reconstructed image a complex process. One of the main consequences of this complexity is the lack of stationarity. It is well understood that fan-beam tomography introduces nonstationary frequencies components and nonstationary noise (Zeng, 2004) by the nature of the scanning geometry. Several methods have been developed to assess the CT noise spectrum in nonstationary conditions (Borsdorf et al., 2008a; Balda et al., 2010; Baek and Pelc, 2010).

Borsdorf et al. (2008a) proposed a non-stationary estimation of noise based on an analytical propagation of the variance throughout the whole reconstruction process (involving interpolations, convolutions, and backprojection) (Borsdorf et al., 2008b). This method provides an estimate of the variance of noise which is further decomposed into vertical and horizontal components to get the anisotropic behavior of noise in the fan-beam reconstruction. The main limitations of this method are the need of a calibrated physical noise model to estimate the noise variance in the fan-beam projections and the need of the raw fan-beam projections that are not typically available after reconstruction. Besides, although the variance is split into two different components (vertical and horizontal), they are assumed to be independent and thus, the method does not provide a truly anisotropic description of noise. The anisotropic limitation is partially avoided in Borsdorf et al. (2009), where the preferred direction is estimated as the direction with the strongest correlation during the backprojection of variance contributions.

It is important to note that the philosophy adopted in Borsdorf et al. (2008a, 2009) is not to provide a comparable level of noise between regions of the same image or even different acquisitions but to reduce the noise from direction estimates of its variance. This means that even with a noise reduction, the non-stationary behavior of noise remains active in the filtered image. This fact evidences that reducing non-stationary noise does not provide a solution for quantitative CT analysis in terms of enabling comparisons between different acquisitions that may show different statistical characterization.

A different approach was adopted by Balda et al. (2010) to provide not only the propagation of noise variance throughout the reconstruction process but also the noise power spectrum. The statistical model that describes the attenuation levels is assumed to be known (following a model or by calibration measurements). Then, certain equally distributed noise is generated with the parameters of the location under study that is reconstructed. The result is a patch with a stationary noise with the estimated noise power spectrum. This approach strongly depends on the underlying noise model adopted for the attenuation observed in the detector. Thus, the noise power spectrum can be substantially biased in scenarios where the underlying model differs due to the response of polychromatic X-ray beams. Besides, the fan-beam projections are required to estimate the noise power spectrum. This methodology provides a way to compare acquisition protocols for CT scanners from different manufacturers when comparing the reconstruction over the same phantom (matching different reconstruction kernels). However, it does not provide a suitable way to enable comparisons between different acquisitions.

Recently, Kim et al. (2016) proposed a methodology based on the IMPACT iterative reconstruction algorithm (Man et al., 2001). That reconstruction method allows the calculation of the local variance in each iteration that can be used to transform the non-stationary noise to a more treatable one. A functional relationship between local variance and local mean is imposed by considering a *conversion factor* defined as the local mean divided by the local variance, which is multiplied to the reconstructed image. Then, the resulting random variable is assumed to have a linear dependence, which can be used to transform the noise distribution to a Gaussian one. Finally, any optimal filter for stationary Gaussian noise can be applied for noise reduction and the transformation is inverted to obtain the denoised image.

The method proposed by Kim et al. (2016) offers an interesting perspective to deal with non-stationary noise. However, it requires the raw data projections before the reconstruction, which are not usually available in standard clinical routine. Besides, the calculation of local statistics is also required. This calculation was done considering identically distributed samples in the local region. This is a strong assumption that obviously provides biased results in locations with different attenuation levels, especially at the edges.

Summarizing, the methods mentioned above show common inconveniences to provide a unifying framework to enable comparisons between images acquired in different conditions:

- They require the projection information from the detectors. That information is not available in all studies.
- They are not designed to address other sources of nonstationarity like changes in the transmission medium due to different body weight distributions.
- They are focused on the noise reduction. The resulting images are still not comparable after noise reduction; the noise remains non-stationary.
- They do not provide a statistical characterization of noise after processing.

In this paper, we propose a complete methodology to avoid these limitations. With this aim, we circumvent the need of projection information by proposing a statistical characterization of noise in the images after reconstruction. The proposal is supported by a statistical exploratory analysis of reconstructed images with different configurations including dose, reconstruction kernels, and manufacturers. As a result, we propose a non-central Gamma as the probabilistic distribution that describes the statistical behavior of noise in all the different situations.

The probabilistic distribution is used to define a statistical mixture model that allows us to account for the partial volume effect in the description of local noise characteristics. The full methodology to estimate the local probabilistic characterization of noise throughout the image is also derived.

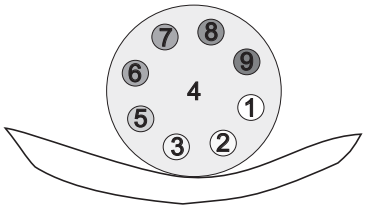

Fig. 1. Scheme of the Cylindrical LSCT 0001 phantom studied. The legend specifies in descending order the mean attenuation levels per material.

Finally, we propose a suitable variance stabilization approach for a mixture of non-central Gamma distributions. Our transformation shows some significant advantages: i) it naturally avoids biases in the estimation of local statistics due to the local characterization of noise; ii) the resulting image shows a stationary noise with known statistical distribution.

The suitability of the proposed methodology is validated with a physical phantom and with clinical images, all of them acquired with different configurations. Our method exhibits a homogeneous response of noise after the stabilization. Besides, the comparison between low dose and high dose shows that, after the proposed stabilization, the local histograms significantly increase their similarities in all the cases. This results confirm that the proposed methodology provides a unifying framework allowing for comparing images acquired with several acquisition parameters. The method was also tested for iterative reconstruction. The results also confirmed its suitability to enable comparisons with different doses, and acquisition protocols.

The paper is structured as follows: Section 2 presents the exploratory data analysis of a phantom acquired with different kernels, doses, and devices. The exploratory analysis carefully tests the functional relationships between the statistics of noise for varying attenuation levels, reconstruction kernels, and doses. The descriptive analysis allows defining some features that a probabilistic distribution should meet to describe the noise statistics (Section 2.3). The proposed distribution is then generalized to deal with heterogeneous tissues using a mixture model, the local estimation of the mixture model that fulfills the distribution features is fully derived in Section 3. The variance stabilizing transformation is presented in Section 4. The validation of the proposed methodology is given in Section 5. Finally, in Section 6 we conclude.

2. Descriptive statistics of attenuation in CT scans

In this section, we perform an exploratory analysis of the statistics of reconstructed CT scans. Our purpose is to provide both qualitative and quantitative behaviors of the statistical distributions that describe the noise in reconstructed images with different reconstruction kernels, doses, and devices. The resulting analysis will provide the evidence for the proposal of a versatile family of distributions that models the overall behavior of noise for different reconstruction kernels, doses, and devices.

2.1. Materials

The exploratory analysis was performed considering the 8-step linearity LSCT 0001 phantom (Kyoto Kagaku, Japan) acquired with different devices and reconstruction kernels. The phantom is schematically described in Fig. 1. It consists of a cylindrical structure made of a homogeneous material that contains other eight concentric tubular structures with different attenuation levels.

The phantom was acquired with two different devices with the following reconstruction protocol:

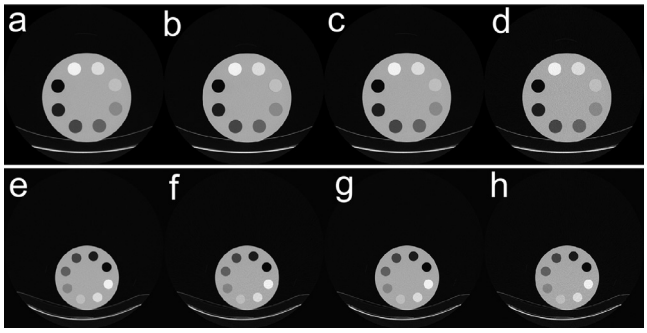

Fig. 2. Axial view of the acquired 3D volumes with different devices (Top: General Electric; Bottom: Siemens) for different kernels and doses. a) STD HD, b) STD LD, c) BONE HD, d) BONE LD, e) B31f HD, f) B31f LD, g) B45f HD, h) B45f LD.

- **General Electric Discovery STE.** Four volumes of size 512 × 512 × 313 were acquired at Brigham and Women's Hospital with different doses (400 mA and 100 mA) and reconstruction kernels (Standard, Bone). All of them with a KVP: 120 kV, Slice thickness 0.625, with software 07MWDVCT36.4. We will refer to these volumes as STD HD, STD LD, BONE HD, BONE LD for the different arrangements of kernels, and doses (HD: high dose; LD: low dose).
- **Siemens Definition.** Similarly, four volumes of size 512 × 512 × 313 were acquired at Brigham and Women's Hospital for the same configurations of doses (400 mA and 100 mA) and reconstruction kernels B31f, B45f. All of them with a KVP: 120 kV, Slice thickness 0.75, with software syngo CT 2007C. Following the same convention as before, we will refer to them as B31f HD, B31f LD, B45f HD, B45f LD.

Fig. 2 shows an example of the acquired images for all the kernels, devices, and doses considered.

2.2. Exploratory data analysis

The analysis of characteristics of noise in CT scans is performed by following the three main strategies proposed by Jones (1986): i) graphical representation, ii) quantitative evidence when possible, and iii) search for simplicity. Thus, to provide an accurate description of the statistical behavior of attenuation levels, the study of data was performed in a set of samples collected from each tissue identified by the numbered regions from 1 to 9 of the CT images as shown in Fig. 1. The samples were acquired by manually selecting a circular area in the axial view laying within each tissue type. More than 20,000 samples were acquired in each region.

In Fig. 3, we show the behavior of the estimated probability density functions (PDFs) obtained from the samples for the different configurations (represented as black dots). At first sight, the distributions look symmetrically centered on their expected value, although a positive skewness appears as the attenuation levels decrease. This effect is due to the lower limit of the attenuation, which is around −1000 HUs. It is also worth to mention that the variance of noise increases as the attenuation levels of the regions increase, which is observed in the way the PDFs decrease their height as their central attenuation level increases. This effect was already observed by Li et al. (2004) for the study of the relationship between variance and mean in the receiving sensors. The precise functional relationship will be carefully studied for each tissue type in terms of its moments until fourth order.

Analysis of mean and variance. The functional relationship between mean and variance is studied by representing the variance with respect to the average attenuation levels for the regions considered. In Fig. 4, the sample variance is plotted against the sample

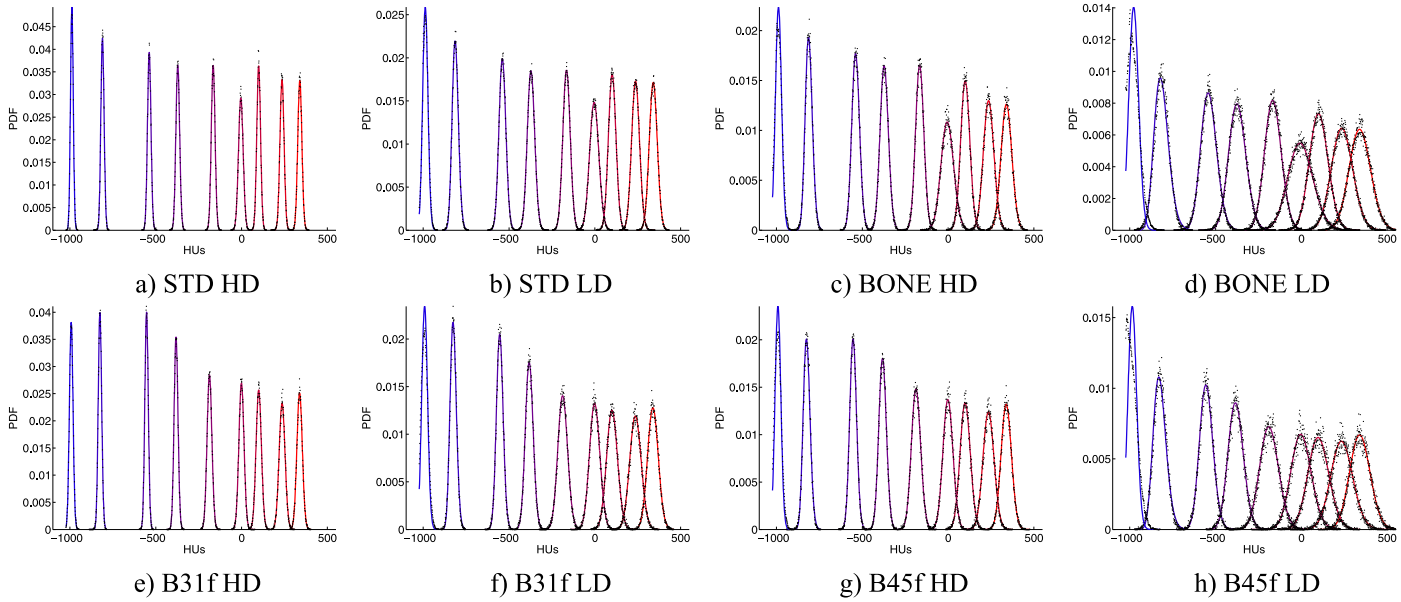

Fig. 3. Empirical probability density functions of samples for the different tissue regions and different arrangements of devices, kernels, and doses. Top: GE, Bottom: Siemens.

