



Universidad de Valladolid

Escuela Técnica Superior de Ingenieros de Telecomunicación
Dpto. de Teoría de la Señal y Comunicaciones e Ing. Telemática

TESIS DOCTORAL

PROBABILISTIC MODELS FOR TISSUE CHARACTERIZATION IN ULTRASONIC AND MAGNETIC RESONANCE MEDICAL IMAGES AND APPLICATIONS

Presentada por Gonzalo Vegas Sánchez-Ferrero para
optar al grado de doctor por la Universidad de Valladolid

Dirigida por:

Marcos Martín Fernández

Doctor por la Universidad de Valladolid

César Palencia de Lara

Doctor en Ciencias Matemáticas

Tesis: **Probabilistic Models for Tissue Characterization in Ultrasonic and Magnetic Resonance Medical Images and Applications**

Autor: **Gonzalo Vegas Sánchez-Ferrero**

Directores: **Marcos Martín Fernández
César Palencia de Lara**

El tribunal nombrado para juzgar la tesis doctoral arriba citada, compuesto por los doctores:

Presidente:

Vocales:

Secretario:

acuerda otorgarle la calificación de

En Valladolid, a

RESUMEN

La caracterización de tejidos resulta de gran importancia en imagen médica pues permite diferenciar entre tejidos sanos y enfermos. Existen multitud de técnicas para proporcionar esta caracterización. Las más comunes son aquellas basadas en reconocimiento de patrones, que examinan características de las texturas que les permitirán extrapolarlas para obtener una clasificación de tejidos. Otras aproximaciones tienen en cuenta los modelos físicos subyacentes del proceso de adquisición y procuran caracterizar los mecanismos que producen las señales recibidas. La principal ventaja de estas técnicas es que su uso no está limitado por su capacidad de aprender patrones de grupos de entrenamiento. Al contrario, dependen de la idoneidad del modelo físico escogido. Esta Tesis estudia la caracterización de tejidos en imágenes ultrasónicas (US) y de resonancia magnética (MRI) a partir de los modelos físicos subyacentes que originan los patrones de naturaleza aleatoria resultantes.

La descripción estadística de la señal de US permite aportar una importante información de los tejidos que reflejan la señal de eco y que presentan un patrón característico llamado *speckle*. Los parámetros de los modelos estadísticos permiten entonces identificar características de los tejidos que sirven como importantes descriptores para lograr una clasificación. Muchos métodos de segmentación, filtrado o clasificación se basan en técnicas Bayesianas que precisan de un modelo estadístico acertado. Esto hace que el campo del modelado de la envolvente de señales de US haya sido muy prolífico en los últimos años. No obstante, la complejidad de la composición de tejidos puede presentar combinaciones de diferentes naturalezas en la celda de resolución. Además, las transformaciones de señal que se llevan a cabo en el proceso de adquisición dificultan una apropiada caracterización probabilística de las señales. Todo esto hace que exista una falta de consenso sobre las distribuciones probabilísticas que modelan los tejidos en US.

A pesar de que la física que gobierna el proceso de adquisición de US y MRI son radicalmente distintas, las distribuciones que modelan el ruido en MRI son similares a las de US. En el caso de MRI la presencia de ruido afecta a la eficacia de los métodos de postprocesado, como pueden ser la segmentación, el registrado o estimación, y ello hace que las aplicaciones de filtrado sean de gran interés. Por tanto, las técnicas de caracterización usadas en US son potencialmente de gran utilidad para fines de filtrado en MRI.

Esta Tesis propone métodos para caracterizar la naturaleza aleatoria de los tejidos en US en diferentes etapas de la adquisición. Con este propósito, las principales suposiciones que originan los modelos probabilísticos son reconsideradas así como la influencia que los procesos de adquisición tienen sobre ellos. Como resultado, se proponen modelos probabilísticos de *speckle* en las etapas de interpolación y compresión logarítmica. Se estudia y modela también la heterogeneidad de los tejidos así como distribuciones que tienen en cuenta la respuesta altamente impulsiva del *speckle*. Además, se proponen algunos métodos matemáticos para caracterizar las distribuciones de probabilidad.

El uso de los modelos propuestos proporciona las bases para desarrollar cuatro aplicaciones: un clasificador probabilístico de tejidos en US, un filtro anisótropo de reducción de *speckle*, un método de estimación de parámetros de compresión logarítmica realista y un método de filtrado en MRI como extensión de los modelos probabilísticos propuestos para US.

ABSTRACT

Tissue characterization in medical images is of paramount importance for diagnostic purposes, since it permits physicians to differentiate between healthy and diseased tissues. Several techniques exist for providing such a characterization. Most common ones make use of pattern recognition algorithms, which learn features from textures and their capability of extrapolation allows them to provide a classification of tissues. Other approaches make use of the underlying physical processes of the acquisition and try to characterize the mechanisms that produced the received signals. The main advantage of these techniques is that their performance is not limited by their capability to learn patterns from training sets. Conversely, they depend on the accuracy of the assumed physical models. The Thesis here proposed studies tissue characterization from this perspective. Specifically, it focuses on the probabilistic characterization of tissues in ultrasonic (US) and magnetic resonance images (MRI), since their underlying physical models give rise to probabilistic models describing the resulting patterns.

The statistical description of US provides an important information of the back-scattered echo from tissues, which exhibits a characteristic pattern known as speckle. The parameters of the statistical models allow identifying the features of tissues and provide important descriptors for classification. Many segmentation, filtering, or classification algorithms rely on a Bayesian approach where an accurate statistical model becomes necessary. As a consequence, modeling the statistics of US envelop signals has been a very active area. However, the complex composition of tissues that may show combinations of different kinds of speckle in the resolution cell, and the transformations of data during the acquisition process, make it difficult to provide a proper probabilistic description of signals. This results in the absence of agreement on the probabilistic distributions that characterize tissues in US.

Though the underlying physics that govern the acquisition of MRI and US are completely different, the distributions that model noise in MRI are similar to those of the ultrasonic speckle. In the case of MRI, the presence of noise affects the performance of post-processing techniques such as segmentation, registration, or estimation and, thus, the application of techniques to remove the noise in MRI are of great interest. Hence, tissue characterization techniques used in US can be potentially used for denoising purposes in MRI.

This Thesis aims at characterizing the random nature of tissues in US imaging at different steps of the acquisition process. For this purpose, the main assumptions of the probabilistic models for speckle are revisited, as well as how they are influenced by the different acquisition stages. As a result, new models for the probability distributions of interpolated, filtered, and log-compressed speckle are proposed. The heterogeneity of tissues is also studied and modeled along with other probabilistic distributions that consider the highly impulsive response of speckle. Additionally, new mathematical tools are developed to characterize probability distributions.

The proposed models provide the bases for four new applications within the scope of US and MRI processing: the probabilistic classification of tissues in US; the reduction of speckle via anisotropic filtering; the estimation of the parameters of log-compressing in US acquisition; and the denoising of MRI as an extension of the probabilistic models devised for US.

AGRADECIMIENTOS

Esta tesis es el resultado de años de trabajo desarrollado en el Laboratorio de Procesado de Imagen de la Universidad de Valladolid. Durante estos años he tenido la oportunidad de colaborar con gente que me ha facilitado llegar hasta este punto y con la que me siento en deuda. Antes que nada, me gustaría agradecer muy sinceramente a César Palencia su apoyo sincero, constante y profundo durante todos estos años. También me siento en deuda con Marcos Martín, por su incansable entusiasmo que ha hecho que en los momentos más difíciles siempre se pudiera sonreír, pues siempre hay muchas más opciones.

No quiero perder la ocasión de recordar también todo el apoyo y la magnífica relación con mis jefes y colegas del “proyecto FIS”, los Doctores Carlos Alberola, Pablo Casaseca y Lucilio Cordero, pues con ellos pude tener un contacto directo con la investigación real. También quisiera recordar a todos mis compañeros: Dani, Javier, Susana, Vero, Lucilio, Emma, Rubén, Miguel ángel, Ariel, Diego y Gonzalo. Con todos ellos he tenido la ocasión de poder colaborar, aprender y mejorar. No me olvido de Antonio; de él he podido aprender, pues es gran maestro; y he podido trabajar, pues es un magnífico compañero; pero todo esto es nada comparado con el amigo que es.

Me gustaría agradecer también el apoyo de mis “hermanos mayores”. Al Dr. Federico Simmross por su apoyo y confianza; al Dr. Rodrigo de Luis por todo lo que he podido vivir con él; al Dr. Pablo Casaseca, de nuevo, por su apoyo en todas mis propuestas; a los Dres. Miguel Bote y Guillermo Vega por sus útiles consejos, todos fueron acertados. Me gustaría agradecer especialmente al Dr. Aja por su confianza, que me ha levantado en los momentos más bajos.

No olvido a los compañeros del GSIC/EMIC, con los que he compartido las vivencias que me han hecho sentir parte de algo durante estos últimos años. Agradezco también a mis amigos incondicionales: Miguel y Jorge, Jorge y Rocío, y Manolo.

A mi familia, que son mi referencia, mi principio y mi fin: mi hermano Rodrigo, mi ejemplo; mi madre Teresa, mi fuerza; mi padre Gonzalo, mis valores.

Finalmente quiero agradecer a Beatriz su sonrisa diaria y su optimismo constante que hace que compartir una vida sea algo único.

ACKNOWLEDGMENTS

Many people have been involved in the completion of this thesis: First of all, I would like to thank to Alejandro Frangi for his outstanding support and guidance in Barcelona, and to Constantine Butakoff. Also thanks to João Sanches and José Seabra for their support in Lisboa. To Petia Radeva, Oriol Rodríguez, Ángel Serrano for all the datasets and information. To Rachid Deriche, Aurobrata Ghosh for their interest and time. To Hans Bosch, Theo van Walsum, Alex Haak and Gerard van Burken for such a great collaboration in Rotterdam. To Hans Knutsson and Mats Andersson. And, finally, to Raúl San Jose for his important advices.

A mi familia

*“Of all the inanimate objects, of all men’s creations,
books are the nearest to us for they contain our very thoughts,
our ambitions, our indignations, our illusions, our fidelity to the truth,
and our persistent leanings to error. But most of all they
resemble us in their precious hold on life.”*

Joseph Conrad, 1857-1924

Contents

Resumen	V
Abstract	VII
1 Introduction and Summary	1
1.1 Problem Statement	3
1.1.1 Ultrasound Image Acquisition	4
1.1.2 The Underlying Physics in Ultrasonic Speckle	6
1.1.3 Classical Probabilistic Models in Ultrasonic Images	10
1.1.4 Probabilistic Models in Magnetic Resonance Imaging	14
1.2 Objectives	17
1.3 Contributions	19
1.4 Methodology	20
1.5 Outline of the Thesis	24
1.6 Conclusions	25
I Theoretical Contributions	29
2 Influence of Interpolation in US images	31
2.1 Introduction	32

2.2	Interpolation Model	33
2.3	Goodness of Fit test	36
2.4	Tissue Classification	38
2.5	Conclusions	38
3	A direct calculation of moments of the Sample Variance	41
3.1	Introduction	42
3.2	Theory	43
3.2.1	The problem of Moments of Sample Variance	43
3.2.2	Sketch of the method	44
3.2.3	Main results	45
3.3	An example: Variance of sample variance	51
3.4	Numerical Implementation	53
3.5	Discussion	55
4	Generalized Gamma Mixture Model	59
4.1	Introduction	60
4.2	Materials and Methods	62
4.2.1	Statistical models for describing the nature of speckle	62
4.2.2	Estimation of parameters of the Generalized Gamma	64
4.2.3	Generalized Gamma Mixture Model	70
4.2.4	Implementation Generalized Gamma Mixture Model	75
4.3	Results and Discussion	77
4.3.1	Performance of the ML method	77
4.3.2	Tissue characterization in real US images	81
4.3.3	Performance of the GGMM methods	84
4.3.4	Potential Applications of the GGMM	89
4.4	Conclusions	95
5	Speckle characterization with Heavy-Tailed Models	97
5.1	Introduction	98

5.2	Probabilistic Models	99
5.2.1	Classical Models	99
5.2.2	HT-Rayleigh Model	100
5.3	α -Stable model for Fully Resolved Speckle	101
5.4	Results	102
5.5	Conclusion	104
II	Applications	107
6	Gamma Mixture Classifier for Plaque Detection in IVUS	109
6.1	Introduction	110
6.2	Materials	115
6.3	Probabilistic model for envelope data	118
6.4	Gamma Mixture Model	120
6.5	GMM Classifier	123
6.5.1	Formulation	124
6.5.2	Training and Classifying	126
6.6	Experimental Results	127
6.6.1	Optimal number of components for each Class	127
6.6.2	Plaque classification according to tissue type	130
6.6.3	Statistical Significance	134
6.6.4	Probability Maps	135
6.7	Conclusions	136
6.A	Expectation Maximization for Mixtures of GMM	138
6.B	Solving $\psi(x) - \log(x) = K$	140
7	Probabilistic-Driven OSRAD	143
7.1	Introduction	144
7.2	Probabilistic Estimation of Tissue Classes in US Images	145
7.3	Probabilistic Directional SRAD	146

7.4	Experiments and Results	149
7.5	Conclusion	150
8	Realistic Log-Compressed law for Ultrasound Image Recovery	153
8.1	Introduction	154
8.2	Log-compression Law Model	155
8.3	Parameter Estimation	157
8.4	Results	158
8.5	Conclusions	159
9	Anisotropic LMMSE denoising of MRI based on statistical models	161
9.1	Introduction	162
9.2	Methods	163
9.2.1	Background	163
9.2.2	Gamma Mixture Model	164
9.2.3	Anisotropic-LMMSE	165
9.3	Results	166
9.4	Conclusions	168
	Bibliography	171

List of Figures

1.1	Scheme of ultrasonic system components	4
1.2	Coordinate system for the scattering in inhomogeneous regions	7
1.3	Multi-coil acquisition process	14
1.4	General scheme of the work proposed including its context, aimed objectives and contributions	18
2.1	Characteristic functions for Gamma, Normal and Interpolated Rayleigh RVs	35
2.2	Fully formed speckle pattern and histograms of Rayleigh esti- mate and reconstructed image	36
2.3	Acceptance rate of the goodness-to-fit test for each distribution.	37
2.4	Probability of passing the goodness-of-fit test for a Gamma in 200 independent experiments for an increasing number of samples.	38
2.5	Real cardiac ultrasound image and simulated image and his- togram fitted with the Rayleigh interpolated model and Pseudo- Rayleigh model	39
3.1	All possible combinations of $M = m + n$ groups of N indexes per group	44
4.1	Some examples of the GG distribution for the parameters of the synthetic dataset.	78
4.2	Results for D_{KL} and D_{KS} for 10^4 samples.	79

4.3	Results for the relative error of the estimates for 10^4 samples . . .	80
4.4	Example of the fitting performance for 10^4 samples	82
4.5	Results for D_{KL} and D_{KS} for 500 samples	83
4.6	Results for the relative error of the estimates for 500 samples	84
4.7	Example of an image of the data set	85
4.8	Results for the relative error of the estimates for 500 samples	86
4.9	Simulation of spatial variant echolucent response of tissue . .	87
4.10	Simulation of spatial variant density of scatterers	88
4.11	Simulation of an anatomic phantom of a kidney scan.	89
4.12	Probability of belonging to each component of the GGMM fitted to the image in Fig. 4.11.(b)	90
4.13	Results for D_{KL} and D_{KS} of the Mixture Models: $GGMM_1$, $GGMM_2$, GMM and NMM	90
4.14	Probability of belonging to each tissue class	92
4.15	Anisotropic behavior of the filter	93
4.16	Results of the POSRAD filter	93
4.17	Undirected graph	94
4.18	Probability of belonging to each tissue class after the LBP . . .	94
4.19	Simple example of the valuable information of the posterior probability obtained from the GGMM with spatial coherence .	95
5.1	Fitting result for alpha-stable and Gaussian hypotheses on the components of the signal	103
5.2	Kullback-Leibler for amplitude signal fitted with HT-Rayleigh, Nakagami and the proposed method	104
6.1	Intravasulcar Ultrasound image in polar and Cartesian coor- dinates	111
6.2	Image reconstruction process	117
6.3	Log-compressed representation of the envelope and KL diver- gence and KS statistic for Rayleigh, Nakagami and Gamma distributions	118
6.4	Gamma, Nakagami and Rayleigh mixture models fitting for dif- ferent kinds of tissue	122

6.5	Conditional probability scheme of the Gamma Mixture Model classifier	125
6.6	Difference rate of D_{KL} and D_{KS} for GMM, RMM and NMM models.	130
6.7	Examples of pre-segmented IVUS images	135
6.8	Probability maps of belonging to each tissue class in the polar image of the example in Fig. 6.7.(a).	136
6.9	Probability maps of belonging to each tissue class in the polar image of the example in Fig. 6.7.(b).	137
7.1	Example of the regions for blood and myocardial tissue and performance for the goodness-of-fit test	146
7.2	Tissue class probability map for a clinical ultrasound image	147
7.3	Synthetic image filtered with the proposed filter	149
7.4	Noisy and filtered 3D ultrasound of a liver and heart	152
8.1	Transfer characteristics of the Log-Compression Law Model.	157
8.2	Image before the interpolation stage and PDF of the fully formed speckle area of the pre-compressed image	158
8.3	Relative error of σ and uniform norm for the realistic log-compression law	160
9.1	Empirical PDF of a pMRI T1 image	164
9.2	Example of dMRI and filtered images	167
9.3	Quantitative comparison with synthetic T1 and T2 for increasing σ_n	168

List of Tables

3.1	Example of the construction of the set \mathcal{A} for $N = 3$ and $M = 4$	47
3.2	Comparison between a composition problem versus an integer partition problem	48
3.3	Maximum number of combinations for each algorithm, to calculate the j -th raw moment of the sample variance, $(E\{V^j\})$, up to order 15.	56
4.1	Results of the t-test for Blood.	83
4.2	Results of the t-test for Myocardial Tissue.	85
4.3	Mean values for D_{KS}	85
5.1	Alpha-stable parameters obtained for real ultrasonic signals .	103
5.2	P-values for the unpaired student's t-test	103
5.3	Kullback-Leibler divergence for each signal component	104
6.1	Data set of 50 different images acquired from 9 different patients.	116
6.2	P-values of the Welch t-test for the case of D_{KL} and D_{KL}	120
6.3	Fitting of different types of tissues	123
6.4	Gamma mixture model, Nakagami mixture model and Rayleigh mixture model fitting for different kinds of tissue.	128

6.5 Preferred number of components and free parameters for each tissue class calculated by using the Bayesian information criterion.	129
6.6 Classification Performance for an increasing number of components and free parameters	131
6.7 Performance of tissue classification	132
6.8 Confusion Matrices.	133
6.9 Mean Rank for the Accuracy of the Classifiers.	134
6.10 Classification Performance for three components of the GMM into three classes.	136
7.1 Results on filtering synthetic 2D image with $\sigma_n = 0.5$	151
8.1 Response of the True logarithmic amplifier TDA8780M for small and large signal.	155
8.2 Results for de-compressed real images.	160

List of Algorithms

1	Algorithm to calculate $E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\}$	54
2	Alternative method for $E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\}$	56
3	Implementation of the $GGMM_1$ method.	75
4	Implementation of the $GGMM_2$ method.	76

Introduction and Summary

In my beginning is my end.
Thomas S. Elliot, 1888–1965.

Among the non-invasive imaging modalities, ultrasound imaging is, probably, the most widespread technique. The main reason of its success is that it provides a low-cost way to help diagnosing and can be used in many medical applications. However, ultrasonic (US) images are characterized by the presence of a peculiar granular pattern, the so-called *speckle*.

This term was adopted from the field of laser optics [Goo75] in the early sixties due to the similarity of the patterns between laser optics and ultrasonics. Although the nature of the speckle in US images stems from a different phenomena, there still share some similarities. Both patterns come from the random interference of many coherent wave components reflected from different microscopic elements. In the case of US, the volume, the number of effective scatterers, and the acquisition process contribute to the formation of speckle [Elt06].

The analysis of backscattered echo from tissues needs a proper description of the US signals. For this purpose, and due to the random nature of the speckle, several statistical models have been proposed in the literature. This characterization can be used either for segmentation [Des09], classification [Sea11] or for filtering the speckle itself [Yu02, AF06, Kri07]. The latter usually considers the speckle as an undesired consequence, since it degrades resolution and adds spatial noise to the image. Thus, filtering is commonly applied as a preprocessing step for further segmentation of regions of interest or to extract relevant measures for physiological analysis.

The statistical description of US signals provide an important information of the backscattered echo from tissues. The parameters of the statistical models allow identifying the features of these tissues and provides important descriptors for classification. Some of the filtering algorithms relay on a Bayesian approach where an accurate statistical model becomes necessary. As a consequence, modeling the amplitude statistics of US sig-

nals has been a very active area. Thus, several statistical models have been proposed in last decades.

Probably the most well-known is the Rayleigh model, which is a one-parameter distribution which describes the so-called fully formed (or developed) speckle. This probabilistic distribution of the amplitude of US signals describes the behavior of speckle when a high number of effective scatterers is present in the resolution cell. However, real images show a deviation from this model. This non-Rayleigh behavior can be due to a small number of scatterers in the resolution cell or when there are some dominant components in the cell. The most commonly accepted distributions that try to model non-Rayleigh distributions are the Rice (fully resolved speckle), K (partially formed speckle) and Homodyned K (partially resolved speckle).

Although those models are based on physical assumptions of the backscattering process, some other distributions have proven to provide a good performance on real images. This is the case of Gamma [Tao06, Nil08] and Nakagami [Sha00] distributions. The first is proposed as a two-parameter distribution that describes the result of interpolated/filtered fully formed speckle and also has shown good results in empirical tests among other distributions [Tao06, Nil08]. The Nakagami distribution proposed by Shankar for the case of US characterization [Sha00] is also a two-parameter distribution which generalizes the Rayleigh distribution. This distribution was adopted from the models proposed to describe the statistics of the returned radar echo.

The capability of the Nakagami distribution to model the backscattering from tissues for fully resolved and fully formed speckle render it one of the most commonly accepted model for tissue characterization. However, the tails of the probabilistic density functions of Nakagami, K, Rayleigh or Gamma do not show the impulsive response of speckle which originates heavier tails. In order to describe this impulsive response, a generalized Nakagami distribution was proposed by Shankar [Sha01]. This is a three-parameter model which has shown a better behavior than the Nakagami or Rayleigh, an expected result since it is a generalization of the other models. However, the generalized Nakagami distribution does not have closed-form Maximum Likelihood (ML) estimates and, thus, it makes their use difficult.

In summary, there is no agreement on the probabilistic distributions that model the different nature of tissues in US. This is due to the complex composition of tissues that can show combinations of different kinds of speckle in the resolution cell. Besides, the probabilistic nature of speckle is affected by every acquisition step before the image is fully acquired.

Other modalities present similar probabilistic behaviors when the noise of the image is analyze. The case of magnetic resonance imaging (MRI) show some important similitudes with respect to the probabilistic models of speckle, though the inner nature of the noise patterns stems from a completely different phenomena. Concretely, the noise in MRI is due to the presence of zero-mean uncorrelated Gaussian noise with equal variance in both the real and imaginary parts of the complex spatial frequency

space called the *k-space*. In these conditions, the noise in absence of signal follows a Rayleigh distribution whereas the Rice distribution describes the noisy signal [Gud95].

The presence of noise in MRI affects the performance of post-processing techniques such as segmentation, registration or classification. Thus, the application of techniques to remove the noise in MRI are of great interest. The probabilistic distributions commonly used for characterizing the observed noisy patterns for both US and MRI are similar. In both cases the Rayleigh and Rice distributions as well as other more general distributions are used for describing the image. Consequently, it seems reasonable that all the techniques that ease the characterization of tissues through the acquisition steps of US imaging can be extended and applied for the case of MRI in order to develop new methods for noise removal.

In this Thesis we intend to characterize the random nature of tissues in US imaging through the different steps of the acquisition process. To that end, it will be necessary to develop some mathematical tools to characterize probabilistic distributions of models with an increasing complexity along the acquisition process of the image. Once the tissues and noise are properly characterized, we propose some theoretical and practical contributions to the field of US and MRI imaging.

1.1 Problem Statement

Filtering or segmenting applications based on Maximum Likelihood and Maximum *a posteriori* approaches need a suitable probability densities estimation in different regions. In the case of US imaging, these approaches are usually derived from the analysis of acoustic physics and the information available of the ultrasound probe. However, having the whole information of the acquisition is not so common and many assumptions must be made during the acquisition process. As an example, images provided by practitioners usually do not include the acquisition parameters as gain and/or contrast adjustment. Additionally, some of the steps of the acquisition process are not known depending on the commercial firm of the ultrasound machine.

A common way to deal with this lack of information on probability densities is to use empirical approximations which fit speckle patterns accurately enough to provide good results for filtering or segmenting. This methodology has been used by Tao *et al.* [Tao06] and Eltoltft [Elt06] for different kind of distributions and offers an empirical methodology to test the goodness of fit of the distributions to real data avoiding the problem of propagating the probabilistic distributions through the whole acquisition process. This approach is a conventional technique in the field of statistical inference to test the goodness of fit for different distributions and appears to be a good choice whenever there is no information regarding the transforming processes that change the probabilistic distributions of the received signals.

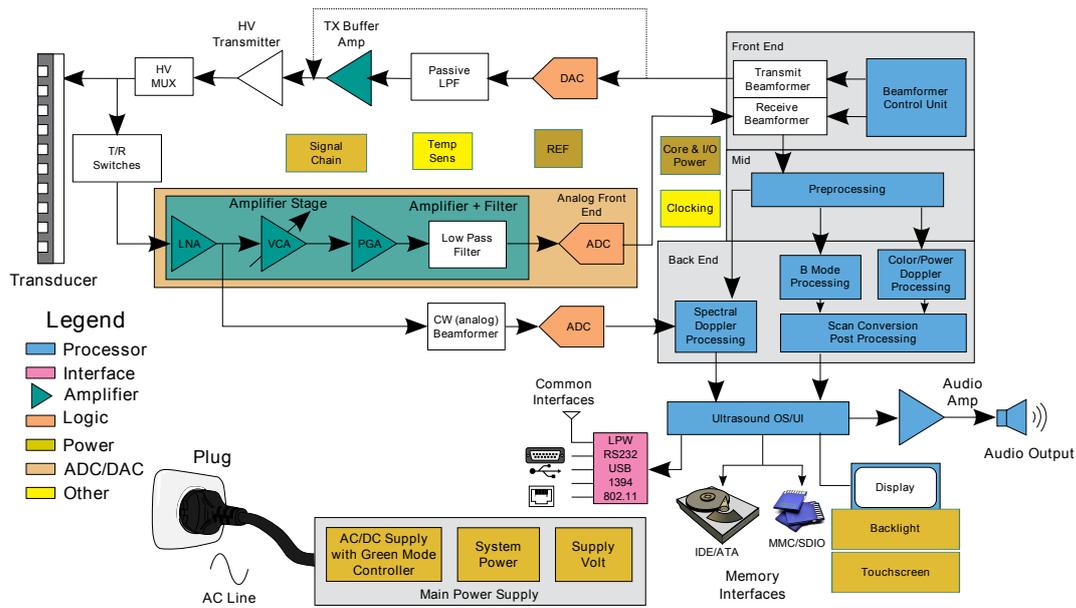


Figure 1.1: Scheme of ultrasonic system components adapted from [Ali08]. This scheme divides processing into three parts: Front-End, Mid-End and Back-end processing. The classical probability distributions model the statistics of speckle without considering Mid-End processing such as Echo Line Averaging, or Back-End Processing such as bilinear interpolation.

In order to provide a relevant study of the probability distributions from both the theoretical and empirical perspectives, all the steps of the acquisition must be studied. In the next section, we explain the steps most commonly used throughout the acquisition process. Then, we provide the fundamentals of speckle from a stochastic perspective and discuss the absence of agreement concerning the probability distributions for speckle characterization. We also introduce an analysis of noise in MRI which links the probabilistic characterization of speckle in US with the probability models derived from the acquisition of MR images. The analysis of speckle and MRI noise will lead us to establish the main objectives of this thesis and the adopted methodology to achieve them.

1.1.1 Ultrasound Image Acquisition

The formation of US images begins with the emission of a pulse packet which travels through the tissue. The backscattering produced by the scatterers in the resolution cell contribute to the change of the pulse shape according to the characteristics of the media, i.e., the number of scatterers as well as their size [Seal1, Sha00].

The ultrasound system focuses sound waves along a given scan line in such a way that the waves constructively add together at the desired focal point. As the pulses propagate through the media, a backscattering process is produced by any object they encounter along their propagation path. The sound waves are focused into a focal point by means of a set of

transducer elements which are energized with time-delayed pulses to produce the corresponding sound waves that propagate through the region of interest. This process to focus the beam is commonly referred to as beamforming. The transducers act either as transmitters or receivers. They can detect any reflected pressure wave produced by a change in tissue density within the region of interest.

The received signals by the piezoelectric elements in the transducer produce low voltage signals that are scaled by using a variable controlled amplifier (VCA) before being sampled by an analog-to-digital converter (ADC). The VCA is configured in such a way that the attenuation of the signal is compensated.

In the receiver beamformer, the signals are scaled and delayed to permit coherent summation of the signals. This new signal represents the beamformed signal for the focal points along a particular scan line. The beamformer operations are usually performed in application-specific integrated circuit (ASIC), field-programmable gate array (FPGA), DSP or a combination of these components [Ali08].

After the beamforming, some preprocessing is performed depending on the US modality and filters are commonly used to reduce band-pass noise. In Doppler mode, the demodulation operation is followed by velocity and turbulence estimation. A color transformation is then applied to represent the velocity and turbulence. In spectral mode, a windowed Fast Fourier Transform (FFT) is applied to the demodulated signal and displayed separately. For B-mode, demodulation followed by envelope detection and log compression is usually applied. The amplitude of the reflected sound waves forms the bases for the ultrasound image.

Envelope detection is used to locate the peaks in the received signal and then logarithmic compression is used to reduce the dynamic range of the received signals for efficient display. The final image can be displayed after envelope detection and logarithmic compression steps for analysis purposes. Usually, a coordinate transformation is needed since the coordinate system of the acquisition is different than the display coordinate system [Ali08].

The US system components are shown in Fig. 1.1 (adapted from [Ali08]). It can be divided into three parts:

- **Front-End Processing.** This part comprises the beamforming and the time gain compensation.
- **Mid-End Processing.** In this stage, the signal is sampled and digitalized. A linear filter is commonly used to improve the signal to noise ratio. This is the so-called Echo Line Averaging. The Envelope extraction is also applied in this stage. It is usually performed by calculating the analytic representation of the signal by means of the Hilbert transform. The magnitude of the resulting complex signal is used as the detected signal for imaging. Additional low pass filtering with decimation or interpolation may be carried out on this signal before presenting it for

further processing. The dynamic range of the signal is then reduced by means of a logarithmic compression.

- **Back-End Processing.** Some operations are still carried out before displaying the information. The exact processing depends on the system configuration. The most common is the scan conversion step, which transforms the samples from polar coordinates to Cartesian coordinates. This step involves a bilinear interpolation technique.

All the described stages of the acquisition change the distribution of speckle and the classical models do not fit its probabilistic behavior any more. Next section describes the fundamentals of probabilistic models in the speckle of the received signals.

1.1.2 The Underlying Physics in Ultrasonic Speckle

The radio-frequency (RF) ultrasound signal in an inhomogeneous random media can be understood as the contribution of multiple ultrasound scatterers represented as vectors with random magnitude and phase in the complex plane. This link between the RF signal and random phasors allow us to derive models that describe the probabilistic behavior of speckle. In this section, we adopt the approach derived by Ng *et al.* [Ng06], which offers an outstanding way to understand the physical assumptions and the consequences in the probabilistic models.

In absence of scatterers, the medium is considered to be uniform with density ρ_0 and adiabatic compressibility κ_0 . The speed c_0 at which acoustic waves travel in this uniform medium is given by

$$c_0 = \frac{1}{\sqrt{\rho_0 \kappa_0}}. \quad (1.1)$$

Ng *et al.* [Ng06] propose to model the presence of scatterers in the medium by adding spatially-dependent terms $\Delta\rho(x)$ and $\Delta\kappa(x)$ to the density and the compressibility, respectively. The total pressure field $P(\mathbf{x}, \omega)$ of the wave propagation follows the partial differential equation

$$\nabla^2 P(\mathbf{x}, \omega) + \left(\frac{\omega}{c_0}\right)^2 P(\mathbf{x}, \omega) = -(\mathbf{S}P)(\mathbf{x}, \omega), \quad (1.2)$$

where \mathbf{S} is the scattering operator defined as

$$\mathbf{S} = \gamma(\mathbf{x}) \left(\frac{\omega}{c_0}\right)^2 - \nabla \cdot \mu(\mathbf{x}) \nabla, \quad (1.3)$$

where the scattering terms are defined as

$$\gamma(\mathbf{x}) = \frac{\Delta\kappa(\mathbf{x})}{\kappa_0}, \quad \mu(\mathbf{x}) = \frac{\Delta\rho(\mathbf{x})}{\rho_0 + \Delta\rho(\mathbf{x})}. \quad (1.4)$$

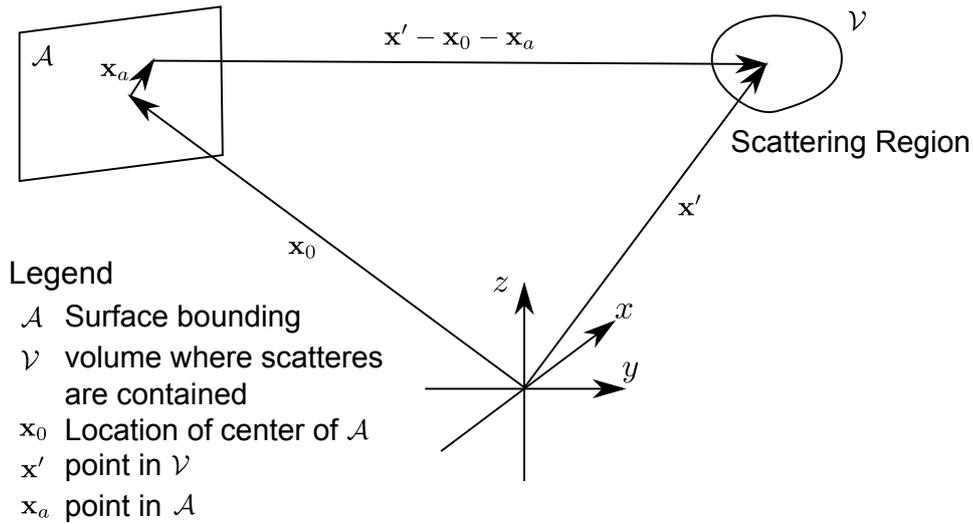


Figure 1.2: Coordinate system for the scattering in inhomogeneous regions adapted from Ng *et al.* [Ng06].

The general solution to Eq. (1.2) can be written as the sum of the solution to the homogeneous equation, $P_i(\mathbf{x}, \omega)$, and the particular solution, $P_s(\mathbf{x}, \omega)$. The physical interpretation of both solutions is that $P_i(\mathbf{x}, \omega)$ is the solution in absence of scatterers and, thus, the *incident pressure field*; whereas $P_s(\mathbf{x}, \omega)$ is the *scattered pressure field*.

The incident pressure field is calculated by assuming that each area element, $d^2\mathbf{x}_a$, of the transducer is infinitesimally small. By applying the Huygen-Fresnel principle¹, the incident field pressure is obtained by summing the spherical wave contribution from each area element. In Fig. 1.2, the notation and the 3D coordinate system for describing the scattering in inhomogeneous regions are shown. The curvature radius of \mathcal{A} is assumed to be large, so \mathcal{A} is considered flat. In these conditions, the incident pressure field is calculated as the Rayleigh integral

$$P_i(\mathbf{x}, \mathbf{x}_0, \omega) = \frac{\rho_0}{2\pi} \int_{\mathcal{A}} j\omega V(\mathbf{x}_a, \omega) \frac{e^{-j\omega/c_0|\mathbf{x}-\mathbf{x}_0-\mathbf{x}_a|}}{|\mathbf{x}-\mathbf{x}_0-\mathbf{x}_a|} d^2\mathbf{x}_a, \quad (1.5)$$

where $V(\mathbf{x}_a, \omega)$ is the temporal Fourier transform of the normal velocity on the transmitting sub-aperture surface. This result can be shortly expressed by introducing a transfer function

$$P_i(\mathbf{x}, \mathbf{x}_0, \omega) = j\omega\rho_0 H_t(\mathbf{x}, \omega), \quad (1.6)$$

where

$$H_t(\mathbf{x}, \omega) = \int_{\mathcal{A}} V(\mathbf{x}_a, \omega) \frac{e^{-j\omega/c_0|\mathbf{x}-\mathbf{x}_0-\mathbf{x}_a|}}{|\mathbf{x}-\mathbf{x}_0-\mathbf{x}_a|} d\mathbf{x}_a. \quad (1.7)$$

The scattered pressure is calculated by using the Green's function method. In this step it is assumed that the waves scattered from the volume \mathcal{V} propagate into an effectively unbounded medium. In this case, the Green's func-

¹The Huygen-Fresnel principle states that each area element contributes with a spherically expanding wave to the incident pressure field.

tion takes the form $\frac{-1}{4\pi|\mathbf{x}-\mathbf{x}'|}e^{-j\omega/c_0|\mathbf{x}-\mathbf{x}'|}$ [Ng06]. Additionally, both $\Delta\rho(\mathbf{x})$ and $\Delta\kappa(\mathbf{x})$ are supposed to be zero outside \mathcal{V} .

So the scattered pressure field is calculated as

$$P_s(\mathbf{x}, \mathbf{x}_0, \omega) = \int_{\mathbb{R}^3} (\mathbf{S}P)(\mathbf{x}', \mathbf{x}_0, \omega) \frac{1}{4\pi|\mathbf{x}-\mathbf{x}'|} e^{-j\omega/c_0|\mathbf{x}-\mathbf{x}'|} d^3\mathbf{x}'. \quad (1.8)$$

The scattering is also supposed to be weak compared to the incident pressure field, so $|P_s(\mathbf{x}, \mathbf{x}_0, \omega)| \ll |P_i(\mathbf{x}, \mathbf{x}_0, \omega)|$. Thus, in Eq. (1.8), one can assume $P(\mathbf{x}, \mathbf{x}_0, \omega) \approx P_i(\mathbf{x}, \mathbf{x}_0, \omega)$. This assumption is referred to as the Born approximation [Jen91] and allow us to express the scattered pressure field in terms of the transfer function as

$$P_s(\mathbf{x}, \mathbf{x}_0, \omega) = j\omega\rho_0 \int_{\mathbb{R}^3} (\mathbf{S}H_t)(\mathbf{x}' - \mathbf{x}_0, \omega) \frac{1}{4\pi|\mathbf{x}-\mathbf{x}'|} e^{-j\omega/c_0|\mathbf{x}-\mathbf{x}'|} d^3\mathbf{x}'. \quad (1.9)$$

For the case of the received voltage, the signal is obtained by summing up the scattered pressure field over the receive sub-aperture and filter it by the electromechanical response of piezoelectric elements. The force perceived in the sub-aperture is defined as $F(\mathbf{x}_0, \omega)$ and can be calculated as the sum over the sub-aperture of an apodization and focusing transfer function, $W(\mathbf{x}, \omega)$ applied to the scattered pressure as follows

$$F(\mathbf{x}_0, \omega) = \int_{\mathcal{A}} W(\mathbf{x}_a, \omega) P_s(\mathbf{x}_0 + \mathbf{x}_a, \mathbf{x}_0, \omega) d^2\mathbf{x}_a. \quad (1.10)$$

Now, defining a receive transfer function, $H_r(\mathbf{x}, \omega)$, for convenience as

$$H_r(\mathbf{x}, \omega) = \int_{\mathcal{A}} W(\mathbf{x}_a, \omega) \frac{1}{4\pi|\mathbf{x}-\mathbf{x}'|} e^{-j\omega/c_0|\mathbf{x}-\mathbf{x}'|} d^2\mathbf{x}_a \quad (1.11)$$

and, by introducing the result obtained in Eq. (1.9), the receive force can be approximated by

$$F(\mathbf{x}_0, \omega) \approx j\omega\rho_0 \int_{\mathbb{R}^3} (\mathbf{S}H_t)(\mathbf{x}' - \mathbf{x}_0, \omega) H_r(\mathbf{x}' - \mathbf{x}_0, \omega) d^3\mathbf{x}'. \quad (1.12)$$

By plugging in the scattering terms defined in Eq. (1.4) and applying the Scatterer operator, the following result is derived (see [Ng06] for more details)

$$F(\mathbf{x}_0, \omega) \approx \frac{j\omega^3\rho_0}{c_0^2} \int_{\mathbb{R}^3} H_t(\mathbf{x}' - \mathbf{x}_0, \omega) H_r(\mathbf{x}' - \mathbf{x}_0, \omega) (\gamma(\mathbf{x}') - \mu(\mathbf{x}')) d^3\mathbf{x}'. \quad (1.13)$$

For a known force over the transducer, the received voltage signal, $R(\mathbf{x}_0, \omega)$, can be calculated by applying the electromechanic transfer function $E_m(\mathbf{x}_0, \omega)$

$$R(\mathbf{x}_0, \omega) = E_m(\mathbf{x}_0, \omega) F(\mathbf{x}_0, \omega). \quad (1.14)$$

Finally, we can calculate the received voltage signal by considering the following definitions

$$f_m(\mathbf{x}) = \frac{\kappa_0 \rho_0^2}{2} \left(\frac{\Delta \kappa(\mathbf{x})}{\kappa_0} - \frac{\Delta \rho(\mathbf{x})}{\rho_0 + \Delta \rho(\mathbf{x})} \right), \quad V_{pe}(\omega) = i\omega^3 E_m(\mathbf{x}), \quad (1.15)$$

where f_m is the scatter field and V_{pe} the electromechanical response. The received voltage signal results is expressed as

$$R(\mathbf{x}_0, \omega) \approx V_{pe}(\omega) H_t(-\mathbf{x}, \omega) H_r(-\mathbf{x}, \omega) \otimes_x f_m(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0}, \quad (1.16)$$

where the symbol \otimes_x stands for the spatial convolution.

In the time domain, the voltage signal can be written as

$$r(\mathbf{x}_0, t) \approx v_{pe}(t) \otimes_t h_{pe}(-\mathbf{x}, t) \otimes_x f_m(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0}, \quad (1.17)$$

where $h_{pe}(-\mathbf{x}, t) = \mathcal{F}^{-1}\{H_t(-\mathbf{x}, \omega)H_r(-\mathbf{x}, \omega)\}$ and \otimes_t stands for time convolution. At this point, the convolution is assumed to be bi-dimensional since the surface \mathcal{A} is restricted to lie on the xy plane. From a signal processing point of view, the electromechanical response $v_{pe}(t)$ and the pulse-echo impulse $h_{pe}(t)$ can be combined into a point spread function $h(\mathbf{x}, t) \approx v_{pe}(t) \otimes_t h_{pe}(-\mathbf{x}, t)$ and the received voltage is calculated as

$$r(x, y, t) \approx \iiint_{\mathbb{R}^3} h(x' - x, y' - y, z', t) f_m(x', y', z') dx' dy' dz'. \quad (1.18)$$

To see the effect in the calculation of the envelope of the signal, the analytic signal is considered by means of the Hilbert transform, $\mathcal{H}_t\{\cdot\}$, in time

$$r_+(x, y, t) \approx \iiint_{\mathbb{R}^3} (h(x' - x, y' - y, z', t) - j\mathcal{H}_t\{h(x' - x, y' - y, z', t)\}) f_m(x', y', z') dx' dy' dz'. \quad (1.19)$$

It is assumed that the center frequency of the RF ultrasound signal and the speed of sound in tissue are reasonably constant. In this case, the complex analytic pulse can be described by its complex envelope

$$h_+(x, y, z, t) = h(x, y, z, t) - j\mathcal{H}_t\{h(x, y, z, t)\} = \tilde{h}(x, y, z, t) e^{j(\omega_0 t - 2k_0 z)}, \quad (1.20)$$

so, the following equation holds

$$r_+(x, y, t) e^{-j\omega_0 t} \approx \iiint_{\mathbb{R}^3} \tilde{h}(x' - x, y' - y, z', t) f_m(x', y', z') e^{-2jk_0 z'} dx' dy' dz'. \quad (1.21)$$

The envelope is then calculated as

$$\tilde{r} = |r_+(x, y, t)| = \left| \iiint_{\mathbb{R}^3} \tilde{h}(x' - x, y' - y, z', t) \tilde{f}_m(x', y', z') dx' dy' dz' \right|, \quad (1.22)$$

where the complex envelope of the point spread function, \tilde{h} , of the imaging system can be viewed as the complex resolution cell which slowly varies with depth (corresponding to time t , since c_0 is assumed to be constant).

Now, the amplitude of the scatter field, \tilde{f} , is assumed to be formed from the sum of all the scatterers within the pulse envelope. This assumption can be mathematically described as a set of Dirac's functions modeling discrete scatterers in the resolution cell

$$f_m(x, y, z) = \sum_{n=0}^{N-1} b_n \delta(x - x_n, y - y_n, z - z_n), \quad (1.23)$$

so the signal, $s(x, y, t) = r_+(x, y, t)$, is simplified to

$$s(x, y, t) = \sum_{n=0}^{N-1} b_n \tilde{h}(x_n - x, y_n - y, z_n, t) e^{-2jk_0 z_n}. \quad (1.24)$$

Finally, for each (x, y) of the probe surface, the received signal is the following sum of phasors

$$S = s(x, y, t) = \sum_{n=0}^{N-1} \alpha_n e^{j\theta_n}. \quad (1.25)$$

This result is equivalent to a random walk due to the random location of the scatterers in the resolution cell [Sha00] and allow us to understand the speckle as a stochastic process. It will serve as a starting point for the development of probabilistic models for tissue characterization. The classical probabilistic models obtained from the analysis of this result is explained in the next section.

1.1.3 Classical Probabilistic Models in Ultrasonic Images

The speckle strongly depends on the effective scatter density, that is, on the effective number and intensity distribution of the scatters in each resolution cell, their size, their shape, their spatial organization as well as the acquisition instrumentation and the tissue attenuation [Sha93, Sha00, Des10]. In the conditions assumed in the previous section and considering the analysis of Eq. (1.25), the resulting speckle noise can be grouped in the following main categories:

- **Fully developed: infinite effective number of scatterers per resolution cell and no deterministic component, modeled by Rayleigh distribution in the case of one-parameter distribution.**

This model arises from the application of the Central Limit Theorem to the sums of Independent and Identically Distributed (IID) random variables. In this case, the phases of Eq. (1.25) are assumed to be uniform random variables in $[0, 2\pi]$. Then, the envelope of the backscattered signal echo, $R = \sqrt{X^2 + Y^2}$ is Rayleigh distributed as [Ray80, Goo75, Wag83, Wag87, Bur78]

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} u(r), \quad (1.26)$$

where $u(\cdot)$ is the Heaviside step function defined as

$$u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \quad (1.27)$$

- **Possibly fully resolved: infinite effective number of scatterers per resolution cell and existence or not of a deterministic component, modeled by Rice distribution.**

Under the assumption of a high number of effective scatterers, but with the presence of resolvable structures in the resolution cell (specular component, C), X and Y become non-zero mean Gaussian distributions. The envelop does no longer follow a Rayleigh distribution but a Rician one as [Ric45, Nak40, Tut88]:

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2+C^2}{2\sigma^2}} I_0\left(\frac{rC}{\sigma^2}\right) u(r), \quad (1.28)$$

where $I_0(\cdot)$ where is the zero-th order modified Bessel function of first kind given by

$$I_0(x) = \frac{1}{\pi} \int_0^\pi \exp(x \cos \theta) d\theta. \quad (1.29)$$

- **Possibly partially developed: arbitrary effective number of scatterers per resolution cell and no deterministic component, modeled by K distribution.**

When the number of scatterers decreases and the Central Limit Theorem cannot be applied, more complicated distributions are proposed to model the distribution of the envelope. Concretely, the K distribution models the case when the number of scatterers is a random variable itself, which is modeled as a Poisson statistical distribution whose local mean is Gamma distributed, this is equivalent to consider σ as Gamma distributed as [Elt06, Jak80, Jak84, Sha93]

$$f_R(r|\sigma) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} u(r) \quad (1.30)$$

and

$$f_\sigma(\sigma) = \frac{1}{2b^2} \frac{1}{\Gamma(\nu+1)} \left(\frac{\sigma}{2b^2}\right) e^{-\frac{\sigma}{2b^2}} u(\sigma). \quad (1.31)$$

So, the probability density function (PDF) of R is

$$f_R(r) = \int f_R(r|\sigma)f_\sigma(\sigma)d\sigma = \frac{2}{b\Gamma(\nu+1)} \left(\frac{\sigma}{2b^2}\right)^{\nu+1} K_\nu\left(\frac{r}{b}\right) u(r), \quad (1.32)$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind

$$K_\nu(x) = \int_0^\infty \exp(-x \cosh t) \cosh(\nu t) dt. \quad (1.33)$$

- **Possibly partially resolved: arbitrary effective number of scatterers per resolution cell and existence or not of a deterministic component, modeled by K-Homodyne distribution.**

A generalization of the previous models appears when a specular component is considered and the number of scatterers, N follows a negative binomial distribution. This is the case of the K-Homodyne distribution [Dut94, Jak87]. Its PDF is given by

$$f_R(r) = r \left(\int \frac{x}{1+x^2\sigma^2/2\nu} J_0(xC) J_0(xr) dx \right) u(r). \quad (1.34)$$

This PDF has no closed expression and, thus, its use is limited.

Note that the most general case of speckle is the possibly partially resolved, which is modeled by the K-Homodyne distribution. The K distribution is a special case of the K-Homodyne distribution with no deterministic component, the Rice distribution is the limiting case corresponding to an infinite effective density of scatterers per resolution cell and the Rayleigh distribution is a special case of the Rice distribution with no deterministic component.

On a completely different approach, Shankar [Sha00] proposed a Nakagami distribution as a “*simpler universal model for tissue characterization*”. Unlike the previously reviewed models, the Nakagami is not based on physical arguments or on a bottom-up modeling of the scattering process. The Nakagami PDF is as follows

$$f_R(r) = \frac{2m^m r^{2m-1}}{\Gamma(m)(2\Omega)^m} e^{-\frac{m}{2\Omega} r^2} u(r). \quad (1.35)$$

This distribution offers good properties to describe the backscattered echo: the Rayleigh distribution is a particular case of the Nakagami ($m = 1$) and, additionally, when $m > 1$ is similar to the Rice distribution. This is the reason that makes the Nakagami distribution one of the commonly accepted distribution for developed speckle and it is also considered as the two-parameter approximation of the true distribution for all the cases (without log-compression or application of filters) [Des11, Sha00, Des10, Sah11]. Nevertheless, this distribution has some limitations. The Nakagami model can not fit the heavier tails of the empirical PDFs due to the impulsive nature of scatterers [Sha01].

Tao *et al.* [Tao06] and Nillesen *et al.* [Nil08] empirically fitted many distributions to real data and shown that speckle is better described by the Gamma distribution. In those works, generic B-mode images having undergone log-compression and filtering were considered [Clo11], though no theoretical justifications were provided for this better fitting.

The distribution of speckle depends not only on the tissues but also on the acquisition process and the post-processing. The transducer center frequency also affects the distributions. Note that, as the central frequency increases, the size of the range cell shrinks and, thus, the number of scatterers in the range cell goes down and one must expect non-Rayleigh statistics. This can be seen as an additional reason for the better fitting of other two-parameter distributions such as the Gamma distribution. Additionally the post-processing techniques as log-compression and filtering also affects the probability distributions of speckle.

In order to describe the impulsive response of scatterers, Shankar [Sha01] proposed a generalized Nakagami distribution which is essentially a generalized Gamma distribution [Sta65]. This is a three-parameter model which has shown a better behavior than the Nakagami or Rayleigh; an expected result since it is a generalization of the other models.

However, the use of the generalized Nakagami distribution is rather difficult since no closed-form for the ML estimates are available. An approximation of its parameters by the method of moments was proposed by Stacy and Mihram [Sta65], where the authors also expressed the difficulties of obtaining ML estimates: “Closed expressions for solutions to the maximum likelihood equations are highly unlikely”. It is important to remark that an inappropriate use of the results by Stacy and Mihram [Sta65] might introduce highly variable estimates, a delicate point which was not considered by Shankar [Sha01], one of the basic references in this topic.

The absence of agreement on the probabilistic models of speckle along the acquisition process makes the characterization of tissues highly difficult in US imaging. This disagreement stems from the deviation of the physical models with respect to real data. There are several reasons that may cause this deviation. On one hand, a very simplified acquisition model does not consider stages as the down-sampling or filters that dramatically change the statistics of the signal. On the other hand, the presence of heavy tailed distributions after the envelope calculation step evidences that the assumptions of the physical model are also very simplified. In this thesis, we study some of the acquisition stages to see their influence on the probabilistic models of speckle. Additionally, new probabilistic models are also proposed in order to describe the highly impulsive response of scatterers in the resolution cell beyond the Central Limit Theorem assumption. From a practical point of view, some applications that take advantage of the probabilistic characterization of tissues are also proposed.

The application of the probabilistic models proposed in this work can also be extended to other modalities such as MRI. The link between the noise of both image modalities is explained in the following section.

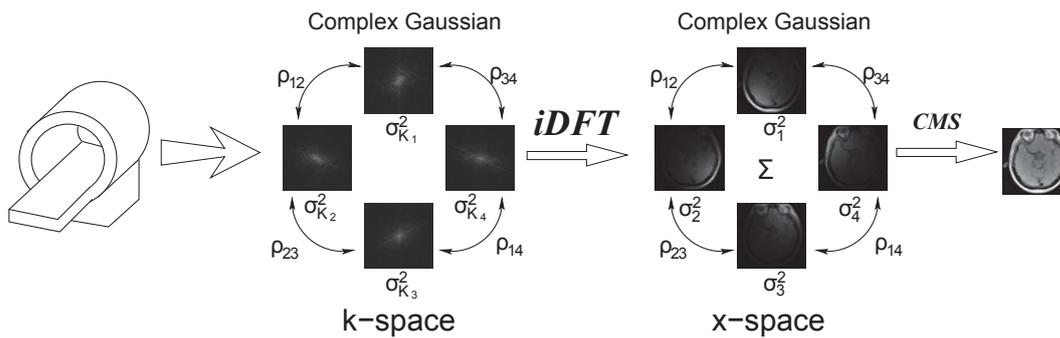


Figure 1.3: Multi-coil acquisition process. The data in both **k**-space and the image domain follow a Gaussian distribution in each coil.

1.1.4 Probabilistic Models in Magnetic Resonance Imaging

Magnetic Resonance imaging is known to be affected by noise during the acquisition process. This noise is an inconvenience that affects not only the visual quality of images, but also affects further processing techniques such as segmentation, registration or fMRI (functional MRI) analysis [McG93, Gud95, AF08b].

Among the most common methods used to reduce the effect of acquisition noise, the most appropriate ones are those that consider the random nature of thermal noise to define the probabilistic model. An accurate noise modeling may be useful not only for filtering purposes, but also for many other processing techniques. For instance, Diffusion Tensor Estimate² have proved to be nearly optimal when the data follows a Rician or a non-central- χ [Sal05] as well as ML and Maximum *a posteriori* (MAP) estimation [And08] or sequential techniques for on-line estimation [Pou08]. The use of an appropriated noise model is crucial in all these methods to attain a statistically correct characterization of the underlying signals.

The noise is usually assumed to be a zero-mean, spatially uncorrelated Gaussian process, with equal variance in both the real and imaginary parts. As a result, in single coil systems the magnitude data in the spatial domain are modeled using a stationary Rician distribution [Gud95]. When multiple (independent) coils are considered, the natural extension of the Rician model yields to a stationary non central- χ distribution, whenever the different images are combined using sum of squares, the variance of noise is the same for all coils, and no correlations exist between them. The derivation of these probabilistic model is as follows:

The **k**-space data is acquired in multiple-shot acquisitions through the repeated application of excitation pulses with a different phase encoding for each readout gradient. Each sampled line of the **k**-space is “frequency encoded”, and the measured signal is uniformly sampled at the desired rate. The primary origin of random fluctuation is due to the so-called thermal noise.

²Diffusion Tensor Imaging has become a very active modality within MRI dealing with fiber pathways estimation in the brain.

Under the assumption that the noise affects equally to all the frequencies without correlation between coils, it can be modeled as a complex Additive White Gaussian Noise (AWGN) process, with zero mean and variance $\sigma_{K_l}^2$ given by [Hen85, Bru91]

$$s_l(\mathbf{k}) = a_l(\mathbf{k}) + n_l(\mathbf{k}; 0, \sigma_{K_l}^2(\mathbf{k})), \quad l = 1, \dots, L, \quad (1.36)$$

with $a_l(\mathbf{k})$ the noise-free signal at the l -th coil (of a total of L coils) and $s_l(\mathbf{k})$ the received signal. If the noise in the RF signal is assumed to be stationary, it makes sense to consider n_l itself stationary, so that we may write

$$n_l(\mathbf{k}; 0, \sigma_{K_l}^2(\mathbf{k})) \equiv n_l(\mathbf{k}; 0, \sigma_{K_l}^2) = n_{l_r}(\mathbf{k}; 0, \sigma_{K_l}^2) + j \cdot n_{l_i}(\mathbf{k}; 0, \sigma_{K_l}^2). \quad (1.37)$$

The complex image domain is obtained as the inverse Discrete Fourier Transform (iDFT) of $s_l(\mathbf{k})$ for each slice and each coil. For the sake of illustration, a pipeline with the distributions involved in single coil acquisitions is depicted in Fig. 1.3. Under the assumption that the data is sampled on a Cartesian lattice, the noise in the complex image domain is still Gaussian for each receiving coil

$$S_l(\mathbf{x}) = A_l(\mathbf{x}) + N_l(\mathbf{x}; 0, \sigma_l^2), \quad l = 1, \dots, L, \quad (1.38)$$

where $N_l(\mathbf{x}; 0, \sigma_l^2) = N_{l_r}(\mathbf{k}; 0, \sigma_{K_l}^2) + j \cdot N_{l_i}(\mathbf{k}; 0, \sigma_{K_l}^2)$. Since the iDFT is applied to each coil, no correlation is assumed among the coils. However, there may be an initial noise correlation between the receiver coils due to electromagnetic coupling [Hay90, Har92]. As a consequence, the noise pattern in the complex image domain may be seen as a complex multivariate (one variable per coil) AWGN process, with zero mean and covariance matrix Σ given by

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2}^2 & \cdots & \sigma_{1,L}^2 \\ \sigma_{2,1}^2 & \sigma_2^2 & \cdots & \sigma_{2,L}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{L,1}^2 & \sigma_{L,2}^2 & \cdots & \sigma_L^2 \end{pmatrix}, \quad (1.39)$$

where $\sigma_{i,j}^2 = \rho_{i,j} \sigma_i \sigma_j$ and $\rho_{i,j}$ is the correlation coefficient between coils i -th and j -th. While this correlation coefficient depends only on the electromagnetic coupling between coils i and j , the variance of noise for each coil may be easily predicted from that in the \mathbf{k} -space as

$$\sigma_l^2 = \frac{1}{|\Omega|} \sigma_{K_l}^2, \quad (1.40)$$

where $|\Omega|$ is the size of the Field of View, i.e., the number of points used in the 2D iDFT.

In the case of a single coil, Eq. (1.38) simplifies to

$$S(\mathbf{x}) = A(\mathbf{x}) + N(\mathbf{x}; 0, \sigma^2), \quad (1.41)$$

with $N(\mathbf{x}; 0, \sigma^2) = N_r(\mathbf{k}; 0, \sigma_K^2) + j \cdot N_i(\mathbf{k}; 0, \sigma_K^2)$ a complex AWGN with zero mean and variance σ^2 . The magnitude signal $M(\mathbf{x})$ is the Rician distributed

envelope of the complex signal, i.e. $M(\mathbf{x}) = |S(\mathbf{x})|$. The Rician PDF is defined as (see Eq. (1.28) for US)

$$f_M(M|A_T, \sigma) = \frac{M}{\sigma^2} \exp\left(-\frac{M^2 + A_T^2}{2\sigma^2}\right) I_0\left(\frac{A_T M}{\sigma^2}\right) u(M), \quad (1.42)$$

where $A_T(\mathbf{x}) = |A(\mathbf{x})|$.

In the image background, where the SNR is zero due to the lack of water-proton density in the air, the Rician PDF simplifies to a Rayleigh distribution with PDF (see Eq. (1.26) for US)

$$f_M(M|A_T, \sigma) = \frac{M}{\sigma^2} \exp\left(-\frac{M^2}{2\sigma^2}\right) u(M). \quad (1.43)$$

In a multiple coil system, if the k-space is fully sampled, the Composite Magnitude Signal (CMS) must be reconstructed from the L complex signals from every coil, $S_l(\mathbf{x})$, with $l = 1, \dots, L$. One of the most used methods is the so-called Sum of Squares (SoS), which can be directly applied over Eq. (1.38) to obtain

$$M_L(\mathbf{x}) = \sqrt{\sum_{l=1}^L |S_l(\mathbf{x})|^2}. \quad (1.44)$$

In an ideal scenario the variance of noise is the same for all the coils, which are assumed to produce uncorrelated samples. The covariance matrix is diagonal with identical eigenvalues: $\Sigma = \sigma^2 \cdot \mathbf{I}_L$, where \mathbf{I}_L is the $L \times L$ identity matrix and $\sigma^2 = \frac{1}{|\Omega|} \sigma_K^2$. Under these assumptions, $M_L(\mathbf{x})$ follows a non-central- χ distribution with PDF given by

$$f_{M_L}(M_L|A_T, \sigma, L) = \frac{A_T^{1-L} M_L^L}{\sigma^2} \exp\left(-\frac{M_L^2 + A_T^2}{2\sigma^2}\right) I_{L-1}\left(\frac{A_T M_L}{\sigma^2}\right) u(M_L), \quad (1.45)$$

with $A_T^2(\mathbf{x}) = \sum_{l=1}^L |A_l(\mathbf{x})|^2$. This distribution particularizes to a Rician one when $L = 1$. In absence of signal, the PDF simplifies to a central- χ PDF as

$$f_{M_L}(M_L|\sigma, L) = \frac{M_L^{2L-1}}{\Gamma(L) 2^{L-1} \sigma^{2L}} \exp\left(-\frac{M_L^2}{2\sigma^2}\right) u(M_L), \quad (1.46)$$

which reduces to a Rayleigh for $L = 1$.

The case of a single coil acquisition is statistically the same as the case of fully formed/resolved speckle in US imaging. In the case of a multi-coil acquisition, the distributions become more general. The correlations between coils may produce a non-stationary noise in images in a similar way as in US images, where the speckle may be non-stationary due to the intrinsic heterogeneity of tissues. This link between both modalities may allow to extend all the methods obtained from US imaging to MRI and vice versa.

With this purpose, the methodology that we follow to achieve the objectives of this thesis is intended to be as general as possible in order to provide mathematical tools and methods for tissue characterization of US and MR medical imaging. The main objectives are stated in the next section.

1.2 Objectives

The characterization of noise in US imaging and MRI described in the previous section evidence the necessity of models that properly describe the probabilistic behavior of the acquired signals, depending on the acquisition protocol. The global objective may be stated as follows:

To provide a general model to characterize the random nature of tissues in US imaging at different stages of the acquisition process. This model should reconsider the assumptions of the classical models in order to cope with the highly impulsive response of speckle.

In order to achieve this goal, some partial objectives can be phrased in the case of US imaging. The partial objectives, the contextual motivations, contributions and applications are depicted in Fig. 1.4 and described next:

[O1] **To analyze the empirical PDFs obtained for different acquisition stages such as interpolation, filtering and log-compression, and propose mathematical models that describe the empirical behavior and highly impulsive response of speckle which originate heavier tails.**

In this partial objective is intended to replicate the experiments performed by Tao *et al.* [Tao06] and Nillesen *et al.* [Nil08], where the classical models were tested in real images in different steps of the acquisition processes. After testing the distributions that better fit the probabilistic behavior of speckle in different tissue classes, some hypothesis about the scattering mechanisms should be derived and tested. This hypothesis will focus on the most common steps of the acquisition processes such as linear filtering, interpolation and log-compression of US images. Additionally, the hypothesis of Gaussianity after assuming a high number of effective scatterers will be tested for real cases that exhibit a heavy-tailed behavior.

[O2] **To develop a mathematical methodology to describe sample statistics probability distributions to characterize the local statistics of images.**

An important tool for approximating distributions of sample statistics is provided by the Edgeworth expansion [Bli98]. This expansion approximates the statistical distribution around the standard normal distribution by means of combinations of Hermite polynomials and coefficients depending on the moments of the distribution, though other polynomials can be used [Roy62, Tan77]. For this reason, a suitable methodology to calculate moments of sample probability distributions in different scenarios becomes necessary. The aim of this goal is to develop a mathematical methodology that eases the characterization of distributions by approximating distributions.

[O3] **To propose mathematical models that describe the non-stationarity of noise and the heterogeneity of tissues per resolution cell.**

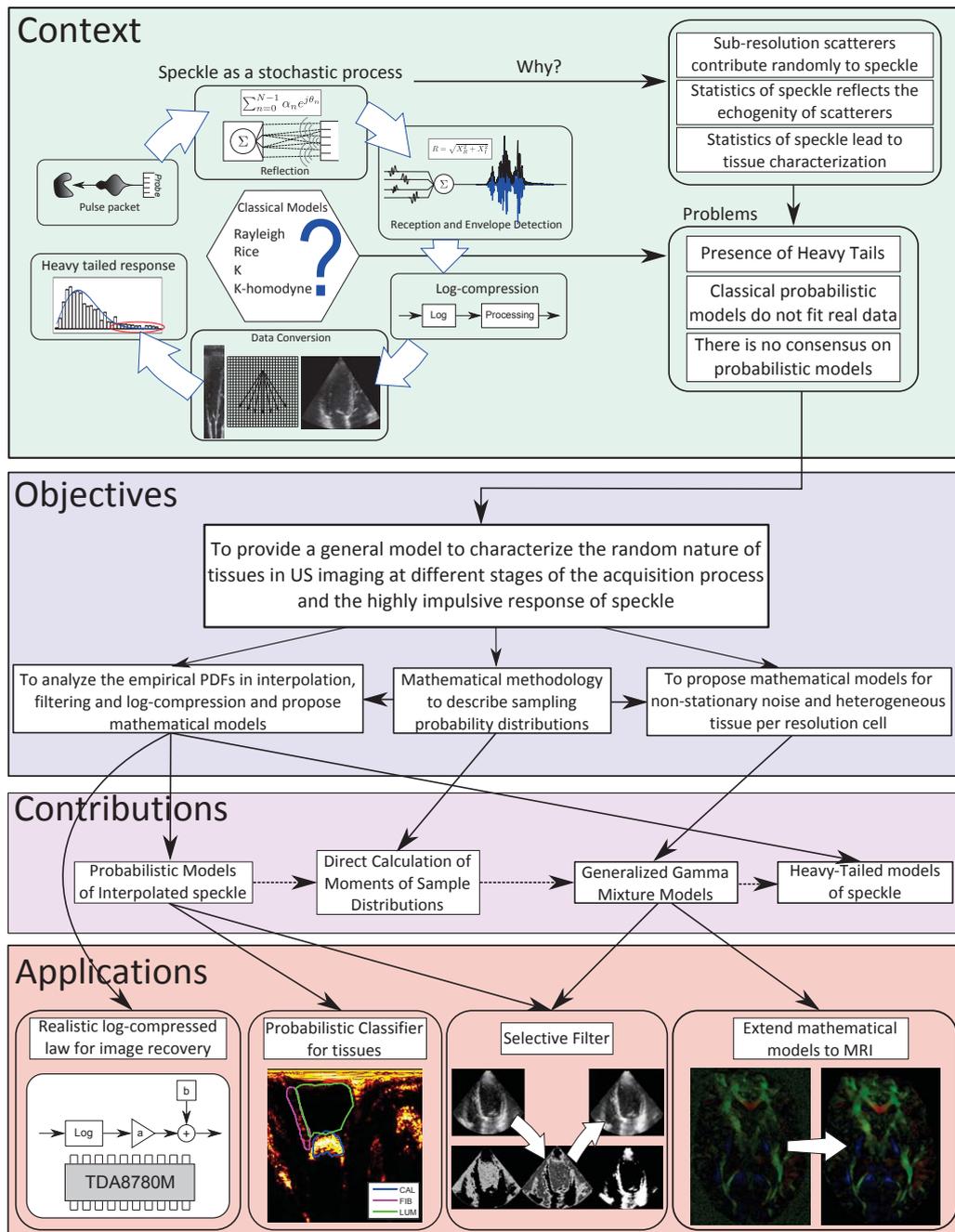


Figure 1.4: General scheme of the work proposed including its context, aimed objectives, contributions and proposed applications. The context shows the acquisition process of US imaging by considering the speckle as a stochastic process for tissue characterization. When the classical models are tested throughout the acquisition process some deviations are detected and not theoretically justified. Hence, new probabilistic models are needed to provide a suitable characterization of tissues. As a result of the main objective, both theoretical contributions and applications are depicted.

The echo-morphology of tissues usually appears as the contribution result of different tissue types (components). Hence, this partial objective focuses on the introduction of mathematical methods to obtain combinations of probabilistic models which describe the multiple

echo-lucid responses in the resolution cell.

All these objectives provide tools to develop some applications which make use of the statistical description of speckle. So, the following applications are also pursued:

[O4] **To propose a classifier which takes into account the probabilistic characterization of tissues and provides the probability of belonging to each tissue class.**

A proper characterization of speckle may offer important features to develop a classifier to distinguish between different tissues for clinical purposes. This objective aims at testing the probabilistic models proposed for classifying tissues.

[O5] **To propose a filter with makes use of the probability of belonging to each tissue.**

The capability of distinguishing tissues offers an important *a priori* knowledge to perform a filter that enhances the features and textures of relevant tissues while the non-relevant are filtered. This objective consists in the proposal of a filter that uses probabilistic characterization to filter non-relevant tissues.

As an additional goal, the MRI noise characterization will be studied in order to extend, if possible, the previous methods to MRI. Then, this objective can be formulated as:

[O6] **To extend the mathematical models developed for US imaging to MRI to characterize tissues for filtering and/or segmenting purposes.**

In this objective, some extensions of the methodologies and models obtained for US will be studied. The links between the characterization of noise in both modalities may allow us to adapt the probabilistic models and applications to specific problems in MRI.

1.3 Contributions

In pursuit of these objectives, the following published contributions were presented:

- Publications in indexed international journals:

[J1] G. Vegas-Sánchez-Ferrero, S. Aja-Fernández, M. Martín-Fernández and C. Palencia, *A Direct Calculation of Moments of the Sample Variance*, *Mathematics and Computers in Simulation*, 82(5): 790–804, 2012.

- [J2] G. Vegas-Sánchez-Ferrero, S. Aja-Fernández, C. Palencia and M. Martín-Fernández, *Generalized Gamma Mixture Model for Ultrasonic Tissue Characterization*, Computational and Mathematical Methods in Medicine, No. 481923, 2012.
- [J3] G. Vegas-Sánchez-Ferrero, J. Seabra, O. Rodríguez-Leor, A. Serrano-Vidal, S. Aja-Fernández, C. Palencia, M. Martín-Fernández, J. Sanches, *Gamma Mixture Classifier for Plaque Detection in Intravascular Ultrasonic Images*, Submitted.
- Publications in Conference Proceedings:
- [C1] G. Vegas-Sánchez-Ferrero, D. Martín-Martínez, S. Aja-Fernández, and C. Palencia, *On the Influence of Interpolation on Probabilistic Models for Ultrasonic Images*, In IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), Rotterdam, The Netherlands, pp. 292–295, April 2010.
- [C2] G. Vegas-Sánchez-Ferrero, F. Simmross Wattenberg, M. Martín-Fernández, C. Palencia, C. Alberola-López, *Caracterización de Speckle con Modelos de Cola Pesada*, XXX Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), San Sebastián, Spain, November 2012.
- [C3] G. Vegas-Sánchez-Ferrero, S. Aja-Fernández, M. Martín-Fernández, A. F. Frangi and C. Palencia, *Probabilistic-Driven Oriented Speckle Reducing Anisotropic Diffusion with Application to Cardiac Ultrasonic Images*, In 13th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Beijing, China, Lecture Notes in Computer Science, Vol. 6361, pp. 518–525, September 2010.
- [C4] G. Vegas-Sánchez-Ferrero, D. Martín-Martínez, Pablo Casaseca-de-la-Higuera, Lucilio Cordero-Grande, S. Aja-Fernández, M. Martín-Fernández and C. Palencia, *Realistic Log-Compressed Law for Ultrasound Image Recovery*, In 18th IEEE International Conference on Image Processing (ICIP), Brussels, Belgium, pp. 2029–2032, September 2011.
- [C5] G. Vegas-Sánchez-Ferrero, A. Tristán-Vega, S. Aja-Fernández, M. Martín-Fernández, C. Palencia and R. Deriche, *Anisotropic LMMSE Denoising of MRI Based on Statistical Tissue Models*, In IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), Barcelona, Spain, pp. 1519–1522, April 2012.

1.4 Methodology

The main objective established in this Thesis emerges from the lack of consensus with respect the probabilistic distributions that describe the scattering processes in US imaging. This problem suggests a set of questions that arises in a natural way: “why the classical models do not hold?” and “which assumptions of classical models are not satisfied?”.

In order to see when the classical models fail, a replication methodology is adopted as a starting point and then the following scientific method was followed [Pop34]:

- **Replication.** Works presented by Tao *et al.* [Tao06] and Nillesen *et al.* [Nil08] was replicated to detect the deviations from the classical models by disproving the following Null Hypothesis: “*Rayleigh and Rice distributions fit fully formed/resolved speckle*”. We expect to obtain some intuition about the processes that change the probabilistic behavior of speckle and to find out which assumptions in the classical models are flawed.
- **Hypotheses.** Once the different stages of the acquisition of US images is empirically studied, some hypotheses were stated. They should take into account both the physical processes of scattering and the preconditioning of the acquired signals. The stated hypotheses will offer some predictions about the behavior of the received signal in different stages of the acquisition process. These predictions must be tested in several conditions with synthetic and real data.
- **Designing Experiments.** The hypotheses will be tested by reproducing the acquisition stages (interpolation, filtering, log-compression, etc.) of the US devices and analyzing the influence on the probabilistic models by statistical test of hypothesis such as the χ^2 goodness-of-fit test or the Kolmogorov-Smirnov test.
- **Analysis.** Once the hypotheses are compared to the predictions, some conclusions can be extracted about the advantage of providing a better description about the probabilistic behavior of speckle. The classification capability of the underlying probability distributions will be tested by defining probabilistic classifiers. Additionally, some speckle selective filters can be proposed. The application of the mathematical methodologies and algorithms may be also applied to MRI under some assumptions. This research line will be studied as well.

The methodology adopted for each of the contributions is the following:

The replication of works presented by Tao *et al.* [Tao06] and Nillesen *et al.* [Nil08] was performed by testing the Null Hypothesis “*Rayleigh and Rice distributions fit fully formed/resolved speckle*” in real-life US non-compressed images with a representative set of distributions proposed in the literature: Gamma, Log-Normal, Rayleigh, Normal, Nakagami, Beta, Rician Inverse Gaussian (RiIG), Rice, Exponential and K. The results we obtained in [C1] and [C3] confirmed those of Tao *et al.* [Tao06] and Nillesen *et al.* [Nil08]: The Gamma distribution showed a better fitting of the probabilistic behavior for fully formed speckle, whereas no differences were appreciated between Gaussian and Rice distributions for fully resolved speckle.

In [C1], we hypothesized that the effect of the interpolation step of the acquisition process modifies the probabilistic behavior of fully formed speckle

towards a Gamma distribution. A set of synthetic experiments were designed in order to test this hypothesis. They confirmed that the interpolation of fully formed speckle give similar results on χ^2 goodness-of-fit tests to the ones obtained by Nillesen *et al.* [Nil08] and Tao *et al.* [Tao06], and also those of our works [C1] and [C3].

As a result, the work presented in [C1] pointed out that a deeper study should be done on the interpolation step of the acquisition process of ultrasonic data. For this purpose, some parametric approaches for probability distributions were studied in order to develop a suitable methodology to model transformations such as the interpolation of random variables. We studied the Edgeworth expansion [Bli98] for approximating distributions of sample statistics. Other expansions were also studied [Roy62, Tan77]. These expansions make use of the moments of the transformed random variable and, thus, the interpolation and other transformations such as linear filtering increased their complexity when higher moments were considered. In [J1] we proposed a direct methodology to obtain higher moments for sample statistics. This methodology allows us to calculate, even in closed form, the moments for combinations of random variables.

The analysis of the results obtained in [C1] led to a better characterization of speckle in important tissues such as vulnerable plaque of arteries. A proper characterization of this plaques provide important information to perform a probabilistic classification scheme by using Gamma distributions on the interpolated images. In [J3] we studied the effect of fitting Gamma distributions in non-interpolated images. We found out that the linear filtering performed in the down-sampling stage of the acquisition also changes the statistics of speckle to a Gamma-like distribution. This was due to the effect of a previous linear filtering usually performed after the envelop calculation to avoid aliasing in the down-sampling stage. We tested the performance of Gamma against the Rayleigh or the Nakagami distributions by several t-test. Results always confirmed the better fitting of Gamma models.

Additionally, in [J3] we proposed a Gamma Mixture Model (GMM) to describe the heterogeneity of tissues. The GMM was compared to a Rayleigh and Nakagami mixture models by means of the Kullback-Leibler divergence, the Kolmogorov-Smirnov statistic and Bayesian Information Criterion. Results demonstrated the better performance for tissue characterization. The capability of the mixture model for characterizing the plaques was used to develop a classifier which provides the probability of belonging to each tissue class. This probability can be used as prior information for several purposes such as classification, segmentation and filtering.

The analysis of the highly impulsive response of scatterers in the resolution cell was studied in two different ways. First, we studied the Generalized Gamma distribution proposed by Shankar in [Sha01]. This distributions fuses both the Gamma and the Nakagami distributions and allow to parametrize the decay of the tails, which is an important advantage when compared to its simplifications –the Gamma and the Nakagami. Nevertheless, this distribution did not have the same acceptance as the Nakagami.

The main reason was that no closed forms for the ML estimates exist and the method commonly used shows important drawbacks such as biases. In [J2] we proposed a simple but robust method to calculate the ML estimates of the Generalized Gamma distribution. The method we propose exhibits a good performance without any appreciable bias and a much smaller variance on the estimates when compared to other methods of the literature. An additional advantage of this methodology is that it leads to a simple formulation of the mixture model with Generalized Gamma (GGMM) distributions. This GGMM provide a more general way to approximate the heterogeneous nature of speckle with the advantage of considering the highly impulsive response of speckle.

We completed the study of the highly impulsive response of scatterers in the resolution cell by revisiting the assumptions commonly adopted for a high number of effective scatterers. Concretely, we checked if the Gaussian hypothesis derived from the application of the Central Limit Theorem can be assumed. Instead of this assumption, we purposed to adopt a more general one which assumes α -stable distributions. This study was carried out in [C2] and results demonstrated that the Gaussian hypothesis is rarely accepted with real data. Statistical tests demonstrated that isotropic α -stable distributions are consistent with real data and they provide a significantly better description of heavy-tailed distributions.

The resulting probability distributions obtained from GMM, GGMM or α -stable models provide an important information for a further post-processing. In [C3] we decided to make use of the probabilistic characterization to define an anisotropic filter which takes into account the probability of belonging to each tissue class. This filter was inspired in the diffusion equation and defines its anisotropy as a function of the structure tensor obtained from the probability of belonging to each tissue class which results in a better definition of the edges of tissues when compared to other similar filters.

The study of the distributions throughout the acquisition protocol was completed by studying the effect of real log-compress devices. The expected distribution of fully formed speckle after the log-compression step is the so-called Fisher-Tippett distribution, which is a double exponential function. However, in real cases, the empirical distribution of speckle is far from the Fisher-Tippett, which evidences that the hypothesis of the compression law should be reconsidered carefully. For instance, the Fisher-Tippett tail for lower values does not appear in real cases. This is probably due to the non-logarithmic behavior of the analog amplifiers for small voltage input. In [C4] we studied the effect of a realistic log-compress law for a true logarithmic amplifier TDA8780M. The contribution of this work is a realistic log-compression law model for ultrasound images based on a real amplifier. This amplifier was chosen since it has 72 dB of true logarithmic dynamic range which is large enough for the dynamic range of the input signal. When we tested the effect of this amplifier on real data, we realized that 12% of samples were in the non-logarithmic regime of the amplifier. Thus, the probabilistic behavior of it should be rather different than the expected Fisher-Tippett. A method for estimating the parameters of the model

in order to recover the pre-compressed image is there proposed.

The methods developed in [J2] and [J3] provided a suitable way to describe the heterogeneity of tissues in US. This important characteristic can be used to overcome an important drawback of the Linear Minimum Mean Squared Error Estimate (LMMSE) filter proposed by [AF08a] for the case of Rician noise (one coils acquisition) and extended to non-central- χ in Brion *et al.* [Bri11] (multiple-coil acquisition). The limitation of LMMSE models relies on the way the local moments are estimated. For those voxels corresponding to image edges, the neighborhoods used typically comprise several different kinds of tissues. Hence, the value estimated for the variance is artificially increased due to the superposition of the effect of noise and the multi-modal distribution of gray levels corresponding to each tissue. As a consequence, edge voxels are not properly denoised, since the noisy pattern is mixed-up with the fine details.

In this context, the main advantage of using Gamma or Generalized Gamma is the good fitting to the non-central- χ distribution as was noticed by Patnaik [Pat49]. This allow us to characterize different tissues in the image and to distinguish the tissues in the local neighborhood where the moments are calculated. Thus, the probability of belonging to each tissue class can be used to calculate local moments conditioned to each of the tissues. As a consequence, the estimation of local variance is not biased by the presence of different tissues.

1.5 Outline of the Thesis

This PhD Thesis is elaborated as a *compendium of publications*. Thus, the following chapters are adapted from the previously mentioned contributions.

In the first part of the Thesis we establish the theoretical background. It comprises the following chapters:

- Chapter 2. The influence of interpolation in fully formed speckle is analyzed. This chapter is adapted from contribution [C1].
- Chapter 3. Some mathematical methods dealing with interrelations of random variables are introduced. This chapter is adapted from contribution [J1].
- Chapter 4. We introduce some mathematical methods for the Generalized Gamma distribution to fit highly impulsive speckle. Additionally, a method to derive the Generalized Gamma Mixture Model is presented. This chapter is adapted from contribution [J2].
- Chapter 5. The highly impulsive response of speckle is studied. The α -stable distributions are tested for real images. This chapter is adapted from contribution [C2].

The second part of the Thesis is devoted to the presentation of applications that make use of probability characterization by probabilistic models:

- Chapter 6. A probabilistic classifier is proposed by means of mixtures of mixtures of Gamma distributions. This chapter is adapted from contribution [J3].
- Chapter 7. We present an anisotropic filter driven by the probability of belonging to each tissue class. This chapter is adapted from contribution [C3].
- Chapter 8. A realistic log-compression law and an estimation method of its parameters are proposed. This chapter is adapted from contribution [C4].
- Chapter 9. An extension of the proposed probability distributions and methods for US is presented for the case of MRI. This chapter is adapted from contribution [C5].

1.6 Conclusions

This Thesis studies the acquisition steps of US images and their effects on the probabilistic behavior of speckle. The classical models commonly accepted were revisited and tested in different steps of the acquisition process. This study resulted in a significant deviation of the probability distributions from the classical models. The main goal was to provide a general model to characterize the random nature of tissues at different stages of the acquisition process and, for this purpose, the assumptions of the classical models were reconsidered.

The analysis of the probability distributions in different stages of the acquisition (as stated in objective [O1]) was studied in the following contributions:

- **[C1] for the interpolation step**, where we showed that the Gamma distribution accurately fits this model. This distribution, though was empirically suggested by other authors, was confirmed to arise from the interpolation of Rayleigh data. Additionally, the probability of acceptance of Gamma distributions achieved for synthetic and real experiments evidences that the interpolation model can be fitted with a Gamma distribution and confirms that interpolation should be taken into account for tissue probabilistic estimation.
- **[C4] for the log-compression step**, where the analysis of a real Log-Compression Law Model showed that 12% of the values of the output signal are in the small signal gain area where the amplifier does not act as a logarithmic compressor. This demonstrates that the compression law should be considered very carefully since many values of the image

are incorrectly decompressed when ideal Log-Compression Laws are used. Additionally, a methodology for decompressing is also proposed.

- **[J3] for the down-sampling step**, where it is shown that the linear filtering performed before down-sampling changes the probability distribution in the same way as the interpolation. Thus, the Gamma distribution approach proposed in [C1] still holds. This distribution was compared to the Nakagami and Rayleigh ones in different scenarios resulting in a better performance of the Gamma for all of them.
- **[J2] for the highly impulsive response of speckle**, where the presence of heavy tails was analyzed for real non-compressed and non-interpolated data. Results demonstrated that the assumption of Gaussianity commonly used fails due to the presence of heavy tails in the in-phase and quadrature components. A model based on α -stable distributions with specular component was proposed. It is consistent with real data and improves the goodness-of-fit in comparison to models which make use of the Gaussian hypothesis and do not consider the specular component.

In order to accomplish objective [O2], some mathematical methods were proposed in [J1]. Two theorems were presented as the main results of [J1]. The first one deals with the way to group all possible combinations of a finite set of indexes. The second one calculates any moment of the sample variance. Some corollaries are also proposed which allow calculating moments of different arrangements of random variables.

In the development of probability models some mathematical methods were needed to deal with arrangements of random variables. In [J2] a simple and robust method for estimating the parameters of Generalized Gamma distributions was proposed. The formulation of the proposed method allows generalizing it for mixture models of Generalized Gamma distributions.

In objective [O3], we pursued the characterization of the heterogeneity of tissues and the non-stationarity of noise. The heterogeneity of tissues was studied and some methods were proposed in [J2] and [J3]. Results of both contributions showed that a better characterization is obtained by using mixture models. In [J3], a classifier based on Gamma Mixture Models was tested for describing the heterogeneous nature of vulnerable plaque in arteries. Results demonstrated the better performance of the proposed method compared to similar and texture-based philosophies. In [J2] the Generalized Gamma Mixture Models were tested in different heterogeneity conditions which justified their use.

In the case of MRI, the non-stationarity of noise and characterization of different tissues was studied in [C5], where the Gamma Mixture Models were used for tissue characterization in parallel MRI. The proposed methodology showed that the inclusion of statistical models lead to a better estimate of local moments by considering the local moments conditioned to the probability of belonging to each tissue class. As a consequence, the proposed

anisotropic LMMSE filter obtains better performance than the conventional ones.

Some potential applications were proposed as a result of a proper characterization. Objective [O4] pursued a classification scheme that takes into account the probability of belonging to each kind of tissue by speckle characterization. For this purpose, we presented a classifier in [J3] which makes use of Gamma Mixture Models. This classifier scheme can be easily generalized to other more general distributions by using the methods derived in contribution [J2]. The proposed classifier provided better results than other *ad hoc* classifiers proposed in the literature.

As an additional potential contribution, a filter which makes use of the probability of belonging to each tissue was stated in objective [O5]. Two different filters were presented: In [C3] we present an anisotropic implementation of the diffusion equation which detects the most probable structures from the probabilistic characterization. The method offers an advantage over the follow-up of statistical properties of the image in each iteration while preserving and enhancing the structures. Experiments on synthetic and real images show that the proposed method preserves edges of the structures better than other state-of-the-art anisotropic methods. In [C5], an anisotropic implementation of the LMMSE is proposed. The anisotropy is achieved by applying probabilistic models to the tissues based on the Gamma Mixture model. This mixture has proved to properly fit the probabilistic behavior of noise in conventional MRI (Rician) and parallel MRI (non-central- χ).

Some extensions to other image modalities such as MRI were intended in the statement of objective [O6]. The methods developed in [J3] were successfully used for modeling the noise in MRI. However, there is still much work to do in analyzing the non-stationarity of noise in parallel MRI and our future research lines will follow this direction.

Many other applications can make use of probabilistic tissue characterization. In the future, we will extend the use of the proposed methods and distributions for developing segmentation algorithms such as Active Shape Models driven by probability maps and speckle tracking methods based on the time coherence of speckle in US sequences.

Several contributions have been proposed to cover all the objectives stated. We hope they may offer a versatile family of distributions and methodologies to characterize speckle in US images for a wide number of applications. We proposed some of them and many others may be developed. Nevertheless, there is still work to do when real devices are considered due to the non-linear regime of devices which may lead to the highly impulsive responses of speckle. Thus, the response of speckle in US imaging is not only due to its physical nature, but also depends on the devices used to measure it.

The methods here proposed were intended to be as general as possible in order to extend their use to other modalities such as MRI, RADAR or

the characterization of fading channels in communications. As an example, the mathematical methods proposed in [J1] also offer a simple way to deal with sample statistics distributions which may be used for developing test of equality of variances and covariances. Other uses are gene classification and DNA analysis [Mur09] or studies on spectrum of nonlinear matter [Sch08].

Part I

Theoretical Contributions

On the influence of Interpolation on Probabilistic Models for Ultrasonic Images

*What we observe is not nature itself,
but nature exposed to our method of questioning.*
Werner Heisenberg, 1901-1976.

Abstract- The influence of the Cartesian interpolation of ultrasound data over the final image statistical model is studied. When fully formed speckle is considered and no compression of the data is done, we show that the interpolated final image can be modeled following a Gamma distribution, which is a good approximation for the weighted sum of Rayleigh variables. The importance of taking into account the interpolation stage to statistically model ultrasound images is pointed out. The interpolation model here proposed can be easily extended to more complex distributions.

Adapted from: G. Vegas-Sánchez-Ferrero¹, D. Martín-Martínez¹, S. Aja-Fernández¹, and C. Palencia², *On the Influence of Interpolation on Probabilistic Models for Ultrasonic Images*, In IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), Rotterdam, The Netherlands, pp. 292–295, April 2010.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Departamento de Matemática Aplicada, Univ. Valladolid.

2.1 Introduction

Some image filtering and segmentation techniques for ultrasound imaging, as those approaches based on maximum likelihood and maximum *a posteriori* [Tao06], rely on an accurate statistical model for the different regions in the image. This model is usually derived from the analysis of the acoustic physics and the information available of the ultrasound probe. However, the whole information during the acquisition process is not usually available, and therefore some suppositions must be considered. For example, images provided by practitioners usually do not include the acquisition parameters as gain and/or contrast adjustment. Additionally, some of the steps of the acquisition process may be unknown, depending on the commercial firm of the ultrasound equipment.

When estimating probability density functions (PDFs) for filtering or segmentation, a common way to deal with the lack of information is to use empirical approximations which accurately fit these PDFs to the speckle patterns. This methodology has been used in literature [Tao06, Elt06] for different kind of distributions. It offers an empirical methodology to test the goodness of fit of the distributions to real data in the last step of the acquisition process, avoiding this way, the problem of propagating the probabilistic model through the whole acquisition process.

Speckle in ultrasound image can be seen as a random process whose statistical features provide information about the tissue. Existence of deterministic component in this process depends on the number of obstacles (scatters) into the resolution cell and their size in comparison with the wavelength of ultrasound signal. Depending on the scatter number density per cell (SND), four types of speckle can be defined: (1) Fully formed: large number of scatters and non-existence of deterministic component, modeled by Rayleigh distribution. (2) Fully resolved: large number of scatters and existence of deterministic component, modeled by Rice distribution. (3) Partially formed: non-large number of scatters and non-existence of deterministic component, modeled by K distribution. (4) Partially resolved: non-large number of scatters and existence of deterministic component, modeled by K-Homodyne distribution.

In this paper we will focus on the first model, i.e. fully formed speckle. In fully formed speckle regions [Bur78], acquired signal can be modeled following a Rayleigh distribution. However, to form the final Cartesian image, these Rayleigh distributed data have to be interpolated. Thus, the resulting image will no longer follow a Rayleigh distribution. Our aim will be to model this final distribution taking into account the interpolation process. Albeit being the simplest of the proposed models, the initial Rayleigh distribution considered can be found in several practical situations, for example in ultrasound imaging of blood. Blood cells behave like tiny randomly distributed scatters, so blood speckle can be classified into fully formed speckle.

If no compression of the data is done (which will be an assumption throughout the paper), we show that the some results provided in litera-

ture [Tao06, Elt06] hold with the interpolation probabilistic model here presented and that this approach can be extended easily to other distributions. We present an exhaustive statistical test using real-life cases for cardiac ultrasonic images which will confirm that the interpolation model accurately fits the data.

The paper is structured as follows: In section 2.2 the interpolating transformation is studied and a distribution is proposed for approaching the theoretical distribution. In section 2.3 the proposed distribution is tested and compared with other works of the literature. In section 2.4 the interpolation model is used for classifying tissues for real and simulated images. Finally, in section 2.5 we conclude analyzing the results.

2.2 Interpolation Model

Ultrasonic images are constructed from a number of acoustic “lines” or vectors usually organized in a sequential pattern [Sza04]. These vectors form lines in the image after conversion by envelope detection. Each line represents a time record of the scattered waves from different depths. The process of image formation begins with a pulse packet emission which travels along the beam vector axis and changes shape according to characteristics of the media. The traveling pulse is scattered by objects placed at a scattering depths and cause delays in the pulse. Reflections are received by the transducer and, considering a constant sound speed, the depths of the scattering objects can be estimated. These intercepted waves are integrated over the surface of the transducer with a suitable weighting and time delays are added for focusing and beamforming. The amplitude of the envelope record is usually logarithmically compressed but this is optional depending on the ultrasonic machine. At this point, when fully formed speckle regions are observed, a Rayleigh probabilistic distribution is often considered [Goo75]. After this step all the lines are interpolated to form a complete Cartesian image from a number of image lines arranged in their geometrical attitude [Sza04].

In this section we discuss the influence of interpolation on the probabilistic model when fully formed speckle regions are considered. Although this strategy can be extended to other distributions, we will pay special attention to the Rayleigh case.

Let $\{X_i\}$ be independent identically distributed (IID) random variables (RVs) which follow a Rayleigh distribution $f_X(x)$

$$f(x) = \frac{x}{\sigma^2} \exp \frac{-x^2}{2\sigma^2}, \quad x \geq 0. \quad (2.1)$$

When a simple scheme of interpolation is considered such the bilinear one in the 2D-case or trilinear in the 3D-case (which is likely to be the one used by the ultrasound machine because of its computational efficiency),

the resultant interpolated value of the pixel can be calculated as

$$Y = \sum_{i=1}^n w_i X_i, \text{ where } \sum_{i=1}^n w_i = 1. \quad (2.2)$$

The resultant PDF of the interpolated RV has no closed expression and several ways for calculating it has been presented in the last years [Bea90]. In this work, we will consider a numerical approach based on quadrature methods due to its simple implementation, since we want to study the behavior of this PDF in order to validate the empirical approaches of the literature.

One simple way to calculate $F_Y(y)$ is to see it as the convolution of the weighted PDFs of independent RV. This way, a closed expression can be obtained when characteristic functions are used since a Rayleigh distribution admits a characteristic function which is known, though not simple [Bea90]

$$\begin{aligned} \phi_X(t) = E\{e^{tX}\} = & 1 + i\sigma t e^{-\sigma^2 t^2/2} \sqrt{\pi} 2 \left(\operatorname{erf} \left(i \frac{\sigma t}{\sqrt{2}} \right) + 1 \right) = \\ & {}_1F_1 \left(1, \frac{1}{2}, \frac{-t^2 \sigma^2}{2} \right) + i \sqrt{\frac{\pi}{2}} t \sigma e^{-t^2 \sigma^2/2}. \end{aligned} \quad (2.3)$$

So, the characteristic function of Y is

$$\phi_Y(t) = \prod_{i=1}^n \phi_i(t), \quad (2.4)$$

where $\phi_i(t)$ is the characteristic function of each $w_i X_i$. Note that w_i affects to X_i in such a way that Y can be considered as the sum of Rayleigh RVs with different σ . The PDF is obtained from Eq. (2.4) by numerical quadrature.

This distribution is not practical to be used in statistical estimation of real data, due to the large number of parameters to estimate. A simplified model must be used. In this paper a Gamma distribution to approximate the exact distribution of the sum of Rayleigh RVs is considered. Note that a Gamma PDF has only two free parameters and the behavior of the tails is *similar* to the PDF of Y . In addition, in literature Gamma PDFs has also been used to model this kind of speckle [Tao06], but no justification has been given.

In Fig. 2.1 we show the characteristic function of a Gamma ϕ_γ , Gaussian ϕ_G and theoretical sum of Rayleigh ϕ for 4 terms and the error committed in the approximation for an increasing number of terms.

In this figure we can see that the characteristic function of a Gamma distribution offers a better behavior than the Normal distribution even when the Central Limit theorem can be applied. Experiments were made considering the same weights of the RVs, however this approach can be done for arbitrary weights and the result still holds (See Fig. 2.1.c).

In order to test this assumption we simulate Speckle based in the acquisition model in the same fashion as it is done in [Per07]. This method

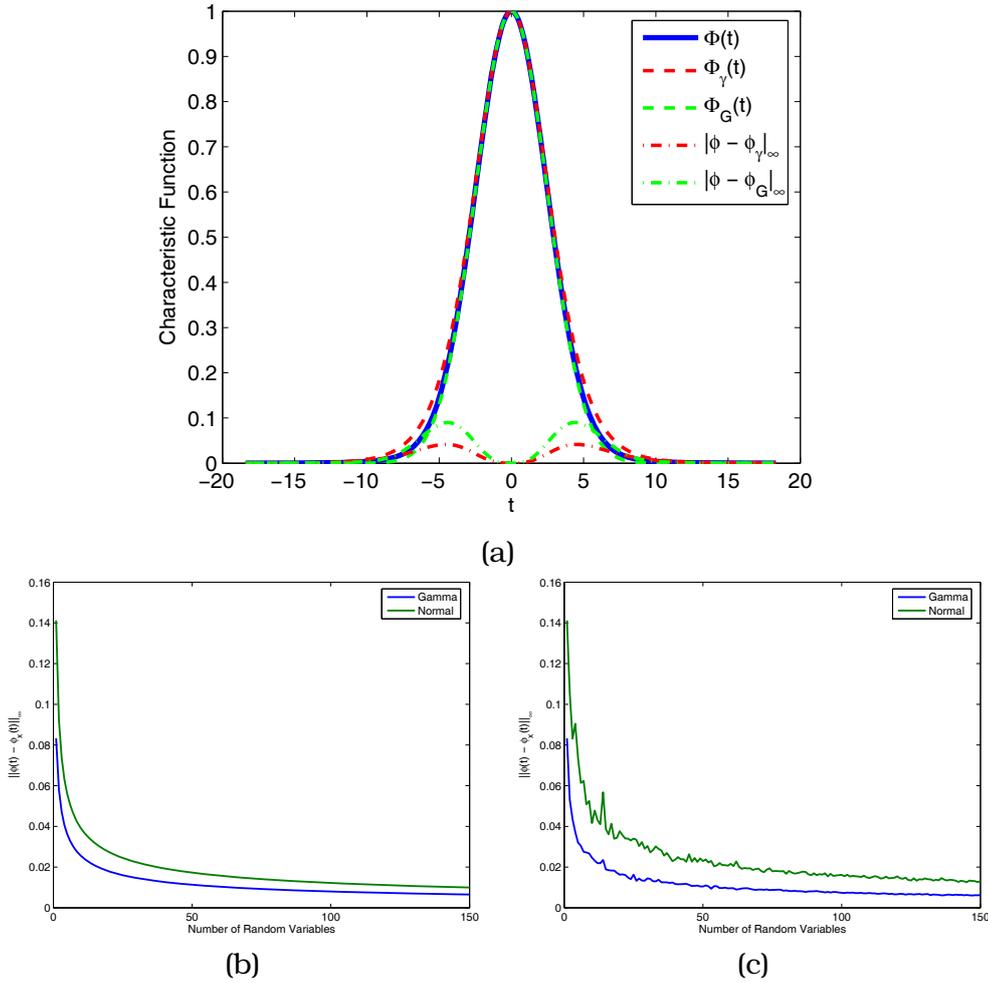


Figure 2.1: Characteristic functions for Gamma, Normal and Interpolated Rayleigh RVs. (a) 4 terms characteristic functions and the error committed (dash-dotted lines). (b) Error of the approach of Gamma and Normal when number of terms increases. (c) Error of the approach of Gamma and Normal when number of terms increases for random weights for each case.

scans an image and records the data in a matrix which is corrupted by means of the speckle formation model of [Bur78] where the tissue is modeled as a collection of scatters that are so numerous and of size comparable to the wavelength. The speckle pattern is obtained by means of random walk which does not assume any statistical distribution in order to avoid any bias of the results.

In Fig. 2.2 we show the reconstructed image when no coherent echoes exist and the number of scatters is high enough to consider the speckle as a fully formed speckle pattern. As it is shown, the histogram of the image before interpolation follows a Rayleigh distribution whereas, in the case of the interpolated image, the result shows clearly that a Gamma distribution accurately fits to the data histogram.

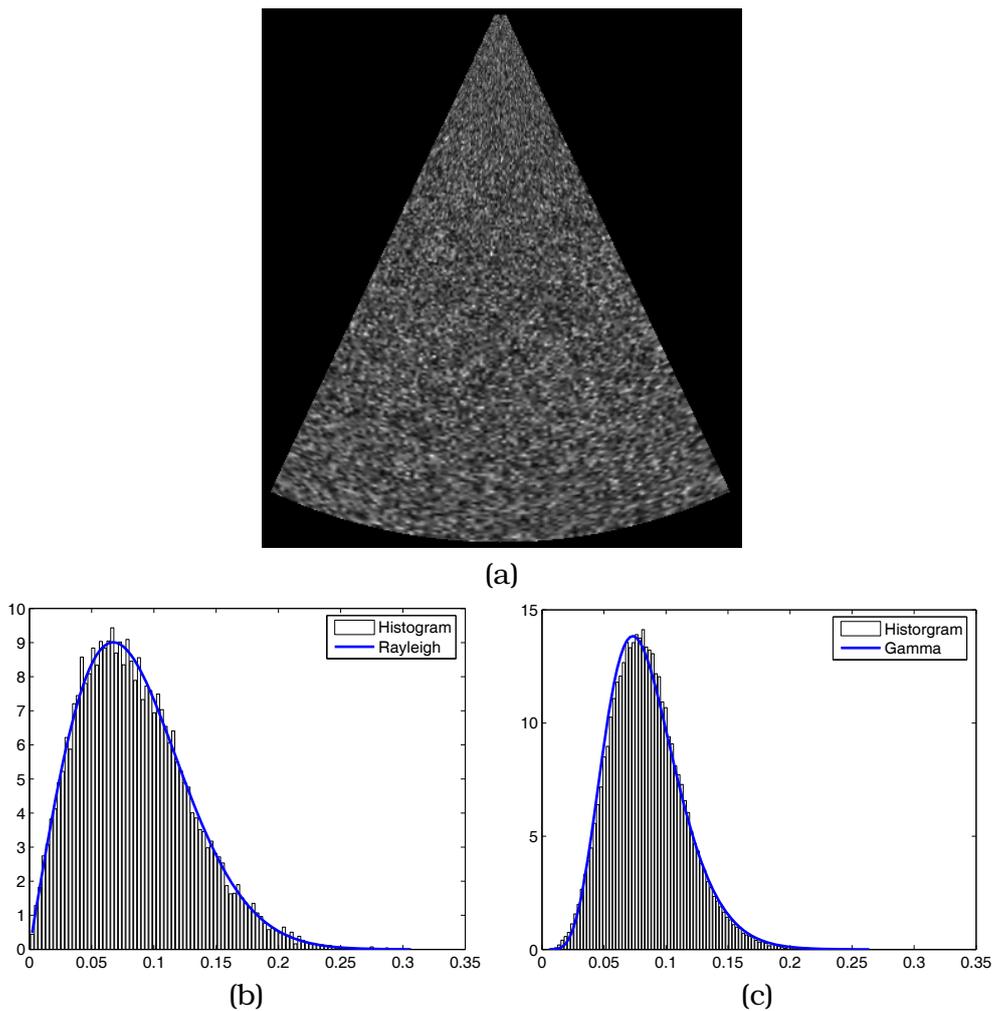


Figure 2.2: (a) Fully formed speckle pattern. (b) Histogram of the received signal and Rayleigh distribution estimate. (c) Histogram of the reconstructed image and Gamma distribution estimate.

2.3 Goodness of Fit test

In this section we study the goodness of fit test in the same way as was done in [Tao06], this test provides the best probability distribution which fits to real data. The study is extended to a representative set of distributions that were proposed in the literature: Gamma, Log-Normal, Rayleigh, Normal, Nakagami, Beta, Rician Inverse Gaussian (RiIG), Rice, Exponential and K. We will show that the Gamma is the distribution that best fits the real data assuming the interpolation model.

The images of the data bank were obtained from a clinical machine *Philips Medical Systems iE33* with the software *PMS5.1 Ultrasound iE33 4.0.1.357* taken to real patients. In this work we have 120 images of size 1024×768 and 8 bits.

In order to obtain a proper statistical model, it is necessary to assume spatial independence between pixels. However, the independence assump-

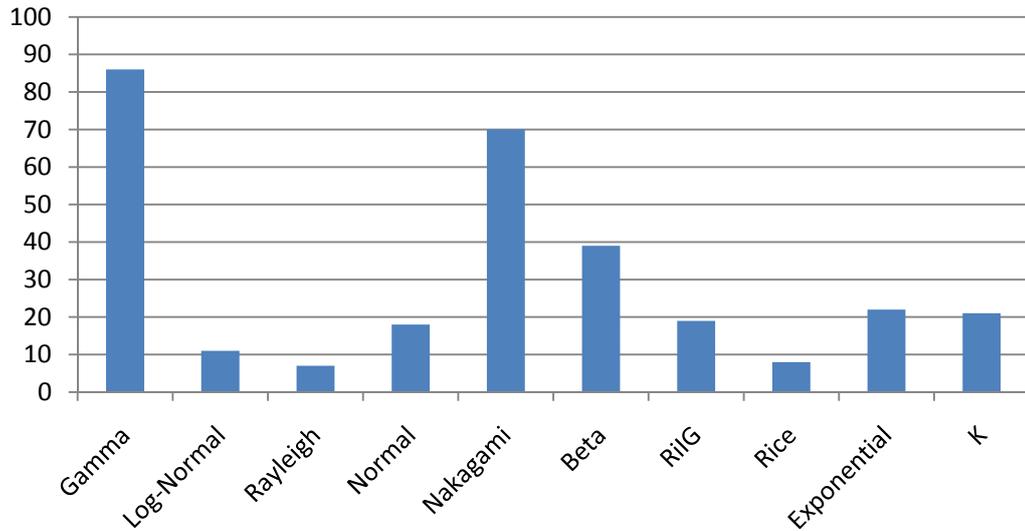


Figure 2.3: Acceptance rate of the goodness-to-fit test for each distribution.

tion does not usually hold. In order to avoid spatial correlation the image is sub-sampled by a factor of 6.

A χ^2 Goodness-to-fit test was done for all the images for an $\alpha = 0.05$ for a representative set of aforementioned distributions in areas with fully formed speckle previously segmented. Results obtained are shown in Fig. 2.3 where a better performance of the Gamma distribution is evident. This result holds with that one obtained in [Tao06] for Gamma distributions and suggest that interpolation of fully formed speckle can be estimated by a Gamma as it was shown before. To confirm this hypothesis we decide to use the χ^2 test for weighted sums of independent Rayleigh RVs for different number of samples and weights. For this test, a bilinear interpolation is considered since real images were obtained from a 2D echography. Weights were chosen to guarantee the contribution of the neighborhood pixels which is the most common situation in the interpolation.

Fig. 2.4 shows the probability of passing the goodness-of-fit test for a Gamma in 200 independent experiments of simulated images for an increasing number of samples. This figure shows that the Gamma distribution has a good behavior for fitting data distribution even for a big number of samples. Segmented areas of the real images (left ventricle) has usually about 500 samples so, for this number of samples, the probability of acceptance Gamma distributions achieved for real and synthetic experiments shows that the interpolation model can be fitted with a Gamma and confirms that interpolation should be taken into account for tissue probabilistic estimation.

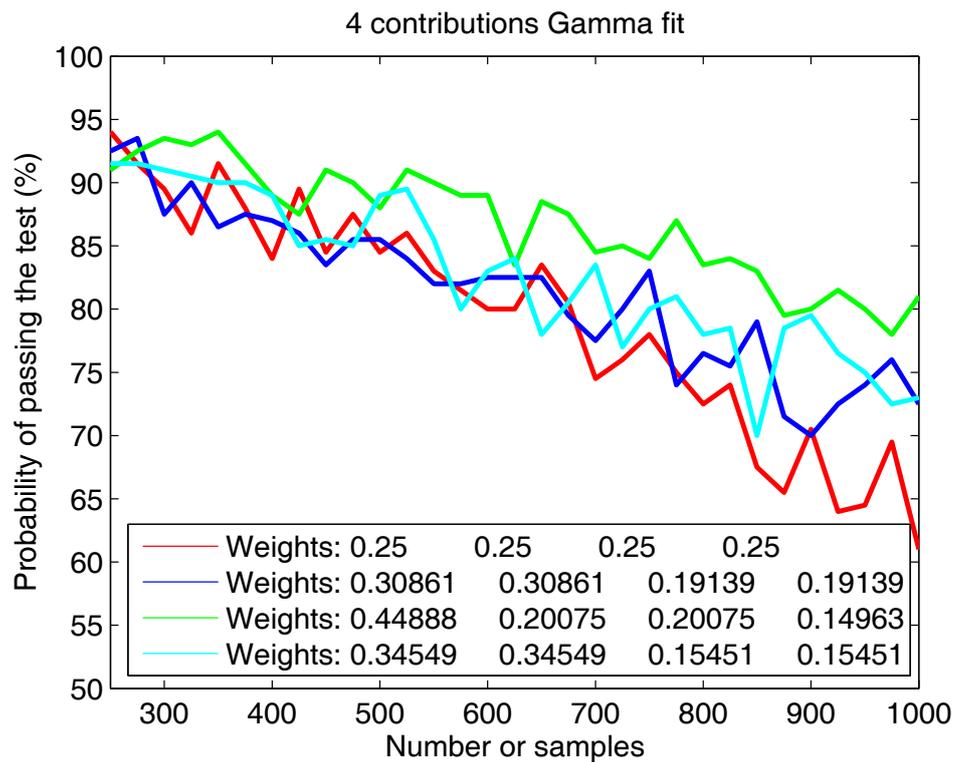


Figure 2.4: Probability of passing the goodness-of-fit test for a Gamma in 200 independent experiments for an increasing number of samples.

2.4 Tissue Classification

In this section we present some results for real and simulated images in order to distinguish between blood and cardiac tissue. For this purpose we will consider a Gamma distribution as a good approach for fully formed speckle for the interpolated image as we saw in previous sections. A normal distribution can be taken for cardiac tissue, this supposition is reasonable when we consider the weighted sum of a Rice or K distribution, though a Gamma distribution is also reasonable for pseudo-Rayleigh speckle. In these cases, the PDF is more symmetric and a Normal provides a good approach. Fig. 2.5 shows a classification by means of distribution fitting for real and simulated images. As we can see, in both cases a Gamma distribution accurately fits the contribution of fully formed speckle.

2.5 Conclusions

In this work we analyze the influence of the interpolation in the probabilistic model of ultrasonic images for the case of fully formed speckle. We show that a Gamma distribution accurately fits to this model. Although the goodness-of-fit of the Gamma was empirically suggested by other authors, in this paper we make clear that this model arises from the interpolation of Rayleigh data. Additionally, we present some synthetic and real results

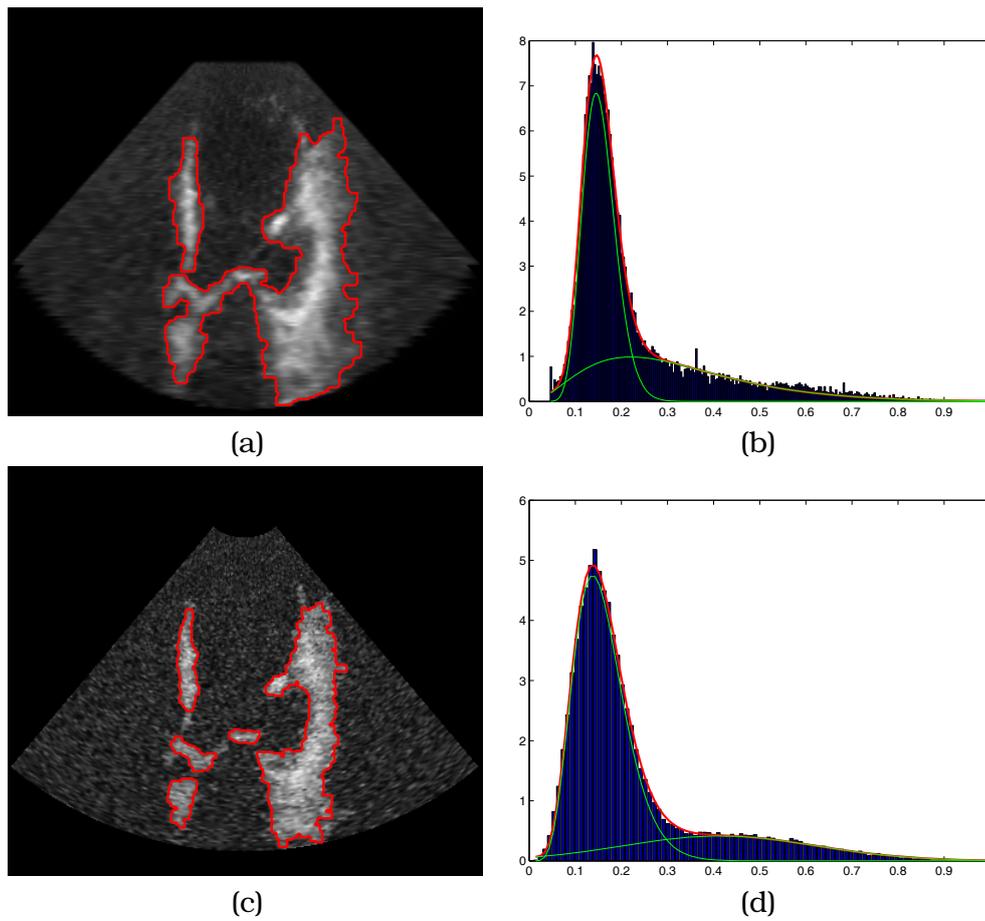


Figure 2.5: (a) and (b): Real cardiac ultrasound image and tissue classification and its histogram fitted with the Rayleigh interpolated model and Pseudo-Rayleigh model. (c) and (d): Simulated image from the real image and its classification for Rayleigh interpolated model and Rice model.

that show that interpolated speckle can be modeled by the same distributions which approximate the theoretical distributions of the interpolated models. These results point out that a deeper study should be done on the interpolation step of the acquisition process of ultrasonic data.

A direct calculation of moments of the Sample Variance

*One doesn't recognize in one's life the really important moments –not until it's too late.
Agatha Christie, 1891-1976.*

Abstract– A systematic method to deal with the interrelations of systems with multi-index quantities (Random Variables) is proposed. The method differs of the well-known Polykays. An application of the theoretical results here presented is the calculation of the moments of the sample variance for general populations in a direct way. The main advantage of the proposed methodology is that no conversion formulae and other complicated Polykays rules are needed. However, the proposed method is compatible with Polykays philosophy and conversion formulae and multiplication rules can be derived by using the theoretical results of this work. For practical purposes, two algorithms for the calculation of the moments of the sample variance are proposed.

Adapted from: G. Vegas-Sánchez-Ferrero¹, S. Aja-Fernández¹, M. Martín-Fernández¹ and C. Palencia², *A Direct Calculation of Moments of the Sample Variance*, *Mathematics and Computers in Simulation*, 82(5): 790–804, 2012.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Departamento de Matemática Aplicada, Univ. Valladolid.

3.1 Introduction

The approximation of the probability density function (PDF) of sample variances from non-normal universes becomes a classical and important problem for studying the deviation from normality on the analysis of variance and covariance [Tan77]. One practical example can be found in [AF10], where authors use the sample (local) variance distribution of Rician and non-central Chi data to model the background/signal areas or magnetic resonance images. Other examples comprise gene classification and DNA analysis [Mur09] and studies in nonlinear matter power spectrum [Sch08].

Some methods have been reported in literature to approximate the PDF of sample variance of general populations, see for instance [Cho04, Cho05, Roy62, Tik65, Tan77]. In order to estimate the PDF, these methods need the computation of higher order moments of the sample variance. This becomes a difficult task, since there is not a closed form to calculate them for general populations. The usual way to do it is by using the Polykays method by Tukey [Tuk56]. However, this is a complex methodology which needs conversion formulae and rules for the product of Polykays that have to be calculated in a non-direct way.

Concretely, Tukey uses the *augmented symmetric functions* for a set of x_1, x_2, \dots, x_n defined as

$$[p_1^{\pi_1}, p_2^{\pi_2}, \dots, p_s^{\pi_s}] = \sum x_i^{p_1} x_j^{p_1} \dots x_q^{p_2} x_r^{p_2} \dots x_u^{p_s} x_v^{p_s}, \quad (3.1)$$

where there are π_1 powers p_1 , π_2 powers p_2 , and son on. All the suffixes are different and the summation takes place over all values of x . Thus, the expression has $n(n-1)(n-2)\dots$ terms in the summation. For example

$$[1^2 23] = \sum x_i^2 x_j x_i^3 x_k; \quad [2^3] = \sum x_i^2 x_j^2 x_k^2.$$

Additionally, Tukey also uses the *monomyal symmetric functions* defined as

$$(p_1^{\pi_1}, p_2^{\pi_2}, \dots, p_s^{\pi_s}) = \frac{[p_1^{\pi_1}, p_2^{\pi_2}, \dots, p_s^{\pi_s}]}{\pi_1! \pi_2! \dots \pi_s!}. \quad (3.2)$$

From these definitions, one can state the following fundamental result

$$E\{[p_1^{\pi_1}, p_2^{\pi_2}, \dots, p_s^{\pi_s}]\} = n(n-1)\dots(n-\rho+1)\mu_{p_1}^{\pi_1}\mu_{p_2}^{\pi_2}\dots\mu_{p_s}^{\pi_s}, \quad (3.3)$$

where $\rho = \sum_{i=1}^s \pi_i$ and μ_i is the i -th raw moment.

The calculation of the moments of sample variance can be defined by expressions including terms with monomial symmetric functions, so the problem is reduced to express the sample variance as sums of augmented symmetric functions. Some tables exist giving these functions in terms of one another. However, the derivation of these tables is not systematic and needs the calculation of power-sums of lower weights. The original tables were heroically (manually) calculated up to 10th order by M.G. Kendall in two different ways and were independently checked by F.N. David in [Dav49].

In this work we propose a systematic method to deal with the interrelations of systems with multi-index quantities (Random Variables) in a different way as it is done with Polykays. Two theorems are proposed that will allow a direct computation of moments of sample variance for general populations. The main advantage of this new methodology is that there is no need of conversion formulae, multiplication tables and other complicated Polykays rules [Tuk56, Tuk50, Dav49]. So, the moments can be calculated in a closed form for any order.

As a result, the calculation of the higher order moments will be straightforward, and the algorithm here proposed will show a computational gain up to order 15.

The paper is structured as follows: Section 3.2 establishes the problem of the moments of the sample variance and presents two theorems as the main results of this work. Additionally, the case of the variance of the sample variance is explained as an example following the theorems presented (section 3.3). Section 3.4 is devoted to the implementation details, where two different implementations of the method are described: a direct one and a more refined method in order to reduce the computations. In Section 3.5, we conclude.

3.2 Theory

3.2.1 The problem of Moments of Sample Variance

Let V be the (unbiased) sample variance of X_1, X_2, \dots, X_N Independent and Identically Distributed (IID) random variables, defined as

$$V = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2 = \frac{1}{N(N-1)} \left(N \sum_{i=1}^N X_i^2 - \left(\sum_{i=1}^N X_i \right)^2 \right). \quad (3.4)$$

The j -th raw moment of the sample variance is

$$\begin{aligned} E \{V^j\} &= E \left\{ \frac{1}{N^j(N-1)^j} \left(N \sum_{i=1}^N X_i^2 - \left(\sum_{i=1}^N X_i \right)^2 \right)^j \right\} = \\ &E \left\{ \frac{1}{N^j(N-1)^j} \sum_{k=0}^j \binom{j}{k} N^k (-1)^{j-k} \left(\sum_{i=1}^N X_i^2 \right)^k \left(\sum_{i=1}^N X_i \right)^{2(j-k)} \right\} = \\ &\frac{1}{N^j(N-1)^j} \sum_{k=0}^j \binom{j}{k} N^k (-1)^{j-k} E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^k \left(\sum_{i=1}^N X_i \right)^{2(j-k)} \right\}. \end{aligned} \quad (3.5)$$

So, according to Eq.(3.5), the problem of the calculation of the raw moments of the sample variance is reduced to the calculus of the expected value of all the combinations of products of sums of two different powers of random variables.

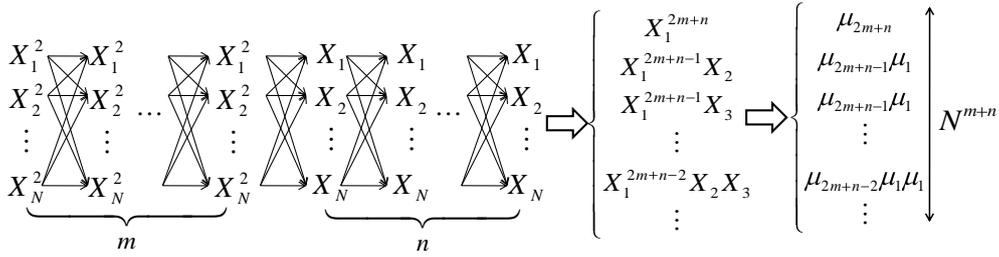


Figure 3.1: This graph represents all possible combinations of $M = m + n$ groups of N indexes per group. Sums of N different indexes are depicted in columns, whereas rows show the product combination. Each path from left to right gives a different combination of products of random variables.