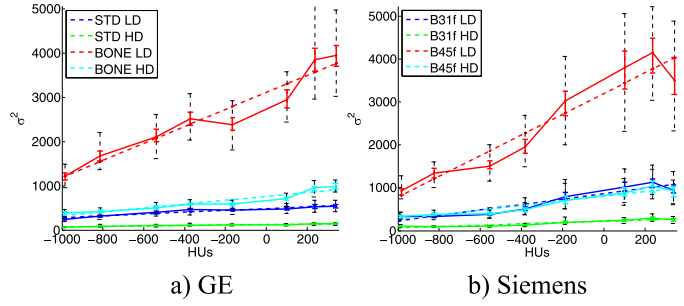

Fig. 4. Functional relationship between variance and mean and its regression line.

Table 1
Results for the linear regression between variance and mean. All the results are strongly significant and confirm that more than 70% of variance is explained by linear regression.

	Pearson's Coef.	R^2	p-value	Hyp.
STD LD	0.8686	0.7544	$< 10^{-4}$	H_1
STD HD	0.8593	0.7384	$< 10^{-4}$	H_1
BONE LD	0.9312	0.8671	$< 10^{-4}$	H_1
BONE HD	0.9133	0.8342	$< 10^{-4}$	H_1
B31f LD	0.8971	0.8047	$< 10^{-4}$	H_1
B31f HD	0.8613	0.7419	$< 10^{-4}$	H_1
B45f LD	0.9097	0.8276	$< 10^{-4}$	H_1
B45f HD	0.8815	0.7771	$< 10^{-4}$	H_1

mean jointly to its regression line. At first glance, the monotonic relationship between variance and mean becomes evident and seems to behave linearly. This is confirmed with an F-test for the null hypothesis *"H_0: The variance and mean do not have a linear relationship."* The results obtained for the regression are shown in Table 1, where the Pearson's correlation coefficient shows a strong linear relationship between mean and variance. Additionally, the coefficient of determination R^2, that accounts for the explained variance shows that more than 75% of the variance is explained by a linear model. Finally, the p-value obtained for the F-test confirms with a strong evidence (p-value $< 10^{-4}$) that both parameters –mean and variance– exhibit a linear relationship and, thus, the null hypothesis can be rejected.

Analysis of Skewness. The skewness for the different tissues was studied by testing the unbiased Fisher-Pearson skewness statistic (Groeneveld and Meeden, 1984), G_1, defined as:

$$G_1 = \frac{n}{(n-1)(n-2)} \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}}, \quad (1)$$

where $\bar{x}$ is the sample mean. If the data follows a Gaussian distribution, this statistic is distributed as a zero-mean Normal distribution with variance (Kendall and Stuart, 1969):

$$\mathrm{Var}\{G_1\} = \frac{6n(n-1)}{(n-2)(n+1)(n+3)} \xrightarrow[n\to\infty]{} \frac{6}{n}. \quad (2)$$

Thus, we can easily test the null hypothesis *"H_0: The data follows a Gaussian distribution"*, by calculating the p-value for the obtained G_1 statistic. In our case, since the number of samples is high, $\mathrm{Var}\{G_1\} \approx 6/n$, so $G_1/\sqrt{6/n}$ follows a standard Gaussian distribution, i.e. $G_1/\sqrt{6/n} \sim \mathcal{N}(0,1)$. The results obtained for the test are shown in Table 2, where the value of $|G_1|$ is represented as well as the hypothesis adopted according to the obtained p-value. Note that the null hypothesis can be rejected with a strong significance for the lowest density tissues (shown in bold letters), which corresponds to the range $[-1000, -898]$ HUs that can be related to air and lung parenchyma.

Analysis of Kurtosis. An analysis of kurtosis allows us to confirm if the distribution of the observed samples may be considered as Gaussian distributed when it is combined with the analysis of skewness. We proceed in a similar way as we did for the skewness with the statistic known as sample excess of kurtosis (Groeneveld and Meeden, 1984), which is defined as:

$$G_2 = \frac{(n+1)}{(n-1)(n-2)(n-3)} \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2} - \frac{3(n-1)^2}{(n-2)(n-3)}. \quad (3)$$

Under null hypothesis *"H_0: The data follows a Gaussian distribution"*, this statistic follows a zero-mean Gaussian distribution with

Table 2
Statistical test for the skewness $|G_1|$ and kurtosis G_2 statistics considering the null hypothesis *"H_0: The data follows a Gaussian distribution"* for different significance levels (H_1: p-value < 0.05, H_1 *: p-value < 0.01, H_1 **: p-value $< 10^{-3}$, H_1 ***: p-value $< 10^{-4}$). Note that the skewness for tissues associated with lower density values shows a very strong evidence of asymmetric behavior. Additionally, the null hypothesis is always rejected for a positive excess of kurtosis.

Skewness	B31f exp		B31f insp		B45f exp		B45f insp		STD exp		STD insp		BONE exp		BONE insp	
Tissue	$\|G_1\|$	Hyp.	$\|G_1\|$	Hyp.	$\|G_1\|$	Hyp.	$\|G_1\|$	Hyp.	G_2	Hyp.	G_2	Hyp.	G_2	Hyp.	G_2	Hyp.
1	0.011	H_0	0.021	H_0	0.031	H_0	0.042	H_1 *	0.020	H_0	0.011	H_0	0.001	H_0	0.003	H_0
2	0.016	H_0	0.023	H_0	0.005	H_0	0.033	H_1	0.012	H_0	0.005	H_0	0.002	H_0	0.030	H_0
3	0.003	H_0	0.014	H_0	0.005	H_0	0.012	H_0	0.011	H_0	0.058	H_1 **	0.004	H_0	0.013	H_0
4	0.025	H_0	0.001	H_0	0.024	H_0	0.007	H_0	0.009	H_0	0.025	H_0	0.010	H_0	0.004	H_0
5	0.014	H_0	0.004	H_0	0.018	H_0	0.003	H_0	0.012	H_0	0.024	H_0	0.029	H_1 *	0.003	H_0
6	0.021	H_0	0.015	H_0	0.019	H_0	0.017	H_0	0.009	H_0	0.010	H_0	0.005	H_0	0.024	H_0
7	0.008	H_0	0.008	H_0	0.011	H_0	0.011	H_0	0.008	H_0	0.066	H_1 ***	0.019	H_0	0.023	H_0
8	0.003	H_0	0.083	H_1 ***	0.004	H_0	0.058	H_1 **	0.005	H_0	0.079	H_1 ***	0.011	H_0	0.026	H_0
9	**0.434**	H_1 ***	**0.071**	H_1 ***	**0.810**	H_1 ***	**0.437**	H_1 ***	**0.199**	H_1 ***	**0.079**	H_1 ***	**0.706**	H_1 ***	**0.359**	H_1 ***

Kurtosis	B31f exp		B31f insp		B45f exp		B45f insp		STD exp		STD insp		BONE exp		BONE insp	
Tissue	$\|G_1\|$	Hyp.	$\|G_1\|$	Hyp.	$\|G_1\|$	Hyp.	$\|G_1\|$	Hyp.	G_2	Hyp.	G_2	Hyp.	G_2	Hyp.	G_2	Hyp.
1	−0.007	H_0	0.004	H_0	0.009	H_0	0.004	H_0	−0.035	H_0	0.056	H_0	−0.009	H_0	0.062	H_0
2	0.007	H_0	−0.002	H_0	**0.081**	H_1	0.034	H_0	0.012	H_0	**0.132**	H_1 ***	0.009	H_0	0.059	H_0
3	**0.079**	H_1	**0.073**	H_1	**0.083**	H_1	**0.063**	H_1	**0.045**	H_1	**0.084**	H_1 *	0.024	H_0	**0.165**	H_1 ***
4	0.061	H_0	0.030	H_0	**0.068**	H_1	−0.018	H_0	−0.033	H_0	0.060	H_0	0.003	H_0	**0.071**	H_1
5	**0.081**	H_1	0.053	H_0	0.044	H_0	0.018	H_0	−0.027	H_0	**0.073**	H_1	0.030	H_0	**0.088**	H_1 *
6	**0.091**	H_1 *	0.011	H_0	**0.070**	H_1	0.000	H_0	0.027	H_0	0.041	H_0	0.035	H_0	−0.012	H_0
7	0.057	H_0	0.049	H_0	0.053	H_0	0.030	H_0	−0.013	H_0	0.051	H_0	**0.044**	H_1	0.012	H_0
8	0.038	H_0	**0.151**	H_1 ***	0.047	H_0	**0.110**	H_1 **	**0.062**	H_1 *	**0.232**	H_1 ***	**0.068**	H_1 *	**0.096**	H_1 *
9	−0.151	H_0	−0.065	H_0	**0.326**	H_1 ***	−0.128	H_0	−0.165	H_0	0.009	H_0	**0.215**	H_1 ***	−0.170	H_0

variance (Groeneveld and Meeden, 1984; Joanes and Gill, 1998):

$$\mathrm{Var}\{G_2\} = \frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)} \xrightarrow[n\to\infty]{} \frac{24}{n}. \quad (4)$$

Thus, for a large number of samples –as it is the case under study– we can assume that $\frac{G_2}{\sqrt{24/n}} \sim \mathcal{N}(0,1)$. We can now calculate the p-values to contrast the null hypothesis for all tissue types. The results obtained for the excess of kurtosis are also shown in Table 2. In this case, the conclusion is more subtle than in the previous case but still useful. Note that all the cases where the null hypothesis can be rejected are due to a positive excess of skewness (shown in bold letters). This means that the data behaves as a leptokurtic distribution whose excess of kurtosis converges to 0 as the mean value increases. This tendency would suggest that the kurtosis-skewness relationship could show a monotonic relationship that may be efficiently described by a family of distributions.

Relationship between Kurtosis and Skewness. The way the empirical distributions converge to a Normal one as the attenuation levels increase can be studied by analyzing the relationship between kurtosis and skewness. The Skewness-Kurtosis plot (Cullen and Frey, 1999) is a representation of the locus of skewness-kurtosis pairs (specifically, kurtosis vs. skewness^2) and allows us to compare the empirical behavior of the data with theoretical distributions. In general, theoretical distributions may be represented as points, lines or surfaces as the parameters of the distribution vary and depending on the kurtosis-skewness relationship. This way, the Gaussian distribution (with null skewness and kurtosis $= 3$) would be represented as a point at location (0, 3), while the Gamma distribution is represented by the line $y = 1.5x + 3$.

In Fig. 5, we show the Skewness-Kurtosis plot for the most extreme cases of each device (soft kernel+HD, sharp kernel+LD), the other ones are similar and were omitted for brevity. The different tissues of the phantom are represented with different colors (darker corresponds to higher attenuation levels). Note that the overall behavior of samples seems to increase linearly and there is a clear convergence to the Gaussian distribution –represented at location (0, 3)– as attenuation levels increase. This result confirms

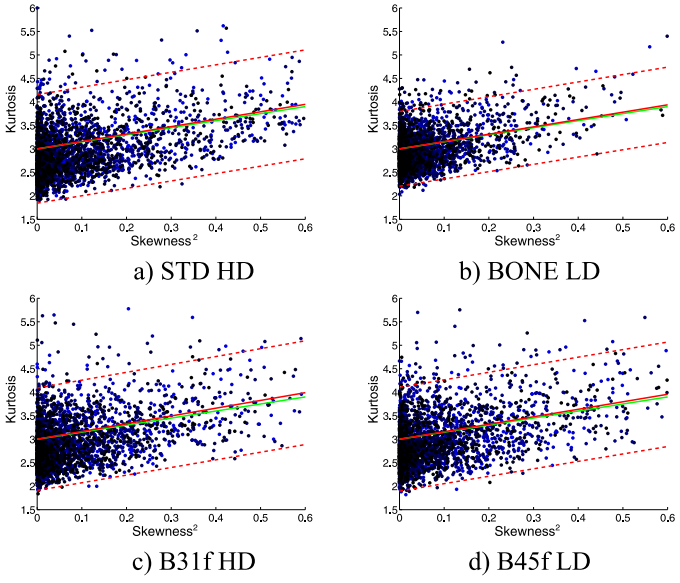

Fig. 5. Skewness-Kurtosis plots for the most extreme cases of each device. Low attenuation levels are represented in brighter blues, whereas higher attenuation levels are represented in darker blues. The regression line constrained to the Gaussian convergence is represented in red and its confidence interval as a dashed line. The green line represents the loci of the *Gamma* distribution defined as $y = 1.5x + 3$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the already mentioned convergence of the skewness to 0 and kurtosis to 3.

Additionally, the leptokurtic behavior of the data is also confirmed by the overall linear increase shown in the figure. To see if this linear increase is significant, we performed a linear regression analysis of the Skewness-Kurtosis pairs constrained to the convergence to (0, 3) previously mentioned. The regression line is also represented in Fig. 5 as the red continuous line and its confidence interval is shown as dashed red lines. The results obtained for the regression analysis are shown

Table 3
Regression analysis for Kurtosis-Skewness pairs constrained to the Gaussian convergence. The evidence of a linear relationship is clear due to the *p*-value $< 10^{-4}$ for all the cases. The Γ behavior of Kurtosis-Skewness pairs cannot be discarded for almost all the cases, obtaining an increment of the total sums of squares of less than a 1%.

	Reg. line	R^2	H_0	H_Γ	ΔTTS%
STD LD	y = 1.47 x + 3	0.1063	$< 10^{-4}$	0.7650	$< 10^{-3}$ %
STD HD	y = 1.58 x + 3	0.3129	$< 10^{-4}$	0.1331	0.12 %
BONE LD	y = 1.56 x + 3	0.1295	$< 10^{-4}$	0.4841	0.03 %
BONE HD	y = 1.46 x + 3	0.2469	$< 10^{-4}$	0.4795	0.03 %
B31f LD	y = 1.56 x + 3	0.2142	$< 10^{-4}$	0.4128	0.03 %
B31f HD	y = 1.65 x + 3	0.3360	$< 10^{-4}$	0.0024	0.42 %
B45f LD	y = 1.59 x + 3	0.2310	$< 10^{-4}$	0.1520	0.10 %
B45f HD	y = 1.61 x + 3	0.3125	$< 10^{-4}$	0.0323	0.21 %

in Table 3. The *p*-values for the null hypotheses H_0: *"There is no linear relationship between kurtosis and skewness²"* and H_Γ : *"The Skewness-Kurtosis are distributed as y = 1.5 x + 3"* are also given jointly with the increment of variance (ΔTTS) due to the approximation of the Gamma Skewness-Kurtosis distribution calculated as:

$$\Delta\text{TTS} = \frac{\sum_i^n (y_i - 1.5x - 3)^2}{\sum_i^n (y_i - \widehat{m}x - 3)^2} - 1, \tag{5}$$

where $\widehat{m}$ accounts for the slope estimation of the regression line.[1] Results in Table 3 show that the null hypothesis H_0 can be discarded with a strong evidence (*p*-values $< 10^{-4}$) for all the cases and the regression lines obtained are close to the one that represents the Gamma distribution. Actually, the test for hypothesis H_Γ shows that there is a very low evidence for discarding the hypothesis for almost all the cases. The only case where the slope significantly differs is the case of B31f HD. However, it is important to note that the increase of error due to the approximation $1.5x + 3$ is 0.42%.

2.3. Proposal

According to the results obtained in the descriptive analysis of data, a proper family of distributions that model the statistical behavior of tissue/noise in CT scans should meet the following features:

1. Positive skewness for low attenuation levels that gradually decreases.
2. Leptokurtic behavior for low attenuation levels that gradually normalizes for higher attenuation levels.
3. Linear relationship between mean and variance.
4. Linear relationship between kurtosis and skewness² with convergence to the Gaussian distribution.

In the light of these characteristics, we propose the non-central Gamma (nc-Γ) distribution defined as a three-parameter distribution:

$$f_X(x|\alpha,\beta,\delta) = \frac{(x-\delta)^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-\frac{x-\delta}{\beta}}, \quad x \geq \delta \text{ and } \alpha,\beta > 0, \tag{6}$$

where $\Gamma(x)$ is the Euler Gamma function defined as $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$, for $x > 0$; α is the shape parameter, β is the scale parameter, and δ is defined as the least attenuation level (typically around −1000 HUs).[2] This distribution meets the whole set of characteristics observed in the exploratory analysis:

1. It has a positive skewness defined as:

$$\text{Skewness}\{\Gamma(\alpha,\beta,\delta)\} = \frac{2}{\sqrt{\alpha}} > 0 \quad \text{for } \alpha > 0 \tag{7}$$

2. The excess of kurtosis is positive and gradually converges to a Gaussian distribution as the mean increases:

$$\text{Excess Kurtosis}\{\Gamma(\alpha,\beta,\delta)\} = \frac{6}{\alpha} > 0 \quad \text{for } \alpha > 0 \tag{8}$$

3. The variance and mean of the nc-*Gamma* variable (i.e. $\mu_X = \alpha\beta + \delta$ and $\sigma_X^2 = \alpha\beta^2$) shows a linear relationship:

$$\sigma_X^2 = \beta\mu_X - \beta\delta \tag{9}$$

4. The linear relationship between kurtosis and skewness² is the same as the observed data:

$$\text{Kurtosis}\{\Gamma(\alpha,\beta,\delta)\} = \frac{3}{2}\text{Skewness}^2\{\Gamma(\alpha,\beta,\delta)\} + 3 \tag{10}$$

In Fig. 3 we show the nc-Gamma distributions (continuous lines) fitted with the maximum likelihood method. Additionally, in Table 4 we show the improvement obtained in the fitting when the Gaussian and a nc-Gamma are compared by means of the Kolmogorov–Smirnov distance of their cumulative distribution function (CDF) calculated as:

$$D_{\text{ks}}(F, F_X) = \max_x |\widehat{F}(x) - F_X(x)| \tag{11}$$

where $\widehat{F}(x)$ is the empirical CDF and F_X is the theoretical distribution fitted to the data. The improvement was calculated as $D_{\text{ks}}(F, F_{\text{Gamma}})/D_{\text{ks}}(F, F_{\text{Gaussian}})$. These results show the better performance of the nc-Gamma distribution even for higher levels of attenuation where the empirical distribution shows an increasing convergence to the Gaussian distribution. Note that the there is always an important improvement which, in some cases, reaches 70%.

3. Localized mixture model for CT images

The limited resolution of CT scans and the heterogeneous nature of some tissues cause the so-called partial volume effect. This effect is shown as an average attenuation level of the different compounding tissues within the pixel resolution. To model properly the combined effect of different tissues a suitable combination of probabilistic models should be applied. We propose to model the partial volume effect as a Mixture Model of Probabilistic distributions. This model generalizes the statistical description of noise when the partial volume effect is considered (Vegas-Sanchez-Ferrero et al., 2014).

As CT scans provide normalized units of the linear attenuation coefficients (Hounsfield units), one can expect some average attenuation levels in clinical scans such as air, fat, water, muscle, blood, bone, etc. These tissues are well characterized and the calibration of CT scans involve the tuning of some of these attenuation levels according to a calibration phantom such as the one shown in Fig. 2.

The prior knowledge about the preexisting tissues in clinical CT scans lead to a more realistic definition of the heterogeneous composition of tissues and can be incorporated in the mixture model. Therefore, attending to the statistical properties of noise for the attenuation levels studied, we propose a mixture model whose components follow nc-Γ distributions and their mean values are located at the average attenuation levels of the expected tissues of clinical CTs (i.e. air, fat, water, muscle, blood, bone, etc.). In our case, and without loss of generality, we will consider those nine attenuation levels shown in Fig. 1. Thus, the nc-Gamma mixture

[1] Note that ΔTTS ∈ [0, 1] since $\widehat{m} = \arg\min_m \sum_{i=1}^n (y_i - mx + 3)^2$.

[2] For the sake of brevity and without introducing confusion, we will apply the same notation $\Gamma(\cdot)$ to the following concepts depending on the parameters used: Euler Gamma $\Gamma(x)$, central Gamma distribution $\Gamma(x|\alpha,\ \beta)$, central Gamma random variable $\Gamma(\alpha,\ \beta)$, non-central Gamma distribution $\Gamma(X|\alpha,\ \beta,\ \delta)$, and non-central Gamma random variable $\Gamma(\alpha,\ \beta,\ \delta)$. Additionally, the notation used from here forth refers to random variables in capital letters and samples of random variables in lower case letters. The expectation operator is denoted as $E\{\cdot\}$.

Table 4
Improvement of the goodness of fit for the nc-Gamma distribution with respect the Gaussian distribution calculated as the ratio of the Kolmogorov–Smirnov distance, $D_{ks}(F, F_{Gamma})/D_{ks}(F, F_{Gaussian})$, and represented as mean (standard deviation).

	B31f LD	B31f HD	B45f LD	B45f HD		STD LD	STD HD	BONE LD	BONE HD
1	0.7621 (0.1968)	0.6223 (0.2342)	0.7526 (0.2299)	0.7036 (0.1999)	1	0.7047 (0.2140)	0.4744 (0.1841)	0.7720 (0.1845)	0.7238 (0.1899)
2	0.7193 (0.1851)	0.6558 (0.2303)	0.7641 (0.2128)	0.7714 (0.1870)	2	0.6776 (0.2117)	0.4661 (0.1885)	0.7670 (0.2002)	0.7157 (0.2316)
3	0.7176 (0.1811)	0.6297 (0.2411)	0.7366 (0.2028)	0.7149 (0.2134)	3	0.6377 (0.2057)	0.4575 (0.2046)	0.7471 (0.1978)	0.7099 (0.1940)
4	0.6847 (0.2206)	0.5911 (0.2054)	0.7205 (0.2165)	0.7348 (0.2055)	4	0.7138 (0.2085)	0.5041 (0.2052)	0.7896 (0.2207)	0.7294 (0.1999)
5	0.7344 (0.2048)	0.5497 (0.2293)	0.7506 (0.2197)	0.7088 (0.2309)	5	0.7009 (0.2172)	0.4723 (0.2156)	0.8079 (0.2194)	0.6916 (0.2140)
6	0.6441 (0.2385)	0.4450 (0.1791)	0.7279 (0.2287)	0.6277 (0.2074)	6	0.6737 (0.2211)	0.4583 (0.2137)	0.7622 (0.2182)	0.6987 (0.2180)
7	0.6378 (0.2226)	0.3955 (0.1850)	0.7160 (0.2162)	0.5907 (0.2186)	7	0.6403 (0.2273)	0.3965 (0.1834)	0.7622 (0.2054)	0.6429 (0.2317)
8	0.6131 (0.2011)	0.4218 (0.1663)	0.7389 (0.2451)	0.6905 (0.2005)	8	0.6163 (0.2238)	0.4542 (0.1951)	0.7509 (0.2374)	0.6497 (0.2287)
9	0.5996 (0.1610)	0.4057 (0.1830)	0.8123 (0.0572)	0.6107 (0.1529)	9	0.4965 (0.1937)	0.3033 (0.1538)	0.9604 (0.0400)	0.7534 (0.1506)

model (nc-ΓMM) proposed is comprised of a maximum of nine components.

We pursue a localized statistical model that may provide the probability of belonging to each of the considered tissue classes per pixel. To estimate this localized model, we propose a global-local approach that allows:

First, to fit the overall statistical behavior using the Expectation-Maximization method to establish suitable priors that can be used for the local refinement; Second, use the global priors as an initial condition for the local approximation, in which a mixture model is used to describe the contribution of each tissue class per pixel.

An additional and important characteristic of the localized mixture model is that it allows removing components whose importance within a local neighborhood becomes negligible. This is a critical issue since otherwise a component which is not present in the local neighborhood would try to fit the histogram even though there is no presence of that tissue in it, giving rise to a non-realistic description of the heterogeneous nature of tissue.

3.1. Global estimate

Let $\mathbf{X} = \{x_i\},\ 1 \le i \le N$ be a set of samples (pixel intensities) of a given region of the CT image. We assume that these samples are independent and identically distributed (IID) random variables (RVs).

The nc-ΓMM considers that these variables result from the contribution of J distributions:

$$p(x|\Theta) = \sum_{j=1}^{J} \pi_j f_X(x|\Theta_j), \tag{12}$$

where Θ_j are the parameters of the PDF. The nc-Gamma PDF is defined as in Eq. (6). We set the values $\delta_j = \delta$ HUs for each j where δ is the minimum value received in the image (usually around -1000 HUs).

Under these conditions and without loss of generality, we will consider the transformation $Y = X - \delta$. This change of variable transforms, by virtue of the formula for the change of random variables (Kendall and Stuart, 1969), a $\Gamma(\alpha, \beta, \delta)$ random variable into its central counterpart with same parameters $\Gamma(\alpha, \beta)$. Therefore the nc-ΓMM, in its turn, becomes a central ΓMM with PDF:

$$p(y) = \sum_{j=1}^{J} \pi_j f_Y(y|\Theta_j), \tag{13}$$

where $\Theta_j = \{\alpha_j, \beta_j\}$ and

$$f_Y(y|\alpha_j, \beta_j) = \frac{y^{\alpha_j - 1}}{\beta_j^{\alpha_j}\Gamma(\alpha_j)} e^{-\frac{x}{\beta_j}}. \tag{14}$$

So, considering the set of samples $\mathbf{Y} = \{y_i\} = \{x_i - \delta\},\ 1 \le i \le N$, the joint distribution of IID samples is given by:

$$p(\mathbf{Y}|\Theta) = \prod_{i=1}^{N} p(y_i|\Theta), \tag{15}$$

where Θ is a vector of the parameters of the ΓMM $\{\pi_1, \cdots, \pi_J, \Theta_1, \cdots, \Theta_J\}$.

The Expectation-Maximization method (Moon, 1996) is applied here to maximize the log-likelihood function when certain hidden discrete random variables, $\mathbf{Z} = \{Z_i\}$, are introduced into the model. These RVs take values in $\{1, \cdots, J\}$ and they reference the membership of each sample, i.e. $Z_i = j$ means sample y_i belongs to the jth distributions with parameters $\Theta_j = \{\alpha_j, \beta_j\}$.

Now, let $\Theta^{(n)}$ be an estimate of the parameters of the mixture at the nth iteration, the expectation step is performed by calculating the expected value of the log-likelihood $\mathcal{L}(\Theta|\mathbf{Y}, \mathbf{Z})$:

$$\mathcal{Q}\left(\Theta|\Theta^{(n)}, \mathbf{Y}\right) = E_{\mathbf{Z}|\Theta^{(n)},\mathbf{Y}}\{\mathcal{L}(\Theta|\mathbf{Y}, \mathbf{Z})\}. \tag{16}$$

In the maximization step, the new estimate $\Theta^{(n)}$ is obtained by maximizing the expectation of the likelihood function $\mathcal{Q}(\Theta|\Theta^{(n)}, \mathbf{Y})$. These steps are iterated until a stop criterion such as $||\Theta^{(n+1)} - \Theta^{(n)}|| < \mathrm{TOL}$ for some pre-established threshold (TOL) is reached.