3.2.2 Sketch of the method

For the sake of simplicity, we will refer to this problem in its general form as

$$E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\}. \quad (3.6)$$

The expansion of the product of two sums will give sums of different arrangements of random variables of different indexes. Since X_i are IID random variables, the expectation of those arrangements are the product of the raw moments of the same order as the exponent of the random variable. Hence, all arrangements with the same number of different indexes are equivalent in terms of its expectations and they can only differ in the power of each random variable, i.e. the order of each moment.

This problem can be seen as a graph (see Fig. 3.1), where each vertex is an index of each sum which can be combined with the indexes of the following sums. Thus, this graph represents all possible combinations of $M = m + n$ groups of N indexes per group. Obviously the number of combinations is N^M .

In order to construct the minimum set of combinations, the method will focus on those arrangements of different number of indexes. This will avoid calculating redundant arrangements of indexes. In the next section we give some theoretical results that provide a methodology to calculate the solution of Eq.(3.6) in a suitable way.

Concretely, two theorems will be demonstrated; the first one allows constructing all the combinations as the disjoint union of sets of s^* different indexes with a certain multiplicity, c_{s^*} , for each index. Note that this theorem performs a suitable way to describe the whole set of combinations in order to make an easier way to treat them, but no distinction is taken into account concerning the exponent of each random variable. The second theorem makes use of the result of the first, and allows us to distinguish between variables with different exponents.

3.2.3 Main results

Theorem 1. *Let A be the set of all possible combinations of M groups of indexes with N indexes per group, and let A_{s^*} be a set of all possible combinations of s^* different indexes where $s^* = \{1, \dots, \min(M, N)\}$. Then, the sets A_{s^*} are mutually disjoint sets of the form*

$$A_{s^*} = \bigcup_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \mathcal{B}_{(\mathbf{s}, \mathbf{c})} \quad (3.7)$$

and A is the union of all these sets

$$A = \bigcup_{s^*=1}^{\min(M, N)} A_{s^*} = \bigcup_{s^*=1}^{\min(M, N)} \bigcup_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \mathcal{B}_{(\mathbf{s}, \mathbf{c})}, \quad (3.8)$$

where \mathcal{S}_{s^*} is the set of the combinations of s^* indexes out of N . \mathcal{C}_{s^*} is the set of all possible solutions of the composition of M as the sum of s^* strictly positive integers, and $\mathcal{B}_{(\mathbf{s}, \mathbf{c})}$ is the set of all permutations without repetition of $\mathbf{s} \in \mathcal{S}_{s^*}$ with multiplicity $\mathbf{c} \in \mathcal{C}_{s^*}$.

Proof. The set \mathcal{S}_{s^*} of s^* different indexes out of N possibilities is constructed from the combinations of N elements in groups of s^* elements, thus, its cardinal is $\binom{N}{s^*}$. From this set, we can construct the set A_{s^*} , of all possible combinations of M elements with s^* different indexes. This is done just by knowing the multiplicity of indexes that must hold the following Diophantine equation

$$c_1 + c_2 + \dots + c_{s^*} = M, \quad (3.9)$$

where c_i with $i = 1, \dots, s^*$ is the non-zero multiplicity of each index.

Note that the solutions of Eq.(3.9) are the same as the problem in Number Theory of the composition of a number, M , as a sum of s^* strictly positive integers. It also can be seen as a partition problem when order is taken into consideration. We call \mathcal{C}_{s^*} to the set of solutions of the form $[c_1, c_2, \dots, c_{s^*}]$ of Eq.(3.9) for s^* different indexes.

Every element of the product set $(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}$ defines a unique set of combinations, $\mathcal{B}_{(\mathbf{s}, \mathbf{c})}$ of $\binom{M}{\mathbf{c}}$ elements which is constructed as the permutations without repetition of indexes \mathbf{s} with multiplicity \mathbf{c} . Note that, for a fix $\mathbf{s} \in \mathcal{S}_{s^*}$, $\mathcal{B}_{(\mathbf{s}, \mathbf{c})}$ is unique because, by construction, each pair (\mathbf{s}, \mathbf{c}) defines a different set $\mathbf{b} \in \mathcal{B}_{(\mathbf{s}, \mathbf{c})}$.

Finally, the set A_{s^*} is constructed as the union of all the disjoint sets $\mathcal{B}_{(\mathbf{s}, \mathbf{c})}$

$$A_{s^*} = \bigcup_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \mathcal{B}_{(\mathbf{s}, \mathbf{c})}. \quad (3.10)$$

So the cardinal of \mathcal{A}_{s^*} can be calculated as follows

$$\begin{aligned} |\mathcal{A}_{s^*}| &= \sum_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} |\mathcal{B}_{(\mathbf{s}, \mathbf{c})}| = \sum_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \binom{M}{\mathbf{c}} = \binom{N}{s^*} \sum_{\mathbf{c} \in \mathcal{C}_{s^*}} \binom{M}{\mathbf{c}} \\ &= \binom{N}{s^*} \sum_{k=0}^{s^*} \binom{s^*}{k} (s^* - k)^M (-1)^k. \end{aligned} \quad (3.11)$$

The last equality is obtained by using the multinomial theorem

$$\sum_{\substack{c_1 + \dots + c_{s^*} = M \\ c_i \geq 0}} \binom{M}{c_1, \dots, c_{s^*}} = (s^*)^M \quad (3.12)$$

in combination with the inclusion-exclusion principle of the sets of non-negative solutions of Eq.(3.9)

$$\begin{aligned} \sum_{\substack{c_1 + \dots + c_{s^*} = M \\ c_i > 0}} \binom{M}{c_1, \dots, c_{s^*}} &= \sum_{\substack{c_1 + \dots + c_{s^*} = M \\ c_i \geq 0}} \binom{M}{c_1, \dots, c_{s^*}} - \\ &- \binom{s^*}{1} \sum_{\substack{c_1 + \dots + c_{(s^*-1)} = M \\ c_i \geq 0}} \binom{M}{c_1, \dots, c_{(s^*-1)}} + \dots + \binom{s^*}{s^* - 1} (-1)^{(s^*-1)} \sum_{\substack{c_1 = M \\ c_i \geq 0}} \binom{M}{c_1}. \end{aligned} \quad (3.13)$$

Hence, for each combination of $s^* = 1, \dots, \min(M, N)$ different indexes, the set of all possible combinations, \mathcal{A} , is

$$\mathcal{A} = \bigcup_{s^*=1}^{\min(M, N)} \mathcal{A}_{s^*} = \bigcup_{s^*=1}^{\min(M, N)} \bigcup_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \mathcal{B}_{(\mathbf{s}, \mathbf{c})} \quad (3.14)$$

and its cardinal is, from Eqs. (3.14) and (3.11)

$$|\mathcal{A}| = \sum_{s^*=1}^{\min(M, N)} \sum_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \binom{M}{\mathbf{c}} = \sum_{s^*=1}^{\min(M, N)} \binom{N}{s^*} \sum_{k=0}^{s^*} \binom{s^*}{k} (s^* - k)^M (-1)^k = N^M. \quad (3.15)$$

□

This theorem allows us to calculate all the combinations of indexes in the following way: since X_i are IID random variables, the index of each variable does not matter but the number of different indexes per combination. So, for each number of different indexes, s^* , we have $\binom{N}{s^*}$ equivalent combinations.

The solutions of the combination problem provide different arrangements of the combinations of indexes. However, it is important to note that many of them are equivalent and there is no need to calculate all of them.

$\left(\sum_{i=1}^3 X_i^2\right)^2 \left(\sum_{i=1}^3 X_i\right)^2$				
s^*	$\mathcal{S}_{s^*}, \mathcal{C}_{s^*}$	$ \mathcal{S}_{s^*} = \binom{N}{s^*},$ $ \mathcal{C}_{s^*} $	$ \mathcal{B}_{(s,c)} = \binom{M}{c}$	$ \mathcal{A}_{s^*} $
$s^* = 1$	$\mathcal{S}_1 = \{(1), (2), (3)\}$ $\mathcal{C}_1 = \{(4)\}$	$ \mathcal{S}_1 = 3$ $ \mathcal{C}_1 = 1$	$\mathcal{B}_{((i),(4))} = \{(i, i, i, i)\}$	$ \mathcal{A}_1 = 3$
$s^* = 2$	$\mathcal{S}_2 = \left\{ \begin{matrix} (1, 2), \\ (1, 3), \\ (2, 3) \end{matrix} \right\}$ $\mathcal{C}_2 = \left\{ \begin{matrix} (1, 3), \\ (2, 2), \\ (3, 1) \end{matrix} \right\}$	$ \mathcal{S}_2 = 3$ $ \mathcal{C}_2 = 3$	$\mathcal{B}_{((i,j),(1,3))} = \left\{ \begin{matrix} (i, j, j, j), \\ (j, i, j, j), \\ (j, j, i, j), \\ (j, j, j, i) \end{matrix} \right\}$ $ \mathcal{B}_{((i,j),(1,3))} = 4$ $\mathcal{B}_{((i,j),(3,1))} = \left\{ \begin{matrix} (i, i, i, j), \\ (i, i, j, i), \\ (i, j, i, i), \\ (j, i, i, i) \end{matrix} \right\}$ $ \mathcal{B}_{((i,j),(3,1))} = 4$ $\mathcal{B}_{((i,j),(2,2))} = \left\{ \begin{matrix} (i, i, j, j), \\ \vdots \end{matrix} \right\}$ $ \mathcal{B}_{((i,j),(2,2))} = 6$	$ \mathcal{A}_2 = 42$
$s^* = 3$	$\mathcal{S}_3 = \{(1, 2, 3)\}$ $\mathcal{C}_3 = \left\{ \begin{matrix} (2, 1, 1), \\ (1, 2, 1), \\ (1, 1, 2) \end{matrix} \right\}$	$ \mathcal{S}_3 = 1$ $ \mathcal{C}_3 = 3$	$\mathcal{B}_{((i,j,k),(2,1,1))} = \left\{ \begin{matrix} (i, i, j, k), \\ \vdots \end{matrix} \right\}$ $ \mathcal{B}_{((i,j,k),(2,1,1))} = 12$	$ \mathcal{A}_3 = 36$
$\mathcal{A} = \bigcup_{s^*=1}^{\min(M,N)} \mathcal{A}_{s^*} = \{(1, 1, 1, 1), (2, 2, 2, 2), (3, 3, 3, 3), (1, 1, 1, 2), \dots\}$				$ \mathcal{A} = N^M = 81$

Table 3.1: Example of the construction of the set \mathcal{A} for $N = 3$ and $M = 4$ for each value of $s^* = \{1, 2, 3\}$ different indexes for $(\sum_{i=1}^3 X_i^2)^2 (\sum_{i=1}^3 X_i)^2$.

For instance, in Table 3.1 an example of the construction of the set \mathcal{A} is presented for $M = 4$ and $N = 4$. Note that when X_i are IID random variables, all the compositions that just differ in order are equivalent. Thus, the set of solutions of the composition problem can be dramatically reduced into a smaller set of solutions with multiplicities for each one. The sets of solutions can be calculated as an integer partition problem where order does not matter and the multiplicities $\sigma(\cdot)$ can be calculated as a multinomial coefficient, we denote the set of solutions of the integer partition problem as $\mathcal{P}_{s^*=1, \dots, 4}$

$$\begin{aligned}
 p_1 + \dots + p_{s^*} &= 4 \\
 \mathcal{P}_{s^*=1, \dots, 4} &= \{\{4\}, \{2, 2\}, \{3, 1\}, \{2, 1, 1\}, \{1, 1, 1, 1\}\} \\
 \sigma(\mathcal{P}_{s^*=1, \dots, 4}) &= \left\{ \binom{1}{1}, \binom{2}{2}, \binom{2}{1, 1}, \binom{3}{1, 2}, \binom{4}{4} \right\}.
 \end{aligned}$$

Table 3.2 shows a comparison between the number of solutions of the composition problem and the partition problem. The number of solutions of the composition problem, $|\mathcal{C}_{s^*=1, \dots, M}|$, can be calculated easily by placing either a plus sign or a comma in each of the $M - 1$ boxes of the array

$$\underbrace{(1 \square 1 \square \dots 1 \square 1)}_M. \tag{3.16}$$

M	Composition Solutions $ \mathcal{C}_{s^*=1,\dots,M} $	Partition Solutions $ \mathcal{P}_{s^*=1,\dots,M} $
10	512	42
20	524288	627
30	536870912	5604
40	549755813888	37338
50	562949953421312	204226
60	576460752303423488	966467
70	590295810358705651712	4087968
80	604462909807314587353088	15796476
90	618970019642690137449562112	56634173
100	633825300114114700748351602688	190569292

Table 3.2: Comparison between the number of solutions of Eq.(3.9) when it is considered a composition problem versus when it is considered as an integer partition problem.

Since every composition of M can be determined by an assignment of pluses and commas, the number of compositions is given by the binomial coefficient $\binom{M-1}{k-1}$ for k parts, where $k = 1, \dots, M$

$$|\mathcal{C}_{s^*=1,\dots,M}| = \sum_{k=1}^M \binom{M-1}{k-1} = 2^{M-1}. \quad (3.17)$$

In the case of the partition problem, $|\mathcal{P}_{s^*=1,\dots,M}|$, one can compute the number of solutions by means of the Hardy-Ramanujan-Rademacher formula [And76].

Obviously, this way for calculating the solutions of Eq.(3.9) is highly more efficient as M increases.

In the following theorem we link the arrangements of indexes with the exponents of each random variable.

Theorem 2. Let X_i be IID random variables with $2m+n$ finite raw moments μ_1, \dots, μ_{2m+n} . Then, the following equality holds

$$E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\} = \sum_{s^*=1}^{\min(M,N)} \binom{N}{s^*} \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sigma(\mathbf{p}) \sum_{j=1}^{\binom{M}{\mathbf{p}}} \prod_{k=1}^{s^*} \mu_{E_{j,k}},$$

where $M = m+n$, \mathcal{P}_{s^*} is the set of solutions of the integer partition problem for M into s^* terms. $\sigma(\mathbf{p})$ is a multinomial coefficient which indicates the number of permutations without repetition of $\mathbf{p} \in \mathcal{P}_{s^*}$, $E_{j,k} = \sum_{i=1}^M \mathbf{e}(i) \delta(\mathbf{b}_j(i) - k)$ with $\mathbf{e} = \underbrace{[2, \dots, 2]}_m, \underbrace{[1, \dots, 1]}_n$, δ is the Kronecker delta and $\mathbf{b}_j \in \mathcal{B}_{s^*,\mathbf{p}}$ is a permutation without repetition of s^* elements with multiplicities the elements of \mathbf{p} .

Proof. First we prove that there is no need to construct the set $\mathcal{B}_{s,c}$ for each $s \in \mathcal{S}_{s^*}$ when dealing with IID random variables. This is an important issue for its numerical implementation.

Let $\mathbf{b} \in \mathcal{B}_{s,c}$ be one of the vector of indexes s with multiplicities c . Prod-

ucts of random variables with indexes \mathbf{s} can be defined as

$$\mathcal{X}(\mathbf{b}, \mathbf{s}) = \prod_{k=1}^{s^*} X_{\mathbf{s}(k)}^{E_k(\mathbf{s})}, \quad (3.18)$$

where

$$E_k(\mathbf{s}) = \sum_{i=1}^M \mathbf{e}(i) \delta(\mathbf{b}(i) - \mathbf{s}(k)) \quad (3.19)$$

and $\mathbf{e} = [\underbrace{2, \dots, 2}_m, \underbrace{1, \dots, 1}_n]$. Let $\mathcal{B}_{s^*, \mathbf{c}}$ be the set of combinations of indexes $[1, 2, \dots, s^*]$ with multiplicities \mathbf{c} . Note that this set of combinations does not depend on the elements of \mathcal{S}_{s^*} but on the number of different indexes s^* . So, for a fixed s^* and \mathbf{c} , a bijection can be established between elements of $\mathcal{B}_{s^*, \mathbf{c}}$ and $\mathcal{B}_{\mathbf{s}, \mathbf{c}}$ in such a way that, for every $\mathbf{b} \in \mathcal{B}_{\mathbf{s}, \mathbf{p}}$, there exists one $\mathbf{b}^* \in \mathcal{B}_{s^*, \mathbf{p}}$ such that

$$E_k(\mathbf{s}) = \sum_{i=1}^M \mathbf{e}(i) \delta(\mathbf{b}(i) - \mathbf{s}(k)) = \sum_{i=1}^M \mathbf{e}(i) \delta(\mathbf{b}^*(i) - k) = E_k \quad (3.20)$$

and, $|\mathcal{B}_{\mathbf{s}, \mathbf{c}}| = |\mathcal{B}_{s^*, \mathbf{c}}|$. Hence,

$$E \{ \mathcal{X}(\mathbf{b}, \mathbf{s}) \} = E \left\{ \prod_{k=1}^{s^*} X_{\mathbf{s}(k)}^{E_k(\mathbf{s})} \right\} = \prod_{k=1}^{s^*} E \{ X_k^{E_k} \} = E \{ \mathcal{X}(\mathbf{b}^*) \}, \quad (3.21)$$

which demonstrates the equality

$$E \left\{ \sum_{\mathbf{a} \in \mathcal{A}_{s^*}} \mathcal{X}(\mathbf{a}) \right\} = E \left\{ \sum_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \sum_{\mathbf{b}_j \in \mathcal{B}_{(\mathbf{s}, \mathbf{c})}} \mathcal{X}(\mathbf{b}_j) \right\} = E \left\{ \binom{N}{s^*} \sum_{\mathbf{c} \in \mathcal{C}_{s^*}} \sum_{\mathbf{b}_j^* \in \mathcal{B}_{(s^*, \mathbf{c})}} \mathcal{X}(\mathbf{b}_j^*) \right\}. \quad (3.22)$$

Now, the construction of $\mathcal{B}_{(s^*, \mathbf{c})} = \{ \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{\binom{N}{s^*}} \}$ becomes direct just by creating all the permutations without repetition of a vector of indexes with multiplicity the elements of \mathbf{c} .

All element $\mathbf{b}_j \in \mathcal{B}_{(s^*, \mathbf{c})}$ has a direct relationship with the exponents of the random variables, i.e. the first m indexes are those of random variables with power 2, whereas the last n indexes are those of power 1. So, the order of the moment of the random variable is the sum of all the powers of the same index

$$\mathcal{X}(\mathcal{B}_{(s^*, \mathbf{c})}) = \left\{ \begin{array}{c} X_1^{E_{1,1}} X_2^{E_{1,2}} \dots X_{s^*}^{E_{1,s^*}} \\ X_1^{E_{2,1}} X_2^{E_{2,2}} \dots X_{s^*}^{E_{2,s^*}} \\ \vdots \\ X_1^{E_{\binom{M}{\mathbf{c}},1}} X_2^{E_{\binom{M}{\mathbf{c}},2}} \dots X_{s^*}^{E_{\binom{M}{\mathbf{c}},s^*}} \end{array} \right\}, \quad (3.23)$$

where

$$E_{j,k} = \sum_{i=1}^M \mathbf{e}(i) \delta(\mathbf{b}_j(i) - k)$$

$$\mathbf{e} = [\underbrace{2, \dots, 2}_m, \underbrace{1, \dots, 1}_n].$$

All in all, from Eq.(3.8) and Eq.(3.22), we have

$$\begin{aligned} E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\} &= E \left\{ \sum_{\mathbf{a} \in \mathcal{A}} \mathcal{X}(\mathbf{a}) \right\} \\ &= E \left\{ \sum_{s^*=1}^{\min(M,N)} \sum_{(\mathbf{s}, \mathbf{c}) \in \mathcal{S}_{s^*} \times \mathcal{C}_{s^*}} \sum_{\mathbf{b}_j \in \mathcal{B}_{(\mathbf{s}, \mathbf{c})}} \mathcal{X}(\mathbf{b}_j) \right\} \quad (3.24) \\ &= \sum_{s^*=1}^{\min(M,N)} \binom{M}{s^*} \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sigma(\mathbf{p}) \sum_{\mathbf{b}_j \in \mathcal{B}_{(s^*, \mathbf{p})}} E \{ \mathcal{X}(\mathbf{b}_j) \}, \end{aligned}$$

where the last equality is obtained due to the equivalence of the compositions that just differ in order.

So, finally, Eq. (3.23) leads to

$$\begin{aligned} \sum_{s^*=1}^{\min(M,N)} \binom{N}{s^*} \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sigma(\mathbf{p}) \sum_{j=1}^{\binom{M}{\mathbf{p}}} E \{ \mathcal{X}(\mathbf{b}_j) \} &= \\ \sum_{s^*=1}^{\min(M,N)} \binom{N}{s^*} \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sigma(\mathbf{p}) \sum_{j=1}^{\binom{M}{\mathbf{p}}} E \left\{ X_1^{E_{j,1}} X_2^{E_{j,2}} \dots X_{s^*}^{E_{j,s^*}} \right\} &= \\ \sum_{s^*=1}^{\min(M,N)} \binom{N}{s^*} \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sigma(\mathbf{p}) \sum_{j=1}^{\binom{M}{\mathbf{p}}} \prod_{k=1}^{s^*} \mu_{E_{j,k}} \quad (3.25) \end{aligned}$$

which completes the proof. □

For the sake of clarity, let us see the example of Table 3.1. For $s^* = 2$ it is clear that

$$\sum_{\mathbf{b} \in \mathcal{B}_{((i,j),(1,3))}} E \{ \mathcal{X}(\mathbf{b}) \} = \sum_{\mathbf{b} \in \mathcal{B}_{((i,j),(3,1))}} E \{ \mathcal{X}(\mathbf{b}) \}$$

and the same happens for $s^* = 3$

$$\sum_{\mathbf{b} \in \mathcal{B}_{((i,j,k),(2,1,1))}} E \{ \mathcal{X}(\mathbf{b}) \} = \sum_{\mathbf{b} \in \mathcal{B}_{((i,j,k),(1,2,1))}} E \{ \mathcal{X}(\mathbf{b}) \} = \sum_{\mathbf{b} \in \mathcal{B}_{((i,j,k),(1,1,2))}} E \{ \mathcal{X}(\mathbf{b}) \}.$$

One of the potentials of Theorem 2 is that it can be easily generalized to more complex products of sums of random variables, since the multiplicity of each sum is defined by \mathbf{e} . Then, the following corollary of Theorem 2 can be established.

Corollary 1. *Let X_i be IID random variables with $t_1 m_1 + t_T m_T$ finite raw moments. Then, the following equality holds*

$$E \left\{ \left(\sum_{i=1}^N X_i^{t_1} \right)^{m_1} \cdots \left(\sum_{i=1}^N X_i^{t_T} \right)^{m_T} \right\} = \sum_{s^*=1}^{\min(M,N)} \binom{N}{s^*} \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sigma(\mathbf{p}) \sum_{j=1}^{\binom{M}{\mathbf{p}}} \prod_{k=1}^{s^*} \mu_{E_{j,k}},$$

where $M = \sum_t m_t$, \mathcal{P}_{s^*} is the set of solutions of the integer partition problem for M into $s^* \in \{1, \dots, M\}$ terms. $\sigma(\mathbf{p})$ is a multinomial coefficient which indicates the number of permutations without repetition of $\mathbf{p} \in \mathcal{P}_{s^*}$, $E_{j,k} = \sum_{i=1}^M \mathbf{e}(i) \delta(\mathbf{b}_j(i) - k)$ with $\mathbf{e} = [\underbrace{t_1, \dots, t_1}_{m_1}, \dots, \underbrace{t_T, \dots, t_T}_{m_T}]$, δ is the Kronecker delta and $\mathbf{b}_j \in \mathcal{B}_{s^*, \mathbf{p}}$ is a permutation without repetition of s^* elements with multiplicities the elements of \mathbf{p} .

The proof can be derived in the same way as was done for Theorem 2.

3.3 An example: Variance of sample variance

To illustrate the results of the previous section with a practical example, we will calculate the variance of the unbiased sample variance

$$\text{Var} \{V\} = E \{V^2\} - E \{V\}^2, \quad (3.26)$$

with $V = \frac{1}{N(N-1)} \left(N \sum_{i=1}^N X_i^2 - \left(\sum_{i=1}^N X_i \right)^2 \right)$ and $N \geq 4$. In this case we need to calculate just two raw moments, i.e. $E \{V\}$ and $E \{V^2\}$

$$E \{V\} = \mu_2 - \mu_1^2$$

$$E \{V^2\} = \frac{1}{N^2(N-1)^2} \sum_{k=0}^2 \binom{2}{k} N^k (-1)^{2-k} E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^k \left(\sum_{i=1}^N X_i \right)^{2(2-k)} \right\} =$$

$$\frac{1}{N^2(N-1)^2} \left[E \left\{ \left(\sum_{i=1}^N X_i \right)^4 \right\} - 2NE \left\{ \left(\sum_{i=1}^N X_i^2 \right) \left(\sum_{i=1}^N X_i \right)^2 \right\} + N^2 E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^2 \right\} \right].$$

The former is straightforward, but note that for the latter moments of products of sums must be calculated. For $E \left\{ \left(\sum_{i=1}^N X_i \right)^4 \right\}$, applying Theorem 2, we have

s^*	\mathcal{P}_{s^*}	$ \mathcal{B}_{(s^*, \mathbf{p})}, \mathcal{B}_{(s^*, \mathbf{p})} = \binom{M}{\mathbf{p}}$	$\mathcal{X}(\mathcal{B}_{(s^*, \mathbf{p})})$	$ \binom{N}{s^*} \sum_{\mathbf{a} \in \mathcal{A}_{s^*}} E\{\mathcal{X}(\mathbf{a})\}$
$s^* = 1$	$\mathcal{P}_1 = \{(4)\}$ $\sigma(\mathcal{P}_1) = \{1\}$	$\mathcal{B}_{(1, (4))} =$ $\{(1, 1, 1, 1)\}$ $ \mathcal{B}_{(1, (4))} = 1$	$\mathcal{X}(\mathcal{B}_{(1, (4))}) = \{X_1^4\}$	$\binom{N}{1} \mu_4$
$s^* = 2$	$\mathcal{P}_2 = \{(3, 1)\}$ $\sigma(\mathcal{P}_2) = \{2, 1\}$	$\mathcal{B}_{(2, (3, 1))} =$ $\left\{ \begin{array}{l} (1, 1, 1, 2) \\ \vdots \\ (1, 1, 2, 2) \end{array} \right\}$ $ \mathcal{B}_{(2, (3, 1))} = 4$ $\mathcal{B}_{(2, (2, 2))} =$ $\left\{ \begin{array}{l} (1, 1, 2, 2) \\ \vdots \end{array} \right\}$ $ \mathcal{B}_{(2, (2, 2))} = 6$	$\mathcal{X}(\mathcal{B}_{(2, (3, 1))}) =$ $\left\{ \begin{array}{l} X_1^3 X_2 \\ \vdots \end{array} \right\}$ $\mathcal{X}(\mathcal{B}_{(2, (2, 2))}) =$ $\left\{ \begin{array}{l} X_1^2 X_2^2 \\ \vdots \end{array} \right\}$	$\binom{N}{2} (8\mu_1 \mu_3 + 6\mu_2^2)$
$s^* = 3$	$\mathcal{P}_3 = \{(2, 1, 1)\}$ $\sigma(\mathcal{P}_3) = \{3\}$	$\mathcal{B}_{(3, (2, 1, 1))} =$ $\left\{ \begin{array}{l} (1, 1, 2, 3) \\ \vdots \end{array} \right\}$ $ \mathcal{B}_{(3, (3, 1))} = 12$	$\mathcal{X}(\mathcal{B}_{(3, (3, 1))}) =$ $\left\{ \begin{array}{l} X_1^2 X_2 X_3 \\ \vdots \end{array} \right\}$	$\binom{N}{3} (36\mu_1^2 \mu_2)$
$s^* = 4$	$\mathcal{P}_4 = \{(1, 1, 1, 1)\}$ $\sigma(\mathcal{P}_4) = \{24\}$	$\mathcal{B}_{(4, (1, 1, 1, 1))} =$ $\left\{ \begin{array}{l} (1, 2, 3, 4) \\ \vdots \end{array} \right\}$ $ \mathcal{B}_{(4, (1, 1, 1, 1))} = 24$	$\mathcal{X}(\mathcal{B}_{(4, (1, 1, 1, 1))}) =$ $\left\{ \begin{array}{l} X_1 X_2 X_3 X_4 \\ \vdots \end{array} \right\}$	$\binom{N}{4} (24\mu_1^4)$

$$E \left\{ \left(\sum_{i=1}^N X_i \right)^4 \right\} = \binom{N}{1} \mu_4 + \binom{N}{2} (8\mu_1 \mu_3 + 6\mu_2^2) + \binom{N}{3} (36\mu_1^2 \mu_2) + \binom{N}{4} (24\mu_1^4).$$

Now for $E \left\{ \left(\sum_{i=1}^N X_i^2 \right) \left(\sum_{i=1}^N X_i \right)^2 \right\}$ we have

s^*	\mathcal{P}_{s^*}	$ \mathcal{B}_{(s^*, \mathbf{p})}, \mathcal{B}_{(s^*, \mathbf{p})} = \binom{M}{\mathbf{p}}$	$\mathcal{X}(\mathcal{B}_{(s^*, \mathbf{p})})$	$ \binom{N}{s^*} \sum_{\mathbf{a} \in \mathcal{A}_{s^*}} E\{\mathcal{X}(\mathbf{a})\}$
$s^* = 1$	$\mathcal{P}_1 = \{(3)\}$ $\sigma(\mathcal{P}_1) = \{1\}$	$\mathcal{B}_{(1, (3))} =$ $\{(1, 1, 1)\}$ $ \mathcal{B}_{(1, (3))} = 1$	$\mathcal{X}(\mathcal{B}_{(1, (3))}) =$ $\{X_1^2 X_1 X_1\}$	$\binom{N}{1} \mu_4$
$s^* = 2$	$\mathcal{P}_2 = \{(2, 1)\}$ $\sigma(\mathcal{P}_2) = \{2\}$	$\mathcal{B}_{(2, (2, 1))} =$ $\left\{ \begin{array}{l} (1, 1, 2) \\ (1, 2, 1) \\ (2, 1, 1) \end{array} \right\}$ $ \mathcal{B}_{(2, (3, 1))} = 3$	$\mathcal{X}(\mathcal{B}_{(2, (2, 1))}) =$ $\left\{ \begin{array}{l} X_1^2 X_1 X_2 \\ X_1^2 X_2 X_1 \\ X_2^2 X_1 X_1 \end{array} \right\}$	$\binom{N}{2} (4\mu_1 \mu_3 + 2\mu_2^2)$
$s^* = 3$	$\mathcal{P}_2 = \{(1, 1, 1)\}$ $\sigma(\mathcal{P}_2) = \{1\}$	$\mathcal{B}_{(2, (1, 1, 1))} =$ $\left\{ \begin{array}{l} (1, 2, 3) \\ (1, 3, 2) \\ \vdots \end{array} \right\}$ $ \mathcal{B}_{(2, (1, 1, 1))} = 6$	$\mathcal{X}(\mathcal{B}_{(2, (1, 1, 1))}) =$ $\left\{ \begin{array}{l} X_1^2 X_2 X_3 \\ X_1^2 X_3 X_2 \\ \vdots \end{array} \right\}$	$\binom{N}{3} (6\mu_1^2 \mu_2)$

$$E \left\{ \left(\sum_{i=1}^N X_i^2 \right) \left(\sum_{i=1}^N X_i \right)^2 \right\} = \binom{N}{1} \mu_4 + \binom{N}{2} (4\mu_1 \mu_3 + 2\mu_2^2) + \binom{N}{3} (6\mu_1^2 \mu_2)$$

and, finally, for $E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^2 \right\}$

s^*	\mathcal{P}_{s^*}	$ \mathcal{B}_{(s^*, \mathbf{p})}, \mathcal{B}_{(s^*, \mathbf{p})} = \binom{M}{\mathbf{p}}$	$\mathcal{X}(\mathcal{B}_{(s^*, \mathbf{p})})$	$ \binom{N}{s^*} \sum_{\mathbf{a} \in \mathcal{A}_{s^*}} E\{\mathcal{X}(\mathbf{a})\}$
$s^* = 1$	$\mathcal{P}_1 = \{(2)\}$ $\sigma(\mathcal{P}_1) = \{1\}$	$\mathcal{B}_{(1, (2))} =$ $\{(1, 1)\}$ $ \mathcal{B}_{(1, (2))} = 1$	$\mathcal{X}(\mathcal{B}_{(1, (2))}) = \{X_1^2 X_1^2\}$	$\binom{N}{1} \mu_4$
$s^* = 2$	$\mathcal{P}_2 = \{(1, 1)\}$ $\sigma(\mathcal{P}_2) = \{1\}$	$\mathcal{B}_{(2, (1, 1))} =$ $\left\{ \begin{array}{l} (1, 2) \\ (2, 1) \end{array} \right\}$ $ \mathcal{B}_{(2, (1, 1))} = 2$	$\mathcal{X}(\mathcal{B}_{(2, (1, 1))}) =$ $\left\{ \begin{array}{l} X_1^2 X_2^2 \\ X_2^2 X_1^2 \end{array} \right\}$	$\binom{N}{2} (2\mu_2^2)$

$$E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^2 \right\} = \binom{N}{1} \mu_4 + \binom{N}{2} (2\mu_2^2).$$

Then, the variance of the sample variance can be written as

$$\text{Var}\{V\} = \frac{1}{N}\mu_4 - \frac{4}{N}\mu_1\mu_3 - \frac{(N-3)}{N(N-1)}\mu_2^2 + \frac{4(2N-3)}{N(N-1)}\mu_1^2\mu_2 - \frac{2(2N-3)}{N(N-1)}\mu_1^4.$$

This result, first derived by Gauss, has been derived recently by quite heuristic algebraic methods [Cho04, Cho05] in terms of zero-mean random variables which was also obtained in [Tuk56]. Quoting [Spe02], deriving this result “is already messy enough to warrant thinking very carefully about the algebraic formulation one adopts, and any desire to obtain more general expressions of the same kind focuses the mind greatly on the same issue”. The contribution of Theorem 2 is that it offers a simple methodology for calculating the sample variance moments which is easy to implement even for higher orders avoiding, as much as possible, the combinatorial explosion. Additionally, Theorem 2 corollary offers the desired general expressions.

3.4 Numerical Implementation

There are two important problems that must be solved in order to calculate arbitrary moments of the sample variance. The first one is the partition problem. This is a very common problem in combinatorics, algebra and many algorithms. For computational purposes one is interested in generating the set of all the partitions of an integer or those that satisfy some conditions. In our case, we are interested in generating all the set of partitions, \mathcal{P}_{s^*} . Many algorithms to generate the set of solutions have appeared in the literature [And76, McK65b, McK65a, Opd10, Sto62, Zog98], so we can use any of them.

On the other hand, the problem of generating the set $\mathcal{B}_{s^*,p}$ is a permutation problem with possible repeated terms. There are also many algorithms to solve this problem [Aho74, Knu97, Rol88, Sed84] so we can also choose any of them.

The direct implementation of $E\left\{\left(\sum_{i=1}^N X_i^2\right)^m \left(\sum_{i=1}^N X_i\right)^n\right\}$ is presented in Algorithm 1. Note that following the results presented in Theorems 1 and 2 this implementation is straightforward. The heaviest computational part of the algorithm is the construction of $\mathcal{B}_{s^*,p}$, since it needs to build all unique combinations of s^* indexes with a particular arrangement of multiplicities p . As an example, for the 5th raw moment of the sample variance, the biggest set $\mathcal{B}_{s^*,p}$ will have 3,628,800 elements and it will cost a considerably amount of time and memory.

In order to reduce the number of elements to calculate, one can realize that the set of all combinations of $\mathcal{B}_{s^*,p}$ can be decomposed into two different groups of m and n elements. The first one refers to those indexes of random variables of power 2: X_i^2 . The second group to those of power 1: X_i . This way, it is clear that every permutation inside these two groups is equivalent in terms of the expectation so, one just need to know the cardinal of all equivalent permutations, which is easily calculated by means of a

Algorithm 1 Algorithm to calculate $E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\}$.

$m \leftarrow$ Power of $(\sum_{i=1}^N X_i^2)^m$
 $n \leftarrow$ Power of $(\sum_{i=1}^N X_i)^n$
 $M \leftarrow m + n$
 $\mathbf{v} \leftarrow$ Vector of M raw moments of X_i
 $\mathbf{e} \leftarrow [2, \dots, 2, 1, \dots, 1]$
 $\mathcal{P}_{s^*=1, \dots, M} \leftarrow$ Sets of all the partitions of M
 $value \leftarrow 0$
for $s^* = 1$ to $\min(M, N)$ **do**
 for $\mathbf{p} \in \mathcal{P}_{s^*}$ **do**
 $\mathcal{B}_{s^*, \mathbf{p}} \leftarrow$ Sets of permutations without repetition
 $E_{j,k} \leftarrow \sum_{i=1}^M \mathbf{e}(i) \delta(\mathbf{b}_j(i) - k)$
 $\sigma(\mathbf{p}) \leftarrow$ Number of unique permutations of \mathbf{p}
 $value \leftarrow value + \binom{N}{s^*} \sigma(\mathbf{p}) \sum_{j=1}^{\binom{M}{\mathbf{p}}} \prod_{k=1}^{s^*} \mathbf{v}(E(j, k))$
 end for
end for
return $value$

multinomial coefficient with multiplicities those of indexes in each group. So, the following corollary can be established.

Corollary 2. Let X_i be IID random variables with $2m + n$ finite raw moments μ_1, \dots, μ_{2m+n} . Then, the following equality holds

$$E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\} = \sum_{s^*=1}^{\min(M, N)} \binom{N}{s^*} \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sigma(\mathbf{p}) \sum_{j=1}^{|\mathcal{D}_{(s^*, \mathbf{p}, m)}|} \prod_{k=1}^{s^*} \mu_{E_{j,k}},$$

where $M = m + n$, \mathcal{P}_{s^*} is the set of solutions of the integer partition problem for M into s^* terms. $\sigma(\mathbf{p})$ is a multinomial coefficient which indicates the number of permutations without repetition of $\mathbf{p} \in \mathcal{P}_{s^*}$. $\mathcal{D}_{(s^*, \mathbf{p}, m)}$ is the set of unique combinations of M elements (s^* indexes with multiplicities those of \mathbf{p}) taken m at time with multiplicities $\mathbf{p} = (p_1, p_2, \dots, p_{s^*})$. $\hat{E}_{j,k}$ is defined as

$$\hat{E}_{j,k} = 2 \times \pi(\mathbf{d}_j)(k) + \pi^*(\mathbf{d}_j)(k),$$

where $\pi(\mathbf{b}_j) = (m_1, \dots, m_{s^*})$ are the multiplicities of each index of $\mathbf{b}_j \in \mathcal{D}_{(s^*, \mathbf{p}, m)}$ and $\pi^*(\mathbf{b}_j) = (p_1 - m_1, \dots, p_{s^*} - m_{s^*})$.

Proof. For each s^* and each \mathbf{p} let $\mathcal{D}_{(s^*, \mathbf{p}, m)}$ be the set of unique combinations of M elements (s^* indexes with multiplicities those of \mathbf{p}) taken m at time with multiplicities $\mathbf{p} = (p_1, p_2, \dots, p_{s^*})$. This set obviously has a maximum cardinal $|\mathcal{D}_{(s^*, \mathbf{p}, m)}| = \binom{M}{m}$ when $N \geq M$, $s^* = M$ and $\mathbf{p} = (1, 1, \dots, 1)$.

From each element $\mathbf{d}_j \in \mathcal{D}_{(s^*, \mathbf{p}, m)}$ we can easily calculate the multiplicities of each index $\pi(\mathbf{d}) = (m_1, m_2, \dots, m_{s^*})$

$$\pi : \mathcal{D}_{(s^*, \mathbf{p}, m)} \longrightarrow \mathbb{N}^{s^*} \\ \mathbf{d} \quad (m_1, \dots, m_{s^*}), \quad (3.27)$$

where

$$m_i = \sum_{k=1}^{s^*} \delta(\mathbf{d}(k) - i) \text{ with } i = 1, \dots, s^*. \quad (3.28)$$

We avoid intentionally the dependence of π with respect to s^* and \mathbf{p} in order to make notation simpler.

Now, let $\mathbf{d}_j^* = \mathcal{D}_{(s^*, \mathbf{p}, m)}^C$ be the complement of each element $\mathbf{d}_j \in \mathcal{D}_{(s^*, \mathbf{p}, m)}$, i.e., the indexes and the multiplicities of the indexes that are not in $\mathbf{d}_j \in \mathcal{D}_{(s^*, \mathbf{p}, m)}$. The multiplicities of \mathbf{d}_j^* can be easily calculated by $\pi(\mathbf{d}_j^*) = \pi^*(\mathbf{d}_j) = (p_1 - m_1, p_2 - m_2, \dots, p_{s^*} - m_{s^*})$

$$\pi^* : \mathcal{D}_{(s^*, \mathbf{p}, m)} \longrightarrow \mathbb{N}^{s^*} \\ \mathbf{d} \quad (p_1 - m_1, \dots, p_{s^*} - m_{s^*}). \quad (3.29)$$

From these sets it is easy to calculate the number of equivalent arrangements since all unique permutations of each group belong to the same sort of random variable (X_i^2 or X_i) and the order of indexes does not matter.

There are $\binom{m}{\pi(\mathbf{d}_j)} \binom{n}{\pi(\mathbf{d}_j^*)}$ equivalent elements for each \mathbf{p} . And now we can avoid using e since the order of the moments per index becomes

$$\hat{E}_{j,k} = 2 \times \pi(\mathbf{d}_j)(k) + \pi^*(\mathbf{d}_j)(k). \quad (3.30)$$

So, finally, for each s^* and \mathbf{p} we have

$$\sum_{j=1}^{\binom{M}{\mathbf{p}}} \prod_{k=1}^{s^*} \mu_{E_{j,k}} = \sum_{j=1}^{|\mathcal{D}_{(s^*, \mathbf{p}, m)}|} \binom{m}{\pi(\mathbf{d}_j)} \binom{n}{\pi^*(\mathbf{d}_j)} \prod_{k=1}^{s^*} \mu_{\hat{E}_{j,k}}. \quad (3.31)$$

□

The method to implement this reduction of combinations is presented in Algorithm 2.

3.5 Discussion

In this paper the problem of the calculation of the moments of the sample variance has been studied. Two theorems are proposed that allow calculating in a direct way the moments of the products of sums of different powers of random variables. The main contribution of these theorems is that there is no need of conversion formulae of multiplication tables as the Polykay

Algorithm 2 Alternative method for $E \left\{ \left(\sum_{i=1}^N X_i^2 \right)^m \left(\sum_{i=1}^N X_i \right)^n \right\}$.

$m \leftarrow$ Power of $(\sum_{i=1}^N X_i^2)^m$
 $n \leftarrow$ Power of $(\sum_{i=1}^N X_i)^n$
 $M \leftarrow m + n$
 $\mathbf{v} \leftarrow$ Vector of M raw moments of X_i
 $\mathcal{P}_{s^*=1, \dots, M} \leftarrow$ Sets of all the partitions of M
 $value \leftarrow 0$
for $s^* = 1$ to $\min(M, N)$ **do**
 for $\mathbf{p} \in \mathcal{P}_{s^*}$ **do**
 $\mathcal{D}_{(s^*, \mathbf{p}, m)} \leftarrow$ set of unique combinations of M elements taken m at time
 (s^* indexes with multiplicities those of \mathbf{p})
 $\sigma(\mathbf{p}) \leftarrow$ Number of unique permutations of \mathbf{p}

 $value \leftarrow value + \binom{N}{s^*} \sigma(\mathbf{p}) \sum_{j=1}^{|\mathcal{D}_{(s^*, \mathbf{p}, m)}|} \binom{m}{\pi(\mathbf{d}_j)} \binom{n}{\pi^*(\mathbf{d}_j)} \prod_{k=1}^{s^*} \mathbf{v}(\hat{E}(j, k))$

 end for
end for
return $value$

j	Algorithm 1. $\max \mathcal{B}_{s^*, \mathbf{p}} $	Algorithm 2. $\max \mathcal{D}_{(s^*, \mathbf{p}, m)} $
1	2	1
2	24	3
3	720	6
4	40320	15
5	3628800	35
6	479001600	84
7	87178291200	210
8	20922789888000	495
9	6402373705728000	1287
10	2432902008176640000	3003
11	1124000727777607680000	8008
12	620448401733239439360000	19448
13	403291461126605635584000000	50388
14	304888344611713860501504000000	125970
15	265252859812191058636308480000000	319770

Table 3.3: Maximum number of combinations for each algorithm, to calculate the j -th raw moment of the sample variance, $(E\{V^j\})$, up to order 15.

philosophy. Additionally, if one is interested in conversion formulae or multiplication tables, they can be derived with the methods here proposed in a direct way.

From a practical point of view, two algorithms for numerical implementation are proposed. The first one is a direct implementation of the main result of this work. This algorithm can be easily generalized for more complex products. The second one is a refined method for the case of two different powers of random variables.

The complexity of both algorithms is related to the number of elements in each of the sets $\mathcal{B}_{s^*,\mathbf{p}}$ and $\mathcal{D}_{(s^*,\mathbf{p},m)}$. These sets can be obtained from algorithms with constant delay (see [Zog98]), this means that the time for obtaining each element of the sets from the previous one can be assumed to be constant and included in the time of the operations inside each loop.

So, considering $N \geq M$, the instructions inside the loops consume a time $T_{(1,s^*)}$ for each element of the sets $\mathcal{B}_{s^*,\mathbf{p}}$ and $\mathcal{D}_{(s^*,\mathbf{p},m)}$. This time can be considered fix since the operations to be done are products in the interval $(1, s^*)$ which, in the worst case, is $T_{(1,M)}$. So, the time consumed in the loops, T_1 and T_2 , for the first and the second algorithm respectively are

$$T_1 = \sum_{s^*}^M \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sum_{j=1}^{|\mathcal{B}_{s^*,\mathbf{p}}|} T_{(1,s^*)} \leq T_{(1,M)} \sum_{s^*}^M \max_p |\mathcal{B}_{s^*,\mathbf{p}}| |\mathcal{P}_{s^*}| \quad (3.32)$$

$$T_2 = \sum_{s^*}^M \sum_{\mathbf{p} \in \mathcal{P}_{s^*}} \sum_{j=1}^{|\mathcal{D}_{(s^*,\mathbf{p},m)}|} T_{(1,s^*)} \leq T_{(1,M)} \sum_{s^*}^M \max_p |\mathcal{D}_{(s^*,\mathbf{p},m)}| |\mathcal{P}_{s^*}|. \quad (3.33)$$

Table 3.3 shows the maximum number of combinations for each algorithm when $N \geq M$, to calculate the j -th raw moment of the sample variance, $(E\{V^j\})$, up to order 15. The higher performance of this implementation becomes evident.

Generalized Gamma Mixture Model for Ultrasonic Tissue Characterization

*Si le poète doit choisir dans les choses (et il doit),
ce n'est pas le beau, mais le caractéristique.
Victor Hugo, 1802–1885.*

Abstract– Several statistical models have been proposed in the literature to describe the behavior of speckle. Among them, the Nakagami distribution has proven to very accurately characterize the speckle behavior in tissues. However, it fails when describing the heavier tails caused by the impulsive response of speckle. The Generalized Gamma (GG) distribution (which also generalizes the Nakagami distribution) was proposed to overcome these limitations. Despite the advantages of the distribution on terms of goodness of fitting, its main drawback is the lack of a closed-form maximum likelihood (ML) estimates. Thus, the calculation of its parameters becomes difficult and not attractive. In this work we propose: (1) a simple but robust methodology to estimate the ML parameters of GG distributions; (2) a Generalized Gamma Mixture Model (GGMM). These mixture models are of great value in ultrasound imaging when the received signal is characterized by different nature of tissues. We show a better speckle characterization is achieved when using GG and GGMM rather than other state-of-the-art distributions and mixture models. Results showed the better performance of the GG distribution in characterizing the speckle of blood and myocardial tissue in ultrasonic images.

Adapted from: G. Vegas-Sánchez-Ferrero¹, S. Aja-Fernández¹, C. Palencia² and M. Martín-Fernández¹, *Generalized Gamma Mixture Model for Ultrasonic Tissue Characterization*, Computational and Mathematical Methods in Medicine, No. 481923, 2012.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Departamento de Matemática Aplicada, Univ. Valladolid.

4.1 Introduction

Among the non-invasive imaging modalities, probably, the most widespread are the ultrasound imaging. The main reason of its success is that it provides a low-cost way to help diagnosing and can be used for many medical applications. However, ultrasonic (US) images are characterized by the presence of a peculiar granular pattern: the so-called *speckle*.

This term was adopted from the field of laser optics [Goo75] in the early sixties due to the similarity of the patterns between laser optics and ultrasonics. Although the nature of the speckle in US images stems from a different phenomena, there still share some similarities. Both patterns come from the random interference of many coherent wave components reflected from different microscopic elements. In the case of US; the volume, the number of effective scatterers, and the acquisition process contribute to the formation of speckle [Elt06].

The analysis of backscattered echo from tissues needs a proper description of the ultrasonic signals. For this purpose, and due to the random nature of the speckle, several statistical models have been proposed in the literature. This characterization can be used either for segmentation [Des09], classification [Sea11] purposes or for filtering the speckle itself [Yu02, AF06, Kri07, VSF10a]. The latter usually considers the speckle as an undesired consequence, since it degrades resolution and adds spatial noise to the image. Thus, filtering is commonly applied as a preprocessing step for further segmentation of regions of interest or to extract relevant measures for physiological analysis.

The statistical description of US signals provide an important information of the backscattered echo from tissues. The parameters of the statistical models allow identifying the features of tissues and provides important descriptors for classification. Some of the filtering algorithms relay on a Bayesian approach where an accurate statistical model becomes necessary. As a consequence, modeling the amplitude statistics of US signals has been a very active area.

Several statistical models have been proposed in the last decades. Probably the most well-known is the Rayleigh model, which is a one-parameter distribution which describes the so-called fully formed (or developed) Speckle. This probabilistic distribution describes the behavior of speckle when a high number of effective scatterers is present in the resolution cell. However, real images show a deviation from this model, this non-Rayleigh behavior can be due to a small number of scatterers in the resolution cell or when there are some dominant components in the cell. The most commonly accepted distributions that try to model non-Rayleigh distributions are the Rice (fully resolved speckle), K (partially formed speckle) and Homodyned K (partially resolved speckle).

Although those models are based in physical assumptions of the back-scattering process, some other distributions have proven to provide a good

performance on real images. This is the case of Gamma [VSF10b, Tao06, Nil08] and Nakagami [Sha00] distributions. The first is proposed as a two-parameter distribution that describes the result of interpolated/filtered fully formed speckle [VSF10b] and also has shown good results in empirical tests among other distributions [Tao06, Nil08]. The Nakagami distribution proposed by Shankar for the case US characterization [Sha00] is also a two-parameter distribution which generalizes the Rayleigh distribution. This distribution was adopted from the models proposed to describe the statistics of the returned echo radar.

The capability of the Nakagami distribution to model the backscattering from tissues for fully resolved and fully formed speckle made it become the most commonly accepted model for tissue characterization. However, the tails of the probabilistic density functions of Nakagami, K, Rayleigh or Gamma do not show the impulsive response of speckle which originate heavier tails. In order to describe this impulsive response, a generalized Nakagami distribution was proposed by Shankar in [Sha01]. This is a three-parameter model which has shown a better behavior than the Nakagami or Rayleigh, an expected result since it is a generalization of the other models. However, the generalized Nakagami distribution does not have closed-form Maximum Likelihood estimates (MLE) and, thus, it makes their use difficult. Note that, though Shankar in [Sha01] said that the MLE can be obtained, the equations used were based in the results from Stacy and Mihram [Sta65], which were calculated by the methods of moments and they also expressed the difficulties of obtaining a MLE: "Closed expressions for solutions to the maximum likelihood equations are highly unlikely". It is important to note that the results of [Sta65] were obtained for the estimation of the Generalized Gamma (GG) distribution which is essentially the same as the Generalized Nakagami Distribution but with another parametrization.

The different nature of tissues is reflected in a different response of the speckle. Hence, a mixture model has shown to be a natural strategy for statistically describing the features of tissues. This approach has been previously used for segmentation purposes in the case of Nakagami Mixture Models (NMM) by Destremes in [Des09], for classification with Rayleigh Mixture Models by Seabra *et al.* in [Sea11], and for filtering with a mixture of Gamma and Gaussian Mixture Models in [VSF10b, VSF10a]. All these approaches make use of the Expectation-Maximization (EM) [Moo96] algorithm to calculate the parameters that better fit the empirical probability distribution function (PDF). This method is particularly useful when the MLE exists since it maximizes the expected value of the log-likelihood function with respect to the condition of the belonging to each tissue class for a given data.

The EM algorithm cannot be easily applied for the calculation of a Generalized Gamma Mixture Model (GGMM) without a MLE. However, some interesting results have been recently published on the calculation of the MLE of the Generalized Gamma which permit and efficient computation of the GGMM.

The aim of this work is to revitalize the use of the Generalized Gamma distribution (also called, Generalized Nakagami Distribution) for tissue characterization. For this purpose, we present two main contributions: first, we propose a simple methodology to calculate the ML estimate which offers robust results comparing to the methods in the literature [Sta65, Gom08, Nou12]. Second, two different methods were proposed for the calculation of the GGMM parameters. Both were developed by applying the EM method in the derivation of the proposed ML method. Results when comparing both methods to the GMM and NMM in real images showed the better fitting of the GGMM. The GGMM provides a posterior probability of belonging to each tissue class which can be of help for further filtering, segmenting or classifying methods.

The rest of the paper is structured as follows: In section 4.2.1 we introduce the distributions most commonly used for characterizing speckle of ultrasonic images. There, the GG distribution is motivated as a suitable generalization of the Gamma and Nakagami distributions which fail in describing the impulsive response of speckle. Then, in section 4.2.2 we analyze the methods proposed in the literature for estimating the parameters of the GG distribution and a simple but robust method is proposed (section 4.2.2). One of the advantages of this method is that it can be easily extended to estimate the parameters of a GGMM by means of the EM algorithm. Section 4.2.3 is devoted to the extension of the ML method to obtain the parameters of the GGMM where two algorithms are proposed. The performance of the ML estimate derived in section 4.2.2 is compared to other state-of-the-art methods in section 4.3.1 for synthetic data and for real cases in section 4.3.2. The performance of the GGMM is analyzed in section 4.3.3, where the GGMM is compared to NMM and GMM. Finally, we propose some applications for the GGMM in section 4.3.4. In section 4.4 we conclude.

4.2 Materials and Methods

4.2.1 Statistical models for describing the nature of speckle

The formation of US images begins with the emission of a pulse packet which travels through the tissue. The backscattering produced by the scatterers in the resolution cell contribute to the change of the pulse shape according to the characteristics of the media, i.e. the number of scatterers as well as their size [VSF10b, Sea11, Sha00].

The contribution of the backscattered echo, $s(t)$ can be treated as a random walk due to the random location of the scatterers in the resolution cell [Sha00]

$$s(t) = \sum_{n=1}^N \alpha_n \cos(\omega_0 t + \phi_n), \quad (4.1)$$

where ω_0 is the mean frequency of excitation and N is the number of effective scatterers in the resolution cell. The phases, ϕ_n , are usually modeled as uniformly distributed in $[0, 2\pi]$ and the amplitude is usually considered to be Normal distributed.

The fully formed speckle model assumes a high number of scatterers, so the *Central Limit Theorem* applies and the backscattered echo can be expressed as

$$s(t) = X \cos(\omega_0 t) + Y \sin(\omega_0 t), \quad (4.2)$$

where X and Y are zero mean identically distributed Gaussian distributions.

Then, the envelop of the backscattered signal echo, $R = \sqrt{X^2 + Y^2}$ is Rayleigh distributed [Goo75, Wag83]

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} u(r), \quad (4.3)$$

where $u(\cdot)$ is the Heaviside step function defined as

$$u(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0. \end{cases} \quad (4.4)$$

Under the assumption of a high number of effective scatterers but with the presence of resolvable structures in the resolution cell (specular component, C), X and Y become non-zero Gaussian distributions. The envelop does no longer follow a Rayleigh distribution but a Rician one [Wag83]

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2+C^2}{2\sigma^2}} I_0\left(\frac{rC}{\sigma^2}\right) u(r), \quad (4.5)$$

where $I_0(\cdot)$ where is the modified Bessel function of first kind.

When the number of scatterers decreases and the Central Limit Theorem cannot be applied, more complicated distributions are proposed to model the distribution of the envelope. Concretely, the K distribution models the case when the number of scatterers is a random variable itself, which is modeled as a Poisson whose local mean is Gamma distributed, this is equivalent to consider σ as gamma distributed [Elt06]

$$f_R(r|\sigma) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} u(r) \quad (4.6)$$

and

$$f_\sigma(\sigma) = \frac{1}{2b^2} \frac{1}{\Gamma(\nu+1)} \left(\frac{\sigma}{2b^2}\right) e^{-\frac{\sigma}{2b^2}} u(\sigma), \quad (4.7)$$

so, the PDF of R is

$$f_R(r) = \int f_R(r|\sigma) f_\sigma(\sigma) d\sigma = \frac{2}{b\Gamma(\nu+1)} \left(\frac{\sigma}{2b^2}\right)^{\nu+1} K_\nu\left(\frac{r}{b}\right) u(r), \quad (4.8)$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind.

A generalization of the previous models appears when a specular component is considered and the number of scatterers, N follows a negative binomial distribution. This is the case of the homodyned-K distribution [Dut94]

$$f_R(r) = r \left(\int \frac{x}{1 + x^2 \sigma^2 / 2\nu} J_0(xC) J_0(xr) dx \right) u(r). \quad (4.9)$$

This PDF has no closed expression and this limits its use.

On a completely different approach, Shankar in [Sha00] proposed a Nakagami distribution as a “simpler universal model for tissue characterization”. Unlike the previously reviewed models, the Nakagami is not based on physical arguments or on a bottom-up modeling of the scattering process. However, it has empirically shown a better performance than the Rayleigh and Rice distributions.

The Nakagami PDF is as follows

$$f_R(r) = \frac{2m^m r^{2m-1}}{\Gamma(m)(2\Omega)^m} e^{-\frac{m}{2\Omega} r^2} u(r). \quad (4.10)$$

This distribution offers good properties to describe the backscattered echo: the Rayleigh distribution is a particular case of the Nakagami ($m = 1$) and, additionally, when $m > 1$ is similar to the Rice distribution. However, this distribution has some limitations. The Nakagami model can not fit the heavier tails of the empirical PDFs due to the impulsive nature of scatterers [Sha01].

In order to describe the impulsive response of scatterers, Shankar proposed in [Sha01] a generalized Nakagami distribution which is essentially the same as a Generalized Gamma distribution [Sta65]. However, this distribution presents some difficulties in the estimation of its parameters, since there are no closed equations for the ML estimates.

In the next section we describe some methods that have been used in the literature with special attention to methods that provide a ML estimate of the GG parameters. Additionally, we propose a simple method to calculate the ML estimates of the parameters. The results obtained in the derivation of this ML method provide the foundations for the development of the Generalized Gamma Mixture Model, which is the main contribution of this work.

4.2.2 Estimation of parameters of the Generalized Gamma

Moments method

This method was proposed by Stacy in [Sta65]. For the derivation of the method the following parametrization was adopted

$$f(x|a, \nu, p) = p \frac{x^{p\nu-1}}{a^{p\nu} \Gamma(\nu)} e^{-\left(\frac{x}{a}\right)^p} u(x), \quad (4.11)$$

where the parameters (a, ν, p) are all positive. This is the definition of the GG distribution hereafter. For a given $p > 0$, all moments $E\{X^r\}$ exist. Now, let Z be the random variable (RV) defined as

$$Z = \log(X/a)^p = p(\log(X) - \log(a)). \quad (4.12)$$

For this RV, the central moments, $\mu_r(\cdot)$, of r-th order are

$$\mu_r(Z) = p^r \mu_r(\log X), \quad (4.13)$$

Additionally, it is easy to show that, given a RV, X , which follows a GG distribution ($X \sim GG(a, \nu, p)$), the following properties hold

$$\begin{aligned} kX &\sim GG(ka, \nu, p), \quad k > 0 \\ X^m &\sim GG(a^m, \nu, p/m), \quad m \neq 0. \end{aligned} \quad (4.14)$$

So, a new RV Z can be defined as $Z = \log(X/a)^p$ where $(X/a)^p$ follows a Gamma distribution of parameter ν . Hence, the log-transformed distribution of the Gamma RV is the following

$$f_Z(z) = f_{GG}(e^z|1, \nu, 1)e^z = \frac{1}{\Gamma(\nu)} \exp(\nu z - \exp(z)), \quad (4.15)$$

where $z \in \mathbb{R}$.