The expectation of the likelihood function with respect to the hidden RVs when data $\mathbf{Y}$ and the previous estimate $\Theta^{(n)}$ are known is:

$$\begin{aligned}
\mathcal{Q}(\Theta|\Theta^{(n)}, \mathbf{Y}) &= E_{Z|\Theta^{(n)},\mathbf{Y}}\{\mathcal{L}(\Theta|\mathbf{Y}, \mathbf{Z})\} \\
&= \sum_{i=1}^{N} E_{Z_i|\Theta^{(n)},y_i}\{\log p(y_i|\Theta_{z_i}, Z_i = z_i) + \log p(Z_i = z_i|\Theta)\} \\
&= \sum_{i=1}^{N}\sum_{j=1}^{J} p(Z_i = j|y_i, \Theta^{(n)})\left(\log p(y_i|\Theta_j) + \log \pi_j\right) \\
&= \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{J} \gamma_{i,j} \log p(y_i|\Theta_j)}_{\Lambda_{\alpha,\beta}} + \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{J} \gamma_{i,j} \log \pi_j}_{\Lambda_{\pi}},
\end{aligned} \tag{17}$$

where $p(Z_i = j|\Theta)$ is the probability of y_i to belong to the jth class, denoted as π_j, and $p(Z_i = j|y_i, \Theta^{(n)})$ is the posterior probability denoted as $\gamma_{i,j}$ and derived by the Bayes theorem:

$$p(Z_i = j|y_i, \Theta^{(n)}) = \frac{p(y_i|\Theta_j^{(n)})p(Z_i = j|\Theta^{(n)})}{p(y_i|\Theta^{(n)})}, \tag{18}$$

where, as in Eq. (12):

$$p(y_i|\Theta^{(n)}) = \sum_{j=1}^{J} p(Z_i = j|\Theta^{(n)})p(y_i|\Theta_j^{(n)}) = \sum_{j=1}^{J} \pi_j f_Y(y_i|\Theta_j^{(n)}). \tag{19}$$

Since Eq. (17) is composed of the convex sum of two independent terms with the same sign, $\Lambda_{\alpha,\beta}$ and Λ_{π}, the maximization

step can be performed independently for each term. In the case of Λ_π, the optimization is performed via *Lagrange Multipliers* to introduce the constraint $\sum_{j=1}^{J}\pi_j = 1$, which ensures the result is a well-defined probability. The Lagrange method of multipliers guarantees a necessary condition for optimality in this problem. The new *Lagrange function* with the Lagrange multiplier, λ, reads:

$$\Lambda_\pi(\lambda) = \sum_{i=1}^{N}\sum_{j=1}^{J}\gamma_{i,j}\log\pi_j + \lambda\left(\sum_{j=1}^{J}\pi_j - 1\right). \tag{20}$$

The optimization with respect π_j gives:

$$\sum_{i=1}^{N}\gamma_{i,j} = -\lambda\pi_j. \tag{21}$$

By summing both terms of the equation over j, we obtain $\lambda = -N$. Finally, the values of $\hat{\pi}_j$ that maximize the Lagrange function (and the likelihood term) are:

$$\hat{\pi}_j = \frac{1}{N}\sum_{i=1}^{N}\gamma_{i,j} = \frac{1}{N}\sum_{i=1}^{N}p(Z_i = j|\Theta). \tag{22}$$

For the optimization of the second term, $\Lambda_{\alpha,\beta}$ we impose the condition of a constant average μ_j for each tissue class, i.e. $E\{Y|\Theta_j\} = \alpha_j\beta_j = \mu_j$. This way the term to be maximized reads:

$$\Lambda_{\alpha,\beta} = \sum_{i=1}^{N}\sum_{j=1}^{J}\gamma_{i,j}\left((\alpha_j - 1)\log y_i - \frac{y_i}{\beta_j} - \alpha_j\log\beta_j - \log\Gamma(\alpha_j)\right) \tag{23}$$

subject to $E\{Y|\Theta_j\} = \alpha_j\beta_j = \mu_j$.

The maximization of Eq. (23) can be obtained by plugging $\alpha_j = \mu_j/\beta_j$ and setting the derivative equal to zero. Therefore, after some algebra, the maximization of Eq. (23) with respect to α_j gives the following result:

$$\log(\alpha_j) - \psi(\alpha_j) = \frac{\sum_{i=1}^{N}\gamma_{i,j}\frac{y_i}{\mu_j}}{\sum_{i=1}^{N}\gamma_{i,j}} - \frac{\sum_{i=1}^{N}\gamma_{i,j}\log\left(\frac{y_i}{\mu_j}\right)}{\sum_{i=1}^{N}\gamma_{i,j}} - 1. \tag{24}$$

where $\psi(x)$ is the digamma function defined as $\psi(x) = \Gamma'(x)/\Gamma(x)$.

This expression has no closed solution. However, it can be obtained by numerical methods since the function $f(x) = \log(x) - \psi(x)$ is well behaved.

From the estimated value $\widehat{\alpha}_j$ that maximizes the log-likelihood, the estimate of $\widehat{\beta}$ is directly obtained from the constraint $\alpha_j\beta_j = \mu_j$.

The method can be summarized in the Algorithm 1.

Algorithm 1 Implementation of the nc-ΓMM.

$\{y_i\}_{i=1}^{N} \leftarrow \{x_i - \delta\}$ ▷ Centering the data
$J \leftarrow$ Number of components
$\Theta^{(0)} \equiv \{\pi_j, \alpha_j, \beta_j\}_{j=1}^{J}$ ▷ Initial guess of paramters
$\gamma_{i,j}^{(0)} \leftarrow p(Z_i = j|y_i, \Theta^{(0)})$ ▷ Eq. (18)
maxIter $\leftarrow$ Maximum number of iterations
Tol $\leftarrow$ Tolerance
Err $\leftarrow \infty$
$n \leftarrow 0$ ▷ Iteration counter
while Err>Tol and n<maxIter **do**
  $n \leftarrow n + 1$
  for j = 1 to J **do**
    $\alpha_j^{(n)} \leftarrow \log(\alpha_j) - \psi(\alpha_j) = \frac{\sum_{i=1}^{N}\gamma_{i,j}\frac{y_i}{\mu_j}}{\sum_{i=1}^{N}\gamma_{i,j}} - \frac{\sum_{i=1}^{N}\gamma_{i,j}\log\left(\frac{y_i}{\mu_j}\right)}{\sum_{i=1}^{N}\gamma_{i,j}} - 1$
    $\beta_j^{(n)} = \mu_j/\alpha_j^{(n)}$
    $\pi_j^{(n)} \leftarrow 1/N\sum_{i=1}^{N}\gamma_{i,j}^{(n-1)}$
    $\gamma_{i,j}^{(n)} \leftarrow p(Z_i = j|y_i, \Theta^{(n-1)})$ ▷ Eq. (18)
  end for
  Err $\leftarrow ||\Theta^{(n)} - \Theta^{(n-1)}||/||\Theta^{(n-1)}||$
end while
return $\Theta^{(n)}$

3.2. Local estimate

The local estimation of parameters per pixel location, $\mathbf{r}$, requires a more subtle analysis of the spatial dependence of $\alpha(\mathbf{r})$ and $\beta(\mathbf{r})$. We assume that the image $X:\Omega\to\mathbb{R}$ follows a non-stationary nc-ΓMM of J components with centrality parameter δ HUs (usually around −1000 HUs), known averages $\{\mu_j\}_{j=1}^{J}$ and spatially variant shape and scale parameters. In order to simplify the notation, the centered image $Y(\mathbf{r}) = X(\mathbf{r}) - \delta$ will be considered hereafter.

We also assume that the spatial variation of both parameters is smoother than the random fluctuation of noise. This assumption allows us to consider that $X(\mathbf{r})$ is locally stationary and, thus, the estimation of parameters could be performed with samples in a local neighborhood, $\eta(\mathbf{r})$, of location $\mathbf{r}\in\Omega$.

Under these conditions, the parameters $\alpha_j(\mathbf{r})$ can be efficiently calculated by considering the first conditioned moment and the first conditioned logarithmic moment of the normalized image $Y(\mathbf{r})/\mu_j$ as follows:

$$E\left\{\frac{Y(\mathbf{r})}{\mu_j}\middle|Z(\mathbf{r}) = j\right\} = 1, \tag{25}$$

$$E\left\{\log\left(\frac{Y(\mathbf{r})}{\mu_j}\right)\middle|Z(\mathbf{r}) = j\right\} = \psi(\alpha(\mathbf{r})) - \log\alpha(\mathbf{r}). \tag{26}$$

The expected values can be approximated by their respective sample conditioned local moments (denoted with the operator $\langle\,\cdot\,|Z(\mathbf{r}) = j\rangle$) as:

$$\left\langle\frac{Y(\mathbf{r})}{\mu_j}\middle|Z(\mathbf{r}) = j\right\rangle = \frac{\sum_{\mathbf{s}\in\eta(\mathbf{r})}P(Z(\mathbf{r}) = j|Y(\mathbf{r}))Y(\mathbf{s})/\mu_j}{\sum_{\mathbf{s}\in\eta(\mathbf{r})}P(Z(\mathbf{r}) = j|Y(\mathbf{r}))}, \tag{27}$$

$$\left\langle\log\left(\frac{Y(\mathbf{r})}{\mu_j}\right)\middle|Z(\mathbf{r}) = j\right\rangle = \frac{\sum_{\mathbf{s}\in\eta(\mathbf{r})}P(Z(\mathbf{r}) = j|Y(\mathbf{r}))\log(Y(\mathbf{s})/\mu_j)}{\sum_{\mathbf{s}\in\eta(\mathbf{r})}P(Z(\mathbf{r}) = j|Y(\mathbf{r}))}. \tag{28}$$

Note that the posterior probability $P(Z(\mathbf{r}) = j|Y(\mathbf{r}))$ plays the same role as the posterior probability $\gamma_{i,j}$ of Eq. (18) with an explicit reference to the location $\mathbf{r}$. Thus, for brevity, we will denote it as $\gamma_j(\mathbf{r})$.

We are interested in those values of $\alpha(\mathbf{r})$ for which both Eqs. (25) and (26) hold. So, considering that $log(x) - \psi(x)$ is an always positive and monotonic decreasing function, and $\Sigma_k\pi_k x_k \geq \Sigma_k\pi_k\log x_k$ holds for any set of positive weights $\{\pi_i\}$ and positive samples $\{x_i\}$, we can subtract Eq. (25) from (26) and we get the following well-defined equality:

$$\log(\alpha_j(\mathbf{r})) - \psi(\alpha_j(\mathbf{r})) = \frac{\sum_{\mathbf{s}\in\eta(r)}\gamma_j(\mathbf{s})\frac{Y(\mathbf{s})}{\mu_j}}{\sum_{\mathbf{s}\in\eta(r)}\gamma_j(\mathbf{s})} - \frac{\sum_{\mathbf{s}\in\eta(r)}\gamma_j(\mathbf{s})\log\left(\frac{Y(\mathbf{s})}{\mu_j}\right)}{\sum_{\mathbf{s}\in\eta(r)}\gamma_j(\mathbf{s})} - 1. \tag{29}$$

This result is the local counterpart of the result obtained in Eq. (24). The same result could be obtained by applying the Expectation-Maximization methodology employed in the previous section with the assumption of local stationarity. In this case, we preferred this equivalent derivation to provide an explicit reference to location $\mathbf{r}$ in the derivation.

The $\beta_j(\mathbf{r})$ and $\pi_j(\mathbf{r})$ parameters can be estimated in the same way as for the global estimate as:

$$\beta_j(\mathbf{r}) = \mu_j/\alpha_j(\mathbf{r}) \tag{30}$$

$$\pi_j(\mathbf{r}) = \frac{1}{|\eta(\mathbf{r})|}\sum_{\mathbf{s}\in\eta(r)}\gamma_j(\mathbf{s}) \tag{31}$$

In conclusion, the local estimation of parameters can be obtained by following an iterative implementation shown in Algorithm 2.

Algorithm 2 Implementation of the local nc-ΓMM.

$Y(\mathbf{r}) \leftarrow X(\mathbf{r}) - \delta$ ▷ Centering the data
$J \leftarrow$ Number of components
$\Theta^{(0)}(\mathbf{r}) \equiv \{\pi_j(\mathbf{r}), \alpha_j(\mathbf{r}), \beta_j(\mathbf{r})\}_{j=1}^{J}$ ▷ Initial guess of paramters
$\gamma_j(\mathbf{r})^{(0)} \leftarrow p(Z(\mathbf{r}) = j|Y(\mathbf{r}), \Theta^{(0)}(\mathbf{r}))$ ▷ Eq. (18)
maxIter ← Maximum number of iterations
Tol ← Tolerance
Err ← ∞
$n \leftarrow 0$ ▷ Iteration counter
while Err>Tol and n<maxIter **do**
  $n \leftarrow n + 1$
  for j = 1 to J **do**
    $\alpha_j^{(n)}(\mathbf{r}) \leftarrow \log(\alpha_j(\mathbf{r})) - \psi(\alpha_j(\mathbf{r})) = \frac{\sum_{\mathbf{s}\in\eta(\mathbf{r})}\gamma_j(\mathbf{s})\frac{\mathbf{Y(s)}}{\mu_j}}{\sum_{\mathbf{s}\in\eta(\mathbf{r})}\gamma_j(\mathbf{s})} - \frac{\sum_{\mathbf{s}\in\eta(\mathbf{r})}\gamma_j(\mathbf{s})\log\left(\frac{Y(\mathbf{s})}{\mu_j}\right)}{\sum_{\mathbf{s}\in\eta(\mathbf{r})}\gamma_j(\mathbf{s})} - 1$
    $\beta_j^{(n)}(\mathbf{r}) = \mu_j/\alpha_j(\mathbf{r})$
    $\pi_j^{(0)}(\mathbf{r}) \leftarrow 1/|\eta(\mathbf{r})|\sum_{\mathbf{s}\in\eta(\mathbf{r})}\gamma_j^{(n-1)}(\mathbf{r})$
    $\gamma_j^{(n)}(\mathbf{r}) \leftarrow p(Z(\mathbf{r}) = j|Y(\mathbf{r}), \Theta^{(0)})$ ▷ Eq. (18)
  end for
  Err ← $||\Theta^{(n)} - \Theta^{(n-1)}||/||\Theta^{(n-1)}||$
end while
return $\Theta^{(n)}$

4. Variance stabilization

The variance-stabilizing transformation (VST) has been historically applied to simplify the analysis of variance of some certain random variable whose variance is related to the mean level of the measurements (Bartlett, 1947). The main goal of a VST is to compensate the change of the variance with respect to the change of the mean value –whenever this relationship is known– to provide a constant variance.

In our case, the VST formulation can be used to seek a transformation of the signal that transforms the process from non-stationary to stationary. This is achieved without an explicit knowledge of the reconstruction geometry or the attenuation medium of the X-ray.

The derivation of the VST is commonly associated with the so-called *delta method*, which links the central limit theorem with the convergence of the transformed random variable $Y = f(X)$ by a differentiable function f. Formally speaking, let X_n be a sequence of random variables that satisfies $\sqrt{n}(X_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ (i.e. convergence in distribution). Then the first order Taylor expansion of $f(X_n)$ around μ is $f(X_n) = f(\mu) + f'(\mu)(X_n - \mu)$ which, reordering terms, gives:

$$\sqrt{n}(f(X_n) - f(\mu)) = f'(\mu)\sqrt{n}(X_n - \mu). \tag{32}$$

Since X_n converges in distribution to a constant, $X_n \xrightarrow{d} \mu$, it also converges in probability and the Slutsky's Theorem (Billingsley, 1995) can be applied to ensure convergence in distribution as $f'(\mu)\sqrt{n}(X_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2(f'(\mu))^2)$. Thus, we conclude:

$$\sqrt{n}(f(X_n) - f(\mu)) \xrightarrow{d} \mathcal{N}(0, \sigma^2(f'(\mu))^2). \tag{33}$$

Now, let us suppose that the variance depends on the mean value, $\sigma^2 = \mathrm{Var}(\mu)$. We are interested in finding a transformation $f_{\mathrm{stab}}(\cdot)$ such that $\sigma^2(f'_{\mathrm{stab}}(\mu))^2$ is a constant. For that purpose, the following differential equation can be set: $\sigma^2(f'_{\mathrm{stab}}(\mu))^2 = C^2$, whose solution provides the expression commonly used to calculate the VST (Bartlett, 1947):

$$f_{\mathrm{stab}}(x) = \int^x \frac{C}{\sqrt{\mathrm{Var}(\mu)}}d\mu, \tag{34}$$

where C is arbitrary constant.

4.1. Goals of the proposed methodology

Attending to the main characteristics of the signal studied in the exploratory data analysis section, the methodology we pursue should meet the following features:

1. **Same average measures**. The average attenuation levels should remain unaltered in order to describe a meaningful measure for quantitative purposes.
2. **Homogeneous variance**. The variance should get stabilized throughout the entire image.
3. **Information from just one slice**. The non-isotropic sampling of CT scans and the particular interpolation schemes of different devices may provide different noise characteristics depending on the location in the longitudinal axis. The methodology developed should be able to stabilize from just one axial slice.

4.2. Variance-stabilizing transformation for Gamma variates

For the case of Gamma distributions (we assume that the data is already centered to its central parameter δ; i.e. for $X \sim \Gamma(\alpha, \beta, \delta)$ we have $Y = X - \delta$, so $Y \sim \Gamma(\alpha, \beta)$ with PDF $f_\Gamma(y|\alpha, \beta)$ (see Eq. (14)), the functional relationship is clear: $\sigma^2 = \beta\mu$. Thus, the variance-stabilizing transformation derived from Eq. (34) trivially becomes:

$$f_{\mathrm{stab}}(y) = C\sqrt{\frac{y}{\beta}} + K, \tag{35}$$

where K and C are arbitrary constants. This result is closely related to the transformation proposed by Box and Cox (1964) to ensure Gaussianity in the transformed variable by means of a power transformation. Actually, this Gaussian behavior can be shown if the square root is carefully analyzed. Note that the transformation $W = \sqrt{Y}$ leads to a generalized Gamma distribution (Stacy, 1962) with PDF:

$$f_W(w) = 2w \cdot f_\Gamma(w^2|\alpha, \beta) = 2\frac{w^{2\alpha-1}}{\beta^\alpha\Gamma(\alpha)}e^{-\frac{w^2}{\beta}}, \quad \text{with } \alpha, \beta > 0. \tag{36}$$

This new variable shows a quadratic exponential decay that contributes to decreasing the kurtosis observed in the exploratory data analysis. Additionally, the exponential term $w^{2\alpha-1}$ increases the order of the derivative in the origin. Thus, the resulting

PDF is more symmetrical and the skewness is consequently decreased. So, after this transformation, the resulting variate shows a Gaussian-like distribution.

Finally, considering that the moments of the generalized Gamma distribution (Vegas-Sanchez-Ferrero et al., 2012) of Eq. (36) are:

$$E\{W^r\} = \beta^{r/2} \sum_{j=1}^{J} \frac{\Gamma(\alpha + r/2)}{\Gamma(\alpha)}, \tag{37}$$

the variance of $W = C\sqrt{Y/\beta}$ is (Tricomi and Erdélyi, 1951):

$$\mathrm{Var}\{W\} = C^2\alpha - C^2\left(\frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)}\right)^2 \xrightarrow[\alpha\to\infty]{} \frac{1}{4}C^2 + \mathcal{O}\left(\frac{1}{\alpha^2}\right), \tag{38}$$

so, if our purpose is to get a constant variance $\mathrm{Var}\{W\} = 1$ and we can assume high values of α, we can set $C = 2$ to get the desired stabilization. Conversely, for lower values of α, an estimate $\widehat{\alpha}$ should be applied in Eq. (38) to get the normalization of variance.

4.3. Variance-stabilizing transformation for ΓMMs

The case of stabilizing a Mixture Model of nc-Gamma distributions cannot be directly derived from eq. (34) as was done before due to the many parameters involved in the mean value of a mixture model. However, the stabilization of each of its components (i.e. Gamma distributions) can be used for this purpose as follows:

Let us suppose the image $X : \Omega \to \mathbb{R}$ is distributed as a nc-ΓMM with central parameters $\delta_j = \delta$ and $Y(\mathbf{r}) = X(\mathbf{r}) - \delta$ the centered Gamma mixture model. If we apply the same transformation derived for Gamma random variables, $W(\mathbf{r}) = \sqrt{Y(\mathbf{r})}$, by virtue of the formula for change of random variables, the new variable will follow a mixture model of generalized Gamma distributions whose PDF now reads:

$$f_W(w) = 2w \cdot \sum_{j=1}^{J} \pi_j f_\Gamma(w^2|\Theta_j) = \sum_{j=1}^{J} \pi_j f_W(w|\Theta_j), \tag{39}$$

where $f_W(\cdot)$ is the generalized Gamma distribution derived in Eq. (36). Now, we can define a similar transformation as the one obtained for Gamma random variables, Eq. (35), with the following desired features: 1) it should transform each component of the mixture model into its corresponding stabilized counterpart, 2) the resulting variable should show the same average value, 3) the variance should be stabilized to a constant value.

The following transformation meets those features:

$$f_{\mathrm{stab}}(Y(\mathbf{r})) = C \cdot \frac{\sqrt{Y(\mathbf{r})} - E\left\{\sqrt{Y(\mathbf{r})}\right\}}{\sqrt{\mathrm{Var}\left\{\sqrt{Y(\mathbf{r})}\right\}}} + E\{Y(\mathbf{r})\}, \tag{40}$$

where C accounts for the homogeneous level of variance desired and the moments of $\sqrt{Y}$ can be estimated taking advantage of the local tissue characterization introduced in Section 3.2 as the operator $\langle \cdot\, | Z(\mathbf{r}) = j\rangle$ in a neighborhood $\eta(\mathbf{r})$ of the sample at location $\mathbf{r}$. Therefore, the local moments are approximated as:

$$E\{Y(\mathbf{r})^k\} = \sum_{j=1}^{J} \pi_j E\{Y(\mathbf{r})^k | Z(\mathbf{r}) = j\} \approx \sum_{j=1}^{J} \pi_j \left\langle Y(\mathbf{r})^k | Z(\mathbf{r}) = j \right\rangle, \tag{41}$$

where, as in Eqs. (27) and (28), the sample conditioned local moments are:

$$\left\langle Y(\mathbf{r})^k | Z(\mathbf{r}) = j \right\rangle = \frac{\sum_{\mathbf{s}\in\eta(\mathbf{r})} y(\mathbf{s})^k \gamma_j(\mathbf{s})}{\sum_{\mathbf{s}\in\eta(\mathbf{r})} \gamma_j(\mathbf{s})}. \tag{42}$$

Note that the stabilizing transformation proposed in Eq. (40) is essentially an affine transformation of the resulting generalized Gamma mixture model. This transformation stabilizes the samples according to their belonging for each tissue class and, therefore, generalizes the stabilization derived for the Gamma variate in Section 4.2. The stabilization method is shown in Algorithm 3.

Algorithm 3 Stabilization for CT scans.

$Y(\mathbf{r}) \leftarrow X(\mathbf{r}) - \delta$ ▷ Centering the data
$\Theta \leftarrow$ Statistical Characterization ▷ Algorithm 2
$\gamma_j(\mathbf{r}) \leftarrow p(Z(\mathbf{r}) = j | Y(\mathbf{r}), \Theta(\mathbf{r}))$ ▷ Eq. (18)
$C \leftarrow$ Homogeneous standard deviation
$\left\langle \sqrt{Y(\mathbf{r})} | Z(\mathbf{r}) = j \right\rangle \leftarrow \sum_{\mathbf{s}\in\eta(\mathbf{r})} \sqrt{y(\mathbf{s})}\gamma_j(\mathbf{s}) / \sum_{\mathbf{s}\in\eta(\mathbf{r})} \gamma_j(\mathbf{s})$ ▷ Eq. (42)
$\langle Y(\mathbf{r}) | Z(\mathbf{r}) = j \rangle \leftarrow \sum_{\mathbf{s}\in\eta(\mathbf{r})} y(\mathbf{s})\gamma_j(\mathbf{s}) / \sum_{\mathbf{s}\in\eta(\mathbf{r})} \gamma_j(\mathbf{s})$ ▷ Eq. (42)
$\widehat{\mathrm{Var}}\{\sqrt{Y(\mathbf{r})}\} \leftarrow \langle Y(\mathbf{r}) | Z(\mathbf{r}) = j \rangle - \left\langle \sqrt{Y(\mathbf{r})} | Z(\mathbf{r}) = j \right\rangle^2$
$Y_{\mathrm{stab}(\mathbf{r})} \leftarrow C \frac{\sqrt{Y(\mathbf{r})} - \left\langle \sqrt{Y(\mathbf{r})} | Z(\mathbf{r}) = j \right\rangle}{\sqrt{\widehat{\mathrm{Var}}\{\sqrt{Y(\mathbf{r})}\}}} + \langle Y(\mathbf{r}) | Z(\mathbf{r}) = j \rangle$ ▷ Eq. (40)
return $Y_{\mathrm{stab}}(\mathbf{r})$

5. Results

In this section we validate the performance of the proposed method to enable comparison between different kernels, doses, and devices from three different scenarios: First, we study the effect of stabilization in the local variance of the phantom. Additionally, we provide an analysis of the sensitivity of prior mean values in the stabilization. We also study the effect of the stabilization in the spatial resolution. The second scenario analyzes the reduction of the inhomogeneities in local variance for common chest CT scans acquired from the same subject. The third scenario considers Low-dose iterative reconstruction compared to high-dose reconstructions.

5.1. Phantom stabilization

In this first experiment, we evaluate the performance of the stabilization method for all the different configurations of the phantom. The stabilization was performed separately in each of the axial slices to get a homogeneous standard deviation of 10 HUs. The local standard deviation was calculated across the longitudinal direction. The results are depicted in Fig. 6 for each of the devices (General Electric and Siemens) where the variance of the stabilized images (lower row) shows a more homogeneous pattern than the observed in the original data (top row).

The stabilization of noise in images should provide same average attenuation levels and same variance. This effect can be confirmed by observing the behavior of the histogram of the phantom in Fig. 7. Note that the histogram shows important differences in the variance of the histograms, not only between tissues but also between reconstructions and devices. On the contrary, the stabilized images show homogeneous variance along tissues, reconstructions, and devices. These results demonstrate that the attenuation levels are now comparable between kernels and devices.