The moment generating function of Z is easily calculated as $E\{e^{tZ}\} = \frac{t+\nu}{\nu}$. Where $E\{\cdot\}$ is the expectation operator. So, the r-th moment of Z is the following

$$E\{Z^r\} = \frac{\Gamma^{(r)}(\nu)}{\Gamma(\nu)} = \Psi^{(r)}(\nu), \quad (4.16)$$

where $\Psi^{(r)}$ is the *polygamma* function defined as

$$\Psi^{(r)}(x) = \frac{d^{m+1}}{dx} \log \Gamma(x). \quad (4.17)$$

Finally, the three first central moments are defined as

$$\begin{aligned} pE\{\log X - \log a\} &= \Psi^{(0)} \\ p^2 \mu_2(\log X) &= \Psi^{(1)} \\ p^3 \mu_3(\log X) &= \Psi^{(2)}. \end{aligned} \quad (4.18)$$

These equations can be used to estimate the parameters of the $GG(a, \nu, p)$; \hat{a} , $\hat{\nu}$, \hat{p} , by approximating the moments by means of the sample moments

$$\begin{aligned} \hat{a} &= \exp(\bar{y} - \Psi^{(0)}(\hat{\nu})) \\ \hat{p} &= -\text{sign}(g_y) \frac{\sqrt{\Psi^{(1)}(\hat{\nu})}}{S_y} \\ -|g_y| &= \frac{\Psi^{(2)}(\hat{\nu})}{(\Psi^{(1)}(\hat{\nu}))^{3/2}}, \end{aligned} \quad (4.19)$$

where $\bar{y} = \frac{1}{N} \sum_{i=1}^N \log x_i$, with $\{x_i\}_{i=1}^N$ the set of samples of X ; S_y^2 is the sample variance of $\{y_i\}_{i=1}^N = \{\log x_i\}_{i=1}^N$, and g_y its sample skewness.

The estimates are derived by means of calculating the value $\hat{\nu}$ from the last equation of eq. (4.19). So, a numerical calculation needs to be performed. In the original article [Sta65], Stacy and Mihran provided a graph representing $\frac{\Psi^{(2)}(\nu)}{(\Psi^{(1)}(\nu))^{3/2}}$ for a range $\nu \in [0.1, 5]$.

This method, though provides a quite straight-forward calculation of the parameters, can provide estimates which are outside the parameter space. Yet, it is highly sensitive to the number of samples.

Heuristic approaches

In order to avoid the problems associated to the moments method, some heuristic methods have been proposed in the literature. As examples, Gomes *et al.* [Gom08] proposed an iterative method which evaluates the best performance of the χ^2 goodness-of-fit test for a fixed p (see the parametrization of Eq. (4.11)). The parameters of the transformed samples $Y = X^p$, which are Gamma distributed, were calculated by the moments method. At the end of the loop, the set of parameters with least p-value is chosen.

This method, present some shortcomings. First, the parameters of the Gamma distributed data were calculated by the moments method, so the problems associated to the moments method are not circumvent. However, even if a good estimate is calculated, the χ^2 goodness-of-fit test depends on the calculation of the estimated PDF which strongly depends on the number of bins considered and the assumption of a sample with sufficient large size.

Other heuristic method is the one presented by Wingo in [Win87]. This method, based on the one proposed by Hager and Bain [Hag70], tries to solve the maximum likelihood equations for the GG distribution. The log-likelihood, \mathcal{LL} , of a RV $X \sim GG(a, \nu, p)$ for the parametrization presented in Eq. (4.11) is

$$\begin{aligned} \mathcal{LL}(a, \nu, p | \mathbf{x}) &= \log \left(\left(\frac{p}{a^{p\nu} \Gamma(\nu)} \right)^n \prod_{i=1}^n x_i^{p\nu-1} e^{-\sum_{i=1}^n (x_i/a)^p} \right) = \\ &= n \log(p) - np\nu \log(a) - n \log(\Gamma(\nu)) + \\ &\quad (p\nu - 1) \sum_{i=1}^n \log x_i - \sum_{i=1}^n \left(\frac{x_i}{a} \right)^p, \quad (4.20) \end{aligned}$$

where $\mathbf{x} = \{x_i\}_{i=1}^n$ is the set of samples.

Now, calculating the derivatives with respect the parameters and setting

it equal to zero, one can obtain the ML equations

$$\begin{aligned}
 a^p &= \frac{1}{n\nu} \sum_{i=1}^n x_i^p \\
 p \sum_{i=1}^n \log\left(\frac{x_i}{a}\right) - n\Psi(\nu) &= 0 \\
 \frac{n}{p} + \nu \sum_{i=1}^n \log\left(\frac{x_i}{a}\right) - \sum_{i=1}^n \left(\frac{x_i}{a}\right)^p \log\left(\frac{x_i}{a}\right) &= 0,
 \end{aligned} \tag{4.21}$$

where $\Psi(x) \equiv \Psi^{(0)}(x) = \frac{\Gamma'(x)}{\Gamma(x)}$.

This system of equations can be reduced to a single non-linear equation with p as the single unknown

$$-\Psi(\nu) + \frac{p}{n} \sum_{i=1}^n \log(x_i) - \log\left(\sum_{i=1}^n x_i^p\right) + \log(n\nu) = 0, \tag{4.22}$$

where

$$\nu = -\left(\frac{p}{n} \sum_{i=1}^n \log(x_i) - p \frac{\sum_{i=1}^n x_i^p \log(x_i)}{\sum_{i=1}^n x_i^p}\right)^{-1} \tag{4.23}$$

and

$$a = \left(\frac{1}{n\nu} \sum_{i=1}^n x_i^p\right)^{1/p}. \tag{4.24}$$

So, the problem is reduced to calculate p from Eq. (4.22). Some authors reported the difficulty of solving this equation with the conventional numerical methods such as Newton-Raphson [Hag70] and conclude that the MLE may not exist.

In [Win87], the author faced the problem by analyzing the effect of inappropriate zero finding algorithms. So, an heuristic method for isolating roots of a general scalar nonlinear equation was proposed. This method makes use of the root-isolation technique proposed in [Jon78], which uses only function values to isolate the roots in a compact interval of the real line.

Though this method can provide a ML estimate of the parameter by solving Eq. (4.22), it has to heuristically divide the intervals where p is searched and calculate whether a root is in it or not by means of the mean value and variance of the function in each of the intervals, so many evaluation of the function are required.

ML approach

A very interesting analysis was recently published by Noufaily and Jones in [Nou12], where an iterative approach is proposed to solve the likelihood equations, Eqs. (4.21), in a way that the individual equations are uniquely

solvable. This result provides a very promising technique for calculating the MLE parameters of the GG.

In that work, the log-likelihood equations were calculated following the reparameterization proposed in [Law80]. Concretely, for a RV X which is distributed, $X \sim GG(a, \nu, p)$, the new RV $Y = \log X$ is calculated, whose PDF is the following

$$f_Y(y) = \frac{p}{\Gamma(\nu)} \frac{e^{yp\nu}}{a^{p\nu}} \exp\left(-\frac{e^{yp}}{a^p}\right) = \frac{\nu^{\nu-1/2}}{\sigma\Gamma(\nu)} \exp(\sqrt{\nu}w - \nu e^{w/\sqrt{\nu}}), \quad (4.25)$$

where $y \in \mathbb{R}$, $w = \frac{y-\mu}{\sigma}$, $\sigma = \frac{1}{p\sqrt{\nu}}$ and $\mu = \log(a) + \frac{1}{p} \log(\nu)$.

So, in the end, the following equations have to be solved

$$\mu = \sigma\sqrt{\nu} \log S_0 \quad (4.26)$$

$$R(\sigma) \equiv \frac{S_0}{S_1} - \bar{Y} - \frac{\sigma}{\sqrt{\nu}} = 0 \quad (4.27)$$

$$T(\nu) \equiv \log(\nu) - \Psi(\nu) - \frac{L}{\sqrt{\nu}} = 0, \quad (4.28)$$

where $L = \frac{\mu - \bar{Y}}{\sigma}$ and $S_j = \frac{1}{n} \sum_{i=1}^n y_i^j \exp\left(\frac{y_i}{\sigma\sqrt{\nu}}\right)$.

The important result of [Nou12] is the demonstration that both Eqs. (4.27) and (4.28) are well-behaved with unique solutions in σ and ν respectively. So, an iterative method can be developed to calculate $\hat{\nu}$ by Eq. (4.28) from an initial guess of the parameters and then $\hat{\sigma}$ by solving Eq. (4.27). Finally, $\hat{\mu}$ is calculated by replacing the previous estimates in Eq. (4.26). These estimates can be used to calculate a new L to compute the new log-likelihood function. By repeating these steps until a desired accuracy, the estimates are achieved [Law80].

This method provides a fundamental result about the behavior of the log-likelihood equations, and guarantees their solution. However, the method does not provide any proof concerning its convergence or the uniqueness of the ML. Yet, this method needs to solve two non-linear equations by numerical techniques whereas the method proposed by Wingo in [Win87], previously described, only needs to solve a linear equation.

The proposed approach

We propose a simple but efficient method to calculate the ML estimates of the GG distribution. The main advantage of the method is that it can be easily implemented and has the same properties of the method of [Nou12], i.e. the equation to solve are well-behaved with unique solution. Additionally, the method just need the calculation of just one non-linear equation and, thus, the computing time is considerably reduced.

The method consists in transforming the RV, $X \sim GG(a, \nu, p)$ by the following transformation $Y = X^{p_0}$ where p_0 is a positive real number. So, the

new PDF of Y is the following

$$f_Y(y) = \frac{p}{p_0} \frac{y^{\frac{p}{p_0}\nu-1}}{a^{p\nu}\Gamma(\nu)} \exp\left(-\frac{y^{\frac{p}{p_0}}}{a^p}\right) u(y). \quad (4.29)$$

Note that this PDF follows a Gamma PDF when $p_0 = p$. Hence, a reasonable way to find the p value is to find the value of p_0 that maximizes the Likelihood of the GG distribution and also maximizes the Gamma distributed RV $Y = X^{p_0}$.

In order to see if this method provides a proper solution, we first demonstrate that the ML estimate of the parameters of the new random variable Y also maximizes the Likelihood of the GG distribution when $p_0 = p$.

First, we calculate the ML estimates of the parameters of Eq. (4.29) for $p_0 = p$, whose log-likelihood is the following

$$\mathcal{L}\mathcal{L}_Y = -np\nu \log(a) - n \log(\Gamma(\nu)) + (\nu - 1) \sum_{i=1}^n \log(y_i) - \sum_{i=1}^n \frac{y_i}{a^p}. \quad (4.30)$$

The maximum with respect to the parameter a is easily calculated by taking the derivative with respect to a and setting it equal to zero

$$a_0^p = \frac{1}{n\nu} \sum_{i=1}^n y_i. \quad (4.31)$$

Finally, the EQ. (4.30) can be maximized with respect ν by introducing the value of a_0

$$\log(\nu) - \Psi(\nu) = \log\left(\frac{1}{n} \sum_{i=1}^n y_i\right) - \frac{1}{n} \sum_{i=1}^n \log(y_i). \quad (4.32)$$

Now, by introducing a_0 in the log-likelihood function of the GG distribution, Eq. (4.20)

$$\mathcal{L}\mathcal{L}_X = n \log(p) - nv \log\left(\frac{1}{nv} \sum_{i=1}^n y_i\right) - n \log(\Gamma(\nu)) + (\nu - \frac{1}{p}) \sum_{i=1}^n \log(y_i) - n\nu. \quad (4.33)$$

Now, by maximizing with respect ν , we obtain the following equation

$$\frac{\partial \mathcal{L}\mathcal{L}_X}{\partial \nu} \equiv -n \log\left(\frac{1}{n} \sum_{i=1}^n y_i\right) + n \log(\nu) + \sum_{i=1}^n \log(y_i) - n\Psi(\nu) = 0 \quad (4.34)$$

and finally, reordering terms, we obtain the same equation for which ν_0 is also a solution.

This result guarantees that there exist always a solution for the ML estimate of the GG distribution $(\hat{a}, \hat{\nu}, \hat{p})$ and the parameters \hat{a} and $\hat{\nu}$ are those obtained for the ML estimate for the transformed RV $Y = X^{\hat{p}}$. Hence, there is always a solution for the MLE for a GG.

Additionally, since the MLE of a Gamma distribution always exist for whatever positive y_i values (the Eq. (4.32) is well-behaved), the problem is

reduced to finding the value p that maximizes \mathcal{LL}_X among the ones that maximize \mathcal{LL}_Y .

The search method for p was implemented by the Nelder-Mead method [Nel65] while the Brent's algorithm was applied for calculating ν [Bre73].

This method does not demonstrate the uniqueness of p as did not any of the methods in the literature. However, in our experience, we agree with Noufaily and Jones [Nou12] that the global maximum of the \mathcal{LL}_X appears to be distant to any other local maximum.

The main advantage of the method here proposed is that it is easy to implement and only one non-linear equation has to be solved, whereas the method of [Nou12] needs to solve two non-linear equations in each iteration and [Win87] method needs several calculations of non-linear equations for each interval considered for the isolation root technique.

4.2.3 Generalized Gamma Mixture Model

An additional advantage of the proposed method for the calculation of the MLE parameters for the GG distribution is that it can be easily adapted for the calculation of the parameters of GG Mixture Models (GGMM).

There were some attempts in the literature to obtain the parameters of a GGMM. Concretely, in [Wah11] they calculated the GGMM by means of the Nelder-Mead and Gradient descent methods for maximizing the log-likelihood. However, that method is strongly sensitive to the number of mixtures since it is just a direct optimization of the log-likelihood score equations of the mixture model.

In this section we derive the GGMM by applying the Expectation-Maximization methodology [Moo96] and combining them with the method used to calculate the MLE of the GG distribution.

Let $X = \{x_i\}, 1 \leq i \leq N$ be a set of samples. These samples are considered to be independent and identically distributed (IID) RVs. Now, the GGMM considers that these samples result from the contributions of J distributions

$$p(x_i|\Theta) = \sum_{j=1}^J \pi_j f_X(x_i|\Phi_j), \quad (4.35)$$

where Φ is a vector of the parameters of the GGMM $(\pi_1, \dots, \pi_J, \Theta_1, \dots, \Theta_J)$ and Θ_j are the parameters of the PDF (in our case the parameters of the GG, represented as a_i, ν_j, p_j).

The joint distribution of IID samples is given by

$$p(X|\Theta) = \prod_{i=1}^N p(x_i|\Theta). \quad (4.36)$$

The EM is applied here to maximize the log-likelihood function when some hidden discrete random variables $Z = \{Z_i\}$, are introduced into the model. These RVs take values in $\{1, \dots, J\}$ and indicate the class for which each sample x_i belongs.

Now, defining $\Theta^{(n)}$ as an estimate of the parameters of the mixture in the n -th iteration, the expectation step is performed by calculating the expected value of the log-likelihood $\mathcal{LL}(\Theta|X, Z)$

$$\mathcal{Q}(\Theta|\Theta^{(n)}) = E_{Z|\Theta^{(n)}}\{\mathcal{LL}(\Theta|X, Z)\}. \quad (4.37)$$

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

The application of the EM algorithm for estimating the parameters of mixture models has been applied for several distributions, see for example [Moo96, Fig02]. However, to the best of our knowledge, this is the first time a mixture model is presented for GG distributions.

In order to derive the estimates of the parameters in each iteration, we first define the joint distribution of IID samples X and the hidden random variables, Z as

$$p(X, Z|\Theta) = \prod_{i=1}^N p(x_i, z_i|\Theta), \quad (4.38)$$

where $p(x_i, z_i|\Theta) = p(x_i|z_i, \Theta)p(z_i|\Theta)$.

Now, the log-likelihood function can be defined in the following way

$$\begin{aligned} \mathcal{LL}(\Theta|X, Z) &= \log(p(X, Z|\Theta)) = \sum_{i=1}^N \log p(x_i, z_i|\Theta) \\ &= \sum_{i=1}^N \log p(x_i|z_i, \Theta) + \sum_{i=1}^N \log p(z_i|\Theta) \\ &= \sum_{i=1}^N \log f_X(x_i|z_i, \Theta) + \sum_{i=1}^N \log \pi_{z_i}. \end{aligned} \quad (4.39)$$

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

$$\begin{aligned} \mathcal{Q}(\Theta|\Theta^{(n)}) &= E_{Z|\Theta^{(n)}}\{\mathcal{LL}(\Theta|X, Z)\} \\ &= \sum_{i=1}^N E_{Z|\Theta^{(n)}, x_i}\{\log f_X(x_i|\Theta) + \log p(z_i|\Theta)\} \\ &= \sum_{i=1}^N \sum_{j=1}^J p(Z_i = j|x_i, \Theta^{(n)}) (\log f_X(x_i|\Theta_{z_j}) + \log \pi_j), \end{aligned} \quad (4.40)$$

where $\pi_j = p(Z_i = j|\Theta)$ is the probability of x_i to belong to the class j .

The probability $p(Z_i = j|x_i, \Theta^{(n)})$ can be calculated by applying the Bayes theorem as

$$p(Z_i = j|x_i, \Theta^{(n)}) = \frac{f_X(x_i|\Theta^{(n)})p(Z_i = j|\Theta^{(n)})}{p(x_i|\Theta^{(n)})}. \quad (4.41)$$

Note that the Eq. (4.39) is composed of two terms, so the maximization step can be done to each term independently. For the term depending on the π_j some constraints have to be considered since they must hold $\sum_{j=1}^J \pi_j = 1$. An optimization via Lagrange multipliers can be done in a straightforward way and they guarantee a necessary condition for optimality in this problem. The new Lagrange function with λ as the Lagrange multiplier is the following

$$\Lambda(\pi, \lambda) = \sum_{i=1}^N \sum_{j=1}^J \gamma_{i,j} \log \pi_j + \lambda \left(\sum_{j=1}^J \pi_j - 1 \right), \quad (4.42)$$

where we introduced $\gamma_{i,j} = p(Z_i = j|x_i, \Theta^{(n)})$ to simplify notation.

Now, calculating the derivative with respect each π_j and setting it equal to 0, the following expression is derived

$$\sum_{i=1}^N \gamma_{i,j} = -\lambda \hat{\pi}_j. \quad (4.43)$$

By summing both terms of the equation over j , we obtain

$$\lambda = -\sum_{i=1}^N \sum_{j=1}^J \gamma_{i,j} = -N$$

and the estimates for the parameters π_j that maximize the Lagrange function (and the likelihood function) are

$$\hat{\pi}_j = \frac{1}{N} \sum_{i=1}^N \gamma_{i,j} = \frac{1}{N} \sum_{i=1}^N p(Z_i = j|x_i, \Theta^{(n)}). \quad (4.44)$$

For the calculation of the maximum of Eq. (4.39) with respect $\Theta_j = (a_j, \nu_j, p_j)$, we first calculate the derivative with respect a_j

$$\frac{\partial}{\partial a_j} \left\{ \sum_{i=1}^N \sum_{j=1}^J \gamma_{i,j} \log f_X(x_i|\Theta_j) \right\} = 0, \quad (4.45)$$

where the log-likelihood of $p(x_i|\Theta_j)$ is the one described in Eq. (4.20) for one sample x_i

$$\begin{aligned} \log f_X(x_i|\Theta_j) &= \log p - p\nu \log(a) - \log \Gamma(\nu) \\ &\quad + (p\nu - 1) \log x_i - \left(\frac{x_i}{a} \right)^p. \end{aligned} \quad (4.46)$$

The result is

$$\hat{a}^{p_j} = \frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_j}}{\nu_j \sum_{i=1}^N \gamma_{i,j}}. \quad (4.47)$$

Now, plugging Eq. (4.47) into Eq. (4.39) and deriving with respect to ν_j and setting it equal to 0

$$\frac{\partial}{\partial \nu_j} \left\{ \sum_{i=1}^N \sum_{j=1}^J \gamma_{i,j} \log f_X \left(x_i \mid \frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_j}}{\nu_j \sum_{i=1}^N \gamma_{i,j}}, \nu_j, p_j \right) \right\} = 0. \quad (4.48)$$

It results in the following equality

$$\log(\nu_j) - \Psi(\nu_j) = \log \left(\frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_j}}{\sum_{i=1}^N \gamma_{i,j}} \right) + \frac{\sum_{i=1}^N \gamma_{i,j} \log(x_i^{p_j})}{\sum_{i=1}^N \gamma_{i,j}}. \quad (4.49)$$

Note that Eq. (4.49) is essentially the same as Eq. (4.32), which is well-behaved and always has a unique solution. Thus, this non-linear equation can be solved by numerical methods in the same way as was performed in the MLE of the GG parameters. In our case, we also used the Brent's algorithm [Bre73].

The interval where the Brent's algorithm is performed can be derived by means of the following property

$$\frac{1}{2\nu_j} < \log(\nu_j) - \Psi(\nu_j) < \frac{1}{\nu_j}. \quad (4.50)$$

So, the desired value of $\hat{\nu}_j$ in the interval

$$\frac{1}{2A} < \hat{\nu}_j < \frac{1}{A}, \quad (4.51)$$

where

$$A = \log \left(\frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_j}}{\sum_{i=1}^N \gamma_{i,j}} \right) + \frac{\sum_{i=1}^N \gamma_{i,j} \log(x_i^{p_j})}{\sum_{i=1}^N \gamma_{i,j}}. \quad (4.52)$$

This property can be found in [Alz97] and was also used in [Nou12] for the calculation of the ML estimates of the GG.

Now, the problem can be stated in the same way as was done for the ML estimate proposed in section 4.2.2. We are interested in the parameter p_j which maximizes the likelihood for the component $j \in [1, J]$. So, for each p_j , the Eqs. (4.47) and (4.49) provides the estimate of a_j and ν_j respectively. By applying the Nelder-Mead algorithm to maximize the log-likelihood function for each component j , as was done for the ML estimates in section 4.2.2, one can obtain the desired ML estimates. We will refer to this method as the $GGMM_1$ method.

It is important to note that the parameter estimates can be also solved by extending the ML method of [Nou12]. For this purpose, the parametrization proposed by Lawless [Law80] can be applied to the mixture model as was explained in section 4.2.2.

The log-likelihood equations to be solved are completely equivalent to Eqs. (4.47) and (4.49) due to the invariance of the ML estimates to the transformation $Y = \log(X)$. However, Lawless' parametrization allows us to extend the results of [Nou12] to the case of GGMM. For the sake of clarity, we rewrite the parametrization

$$\begin{aligned}\sigma_j &= \frac{1}{p_j \sqrt{\nu_j}} \\ \mu_j &= \log(a_j) + \frac{1}{p_j} \log(\nu_j) \\ k_j &= \nu_j.\end{aligned}\tag{4.53}$$

With this parametrization, Eq.(4.47) becomes

$$\hat{\mu}_j = \sqrt{k_j} \sigma_j \log(\tilde{S}_0),\tag{4.54}$$

where

$$\tilde{S}_r = \frac{\sum_{i=1}^N \gamma_{i,j} y_i^r e^{\left(\frac{y_i}{\sigma_j \sqrt{k_j}}\right)}}{\sum_{i=1}^N \gamma_{i,j}}.\tag{4.55}$$

So, in the case of the parameter σ_j which maximizes the log-likelihood of $Y = \log(X)$

$$\frac{\partial}{\partial \sigma_j} \left\{ \sum_{i=1}^N \sum_{j=1}^J \gamma_{i,j} \log f_Y(y_i | \hat{\mu}_j, \sigma_j, k_j) \right\} = 0.\tag{4.56}$$

It results in

$$\frac{\tilde{S}_1}{\tilde{S}_0} - \frac{\sigma_j}{\sqrt{k_j}} - \frac{\sum_{i=1}^N \gamma_{i,j} y_i}{\sum_{i=1}^N \gamma_{i,j}} = 0.\tag{4.57}$$

This equations is well-behaved and all the theoretical demonstrations obtained in [Nou12] still hold: it is monotone decreasing and, when $\lim \sigma_j \rightarrow 0$, the function takes the value

$$y_{\max} - \frac{\sum_{i=1}^N \gamma_{i,j} y_i}{\sum_{i=1}^N \gamma_{i,j}} > 0.\tag{4.58}$$

As a conclusion, Eq. (4.57) has always a positive solution for any μ_j and k_j . Additionally, due to the invariance of the ML estimates for the transformation $Y = \log(X)$, there always exist a p_j for any a_j and ν_j .

The solution is in the interval

$$0 < \hat{\sigma}_j < \sqrt{k_j} \left(y_{\max} - \frac{\sum_{i=1}^N \gamma_{i,j} y_i}{\sum_{i=1}^N \gamma_{i,j}} \right),\tag{4.59}$$

so, the value can be calculated by any numerical method. We used here also the Brent's algorithm.

Finally, from an initial guess of p_j one can calculate $k_j \equiv \nu_j$ from Eq. (4.49) and then use it to calculate the estimate of σ_j from Eq. (4.57), in an iterative way until a desired tolerance is reached.

This methodology generalizes the proposed method of [Nou12] for the case of GGMM and we will refer to it as the $GGMM_2$ method.

4.2.4 Implementation Generalized Gamma Mixture Model

In this section we detail the implementation of both of the proposed methods for the GGMM.

In the algorithm 3 the Nelder-Mead method [Nel65] was used for the calculation of $p_j^{(n)}$ and the Brent's algorithm [Bre73] for $\nu_j^{(n)}$ in the interval given in Eq. (4.51).

In the case of Algorithm 4, the Brent's algorithm [Bre73] was used for calculating σ and k in the intervals of Eqs. (4.59) and (4.51) respectively.

The computational complexity of the previous GGMM methods when compared to the calculation of a simple GG depends on the number of components, J , assumed by the model. In each iteration of the EM algorithm, the expected parameters of each component have to be calculated. So, if the time consumed to estimate a GG is T , the calculation of the expected parameters of the mixture is $J \times T$.

When other mixture models such as RMM, NMM and GMM are consid-

Algorithm 3 Implementation of the $GGMM_1$ method.

```

 $\{x\}_{i=1}^N \leftarrow$  Samples
 $J \leftarrow$  Number of components
 $\Theta^{(0)} \equiv \{\pi_j^{(0)}, a_j^{(0)}, \nu_j^{(0)}, p_j^{(0)}\}_{j=1}^J \leftarrow$  Initial guess of parameters for each component
 $\gamma_{i,j}^{(0)} \leftarrow p(Z_i = j | x_i, \Theta^{(0)})$ 
 $maxIter \leftarrow$  Maximum number of iterations
 $Tol \leftarrow$  Tolerance
 $err \leftarrow \infty$ 
 $n \leftarrow 0$  Iterations counter
while  $err > Tol$  &  $n < maxIter$  do
   $n \leftarrow n + 1$ 
  for  $j = 1 \rightarrow J$  do
     $p_j^{(n)} \leftarrow \arg \max_{p_j} \left\{ \sum_{i=1}^N \gamma_{i,j}^{(n-1)} \log f_X \left( x_i \mid \frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_j}}{\nu_j(p_j) \sum_{i=1}^N \gamma_{i,j}} \right) \right\}$ 
     $\nu_j^{(n)} \leftarrow \log(\nu_j) - \Psi(\nu_j) = \log \left( \frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_j^{(n)}}}{\sum_{i=1}^N \gamma_{i,j}} \right) + \frac{\sum_{i=1}^N \gamma_{i,j} \log(x_i^{p_j^{(n)}})}{\sum_{i=1}^N \gamma_{i,j}}$ 
     $a_j^{(n)} \leftarrow \left( \frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_j^{(n)}}}{\nu_j^{(n)} \sum_{i=1}^N \gamma_{i,j}} \right)^{1/p_j^{(n)}}$ 
     $\pi_j^{(n)} \leftarrow \frac{1}{N} \sum_{i=1}^N \gamma_{i,j}^{(n-1)}$ 
     $\gamma_{i,j}^{(n)} \leftarrow p(Z_i = j | x_i, \Theta^{(n)})$ 
  end for
   $err \leftarrow \|\Theta^{(n)} - \Theta^{(n-1)}\| / \|\Theta^{(n-1)}\|$  Evaluate the relative tolerance
end while
return  $\Theta^{(n)}$ 

```

Algorithm 4 Implementation of the $GGMM_2$ method.

```

 $\{x\}_{i=1}^N \leftarrow$  Samples
 $\{y\}_{i=1}^N \leftarrow \{\log(x_i)\}_{i=1}^N$  Transformed samples
 $J \leftarrow$  Number of components
 $\Theta^{(0)} \equiv \{\pi_j^{(0)}, a_j^{(0)}, \nu_j^{(0)}, p_j^{(0)}\}_{j=1}^J \leftarrow$  Initial guess of parameters for each component
 $\gamma_{i,j}^{(0)} \leftarrow p(Z_i = j|x_i, \Theta^{(0)})$ 
 $maxIter \leftarrow$  Maximum number of iterations
 $Tol \leftarrow$  Tolerance
 $err \leftarrow \infty$ 
 $n \leftarrow 0$  Iterations counter
 $maxIterML \leftarrow$  Maximum number of iterations for the ML algorithm
 $TolML \leftarrow$  Tolerance for the ML
while  $err > Tol$  &  $n < maxIter$  do
   $n \leftarrow n + 1$ 
  for  $j = 1 \rightarrow J$  do
     $(a_{aux}, \nu_{aux}, p_{aux}) \leftarrow (a_j^{(n-1)}, \nu_j^{(n-1)}, p_j^{(n-1)})$ 
     $errML \leftarrow \infty$ 
     $m \leftarrow 0$  Iterations counter for ML
     $\mathcal{L}\mathcal{L}_j^{(0)} \leftarrow$  Calculate the log-likelihood for the  $j$ -th component
    while  $errML > TolML$  &  $m < maxIterML$  do
       $m \leftarrow m + 1$ 
       $\sigma = 1/(p_{aux}\sqrt{\nu_{aux}})$ 
       $k \leftarrow \log(\nu_j) - \Psi(\nu_j) = \log\left(\frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_{aux}}}{\sum_{i=1}^N \gamma_{i,j}}\right) + \frac{\sum_{i=1}^N \gamma_{i,j} \log(x_i^{p_{aux}})}{\sum_{i=1}^N \gamma_{i,j}}$ 
       $\tilde{S}_0 \leftarrow \frac{\sum_{i=1}^N \gamma_{i,j} \exp\left(\frac{y_i}{\sigma\sqrt{k}}\right)}{\sum_{i=1}^N \gamma_{i,j}}$ 
       $\tilde{S}_1 \leftarrow \frac{\sum_{i=1}^N \gamma_{i,j} y_i \exp\left(\frac{y_i}{\sigma\sqrt{k}}\right)}{\sum_{i=1}^N \gamma_{i,j}}$ 
       $\sigma \leftarrow \frac{\tilde{S}_1}{\tilde{S}_0} - \frac{\sigma}{\sqrt{k}} - \frac{\sum_{i=1}^N \gamma_{i,j} y_i}{\sum_{i=1}^N \gamma_{i,j}} = 0$ 
       $\nu_{aux} \leftarrow k$ 
       $p_{aux} \leftarrow 1/(\sigma\sqrt{k})$ 
       $a_{aux} \leftarrow \left(\frac{\sum_{i=1}^N \gamma_{i,j} x_i^{p_{aux}}}{\nu_{aux} \sum_{i=1}^N \gamma_{i,j}}\right)^{1/p_{aux}}$ 
       $\mathcal{L}\mathcal{L}_j^{(m)} \leftarrow$  Calculate the log-likelihood for the  $j$ -th component
       $errML \leftarrow \|\mathcal{L}\mathcal{L}_j^{(m)} - \mathcal{L}\mathcal{L}_j^{(m-1)}\|/\|\mathcal{L}\mathcal{L}_j^{(m-1)}\|$ 
    end while
     $(a_j^{(n)}, \nu_j^{(n)}, p_j^{(n)}) \leftarrow (a_{aux}, \nu_{aux}, p_{aux})$ 
     $\pi_j^{(n)} \leftarrow \frac{1}{N} \sum_{i=1}^N \gamma_{i,j}^{(n-1)}$ 
     $\gamma_{i,j}^{(n)} \leftarrow p(Z_i = j|x_i, \Theta^{(n)})$ 
  end for
   $err \leftarrow \|\Theta^{(n)} - \Theta^{(n-1)}\|/\|\Theta^{(n-1)}\|$  Evaluate the relative tolerance
end while

```

ered, the complexity of the EM method is similar to the GGMM. Note that both the GMM and the NMM need to solve a non-linear equation similar to Eq. (4.49) so the consumed time of the solution is the same. The additional cost of calculating the GGMM parameters is due to the calculation of the estimate of the parameter $p_j^{(n)}$. If we define the time to solve Eqs. (4.49) and (4.47) as T_1 , and T_2 as the time consumed for solving $p_j^{(n)}$, the computational time for a simple GG (T_{GG}) and a GGMM of J components (T_{GGMM}) would

be

$$T_{GG} = T_1 + T_2$$

$$T_{GGMM} = J \cdot (T_1 + T_2).$$

The estimated times in a Matlab (R2011a) implementation running in an ASUS G53SW laptop (Intel Core i7 2630QM Processor, 2.2 GHz, 8 GB RAM) were: $T_1 = 1.637 \text{ ms}$ and $T_2 = 0.2056 \text{ s}$.

4.3 Results and Discussion

4.3.1 Performance of the ML method

In this section we show the performance of the proposed methods for calculating the parameters of a GG distribution. For this purpose, we performed 200 synthetic experiments and tested the methods presented in section 4.2.2. Concretely, we tested the method of Stacy and Mihram [Sta65], Gomes *et al.* [Gom08], Noufaily and Jones [Nou12], and our proposed method of section 4.2.2. We will refer to them as *Stacy*, *Gomes*, *Noufaily* and *proposed* methods respectively.

The synthetic data was calculated in the same way as was done in [Nou12]: a set of gamma-distributed random samples are generated by means of the method proposed in [Mar00] and the GG-distributed data are obtained by taking the $1/p$ -th power of the samples. The parameters of the GG distribution were also calculated from sets of parameters in a reasonable dynamic range. The scale parameter a was set to 1 in all the experiments since this parameter just affects to the scale of the data. Both, the p parameter and the ν parameter were obtained from random samples of a uniform RV in the interval $[0.3, 5]$.

We choose this interval since lower values than 0.3 make the distribution to take values that tend to infinity as ν get closer to 0. This is an unrealistic situation when real images are considered. Additionally, when p takes lower values, the tail becomes heavier and the shape of the distribution also becomes unrealistic. These effects are shown in Fig. 4.1, where some examples of the PDFs of the GG distribution are depicted.

The number of iterations for the proposed method and for the method of [Nou12] was set to 100 and the tolerance function to 10^{-8} . The number of bins where the χ^2 -test was performed in the method of [Gom08] was 150 and the number of samples per experiment was 10^4 . The comparisons of the methods were performed by comparing the goodness-to-fit of each distribution by means of two different measures: Kullback-Leibler divergence (KL) and Kolmogorov-Smirnov (KS) statistic. The former is a non-symmetric

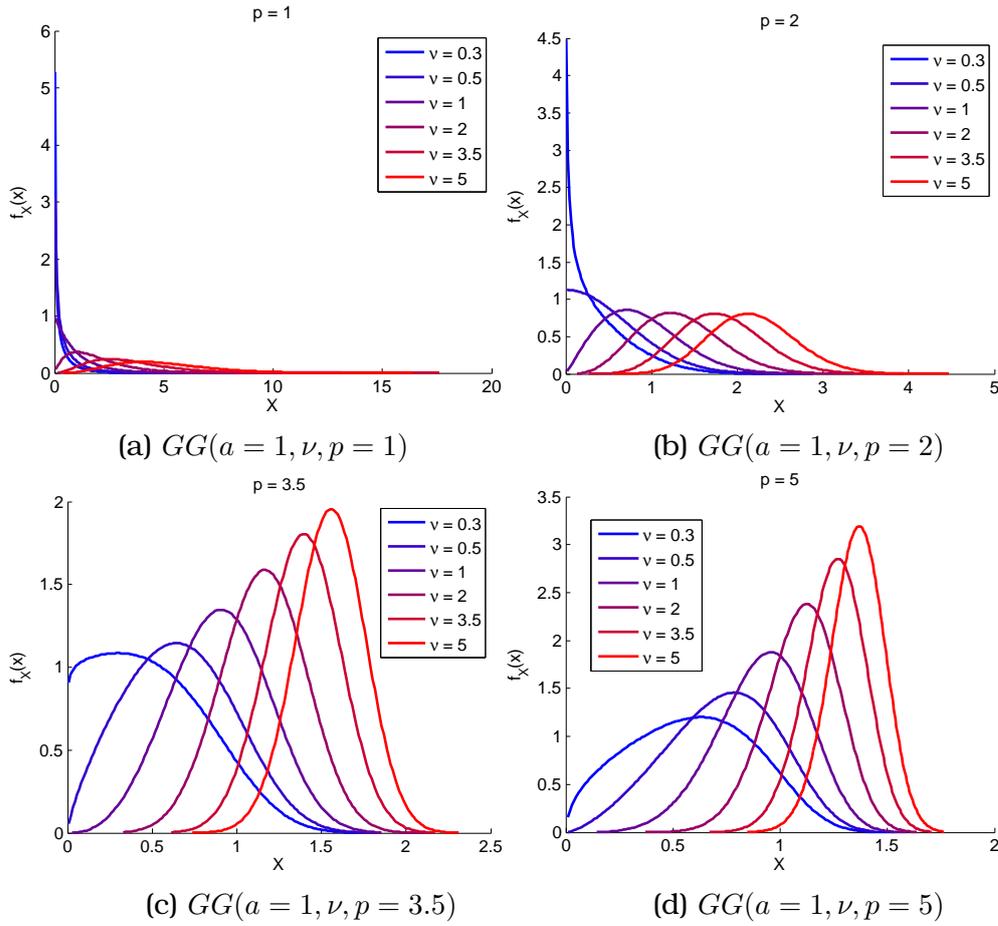


Figure 4.1: Some examples of the GG distribution for the parameters of the synthetic dataset.

measure of the difference between two probability distributions defined as

$$D_{KL}(p_n, f_X) = \sum_{i=1}^N p_n(i) \log \frac{p_n(i)}{f_X(i)}, \quad (4.60)$$

where p_n is the empirical PDF estimate and f_X is the theoretical distribution (the GG distribution). For the empirical estimate of the PDF, the number of bins of the histogram was set to 150.

The Kolmogorov-Smirnov statistic is the uniform norm of the cumulative distribution function (CDF), defined as

$$D_{KS} = \sup |\hat{F}(i) - F_X(i)|, \quad (4.61)$$

where \hat{F} is the empirical CDF of data and F_X the theoretical CDF. The KS measure was chosen since it does not depend on the PDF estimate and can be calculated with a few number of samples. Additionally, the Glivenko-Cantelli theorem states that, if the samples are drawn from distribution F_X , then D_{KS} converges to 0 almost surely [Dud99].

In Fig. 4.2 the results for both measures are depicted. It is clear that the moments method of Stacy gives poorer results than the other methods

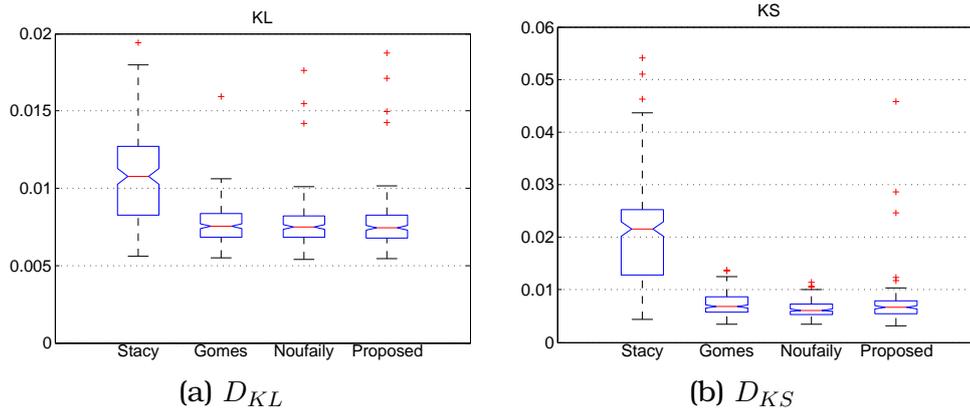


Figure 4.2: Results for D_{KL} and D_{KS} for 10^4 samples. Methods: Stacy [Sta65], Gomes *et al.* [Gom08], Noufaily [Nou12] and the proposed one of section 4.2.2.

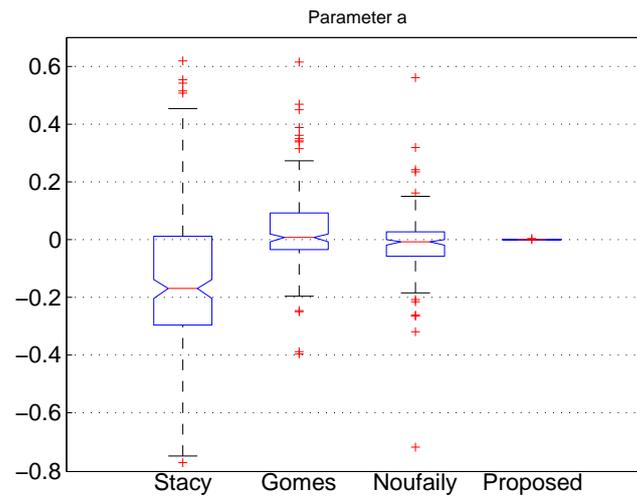
for both measures. This result was expected since the moments method depends on moments of third order, so the variance of the estimates becomes higher. The rest of the methods performed well for both measures. In the case of the D_{KL} they fit practically the same while, in the case of D_{KS} , there are some better results for the method of Noufaily and the proposed one. This is the effect of the approximation of the PDF for the χ^2 test performed by the method of Gomes: it calculates the best set of parameters for an approximation of the empirical PDF which depends on the number of bins and the number of samples of the dataset. So, as the number of samples is reduced or the number of bins is reduced, the estimate becomes worse.

In order to see the effect of this we also show in Fig. 4.3 the relative error¹ of the estimates for all the methods. In the figure, the whiskers show the dynamic range of the data which is not consider an outlier. So, though the method of Gomes provides good fitting, the variance of the estimates is higher than the method of Noufaily and the proposed one. At first sight, the results of Fig. 4.3 demonstrate the better performance of the proposed method in terms of variance of the ML estimates with no appreciable bias in the estimates.

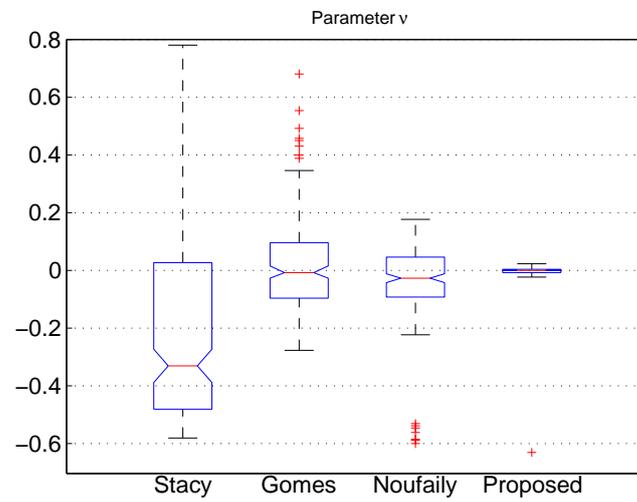
An example of the fitting performance of the methods is shown in Fig. 4.4 where the PDFs obtained with the methods are depicted as well as the absolute error and the relative error of the PDFs.

Following, we analyze the dependence of the estimates with the number of samples. The same experiment is repeated considering 500 samples. The results of both goodness-to-fit measures are shown in Fig. 4.5, and the relative error of the estimates are depicted in Fig. 4.6. The performance for the D_{KL} measure is similar for all the methods. However, note that the value is considerably higher than the obtained for the case of 10^4 samples, this effect is caused by the difficulties of estimating the PDF with so few samples. Since the Gomes algorithm is based in the χ^2 test, it is expected that its performance decreases and the variance of the parameter estimates

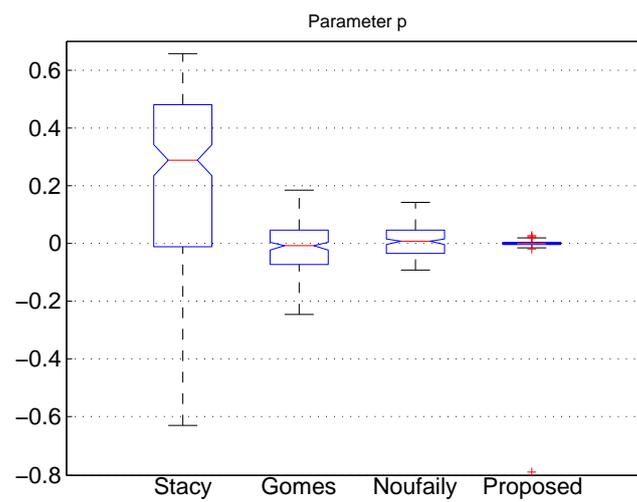
¹The relative error of an estimate $\hat{\theta}$ is calculated as $\epsilon_{rel} = \frac{\|\theta - \hat{\theta}\|}{\theta}$, while the absolute error is $\epsilon_{abs} = \|\theta - \hat{\theta}\|$.



(a)



(b)



(c)

Figure 4.3: Results for the relative error of the estimates for 10^4 samples. Methods: Stacy [Sta65], Gomes *et al.* [Gom08], Noufaily [Nou12] and the proposed one of section 4.2.2.

increases. In the case of the D_{KS} measure, the performance of all methods is comparatively equal to the case of 10^4 samples but a higher variability is observed in the Gomes method due to the sensitivity to the number of samples.

The better performance of Noufaily and the proposed methods is seen in Fig. 4.6 where the variability of the Noufaily method did not increase dramatically as the Gomes method did. The proposed method also presented a very low variance of the parameter estimates with no appreciable bias. In the light of these results we can conclude that the proposed method is robust with respect the number of samples and it does not introduce any appreciable bias in the parameter estimates. The goodness-of-fit performance of both the Noufaily and the proposed method are similar, though the estimates are more accurate with the proposed method. This can be due to the better convergence of the Nelder-Mead method than the algorithm of the Noufaily method.

4.3.2 Tissue characterization in real US images

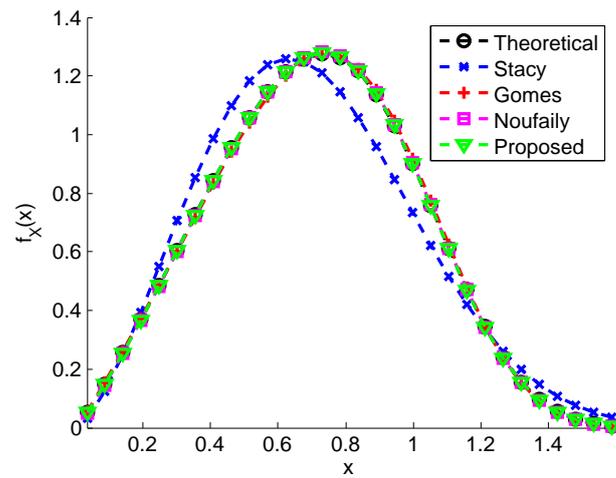
In this section we test the performance of the GG distribution for characterizing tissues of real images. For this purpose, we used a set of 518 real US images (584×145 , 8 bits) obtained from 3 human subjects by means of a clinical machine GE Vivid 7 echographic system (GE Vingmed Ultrasound A.S., Horten, Norway). The images were acquired before the Cartesian rearrangement².

In Fig. 4.7, an example of a real US images is shown with its Cartesian rearrangement. The red contour is the segmented areas of blood which are considered in the study, while the green contour is the segmented areas of tissue. The intersection of both regions was rejected in the study.

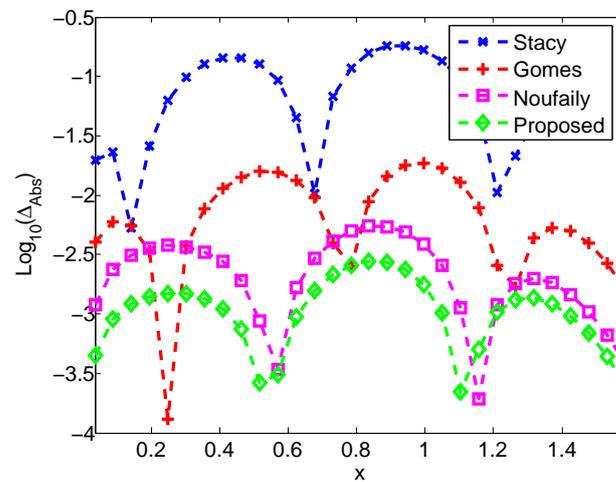
Additionally, the histogram of the image was depicted for the blood region as well as the fitted distributions most commonly used to characterize tissue. From the whole data set, a total number of 3185 regions were segmented for myocardial tissue while 1960 were segmented as blood. The sizes of regions vary depending on the tissue. However, it is high enough to provide a good estimate of the parameters. For instance, the segmented region of Fig. 4.7 has 18250 samples for blood and 5529 for tissue.

In the case of Fig. 4.7, the lower value of the histogram shown is 19 since the intensity values in the blood area was in the interval $[19, 156]$. The number of bins used for the representation of the histogram was set to 20 equally spaced in that interval.

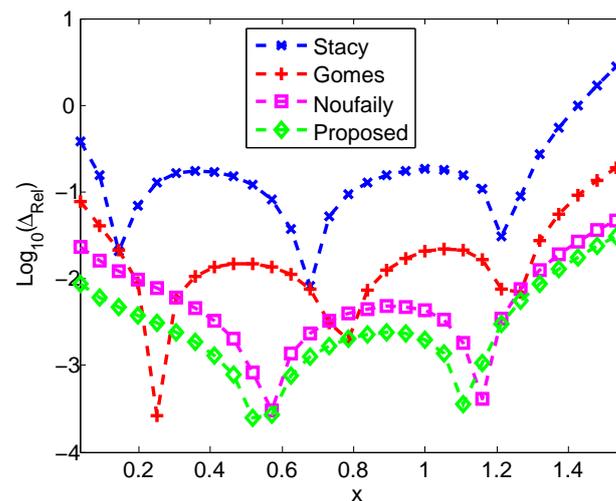
²The image collection was supervised by specialists Marta Sitges and Etelvino Silva from (Hospital Clinic; IDIBAPS; Universitat de Barcelona, Spain). The subjects were volunteers for a study of the reconstruction process of ultrasonic images. The acquisition was done in the Hospital Clinic of Barcelona with its approval. The images were provided by Nicolas Duchateau (CISTIB - Universitat Pompeu Fabra, Ciber-BBN, Barcelona, Spain) and Bart Bijmens (Instituco Catalana de Recerca i Estudis Avan cats (ICREA), Spain). The authors want to thank all of them for providing the images.



(a)



(b)



(c)

Figure 4.4: Example of the fitting performance for 10^4 samples. Methods: Stacy [Sta65], Gomes *et al.* [Gom08], Noufaily [Nou12] and the proposed one of section 4.2.2. (a) Probability Density Functions, (b) Absolute error of the PDFs, (c) Relative error of the PDFs.

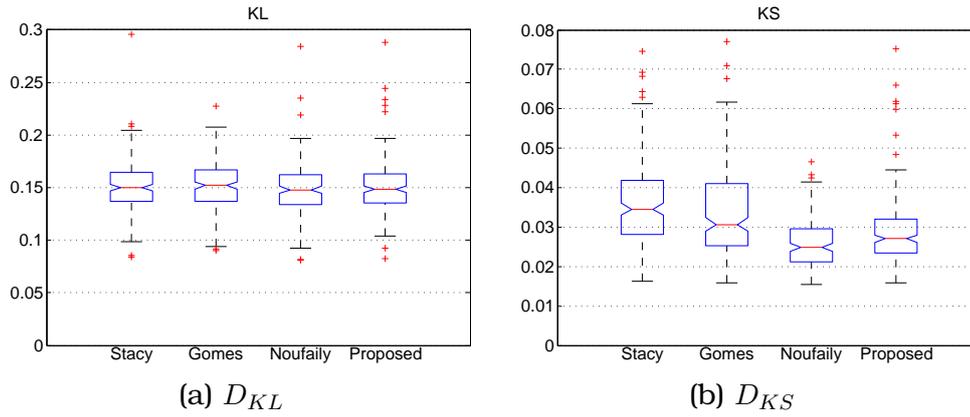


Figure 4.5: Results for D_{KL} and D_{KS} for 500 samples. Methods: Stacy [Sta65], Gomes *et al.* [Gom08], Noufaily [Nou12] and the proposed one of section 4.2.2.

The performance of the GG distributions was tested by estimating the PDFs for both tissue classes (myocardial tissue and blood) for the following distributions: Exponential, Rayleigh, Weibull, Normal, Nakagami, Gamma, GG. The PDFs were compared by means of both the D_{KL} and the D_{KS} measures. The results of the comparison is depicted in Fig. 4.8 where the better performance of the Gamma, Nakagami and GG becomes clear. In order to see whether these measures are statistically significant we carried out a Welch t-test for the Gamma, Nakagami and GG distributions for the D_{KS} measures. This test was chosen since no equal variance should be assumed and the D_{KS} since it does not depend on the empirical PDF estimate but just on the samples. The assumed hypothesis H_0 is that “both distributions have the same mean”, H_1 indicates that the null hypothesis can be rejected at a 5% of significance level.

The results are shown in tables 4.1 and 4.2. Note that all the null hypothesis were rejected but just one: Myocardial Tissue. In that case, the difference of the mean value of the Gamma and the GG is not statistically significant. The mean values of the D_{KS} are represented in table 4.3 where the lower mean value of the GG for both tissues can be appreciated. The results of the t-test of tables 4.1 and 4.2, and the lower mean values of the D_{KS} evidence the better performance of the GG than the rest of the distributions, with the exception of the myocardial tissue, where a Gamma distribution offers the same performance.

Blood	p-value	Hypothesis
Nakagami vs. Gamma	$< 10^{-15}$	H_1
Gamma vs. GG	$3.38 \cdot 10^{-7}$	H_1
Nakagami vs. GG	$< 10^{-15}$	H_1

Table 4.1: Results of the t-test for Blood.

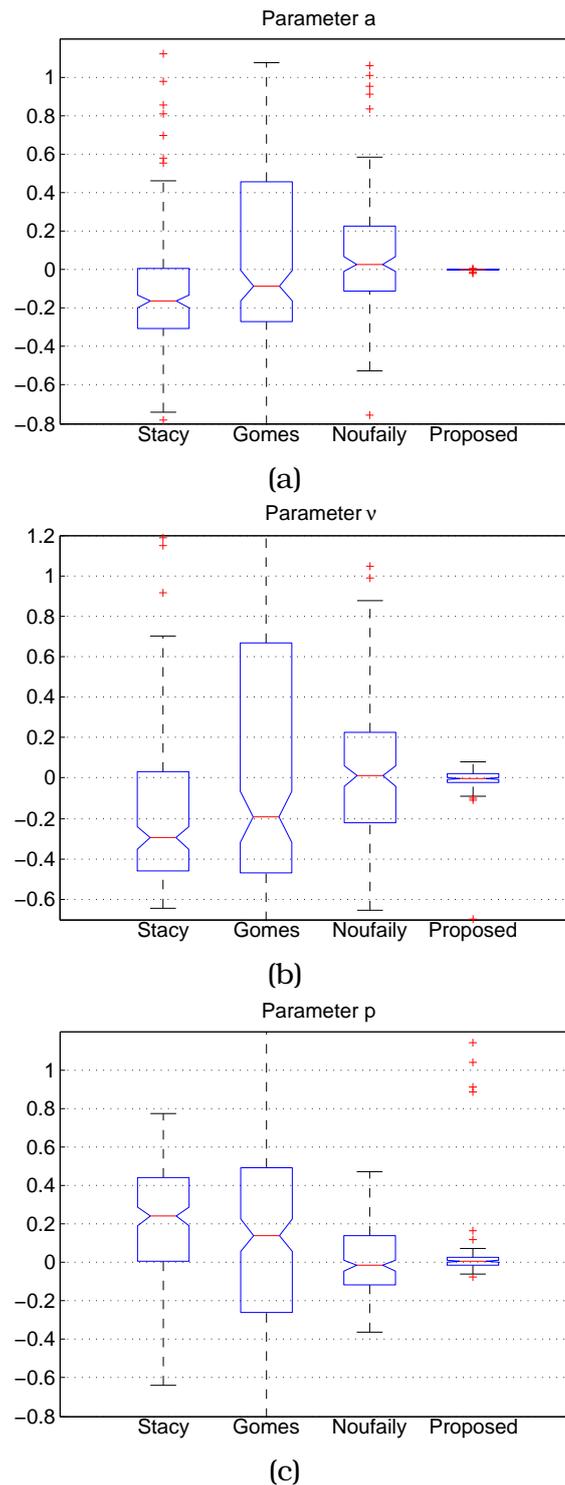


Figure 4.6: Results for the relative error of the estimates for 500 samples. Methods: Stacy [Sta65], Gomes *et al.* [Gom08], Noufaily [Nou12] and the proposed one of section 4.2.2.

4.3.3 Performance of the GGMM methods

In this section we test the performance of the proposed GGMM methods in three different scenarios. First we test the necessity of using more than a

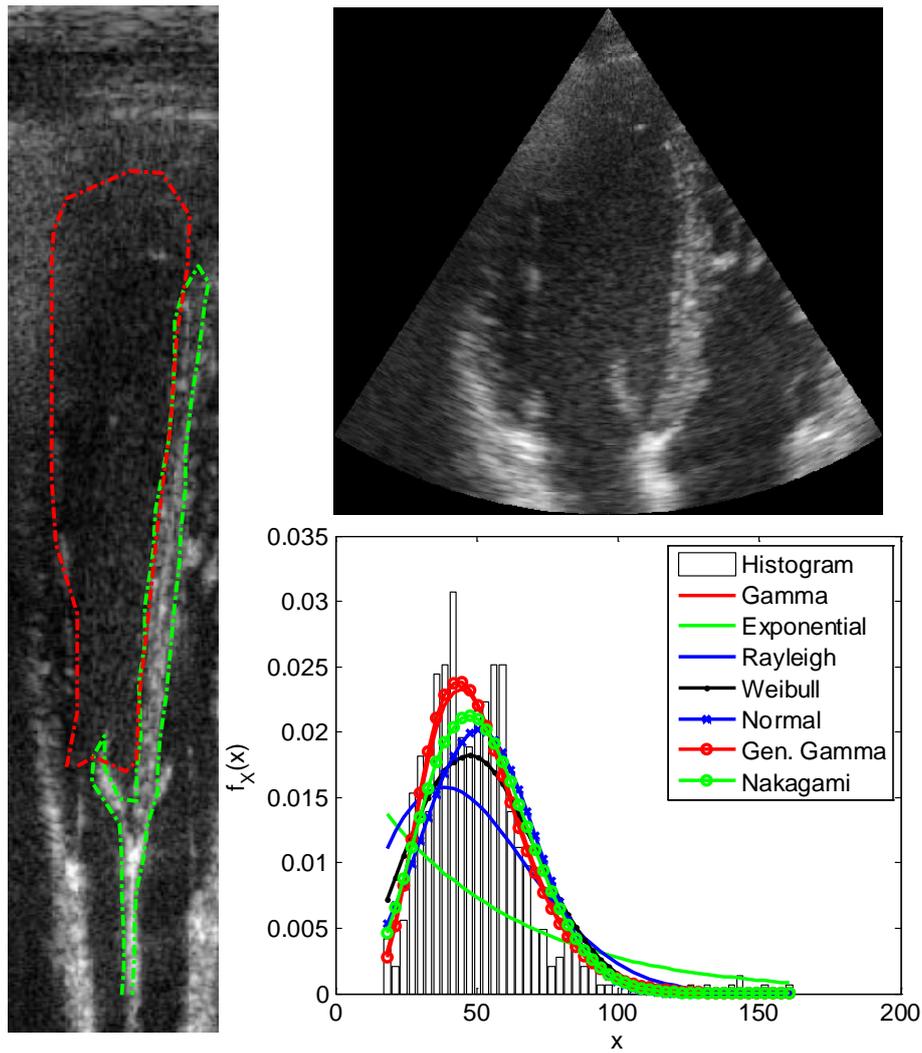


Figure 4.7: Example of an image of the data set. The red contour is the segmented areas of blood which are considered in the study, while the green contour is the segmented areas of tissue. The intersection of both regions was rejected in the study.

Myocardial Tissue	p-value	Hypothesis
Nakagami vs. Gamma	3.2410^{-4}	H_1
Gamma vs. GG	0.96	H_0
Nakagami vs. GG	2.7410^{-4}	H_1

Table 4.2: Results of the t-test for Myocardial Tissue.

	Nakagami	Gamma	Generalized Gamma
Blood	$5.5626 \cdot 10^{-2}$	$4.4970 \cdot 10^{-2}$	$4.2860 \cdot 10^{-2}$
Myocardial	$5.7711 \cdot 10^{-2}$	$5.5665 \cdot 10^{-2}$	$5.5644 \cdot 10^{-2}$

Table 4.3: Mean values for D_{KS} .

simple GG for describing tissues with an increasing echolucent response of the effective scatterers. The case of a variation of the number of effective scatterers is also considered. This behavior can be found in structures with

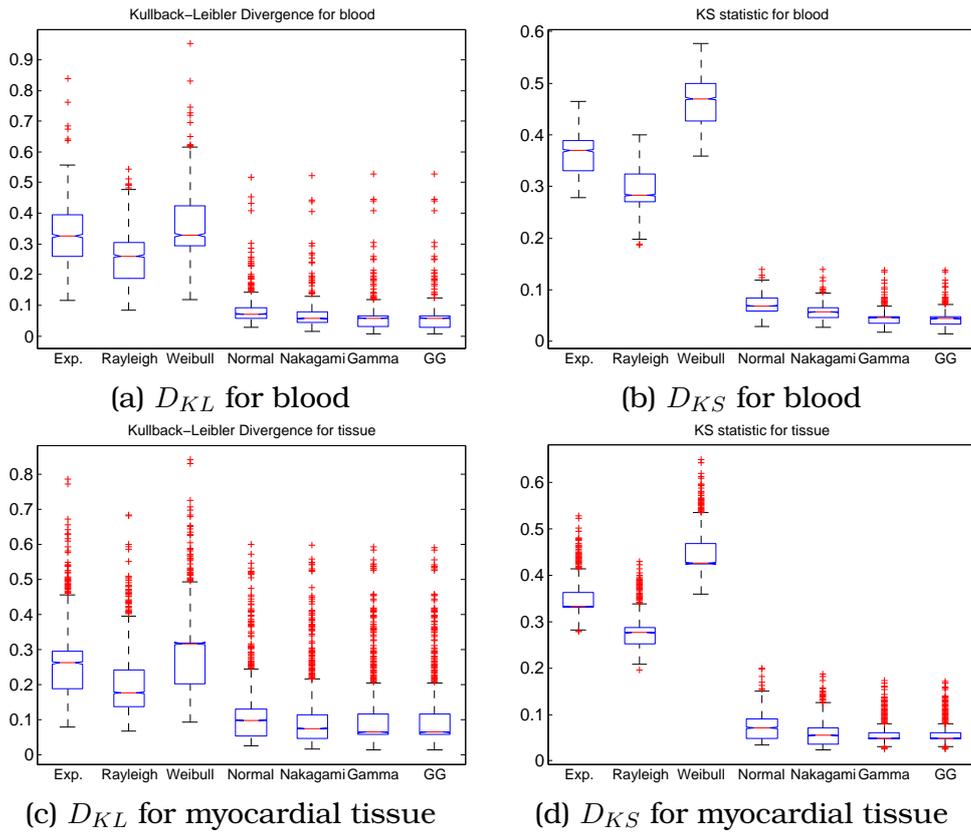


Figure 4.8: Results for the relative error of the estimates for 500 samples. Methods: Stacy [Sta65], Gomes *et al.* [Gom08], Noufaily [Nou12] and the proposed one of section 4.2.2.

an increasing deterministic response that changes the speckle nature from fully formed speckle to fully resolved speckle. The variation of the number of effective scatterers can be found in structures which change their scattering cross section.

In order to simulate B-mode US images, we followed the same methodology proposed in [VSF10b]. This method scans an image and records the data in a matrix which is corrupted by means of the speckle formation model of Eq. (4.1) where the tissue is modeled as a collection of scatterers of size comparable to the wavelength. The speckle pattern is obtained by means of random walk which does not assume any statistical distribution in order to avoid any bias of the results. The Cartesian arrangement is obtained by means of linear interpolation of the corrupted samples.

As a first example, we simulate an increasing echolucent tissue which varies its intensity from 0 to 255 from left to right. The sampling process and the resulting B-mode image are shown in Fig. 4.9.(b). The number of samples were set to 50 angular samples and 100 radial samples, represented as red points. The amplitude of each scatterer is defined as a Normal distributed RV with zero mean and $\sigma = 8$. Note that, along with the variation of intensities from left to right, a specular component of the speckle will appear. The number of scatterers was set 20 in order to simulate fully formed speckle in regions with low echolucent response and fully resolved

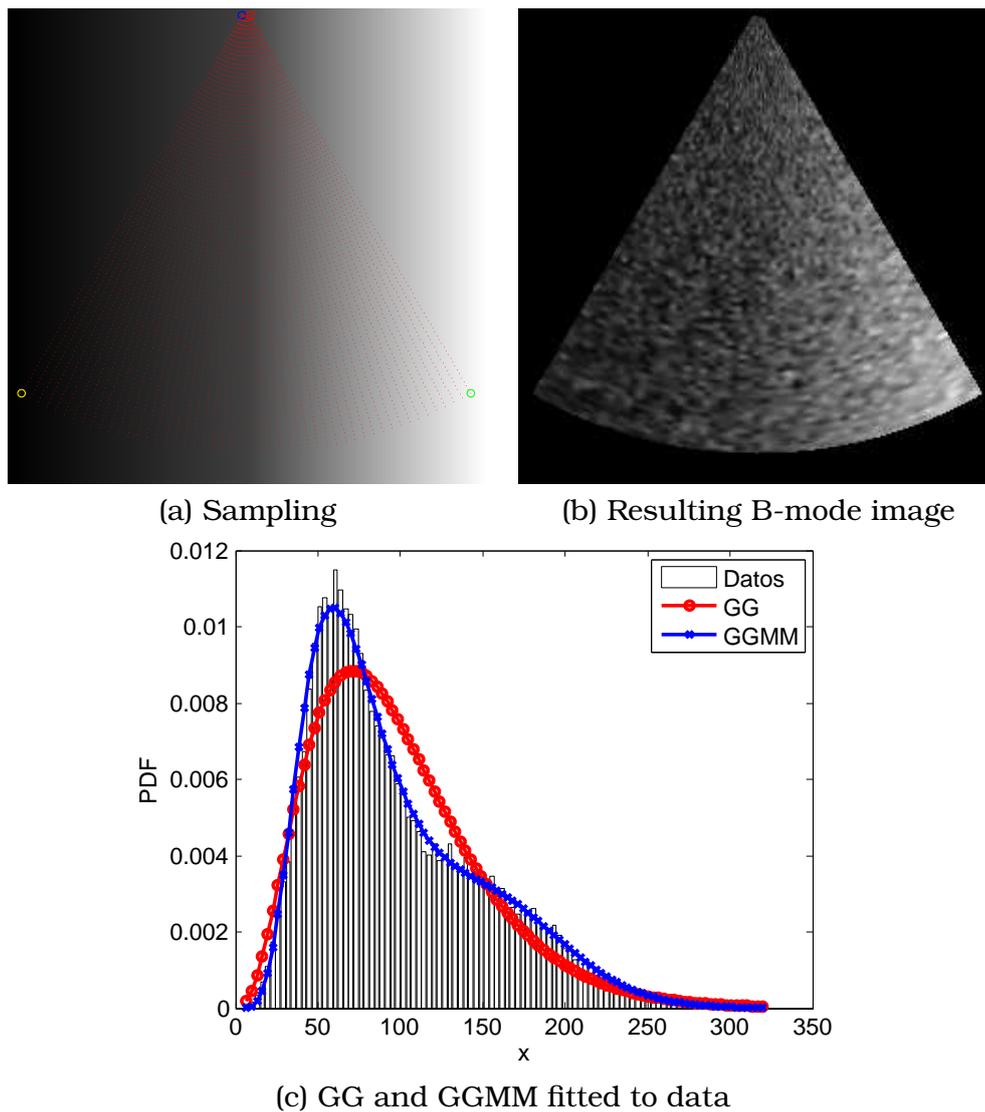


Figure 4.9: Simulation of spatial variant echolucent response of tissue. (a) Sampling of an increasing echolucent tissue. (b) Resulting B-mode image obtained by corrupting the samples by a random walk process of 20 scatterers per resolution cell, in order to simulate fully formed speckle in regions with low echolucent response and fully resolved speckle in regions with high echolucent response. (c) Histogram of (b) and a GG and GGMM with 2 components.

speckle in regions with high echolucent response. The resulting B-mode image is represented in 4.9.(b).

The fitted GG and GGMM with 2 components depicted in Fig. 4.9.(c) show that one simple GG fails to model the probabilistic behavior of a spatially variant echolucent tissue, while a GGMM with 2 components properly describes the echolucent variation.

As an additional experiment, in Fig. 4.10 we represent the spatial variation of the number of effective scatterers. The simulation was performed with the same sampling parameters as was done in the previous experiment. In this case the echolucid response was set to be homogeneous with

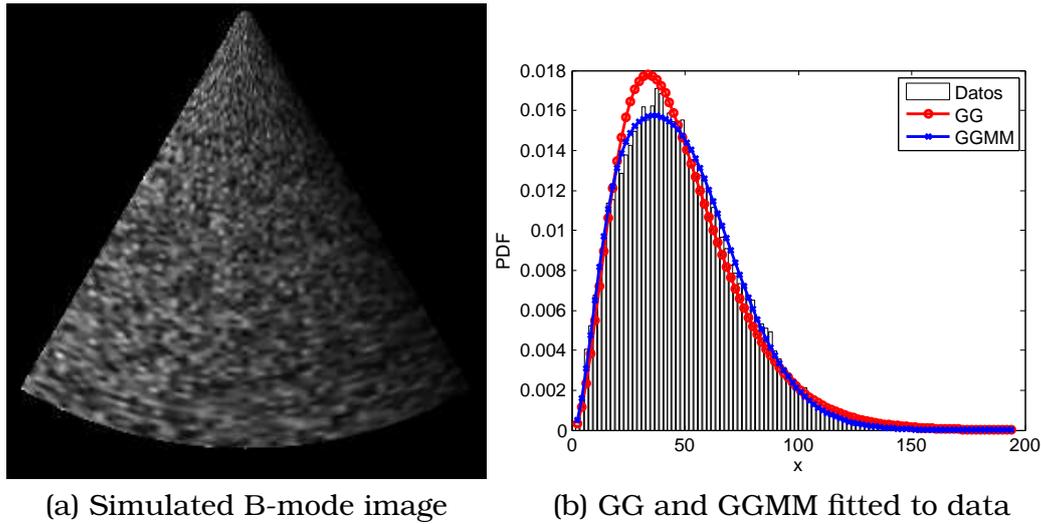


Figure 4.10: Simulation of spatial variant density of scatterers. The number of scatterers per resolution cell decreases from left to right in order to simulate fully formed speckle in regions with low density and partially resolved speckle in regions with high density.

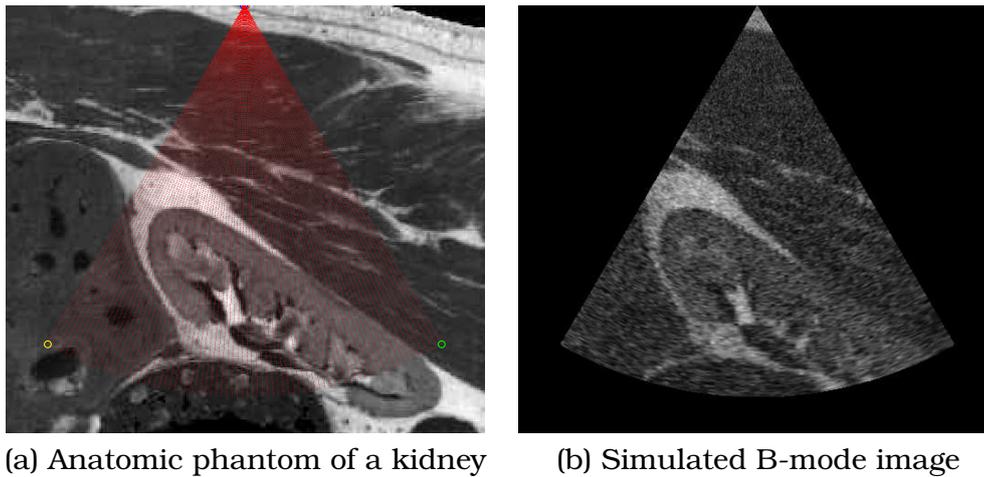
no deterministic component. Thus, the nature of the speckle changes from fully formed speckle to partially formed speckle. The number of scatterers decreases from left to right from 256 to 1. The amplitude of each scatterer is defined as a Normal distributed RV with zero mean and $\sigma = 8$.

The speckle PDF in this case becomes more impulsive in areas with more effective scatterers (left part of Fig. 4.10.(a)), this behavior is observed in a lower decay of the tail. Both the simple GG and GGMM with 2 components were calculated from the data and are depicted in Fig. 4.10.(b). In that figure, the fitting of a simple GG clearly shows that one component does not suffice to describe a spatial variation of number of effective scatterers.

In last synthetic experiment for testing the necessity of GGMM we simulate an anatomic phantom of a kidney scan. For this purpose, we used the artificial kidney scan proposed by Jensen [Jen96]. The image can be downloaded from the Field II website³. The sampling of the kidney and the resulting B-mode image are represented in Fig. 4.11. In this case, a GGMM with 4 components was used to fit the PDF of the image. The probability of belonging to each component is represented in Fig. 4.12 where the differentiation of tissues can be easily observed. In this case, a lower number of components fail to describe the kidney and the surrounding tissue which have a similar speckle response.

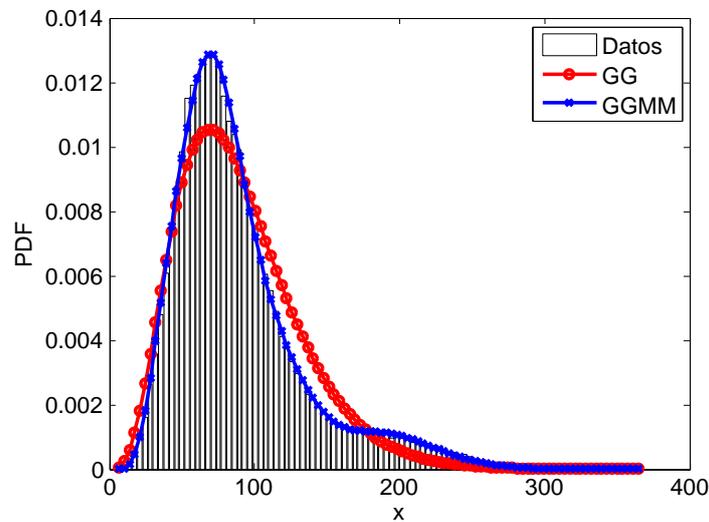
For testing the performance of the proposed GGMM methods with real data, we use the same data set used in the previous section. The number of components is set to two: blood and myocardial tissue. In order to compare the performance of the GGMM methods, we also fit a Gamma Mixture Model and a Nakagami Mixture model to the data [Des09, Web00, Cop03]. Both the D_{KL} and the D_{KS} were calculated for the mixture models in each image. The number of iterations for each mixture model was set to 100 and the

³<http://field-ii.dk/>



(a) Anatomic phantom of a kidney

(b) Simulated B-mode image



(b) GG and GGMM fitted to data

Figure 4.11: Simulation of an anatomic phantom of a kidney scan.

tolerance to 10^{-8} .

The lower values of D_{KL} and D_{KS} shown in Fig. 4.13 evidences the better characterization of the GGMM when compared to the NMM or the GMM. These results were expected due to the results of the previous section. Again, the t-tests were performed to the D_{KS} measure of the data. All the mixtures were statistically different with the exception of the $GGMM_1$ and $GGMM_2$. In that case, a p-value of 0.4906 was obtained. These results show once more that the GG can characterize better than other commonly accepted distributions and the differences are significant.

4.3.4 Potential Applications of the GGMM

A proper characterization of the speckle by means of suitable distributions can be used to guide segmentation algorithms as the one in [Des09]. The parameters of the mixture model can be used as features for developing a

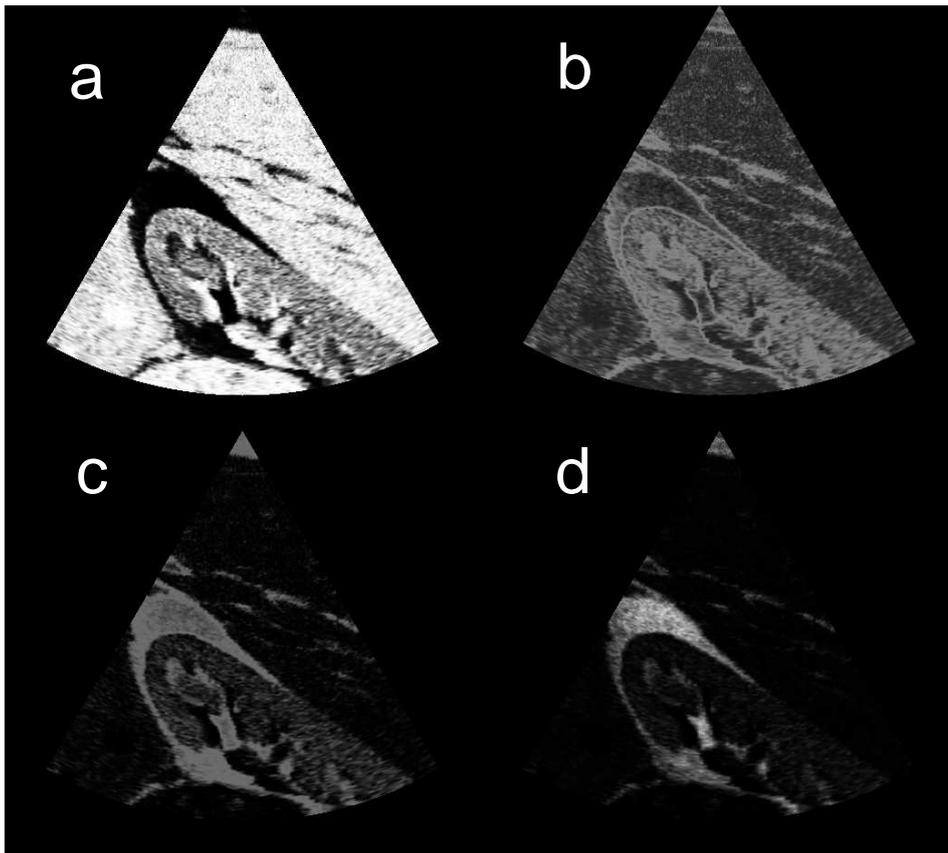


Figure 4.12: Probability of belonging to each component of the GGMM fitted to the image in Fig. 4.11.(b). The components are sorted in increasing mean value.

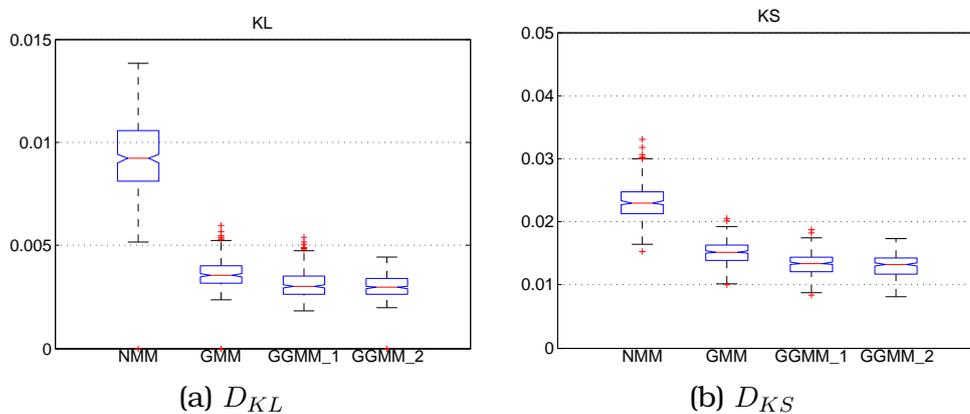


Figure 4.13: Results for D_{KL} and D_{KS} of the Mixture Models: $GGMM_1$, $GGMM_2$, GMM and NMM .

classifier as was done in [Seal1]. Furthermore, some filters use the probability of belonging to each tissue class. As an example of the performance of the GGMM, we show the some results of the Probabilistic-Driven Oriented Speckle Reducing Anisotropic Diffusion (POSRAD) [VSF10a].

This last filter includes the probability of belonging the each tissue class and adapts the diffusion tensor. Concretely, it calculates the structure tensor of the posterior probability and detects the most probable edges of the

image. This information is used to define the diffusion tensor which provides a better behavior in the boundaries of the image.

The structure tensor of the probability density function for each tissue class is calculated as

$$T_j(x_i) = G_\sigma * (\nabla_\sigma p(Z_i = j|x_i, \Theta) \cdot \nabla_\sigma p(Z_i = j|x_i, \Theta)^T), \quad (4.62)$$

where G_σ is a Gaussian kernel of standard deviation σ , and $\nabla_\sigma p(Z_i = j|x_i, \Theta)$ is the gradient of the probability density function for each tissue class filtered with a Gaussian kernel of standard deviation σ . Finally, let $\lambda_1^j \geq \lambda_2^j$ be the eigenvalues and $(\mathbf{v}_1^j, \mathbf{v}_2^j)$ their respective eigenvectors. The local orientation of the maximal variation of probability of the class \mathcal{C}_j is given by \mathbf{v}_1^j , and the local orientation of the minimal variation is given by \mathbf{v}_2^j .

Let consider the following diffusion equation

$$\begin{cases} u(x, 0) & = & u_0 \\ \frac{\partial u}{\partial t} & = & \text{div}(D\nabla u), \end{cases} \quad (4.63)$$

where the matrix D is the diffusion tensor which can be described by its eigenvectors $(\mathbf{v}_1, \mathbf{v}_2)$ and eigenvalues λ_1, λ_2 .

Given a diffusion tensor, D , the diffusion of the intensity values of the image is performed in the direction of eigenvectors with different diffusion coefficients. For each eigenvector, its eigenvalue defines the diffusion coefficient and, thus, an anisotropic diffusion can be achieved.

As an example, when one eigenvalue is equal to 1 and the other one is 0, a complete anisotropic diffusion is obtained, since the intensity values diffuse in the direction of the eigenvector associated to the eigenvalue equal to 1. This would be the desired behavior of a filter in regions where structures must be preserved. When both eigenvalues are equal to 1, the diffusion process becomes isotropic and the intensity levels diffuse equally in all directions. This case would be the desired behavior for homogeneous regions where no structures must be preserved.

The POSRAD philosophy makes use of the structure tensors determined out of the probability maps to obtain the most probable structures. In that case, the diffusion filter should be anisotropic. When no probable structures are detected, the diffusion should be isotropic.

Since we have J structure tensors (each tissue class with probability density function), we choose the eigenbase of the structure tensor with maximal λ_1^j : $\hat{j} = \arg \max_j (\lambda_1^j)$. This base gives the orientation of the maximal variation of probability among all the classes.

The interpretation of this choice is that we choose as boundary the one with the maximal gradient of the probability density function over all tissue classes. This way, the most probable boundary is preserved in the filtering process. In the basis of $T_{\hat{j}}$, namely $(\mathbf{e}_1, \mathbf{e}_2)$, the diffusion matrix D is defined

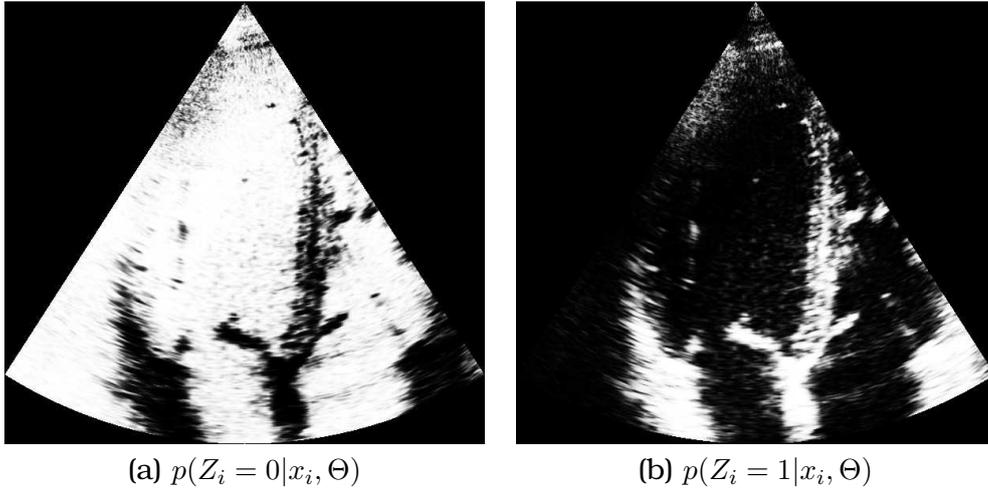


Figure 4.14: Probability of belonging to each tissue class, where the class 0 describes the blood and the class 1 describes the myocardial tissue.

as

$$D = E \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} E^T \quad (4.64)$$

where

$$\begin{aligned} \lambda_1 &= 1 - \|\nabla_{\mathbf{e}_1, \sigma} p(Z_i = j|x_i, \Theta)\|_2 \\ \lambda_2 &= 1 \end{aligned} \quad (4.65)$$

and $\|\cdot\|_2$ is the 2-norm, $\nabla_{\mathbf{e}_i, \sigma}$ is the directional derivative in the direction \mathbf{e}_i filtered with a Gaussian kernel with a standard deviation σ , and E is the matrix whose columns are the eigenvectors ($\mathbf{e}_1, \mathbf{e}_2$). This definition performs a diffusion filtering in the direction of the minimal variation of probability (\mathbf{e}_2) while preserves the maximal variation of probability since $\|\nabla_{\mathbf{e}_1, \sigma} p(Z_i = j|x_i, \Theta)\|_2$ will have a value closed to 1. Note that the discrete approximations of $\|\nabla_{\mathbf{e}_1, \sigma} p(Z_i = j|x_i, \Theta)\|_2$ is bounded in $[0, 1]$, thus $\lambda_1 \in [0, 1]$.

In Fig. 4.14 we show the probability of belonging to each tissue class, $p(Z_i = j|x_i, \Theta)$, provided by the GGMM method (the $GGMM_2$ was used for this example). All the figures of the example are represented in their Cartesian arrangement in order to ease visualization of fine structures. Note that the structures are clearly identified by each posterior probability of each tissue class and the filter can perform an efficient anisotropic diffusion. To see this, in Fig. 4.15 we represent λ_1 , which describes the anisotropic behavior of the filter. When $\lambda_1 = 1$, the filter acts like a conventional isotropic filter whereas the pure anisotropic behavior is carried out when $\lambda_1 = 0$.

Finally, the resulting image after 40 iterations is depicted in Fig. 4.16 in comparison to the original one.

As a final application of the GGMM, one can make use of the pixel-wise probability of belonging to each tissue class to obtain a spatially coherent probability by introducing a undirected graph where the nodes (each pixel of the image) represent a random variable and the edges of the graph represent the relationships between nodes as it is represented in Fig. 4.17. The problem is reduced to find the labels for each node by taking into account the

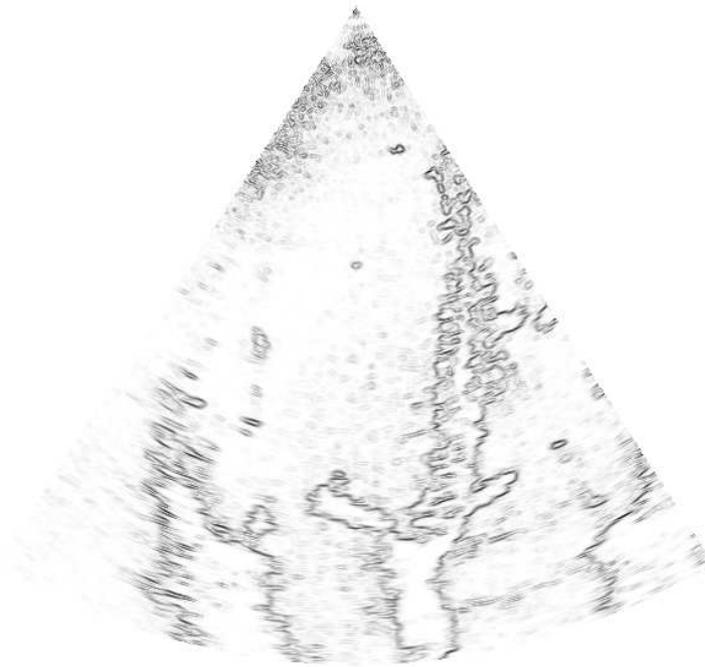


Figure 4.15: Anisotropic behavior of the filter. The most probably edges of the image are described by the lower values of $\lambda_1 \in [0, 1]$.

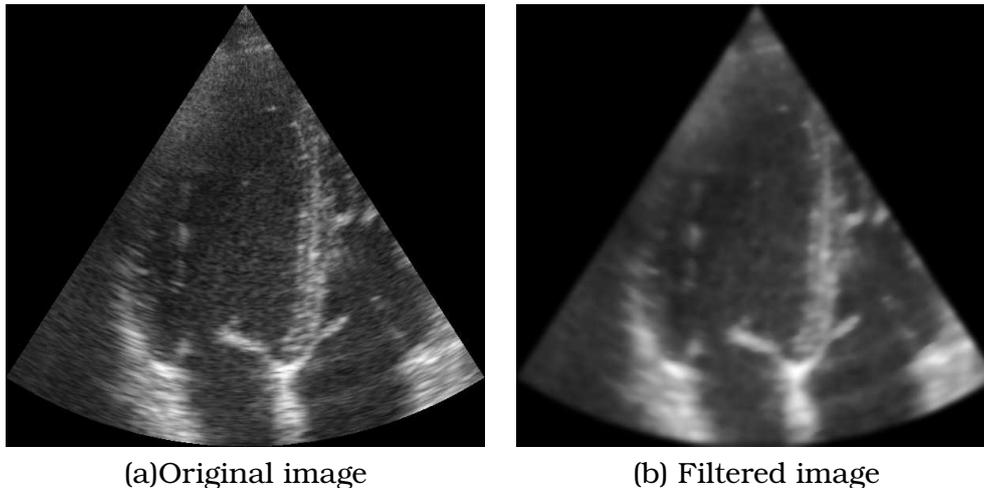


Figure 4.16: Results of the POSRAD filter. The anisotropic behavior of the filter is appreciated in the preserved details of the myocardial tissues.

relationships between nodes of the local neighborhood (the Markov property is assumed). This problem, though is intractable in terms of direct probabilistic inference, can be solved by means of the Loopy Belief Propagation (LBP) algorithm introduced by Pearl in [Pea82]. This algorithm performs approximate inference of a graphical model. Although LBP does not guarantee to converge due to the presence of loops in the graph, however it has shown good experimental results and is commonly used [Ksc01].

In the end, the problem is faced as a discrete MRF where the labels, Z , are each tissue class and the nodes are the pixels of the image. The energy

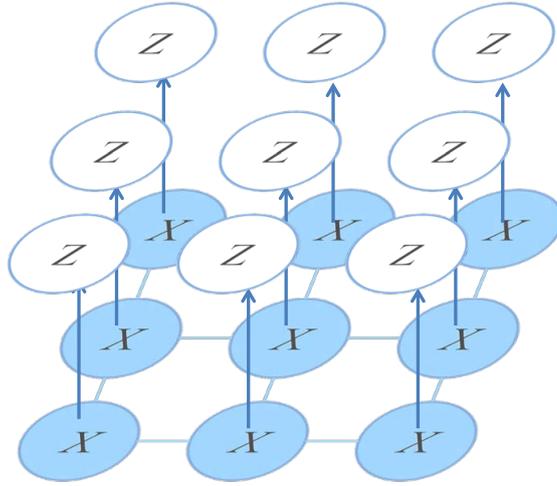


Figure 4.17: Undirected graph. The nodes represent a random variable X and the edges relationships between nodes. Each random variable can be classified as a tissue class J .

to be minimize by the LBP method can be defined as

$$V(Z) = \sum_{i=1}^N V_1(Z_i) + \sum_{k \in \eta(i)} V_2(Z_i, Z_k), \quad (4.66)$$

where $\eta(i)$ is the neighborhood of the i -th node, $V_1(Z_i) = -\log p(Z_i = j|x_i, \Theta)$ and

$$V_2(Z_i, Z_k) = - \sum_{k \in \eta(i)} \log p(Z_k = z_i|x_k, \Theta). \quad (4.67)$$

The output of the LBP is a belief of node i belongs to class $Z = j$. Thus, the probability with spatial coherence can be directly obtained from the outputs of the LBP algorithm. In Fig. 4.18 are represented the probability of each tissue class when the spatial coherence is introduced.

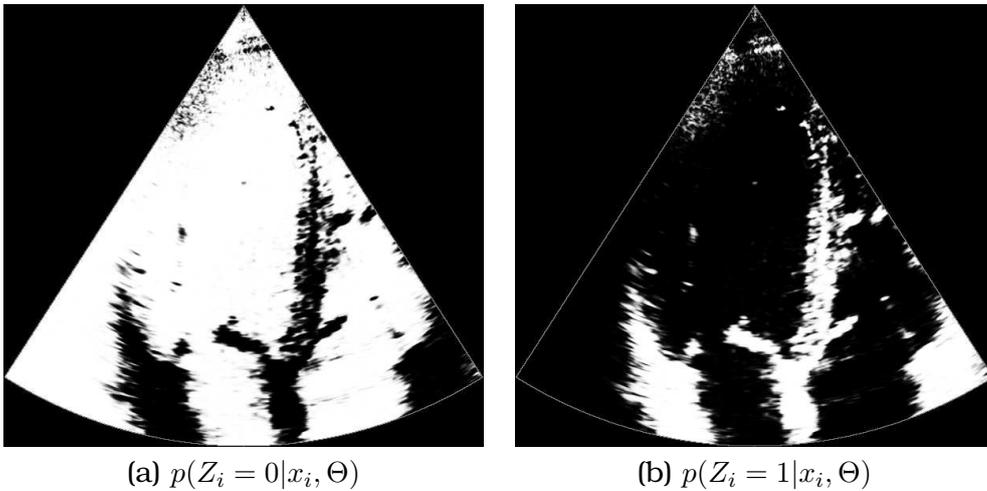


Figure 4.18: Probability of belonging to each tissue class after the LBP, where the class 0 describes the blood and the class 1 the myocardial tissue.

This coherent probability maps can be of help for classifying purposes or as prior information for segmentation algorithms. The valuable information that they provide can be seen in a simple experiment in which we consider the classification of two tissues (blood and myocardial tissue) and we compare the results with the k-means algorithm applied to the original image and a simple classifier consisting in assign the class with maximum posterior probability. The results of this example are shown in Fig. 4.19 where the identification of the myocardial tissue is clearly obtained by the posterior probability of the GGMM whereas the k-means method cannot properly define a contour of each tissue.

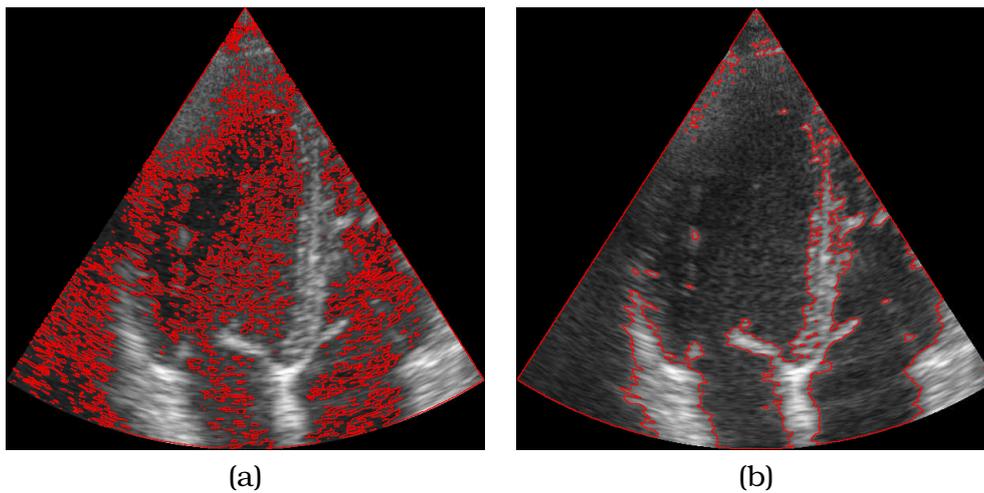


Figure 4.19: Simple example of the valuable information of the posterior probability obtained from the GGMM with spatial coherence. (a) Classification with k-means, (b) Classification with the GGMM probability maps after LBP.

4.4 Conclusions

Throughout this work we have analyzed the advantages of using a GG distribution for characterizing the speckle in ultrasound images. This distribution offers a suitable way to deal with the impulsive behavior of speckle which causes heavier tails in the distributions. Additionally, the GG is a natural generalization for many distributions commonly used to characterize the speckle: Rayleigh, Gamma, Nakagami, Weibull, Exponential, Rician [Sha01]. Thus, it has all the advantages of these distributions and avoids some of their generalization problems.

Although some approaches have used this distribution in the literature, the inconveniences of estimating its parameters make this option thorny and not attractive. The problems stems from the inaccurate estimate of the moments method proposed in [Sta65] and the impossibility of obtaining a closed-form ML estimates. Some solutions have been recently proposed such as heuristic methods [Gom08], which are strongly dependent on the number of samples, and iterative methods [Nou12] which depends on the initial condition.

In this work we have proposed a simple methodology to calculate the ML estimate which offers robust results comparing to the methods in the literature [Sta65, Gom08, Nou12]. It is worth to mention that the fundamentals of the ML method of [Nou12] and the proposed one are the same since both try to find the solution of three simultaneous non-linear equations. However, the different optimization technique makes the proposed method more robust. Additionally, the performance for describing speckle was tested in a set of 518 real US images of the heart, in which 3185 regions were manually segmented for myocardial tissue and 1960 for blood.

Results with t-tests applied to the D_{KS} goodness-of-fit measure demonstrated the better behavior of the GG in most of the cases and in those cases where there were not statistical difference, the other distribution is a particularization of the GG.

The formulation of the proposed method allows to generalize this methodology to a GGMM. These mixture models are of great value due to the different nature of the echogenic response of tissues in the received signal. Two different methods were proposed for the calculation of the GGMM parameters $GGMM_1$ and $GGMM_2$. Both were developed by applying the EM method in the derivation of the proposed ML method, the optimization technique for $GGMM_1$ follows the same approach used for the proposed ML method. The $GGMM_2$ method makes use of the optimization technique proposed by [Nou12]. Results when comparing both methods to the GMM and NMM in real images showed the better fitting of the GGMM. No statistical differences were detected between $GGMM_1$ and $GGMM_2$.

Through this article we showed the better behavior of the GGMM methods when compared to the RMM, NMM and GMM for the case of cardiac imaging. The potentials of mixture models have proven a good classification performance in intravascular ultrasonic images for RMM [Sea11]. Additionally, the NMM showed good results for segmentation in carotid arteries [Des09]. In the case of filtering Cardiac imaging, the mixture models have also shown good results [VSF10b].

We think the GGMM methods here proposed can be used with good results in the aforementioned modalities since they generalize the RMM, NMM and GMM in a natural way and allow to describe heavier tails of the PDFs that the RMM, NMM and GMM fail to fit. Many other US modalities such as breast, liver and kidney should be considered. We hope the proposed GGMM methods can encourage future research for tissue characterization in those different US modalities.

Finally, we want to recall that the potential applications of GGMM do not confine to those proposed in this paper. We hope the results of this work can revive the use of the GG distribution and its extension, the GGMM, in many other areas.

Speckle characterization with Heavy-Tailed Models

*All possible definitions of probability
fall short of the actual practice.
William Feller, 1906–1970.*

Abstract– Several probabilistic models have been used in the last decades to describe the speckle pattern of US images. The most common accepted models have been the Rayleigh because of its simplicity and the Nakagami due to its better empirical goodness-of-fit performance. However, these models fail to fit the heavy tailed behavior of the probability distributions due to the highly impulsive response of scatterers. In this work, a probabilist model, which describes the heavy tailed behavior of speckle and considers the specular component, is presented. Results show that the hypotheses considered in this work fit remarkably better than those of other models.

Adapted from: G. Vegas-Sánchez-Ferrero¹, F. Simmross Wattenberg¹, M. Martín-Fernández¹, C. Palencia², C. Alberola-López¹, *Caracterización de Speckle con Modelos de Cola Pesada*, XXX Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB), San Sebastián, Spain, November 2012.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Departamento de Matemática Aplicada, Univ. Valladolid.

5.1 Introduction

Among non-invasive techniques, ultrasonic imaging is probably the most extended. The main reason of its success stems from its low cost for diagnostic support and its wide use in medical protocols. However, US images show an intrinsic granular pattern in this modality which make its interpretation difficult. It is the so-called *speckle*. This pattern is due to the random interferences of multiple coherent components produced by microscopic elements into the medium. Concretely, in the case of US, the resolution cell, the number of effective scatterers and the acquisition protocol are the elements which participate in the speckle formation [Sha00].

The analysis of the incoming echo wave leads to a characterization of tissues for medical analysis. Thus, and due to the random nature of speckle, several probabilistic models have been proposed in the literature. This characterization has been successfully used in applications such as segmentation, classification and filtering. However, the performance of these methods strongly depends on the probabilistic model used for describing tissues in US. For this reason, probabilistic modeling of amplitude signals in US has been a prolific topic in the last years.

Probably, the most celebrated model has been the one-parameter Rayleigh distribution, which describes the so-called fully formed speckle. This pattern appears whenever there are a high number of effective scatterers in the resolution cell. However, real images show a deviation from this theoretical model due to a lower number of scatterers in the resolution cell or the predominance of some components in the cell. The distributions most commonly used are the Rice distribution, when specular component is considered (fully resolved), K distribution (partially formed) and K-homodyne (partially resolved).

Though all these models stems from physical suppositions on the scattering phenomena of random media, other distributions have shown its empirical effectiveness in real images. This is the case of Gamma [VSF10b] and Nakagami [Sha00] distributions. The Gamma is a two-parameter distribution which describes the resulting distribution of interpolated fully formed speckle. The Nakagami was proposed as a two-parameter distribution that generalizes the Rayleigh and approximates the Rice distribution.

The capability of the Nakagami distribution to model fully formed and fully resolved speckle make it the most accepted model. However, the tails of the Nakagami, K, Rayleigh or Gamma distributions do not show the impulsive response which appears in speckle. This response is produced as a result of the highly impulsive nature of scatterers causing the huge dynamic range of the signal, which is appreciated in the heavy tails of distribution of the amplitude signal.

One of the distributions proposed to fit the impulsive response of speckle is the Heavy-tailed Rayleigh model (HT-Rayleigh) [Kur04]. Nevertheless, this model does not consider the presence of specular scattering and, thus, it is

not able to describe the fully resolved speckle.

In this work, we propose an α -stable (AS) model which describes the impulsive behavior that causes the heavy tails and also takes into account the specular components of fully resolved speckle. Thus, it is a generalization of the HT-Rayleigh model proposed in [Kur04] and also generalizes in a natural way the Rayleigh and Rice models, since the Gaussian hypothesis is a particular case of the proposed.

The main contribution of this work is two-fold: **1)** We show that the AS hypothesis can be applied to US real images and it fits remarkably better than the Gaussian hypothesis commonly used. **2)** We show that the HT-Rayleigh model is insufficient for characterizing fully resolved speckle, since the null specular component is not always admissible.

The experiments were carried out using 30 plaques of coronary arteries which were histologically analyzed and classified as calcified tissue, where the specular component is not negligible. Additionally, 50 Lumen regions were analyzed, where there are no relevant specular structures.

5.2 Probabilistic Models

5.2.1 Classical Models

Image acquisition in US begins with the emission of packets of echo pulses that travel along the tissue. The scattering phenomena is produced by the effective number of scatterers in the resolution cell. They contribute in the change of shape of the pulse depending on the tissue characteristics such as the number of effective scatterers and their shapes [Sha00, VSF10b].

The echo model after the scattering, $s(t)$, is usually treated as a “*random walk*” process due to the random location of scatterers in the resolution cell [Sha00]

$$s(t) = \sum_{n=1}^N \alpha_n \cos(\omega_0 t + \phi_n), \quad (5.1)$$

where ω_0 is the average excitation frequency and N is the effective number of scatterers within the resolution cell. The phases, ϕ_n , are usually modeled as uniform distributed in $[0, 2\pi]$ and the amplitude is considered to follow a Normal distribution.

Fully formed speckle considers a high number of scatterers in the resolution cell and, by applying the Central Limit Theorem, the signal after the scattering can be expressed as

$$s(t) = X_r \cos(\omega_0 t) + X_i \sin(\omega_0 t), \quad (5.2)$$

where X_r and X_i are the in-phase and quadrature components respectively and follow independent and identically distributed Gaussian distributions

of zero mean. Thus, the envelop of the received signal is calculated as $R = \sqrt{X_r^2 + X_i^2}$ and follows a Rayleigh distribution [Wag83]

$$f_R(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, \quad r \geq 0. \quad (5.3)$$

On the other hand, following a different argument, Shankar in [Sha00] proposed a Nakagami distribution as a “simple and universal model for tissue characterization”. This model is not based on physical assumptions as those mentioned in the introduction. It is based on an empirical goodness-of-fit performance compared to the Rayleigh and Rice.

The PDF of the Nakagami distribution is the following

$$f_R(r) = \frac{2m^m r^{2m-1}}{\Gamma(m)(2\Omega)^m} e^{-\frac{m}{2\Omega} r^2} u(r). \quad (5.4)$$

This distribution shows good properties for describing the behavior of speckle because the Rayleigh is a particular case of it ($m = 1$). Additionally, when $m > 1$ it is similar to the Rice distribution. However, this distribution shows also some limitations since it is not able to fit the heavy tails appearing in real images due to the impulsive nature of speckle [Sha01].

5.2.2 HT-Rayleigh Model

This model was firstly proposed for the case of synthetic aperture radar images and was extended to the case of US images because both modalities are based on similar assumptions [Kur04]. The main advantage of this model is that it tries to describe the impulsive response of speckle by applying the AS hypothesis to the in-phase and quadrature components received in the probe.

Its derivation starts from the characteristic function of a bi-variant two-parameter AS distribution

$$\psi(t_1, t_2) = e^{-\gamma|\mathbf{t}|^\alpha}, \quad (5.5)$$

where t_1 and t_2 are the elements of the vector \mathbf{t} , and $|\mathbf{t}| = \sqrt{t_1^2 + t_2^2}$. The relation between the characteristic function and the PDF can be established by means of its Fourier transform

$$f_{X_1, X_2}(x_1, x_2) = \frac{1}{(2\pi)^2} \int_{t_1} \int_{t_2} e^{-\gamma|\mathbf{t}|^\alpha} e^{-j2\pi(x_1 t_1 + x_2 t_2)} dt_1 dt_2. \quad (5.6)$$

Finally, the PDF of the envelop obtained from the random variable $\mathbf{X} = (X_1, X_2)$ can be calculated by using polar coordinates and marginalizing with respect the angular variables

$$f_R(r) = \int_0^{2\pi} r f_{X_1, X_2}(r \cos \phi, r \sin \phi) d\phi, \quad (5.7)$$

where $X_1 = R \cos(\Phi)$ and $X_2 = R \sin(\Phi)$. In [Kur04] this expression is calculated by means of numerical quadrature methods, since the series expansion are unstable.

This model solves the main problem showed by the Rayleigh and Nakagami models, since it considers the highly impulsive response of speckle which originates heavy tails. However, it does not consider the specular component in the scattering model and, thus, shows an important drawback for tissue characterization with a relevant specular component; which is the case of muscular or calcified tissues.

5.3 α -Stable model for Fully Resolved Speckle

In this section we propose a more general α -stable distribution which accounts for the specular contribution of tissues.

The underlying hypothesis of classical models is the Gaussianity in the real and imaginary components of the *random walk* model. In the case of the HT-Rayleigh, both the real and imaginary components are supposed to be IID with a null specular component. In order to avoid this assumptions and to analyze the assessment of the hypothesis, we propose a model that generalizes the one proposed in [Kur04]. We consider specular components and different scattering parameters for each component. The characteristic function of an α -stable random variable of 4 parameters is the following

$$\psi(t) = \begin{cases} e^{j\delta t - \gamma^\alpha |t|^\alpha (1 + j\beta \text{sign}(t) \tan(\alpha\pi/2))}, & \text{si } \alpha \neq 1 \\ e^{j\delta t - \gamma^\alpha |t|^\alpha (1 + j\beta \text{sign}(t) \frac{2}{\pi} \log |t|)}, & \text{si } \alpha = 1, \end{cases} \quad (5.8)$$

where δ is the localization parameter, which models the specular component of the random variable and α takes values in $(0, 2]$. In the case of $\alpha = 2$ the distribution becomes Gaussian and, thus, this model generalizes in a natural way the Gaussian hypothesis. The parameter β is the asymmetry parameter and takes values in $[-1, 1]$. This value is expected to be around zero since no anisotropic response is assumed. The γ parameter is the so-called dispersion parameter, which is the counterpart of the standard deviation of the Gaussian case.

If the asymmetry parameter, β , is shown to be negligible and the dispersion of both real and imaginary components are the same. Thus, we can consider the characteristic function of a bi-variant and symmetric α -stable

$$\psi(t_1, t_2) = e^{j\delta_1 t_1 + j\delta_2 t_2 - \gamma^\alpha |t|^\alpha}. \quad (5.9)$$

This expression generalizes the HT-Rayleigh distribution since it considers the specular component of X_r and X_i . The derivation of the PDF for this case can be calculated by means of the Fourier transform

$$f_{X_r, X_i}(x_1, x_2) = \frac{1}{(2\pi)^2} \int_{t_1} \int_{t_2} \psi(t_1, t_2) e^{-jx_1 t_1 - jx_2 t_2} dt_1 dt_2. \quad (5.10)$$

Finally, the envelop of the PDF of the random variable $\mathbf{X} = (X_r, X_i)$ can be calculated by using polar coordinates and marginalizing with respect the angular variable

$$f_R(r) = \int_0^{2\pi} r f_{X_r, X_i}(r \cos \phi, r \sin \phi) d\phi, \quad (5.11)$$

where $X_r = R \cos(\Phi)$ and $X_i = R \sin(\Phi)$. This expression can be calculated by numerical methods in the same way as was done in [Kur04].

5.4 Results

The experiments were carried out using intra-vascular US signals from coronary arteries acquired after the gain compensation step. These signals were directly take from the probe before the envelop calculation and, thus, no log-compression or interpolation to Cartesian coordinates was performed¹. A set of 30 calcified plaques with histological validation and 50 lumen regions were considered in the study. They were acquired with an IVUS Galaxy II device (Boston Scientific) by using an Atlantis SR Pro 40MHz catheter (Boston Scientific). The sampling of data was performed at 200 MHz with an acquisition 12-bit Acquisis card.

We performed two different experiments for two different objectives:

1) To test the AS hypothesis in real data. For this purpose, we use the general model stated in Eq. (5.8) for each component (in-phase and quadrature) of the acquired signal. In doing so, the AS model of Eq. (5.8) is fitted to each component by means of the Nolan's algorithm [Nol97]. The parameters obtained for both components are compared in order to test the isotropic nature of data (i.e. same α and γ for both components) and the Gaussian model is also tested. This experiment allow us to test if the Gaussianity hypothesis is reasonable and if the proposed model of Eq. (5.9) is consistent.

2) To test the goodness-of-fit of the proposed model in both kinds of tissues. The model of Eq. (5.11) is compared to the HT-Rayleigh and the Nakagami models. This experiment will show if a significant difference between fitting PDFs of tissues by using the α -stable hypothesis versus the Gaussian hypothesis exists.

As a result of the first experiment, in table 5.1 the difference between parameters α of the in-phase and quadrature components is shown. That table shows the high similarity between AS parameters of both components. An unpaired student's t-test was performed with equal mean values as the null hypothesis, H_0 . The statistical significance was set to 5%. P-values and the accepted hypothesis is shown in table 5.2, where one can appreciate

¹The images belong to post-mortem arteries acquired in the Hospital Germans Trias i Pujol of Badalona, Spain, with its approval under the supervision of Dr. Josepa Mauri.

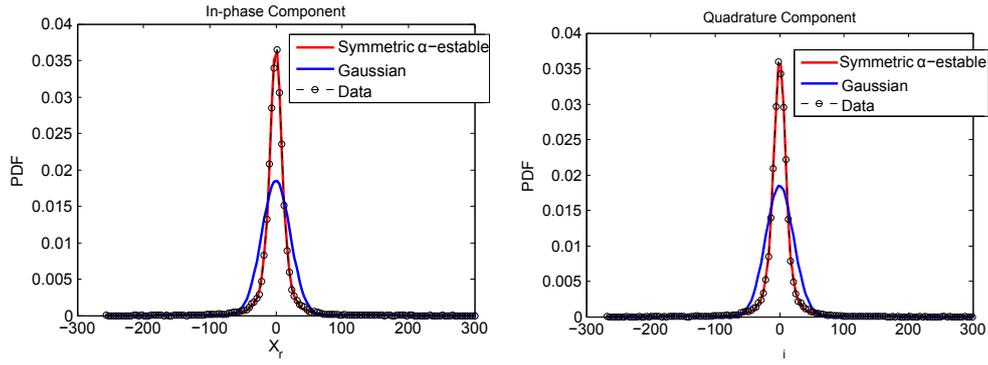


Figure 5.1: Comparison between the proposed AS model and the Gaussian model for the in-phase and quadrature signals for the case of Lumen.

	Lumen		Calcified Plaque	
	X_r	X_i	X_r	X_i
α	1.18 ± 0.18	1.18 ± 0.18	1.16 ± 0.22	1.13 ± 0.19
β	-0.01 ± 0.01	-0.01 ± 0.0	-0.001 ± 0.02	-0.005 ± 0.02
γ	30.48 ± 26.38	30.42 ± 26.30	341.62 ± 408.17	337.96 ± 406.44
δ	0.09 ± 0.17	0.14 ± 0.24	-0.53 ± 7.36	0.75 ± 6.77

Table 5.1: Parameters obtained for the AS distribution (mean \pm standard deviation).

	Lumen		Calcified Plaque	
	p-value	Hypothesis	p-value	Hypothesis
α	0.9519	H_0	0.6084	H_0
β	0.3316	H_0	0.4869	H_0
γ	0.9906	H_0	0.9642	H_0

Table 5.2: P-values for the unpaired student's t-test.

that all the null hypotheses were accepted. Thus, same mean value should be assumed.

On the other hand, the parameter β obtained for both tissues shows, in the worst case, a relative error of 0.3% with respect to the null value, which makes the symmetric assumption reasonable with an error lower than 1%. This result evidences that the bivariate PDF of the received signal shows a radial symmetry with respect to the mean and, thus, the proposed model of Eq. (5.9) is consistent in both kinds of tissues.

The Gaussian hypothesis was tested by using the Kullback-Leibler divergence (D_{KL}). This measure takes lower values when the PDFs are similar. The results are shown in Fig. 5.1 for the case of Lumen. In the case of calcified plaque, the fitting is similar and we omit it. The results obtained for the D_{KL} are shown in table 5.3 where the better performance of the AS hypothesis becomes evident when compared to the Gaussian hypothesis.

The influence of the specular component was tested by means of a t-test

	Lumen		Calcified Plaque	
	X_r	X_i	X_r	X_i
Gaussian	0.4392	0.3727	0.2221	0.2170
α -stable	0.0191	0.0237	0.0476	0.0445

Table 5.3: Kullback-Leibler divergence for each signal component.

where the null hypothesis consists in considering a $\delta = \sqrt{\delta_r^2 + \delta_i^2}$ equal to zero, where δ_r and δ_i are location parameters for both components (this parameter is the deviation from the null specular component of the amplitude signal and, thus, is equivalent to the amplitude of the specular component in the resolution cell).

Results obtained rejected the null hypothesis with p-values $2.21 \cdot 10^{-6}$ in the case of Lumen and $3.19 \cdot 10^{-6}$ in the case of calcified plaque. This result evidences that the specular component must be considered and, concretely, in the case of calcified plaque doubles that one of lumen.

Finally, the proposed model was fitted to the PDF envelope of the signal as well as the HT-Rayleigh and the Nakagami. They were compared by using the D_{KL} measure. Results are shown in Fig. 5.2 where the better results of the proposed method can be appreciated.

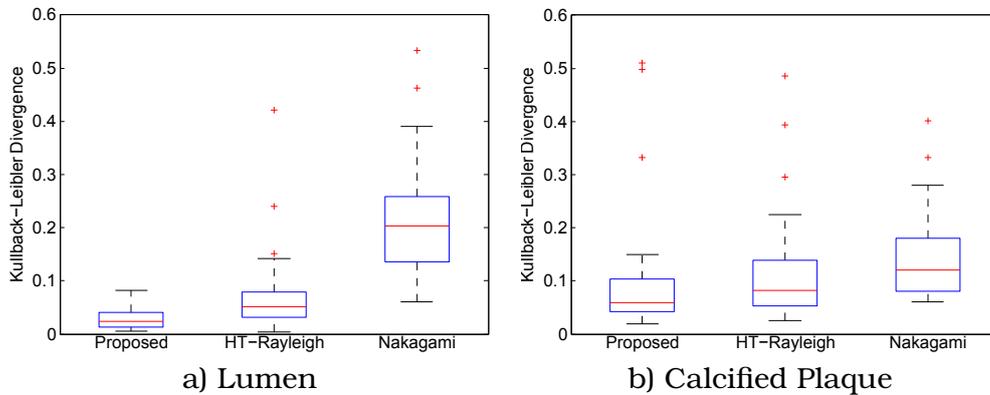


Figure 5.2: Kullback-Leibler for amplitude signal fitted with HT-Rayleigh, Nakagami and the proposed method.

5.5 Conclusion

In this work we present a new distribution which models the fully resolved speckle based on the envelop of AS distributions with specular component. This model is the generalization of classical models and the HT-Rayleigh. Results demonstrate that the proposed model is consistent with real data and improves the goodness-of-fit in comparison to models which make use of the Gaussian hypothesis and do not consider the specular component.

In our opinion, this work proposes a very versatile family of distributions

which accurately model the probabilistic behavior of speckle in US images and, due to its capability to describe tissue characteristics, may be used in potential applications such as filtering, segmentation or classification.

Part II

Applications

Gamma Mixture Classifier for Plaque Detection in Intravascular Ultrasonic Images

Evolution is a change from an indefinite incoherent homogeneity to a definite coherent heterogeneity.
Herbert Spencer, 1820–1903

Abstract– Carotid and coronary vascular accidents are mostly caused by vulnerable plaques. Detection and characterization of vulnerable plaques are important for early disease diagnosis and treatment. For this purpose, the echo-morphology and composition have been studied. Several distributions have been used to describe ultrasonic data depending on tissues, acquisition conditions and equipment. Among them, The Rayleigh distribution is the one-parameter model used to describe the raw envelope RF Ultrasound signal for its simplicity, whereas the Nakagami distribution (a generalization of the Rayleigh distribution) is the two-parameter model which is commonly accepted. However, it fails to describe B-Mode images or Cartesian interpolated or sub-sampled RF images because linear filtering changes the statistics of the signal.

In this work a *Gamma Mixture Model* (GMM) is proposed to describe the sub-sampled/interpolated RF images and it is shown that the parameters and coefficients of the mixture are useful descriptors of speckle pattern for different types of plaque tissues. This new model outperforms recently proposed probabilistic and textural methods with respect to plaque description and characterization of echogenic contents. Classification results provide an overall accuracy of 86.56% for four classes and 95.16% for three classes. These results evidence the classifier usefulness for plaque characterization. Additionally, the classifier provides probability maps according to each tissue type, which can be displayed for inspecting local tissue composition, or used for automatic filtering and segmentation.

Adapted from: G. Vegas-Sánchez-Ferrero¹, J. Seabra³, O. Rodríguez-Leor⁴, A. Serrano-Vida⁵, S. Aja-Fernández¹, C. Palencia², M. Martín-Fernández¹, J. Sanches³, *Gamma Mixture Classifier for Plaque Detection in Intravascular Ultrasonic Images*, Submitted.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Departamento de Matemática Aplicada, Univ. Valladolid.

³Institute for Systems and Robotics, Bioengineering Department, Univ. of Lisbon.

⁴Hospital Universitario Germans Trias i Pujol, Badalona.

⁵Hospital General de Granollers.

6.1 Introduction

Vulnerable plaque consists of a collection of blood cells and cholesterol in the wall artery which is prone to cause cardiovascular problems such as heart attack and brain stroke [Spa04, Esc06]. Generally, these lesions feature a thin fibrous cap over a soft lipid pool. The presence of these structures in the arterial wall causes a high mechanical stress in the arterial wall and, eventually, leads to rupture of the vulnerable plaque and clots. Stenosis is also one of the problems derived from the presence of plaques.

Among the imaging techniques commonly used for detecting atherosclerotic plaques in the coronary arteries, the invasive coronary angiography has been considered as the standard. However, most patients with acute coronary syndromes have minimal or mild coronary lumen obstruction detected by angiography and this modality has proved to have a limited ability to accurately measure the degree of stenosis and to characterize plaque morphology [Esc06]. Given these limitations, the importance of detecting stenosis areas and the presence of different kinds of plaque becomes evident. Instead of angiography, Intravascular Ultrasonography (IVUS) has demonstrated to provide clear visualization of arterial wall inner morphology and turns out to be a convenient alternative method for assessing the severity of morphology lesions [Esc06].

The technical procedure of acquiring IVUS data consists of introducing a catheter similar to the standard catheters employed in coronary angioplasty. The catheter is inserted inside the artery and moved until it reaches the artery segment to be studied. A rotating piezoelectric transducer transmits acoustic pulses and collects the A-lines that correspond to the reflected echoes along the depth, ρ , for each direction θ . The result is a polar representation of a 360° cross-sectional view. This image is interpolated and geometrically arranged to build the Cartesian image. In Fig. 6.1 an example of an IVUS image in polar and Cartesian coordinates is depicted.