5.2. Sensitivity of prior mean values

During statistical characterization of tissues (such as bone, blood, muscle), their mean values are assumed to be known. This is a reasonable assumption since these tissues are well characterized and the calibration of CT scans involve the tuning of some of these attenuation levels as previously mentioned. Some other tissues, however, do not necessarily fit the attenuation levels of

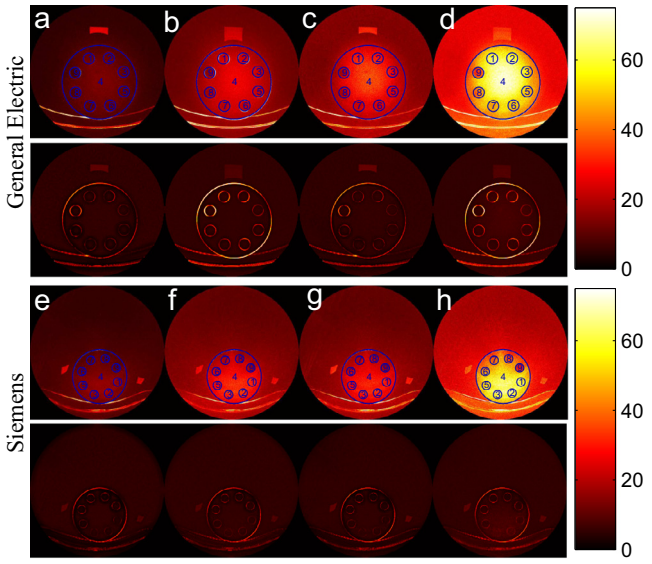

Fig. 6. Standard deviation of the stabilized data for different kernels and doses to a homogeneous standard deviation of $C = 10$ HUs. a) STD HD, b) STD LD, c) BONE HD, d) BONE LD, e) B31f HD, f) B31f LD, g) B45f HD, h) B45f LD.

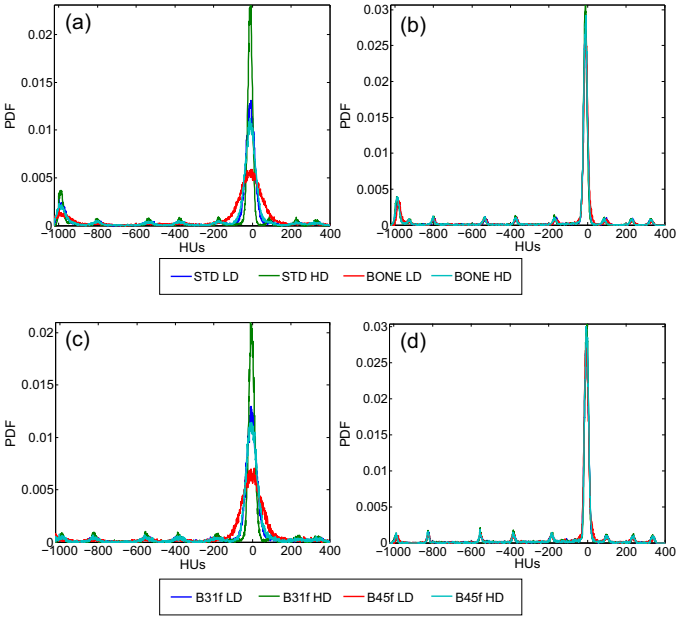

Fig. 7. Histogram comparison of the phantom before and after stabilization for the different configurations: general electric acquisition (a) and (b) and for the Siemens acquisitions (c) and (d). Note that the variance observed in non-stabilized images (a) and (c) is clearly heteroscedastic, whereas the stabilized variance in (b) and (d) becomes homoscedastic.

any of the mean values established and, therefore, the stabilization of variance could be negatively affected.

Our method circumvents this problem by calculating the local moments, Eq. (41), considering the effect of different tissues within the local neighborhood, and heterogeneous compositions that may deviate from the previously established mean attenuation of tissues. The normalization of weights according to the posterior probabilities in Eq. (42) provides a convex weighting of sample intensities, which results in an effective anisotropic estimate of local moments. Even in the cases with intensities significantly deviated from the prior mean values, the local weights will contribute for the anisotropic estimation of local moments. In the worst-case scenario, the differences between tissues will be negligible, and

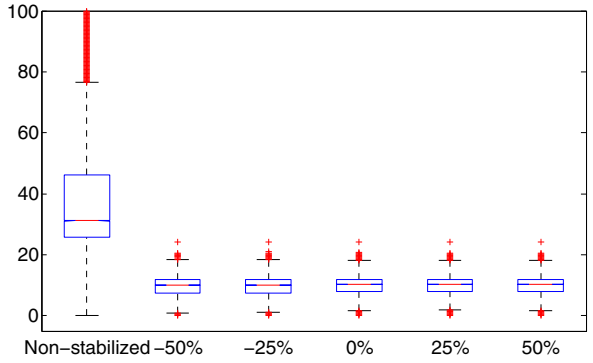

Fig. 8. Box-plot of the local standard deviation for different bias in the prior mean for region 4 of the phantom.

the local weighting would become more isotropic. This situation would be only possible if the prior means are too close and the tissues within the local neighborhood show a poor contrast.

For the sensitivity analysis, we considered the GE phantom acquisitions with the configuration with the highest variance (BONE LD) to enhance deviations from the ideal result. We intentionally introduced a bias in the prior mean value for the substrate of the phantom (designed as region 4 in Fig. 1 because it is the tissue type showing more dynamic range in the local variance and it is also in contact with all the rest tissue types. Five different biases (−50%, −20%, 0%, 25%, and 50%) were considered in terms of the distance with its next or previous prior mean. The stabilization of the standard deviation was set to $C = 10$ HU. The results depicted in Fig. 8 clearly show the robustness even for strong deviations from the expected value and the suitability of the estimation of local moments proposed in Eq. (42).

5.3. Influence of the Stabilization on the Spatial Resolution

We study the effect of the stabilization technique on the spatial resolution by measuring the line-spread function (LSF) of the system, or equivalently, by its frequency transformation, the modulation transfer function (MTF) (Judy, 1976). The MFT is the standard way to assess a scanner resolution according to ISO 15708-2.[3]

We employed the same phantom studied in the exploratory analysis since it fits the requirements of the standard ISO 15708-2. We considered the edge profile of the bigger cylinder (region 4) and background as indicated in Fig. 9 (left column). The image was stabilized with $C = 0$, which is the worst-case scenario in which the stabilized image is the estimated signal and, thus, the edges are more prone to smoothing effects. The edge response function is computed as the average of the whole set of samples in the z-axis as suggested in the standard ISO 15708-2 and it is shown in Fig. 9 for the case of GE (the Siemens case showed similar results and was omitted for brevity). We superimposed the intensities for which the 10% and 90% of the edge response is reached. These margins serve as a robust measure to compare the differences between modulation transfer functions quantitatively. At first glance, the differences between the 10% − 90% margins are indistinguishable and clearly evidence the good preservation of resolution after the stabilization. The LSF, calculated from the edge response function by the absolute value of finite differences, shows a similar support in the edge location in all the cases. Finally, the MTF, calculated as the normalized discrete Fourier transform of the LSF, exhibits the high similarity between non-stabilized and stabilized

[3] ISO 15708-2, "Non-destructive testing–Radiation methods–Computed tomography. Part 2: Examination practices", www.iso.org/iso.

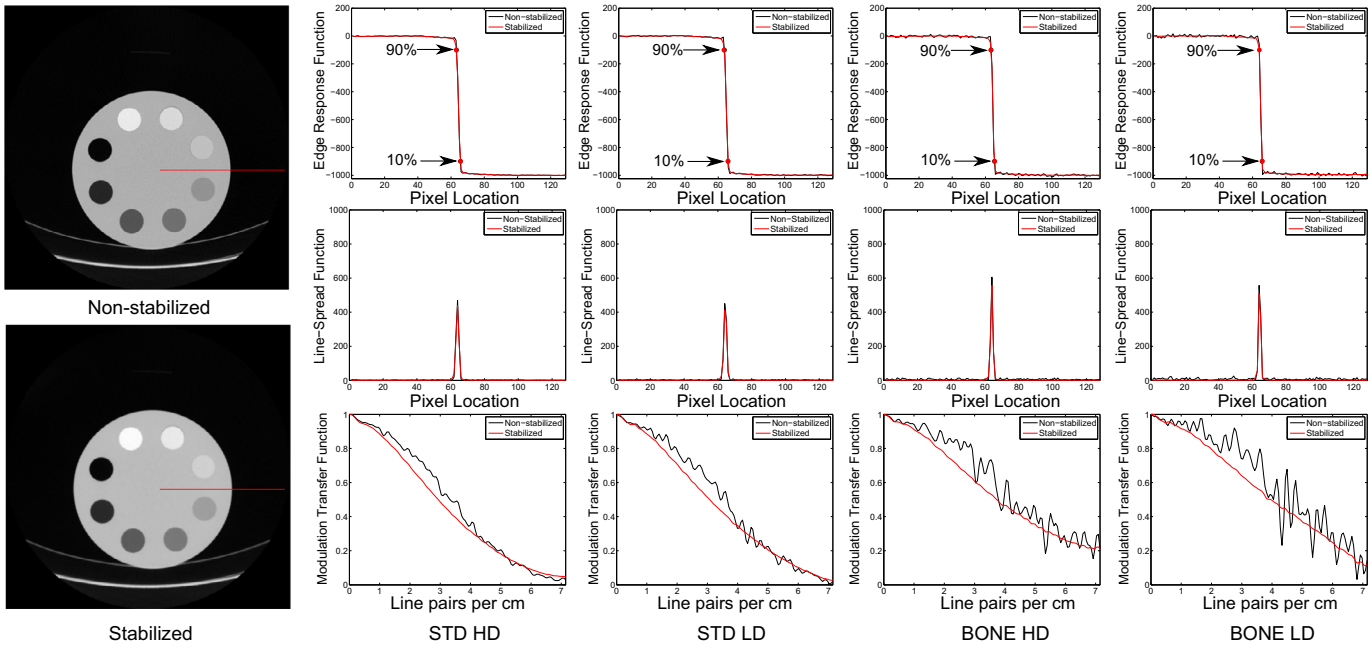

Fig. 9. Influence of the stabilization scheme in the spatial resolution for the GE acquisition. Top row: Edge Response Function. Mid row: Line-Spread Function. Bottom row: Modulation Transfer Function. The strong similarity between Modulation Transfer Functions shows that the resolution is not critically compromised after the stabilization.

Table 5
Distance of the 10% − 90% of the edge response in pixel units for stabilized images with $C = 0$. The differences between both measures is always below 0.5 pixels.

GE	Normal	Stabilized	Siemens	Normal	Stabilized
STD LD	2.14	2.50	B31f LD	1.99	2.31
STD HD	2.38	2.63	B31f HD	2.12	2.36
BONE LD	1.67	1.83	B45f LD	1.36	1.54
BONE HD	1.99	2.32	B45f HD	1.36	1.52

images. A slightly lower performance is, however, observed for medium spatial frequencies (2 to 4 line pairs per cm). To analyze the importance of this fact, we provide a more detailed analysis in Table 5, where we show the 10% − 90% distance in pixels. Note that the difference between both results is always lower than 0.5 pixels, meaning that the 10% − 90% rise differences between non-stabilized and stabilized images are not critically compromised.

If we consider a higher level of variance in the stabilized image, say $C = 30$, the MTF becomes more similar to the non-stabilized case, and so it does the 10% − 90% distance. These results confirm the suitability of the statistical characterization for the calculation of accurate local moments and show an excellent trade-off between resolution and stabilization.

Note that the method was intentionally designed to preserve the resolution of the original image after the stabilization. This means that the resolution will still depend on the reconstruction kernel and, therefore, the comparisons between images reconstructed with different kernels are still limited to their differences in spatial resolution. This fact would not remarkably affect to studies that analyze the evolution of the disease according to the attenuation levels in certain regions of interest, where the statistical properties are of most importance for comparison purposes instead of the delineation of structures.

5.4. Clinical CT scans stabilization

The stabilization performance was also evaluated in conventional chest CT scans obtained with different doses and kernels.

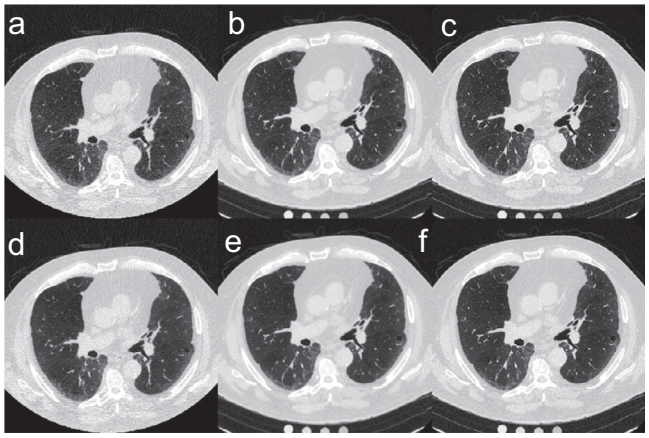

Fig. 10. CT real images acquired with different doses and kernels for a window level of $[-1024, 200]$ HUs: a) B31f LD, b) B31f HD, c) B45f HD) and their respective stabilized images d) B31f LD, e) B31f HD, f) B45f HD).

The images were acquired from one subject that underwent CT scanning with two different radiation exposures (high dose and low dose). For the high dose, the data was reconstructed with two different kernels, one smooth (B31f) and the other sharp (B45f). The low dose data was reconstructed with the same smooth kernel.[4] An axial view of the volumes is shown in Fig. 10 jointly to the stabilized images for a window level of $[-1024, 200]$ HUs. Note that the lower dose acquisition exhibits a much higher variance of noise due to photon starvation, especially in the lower parts of the lungs, where bony tissues increase the noise.

As opposed to the phantom case, we cannot directly calculate the local standard deviation for the evaluation due to the lack of multiple acquisitions. For this reason, we used the method proposed by Aja-Fernandez et al. (2015) to get an estimate of the local standard deviation. In Fig. 11 we show the local standard

[4] The data was obtained at Brigham and Women's Hospital (Boston, MA, USA) with the approval of its ethical committee and informed consent of the subject.