The acoustic response of different kinds of plaque is qualitatively known: Lipidic plaque presents low echolucent response; Fibrous plaque presents intermediate level echogenicity; Calcified plaque is hyperechogenic and usually presents an acoustic shadow due to the series of echoes created by

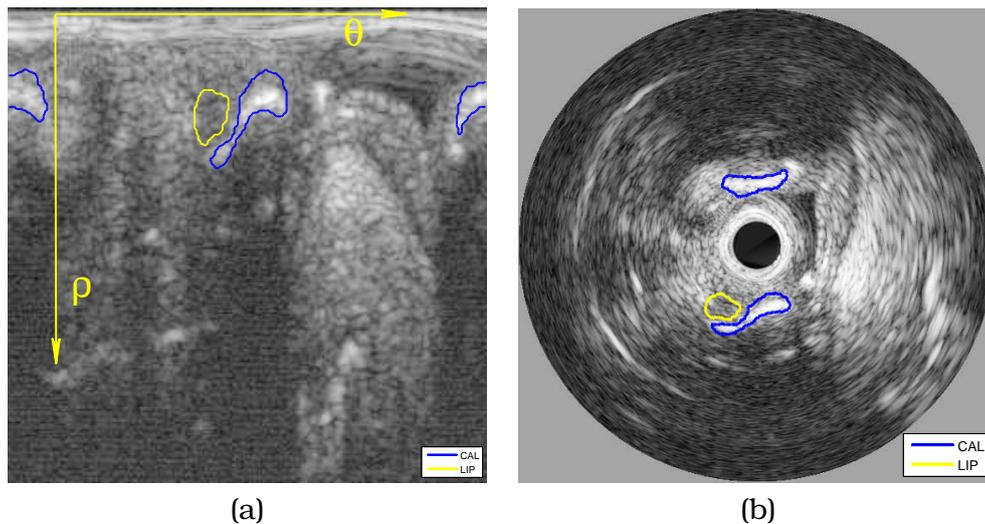


Figure 6.1: IVUS image in polar coordinates (a) and Cartesian coordinates (b) with the presence of lipidic (LIP) and calcified (CAL) plaques that were histologically identified.

multiple reflections within a small but highly reflective tissue [Esc06, Wil98, EB95, EB96].

Although this qualitative characterization of the plaques offers an intuitive interpretation of IVUS images, an important effort has been done to understand the echo-morphology and pathological evolution [Mat00]. Quantitative characterization of plaques allows developing or refining methods for plaque detection, risk predictions and potentially suggesting different therapies.

In order to obtain a quantitative characterization of the ultrasonic response of plaques, the physics of ultrasound imaging should be taken into account. Basically, the process of image formation in medical ultrasound begins with a pulse packet emission which travels along the beam vector axis and changes shape according to characteristics of the media.

The traveling pulse is scattered by objects placed at different scattering depths and cause delays in the pulse. The backscattered (received) signal is corrupted by a characteristic granular pattern noise called *speckle* which depends on the number of scatterers per resolution cell as well as their size [VSF10b, Sea10b]. This type of multiplicative noise, in the sense that its variance depends on the underlying signal, is observed in other modalities using coherent radiation such as LASER [Goo75] and Synthetic-aperture Radar (SAR) [Cop03].

Speckle mainly depends on the micro-structure of the tissues and thus its statistics can be used as tissue histological descriptors [Thi03]. These statistics strongly depend on the effective scatter density, that is, on the effective number and intensity distribution of the scatterers in each resolution cell, their size, their shape, their spatial organization as well as the acquisition instrumentation and the tissue attenuation [Sha93, Sha00, Des10]. The resulting *speckle* noise can be grouped in the following main classes:

- Fully developed: infinite effective number of scatterers per resolution cell and no deterministic component, modeled by Rayleigh distribution in the case of one-parameter distribution [Wag87, Ray80, Wag83, Bur78].
- Possibly fully resolved: infinite effective number of scatterers per resolution cell and existence or not of a deterministic component, modeled by Rice distribution [Ric45, Nak40, Tut88].
- Possibly partially developed: arbitrary effective number of scatterers per resolution cell and no deterministic component, modeled by K distribution [Jak84, Jak80, Sha93].
- Possibly partially resolved: arbitrary effective number of scatterers per resolution cell and existence or not of a deterministic component, modeled by K-Homodyne distribution [Dut94, Jak87].

Fully developed speckle is the most common model for speckle formation. It considers a tissue or region composed of a large number of scatterers, acting as echo diffusers. These scatterers arise from inhomogeneity and structures approximately equal to or smaller in size than the wavelength of the ultrasound, such as tissue parenchyma, where there are changes in acoustic impedance on a microscopic level within the tissue. Under this condition, pixel intensities in envelope data were usually modeled by Rayleigh probability density functions (PDFs) [VSF10b, Bur78, Elt06].

Note that the most general case of speckle is the Possibly partially resolved, which is modeled by a K-Homodyne distribution. The K-distribution is a special case of the K-Homodyne distribution with no deterministic component. The Rice distribution is the limiting case corresponding to an infinite effective density. The Rayleigh distribution is a special case of the Rice distribution with no deterministic component.

Other distributions have been proposed for speckle characterization. Probably, the most noticeable distribution is the Nakagami proposed in [Nak60]. This distribution has two parameters and can be considered as a generalization of the Rayleigh distribution. In [Sha00], a model based on Nakagami distributions is proposed for the characterization of backscattered echo. This model is motivated from the fact that the Nakagami distribution generalizes the Rayleigh distribution and also appears to be similar to Rician distribution, which is also a generalization of the Rayleigh (see [Sha00]). This is the reason that makes the Nakagami distribution the commonly accepted distribution for developed speckle and it is also considered as the two-parameter approximation of the true distribution for all the cases (without log-compression or application of filters) [Des11, Sha00, Des10, Sah11].

The distribution of speckle depends not only on the tissues but on the acquisition process and the post-processing. The transducer center frequency also affects to the distributions. Note that, as the central frequency

increases, the size of the range cell shrinks and, thus, the number of scatterers in the range cell go down and one must expect non-Rayleigh statistics. This can be seen as an additional reason for the better fitting of other two-parameter distributions such as the Gamma distribution. Additionally the post-processing techniques as log-compression and filtering also affects to the probability distributions of speckle.

Plaque echo-morphology is the contribution result of different tissue types (components). The lipidic plaque usually presents a fibrotic cap which has different acoustic response and thus different distributions [Esc06]. Additionally, accumulation of blood cells (macrophages) within plaques may change their probabilistic models. Hence, a mixture model becomes an opportune strategy for statistically describing the echo-morphology of the plaque.

The Rayleigh mixture model (RMM) was first proposed in [Sea11, Sea10b] for plaque characterization and classification. In that work, a RMM was obtained by Expectation-Maximization method [Dem77, Moo96]. Three kinds of plaque were considered in that study: Fibrotic, Lipidic and Calcified. The RMM parameters were estimated for each kind of plaque and were used, in combination to other textural features, to provide a descriptor of plaque composition. On the other hand, a Nakagami mixture model (NMM) was proposed in [Des11] for segmentation of arteries. This approach uses the Nakagami as a generalization of Rayleigh distribution as a good candidate to characterize the speckle.

The Rayleigh has been an accepted assumption for fully developed speckle [Sha00, Elt06, Tao06, Nil08]. Its generalization by means of the Nakagami distribution has been used as an approximate general model for the echo envelope [Sha00]. However, in the presence of down-sampling with interpolation, the reported tests indicate that Rayleigh or Nakagami models do not fit as well the data as the Gamma distribution.

In [Tao06, Nil08], many distributions were empirically fitted to real data and it is shown that speckle is better described by the Gamma distribution. In those works generic B-mode images having undergone log-compression and filtering were considered [Clo11], though no theoretical justifications were provided for this better fitting.

In [VSF10b] the Gamma performance was compared to other distributions when interpolated fully developed speckle was considered. Experimental tests have shown the superiority of the Gamma distribution over the Rayleigh and Nakagami for describing US data - 85% of the fully developed speckle areas passed the χ^2 test when a Gamma distribution was fitted, whereas 70% and less than 10% passed in the Nakagami and Rayleigh cases respectively.

The interpolation operation performed in the A-lines of the raw RF signal to re-sample the data and equalize the resolution in both dimensions, angle and depth seems to be the key element to explain why Gamma describes better the data than the Rayleigh or Nakagami distributions. The

interpolation process can be formulated as linear filter that linearly combines different pixels that are Rayleigh distributed. As shown in [VSF10b], a linear combination of Rayleigh random variables can be accurately fitted by Gamma distributions.

Note that the interpolation process consists of a weighted sum of values that, in the case of Rayleigh distributed data, results in a different random variable. Hence, not only interpolation processes but every linear filter applied to a Rayleigh distributed data is a weighted sum of Rayleigh random variables, which is better described by a Gamma random variable than a Rayleigh.

In [Sha03] the authors presented a bimodal gamma distribution with five parameters to model the statistics of the pixels in the gray-level (B-mode) images. The parameters of the distribution were evaluated for regions containing plaque using curve-fitting techniques. In that work just two gamma distributions were used due to the limitations of curve fitting techniques applied. The model showed good fitting properties for “hard” (calcified) plaque and soft plaque (“lipids, cellular components, loose connective tissue”).

In [Ata11], some mixtures of Gamma distributions are also suggested, this time for wireless channels, where they show a better fitting results than Rayleigh or Nakagami distributions.

A common stage of the acquisition process of US images is to down-sample the acquired signal in order to provide an isotropic image resolution. This re-sampling stage usually involves an interpolation stage where linear filtering is applied. In these conditions, the results obtained in [VSF10b, Elt06, Tao06] still hold and the Gamma distribution better describe US RF envelope down-sampled data than the Rayleigh or Nakagami distributions.

The aim of this work is threefold: i) Propose a *Gamma Mixture Model* (GMM) to describe the interpolated/re-sampled RF envelope US data; ii) Based on the parameters and coefficients of this mixture, design and train a classifier to discriminate Calcified, Lipidic, Fibrotic and Lumen regions within atherosclerotic plaques. iii) Provide probability maps which can be of help for physicians or for automatic post-processing techniques such as filtering or segmenting methods.

GMM and a method based on RMM [Sea11] are compared in terms of goodness of fit and classification accuracy. Comparison results showed that the GMM outperforms the RMM in terms of goodness of fit as well as accuracy. Besides the approaches by probabilistic speckle characterization, some methods have been proposed based on textural analysis. These methods usually consider Autoregressive Models, spectral features or wavelet coefficients [Tak10, Cio10, Sea11, Nai02]. The proposed classifier is also compared with a recently published method based on textural features [Cio10]. Results showed that the proposed method outperforms both the textural-based and the RMM classifier in 5% and 22.7% of accuracy respectively.

The rest of the paper is structured as follows: In Section 6.2 we describe

the dataset, the acquisition protocols and histological validation for plaques. In Section 6.3, we test the capability of the most commonly used distributions for describing the probabilistic behavior of speckle (i.e. Rayleigh, Nakagami, Gamma). In Section 6.4, the Gamma Mixture Model is proposed. Section 6.5 is devoted to the GMM classifier, where the mixture of GMMs is proposed and derived. The experiments are presented in Section 6.6; first, the optimal number of component for each tissue class is analyzed; second, the proposed classification method is compared with those of [Sea11] and [Cio10]; third, the statistical significance of the classification results is tested by means of the Friedman test and Bonferroni-Dunn test; fourth, probability maps are presented. Finally, we conclude in Section 6.7.

6.2 Materials

Having a golden standard to test the performance of classifiers is of great importance, however the only way to assess the nature of plaques and its location is by histological validation. In this work, the classification is tested with the golden standard obtained with the methodology recently presented in [Cio10]. The IVUS dataset consists of 9 post-mortem coronary arteries obtained from 9 different patients. All the patients died for accidental death, so there was not indication about cardiac problems. The relatives of the deceased gave the consent to use their arteries. The samples of the coronary arteries were obtained in the University Hospital “Germans Trias i Pujol” (Badalona, Spain) with the approval of the its ethical committee. The analysis of the post-mortem arteries was performed at the “hemodynamic” department of the same hospital, under the supervision of Dr. Josepa Mauri and Dr. Oriol Rodriguez-Leor.

From these arteries, 50 different images with the presence of plaques of different nature were selected. Then, the arteries are sliced in order to characterize plaques by histological analysis. The histological analysis was performed in the General Hospital of Granollers (Spain) by the pathologist Dr. Angel Serrano-Vida.

The acquisition of the images was performed in the following way: The artery is separated from the heart fixed in a mid-soft plane and filled (using a catheter) with physiological saline solution at constant pressure (around 120 mmHg), simulating blood pressure. References of distal, proximal, left and right positions are marked. The probe is introduced through the catheter and RF data are acquired in correspondence of plaques.

Real-time Radio-Frequency (RF) data acquisition was performed with the Galaxy II IVUS Imaging System (BostonScientific) using a catheter Atlantis SR Pro 40MHz (Boston Scientific). To collect and store the RF data, the imaging system has been connected to a workstation equipped with a 12-bit Acquiris acquisition card with a sampling rate of 200MHz.

The RF data for each frame is arranged in a matrix of $N \times M$ samples,

where $M = 1024$ is the number of samples per A-line, and $N = 256$ is the number of positions assumed by the rotational ultrasound probe.

The histological validation of plaques comprises the following steps: Vessels are cut in correspondence with previously marked positions and plaque composition is determined by histological analysis. A correspondence between detected plaques by histology and respective IVUS image is established by means of reference positions set by an expert interventionist (Dr. Oriol Rodriguez-Leor) in cooperation with the pathologist (Dr. Angel Serrano-Vida).

With the purpose of preserving a reliable correspondence between histological tissue and regions of the IVUS image, the medical team manually performs the plaque labeling task discarding pairs of images in which a correspondence can not be obtained.

Finally, the dataset comprises 50 different images obtained from 9 different arteries (patients). All of them present a segmentation of the Lumen (50 Lumen regions, one per image), a set of 69 plaques was identified in the images and histologically characterized in the following types: 30 Calcified, 14 Lipidic, 25 fibrotic. Table 6.1 shows the distribution of plaques among the patients (arteries) in the dataset.

Patient (Artery)	# Images	Lipidic	Fibrotic	Calcified	Lumen
1	8	5	0	5	8
2	6	2	0	6	6
3	8	0	7	3	8
4	2	2	1	2	2
5	5	0	3	5	5
6	2	2	0	0	2
7	3	2	2	1	3
8	10	1	6	8	10
9	6	0	6	0	6
TOTAL:	50	14	25	30	50

Table 6.1: Data set of 50 different images acquired from 9 different patients.

In order to provide comparable features between patients, we applied the following acquisition protocol: the IVUS images have been directly reconstructed from the raw RF signals rather than using the ones produced by the IVUS equipment. The image reconstruction algorithm used in this work is the one described in [Cio10] and it is shown in Fig. 6.2. The process comprises the following stages:

1. Time Gain Compensation, with $TGC(r) = 1 - e^{-\beta r}$ where $\beta = \log 10^{\alpha f/20}$, α parameters is the attenuation coefficient for biological soft tissues ($\alpha \approx 0.8\text{dB/MHz.cm}$ for $f = 40\text{MHz}$ [Cio10]), f is the central frequency of the transducer in MHz and r is the radial distance from the catheter in cm.;

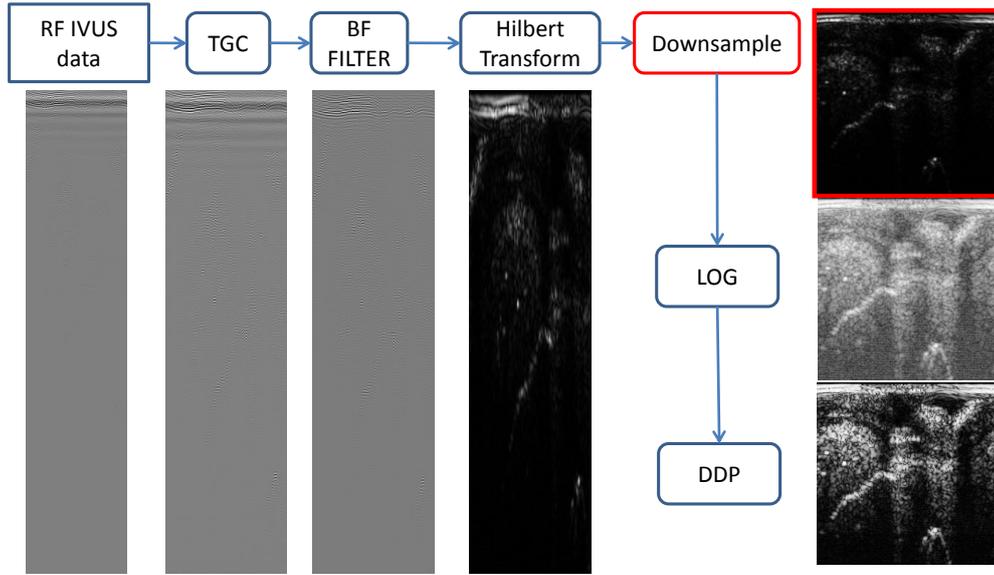


Figure 6.2: Image reconstruction process. A Time Gain Compensation (TGC) operation is applied to the RF IVUS data acquired. The Envelope is recovered by a Hilbert Transform. A down-sample stage is applied to obtain isotropic resolution, all the analysis of this work is applied after this stage where the Gamma assumption is applied. Log-compression and Digital Development Process (DDP) stages are usually applied for visualization.

2. Butterworth Band pass filter with cut frequencies $f_L = 20\text{MHz}$ and $f_u = 60\text{MHz}$.
3. Envelope recovering with Hilbert Transform.
4. Down-sampling of the image in order to obtain isotropic resolution with linear filtering.
5. Log-compression.
6. Digital Development Process (DDP). A non-linear adjustment of gain and edge-emphasis process to enhance the tissue visualization.

After this reconstruction process, the IVUS displayed image can be easily obtained by interpolating polar coordinates into Cartesian coordinates, resulting in a non-compressed, 256×256 pixels image (cf. Fig. 6.1.a where the polar coordinate image is shown and Fig. 6.1.b where the interpolation into Cartesian coordinates is depicted).

The traditional displayed IVUS image is obtained from the polar representation (ρ, θ) by interpolating in a rectangular (Cartesian) grid, (i, j) . However, in this work, the image used for feature extraction and classification is the non-compressed polar one obtained after the Down-sampling step (c.f. Fig. 6.2). This stage of the reconstruction process involves linear filtering (in order to down-sample by factor without aliasing) and, thus, Rayleigh or Nakagami models do not fit as well the data.

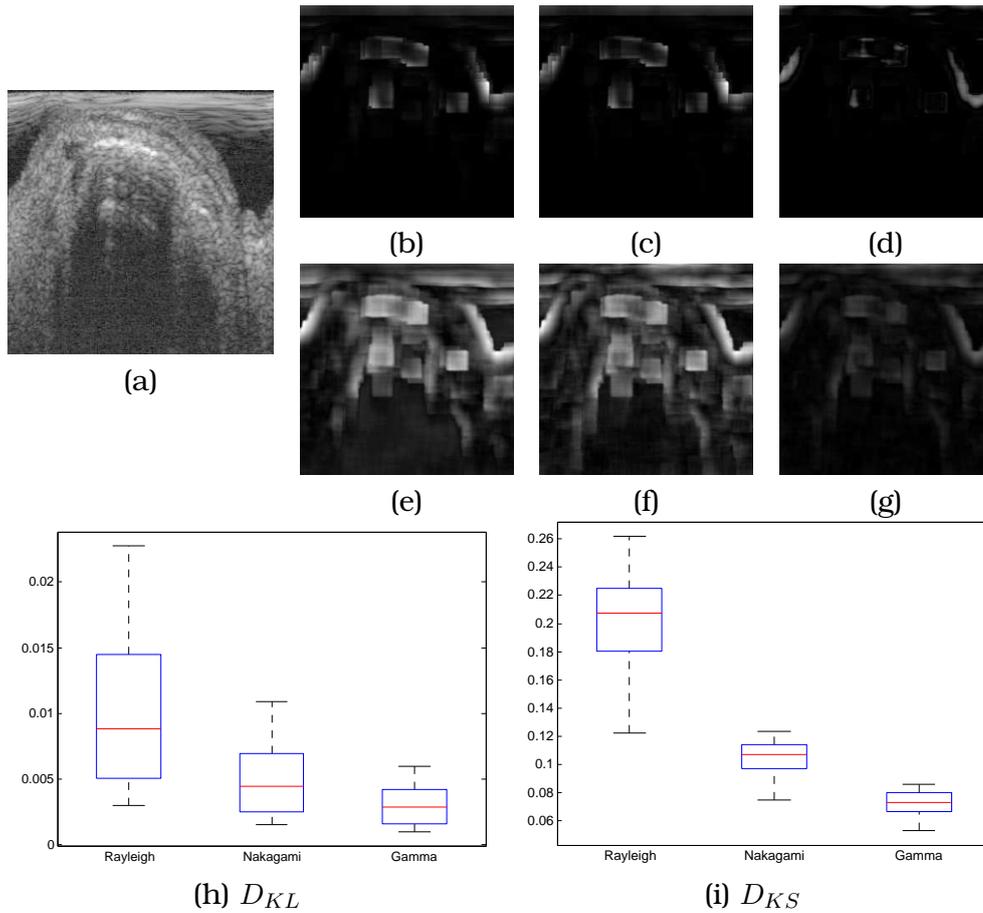


Figure 6.3: Log-compressed representation of the envelope (a). Kullback-Leibler divergence for Rayleigh (b), Nakagami (c) and Gamma (d) distributions. Uniform norm of the cumulative distribution function for Rayleigh (e), Nakagami (f) and Gamma distributions (g). Box-plots for both measures D_{KL} and D_{KS} are represented in (h) and (i) respectively. Welch t-test results show that populations are statistically different and thus the Gamma Distributions fits better than Rayleigh or Nakagami distributions.

6.3 Probabilistic model for envelope data

In this section the performance of Rayleigh, Nakagami and Gamma distributions are tested as candidates for describing the probabilistic behavior of speckle in the preprocessing stage of the envelope image formation process.

The performance test of both distributions is carried out after the down-sampling stage (see Fig. 6.2), where the Nakagami distribution is commonly accepted. All envelope images of the data set (50 in total) were tested with two different measures: Kullback-Leibler divergence and Uniform Norm of the cumulative distribution function (CDF). The former is a non-symmetric measure of the difference between two probability distributions defined as

$$D_{KL}(p_n, f_X) = \sum_{i=1}^N p_n(i) \log_2 \frac{p_n(i)}{f_X(i)}, \quad (6.1)$$

where p_n is the empirical PDF estimate and f_X is the theoretical distribution: Rayleigh, Gamma or Nakagami.

Note that the binary logarithm was chosen in this definition. We decided to use this base since the calculation of Kullback-Leibler divergence in other works like [Seal1] was provided in that base.

Instead of using the formulation of Eq. (6.1) we used the following symmetrized form

$$D_{KL} = \frac{1}{2}(D_{KL}(p_n, f_X) + D_{KL}(f_X, p_n)). \quad (6.2)$$

The measure D_{KL} provides a symmetric way to measure the similitude between both PDFs. Since both the $D_{KL}(p_n, f_X)$ and $D_{KL}(f_X, p_n)$ are non-negative this measure is more restrictive than that one of Eq. (6.1).

The empirical PDF was estimated by means of the histogram of the neighborhood of the pixel under study. The chosen neighborhood size is 11×11 the number of bins of the histogram is $n = 30$ equally spaced and smoothed with a Gaussian kernel (see [Har90] for more details). The smoothing process reduces the dependence of the PDF approximation on the number of bins used. Parameters of Rayleigh and Gamma PDFs correspond to the maximum likelihood estimates of the data in the neighborhood of the pixel studied, parameters of the Nakagami distribution were calculated as in [Des09].

The uniform norm of the cumulative distribution function, also called Kolmogorov-Smirnov (KS) statistic, is defined as

$$D_{KS} = \sup |\hat{F}_n(i) - F_X(i)|, \quad (6.3)$$

where \hat{F}_n is the empirical CDF of data and F_X is the CDF of the theoretical distributions.

This last metric was chosen since it does not depend on the number of bins used for the PDF estimate and can be calculated with a few number of samples. Additionally, the Glivenko-Cantelli theorem states that, if the samples are drawn from distribution F_X , then D_{KS} converges to 0 almost surely [Dud99].

As an example, in Fig. 6.3.a a down-sampled envelope image is depicted (it is log-compressed only for easing the visualization); Fig. 6.3.b-d show the Kullback-Leibler divergence, D_{KL} , computed for the Rayleigh, Nakagami and Gamma distributions respectively. Fig. 6.3.e-g show the D_{KS} measure for Rayleigh, Nakagami and Gamma distributions respectively. This example shows a better performance of Gamma distribution for both measures. The darker the image, the better the performance for both measures.

In order to provide a quantitative result of the performance of the Gamma distribution in contrast to the Rayleigh and Nakagami ones, a Welch t-test was performed for both measures, D_{KL} and D_{KS} , in the following way: for each of the images of the dataset, both measures were calculated in each neighborhood (with size 11×11). For each image, the average value was calculated and, thus, the Welch t-test is performed with 50 samples. This

test was chosen since no equal variance should be assumed. This test is performed considering pairs of distributions (Gamma vs. Rayleigh, Gamma vs. Nakagami, and Rayleigh vs. Nakagami). The null hypothesis considers that both population means are the same. The box-plot of both measures is depicted in Fig. 6.3.h-i.

P-values of the Welch t-test for the case of KL divergence and uniform norm of the CDF are shown in Table 6.2. Note that all values are negligible and the null hypothesis has to be rejected in all cases. Consequently, these three distributions fit in a different way, and the Gamma is the one with best fitting for both measures.

	Rayleigh	Nakagami	Gamma
Rayleigh	X	$7.54 \cdot 10^{-35}$	$7.9 \cdot 10^{-38}$
Nakagami	$7.6 \cdot 10^{-17}$	X	$3.67 \cdot 10^{-46}$
Gamma	$3.47 \cdot 10^{-17}$	$1.69 \cdot 10^{-16}$	X

Kullback-Leibler Divergence
 Kolmogorov-Smirnov Statistic

Table 6.2: P-values of the Welch t-test for the case of D_{KL} and D_{KL} .

This result is a strong confirmation that Gamma Distribution better describes the probabilistic nature of speckle when internal preprocessing such as linear filtering (to avoid aliasing in the down-sampling stage) are taking place and confirms the result obtained in [VSF10b, Elt06, Tao06]. Note that the Nakagami distribution was introduced as an approximate model for the echo envelope.

6.4 Gamma Mixture Model

In this section the Gamma Mixture Model is proposed and the method for the computation of its parameters and coefficients is described. The aim of using GMMs is that the echo-morphology may result from the contribution of different echogenic components of the plaque that follow different distributions. Under the assumption of gamma distributed speckle, the GMM arises in a natural way.

Let $\mathbf{X} = \{x_i\}$, $1 \leq i \leq N$ be a set of samples (pixel intensities) of a given region of the ultrasound image. These samples can be considered as independent and identically distributed (IID) random variables (RVs). This assumption is taken since the down-sampling stage reduces the possible correlation between neighboring pixels. The GMM considers that these variables result from the contributions of J distributions¹

$$p(x_i|\Theta) = \sum_{j=1}^J \pi_j f_X(x_i|\Theta_j), \quad (6.4)$$

¹The 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 $E\{\cdot\}$.

where Θ is a vector of the parameters of the GMM $(\pi_1, \dots, \pi_J, \Theta_1, \dots, \Theta_J)$ and Θ_j are the parameters of the PDF (in our case the parameters of a Gamma distribution are represented as α_j and β_j). The Gamma PDF is defined as

$$f_X(x|\alpha, \beta) = \frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-\frac{x}{\beta}}, \quad x \geq 0 \text{ and } \alpha, \beta > 0, \quad (6.5)$$

where $\Gamma(x)$ is the Euler Gamma function defined as $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$, for $x > 0$. The condition $\sum_{j=1}^J \pi_j = 1$ must hold to guarantee that $p(x_i|\Theta)$ is a well defined probability distribution.

The joint distribution of IID samples is given by

$$p(\mathbf{X}|\Theta) = \prod_{i=1}^N p(x_i|\Theta). \quad (6.6)$$

The Expectation-Maximization method is applied here to maximize the log-likelihood function when some hidden discrete random variables, $\mathbf{Z} = \{Z_i\}$, are introduced to the model. These RVs take values in $\{1, \dots, J\}$, their meaning is that the sample x_i belongs to the distribution class j when $Z_i = j$.

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

$$\mathcal{Q}(\Theta|\Theta^{(n)}, \mathbf{X}) = E_{\mathbf{Z}|\Theta^{(n)}, \mathbf{X}} \{\mathcal{L}(\Theta|\mathbf{X}, \mathbf{Z})\}. \quad (6.7)$$

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

The application of the EM algorithm for general distributions is not new, see for example [Moo96, Fig02]. In the case of a Gamma Mixture Model, it was firstly derived by Webb in [Web00]; another similar derivation was obtained in [Des09, Des11]. For the sake of completeness and to introduce the notation used to derive the classifier, the EM algorithm is explained in [Des11] and [Web00] which results in the following equality for the weights

$$\hat{\pi}_j = \frac{1}{N} \sum_{i=1}^N \gamma_{i,j} = \frac{1}{N} \sum_{i=1}^N p(Z_i = j|x_i, \Theta^{(n)}), \quad (6.8)$$

where $\gamma_{i,j} = p(Z_i = j|x_i, \Theta^{(n)})$ to make notation simpler, and it can be derived by the Bayes theorem as

$$\gamma_{i,j} = p(Z_i = j|x_i, \Theta^{(n)}) = \frac{p(x_i|\Theta_j^{(n)})p(Z_i = j|\Theta^{(n)})}{p(x_i|\Theta^{(n)})}, \quad (6.9)$$

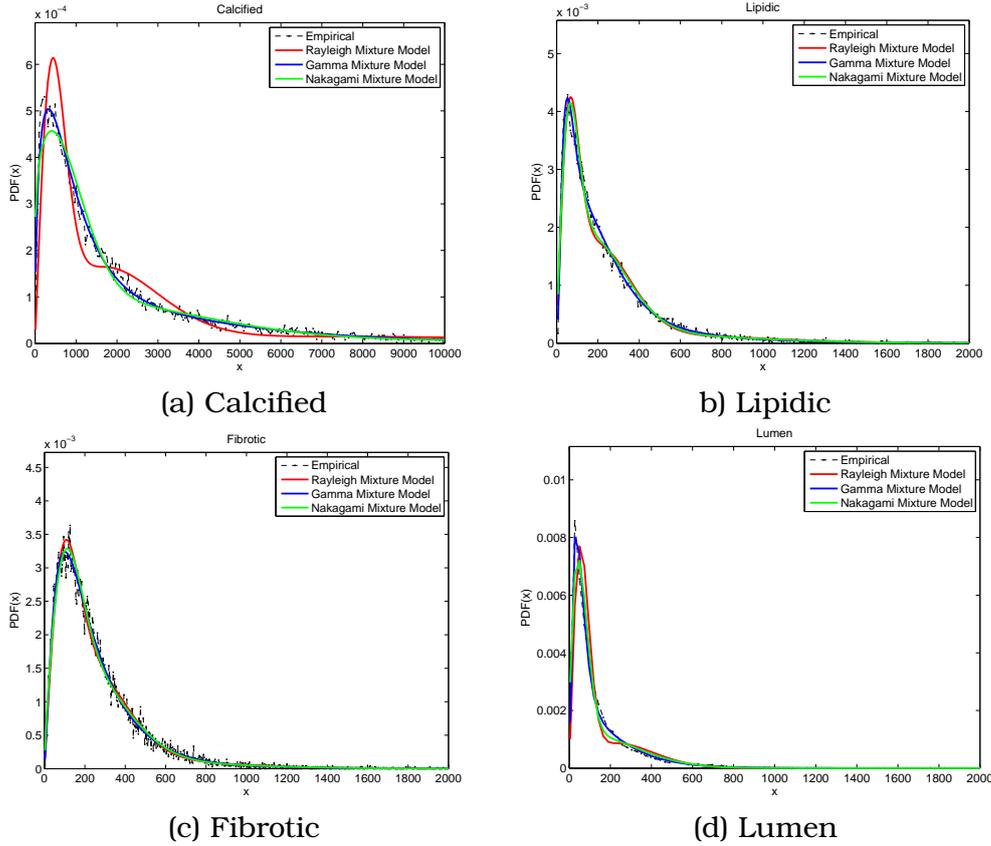


Figure 6.4: Gamma, Nakagami and Rayleigh mixture models fitting for different kinds of tissue. In both methods the tolerance was fixed to $\|\Theta^{(n+1)} - \Theta^{(n)}\| < 10^{-4}$ with a maximum number of iterations of 1000 and three components in each mixture.

where $p(Z_i = j | \Theta^{(n)}) = \pi_j^{(n)}$.

The estimates of the Gamma distribution parameters results in the following equalities (see [Des09, Web00])

$$\log(\hat{\alpha}_j) - \psi(\hat{\alpha}_j) = \log\left(\frac{\sum_i^N \gamma_{i,j} x_i}{\sum_i^N \gamma_{i,j}}\right) - \frac{\sum_i^N \gamma_{i,j} \log x_i}{\sum_i^N \gamma_{i,j}}. \quad (6.10)$$

Note that this is the case of uniform prior weights proposed in [Web00], which is a special case of the Dirichlet prior weights proposed in [Des09]

$$\hat{\beta}_j = \frac{1}{\hat{\alpha}_j} \frac{\sum_{i=1}^N \gamma_{i,j} x_i}{\sum_{i=1}^N \gamma_{i,j}}, \quad (6.11)$$

where $\psi(x)$ is the *Digamma* function defined as $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$.

The method is applied in the following way:

1. A first estimate of the hidden variables is obtained by means of any clustering method (k-means for example). For each cluster $j = \{1, \dots, J\}$, the parameters of the distribution $\Theta_j^{(0)}$ can be calculated by means of the moments method or Maximum likelihood methods. Weights, $\hat{\pi}_j^{(0)}$,

can be calculated either as $1/J$ or as the number of elements of each cluster divided by the total number of elements. Set $n = 0$.

2. Expectation Step. $\gamma_{i,j}$ is calculated from eq. (6.9).
3. Maximization Step. The estimate for $\hat{\alpha}_j$ is calculated from eq. (6.10). and $\hat{\beta}_j$ is calculated from eq. (6.11). $\hat{\pi}_j$ is calculated from eq. (6.8).
4. $n = n + 1$
5. Go to step 2 until $\|\Theta^{(n)} - \Theta^{(n-1)}\| < \text{TOL}$ is satisfied.

As an example, in Fig. 6.4, the GMM is used to fit different tissue types and compared to the output of the NMM proposed in [Des09] and the RMM proposed in [Sea10b]. Specifically, three mixture components were used in the fitting process for GMM, NMM and RMM. At first sight, the performance of the gamma mixture is better. Quantitative results calculated with D_{KL} and D_{KS} for all the data set are provided in Table 6.3. In this case, numerical results show that plaques, which can be composed by different echolucent sources or echogenic structures such as fibrous caps, are properly fitted with a gamma mixture model and, potentially, the parameters of the mixture model could be a good descriptor of the tissue class under the operations of down-sampling and interpolation. Nevertheless, an extended study of the behavior of the GMM, NMM and RMM is presented in section 6.6.

Tissue	D_{KL}			D_{KS}		
	GMM	NMM	RMM	GMM	NMM	RMM
Calcified	$4.37 \cdot 10^{-5}$	$1.06 \cdot 10^{-4}$	$2.65 \cdot 10^{-4}$	0.0080	0.0180	0.0359
Lipidic	$3.29 \cdot 10^{-3}$	$5.03 \cdot 10^{-3}$	$5.78 \cdot 10^{-3}$	0.0054	0.0114	0.0161
Fibrotic	$3.97 \cdot 10^{-3}$	$4.57 \cdot 10^{-3}$	$4.91 \cdot 10^{-3}$	0.0038	0.0078	0.0117
Lumen	$1.13 \cdot 10^{-2}$	$2.41 \cdot 10^{-2}$	$4.85 \cdot 10^{-2}$	0.0167	0.0341	0.0894

Table 6.3: Fitting of different types of tissues using GMM, NMM and RMM.

6.5 GMM Classifier

In this section a Classifier based on GMMs priors is derived. This classifier is inspired in the work of [Cop03], where a Bayesian gamma mixture model approach is proposed. There are some important differences between that method and the one here presented. The first one is that the method in [Cop03] adopted a Bayesian approach for deriving the model. This methodology considers the parameters of the GMMs as random variables and thus needs the definition of some prior distributions to characterize them. These distributions require their parameters to be manually initialized which clearly affects, and may bias, the mixture model.

Instead of considering prior distributions which can be inaccurate and may bias the results of the model, we apply an Expectation-Maximization approach without prior on the parameters. The initial values of the GMMs parameters are obtained from the training set according to each tissue class. The final labeling of the GMMs is then obtained according to the closest initial GMMs (with respect to the D_{KS}) for each tissue class and the probability of belonging to each tissue class is obtained by the Bayes Theorem.

6.5.1 Formulation

We state the problem as a mixture model of GMMs in the following way: Let $\mathbf{X} = \{x_i\}$, $1 \leq i \leq N$ be a set of samples (pixel intensities) of a given region of the ultrasound image. These samples are IID random variables. The GMM considers that these variables result from the contributions of J tissue classes and each PDF comprises R_j components for each $j \in \{1, 2, \dots, J\}$. This mixture model takes into account the presence of different echolucent responses of plaques

$$p(x_i|\Theta) = \sum_{j=1}^J \nu_j \sum_{r=1}^{R_j} \pi_{j,r} f_X(x_i|\Theta_{j,r}), \quad (6.12)$$

where Θ is a vector consisting of all the GMMs parameters, i.e. ν_j , $\pi_{j,r}$ and $\Theta_{j,r} = (\alpha_{j,r}, \beta_{j,r})$.

In eq. (6.12), two conditions must be imposed to the component weights in order to assure that eq. (6.12) is well defined as a true PDF

$$\sum_{j=1}^J \nu_j = 1 \text{ and } \sum_{r=1}^{R_j} \pi_{j,r} = 1, \quad \text{for each } j = \{1, \dots, J\}. \quad (6.13)$$

Now, the joint distribution, under the assumption of IID random variables is the product over all the samples as in eq. (6.6).

In order to apply the expectation-maximization method, two hidden discrete RVs vectors are introduced $\mathbf{Z} = \{Z_i\}$ and $\mathbf{W} = \{W_{i,j}\}$. The former, \mathbf{Z} , takes values onto the set of all possible tissue classes, i.e. $\{1, \dots, J\}$ and $Z_i = j$ means that sample x_i belongs to the tissue class j . The latter, $W_{i,j}$, takes values over $\{1, \dots, R_j\}$ and the meaning of $W_{i,j} = r$ is that sample x_i belongs to the r component of the j class tissue.

Both RVs are defined in such a way that there is an implicit relationship which makes sense from the point of view of conditional probability. For instance, $p(Z_i = j|\Theta)$ is the probability of belonging to tissue class j when the parameters of the mixtures are known and, in the mixture model of eq. (6.12), can be identified as ν_j . On the other hand, $p(W_{i,j} = r|Z_i = j, \Theta)$ is identified as $\pi_{j,r}$. The hierarchical relationship between Z_i and $W_{i,j}$ is shown in Fig. 6.5 where tissue classes are denoted as \mathcal{C}_j and each component of each tissue class is $\mathcal{B}_{j,r}$. Note that this relationship allows to consider the sample x_i as a contribution of each component within each

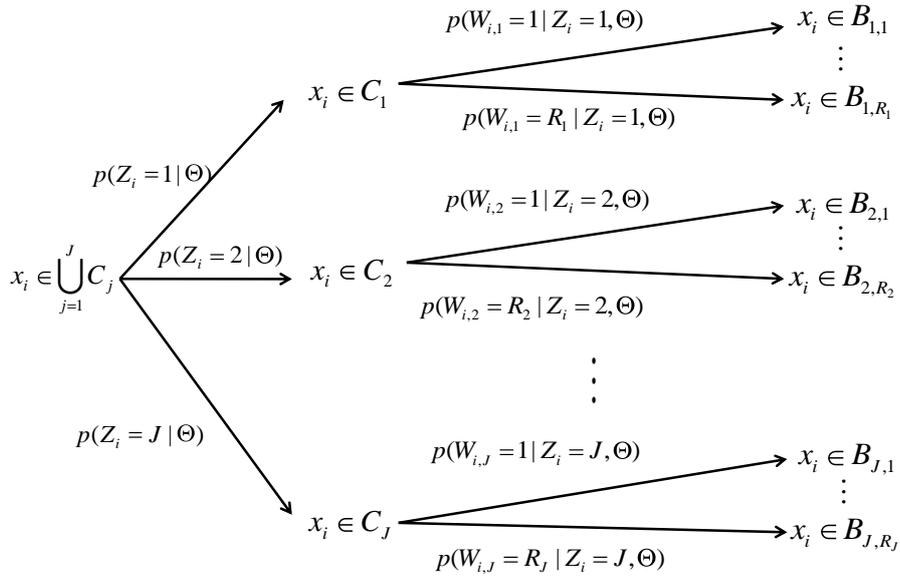


Figure 6.5: Conditional probability scheme. Each sample x_i may belong to tissue class, C_j with probability $p(Z_i = j | \Theta)$, and to the component $B_{j,r}$ with probability $p(W_{i,j} = r | Z_i = j, \Theta)p(Z_i = j | \Theta)$.

tissue class, which is a desired property in order to reflect the mixed composition of plaques. A detailed explanation on the EM method is presented in Appendix 6.A.

Supposing the previous estimate for the parameters of the mixture model ($\Theta^{(n)}$) is known, the log-likelihood function with both hidden variables is the following (see Appendix 6.A)

$$\mathcal{L}(\Theta | \mathbf{X}, \mathbf{Z}, \mathbf{W}) = \sum_{i=1}^N \log f_X(x_i | \Theta_{z_i, w_{i,z_i}}) + \sum_{i=1}^N \log p(w_{i,z_i} | \Theta) p(z_i | \Theta). \quad (6.14)$$

In the expectation step, the expectation is calculated over \mathbf{Z} and \mathbf{W} for known data \mathbf{x} and a previous estimate $\Theta^{(n)}$

$$\mathcal{Q}(\Theta | \Theta^{(n)}, \mathbf{X}) = E_{\mathbf{Z}, \mathbf{W} | \Theta^{(n)}, \mathbf{X}} \{ \mathcal{L}(\Theta | \mathbf{X}, \mathbf{Z}, \mathbf{W}) \} = \sum_{i=1}^N \sum_{j=1}^J \sum_{r=1}^{R_j} \gamma_{i,j,r} (\log f_X(x_i | \alpha_{j,r}, \beta_{j,r}) + \log(\nu_j \pi_{j,r})), \quad (6.15)$$

where $\gamma_{i,j,r} = p(Z_i = j, W_{i,j} = r | x_i, \Theta^{(n)})$, which can be easily calculated by means of the Bayes theorem

$$\gamma_{i,j,r} = \frac{f_X(x_i | \Theta_{j,r}^{(n)}) \nu_j^{(n)} \pi_{j,r}^{(n)}}{\sum_{k=1}^J \sum_{s=1}^{R_j} f_X(x_i | \Theta_{k,s}^{(n)}) \nu_k^{(n)} \pi_{k,s}^{(n)}}. \quad (6.16)$$

The updated weights can be calculated by the following equations

$$\hat{\nu}_j = \frac{1}{N} \sum_{i=1}^N \sum_{r=1}^{R_j} \gamma_{i,j,r}, \quad \hat{\pi}_{j,r} = \frac{\sum_{i=1}^N \gamma_{i,j,r}}{\sum_{i=1}^N \sum_{s=1}^{R_j} \gamma_{i,j,s}}. \quad (6.17)$$

The updated parameters of the Gamma mixtures can be calculated with the following equations

$$\log(\hat{\alpha}_{j,r}) - \psi(\hat{\alpha}_{j,r}) = \log\left(\frac{\sum_{i=1}^N \gamma_{i,j,r} x_i}{\sum_{i=1}^N \gamma_{i,j,r}}\right) - \frac{\sum_{i=1}^N \gamma_{i,j,r} \log(x_i)}{\sum_{i=1}^N \gamma_{i,j,r}} \quad (6.18)$$

$$\hat{\beta}_{j,r} = \frac{1}{\hat{\alpha}_{j,r}} \frac{\sum_{i=1}^N \gamma_{i,j,r} x_i}{\sum_{i=1}^N \gamma_{i,j,r}} \quad (6.19)$$

6.5.2 Training and Classifying

The training process is performed by estimating the prior parameters, $\Theta_{j,r}^{(0)}$ for the training set by fitting the GMMs explained in section 6.4 for each tissue class. Then, the mixtures of GMMs are fitted to the test set, $\mathbf{y} = \{y_i\}$ by the method explained in section 6.5.1 in the following way

1. Set $\Theta_{j,r}^{(0)}$ the prior parameters obtained by fitting the GMMs for the training set and for each tissue class. Initial weights are calculated as $\hat{\pi}_{j,r}^{(0)} = \frac{1}{R_j}$ and $\hat{\nu}_j^{(0)} = \frac{1}{j}$. Set $n = 0$.
2. Expectation Step. $\gamma_{i,j,r}$ is calculated from eq. (6.16).
3. Maximization Step. $\hat{\beta}_{j,r}$ and $\hat{\alpha}_{i,j}$ are calculated from eq. (6.19) and eq. (6.18). $\hat{\nu}_j$ and $\hat{\pi}_{j,r}$ are calculated from eq. (6.17).
4. $n = n + 1$.
5. Go to step 2 until $\|\Theta^{(n)} - \Theta^{(n-1)}\| < \text{TOL}$ is satisfied.

The GMMs fitting can have more than one solution [Tit92]. Actually, the parameters of the GMMs can be interchanged without affecting the fit with the data. To circumvent this problem, we compare by D_{KS} with respect the initial one for each tissue class. Finally, the labels are assigned by considering the lower D_{KS} between the final distributions and the initial ones.

The posterior probability of belonging to each tissue class is then obtained by simply calculating the probability of $Z_i = j$ in the test set

$$p(\mathbf{Z} = j | \Theta, \mathbf{y}) = \prod_{i=1}^N \frac{\sum_{r=1}^{R_j} \pi_{j,r} f_X(y_i | \Theta_{j,r}) \nu_j}{\sum_{k=1}^J \sum_{r=1}^{R_k} \pi_{k,r} f_X(y_i | \Theta_{k,r}) \nu_k}. \quad (6.20)$$

Finally, the selected class is that one with the highest posterior probability.

6.6 Experimental Results

In this section, two experiments are described. The first one is related to the selection of the number of elements of the GMM which better fits to each tissue class. This experiment is also performed with RMM and NMM in order to see which one of the Gamma, Nakagami or Rayleigh distribution better fits the ultrasound data under the operations of down-sampling with interpolation.

The second experiment aims at classifying plaques into four different tissue classes: Calcified, Lipidic, Fibrotic and Lumen.

At the end of these tests, some examples of probability maps for each tissue class are shown. We believe that these maps provide useful information for manual analysis of plaque composition and we foresee its usage as a pre-processing step for automatic segmentation.

6.6.1 Optimal number of components for each Class

Here, the GMM, NMM and RMM have been applied to every pre-segmented plaque using an increasing number of components. We have used D_{KL} and D_{KS} to investigate the performance of GMM, NMM and RMM with respect to the number of components for each class. Results are shown in Table 6.4.

Note that results of both measures evidence a better fit for the GMM for every tissue class. In all cases (GMM, RMM and NMM) the performance is better as the number of components increases but the GMM approach still remains as a better approach in nearly all the cases.

These results can be used in order to select the number of components of the mixture models. Note that, as the number of components increases, the measure decreases with a slower rate. In Fig. 6.6 the rate of decreasing of both measures for GMM, RMM and NMM are depicted for each tissue class. This rate is calculated as $f_{rate}(n) = D(n-1) - D(n)$ where $D(n)$ is D_{KL} or D_{KS} measures and n is the number of components of the mixture model. This criterion takes into account the lowest number of components per tissue class for which the goodness of fit does not significantly improve when another component is added to the mixture model. Note that the global maximum is obtained for $n = 3$. From that value, the differences are lower than that observed for 3 components.

Other selection criterion could be applied. We also applied the Bayesian Information Criterion (BIC) [Sch78] to determine the preferred number of components for each tissue class. This criterion provides the preferred model for a finite set of models, which is the case. It is based on calculating the log-likelihood and penalizing distributions with a higher number of free parameters to avoid over-fitting.

The number of free parameters used for the RMM is $2 \times \#components - 1$,

2 components		D_{KL}			D_{KS}		
Tissue	GMM	NMM	RMM	GMM	NMM	RMM	
Calcified	$6.22 \cdot 10^{-5}$	$2.26 \cdot 10^{-4}$	$9.23 \cdot 10^{-4}$	0.0126	0.0331	0.1087	
Lipidic	$6.32 \cdot 10^{-3}$	$1.26 \cdot 10^{-2}$	$3.36 \cdot 10^{-2}$	0.0198	0.0410	0.1059	
Fibrotic	$4.26 \cdot 10^{-3}$	$7.58 \cdot 10^{-3}$	$1.21 \cdot 10^{-2}$	0.0085	0.0257	0.0445	
Lumen	$1.73 \cdot 10^{-2}$	$5.71 \cdot 10^{-2}$	$3.94 \cdot 10^{-1}$	0.0277	0.0768	0.4110	

3 components		D_{KL}			D_{KS}		
Tissue	GMM	NMM	RMM	GMM	NMM	RMM	
Calcified	$4.37 \cdot 10^{-5}$	$1.06 \cdot 10^{-4}$	$2.65 \cdot 10^{-4}$	0.0080	0.0180	0.0359	
Lipidic	$3.29 \cdot 10^{-3}$	$5.03 \cdot 10^{-3}$	$5.78 \cdot 10^{-3}$	0.0054	0.0114	0.0161	
Fibrotic	$3.97 \cdot 10^{-3}$	$4.57 \cdot 10^{-3}$	$4.91 \cdot 10^{-3}$	0.0038	0.0078	0.0117	
Lumen	$1.13 \cdot 10^{-2}$	$2.41 \cdot 10^{-2}$	$4.85 \cdot 10^{-2}$	0.0167	0.0341	0.0894	

4 components		D_{KL}			D_{KS}		
Tissue	GMM	NMM	RMM	GMM	NMM	RMM	
Calcified	$4.24 \cdot 10^{-5}$	$4.46 \cdot 10^{-5}$	$9.89 \cdot 10^{-5}$	0.0079	0.0072	0.0141	
Lipidic	$3.29 \cdot 10^{-3}$	$3.98 \cdot 10^{-3}$	$4.26 \cdot 10^{-3}$	0.0055	0.0076	0.0083	
Fibrotic	$3.96 \cdot 10^{-3}$	$4.23 \cdot 10^{-3}$	$4.07 \cdot 10^{-3}$	0.0032	0.0062	0.0036	
Lumen	$6.44 \cdot 10^{-3}$	$2.39 \cdot 10^{-2}$	$1.18 \cdot 10^{-2}$	0.0043	0.0339	0.0126	

5 components		D_{KL}			D_{KS}		
Tissue	GMM	NMM	RMM	GMM	NMM	RMM	
Calcified	$3.08 \cdot 10^{-5}$	$4.36 \cdot 10^{-5}$	$9.89 \cdot 10^{-5}$	0.0042	0.0071	0.0141	
Lipidic	$3.09 \cdot 10^{-3}$	$3.79 \cdot 10^{-3}$	$3.88 \cdot 10^{-3}$	0.0044	0.0070	0.0060	
Fibrotic	$3.95 \cdot 10^{-3}$	$4.23 \cdot 10^{-3}$	$4.03 \cdot 10^{-3}$	0.0029	0.0062	0.0033	
Lumen	$6.44 \cdot 10^{-3}$	$9.76 \cdot 10^{-3}$	$1.12 \cdot 10^{-2}$	0.0044	0.0117	0.0110	

6 components		D_{KL}			D_{KS}		
Tissue	GMM	NMM	RMM	GMM	NMM	RMM	
Calcified	$3.07 \cdot 10^{-5}$	$4.25 \cdot 10^{-5}$	$4.77 \cdot 10^{-5}$	0.0041	0.0068	0.0036	
Lipidic	$2.94 \cdot 10^{-3}$	$3.39 \cdot 10^{-3}$	$3.68 \cdot 10^{-3}$	0.0029	0.0057	0.0065	
Fibrotic	$3.95 \cdot 10^{-3}$	$4.13 \cdot 10^{-3}$	$4.01 \cdot 10^{-3}$	0.0031	0.0054	0.0031	
Lumen	$6.18 \cdot 10^{-3}$	$9.39 \cdot 10^{-3}$	$8.90 \cdot 10^{-3}$	0.0031	0.0110	0.0118	

7 components		D_{KL}			D_{KS}		
Tissue	GMM	NMM	RMM	GMM	NMM	RMM	
Calcified	$3.08 \cdot 10^{-5}$	$4.30 \cdot 10^{-5}$	$4.77 \cdot 10^{-5}$	0.0041	0.0069	0.0036	
Lipidic	$2.93 \cdot 10^{-3}$	$3.40 \cdot 10^{-3}$	$3.66 \cdot 10^{-3}$	0.0026	0.0059	0.0066	
Fibrotic	$3.94 \cdot 10^{-3}$	$4.07 \cdot 10^{-3}$	$4.00 \cdot 10^{-3}$	0.0024	0.0047	0.0029	
Lumen	$6.18 \cdot 10^{-3}$	$6.83 \cdot 10^{-3}$	$8.90 \cdot 10^{-3}$	0.0031	0.0052	0.0118	

Table 6.4: Gamma mixture model, Nakagami mixture model and Rayleigh mixture model fitting for different kinds of tissue.

since each Rayleigh has one parameter and the number of free weights is $\#components - 1$ because the sum of weights must be 1. In the case of GMM and NMM, it is $3 \times \#components - 1$ for the same reason (both Gamma and Nakagami have 2 parameters).

The BIC criterion is obtained by means of the following formula

$$BIC = -2 \log(L) + k \log(n), \quad (6.21)$$

Calcified	GMM	NMM	RMM
Components	5	5	6
Free Parameters	14	14	11
D_{KL}	$3.08 \cdot 10^{-5}$	$4.36 \cdot 10^{-5}$	$4.77 \cdot 10^{-5}$
Lipidic	GMM	NMM	RMM
Components	3	6	6
Free Parameters	8	17	11
D_{KL}	$3.29 \cdot 10^{-3}$	$3.39 \cdot 10^{-3}$	$3.68 \cdot 10^{-3}$
Fibrotic	GMM	NMM	RMM
Components	3	4	4
Free Parameters	8	11	7
D_{KL}	$3.97 \cdot 10^{-3}$	$4.23 \cdot 10^{-3}$	$4.07 \cdot 10^{-3}$
Lumen	GMM	NMM	RMM
Components	4	6	7
Free Parameters	11	17	13
D_{KL}	$6.44 \cdot 10^{-3}$	$9.39 \cdot 10^{-3}$	$8.9 \cdot 10^{-3}$

Table 6.5: Preferred number of components and free parameters for each tissue class calculated by using the Bayesian information criterion.

where L is Likelihood function for the estimated model, k is the number of free parameters of the model and n is the number of samples.

The obtained preferences for each tissue class are shown in Table 6.5 as well as the number of free parameters and the D_{KL} . These results show the goodness of fit of the mixture models penalizing loss of degree of freedom by having more parameters in the fitted model. In all the cases the D_{KL} is lower than that obtained by the NMM and the RMM. The Gamma distribution provides the least number of components for describing all tissue classes. It also provides the least number of free parameters with the exception of Calcified and Fibrotic tissues. In the case of Calcified tissue, one can choose a GMM of 3 components (8 free parameters) which provides a $D_{KL} = 4.37 \cdot 10^{-5}$. This value is still lower than the one obtained for the RMM and almost equal to NMM and allows us to simplify the mixture model to prevent from over-fitting in the training step. If we decrease the number of components (and free parameters) of NMM and RMM, the D_{KL} considerably increases (see Table 6.4) and the results are worse than the ones obtained for the GMM.

For the case of Fibrotic tissue, the RMM requires less free parameters to be calculated. However, the lower number of components of the GMM avoids the effect over-fitting during the classification step as we show in a classification experiment presented in a following section (Table 6.6).

The lower preferred number of components in the GMM and NMM models compared to the RMM show the better fitting to the data. However,

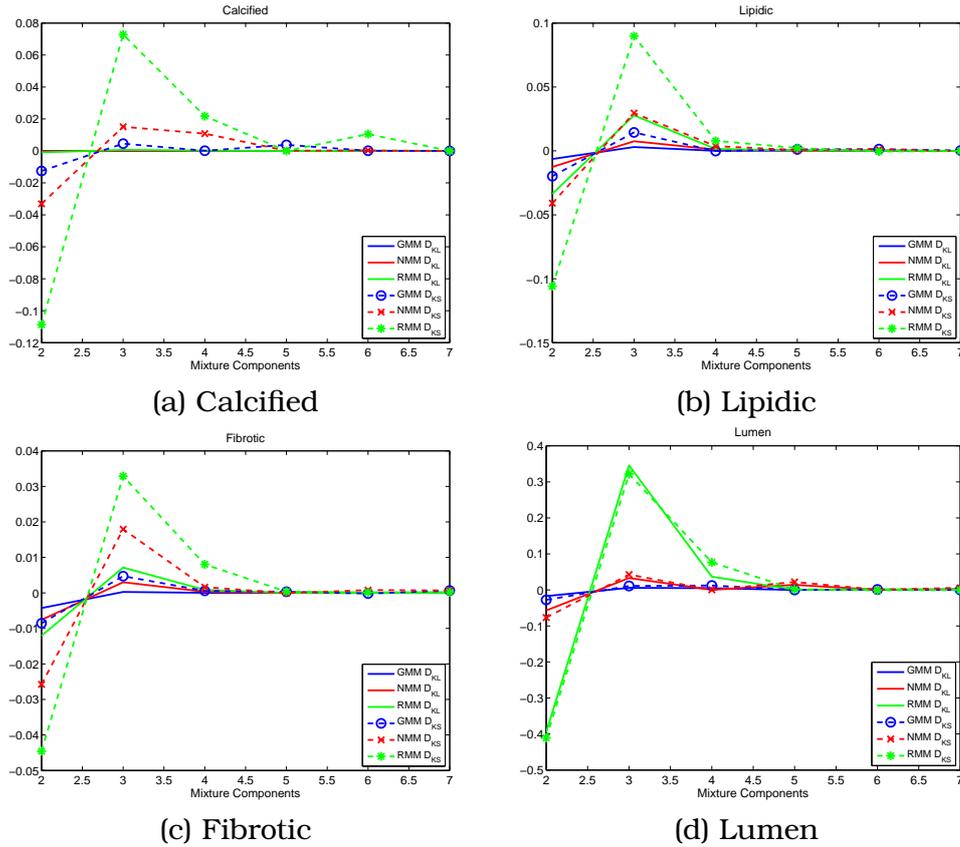


Figure 6.6: Difference rate of D_{KL} and D_{KS} for GMM, RMM and NMM models.

considering the pictures shown in Fig. 6.6 the incremental gain of the fitting measures is negligible for so many components and the BIC criterion can be relaxed to 3 components. Nevertheless, in the next section we classify with an increasing number of components and the best cases of each mixture model are considered for comparisons.

6.6.2 Plaque classification according to tissue type

In this experiment, the classifier explained in section 6.6 is applied for the whole data with the leave-one-patient-out (LOPO) cross-validation technique. This method performs the classification by excluding all the images that come from the same patient of the plaque which is being classified. This is a technique for assessing the statistical independence of classification and, thus, avoids the problem of correlation between images from the same patient in the validation stage.

The classification is performed by partitioning the data into two complementary subsets. One of the sets contains all the images from the same patient and it is used for validation (validation set), while its complementary subset (the rest of patients) is considered as the analysis set and is used for the training step (training set).

Components (Free Parameters)						
GMM	2 (5)	3 (8)	4 (11)	5 (14)	6 (17)	7 (20)
Performance	73.95%	86.56%	82.35%	81.51%	73.11%	75.63%
NMM	2 (5)	3 (8)	4 (11)	5 (14)	6 (17)	7 (20)
Performance	77.31%	83.19%	79.83%	78.15%	78.99%	76.13%
RMM	2 (3)	3 (5)	4 (7)	5 (9)	6 (11)	7 (13)
Performance	31.93%	63.86%	68.07%	71.42%	71.27%	70.79%

Table 6.6: Classification Performance for an increasing number of components and free parameters per tissue of GMM, NMM and RMM.

The training set is used to obtain the initial values of the GMMs parameters according to each tissue class. Then, the Expectation-Maximization method is applied to the image under study which belongs to the validation set and the GMMs fit to the whole image. The resulting GMMs are used to obtain the probability of belonging to each tissue class, which is obtained by the Bayes Theorem.

Table 6.6 shows the classification performance for an increasing number of components and free parameters for each mixture model in terms of global accuracy, $A = \frac{TP+TN}{TP+TN+FP+FN}$, where FN and FP stand for *False Negatives* and *False Positives* respectively; TN and TP stand for *True Negatives* and *True Positives*. Note that the highest value of accuracy is reached for three components both for the GMM and the NMM. This confirms the relation between the classification and the fitting measures and evidences that the GMMs and NMMs are good descriptors of the behavior of the probabilistic nature of speckle in each tissue class for classification purposes. It also confirms the selection of number of components and free parameters performed in the previous section.

In order to provide a comparison between the proposed method (GMM Classifier) and other methodologies, three more methods are considered for plaque classification. The first one is a Rayleigh Mixture classifier which can be easily derived following the formulation of section 6.5.1, we will refer to this method as the *RMM Classifier*. This classifier does not exactly follow the same philosophy as the proposed in [Seal1], the reason is that we are interested not only in the classification but also in providing probability maps for further post-processing and can be of help for physicians in their plaque detections or diagnosis. The RMM classifier presented in [Seal1] does not adapt the RMMs to each tissue class in the image under study and, thus, the reliability of the fitting is reduced. The second is the Nakagami Mixture Model classifier (*NMM Classifier*) obtained in the same way as the RMM Classifier. A comparison between the GMM Classifier vs. the RMM and the NMM ones will confirm once more that GMM not only describes better the speckle patterns but also provides more discriminative power as a classifier.

The third method considered for comparison is the one of [Cio10] where a set of 35 textural features directly extracted from the envelope data are used. The multi-class classification is tackle by a combination of binary classifiers

	RMM Classifier	NMM Classifier	Textural Classifier	GMM Classifier
A	71.42	83.19	81.51	86.56
S_{Cal}	96.67	96.67	96.67	96.67
S_{Lip}	50.00	35.71	57.14	42.86
S_{Fib}	20.00	72.00	48.00	76.00
S_{Lum}	88.00	94.00	96.00	98.00
K_{Cal}	93.33	94.59	98.55	96.10
K_{Lip}	81.25	93.07	87.25	95.10
K_{Fib}	88.89	94.19	93.41	93.33
K_{Lum}	95.35	92.86	96.08	96.43
P_{Cal}	87.88	87.88	96.67	90.63
P_{Lip}	28.00	41.67	38.10	54.55
P_{Fib}	33.33	78.26	66.67	76.00
P_{Lum}	95.65	92.16	96.00	96.08

Table 6.7: Performance of tissue classification (Accuracy (A), Sensitivity (S), Specificity (K), Precision (P)).

into the *Error-Correcting Output Codes* (ECOC) framework [Die95]. Each binary classification is obtained by means of *Adaptive Boosting* [Sch97] where the weak classifiers are *decisions stumps* [Sea11]. We refer to this classifier as *Textural Classifier*.

Results of the classification are shown in table VI. This table was calculated by considering the whole data set for classification. Specificity was calculated as $K = \frac{TN}{TN+FP}$; Sensitivity, $S = \frac{TP}{TP+FN}$ and Precision $P = \frac{TP}{TP+FP}$. Note that in all the classifiers the performance of detecting lipidic plaque is poorer than the others. In order to see better the differences between the methods, the confusion matrices are shown in table 6.8.

The good performance for Calcified classification becomes clear for all the classifiers since all of them classify correctly 29 out of 30 Calcified plaques. Regarding Lumen, the best are the Textural and GMM classifiers with a correct classification of 48 and 49 out of 50 respectively.

In the case of Fibrotic plaques, the best classifications are obtained with the NMM and GMM classifiers with a noticeable difference respect to the Textural classifier and RMM classifier. This result evidences the discriminative power of the GMM and NMM classifiers between Fibrotic and Lipidic tissues and how the GMM and NMM properly describe the nature of the speckle in both cases.

A very interesting case appears with RMM classifier for Fibrotic tissue. Note that this classifier shows a clear bias to the Lipidic tissue. The reason is that the PDF of Lumen shows a heavy tailed distribution which is very difficult to model with Rayleigh distributions. Thus, more Rayleigh components are needed for modeling the tail and those of Lipidic and Fibrotic tissues contribute for that purpose. This results in 6 Lumen plaques misclassified as Fibrotic and 18 Fibrotic plaques misclassified as Lipidic.

Actual	Predicted			
	Calcified	Lipidic	Fibrotic	Lumen
Calcified	29	0	1	0
Lipidic	2	7	3	2
Fibrotic	2	18	5	0
Lumen	0	0	6	44

(a) RMM Classifier

Actual	Predicted			
	Calcified	Lipidic	Fibrotic	Lumen
Calcified	29	0	1	0
Lipidic	2	5	4	3
Fibrotic	2	4	18	1
Lumen	0	3	0	47

(b) NMM Classifier

Actual	Predicted			
	Calcified	Lipidic	Fibrotic	Lumen
Calcified	29	0	1	0
Lipidic	1	8	4	1
Fibrotic	0	12	12	1
Lumen	0	1	1	48

(c) Textural Classifier

Actual	Predicted			
	Calcified	Lipidic	Fibrotic	Lumen
Calcified	29	0	1	0
Lipidic	1	6	5	2
Fibrotic	2	4	19	0
Lumen	0	1	0	49

(d) GMM Classifier

Table 6.8: Confusion Matrices.

Regarding Lipidic plaques, poorer results are observed for all the classifiers. The best performances are obtained by the Textural classifier. However, all classifiers present some difficulties in distinguishing between Lipidic and Fibrotic. This is an expected result since plaques usually present a mixed nature.

The overall accuracy shows that the GMM classifier provides an increase of 5% with respect to the Textural classifier and 3% with respect the NMM classifier.

Classifier	RMM	NMM	Textural	GMM
Mean rank	3.2	2.3	2.6	1.9

Table 6.9: Mean Rank for the Accuracy of the Classifiers.

6.6.3 Statistical Significance

The statistical significance of these results is studied by means of the *Friedman* and *Bonferroni-Dunn* tests [Dem06]. These tests were selected since no equal variance between results can be assumed and the Gaussian hypothesis is not guaranteed.

The fractional ranking of each separate classification test for each plaque is shown in Table 6.9. Fractional ranking consists of assigning distinct ordinal numbers to each classifier even if they are equal (where the assignment is arbitrarily done). Then, for those classifiers that are compared equally, the rank average is assigned to each one. Note that the best rank is obtained for the GMM classifier followed by the Textural classifier, the GMM and finally the RMM classifier. This result holds with the overall accuracy results of Table 6.7.

The *Friedman* test states that the null-hypothesis is that the differences on the measured classification performance are due to randomness. In order to reject or not this hypothesis, the *Friedman statistic value* is calculated as

$$\chi_F^2 = \frac{12N}{k(k+1)} \left(\sum_j r_j^2 - \frac{k(k+1)^2}{4} \right), \quad (6.22)$$

where $k = 4$ is the number of classifiers and $N = 119$ is the number of samples. In our case the result is $\chi_F^2 = 64.26$. However, a more thorough statistic is obtained by means of the *Iman-Davenport* correction of the *Friedman statistic*

$$F_F = \frac{(N-1)\chi_F^2}{N(k-1) - \chi_F^2}. \quad (6.23)$$

In this case, the value obtained is $F_F = 25.90$. The statistic F_F is distributed by a *Fisher-Snedecor distribution* with parameters $(k-1) = 3$ and $(k-1)(N-1) = 354$, so the critical value of $F_{3,\infty}$ for a confidence of 95% is 2.68. The statistic value obtained is much higher than this critical value so the null-hypothesis can be rejected and we can conclude that differences of the classifications are not due to randomness.

In order to see whether the GMM classifier statistically improves the results of the Textural classifier, the *Bonferroni-Dunn* test is applied. This test states that the performance of two classifiers is significantly different if the corresponding average ranks are higher than the following *critical difference*

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}}, \quad (6.24)$$

where critical values q_α are based on the Studentized range statistic [Dem06]. In our case, $q_\alpha = 2.128$ for a confidence of 90% and $q_\alpha = 2.394$ for a confi-

dence of 95%. Then, the *critical difference* is $CD_{90} = 0.36$ and $CD_{95} = 0.40$ for a confidence of 90% and 95% respectively. This distance is smaller than the differences between the GMM and the NMM so we can conclude that the GMM is significantly better than the others with a confidence of 95%. The NMM and the Textural classifiers are not significantly different between each other and all of them are significantly better than the RMM classifier.

6.6.4 Probability Maps

Two examples of the probability of belonging to each tissue class are analyzed. Both examples are depicted in Fig. 6.7. These examples show the way the plaques are represented in polar images. Note that the lipidic and fibrotic cases are difficult to be identified since both represent similar speckle pattern. Calcified plaques present a shadow below them and their identification is easier.

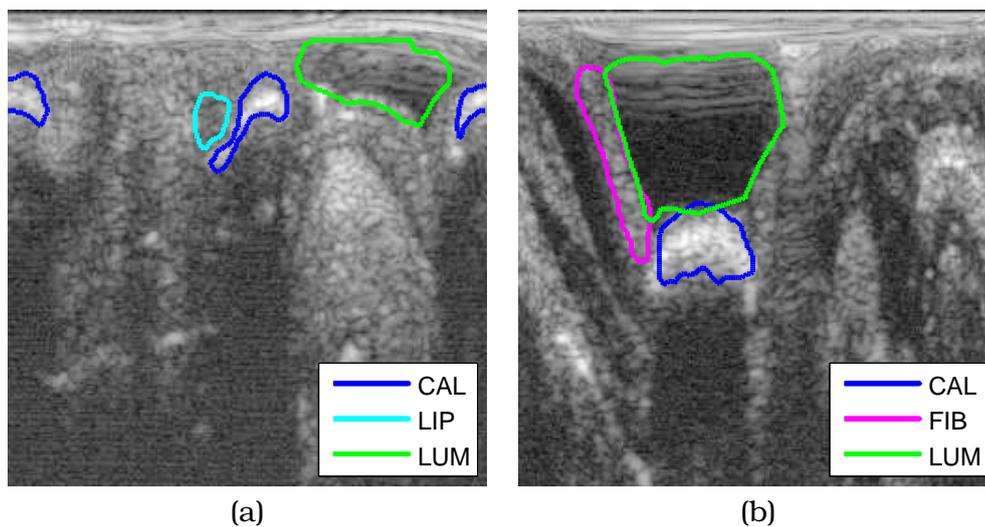


Figure 6.7: Examples of pre-segmented IVUS images. Log compression was applied for a better visualization. The example image (a) shows two calcified plaques and a small lipidic plaque. The example (b) presents a fibrotic and a calcified plaque.

The probability of belonging to each tissue class for the example shown in Fig. 6.7.a are depicted in Fig. 6.8. Note that in Fig. 6.8.a, the probability of belonging to the calcified tissue is almost 1 in all the regions labeled as calcified. The upper region with high probability of belonging to calcified class is due to the artifacts of the catheter. These artifacts do not suppose a problem since they can be easily detected or each image can be cropped according to the diameter of the catheter. Lumen tissue has also high probabilities through the whole segmented area and thus it is almost always correctly classified. The most problematic tissues are the fibrotic and lipidic ones. In these cases the distributions of tissues are quite overlapped and, for these two tissues, the probabilities are comparable. These effects are shown in table 6.10, where the classification was performed with three components and three tissue classes: Calcified, lipidic/fibrotic and Lumen.

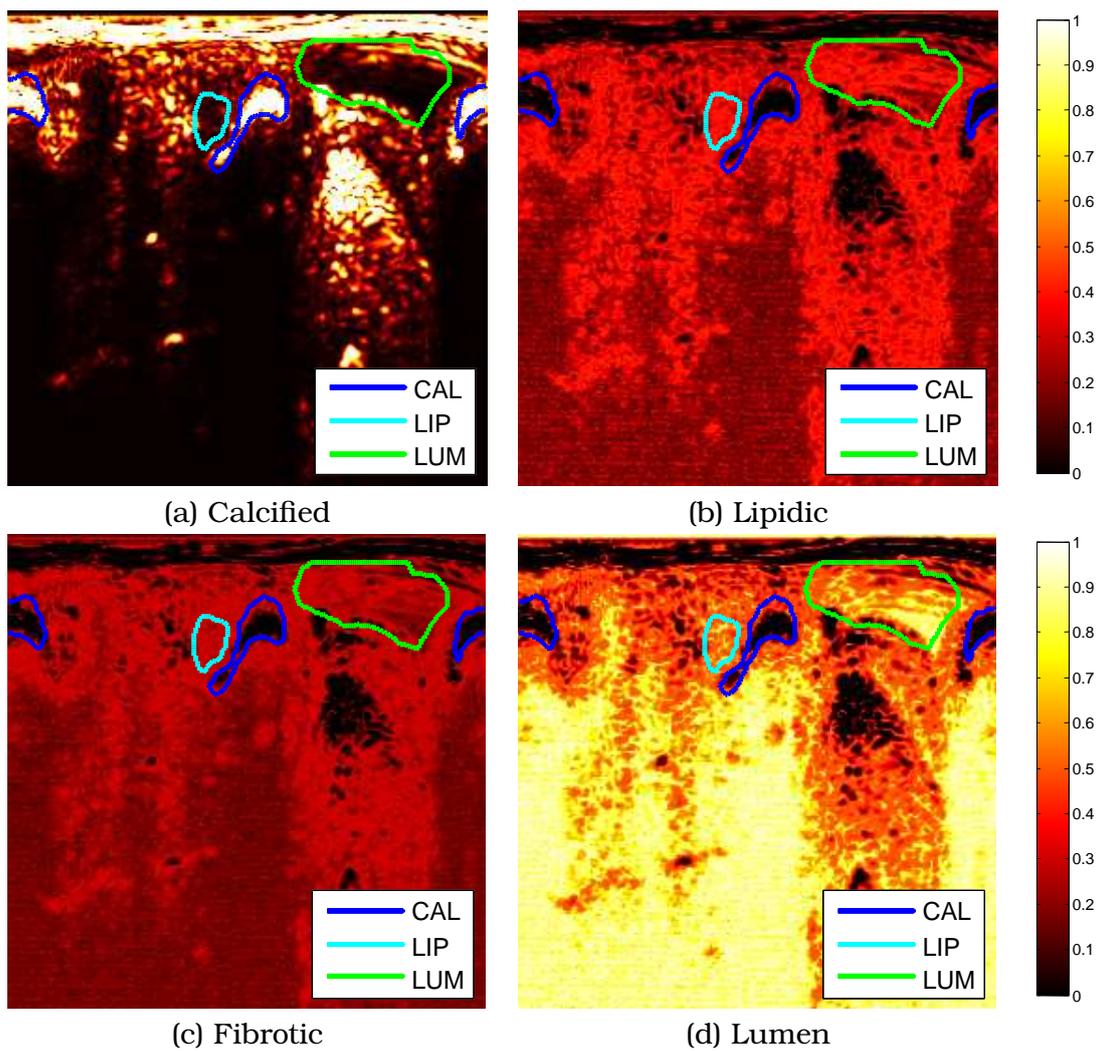


Figure 6.8: Probability maps of belonging to each tissue class in the polar image of the example in Fig. 6.7.(a).

# components = 3	Calcified	Lipidic/Fibrotic	Lumen
Precision	92.59	93.94	97.67
Sensitivity	96.15	91.18	97.67
Total Accuracy: 95.1456			

Table 6.10: Classification Performance for three components of the GMM into three classes.

The same observation is reinforced by considering the probability maps computed from the example depicted in Fig. 6.7.(b). These are shown in Fig. 6.9.

6.7 Conclusions

This work proposes a plaque characterization method for IVUS images based on the probabilistic behavior of speckle in each tissue class in different

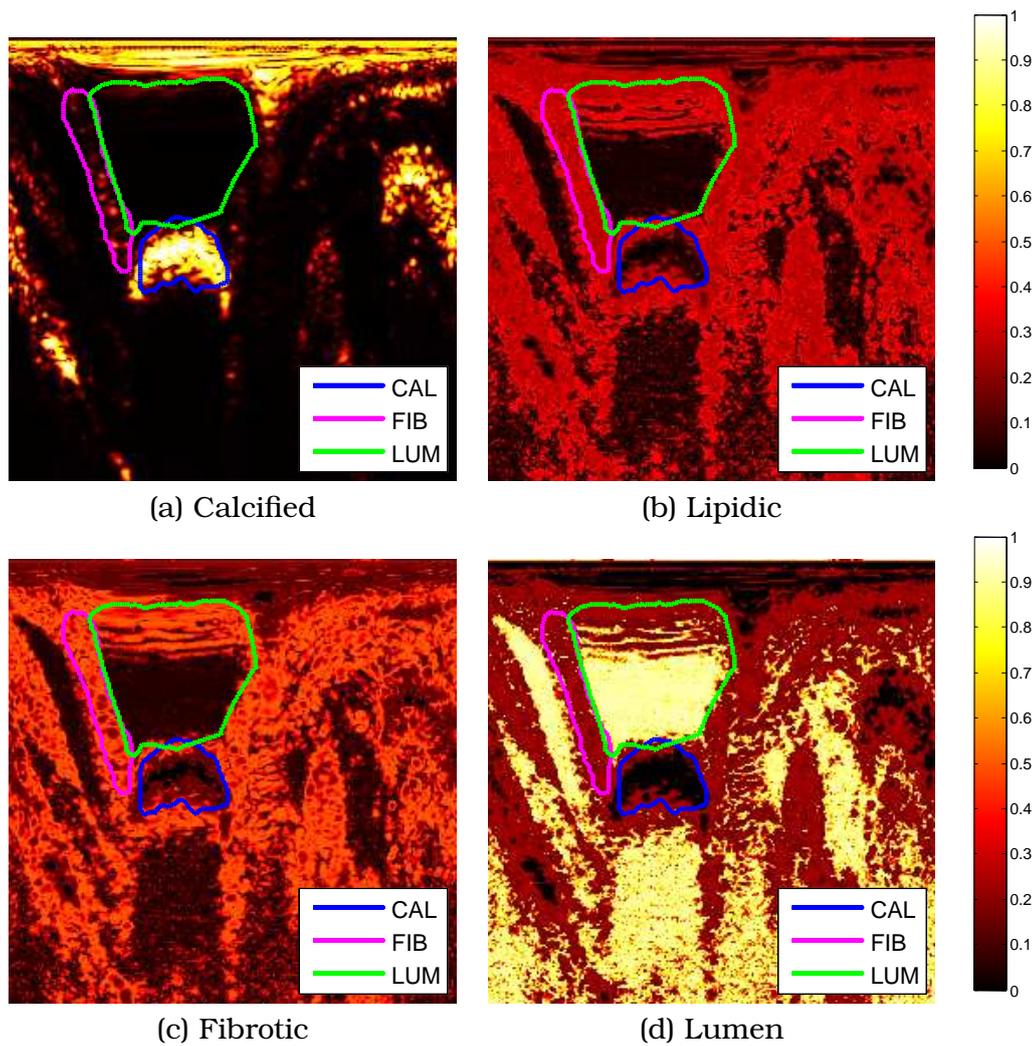


Figure 6.9: Probability maps of belonging to each tissue class in the polar image of the example in Fig. 6.7.(b).

tissue types. A Gamma distribution is assumed for the probabilistic description of the speckle because it has shown better performance than the Rayleigh and Nakagami distributions under the operations of down-sampling and interpolation of the echo envelope. Since each plaque type may present different echogenic content, a GMM is adopted for modeling each class. It was shown that the proposed GMM outperforms the Rayleigh mixture approach of [Sea11].

A classifier is proposed in order to obtain the posterior probability according to each tissue class. This classifier makes use of the GMM obtained for each test data which has been previously initialized with the GMM obtained during training. Four tissue classes were considered in this study: Calcified, lipidic, fibrotic and lumen. A leave-one-patient-out cross-validation technique was applied for validating the performance of the classifier providing an overall accuracy of 86.56%. Statistical comparison with other methods showed that the proposed classifier is significantly better than the others. The overall gain of accuracy of the proposed method

compared to the best of the others is about 5%.

The most problematic classes for classifying were Lipidic and Fibrotic tissues. This is due to the similarity between both speckle patterns. A classification using three classes was also provided, considering fibrotic and lipidic as a joint class. In that case, an overall accuracy of 95.15% was obtained. This result evidences that the proposed classifier offers a useful way to help physicians for detecting and diagnosing plaques.

Apart from the classification results, this classification scheme offers the possibility of representing probability maps of belonging to each tissue class. These probability maps can be codified into color-maps in order to offer useful information for physicians. Additionally, probability maps can be useful for filtering and automatic segmentation purposes [VSF10a].

It is important to notice that this work performs the validation with histological validated plaque. It would be desirable to complete the study with the analysis of plaque in live subjects. However, due to the difficulties of having a golden standard in live subjects this study, the study with live subjects will be carried out as a future work.

6.A Expectation Maximization for Mixtures of GMM

The joint distribution of IID samples \mathbf{X} and the hidden random variables, \mathbf{Z} and \mathbf{W} is given by

$$p(\mathbf{X}, \mathbf{Z}, \mathbf{W} | \Theta) = \prod_{i=1}^N p(x_i, z_i, w_{i,j} | \Theta), \quad (6.25)$$

where

$$\begin{aligned} p(x_i, z_i, w_{i,j} | \Theta) &= p(x_i | z_i, w_{i,j}, \Theta) p(w_{i,j} | z_i, \Theta) p(z_i | \Theta) \\ &= p(x_i | w_{i,z_i}, \Theta) p(w_{i,z_i} | \Theta) p(z_i | \Theta) \\ &= f_X(x_i | \Theta_{z_i, w_{i,z_i}}) \pi_{z_i, w_{i,z_i}} \nu_{z_i}. \end{aligned} \quad (6.26)$$

Note that, the following equation holds

$$p(x_i, \Theta) = \sum_{z_i} \sum_{w_{i,z_i}} f_X(x_i | \Theta_{z_i, w_{i,z_i}}) \pi_{z_i, w_{i,z_i}} \nu_{z_i} = \sum_{j=1}^J \nu_j \sum_{r=1}^{R_j} \pi_{j,r} f_X(x_i | \Theta_{j,r}). \quad (6.27)$$

Now, the log-likelihood function function can be defined in the following way

$$\begin{aligned} \mathcal{L}(\Theta | \mathbf{X}, \mathbf{Z}, \mathbf{W}) &= \sum_{i=1}^N \log p(x_i, z_i, w_{i,j} | \Theta) \\ &= \sum_{i=1}^N \log p(x_i | \Theta_{z_i, w_{i,z_i}}) + \sum_{i=1}^N \log p(w_{i,z_i} | \Theta) p(z_i | \Theta) \\ &= \sum_{i=1}^N \log f_X(x_i | \Theta_{z_i, w_{i,z_i}}) + \sum_{i=1}^N \log \pi_{z_i, w_{i,z_i}} + \sum_{i=1}^N \log \nu_{z_i}. \end{aligned} \quad (6.28)$$

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

$$\begin{aligned} \mathcal{Q}(\Theta|\Theta^{(n)}, \mathbf{X}) &= E_{\mathbf{Z}, \mathbf{W}|\Theta^{(n)}, \mathbf{X}}\{\mathcal{L}(\Theta|\mathbf{X}, \mathbf{Z}, \mathbf{W})\} = \\ &= \sum_{i=1}^N \sum_{j=1}^J \sum_{r=1}^{R_j} \gamma_{i,j,r} (\log p(x_i|\alpha_{j,r}, \beta_{j,r}) + \log(\nu_j \pi_{j,r})), \end{aligned} \quad (6.29)$$

where $\gamma_{i,j,r} = p(Z_i = j, W_{i,j} = r|x_i, \Theta^{(n)})$, which can be easily calculated by means of the Bayes theorem

$$\gamma_{i,j,r} = \frac{f_X(x_i|\Theta_{j,r}^{(n)})\nu_j^{(n)}\pi_{j,r}^{(n)}}{\sum_{s=1}^J \sum_{r=1}^{R_j} f_X(x_i|\Theta_{s,r}^{(n)})\nu_s^{(n)}\pi_{s,r}^{(n)}}. \quad (6.30)$$

The maximization process can be done in each of two terms separately. Then, the maximization step can be performed in the same way as was done in section 6.4. So, for the term with a dependence on the parameters ν_j and $\pi_{j,r}$ in eq. (6.29), one can establish a *Lagrange function* with *Lagrange multipliers* λ , μ_j for $j = \{1, \dots, J\}$. The constraints are those expressed in eq. (6.13). The *Lagrange function* is the following

$$\begin{aligned} \Lambda(\nu, \pi, \lambda, \mu) &= \sum_{i=1}^N \sum_{j=1}^J \sum_{r=1}^{R_j} \gamma_{i,j,r} \log(\nu_j \pi_{j,r}) \\ &+ \lambda \left(\sum_{j=1}^J \nu_j - 1 \right) + \sum_{j=1}^J \mu_j \left(\sum_{r=1}^{R_j} \pi_{j,r} - 1 \right). \end{aligned} \quad (6.31)$$

Taking the derivative with respect to ν_j , for a fixed j , the following relation is deduced

$$\nu_j \lambda = - \sum_i^N \sum_r^{R_j} \gamma_{i,j,r}, \quad (6.32)$$

so, summing over $j = \{1, \dots, J\}$, one gets $\lambda = -N$. So the value of ν_j that maximizes $\Lambda(\nu, \pi, \lambda, \mu)$ is

$$\nu_j^{(n+1)} = \frac{1}{N} \sum_{i=1}^N \sum_{r=1}^{R_j} \gamma_{i,j,r}. \quad (6.33)$$

The calculation of $\pi_{j,r}$ is performed in the same way, the result is

$$\pi_{j,r}^{(n+1)} = \frac{\sum_{i=1}^N \gamma_{i,j,r}}{\sum_{i=1}^N \sum_{s=1}^{R_j} \gamma_{i,j,s}}. \quad (6.34)$$

Now, regarding the term of eq. (6.29) which depends on the parameters of the distribution $\alpha_{j,r}$ and $\beta_{j,r}$, the process of deriving these parameters is

analog to the GMM model explained in section 6.4. It consists in solving the following equations

$$\begin{aligned}\hat{\beta}_{j,r} &= \frac{1}{\hat{\alpha}_{j,r}} \frac{\sum_i^N \gamma_{i,j,r} x_i}{\sum_i^N \gamma_{i,j,r}} \\ \log(\hat{\alpha}_{j,r}) - \psi(\hat{\alpha}_{j,r}) &= \log\left(\frac{\sum_i^N \gamma_{i,j,r} x_i}{\sum_i^N \gamma_{i,j,r}}\right) - \frac{\sum_i^N \gamma_{i,j,r} \log(x_i)}{\sum_i^N \gamma_{i,j,r}}.\end{aligned}\quad (6.35)$$

From the estimated value $\hat{\alpha}_{j,r}$ that maximizes the log-likelihood, the estimate of $\hat{\beta}_{j,r}$ can be calculated. See Appendix 6.B for solving this equation.

6.B Solving $\psi(x) - \log(x) = K$

In this appendix we demonstrate that the function $g(x) = \psi(x) - \log(x)$ is a strictly increasing function when $x > 0$. For this purpose we make use of the demonstration of the convexity of $f(x) = \log \Gamma(x) - x \log x$ for $x > 0$ (see [Mer96]).

The property of convexity of $f(x)$ just guarantees that the first derivative of $f(x)$ is monotonically non-decreasing. Thus, a more detailed analysis should be done. To this end, we recall the well known expansion of the second derivative of $\log \Gamma(x)$ (see [Mer96])

$$\frac{d^2}{dx^2} \log \Gamma(x) = \frac{1}{x^2} + \frac{1}{(x+1)^2} + \frac{1}{(x+2)^2} + \dots, \quad (6.36)$$

with $x \neq 0, -1, -2, \dots$.

Making use of eq. (6.36), the second derivative of $f(x)$ can be expressed as

$$f''(x) = \frac{d^2}{dx^2} \log \Gamma(x) - \frac{1}{x} = \sum_{k=0}^{\infty} \frac{1}{(x+k)^2} - \frac{1}{x}. \quad (6.37)$$

The second term can be introduced in the sum by means of expressing it as the following telescoping series

$$\frac{1}{x} = \sum_{k=0}^{\infty} \frac{1}{x+k} - \frac{1}{x+k+1}. \quad (6.38)$$

This yields the following result

$$f''(x) = \sum_{k=0}^{\infty} \frac{1}{(x+k)^2} - \frac{1}{x+k} + \frac{1}{x+k+1} = \sum_{k=0}^{\infty} \frac{1}{(k+x)^2(1+k+x)} > 0, \quad (6.39)$$

for $x > 0$. This result demonstrates the convexity of $f(x)$ and also that $f'(x) = \psi(x) - \log(x) - 1$ is a strictly monotonic increasing function and so it is $g(x) = \psi(x) - \log(x)$.

We are interested in calculating the value \hat{x} that follows eq. (6.35) that, without loss of generality, can be stated as

$$\log(x) - \psi(x) = K, \quad (6.40)$$

where K is the second term of the eq. (6.35). Note that this constant K is positive unless all elements x_i are identical (a case that does not occur in practice) due to the Jensen's inequality

$$\log \left(\frac{\sum_i^N \gamma_i x_i}{\sum_i^N \gamma_i} \right) \geq \frac{\sum_i^N \gamma_i \log(x_i)}{\sum_i^N \gamma_i}, \quad (6.41)$$

for any $x_i > 0$ and $\gamma_i > 0$.

An interval where the function $x \rightarrow \log(x) - \psi(x) - K$ changes its sign can be derived just by applying the following result (see [Alz97])

$$\frac{1}{2x} < \log(x) - \psi(x) < \frac{1}{x}. \quad (6.42)$$

Hence, the solution $\hat{x} \in (\frac{1}{2K}, \frac{1}{K})$ and any root-finding method can be efficiently used.

Probabilistic-Driven Oriented Speckle Reducing Anisotropic Diffusion with Application to Cardiac Ultrasonic Images

The law of conservation of energy tells us we can't get something for nothing, but we refuse to believe it.
Isaac Asimov, 1920–1992.

Abstract– A novel anisotropic diffusion filter is proposed in this work with application to cardiac ultrasonic images. It includes probabilistic models which describe the probability density function (PDF) of tissues and adapts the diffusion tensor to the image iteratively. For this purpose, a preliminary study is performed in order to select the probability models that best fit the statistical behavior of each tissue class in cardiac ultrasonic images. Then, the parameters of the diffusion tensor are defined taking into account the statistical properties of the image at each voxel. When the structure tensor of the probability of belonging to each tissue is included in the diffusion tensor definition, a better boundaries estimates can be obtained instead of calculating directly the boundaries from the image. This is the main contribution of this work. Additionally, the proposed method follows the statistical properties of the image in each iteration. This is considered as a second contribution since state-of-the-art methods suppose that noise or statistical properties of the image do not change during the filter process.

Adapted from: G. Vegas-Sánchez-Ferrero¹, S. Aja-Fernández¹, M. Martín-Fernández¹, A. F. Frangi² and C. Palencia³, *Probabilistic-Driven Oriented Speckle Reducing Anisotropic Diffusion with Application to Cardiac Ultrasonic Images*, In 13th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Beijing, China, Lecture Notes in Computer Science, Vol. 6361, pp. 518–525, September 2010.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Center for Computational Imaging & Simulation Technologies in Biomedicine, Univ. Pompeu Fabra

³Departamento de Matemática Aplicada, Univ. Valladolid.

7.1 Introduction

Ultrasound is a non-invasive imaging modality and a low-cost way to help diagnosing, which is widespread for many medical applications. However, ultrasonic (US) images are characterized by the presence of speckle, which is a granular pattern that degrades resolution and adds spatial noise to the image [Nil08]. Therefore, many speckle removal filters have been proposed in the literature [Yu02, AF06, Kri07].

Speckle in ultrasound images can be seen as a random process whose statistical features depend on the tissue class. The existence of a deterministic component due to specular reflections of the echo-pulse depends on the number of obstacles of the tissue (scatters) into the resolution cell and their size in comparison with the wavelength of ultrasound signal. Thus, many of the speckle removal filters are based on local statistics.

Among those filters, we focus on anisotropic diffusion based filters since the proposed filter can be considered as a probabilistic extension of them. One of the most popular approaches is Speckle Reducing Anisotropic Diffusion (SRAD) [Yu02], which extends the well known Perona and Malik's anisotropic diffusion avoiding to threshold the norm of the gradient. Instead of that threshold, an estimate of the coefficient of variation of noise is used. Results and stability depends on a good estimation of the local statistics as it was demonstrated in [AF06].

In the case of Oriented Speckle Reducing Anisotropic Diffusion (OSRAD) [Kri07], no logarithmic compression is supposed and the speckle statistics are assumed to be modeled by a Rician distribution. In the case of high SNR, speckle statistics are approximated by a Gaussian distribution. This is a reasonable assumption when resolution cells have a large number of scatters and there exists a deterministic component, which is the case of the myocardial tissue for the raw signal obtained by the transducer. However, besides logarithmic compression, an interpolation process is performed to reconstruct the image and, therefore, the probabilistic model is affected. Additionally, the diffusion schemes are iterative, so the Rician distribution supposition does not hold in each iteration. So, because of logarithmic compression and interpolation, the statistics of speckle in reconstructed images do not follow the original distributions of the raw signal. Hence, many recent works study different probability models in order to obtain a suitable model to describe the statistical behavior of the speckle [Nil08, Tao06, Elt06].

In this work, we propose a novel anisotropic diffusion filter with application to cardiac ultrasonic images, which includes the probability models

that describe PDFs of different tissue classes and adapt iteratively to the filtered image. Probability models for each tissue class were selected from a training database using real-life cases for cardiac ultrasonic images. A mixture PDF model is adopted for fitting probabilistic models of tissue classes to the histogram of the whole image. With this methodology, a Bayesian scheme can be applied to estimate the probability of belonging to each of the tissue classes for every image voxel.

When the structure tensor of the probability of belonging to each tissue is combined in the definition of the diffusion tensor, better boundaries estimates can be obtained instead of directly calculating them from the image. The structure tensor of the probability of belonging to each tissue allows defining a diffusion tensor, which takes into consideration the boundaries between different tissue classes for filtering purposes. This is the main contribution of this work and, to the best of our knowledge, no similar approach has been considered in the literature for this purpose.

The paper is organized as follows. In section 7.2, the statistical study is explained and the distributions are selected. Section 7.3 describes the filtering method. Section 7.4 presents some experiments and results on synthetic and real images. In section 7.5, we conclude analyzing the results.

7.2 Probabilistic Estimation of Tissue Classes in US Images

In this section, the goodness-of-fit of several probabilistic distributions is presented. The study is extended to a representative set of distributions that were proposed in the literature: Gamma [Nil08, Tao06], Log-Normal [Tao06], Rayleigh [Nil08, Tao06, Elt06], Normal [Tao06], Nakagami [Nil08, Elt06], Beta, Rician Inverse Gaussian (RiIG) [Elt06], Rice [Wag83], Exponential and K [Nil08]. A set of 120 two-dimensional images of size 1024×768 and 8 bits per pixel were obtained from a clinical machine Philips iE33 ultrasonographic system¹ with the software PMS5.1 Ultrasound iE33 4.0.1.357 scanned from human subjects.

The methodology used to carry out the study is the same used in [Nil08, Tao06] where a χ^2 goodness-to-fit test is performed. We used a significance value $\alpha = 0.05$ for the test. In order to avoid spatial correlation, the image is sub-sampled by a factor of 6. Two different tissue classes were considered: blood and myocardium.

In Fig. 7.1, an example of the analyzed tissues is presented as well as the results of passing the goodness-of-fit test for blood and myocardial tissue. In Fig. 7.1.b we can see the better performance of the gamma distribution in the case of blood, this result holds with that one obtained in [Tao06]. In the case of the myocardial tissue (Fig. 7.1.c), Rice distribution and Normal distributions are the ones with the best fit. This result is coherent to the in-

¹Philips Healthcare, Andover, MA, USA.

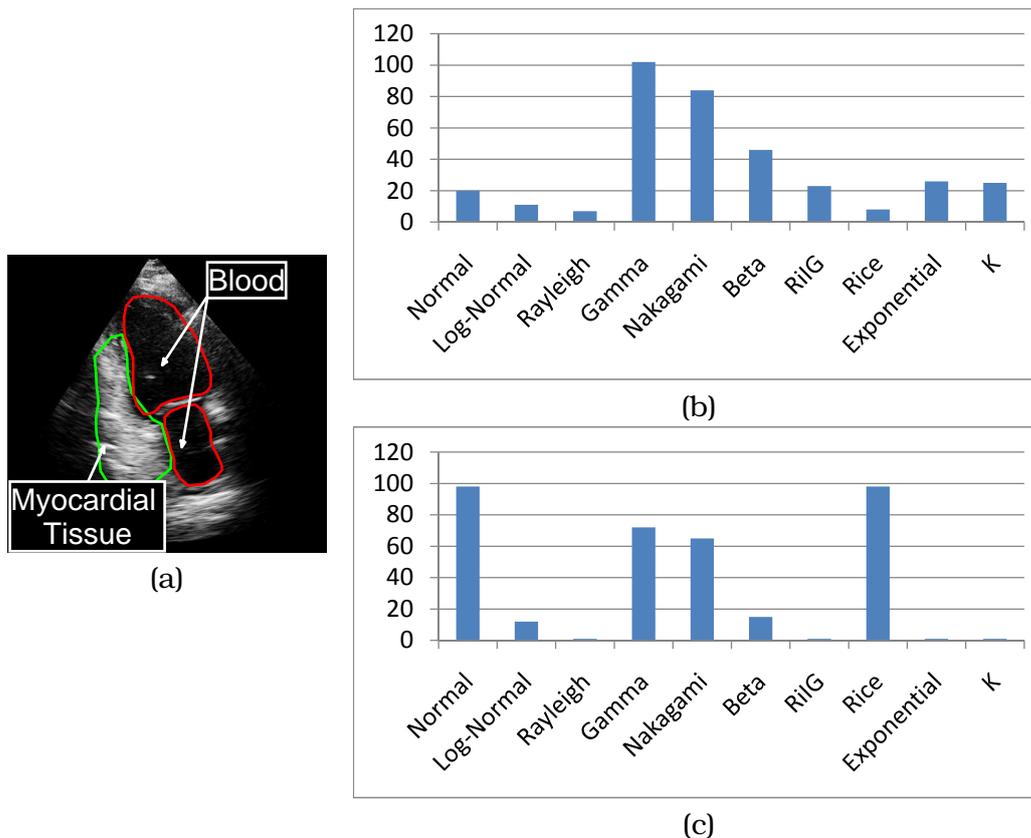


Figure 7.1: (a) Example of the regions for blood and myocardial tissue. (b) Blood. (c) Myocardial Tissue. For blood tissue class, the Gamma distribution performs best. For myocardial tissue, the best performance is achieved for Rice or normal distributions.

terpretation of quasi periodic scatterers, which gives rise to a Rician model of speckle [Wag83] and confirms the hypothesis of the OSRAD filter [Kri07]. Due to the simplicity of estimating Normal parameters for estimating mixtures of PDFs, we chose the Normal distribution as a good candidate for this tissue class.

7.3 Probabilistic Directional SRAD

When the histogram of the ultrasonic image is considered as the contribution of different kinds of tissues which follow different PDFs, the mixture of PDFs arises as a natural way to fit the histogram. This way, when parametric distributions are considered, the parameters of each independent variable have to be calculated. The most widespread method for this purpose is the Expectation-Maximization (EM) algorithm, which allows estimating the parameters of the mixture. We decided to adopt the unsupervised learning of mixture models method [Fig02], since it is capable of selecting the number of mixture components from a maximum (k_{\max}) and it simultaneously integrates model selection and estimation.

Let X be a random variable (and x a sample) that follows a finite mixture

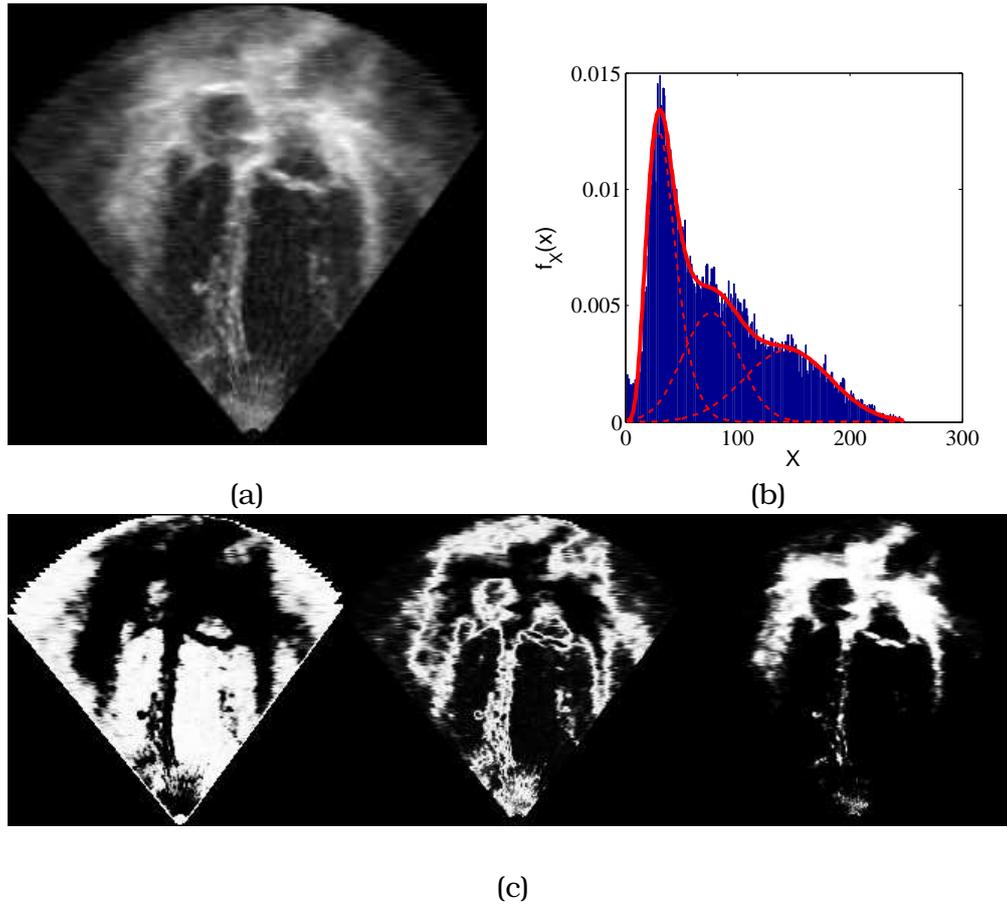


Figure 7.2: (a) Clinical ultrasound image. (b) Finite Mixture Model. (c) Tissue class probability map for each voxel of the image. The good fit of the mixture model of the image (a) becomes clear as well as the structure of the ownership to each class where the boundaries of the left and right ventricles are clearly defined.

PDF of K components where $K \in [1, k_{\max}]$. Its PDF can be written as follows

$$f_X(x) = \sum_{k=1}^K a_k f_{X_k}(x | \Theta_k), \quad (7.1)$$

where $a_k > 0$ and $\sum_{k=1}^K a_k = 1$ are the probabilities of belonging to each class of the mixture, and Θ_k are the parameters of the class \mathcal{C}_k , which follow a Gaussian or Gamma distribution. The probability of belonging to each class for each voxel (i.e. each sample x) is the following

$$P(x \in \mathcal{C}_k | X = x, \Theta_k) = \frac{P(X = x | x \in \mathcal{C}_k, \Theta_k) P(x \in \mathcal{C}_k | \Theta_k)}{\sum_{i=1}^K P(X = x | x \in \mathcal{C}_i, \Theta_i) P(x \in \mathcal{C}_i | \Theta_i)}. \quad (7.2)$$

This expression allows us to calculate a set of images that indicate the probability of belonging to each class for every voxel and can be seen as the partial volume contribution of each class. Fig 7.2 shows an ultrasound image, its finite mixture PDF, and the probability ownership to each class. In Fig 7.2.b a good fit of the mixture model of the image in Fig. 7.2.a becomes clear as well as the structure the ownership to each class (Fig. 7.2.c).

The main advantage of this set of images is that a structure tensor can be derived from each probability image, which is able to extract the directions of minimal probability change even when low gradients are presented. This way, an anisotropic filter can be defined that takes the advantage of these directions in the probability map for each tissue class. We will refer to it as POSRAD (Probabilistic Oriented Speckle Reducing Anisotropic Diffusion). This filter makes use of the extension of the scalar component of the diffusion equation to a diffusion matrix which takes the advantage of the local tissue orientation as it was proposed in [Kri07, Kri09a]. However, in our case we calculate the local orientation of the probability density function for each tissue class and combine them in the following way.

As in [Kri09a], we calculate the structure tensor of the probability density function for each tissue class as

$$T_k = G_\sigma * (\nabla_\sigma P(x \in \mathcal{C}_k | X = x, \Theta_k) \cdot \nabla_\sigma P(x \in \mathcal{C}_k | X = x, \Theta_k)^T), \quad (7.3)$$

where G_σ is a Gaussian kernel of standard deviation σ , and $\nabla_\sigma P(x \in \mathcal{C}_k | X = x, \Theta_k)$ is the gradient of the probability density function for each tissue class filtered with a Gaussian kernel of standard deviation σ . Finally, let $\lambda_1^k \geq \lambda_2^k \geq \lambda_3^k$ be the eigenvalues and $(\mathbf{v}_1^k, \mathbf{v}_2^k, \mathbf{v}_3^k)$ their respective eigenvectors. The local orientation of the maximal variation of probability of the class \mathcal{C}_k is given by \mathbf{v}_1^k , and the local orientation of the minimal variation is given by \mathbf{v}_3^k .

Let consider the following diffusion equation

$$\begin{cases} u(x, 0) &= u_0 \\ \frac{\partial u}{\partial t} &= \text{div}(D\nabla u), \end{cases} \quad (7.4)$$

where the matrix D is expressed in a diagonal form with the eigenvectors $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ and eigenvalues $\lambda_1, \lambda_2, \lambda_3$. Since we have K structure tensors (each tissue class with probability density function), we choose the eigenbase of the structure tensor with maximal λ_1^k : $\hat{k} = \arg \max_k (\lambda_1^k)$. This base gives the orientation of the maximal variation of probability among all the classes. The interpretation of this choice is that we choose as boundary the one with the maximal gradient of the probability density function over all tissue classes. This way, the most probable boundary is preserved in the filtering process. In the basis of $T_{\hat{k}}$, namely $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, the diffusion matrix D is defined as

$$D = E \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} E^T \quad \text{where} \quad \begin{aligned} \lambda_1 &= 1 - \|\nabla_{\mathbf{e}_1, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2 \\ \lambda_2 &= 1 - \|\nabla_{\mathbf{e}_2, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2 \\ \lambda_3 &= 1, \end{aligned} \quad (7.5)$$

where $\|\cdot\|_2$ is the 2-norm, $\nabla_{\mathbf{e}_i, \sigma}$ is the directional derivative in the direction \mathbf{e}_i filtered with a Gaussian kernel with a standard deviation σ , and E is the matrix whose columns are the eigenvectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Note that this definition performs a diffusion filtering in the direction of the minimal variation of probability (\mathbf{e}_3) while preserves the maximal variation of probability since $\|\nabla_{\mathbf{e}_1, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2$ will have a value closed to 1. Note that the discrete approximations of $\|\nabla_{\mathbf{e}_1, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2$ and $\|\nabla_{\mathbf{e}_2, \sigma} P(x \in \mathcal{C}_{\hat{k}} | X = x, \Theta_{\hat{k}})\|_2$ are bounded in $[0, 1]$, thus $\lambda_1, \lambda_2 \in [0, 1]$.

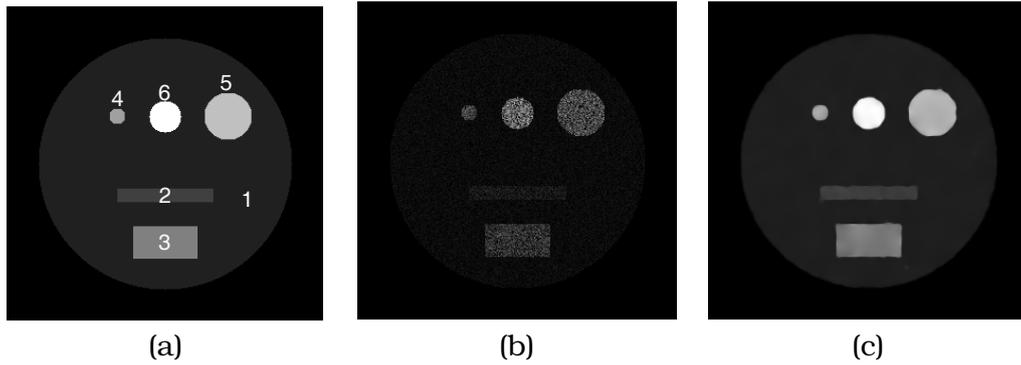


Figure 7.3: (a) Original synthetic image of [Kri07]. (b) Noisy image with multiplicative noise with $\sigma = 0.5$. (c) Filtered image with the proposed filter.

The main advantage of this definition is that, in homogeneous areas D becomes isotropic and in the presence of boundaries it becomes more anisotropic and with a main orientation aligned along the boundaries thus preserving the main contours. Additionally, the evolution of the probabilistic models can be followed through the iterative process since the finite mixture model is re-estimated in each iteration. This is an important advantage of the presented methodology over other state-of-the-art filters, which assume that the statistics of the image/noise do not change during the filtering process. Since no assumptions can be made about the changes of the statistics of the image, a finite mixture of Gaussian distributions is adopted during the filtering process except for the first iteration.

7.4 Experiments and Results

In this section we present some experiments with synthetic and real ultrasound images in order to compare our technique to other state-of-the-art algorithms. First we consider the 2D synthetic image presented in [Kri07] which is publicly available² and allows us to compare the proposed method to other algorithms. This image is presented in Fig 7.3.a which has 6 regions of constant intensity. The image is corrupted with a Gaussian multiplicative noise with $\sigma_n = 0.5$ (Fig. 7.3.b). In Fig. 7.3.c the filtered image with the proposed method is presented. The parameters of the filter were: 200 iterations, time step 0.5, $\sigma = 1$ and $k_{\max} = 6$. No suppositions were made concerning the probability distributions of the tissue classes, so a Gaussian mixture model was used.

Since these parameters are the same used by OSRAD, a direct comparison with the results in [Kri07] can be established. We use the same quality measures for validation: M_f and figure of merit (FOM)³, see [Kri07] for more details. Both values are comprised between 0 and 1. The highest the value, the better the filter. Results presented in Table 7.1 show that the

²<http://serdis.dis.ulpgc.es/~krissian/HomePage/Demos/OSRAD/OSRAD.html>

³The quality measure M_f was recalculated since it measures the quality of the filter compared to the others filters.

best FOM and M_f is achieved by our method. These results show that a better performance can be achieved when variations of the probability for each class is taken into consideration since the diffusion tensor becomes more anisotropic near the most probable edges while homogeneous regions are filtered in an isotropic way.

We now consider a 3D ultrasound image of a liver generated using the Stradwin v3.8 software (Cambridge University, Cambridge, UK), which has been used commonly in the literature [Kri07]. These data were acquired by a freehand system and are available on Cambridge University⁴. The image is a volume with $201 \times 193 \times 142$ voxels with an isotropic resolution of $0.5 \times 0.5 \times 0.5$ mm³. Both, the original and the filtered image are presented in Fig 7.4.(a-b). Visual inspection shows that a good edge preservation is obtained with an efficient noise removal. The parameters for this experiment were: 200 iterations, time step 0.5, $\sigma = 1$ and $k_{\max} = 6$ with a Gaussian mixture model.

As a final experiment, we consider a real 3D ultrasonic image of the heart obtained with the same clinical machine of section 7.2. In this experiment we use the finite mixture model of Gaussian and Gamma distributions, 100 iterations, time step 0.5, $\sigma = 1$ and $k_{\max} = 4$. Results are presented in Fig. 7.4.(c-d) where a good edge preservation is observed. An efficient noise removal is achieved in the left ventricle when compared to the noisy image.

7.5 Conclusion

In this work we have presented a new diffusion filter that takes into account the statistical properties of the image. This methodology offers an advantage over the follow-up of statistical properties of the image in each iteration while preserving and enhancing the structures. Experiments on synthetic and real images show that the proposed method preserves edges of the structures better than others and performs a good restoration.

⁴<http://mi.eng.cam.ac.uk/~rwp/stradwin/>

Method	R1	R2	R3	R4	R5	R6	M_f	FOM
Initial	10	20	40	50	60	80		1
Noisy image	10.03 ± 5.00	2.26 ± 9.57	39.39 ± 20.23	54.86 ± 22.09	58.72 ± 29.55	81.12 ± 40.59		0.193
Median	10.00 ± 0.65	18.58 ± 1.96	37.42 ± 4.77	38.42 ± 7.53	57.40 ± 4.89	75.80 ± 9.77	0.01	0.434
Lee	10.09 ± 1.42	19.34 ± 1.93	38.55 ± 5.64	48.66 ± 7.27	56.18 ± 7.20	77.82 ± 11.96	0.02	0.224
Kuan	10.11 ± 1.42	19.33 ± 1.90	38.53 ± 5.47	48.64 ± 7.26	57.13 ± 6.93	77.71 ± 11.52	0.02	0.222
P&M	9.97 ± 2.39	19.95 ± 9.10	39.38 ± 20.17	54.90 ± 22.02	58.72 ± 29.53	81.12 ± 40.58	0.00	0.323
homotopic P&M	9.55 ± 1.53	18.56 ± 6.00	37.80 ± 13.56	51.46 ± 11.37	56.31 ± 19.87	77.85 ± 26.69	0.00	0.533
Catte et al.	10.02 ± 0.89	20.02 ± 1.22	39.39 ± 10.36	54.86 ± 19.65	58.72 ± 17.97	81.12 ± 34.55	0.00	0.626
Flux	10.01 ± 0.61	20.13 ± 0.93	39.58 ± 4.08	55.12 ± 4.93	58.90 ± 5.49	81.36 ± 8.82	0.33	0.693
Rudin et al.	10.01 ± 0.87	15.11 ± 0.41	37.49 ± 2.08	47.16 ± 3.15	56.97 ± 2.62	79.10 ± 6.64	0.04	0.852
Rudin et al. attach	9.85 ± 0.89	15.04 ± 0.42	37.44 ± 2.10	47.14 ± 3.17	56.93 ± 2.63	79.07 ± 6.67	0.03	0.846
SRAD, expl., N_4	10.12 ± 0.93	19.95 ± 1.60	39.66 ± 4.63	53.82 ± 5.07	58.89 ± 7.07	81.45 ± 10.44	0.17	0.361
SRAD, expl., N_4	10.11 ± 0.98	20.00 ± 1.55	39.64 ± 5.03	54.15 ± 4.36	58.96 ± 7.25	81.48 ± 10.53	0.16	0.347
SRAD, impl., N_4	10.47 ± 0.59	20.66 ± 0.90	41.36 ± 2.74	57.51 ± 2.96	60.89 ± 2.06	84.89 ± 4.33	0.24	0.375
SRAD, impl., N_4	10.42 ± 0.39	20.42 ± 0.40	41.22 ± 2.54	59.14 ± 0.81	60.69 ± 1.02	84.13 ± 1.37	0.43	0.615
SRAD, impl., N_4	10.12 ± 0.93	19.94 ± 1.64	39.64 ± 4.58	53.67 ± 4.89	58.86 ± 6.95	81.37 ± 10.27	0.18	0.387
DPAD, expl., N_9	10.11 ± 0.97	20.00 ± 1.55	39.63 ± 4.93	54.03 ± 4.15	58.94 ± 7.08	81.43 ± 10.22	0.17	0.341
DPAD, impl., N_4	10.48 ± 0.38	20.65 ± 0.66	41.54 ± 2.19	57.43 ± 3.06	60.61 ± 2.13	85.20 ± 3.72	0.32	0.524
DPAD, impl., N_9	10.44 ± 0.35	20.43 ± 0.54	41.09 ± 2.00	58.35 ± 1.11	60.89 ± 0.86	85.00 ± 2.03	0.45	0.731
OSRAD	10.28 ± 0.51	20.09 ± 0.82	40.38 ± 2.38	54.95 ± 1.69	59.64 ± 3.73	83.52 ± 4.26	0.47	0.768
POSRAD	10.09 ± 0.68	19.91 ± 0.85	39.49 ± 1.53	51.90 ± 3.14	58.11 ± 2.11	79.57 ± 3.95	0.55	0.865

Table 7.1: Results on filtering synthetic 2D image with $\sigma_n = 0.5$, presented as Mean ± Standard Deviation of the intensity in each region (R1 - R6) of the Fig. 7.3.a.

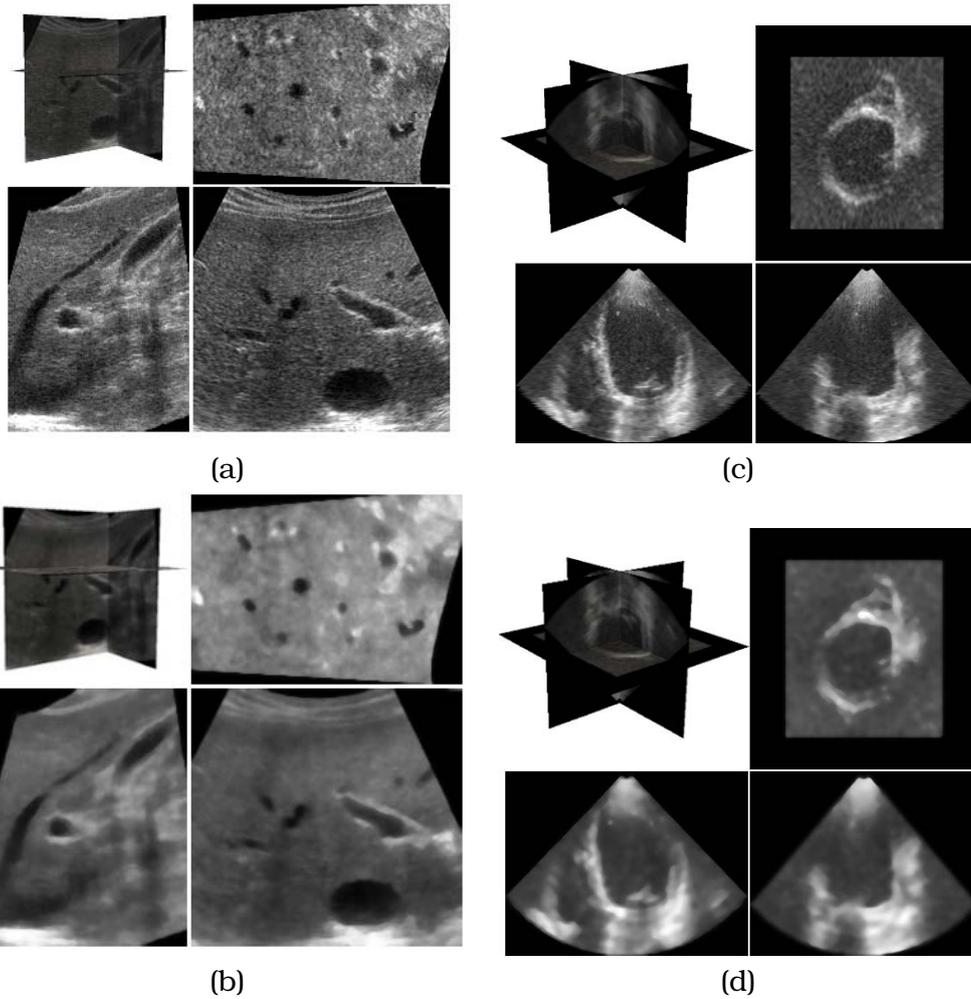


Figure 7.4: (a) Noisy 3D ultrasound of a liver. (b) Filtered 3D ultrasound of a liver. (c) Noisy 3D ultrasound of a Heart. (d) Filtered 3D ultrasound of a Heart. The proposed method achieve a good edge preservation with an efficient noise removal in both cases.

Realistic Log-Compressed law for Ultrasound Image Recovery

*Dans les champs de l'observation,
le hasard ne favorise que les esprits préparés.
Luis Pasteur, 1822–1895.*

Abstract– A realistic log-compressed law for ultrasound images based on a real device is proposed. The model takes into account the linear behavior of the logarithmic amplifier for small signal gain which transforms image values in a different way as the classical models do. Additionally, for recovery purposes, a method for the estimation of the compression parameters is also proposed when a realistic log-compressed law is considered. Results with synthetic images show that the proposed method achieved a consistent Rayleigh parameter estimate with a very low error. Experiments with real images show that the inversion method is consistent through the whole acquisition process when parameters of the logarithmic amplifier are assumed constant.

Adapted from: G. Vegas-Sánchez-Ferrero¹, D. Martín-Martínez¹, Pablo Casaseca-de-la-Higuera¹, Lucilio Cordero-Grande¹, S. Aja-Fernández¹, M. Martín-Fernández¹ and C. Palencia², *Realistic Log-Compressed Law for Ultrasound Image Recovery*, In 18th IEEE International Conference on Image Processing (ICIP), Brussels, Belgium, pp. 2029–2032, September 2011.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Departamento de Matemática Aplicada, Univ. Valladolid.

8.1 Introduction

Speckle in ultrasound images can be seen as a random process whose statistical features depend on the tissue class. The existence of a deterministic component due to specular reflections of the echo-pulse depends on the number of obstacles of the tissue (scatters) into the resolution cell and their size in comparison with the wavelength of ultrasound signal. Thus, the estimation of the probability density function (PDF) of different tissue classes can be practically applied in tissue filtering and segmentation, using maximum likelihood and maximum *a posteriori* algorithms [Sea10a, Sea10b, VSF10b, VSF10a]. These approaches are usually derived from the analysis of acoustic physics and the information available of the ultrasound probe. However, the whole acquisition information is not usually available and therefore suppositions must be made. For instance, images provided by practitioners usually do not include parameters as gain and/or contrast adjustment. Additionally, some of the steps of the acquisition process are not known depending on the commercial firm of the ultrasound machine.

Some approaches have been proposed in order to deal with this lack of information on PDFs. The most widespread method is to use empirical approximations for fitting speckle patterns accurately enough to provide good results for filtering or segmentation. This methodology has been used in [Tao06, Elt06] for different kind of distributions and extended to some more distributions in [VSF10b].