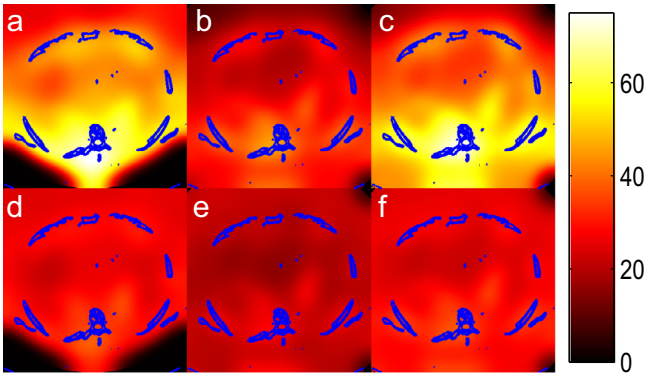

Fig. 11. Local standard deviation estimated from real CT images: a) B31f LD, b) B31f HD, c) B45f HD) and their respective stabilized images d) B31f LD, e) B31f HD, f) B45f HD). The contours of bony structures are over-imposed in blue for anatomical reference.

deviation estimated with the contours of the bony structures over-imposed (in blue) to have an anatomical reference. Note that the stabilized image reduces the variance of noise in those regions leading to a more accurate comparison between low dose and high dose images.

The visual effect of the noise stabilization is depicted in Fig. 12, were the acquisition reconstructed with a B45f kernel is represented along with its stabilized image with standard deviation set to 20. Note that the stabilized image provides a much clearer differentiation of soft organs that were affected by a higher variance of noise. Two different regions were zoomed to show better this effect. Fig. 12 also includes the histogram of noise estimated during the stabilization process. The stabilized histogram of noise (in blue) exhibits a Gaussian shape and a much lower variance as a result of the homogeneous response after stabilizing. Conversely, the non-stabilized noise has a much higher variance due to the contributions of noise with different variances. This qualitative behavior is confirmed when the kurtosis and skewness of both estimated noises are compared. The non-stabilized noise has a skewness of 0.47, and a kurtosis $=$ 6.90, confirming a clearly skewed and leptokurtic behavior. On the other hand, the stabilized version exhibits a skewness of -0.108 and a kurtosis of 3.14, which is much closer to the desired Gaussian behavior of noise (in green).

The performance of the stabilizing method was quantitatively evaluated by comparing local histograms with the two-sample Kolmogorov–Smirnov distance. The comparison considered the following configurations: dose (LD B31f vs. HD B31f), kernels (HD B31f vs. HD B45f) and both dose and kernel (LD B31f vs. HD B45f). For this purpose, we divided the images shown in Fig. 10 into a set of 165 non-overlapping neighborhoods of size 31×31 from each acquisition and their respective empirical cumulative distribution functions were compared according to the Kolmogorov–Smirnov distance introduced in Eq. (11). This distance allows us to test statistically whether both distributions come from the same random variable.

The quantitative evaluation was done for a stabilized standard deviation of $C = 40$ HUs (see Eq. (40)), which is the median of the local standard deviation of the original images. This value was set to account for the fine details that may be missed during the estimation of the expected values in Eqs. (40)–(42). The mean value of histograms was removed to avoid potential biases non-related to local variance.

The results obtained for the two-sample Kolmogorov–Smirnov distances are shown in Fig. 13, where the boxplots corresponding to the stabilized images exhibit a remarkable decrease in the distance. This difference is statistically significant in all the comparisons for the two-sample Welch t-test (p-value $< 10^{-4}$). The ratio of neighborhoods with indistinguishable distributions considering a statistical significance of 0.001 are shown in Table 6 where the improvement becomes evident and, in some cases, doubles the performance (from 22.21% to 55.15% in the worst case). The results in HD for both kernels is improved up to 100% which shows that the differences due to variance are avoided, making HD reconstructions completely comparable.

Table 6
Ratio of neighborhoods with statistically indistinguishable distributions obtained with a two-sample Kolmogorov–Smirnov test.

	LD B31f vs. HD B31f	LD B31f vs. HD B45f	HD B31f vs. HD B45f
Normal	21.21%	48.48%	80.00%
Stabilized	55.15%	72.73%	100.00%

Table 7
Ratio of neighborhoods with statistically indistinguishable distributions obtained with a two-sample Kolmogorov–Smirnov test.

	LD I44f vs. HD B31f	LD I44f vs. HD B45f
Normal	52.12%	80.61%
Stabilized	80.61%	96.97%

5.5. Iterative reconstruction

The proposed methodology was also tested for images reconstructed with an iterative algorithm. The same low-dose acquisition considered before was reconstructed by the Siemens iterative reconstruction method SAFIRE (Sinogram Affirmed Iterative Reconstruction) denoted as I44f. This reconstruction behaves as a sharp kernel and the overall effect in the anatomy is comparable to B45f. Then the nc-ΓMM approach was applied to the resulting image and stabilized. The estimated local standard deviation is shown in Fig. 14. Although the reduction of spatial variance is less effective, this case also shows a significant improvement in the homogeneity of local variance. This effect was also quantified with an analysis of histogram similitude as before (a set of 165 non-overlapping neighborhoods of size 31×31 and Kolmogorov–Smirnov distance, $C = 40$ HUs). The results of the histogram comparison shown in Fig. 15 again confirm the importance of stabilization before histogram comparison. The differences between stabilized and non-stabilized images are statistically significant (Welch t-test, p-value $< 10^{-3}$ for all the cases).

The ratio of neighborhoods with indistinguishable distributions considering the two-sample Kolmogorov–Smirnov statistical test with a significance of 0.001 are shown in Table 7, where the improvement after stabilization is clearly confirmed once again. Note that the performance of the comparison between LD and HD B45f stands out because the details provided by sharper kernels are also preserved by the iterative reconstruction, which results in an increase of the similarity. In the case of smooth kernels such as B31f, some details are lost due to the low-pass filter effect of the smooth kernel. Nevertheless, the stabilization raises this ratio until 80.61%, which is more than the ones obtained for LD reconstruction with the B31f kernel. This means that the stabilization is also effective for iterative reconstruction and takes advantage of the better resolution provided by iterative schemes. Besides, the comparison between the iterative LD reconstruction and the B45f HD reconstruction reaches 96.97%. This result clearly shows the suitability of the proposed methodology to enable comparisons with different reconstruction methods and different doses.

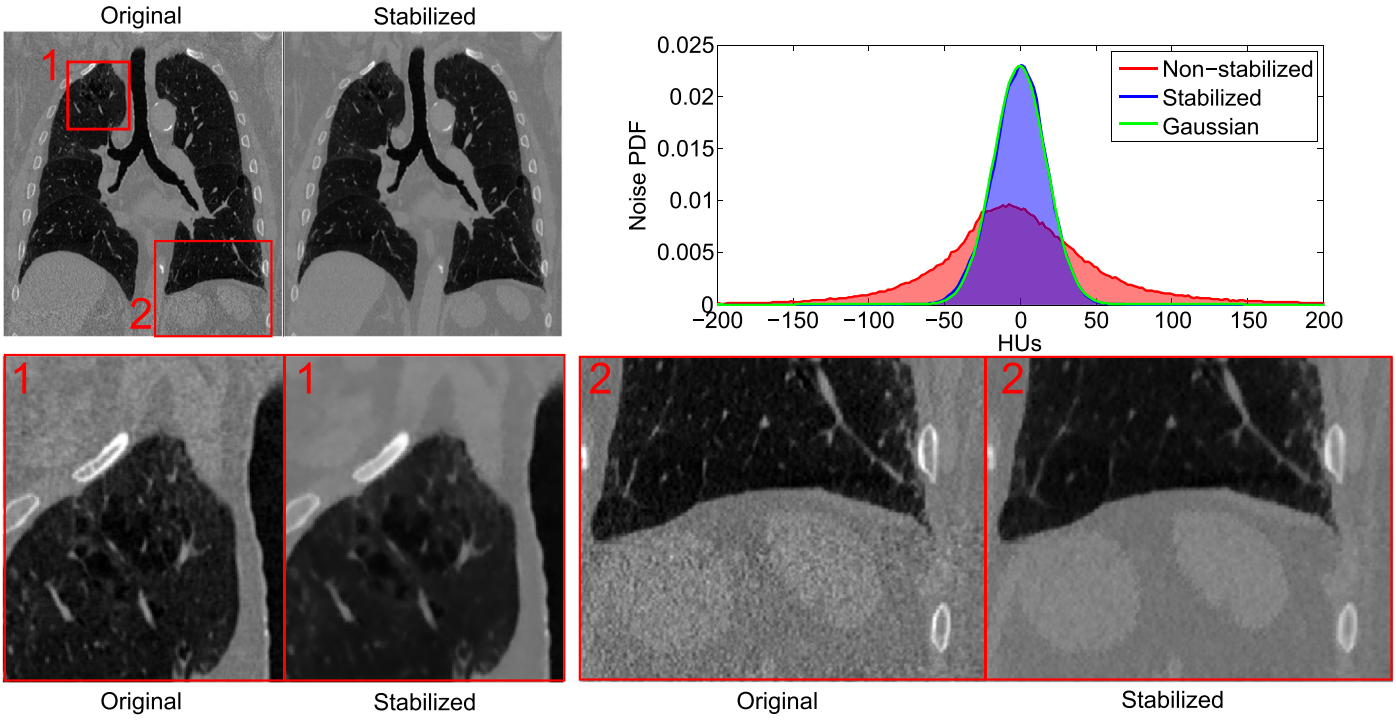

Fig. 12. Noise stabilization effect in clinical CT scans. The variance of noise increases dramatically in the upper and lower parts of the image (zoomed regions 1 and 2), causing a noticeable reduction of contrast in soft organs and lung parenchyma. The stabilized image shows a more homogeneous response of noise and increases the contrast between tissues. The stabilized histogram of noise shows a behavior much closer to a Gaussian distributed noise (in green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

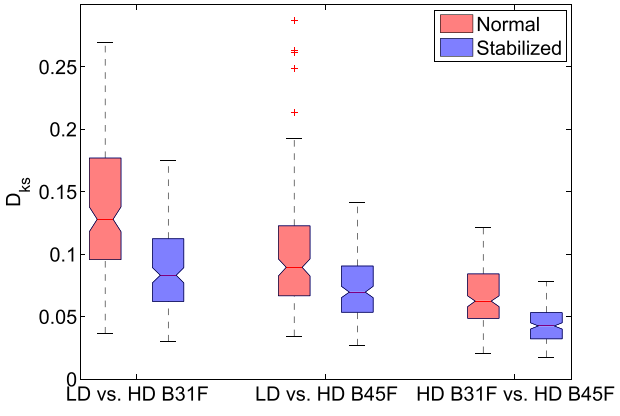

Fig. 13. Kolmogorov–Smirnov distance for the two-sample tests obtained from non-stabilized images (red) and stabilized images (blue). The improvement provided by the stabilization is strongly significant (p-value $< 10^{-4}$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

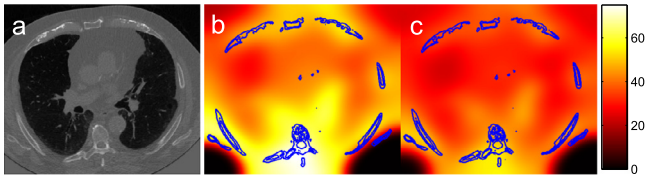

Fig. 14. Iterative reconstruction (a) and local standard deviation estimated for iterative reconstruction before (b) and after stabilization (c).

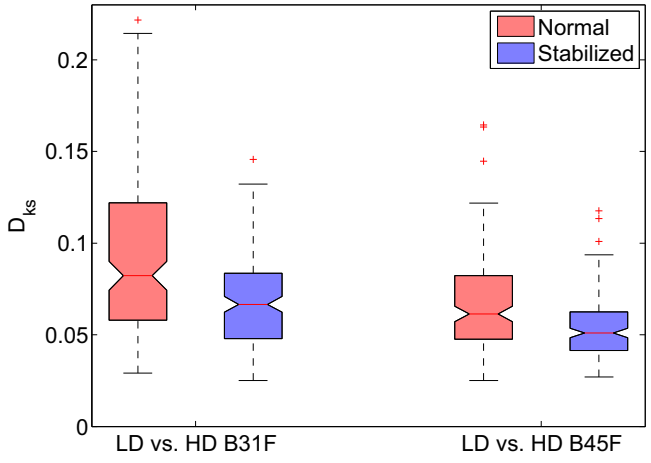

Fig. 15. Kolmogorov–Smirnov distance for the two-sample tests obtained from non-stabilized images (red) and stabilized images (blue) for the case of Iterative reconstruction. The improvement provided by the stabilization is strongly significant (p-value $< 10^{-4}$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6. Discussion and conclusions

This paper analyzes the effective statistical response of attenuation levels in CT scans in different arrangements including, reconstruction kernels, doses, and devices. The statistical analysis was carried out through an exhaustive exploratory data study including graphical representations, quantitative analysis up to 4-th order moments and their functional relationships among them. The study showed a leptokurtic behavior of noise, with a linear relationship between mean and variance. The noise also shows a convergence to a Gaussian noise as the average attenuation

level increases. Finally, the skewness-kurtosis plot of noise can be explained by a linear regression that fits extraordinarily well to a Gamma-like distribution. This analysis lead us to propose a parametric distribution –the non-central Gamma– that accurately models the statistical behavior regardless of dose, kernel or device. The suitability of this distribution was confirmed with different statistical tests involving the moments of the observed samples and their functional relationship (variance vs. mean, kurtosis vs. skewness).

Having a statistical model that properly describes the data paves the way for a more accurate treatment of the signal properties observed in the image. In the case of CT scans, the non-homogeneous behavior of noise is an undesired effect that depends on several factors (dose, tissue, artifacts due to photon starvation, etc.). These factors introduce serious difficulties in the comparison of scanners when the progression of a disease or cross-sectional studies are performed. Actually, the same acquisition reconstructed with different kernels may show important discrepancies in the signal due to the effect of noise with different variances.

This work extends the proposed statistical distribution to model the non-homogeneous response of tissues (due to fluctuations in the variance) with a non-central Gamma Mixture Model. This model can cope with the partial volume effect and provides an accurate description of the local statistical properties of noise. The approach chosen for the estimation of local parameters is from coarse-to-fine, providing first a global set of parameters describing the statistics of the image that will be used for the local estimation. Both the global and the local implementations are presented.

Additionally, this characterization of noise can be stabilized to provide a homogeneous variance with the same attenuation levels as in the original image. As a result, the stabilized image presents the same attenuation levels with a more homogeneous response of noise that unveils details of soft tissues and enables the comparison of images acquired with different protocols. The evaluation of the proposed methodology is carried out in three different scenarios: first, we stabilize a phantom image acquired with different doses, reconstruction kernels, and devices. The stabilizing method successfully transformed the heteroscedastic image into a homoscedastic image (Figs. 6 and 7). The effect of stabilization was also tested in clinical images, for which different images of the same subject were acquired for different doses and reconstruction kernels. In that case, the stabilization of variance was also achieved (Fig. 11), showing a remarkable improvement in contrast for soft tissues and a more accurate comparison between images which is statistically significant (p-value $< 10^{-4}$). Actually, in some cases, the stabilization doubles the performance of the non-stabilized images (see Table 6).

The suitability of the proposed methodology is also confirmed for iterative reconstruction and also shows a significant improvement in the statistical similitudes of images after the stabilization (p-value $< 10^{-4}$). This result in especially relevant since it raises the power of the statistical comparison up to the levels of high dose acquisitions (from 52.12% to 80.61% in the worst case).

Our variance stabilization method is designed to transform the non-homogeneous nature of noise to a more treatable homogeneous and Gaussian-like distribution of noise without compromising the signal of the original image. With this aim, the estimation techniques used in the calculation of local moments, Eqs. (41) and (42), take advantage of the local statistical characterization to provide more accurate estimates that are not affected by the presence of different tissue types in the same neighborhood. Besides, our resolution experiments support that the stabilization does not compromise the details of fine structures since the homogeneous variance factor, C, can be set to a level that preserves any loss of detail that is not preserved in the mean image, $E\{Y(\mathbf{r})\}$.

During the stabilization, we make use of local moment estimates according to a non-stationary noise model derived. The local mean obtained from the conditioned local moments can be seen as a robust estimate of the signal level and, thus, a denoised version of the original image. As a consequence, for low values of C in Eq. (42) ($C = 0$ in an extreme case), the stabilization may be seen as a denoising methodology with some advantages. First, the local expectation successfully preserves the spatial resolution as was shown in Section 5.3. Second, the preservation of details and structures that have been considered as noise by the local expectation operator can be preserved just by increasing the C value. Finally, the resulting homogeneous Gaussian-like noise obtained after stabilization allows other methods designed for additive Gaussian noise to work in optimal conditions.

The proposed statistical characterization has shown its suitability for modeling non-homogeneous tissues according to a consistent noise model across kernels, doses, and devices. However, it is worth mentioning that the correlation of noise has been avoided intentionally for the sake of simplicity. This simplification does not show significant consequences in the estimation of posterior probability maps and, therefore, the estimation of local moments is not affected. However, the extension of the proposed statistical framework to correlated noise models would benefit other applications such as image reconstruction (if sinograms are provided) or denoising.

Our work provides methodological contributions that are amenable to be combined with preexisting reconstruction approaches. Some iterative reconstruction methods, such as DART (Batenburg and Sijbers, 2011) or SAFIRE (Grant and Raupach, 2012), exploit an iterative refinement of the image in the reconstructed domain that can be used as prior for estimating a new back-projection. We believe that our methodology makes use of a probabilistic distribution of tissue components that could be utilized as a probabilistic representation of the reconstructed image to generalize iterative approaches whenever sinograms are available.

We believe the results provided in this paper will enable the preprocessing of CT images before QI analysis tasks are performed to control the precision of the extracted biomarkers due to heteroscedastic conditions. Our results can be especially relevant to those studies that track the disease progression. The changes in transmission medium due to weight loss or gain that typically happens in longitudinal studies or oncological situations makes the issue of noise stabilization really important.

Acknowledgments

Gonzalo Vegas-Sánchez-Ferrero acknowledges Consejería de Educación, Juventud y Deporte of Comunidad de Madrid and the People Programme (Marie Curie Actions) of the European Union Seventh Framework Programme (FP7/2007– 2013) for REA grant agreement no. 291820. This study was supported by the Ministry of Economy and Competitiveness (project TEC-2013-48251-C2-2-R) and the National Institutes of Health NHLBI awards R01HL116931 and R01HL116473.

References

Abramson, R., Burton, K., Yu, J., Scalzetti, E., Yankeelov, T., Rosenkrantz, A., Mendiratta-Lala, M., Bartholmai, B., Ganeshan, D., Lenchik, L., Subramaniam, R., 2015. Methods and challenges in quantitative imaging biomarker development. Acad. Radiol. 22 (1), 25–32. doi:10.1016/j.acra.2014.09.001.

Aerts, H.J.W.L., Velazquez, E.R., Leijenaar, R.T.H., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M.M., Leemans, C.R., Dekker, A., Quackenbush, J., Gillies, R.J., Lambin, P., 2014. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat. Commun. 5 SP –, 4006.

Agatston, A.S., Janowitz, W.R., Hildner, F.J., Zusmer, N.R., Viamonte, M., Detrano, R., 1990. Quantification of coronary artery calcium using ultrafast computed tomography.. J. Am. College Cardiol. 15 (4), 827–832.

Aja-Fernandez, S., Pieciak, T., Vegas-Sanchez-Ferrero, G., 2015. Spatially variant noise estimation in MRI: a homomorphic approach. Med. Image Anal. 20 (1), 184–197.

Baek, J., Pelc, N.J., 2010. The noise power spectrum in CT with direct fan beam reconstruction. Med. Phys. 37 (5).

Balda, M., Heismann, B., Hornegger, J., 2010. Non-stationary ct image noise spectrum analysis.. In: Bildverarbeitung für die Medizin, pp. 191–195.

Bartlett, M.S., 1947. The use of transformations. Biometrics 3 (1), 39–52.

Batenburg, K.J., Sijbers, J., 2011. DART: a practical reconstruction algorithm for discrete tomography. IEEE Trans. Image Process. 20 (9), 2542–2553. doi:10.1109/TIP.2011.2131661.

Billingsley, P., 1995. Probability and Measure, 3 Wiley.

Boogers, M.J., Schuijf, J.D., Kitslaar, P.H., van Werkhoven, J.M., de Graaf, F.R., Boersma, E., van Velzen, J.E., Dijkstra, J., Adame, I.M., Kroft, L.J., de Roos, A., Schreur, J.H., Heijenbrok, M.W., Jukema, J.W., Reiber, J.H., Bax, J.J., 2010. Automated quantification of stenosis severity on 64-slice cta comparison with quantitative coronary angiography. JACC 3 (7), 699–709. doi:10.1016/j.jcmg.2010.01.010.

Borsdorf, A., Kappler, S., Raupach, R., Hornegger, J., 2008. Analytic noise propagation for anisotropic denoising of CT images. In: 2008 IEEE Nuclear Science Symposium Conference Record, pp. 5335–5338. doi:10.1109/NSSMIC.2008.4774438.

Borsdorf, A., Kappler, S., Raupach, R., Hornegger, J., 2008. Analytic noise-propagation in indirect fan-beam fbp reconstruction. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2701–2704. doi:10.1109/IEMBS.2008.4649759.

Borsdorf, A., Kappler, S., Raupach, R., Noo, F., Hornegger, J., 2009. Local orientation-dependent noise propagation for anisotropic denoising of ct-images. In: 2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC), pp. 2472–2475. doi:10.1109/NSSMIC.2009.5402090.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. R. Stat. Soc.. Ser. B (Methodol.) 211–252. doi:10.2307/2287791.

Buckler, A.J., Bresolin, L., Dunnick, N.R., Sullivan, D.C., the Group, F., 2011. Quantitative imaging test approval and biomarker qualification: Interrelated but distinct activities. Radiology 259 (3), 875–884. doi:10.1148/radiol.10100800.

Cullen, A.C., Frey, H.C., 1999. Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press, New York.

Grant, K., Raupach, R., 2012. SAFIRE : sinogram affirmed iterative reconstruction. White Paper 1–8. URL www.usa.siemens.com/healthcare.

Groeneveld, R.A., Meeden, G., 1984. Measuring Skewness and Kurtosis. J. R. Stat. Soc.. Ser. D (Stat.) 33 (4).

Hsieh, J., 2003. Computed tomography : principles, design, artifacts, and recent advances. Bellingham, Wash. SPIE Hoboken, N. J. John Wiley & Sons.

Joanes, D.N., Gill, C.A., 1998. Comparing measures of sample skewness and kurtosis. J. R. Stat. Soc (Stat.) 47 (1), 183–189. doi:10.1111/1467-9884.00122.

Jones, L.V.(Ed.), 1986. The collected works of John W. Tukey. Volume III. Philosophy and principles of data analysis: 1949–1964. With a biography of Tukey by Frederick Mosteller.

Judy, P.F., 1976. The line spread function and modulation transfer function of a computed tomographic scanner.. Med. Phys. 3 (4), 233–236.

Kendall, M.G., Stuart, A., 1969. The Advanced Theory of Statistics, Volume 1: Distribution Theory, third Charles Griffin & Company Limited, Bucks, UK.

Kim, J.H., Chang, Y., Ra, J.B., 2016. Denoising of polychromatic ct images based on their own noise properties. Med. Phys. 43 (5), 2251–2260. http://dx.doi.org/10.1118/1.4945022.

Li, T., Li, X., Wang, J., Wen, J., Lu, H., Hsieh, J., Liang, Z., 2004. Nonlinear sinogram smoothing for low-dose X-ray CT. IEEE Trans. Nuclear Sci. 51 (5 II), 2505–2513. doi:10.1109/TNS.2004.834824.

Man, B.D., Nuyts, J., Dupont, P., Marchal, G., Suetens, P., 2001. An iterative maximum-likelihood polychromatic algorithm for CT. IEEE Trans. Med. Imaging 20 (10), 999–1008. doi:10.1109/42.959297.

Millner, M.R., Payne, W.H., Waggener, R.G., McDavid, W.D., Dennis, M.J., Sank, V.J., 1978. Determination of effective energies in ct calibration. Med. Phys. 5 (6), 543–545. http://dx.doi.org/10.1118/1.594488.

Moon, T., 1996. The expectation-maximization algorithm. IEEE Signal Process. Mag. 13 (6), 47–60.

Müller, N.L., Staples, C.A., Miller, R.R., Abboud, R.T., 1988. Density mask. An objective method to quantitate emphysema using computed tomography. Chest 94 (4), 782–787.

Mulshine, J.L., Gierada,David S III, S.G.A., Avila, R.S., Yankelevitz, D.F., Kazerooni, E.A., Mcnitt-gray, M.F., Buckler, A.J., Sullivan, D.C., 2015. Role of the quantitative imaging biomarker alliance in optimizing CT for the evaluation of lung cancer screen e detected nodules. J. Am. College Radiol. 12 (4), 390–395. doi:10.1016/j.jacr.2014.12.003.

Stacy, E.W., 1962. A generalization of the gamma distribution. Ann. Math. Stat. 33, 1187–1192. doi:10.1214/aoms/1177704481.

Tricomi, F.G., Erdélyi, A., 1951. The asymptotic expansion of a ratio of gamma functions. Pacific J. Math. 1 (1), 133–142.

Uppot, R.N., Sahani, D.V., Hahn, P.F., Gervais, D., Mueller, P.R., 2007. Impact of obesity on medical imaging and image-guided intervention.. AJR. Am. J. Roentgenol. 188 (2), 433–440.

Vegas-Sanchez-Ferrero, G., Aja-Fernandez, S., Palencia, C., Martin-Fernandez, M., 2012. A generalized gamma mixture model for ultrasonic tissue characterization. Comput. Math. Methods Med. 2012. doi:10.1155/2012/481923.

Vegas-Sanchez-Ferrero, G., Seabra, J., Rodriguez-Leor, O., Serrano-Vida, A., Aja-Fernandez, S., Palencia, C., Martin-Fernandez, M., Sanches, J., 2014. Gamma mixture classifier for plaque detection in intravascular ultrasonic images. IEEE Trans. Ultrason., Ferroelectr., Freq. Control 61 (1), 44–61. doi:10.1109/TUFFC.2014.6689775.

Whiting, B.R., 2002. Signal statistics of X-ray computed tomography. In: SPIE 4682, Medical Imaging 2002: Physics of Medical Imaging, pp. 53–60.

Whiting, B.R., Massoumzadeh, P., Earl, O.A., O'Sullivan, J.A., Snyder, D.L., Williamson, J.F., 2006. Properties of preprocessed sinogram data in x-ray computed tomography. Med. Phys. 33 (9), 3290. doi:10.1118/1.2230762.

Zeng, G.L., 2004. Nonuniform noise propagation by using the ramp filter in fan-beam computed tomography. IEEE Trans. Med. Imaging 23 (6), 690–695.