In this article we focus on the influence of the log-compression step on the probabilistic distributions of the speckle. This is a very important problem since many filtering and segmentation methods are based on PDF estimation and all the analysis of acoustic physics is lost in the pre-processing steps. Some methods has been presented to overcome this problem [Sea10b, Sea08, San03, Pra03]. These methods are based on the supposition that fully formed speckle can be modeled by a Rayleigh distribution. This is the case of large number of scatterers into the resolution cell and non-existence of deterministic component, which is a common accepted hypothesis which was first presented by Goodman in [Goo75].

The aforementioned works assume the following logarithmic law: $y = \alpha \log(1 + x) + \beta$, where α and β are the unknown parameters which respectively account for the contrast and brightness. When this transformation is performed on a Rayleigh distributed data, the resultant distribution becomes a Fisher-Typpet distribution (double exponential) [San03].

However, in real cases, the behavior of the log-compressed fully formed speckle regions is far from the Fisher-Typpet distribution and evidences that the hypothesis of the compression law should be reconsidered carefully (Fig. 8.1.(b)). For instance, the Fisher-Typpet tail for lower values distribution does not appear in real cases. This is probably due to the non-logarithmic behavior of the analog amplifiers for small voltage input. The non-logarithmic response of the amplifier, which specially affects to fully

Symbol	Parameter	Min.	Typ.	Max.
$R_g(k\Omega)$		0	3.3	∞
$V_{os}(mV)$	No input	-60	40	140
$V_{out}(mV)$	$V_{in} = 60\mu V$	45	80	115
	$V_{in} = 400\mu V$	200	245	290
	$V_{in} = 3mV$	365	440	495
	$V_{in} = 25mV$	530	610	690
	$V_{in} = 200mV$	680	780	880
	$V_{in} = 300mV$	710	820	930
	$V_{in} =$ $x \in (0, 60)\mu V$	$x \frac{45mV}{60\mu V}$	$x \frac{80mV}{60\mu V}$	$x \frac{115mV}{60\mu V}$

Table 8.1: Response of the True logarithmic amplifier TDA8780M for small and large signal.

formed speckle, produces lower values of the signal.

The main contribution of this work is a realistic log-compression law model for ultrasound images based on real amplifiers. The main advantage of this model is that the non-logarithmic behavior of the amplifier is considered and, to the best of our knowledge, no similar approach has been considered in the literature. Additionally, we present a method for estimating the parameters of the model in order to recover the pre-compressed image.

8.2 Log-compression Law Model

In this section we discuss the influence of the logarithmic compression on the probabilistic model when fully formed speckle regions are observed and a real logarithmic amplifier is considered. Specifically, the *True logarithmic amplifier* TDA8780M¹ is analyzed. It has a $72dB$ true logarithmic dynamic range which is large enough for the dynamic range of the input signal and the datasheet provides information for different values of gain and offset.

The logarithmic amplifier works as follows: the differential output from the true logarithmic amplifier is converted internally to a single-ended output in which the DC output level is set by an externally-supplied reference voltage. The gain adjustment can be performed by an off-chip resistor, R_g . The parameters provided by the datasheet are summarized in table 8.1.

In order to properly model the transfer characteristics of the amplifier, a continuous model is defined that guarantees uniform convergence of the values provided in the datasheet for $R_g \in [0, \infty)$. For this purpose, a least squares approximation is performed for the transfer characteristics for each

¹Available in <http://www.alldatasheet.com/datasheet-pdf>

value of R_g provided in the datasheet for large signal gain²

$$J = \sum_{i=1}^{|V_{in}|} (a_{ls}[R_g](V_{in}[i] - V_{in}[1]) + b_{ls}[R_g] - V_{out}[R_g, i])^2, \quad (8.1)$$

so, the transfer characteristics for large signal is the following³

$$\begin{aligned} V_{R_g, out} &= a_{ls}[R_g](V_{in}[i] - V_{in}[1]) + b_{ls}[R_g] \\ b_{ls}[R_g] &= V_{out}[R_g, 1] \\ a_{ls}[R_g] &= -\frac{\sum_{i=1}^{|V_{in}|} (b_{ls}[R_g] - V_{out}[R_g, i])(V_{in}[i] - V_{in}[1])}{\sum_{i=1}^{|V_{in}|} (V_{in}[i] - V_{in}[1])^2}. \end{aligned} \quad (8.2)$$

Now, the generalized continuous model is defined as

$$V_{out}(R_g, x) = \begin{cases} a_{sg}(R_g)x, & x \in [0, V_{max}]\mu V \\ a_{ls}(R_g)x^{dBm} + b_{ls}(R_g), & x^{dBm} \in (V_{max}^{dBm}, \infty), \end{cases} \quad (8.3)$$

where x^{dBm} is the input voltage in dBm which is calculated for $R_0 = 50\Omega$ as $x^{dBm} = 20 \log(x) - 10 \log(R_0) + 30$. $V_{max} = 60\mu V$ is the maximum voltage for small signal gain. The slope value for small signal gain, $a_{sg}(R_g)$, is defined as

$$\begin{aligned} a_{sg}(R_g) &= \frac{K_1}{K_2 + R_g} + K_3 \\ K_1 &= \frac{(a_{sg}[0] - a_{sg}[\infty])(a_{sg}[3.3k\Omega] - a_{sg}[\infty])}{(a_{sg}[0] - a_{sg}[3.3k\Omega])} 3.3k\Omega \\ K_2 &= \frac{(a_{sg}[0] - a_{sg}[\infty])}{(a_0 - a_{R0})} 3.3k\Omega \\ K_3 &= a_{sg}[\infty]. \end{aligned} \quad (8.4)$$

The slope of the large signal gain, $a_{ls}(R_g)$, and its offset $b_{ls}(R_g)$ are calculated in the same way

$$a_{ls}(R_g) = \frac{K_4}{K_5 + R_g} + K_6, \quad b_{ls}(R_g) = \frac{K_7}{K_8 + R_g} + K_9,$$

where the constants have the same expression as eqs. (8.4) calculated with $a_{ls}[R_g]$. The slope of small and large signal were defined in that way because is the most common way to adapt gains in analog circuits by voltage dividers.

The transfer characteristics of the Log-Compression Law Model are presented in Fig. 8.1.(a). Additionally, Fig. 8.1.(b) shows the histogram of the output voltage of the logarithmic amplifier calculated for a real image and it turns out that 12% of the values of the output signal are in the small signal gain area. This demonstrates that the compression law should be taken very carefully since many values of the image are incorrectly decompressed when ideal Log-Compression Laws, as the ones of [Sea10a, Sea08, San03, Pra03], are considered.

² $R_g = \{0, 3.3k\Omega, \infty\}$ and $V_{in} = \{60\mu V, 400\mu V, 3mV, 25mV, 200mV, 300mV\}$

³Note that discrete functions are represented with brackets whereas continuous functions use parenthesis.

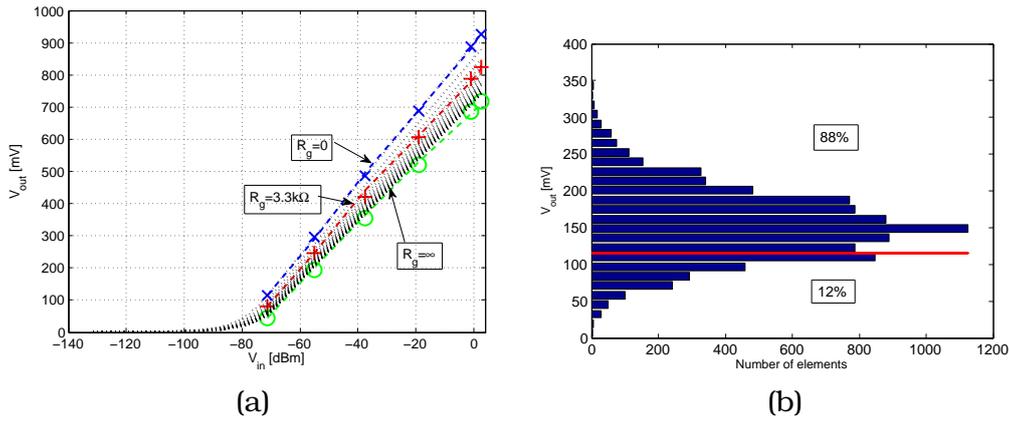


Figure 8.1: Transfer characteristics of the Log-Compression Law Model. (a) Continuous lines are the least squares approximation of large signal gain for each $R_g = \{0, 3.3k\Omega, \infty\}$. Black dotted lines are some examples of the continuous model for different values of R_g . (b) Histogram of a fully formed speckle area of a real image. 12% of the values are below the threshold of logarithmic compression.

8.3 Parameter Estimation

In this section a method for recovering the pre-compressed image is presented. Two values should be estimated to perform the inversion of the transfer function of the continuous model presented in section 8.2: R_g and V_{os} . However, V_{os} is the reference voltage and does not change the transfer function so we can assume that V_{os} is the minimum value of the output voltage before the analog-to-digital converter. Hence, V_{out} can be obtained in the following way

$$V_{out} = I \frac{DR}{255} - V_{os}, \quad (8.5)$$

where DR is the dynamic range of V_{out} (in our case $930 - (-60)mV$), I is the intensity value of the image and $V_{os} = \min(I \frac{DR}{255})$.

The inversion of the transfer characteristic function can be easily calculated by the following expression

$$f^{-1}(R_g, x) = \begin{cases} \frac{x}{a_{sg}(R_g)} & x \in [0, V_{max}a_{sg}(R_g)] \\ \frac{x - b_{gs}(R_g)}{a_{sg}(R_g)} - 30 + 10 \log(50) & \\ 10 \frac{a_{sg}(R_g)}{20} & x \in (V_{max}a_{sg}(R_g), \infty). \end{cases} \quad (8.6)$$

The inversion performed by this expression is showed in Fig. (8.2), where one can see the real image before interpolation (a), the pre-compressed image calculated for $R_g = 3.3k\Omega$ (b), the reconstructed image after the interpolation stage (c) and PDF of the fully formed speckle area selected in (a).

The estimation is performed by minimizing the uniform norm of the difference between the Cumulative Distribution Function (CDF), $F_X(x)$, for the Rayleigh distribution estimation and the empirical CDF, $E(x|R_g)$

$$\hat{R}_g = \arg \min_{R_g} \{ \|F_X(x) - E_X(x|R_g)\|_{\infty} \}. \quad (8.7)$$

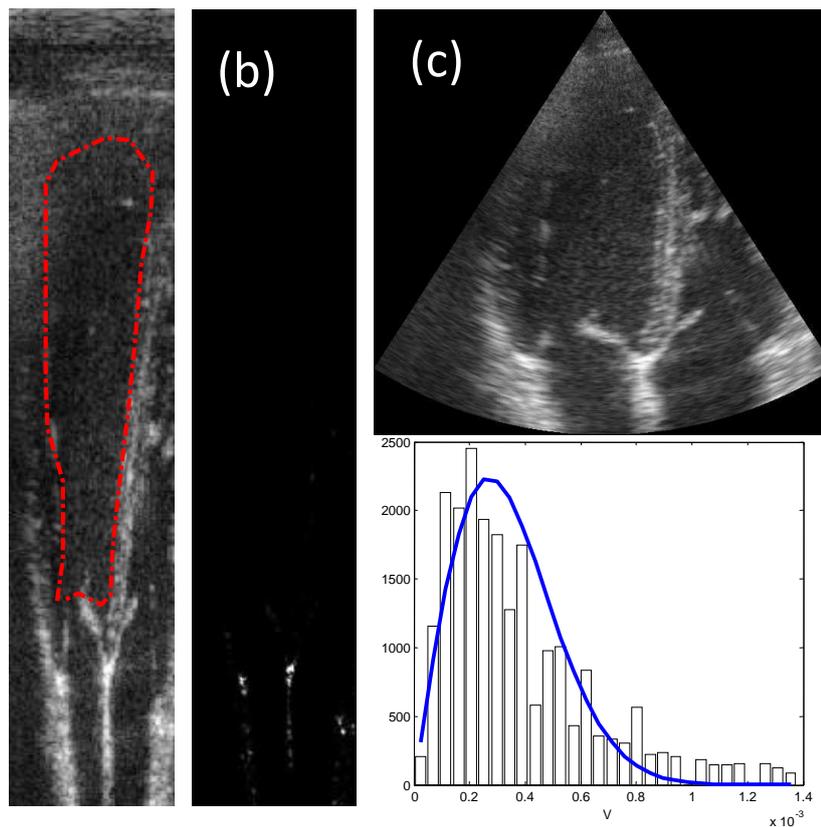


Figure 8.2: (a) Real image before the interpolation stage, (b) pre-compressed image ($R_g = 3.3k\Omega$), (c) Reconstructed image after the interpolation stage. (d) PDF of the fully formed speckle area selected in (a) in the pre-compressed image.

We decide to minimize this norm since other measures, such as χ^2 tests or Kullback-Leibler divergence, need a good PDF estimate. A proper PDF is difficult to obtain since the dynamic range of pre-compressed data is too large and the histogram analysis becomes unfeasible. Instead of parametric approaches, a Kolmogorov-Smirnov test could be used since it is a general nonparametric method for quantifying distance between the empirical CDF and the CDF of the reference distribution. However, the Kolmogorov-Smirnov test is biased when the parameter of the reference distribution is also estimated. This norm provides a similarity measure between distributions which can be calculated in a non-parametric way.

8.4 Results

Our experiments were performed by using a data bank of 574 images (584×145 , 8 bits) images obtained from 4 patients by means of a clinical machine *GE Vivid 7 echographic system (GE Vingmed Ultrasound A.S., Horten, Norway)*. The images were obtained before the interpolation stage of the acquisition process. All the fully formed speckle areas of the images were

manually segmented⁴.

In this section we present two different experiments. First, we test the performance of the estimators for synthetic images that are generated as random Rayleigh distributed data with parameter σ . In order to choose the σ parameter similar to real cases, the maximum and minimum mean value (μ_{min} and μ_{max} respectively) of the fully formed speckle areas is calculated for all the real images. The pre-compressed image is obtained for both extreme values $R_g = 0$ and $R_g = \infty$. This way, the dynamic range of σ can be estimated just knowing that the mean value of a Rayleigh random variable is $E\{X\} = \sigma\sqrt{\frac{\pi}{2}}$. Hence, dynamic range is $\sigma \in \left[\mu_{min}\sqrt{\frac{2}{\pi}}, \mu_{max}\sqrt{\frac{2}{\pi}}\right]$. In our case is $[1.5 \cdot 10^{-5}, 8.7 \cdot 10^{-4}]$.

In Fig. 8.3.(a), the relative error of σ , $\epsilon_\sigma = \frac{|\sigma - \hat{\sigma}|}{\sigma}$, is represented for $R_g \in [0, 1000]k\Omega$ and σ . Fig 8.3.(b) shows the uniform norm $\|F_X(x) - E_X(x|R_g)\|_\infty$. Note that the maximum relative error in the estimate is under 9%, which is a very good estimation when we compare with the original PDF (see some examples in Fig. 8.3.(c)). Additionally, the difference between empirical and theoretical CDF is below 0.0049 which evidences the accuracy of the inversion. In Fig 8.3.(d) the uniform norm is represented for the inverted compression when the non-logarithmic regime is considered. Note that the maximum error (0.1290) is 26 times higher than the obtained with the realistic compression and the minimum error is 0.0012, still higher. The importance of a realistic model for log-compressed data should remain evident in the light of this result.

In the second experiment, all the data bank of real images was decompressed with the proposed method in order to see the consistence of the parameters of the logarithmic amplifier through the whole acquisition. This validation methodology was chosen since we have no pre-compressed data yet, future work will deal with pre-compressed and compressed images. It is supposed that the logarithmic amplifier parameters are constant through the whole acquisition, though may be different for each patient. Results for each patient are shown in table 8.2. There one can see that the low deviations of R_g and σ show stable values in the dynamic range of each value $[0, \infty)$ and $[1.5 \cdot 10^{-5}, 8.7 \cdot 10^{-4}]$, respectively. This confirms that the estimation method is consistent.

8.5 Conclusions

In this work we propose a realistic log-compressed law for ultrasound images based on a real device. This is the main contribution of this work and, to the best of our knowledge, no similar approach has been considered in the literature for this purpose.

⁴The authors would like to thank Marta Sitges, Etelvino Silva (Hospital Clinic; IDIBAPS; Universitat de Barcelona, Spain), Bart Bijmens (Instituto Catalana de Recerca i Estudis Avançats (ICREA) Spain) Nicolas Duchateau (CISTIB - Universitat Pompeu Fabra, CIBERBBN, Barcelona, Spain) for providing the images

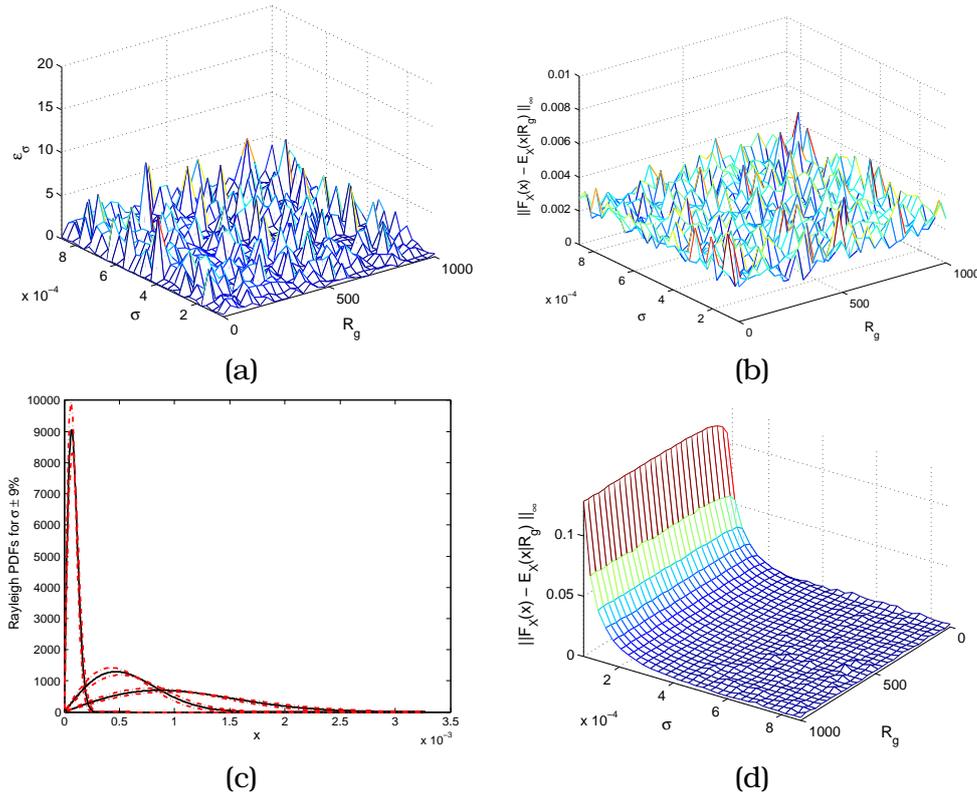


Figure 8.3: (a) Relative error of σ . (b) uniform norm for the realistic log-compression law. (c) Some examples of Rayleigh PDFs for σ (black solid lines) and $\sigma \pm 9\%$ (dashed lines). (d) uniform norm for the ideal log-compression law.

Patient	R_g	$\sigma \cdot 10^{-5}$	$\ F_X - F_E\ _\infty$
1	$3.43 \pm 2.98k\Omega$	2.40 ± 0.343	0.09 ± 0.03
2	$0.09 \pm 0.03k\Omega$	1.80 ± 0.168	0.17 ± 0.02
3	$0.17 \pm 0.02k\Omega$	1.88 ± 0.222	0.13 ± 0.02
4	$0.13 \pm 0.02k\Omega$	2.80 ± 0.262	0.05 ± 0.01

Table 8.2: Results for de-compressed real images.

Additionally, a method for the estimation of the compression parameters is proposed for recovery purposes. This method is based on the minimization of the uniform norm of the difference of the empirical CDF after de-compressing the image, and the estimated Rayleigh CDF.

Results with synthetic images showed that the proposed method achieved a proper Rayleigh parameter estimate with low errors. Additionally, the error committed when the realistic model is not taken into consideration demonstrates that the inverse transformation obtains very different PDFs. Experiments with real images showed that the inversion method is consistent through the whole acquisition process when parameters of the logarithmic amplifier are supposed constant through the process.

Anisotropic LMMSE denoising of MRI based on statistical tissue models

Nature appears not to have intended that any flower should be fertilized by its own pollen.
Christian Konrad Sprengel, 1750–1816.

Abstract- Linear Minimum Mean Squared Error Estimation is a simple, yet powerful denoising technique within MRI. It is based on the computation of the mean and variance of the data being filtered according to a noise model assumed, which is usually accomplished by calculating local moments over squared neighborhoods. When these neighborhoods are centered in pixels corresponding to image contours, the estimation is not accurate due to the presence of two or more tissues with different statistical properties. We overcome this limitation by introducing an anisotropic Linear Minimum Mean Squared Error Estimation (LMMSE) scheme: the gray levels of each tissue in the MRI volume are modeled as a Gamma-mixture, such that we can discriminate between the different matters to construct anisotropic neighborhoods containing only one kind of tissue. The potential of the Gamma distribution relies on its ability to fit both the Rician distribution traditionally used to model the noise in MRI and the non-central Chi noise found in modern parallel MRI systems.

Adapted from: G. Vegas-Sánchez-Ferrero¹, A. Tristán-Vega¹, S. Aja-Fernández¹, M. Martín-Fernández¹, C. Palencia² and R. Deriche³, *Anisotropic LMMSE Denoising of MRI Based on Statistical Tissue Models*, In IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), Barcelona, Spain, pp. 1519–1522, April 2012.

¹Laboratorio de Procesado de Imagen, Univ. Valladolid.

²Departamento de Matemática Aplicada, Univ. Valladolid.

³Athena Project Team, INRIA Sophia Antipolis, France

9.1 Introduction

Magnetic Resonance Imaging (MRI) is one of the medical imaging modalities that exhibits the largest Signal to Noise Ratio (SNR). Even so, noise-induced artifacts in MRI are not negligible, degrading image quality in both qualitative and quantitative applications. On one hand, several new acquisition technologies have appeared lately, aiming to speed-up the acquisition of MRI, and to reduce phase distortions originated by the strong sensitizing gradients used in Diffusion Weighted Imaging (DWI). These methods are commonly referred to as Parallel Magnetic Resonance Imaging (pMRI), and are based on the simultaneous acquisition of different parts of the sub-sampled k -space by different receiving coils. A well-known side effect of this methodology is the worsening of the final SNR, which is typically modeled through the g -factor [AF11]. On the other hand, imaging modalities such as DWI demand much larger amounts of data compared to anatomical MRI, and are also noisier. As a consequence, direct procedures for boosting the SNR in MRI, mainly the averaging of multiple independent repetitions of the volume, are no longer practical.

Given these concerns, the need for *a posteriori* denoising schemes have become clear to many authors. General purpose and Gaussian-based filters have been used, though Rician-adapted techniques have been proven more adequate for quantitative pMRI. They include the conventional approach [McG93], Rician Maximum Likelihood estimation [Sij04], Expectation-Maximization (EM) [Mar95], wavelets [Now99], anisotropic diffusion [Kri09b], and Unbiased Non-Local Means [WD07].

Among these approaches, the LMMSE estimator for Rician data first described in [AF08a] has shown to outperform the most common techniques. This method is a classical estimation theory approach: provided we can characterize the expected value and covariance matrix of both the signal to estimate and the noise, it obtains the best linear estimation (i.e. that with minimum variance) of the data. These parameters may be easily inferred from local moments computed as sample estimates inside image neighborhoods. Besides, the linear nature of the filter allows a very fast implementation, and moreover it may be adapted to cope with noise models deviating from Rician, mainly the non-central Chi ($nc-\chi$) [Bri11], arising with certain modern machinery and pMRI protocols [AF11].

An important limitation of LMMSE models relies on the way the local moments are estimated. For those voxels corresponding to image edges, the neighborhoods used typically comprise several different kinds of tissues. Hence, the value estimated for the variance is artificially increased due to the superposition of the effect of noise and the multi-modal distribution of gray levels corresponding to each matter. As a consequence, edge voxels are not properly denoised, since the noisy pattern is mixed-up with the fine details (gradients).

In this work we propose to estimate the probabilistic response of each tissue by Gamma distributions for pMRI protocols. These distributions pro-

vide a good fitting as was demonstrated by Patnaik in [Pat49]. With this estimation, the probability of belonging to each tissue class can be obtained and, thus, the local moments used for the calculation of the LMMSE can be derived as the local moments conditioned to each of the tissues. As a consequence, the estimation of local variance is not biased by the presence of different tissues because only those pixels belonging to each tissue class contribute to the calculation of moments resulting in an anisotropic denoising. Results demonstrate that this approach outperforms isotropic LMMSE methods [Bri11].

9.2 Methods

9.2.1 Background

In the case of multiple-coil acquisition, the composite magnitude image can be obtained using Sums of Squares (SoS) as $M_L = \sqrt{\sum_{l=1}^L |S_l|^2}$, where each coil signal is modeled as $S_l = A_l + N_l(\sigma_n^2)$ where N_l is a complex Gaussian process with zero mean and variance σ_n^2 . Defining $A_T = \sqrt{\sum_{l=1}^L |A_l|^2}$ and assuming the noise components are independent and identically distributed (IID), the magnitude signal M_L follows a nc- χ distribution [Bri11]. When the square of the magnitude signal is considered, M_L^2 , the PDF trivially becomes a nc- χ^2 square distribution (nc- χ^2).

The LMMSE filter in [AF08a] and the extended version of [Bri11] rely on the following expression

$$\widehat{A_T^2} = E\{A_T^2\} + \mathbf{C}_{A_T^2 M_L^2} \mathbf{C}_{M_L^2 M_L^2}^{-1} (M_L^2 - E\{M_L^2\}), \quad (9.1)$$

where $\mathbf{C}_{M_L^2 M_L^2}$ is the covariance of M_L^2 and $\mathbf{C}_{A_T^2 M_L^2}$ is the cross-covariance. The Characteristic function of a nc- χ^2 is

$$\varphi_{M_L^2}(t) = E\{e^{itM_L^2}\} = \frac{1}{(1 - 2it\sigma_n^2)^L} e^{\frac{iA_T^2 t}{1 - 2it\sigma_n^2}}. \quad (9.2)$$

Note that this result gives rise to a Gamma distribution in the absence of the original signal ($A_T = 0$). The PDF, f_X , and Characteristic function, φ_X , of a Gamma distributed random variable, X , of parameters α and β are

$$f_X(x|\alpha, \beta) = \frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-\frac{x}{\beta}} u(x), \quad \varphi_X(t) = \frac{1}{(1 - i\beta t)^\alpha}, \quad (9.3)$$

where $\alpha, \beta > 0$; $\Gamma(x)$ is the Euler Gamma function.

This result was noticed by Patnaik in [Pat49], where the approximation of nc- χ^2 by a Gamma distribution, was firstly proposed. As an example, in Fig. 9.1.a the Gamma PDF approximation is depicted for a set of synthetic images of size 512×512 obtained by SoS with incremental $A_T \in [0, 100]$,

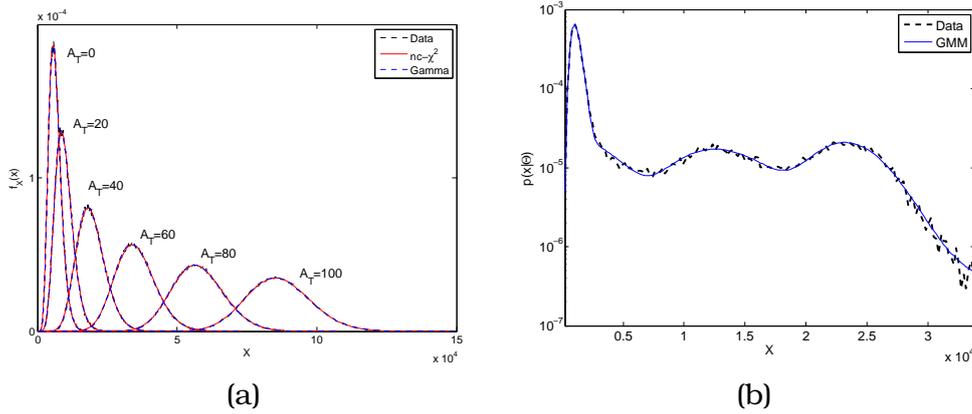


Figure 9.1: (a) M_L^2 PDF for $A_T \in [0, 100]$, $L = 8$, and $\sigma_n = 20$. (b) Empirical PDF of a pMRI T1 image ($L = 8$ and $\sigma_n = 70$) and the proposed GMM for 5 elements in the mixture.

number of coils $L = 8$, and $\sigma_n = 20$. Note that the theoretical PDF of the $nc-\chi^2$ and the Gamma approach are almost indistinguishable and both follow the empirical PDF of M_L^2 .

This result shows the potential of the Gamma distribution for fitting the response of both the Rician distributed noise traditionally used to model conventional MRI and $nc-\chi$ distributed noise found in modern parallel MRI systems. We propose to estimate the probabilistic response of each tissue by a Gamma distributions. This way, the local moments required for the LMMSE can be calculated as the local moments conditioned to each of the tissues. In order to find the distributions of different tissues of the image, a Gamma mixture model is proposed and explained in the next section.

9.2.2 Gamma Mixture Model

The mixture of Gamma distributions (GMM) is calculated by means the EM algorithm [Mar95] in the following way¹: Let $\mathbf{X} = \{x_i\}$, $1 \leq i \leq N$ be a IID set of samples (pixel intensities of M_L^2). The GMM considers that these variables result from the contributions of J distributions: $p(x_i|\Theta) = \sum_{j=1}^J \pi_j f_X(x_i|\Theta_j)$, where Θ is a vector of the parameters of the GMM (π_j, Θ_j) and Θ_j are the parameters of the PDF, α_j and β_j and $\sum_{j=1}^J \pi_j = 1$.

The method is applied in the following way:

1. A first estimate of the hidden variables is obtained by k-means. For each cluster, j parameters $\Theta_j^{(0)}$ and $\hat{\pi}_j^{(0)}$, are calculated from samples.
2. Bayesian Inference Step. Calculate $\gamma_{i,j}$ as

$$p(Z_i = j|x_i, \Theta^{(n)}) = \gamma_{i,j} = \frac{\pi_j^{(n)} p(x_i|\Theta_j^{(n)})}{p(x_i|\Theta^{(n)})}. \quad (9.4)$$

¹This method maximizes the log-likelihood function when some hidden discrete random variables (RVs), $\mathbf{Z} = \{Z_i\}$, are introduced to the model. These RVs take values in $\{1, \dots, J\}$, their meaning is that the sample x_i belongs to the distributions class j when $Z_i = j$.

3. Expectation-Maximization Step. The estimate for $\hat{\alpha}_j$, $\hat{\beta}_j$ and $\hat{\pi}_j$ are calculated by the following equations

$$\log(\hat{\alpha}_j) - \psi(\hat{\alpha}_j) = \log\left(\frac{\sum_i^N \gamma_{i,j} x_i}{\sum_i^N \gamma_{i,j}}\right) - \frac{\sum_i^N \gamma_{i,j} \log x_i}{\sum_i^N \gamma_{i,j}} \quad (9.5)$$

$$\hat{\pi}_j = \frac{1}{N} \sum_{i=1}^N \gamma_{i,j}, \quad \hat{\beta}_j = \frac{1}{\hat{\alpha}_j} \frac{\sum_{i=1}^N \gamma_{i,j} x_i}{\sum_{i=1}^N \gamma_{i,j}}. \quad (9.6)$$

4. Go to step 2 until $\|\Theta^{(n)} - \Theta^{(n-1)}\| < \text{Tolerance}$.

As an example of the GMM, in Fig. 9.1.b the GMM is fitted to a pMRI T1 of Fig. 9.3 acquired with $L = 8$ and $\sigma_n = 70$ and the proposed GMM for 5 elements in the mixture.

9.2.3 Anisotropic-LMMSE

In [Bri11] it is shown that for uncorrelated and independent Gaussian noises the variance and covariance terms, $C_{M_L^2, M_L^2}$ and $C_{A_T^2, M_L^2}$ can be calculated as

$$C_{M_L^2, M_L^2} = E\{M_L^4\} - E\{M_L^2\}^2 \quad (9.7)$$

$$C_{A_T^2, M_L^2} = E\{A_T^4\} + 2L\sigma_n^2 E\{A_T^2\} - E\{A_T^2\}E\{M_L^2\}, \quad (9.8)$$

where the 2nd and 4th moments of A_T can be derived by the calculation of the same order moments of M_L^2 which is a nc- χ^2 distribution

$$E\{A_T^4\} = E\{M_L^4\} - 4(E\{M_L^2\} - 2\sigma_n^2)(L+1)\sigma_n^2 - 4L(L+1)\sigma_n^4 \quad (9.9)$$

$$E\{A_T^2\} = E\{M_L^2\} - 2L\sigma_n^2. \quad (9.10)$$

These moments can be calculated by assuming local ergodicity as was done in [AF08a, Bri11]. Hence, the expectation, $E\{\cdot\}$ can be replaced by local spatial mean, $\langle \cdot \rangle$, calculated in a neighborhood. Now, the LMMSE can be calculated just as a function of the local moments of M_L^2 , $\langle M_L^2 \rangle$ and $\langle M_L^4 \rangle$.

The calculation of the components for the mixture model provides important information related to the probability of belonging to each tissue class. This information can be taken into account by calculating the conditioned mean for each tissue class in the local neighborhood, $\eta(x, y)$. When the moments are calculated in this way, only those pixels belonging to each tissue class contribute to the estimate of each local moment, resulting in a more accurate estimate of local moments. This result can be considered as a non isotropic way to calculate the local moments since each pixel is not equally treated when the local moments are calculated.

By introducing the probability of belonging to each tissue class in the calculation of the local moments, they are calculated as

$$\langle M_L^2(x, y) \rangle = \sum_{j=1}^J P(Z = j) \langle M_L^2(x, y) | Z = j \rangle, \quad (9.11)$$

where the local moment for each class is

$$\langle M_L^2 | Z = j \rangle = \frac{\sum_{(x',y') \in \eta(x,y)} M_L^2(x',y') p(Z(x',y') = j | M_L^2, \Theta^{(n)})}{\sum_{(x',y') \in \eta(x,y)} p(Z(x',y') = j | M_L^2, \Theta^{(n)})}, \quad (9.12)$$

with

$$p(Z(x',y') = j | M_L^2, \Theta^{(n)}) = \frac{\pi_j \cdot p(M_L^2 | \Theta_j^{(n)})}{p(M_L^2 | \Theta^{(n)})}. \quad (9.13)$$

The weight $P(Z = j)$, is the probability of belonging to the j -th tissue class for the neighborhood of (x, y) and is calculated as

$$P(Z(x, y) = j) = \frac{\prod_{(x',y') \in \eta(x,y)} p(Z(x',y') = j | M_L^2, \Theta^{(n)})}{\sum_{j=1}^J \prod_{(x',y') \in \eta(x,y)} p(Z(x',y') = j | M_L^2, \Theta^{(n)})}. \quad (9.14)$$

9.3 Results

For the sake of validation, two experiments are considered to test the performance of the anisotropic LMMSE in comparison to the isotropic nc- χ LMMSE [Bri11]. In the first one, 3D synthetic T1 and T2 images from Brain-Web MR [Col98] with intensity values in $[0 - 255]$ and 1 mm Slice thickness were considered. A simulation of 8-coil system was performed for an increasing $\sigma_n \in [10, 100]$ in each coil [AF11]. The image was reconstructed by using SoS. A neighborhood of $5 \times 5 \times 5$ for both the proposed and the nc- χ LMMSE filter in [Bri11] was used. The number of elements in the GMM was 5.

Results for the experiment are shown in Fig. 9.3, in the top row, the T1 image is shown as well as the noisy image (Fig. 9.3.b) for $\sigma_n = 70$. The denoised image obtained by the isotropic nc- χ LMMSE is shown in Fig. 9.3.c. Note that this filter retains noise in regions with the presence of several tissues. the proposed method is shown in Fig. 9.3.d where the better performance in areas with more than one tissue can be appreciated at first sight. An analogous conclusion can be obtained for T2 images represented in the lower row of Fig. 9.3.

The quantitative evaluation was performed using the mean squared error (MSE) and the mean structural similarity² (MSSIM) [Wan04]. Twenty independent trials were considered for each value of σ_n . The mean of both measures is represented in columns Fig. 9.3.e-f for T1 and T2.

Though the proposed method outperforms the isotropic one for both measures, for the MSE both methods do not exhibit a big difference. This is so because homogeneous areas are equally treated by both methods and

²This gives a similarity value in $[0, 1]$.

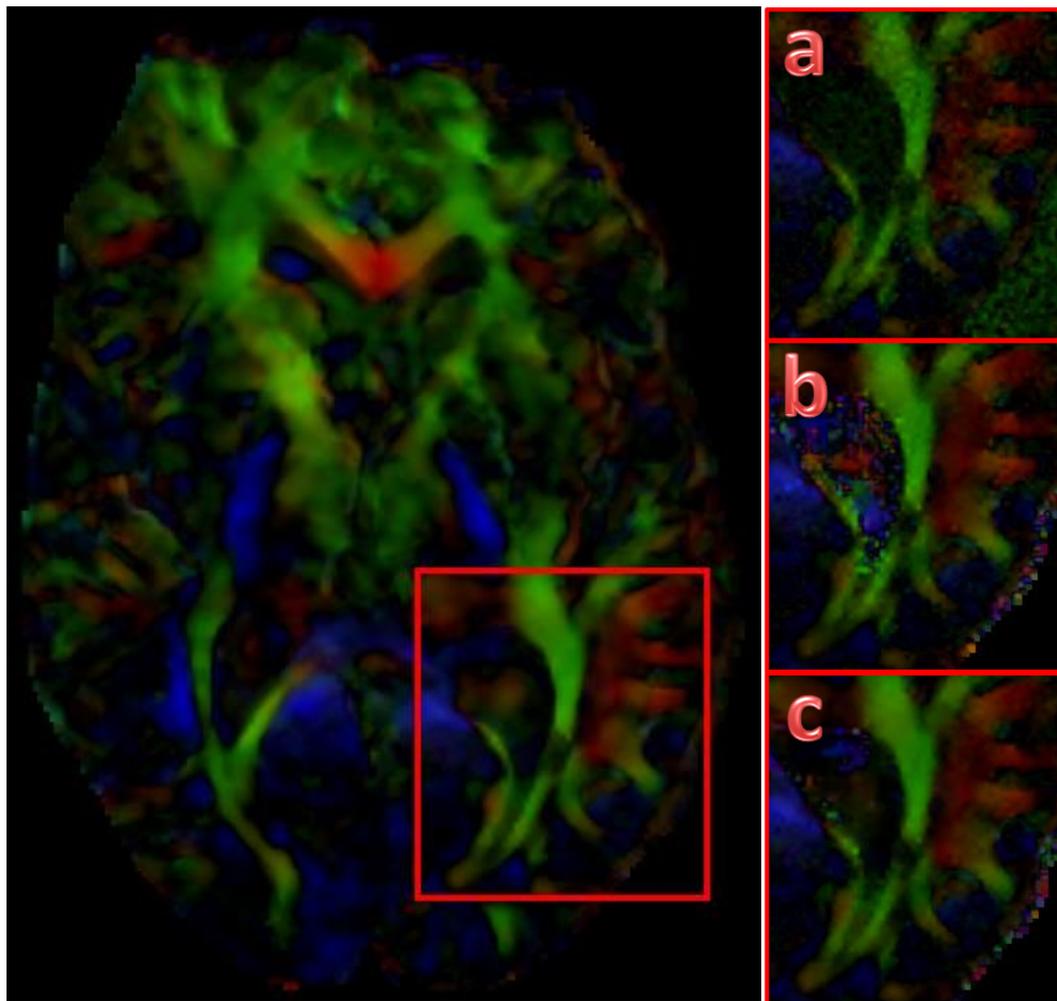


Figure 9.2: (a) Noisy (SNR=15db) (b) $nc\text{-}\chi$ LMMSE (c) Proposed.

pixels in the edge of different tissues are not explicitly reflected by the MSE. Nevertheless, for the MSSIM measure, the better performance of the anisotropic filter becomes evident as we increase σ_n . This result confirms the ability for detecting the structures of the image leading to a better estimate of local moments and, thus, overcoming the main problem of LMMSE methods.

For the second experiment, a 15 gradients DWI phantom presented in [TV09] was considered. A simulation of 8-coil system was performed for a SNR=15 dB. In Fig. 9.2 the fractional anisotropy colored by orientation is represented. The zoomed area shows a better definition of areas with different orientations. Quantitatively, the MSE was calculated for the mean FA; noisy: 0.0271, $nc\text{-}\chi$ LMMSE: 0.0124, and the proposed: 0.0055. Additionally, the MSE of the trace of the diffusion tensor gives the following results; noisy: $7.38 \cdot 10^{-6}$, $nc\text{-}\chi$ LMMSE: $8.23 \cdot 10^{-7}$, and the proposed: $5.7 \cdot 10^{-7}$. Both measures show the better performance of the proposed method for the calculation of the diffusion tensor.

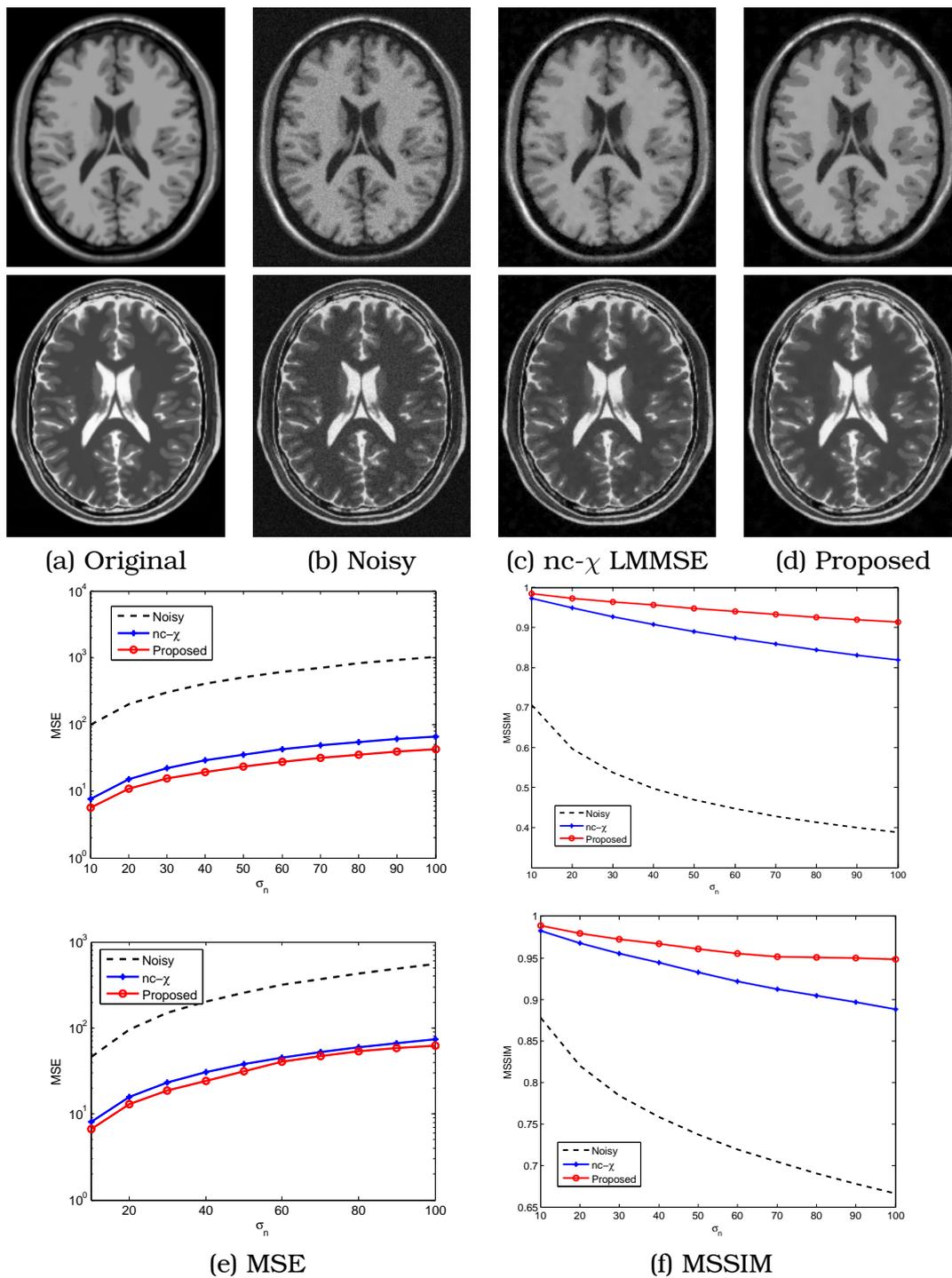


Figure 9.3: Quantitative comparison with synthetic T1 (first row) and T2 (second row) for increasing σ_n .

9.4 Conclusions

An anisotropic LMMSE method is proposed. The anisotropy is achieved by applying statistical models to the tissues based on a GMM model. This mixture have proved to properly fit the probabilistic behavior of noise in

conventional MRI (Rician) and pMRI ($nc-\chi$). Experiments have shown that the inclusion of statistical models lead to a better estimate of local moments by considering the local moments conditioned to the probability of belonging to each tissue class.

Bibliography

- [AF06] S. Aja-Fernandez and C. Alberola-Lopez. On the Estimation of the Coefficient of Variation for Anisotropic Diffusion Speckle Filtering. *IEEE Transactions on Image Processing*, 15(9):2694–2701, 2006.
- [AF08a] S. Aja-Fernandez, C. Alberola-Lopez, and C.-F. Westin. Noise and Signal Estimation in Magnitude MRI and Rician Distributed Images: A LMMSE Approach. *IEEE Transactions on Image Processing*, 17(8):1383–1398, 2008.
- [AF08b] S. Aja-Fernandez, M. Niethammer, M. Kubicki, M. E. Shenton, and C.-F. Westin. Restoration of DWI Data using a Rician LMMSE Estimator. *IEEE Transactions on Medical Imaging*, 27(10):1389–1403, 2008.
- [AF10] S. Aja-Fernandez, G. Vegas-Sanchez-Ferrero, and A. Tristan-Vega. About the Background Distribution in MR Data: A Local Variance Study. *Magnetic Resonance Imaging*, 28(5):739–752, 2010.
- [AF11] S. Aja-Fernandez, A. Tristan-Vega, and S. Hoge. Statistical Noise Analysis in GRAPPA Using a Parametrized Non-Central Chi Approximation Model. *Magnetic Resonance in Medicine*, 65(4):1195–1206, 2011.
- [Aho74] A. V. Aho and J. E. Hopcroft. *The Design and Analysis of Computer Algorithms*. Addison-Wesley Longman Publishing Co. Inc., Boston, MA, USA, 1974.
- [Ali08] M. Ali, D. Magee, and U. Dasgup. Signal Processing Overview of Ultrasound Systems for Medical Imaging. Tech. rep., Texas Instruments, 2008.
- [Alz97] H. Alzer. On Some Inequalities for the Gamma and Psi Functions. *Mathematics of Computation*, 66(217):373–389, 1997.
- [And76] G. E. Andrews. *The Theory of Partitions, Encyclopedia of Mathematics and Its Applications*. Addison-Wesley, Boston, MA, USA, 1976.
- [And08] J. L. Andersson. Maximum *a posteriori* Estimation of Diffusion Tensor Parameters using a Rician Noise Model: Why, How and But. *NeuroImage*, 42(4):1340–1356, 2008.
- [Ata11] S. Atapattu, C. Tellambura, and H. Jiang. A Mixture Gamma Distribution to Model the SNR of Wireless Channels. *IEEE Transactions on Wireless Communications*, 10(12):4193–4203, 2011.

- [Bea90] N. C. Beaulieu. An Infinite Series for the Computation of the Complementary Probability Distribution Function of a Sum of Independent Random Variables and its Application to the Sum of Rayleigh Random Variables. *IEEE Transactions on Communications*, 38(9):1463–1474, 1990.
- [Bli98] S. Blinnikov and R. Moessner. Expansions for Nearly Gaussian Distributions. *Astronomy and Astrophysics Supplement Series*, 130(1):193–205, 1998.
- [Bre73] R. P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1973.
- [Bri11] V. Brion, C. Poupon, O. Riff, S. Aja-Fernandez, A. Tristan-Vega, J.-F. Mangin, D. L. Bihan, and F. Poupon. Parallel MRI Noise Correction: An Extension of the LMMSE to Non Central Distributions. In G. Fichtinger, A. Martel, and T. Peters, eds., *14th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, vol. 6892 of *Lecture Notes in Computer Science*, pp. 226–233. Toronto, ON, Canada, September 2011.
- [Bru91] M. E. Brummer, R. M. Mersereau, R. L. Eisner, and R. R. J. Lewine. Automatic Detection of Brain Contours in MRI Data Sets. In A. C. F. Colchester and D. J. Hawkes, eds., *12th International Conference on Information Processing in Medical Imaging (IPMI)*, vol. 511 of *Lecture Notes in Computer Science*, pp. 188–204. Wye, UK, July 1991.
- [Bur78] C. B. Burckhardt. Speckle in Ultrasound B-Mode Scans. *IEEE Transactions on Sonics and Ultrasonics*, 25(1):1–6, 1978.
- [Cho04] E. Cho and M. J. Cho. The Variance of Sample Variance from a Finite Population. In *Joint Statistical Meeting*, pp. 3345–3350. Toronto, ON, Canada, August 2004.
- [Cho05] E. Cho and M. J. Cho. The Variance of Sample Variance from a Finite Population. *The International Journal of Pure and Applied Mathematics*, 21(3):389–396, 2005.
- [Cio10] F. Ciompi, O. Pujol, C. Gatta, O. Rodriguez-Leor, J. Mauri-Ferre, and P. Radeva. Fusing *in-vitro* and *in-vivo* Intravascular Ultrasound Data for Plaque Characterization. *International Journal on Cardiovasc Imaging*, 26(7):763–779, 2010.
- [Clo11] G. Cloutier and F. Destrempes. Response to the Letter to the Editor-in-Chief on Manuscript Entitled: “A Critical Review and Uniformized Representation of Statistical Distributions Modeling the Ultrasound Echo Envelope”. *Ultrasound in Medicine and Biology*, 37(4):675–676, 2011.
- [Col98] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and Construction of a Realistic Digital Brain Phantom. *IEEE Transactions on Medical Imaging*, 17(3):463–468, 1998.
- [Cop03] K. Copsey and A. Webb. Bayesian Gamma Mixture Model Approach to Radar Target Recognition. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1201–1217, 2003.
- [Dav49] F. N. David and M. G. Kendall. Tables of Symmetric Functions – Part I. *Biometrika*, 36(3/4):431–449, 1949.
- [Dem77] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society: Series B*, 39(1):1–38, 1977.

- [Dem06] J. Demar. Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research*, 7:1–30, 2006.
- [Des09] F. Destrempes, J. Meunier, M. F. Giroux, G. Soulez, and G. Cloutier. Segmentation in Ultrasonic B-Mode Images of Healthy Carotid Arteries Using Mixtures of Nakagami Distributions and Stochastic Optimization. *IEEE Transactions on Medical Imaging*, 28(2):215–229, 2009.
- [Des10] F. Destrempes and G. Cloutier. A Critical Review and Uniformized Representation of Statistical Distributions Modeling the Ultrasound Echo Envelope. *Ultrasound in Medicine and Biology*, 36(7):1037–1051, 2010.
- [Des11] F. Destrempes, J. Meunier, M. F. Giroux, G. Soulez, and G. Cloutier. Segmentation of Plaques in Sequences of Ultrasonic B-Mode Images of Carotid Arteries Based on Motion Estimation and a Bayesian Model. *IEEE Transactions on Biomedical Engineering*, 58(8):2202–2211, 2011.
- [Die95] T. G. Dietterich and G. Bakiri. Solving Multiclass Learning Problems via Error-Correcting Output Codes. *Journal of Artificial Intelligence Research*, 2(1):263–286, 1995.
- [Dud99] R. M. Dudley. *Uniform Central Limit Theorems*. Cambridge University Press, Cambridge, UK, first ed., 1999.
- [Dut94] V. Dutt and J. F. Greenleaf. Ultrasound Echo Envelope Analysis using a Homodyned-K Distribution Signal Model. *Ultrasonic Imaging*, 16(4):265–287, 1994.
- [EB95] N. M. El-Barghouty, G. Geroulakos, A. Nicolaidis, A. Androulakis, and V. Bahal. Computer-Assisted Carotid Plaque Characterisation. *European Journal of Vascular and Endovascular Surgery*, 9(4):389–393, 1995.
- [EB96] N. M. El-Barghouty, T. Levine, S. Ladva, A. Flanagan, and A. Nicolaidis. Histological Verification of Computerised Carotid Plaque Characterisation. *European Journal of Vascular and Endovascular Surgery*, 11(4):414–416, 1996.
- [Elt06] T. Eltoft. Modeling the Amplitude Statistics of Ultrasonic Images. *IEEE Transactions on Medical Imaging*, 25(2):229–240, 2006.
- [Esc06] E. Escolar, G. Weigold, A. Fuisz, and N. J. Weissman. New Imaging Techniques for Diagnosing Coronary Artery Disease. *Canadian Medical Association Journal*, 174(4):487–95, 2006.
- [Fig02] M. A. T. Figueiredo and A. K. Jain. Unsupervised Learning of Finite Mixture Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):381–396, 2002.
- [Gom08] O. Gomes, C. Combes, and A. Dussauchoy. Parameter Estimation of the Generalized Gamma Distribution. *Mathematics and Computers in Simulation*, 79(4):955–963, 2008.
- [Goo75] J. W. Goodman. *Laser Speckle and Related Phenomena*, vol. 9-75 of *Topics in Applied Physics*, chap. Some fundamental properties of laser speckle, pp. 1145–1150. Springer, Heidelberg, Germany, 1975.
- [Gud95] H. Gudbjartsson and S. Patz. The Rician Distribution of Noisy MRI Data. *Magnetic Resonance in Medicine*, 34(6):910–914, 1995.
- [Hag70] H. W. Hager and L. J. Bain. Inferential Procedures for the Generalized Gamma Distribution. *Journal of the American Statistical Association*, 65(332):1601–1609, 1970.

- [Har90] W. Hardle. *Smoothing Techniques: With Implementation in S*. Springer Series in Statistics. Springer, Heilderberg, Germany, 1990.
- [Har92] M. D. Harpen. Noise Correlations Exist for Independent RF Coils. *Magnetic Resonance in Medicine*, 23(2):394–397, 1992.
- [Hay90] C. E. Hayes and P. B. Roemer. Noise Correlations in Data Simultaneously Acquired from Multiple Surface Coil Arrays. *Magnetic Resonance in Medicine*, 16(2):181–191, 1990.
- [Hen85] R. M. Henkelman. Measurement of Signal Intensities in the Presence of Noise in MR Images. *Medical Physics*, 12(2):232–233, 1985.
- [Jak80] E. Jakeman. On the Statistics of K-Distributed Noise. *Journal of Physics A: Mathematical and General*, 13(1):31–48, 1980.
- [Jak84] E. Jakeman. Speckle Statistics with a Small Number of Scatterers. *Optical Engineering*, 23(4):453–461, 1984.
- [Jak87] E. Jakeman and R. J. A. Tough. Generalized K Distribution: A Statistical Model for Weak Scattering. *Journal of the Optical Society of America A*, 4(9):1764–1772, 1987.
- [Jen91] J. A. Jensen. A Model for the Propagation and Scattering of Ultrasound in Tissue. *The Journal of the Acoustical Society of America*, 89(1):182–190, 1991.
- [Jen96] J. A. Jensen. FIELD: A Program for Simulating Ultrasound Systems. In *10th Nordicbaltic Conference on Biomedical Imaging*, pp. 351–353. Tampere, Finland, June 1996.
- [Jon78] B. Jones, W. G. Waller, and A. Feldman. Root Isolation Using Function Values. *BIT Numerical Mathematics*, 18(3):311–319, 1978.
- [Knu97] D. E. Knuth. *The Art of Computer Programming*, vol. 1: Fundamental Algorithms. Addison Wesley Longman Publishing Co. Inc., Redwood City, CA, USA, third ed., 1997.
- [Kri07] K. Krissian, C.-F. Westin, R. Kikinis, and K. G. Vosburgh. Oriented Speckle Reducing Anisotropic Diffusion. *IEEE Transactions on Image Processing*, 16(5):1412–1424, 2007.
- [Kri09a] K. Krissian and S. Aja-Fernandez. Noise-Driven Anisotropic Diffusion Filtering of MRI. *IEEE Transactions on Image Processing*, 18(10):2265–2274, 2009.
- [Kri09b] K. Krissian and S. Aja-Fernandez. Noise-Driven Anisotropic Diffusion Filtering of MRI. *IEEE Transactions on Image Processing*, 18(10):2265–2274, 2009.
- [Ksc01] F. R. Kschischang, B. J. Frey, and H. A. Loeliger. Factor Graphs and the Sum-Product Algorithm. *IEEE Transactions on Information Theory*, 47(2):498–519, 2001.
- [Kur04] E. E. Kuruoglu and J. Zerubia. Modeling SAR Images with a Generalization of the Rayleigh Distribution. *IEEE Transactions on Image Processing*, 13(4):527–533, 2004.
- [Law80] J. F. Lawless. Inference in the Generalized Gamma and Log Gamma Distributions. *Technometrics*, 22(3):409–419, 1980.
- [Mar95] T. L. Marzetta. EM Algorithm for Estimating the Parameters of Multivariate Complex Rician Density for Polarimetric SAR. In *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, pp. 3651–3654. Detroit, MI, USA, May 1995.

- [Mar00] G. Marsaglia and W. W. Tsang. A Simple Method for Generating Gamma Variables. *ACM Transactions on Mathematical Software*, 26(3):363–372, 2000.
- [Mat00] M. I. Matsagas, S. N. Vasdekis, A. G. Gugulakis, A. Lazaris, M. Foteinou, and M. N. Sechas. Computer-Assisted Ultrasonographic Analysis of Carotid Plaques in Relation to Cerebrovascular Symptoms, Cerebral Infarction, and Histology. *Annals of Vascular Surgery*, 14(2):130–137, 2000.
- [McG93] G. McGibney and M. R. Smith. An Unbiased Signal-to-Noise Ratio Measure for Magnetic Resonance Images. *Medical physics*, 20(4):1077–1078, 1993.
- [McK65a] J. K. S. McKay. Algorithm 262: Number of Restricted Partitions of N . *Communications of the ACM*, 8(8):493, 1965.
- [McK65b] J. K. S. McKay. Algorithm 264: Map of Partitions into Integers. *Communications of the ACM*, 8(8):493, 1965.
- [Mer96] M. Merkle. Logarithmic Convexity and Inequalities for the Gamma Function. *Journal of Mathematical Analysis and Applications*, 203(2):369–380, 1996.
- [Moo96] T. K. Moon. The Expectation-Maximization Algorithm. *IEEE Signal Processing Magazine*, 13(6):47–60, 1996.
- [Mur09] C. Murie, O. Woody, A. Lee, and R. Nadon. Comparison of Small N Statistical Tests of Differential Expression Applied to Microarrays. *BMC Bioinformatics*, 10(1):45, 2009.
- [Nai02] A. Nair, B. D. Kuban, E. M. Tuzcu, P. Schoenhagen, S. E. Nissen, and D. G. Vince. Coronary Plaque Classification with Intravascular Ultrasound Radiofrequency Data Analysis. *Circulation*, 106(17):2200–2206, 2002.
- [Nak40] M. Nakagami. Study of the Resultant Amplitude of Many Vibrations whose Phases and Amplitudes are at Random. *Journal of the Institute of Electronics and Communication Engineers of Japan*, 24(202):17–26, 1940.
- [Nak60] M. Nakagami. The m -Distribution, a General Formula of Intensity of Rapid Fading. In W. C. Hoffman, ed., *Symposium on Statistical Methods in Radio Wave Propagation*, pp. 3–36. Los Angeles, CA, USA, 1960.
- [Nel65] J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, 1965.
- [Ng06] J. Ng, R. Prager, N. Kingsbury, G. Treece, and A. Gee. Modeling Ultrasound Imaging as a Linear, Shift-Variant System. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 53(3):549–563, 2006.
- [Nil08] M. M. Nillesen, R. G. Lopata, I. H. Gerrits, L. Kapusta, J. M. Thijssen, and C. L. de Korte. Modeling Envelope Statistics of Blood and Myocardium for Segmentation of Echocardiographic Images. *Ultrasound in Medicine and Biology*, 34(4):674–680, 2008.
- [Nol97] J. P. Nolan. Parameter Estimation and Data Analysis for Stable Distributions. In *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 443–447. Pacific Grove, CA, USA, November 1997.

- [Nou12] A. Noufaily and M. Jones. On Maximization of the Likelihood for the Generalized Gamma Distribution. *Computational Statistics*, pp. 1–13, 2012.
- [Now99] R. D. Nowak. Wavelet-based Rician noise removal for Magnetic Resonance Imaging. *IEEE Transactions on Image Processing*, 8(10):1408–1419, 1999.
- [Opd10] J. Opdyke. A Unified Approach to Algorithms Generating Unrestricted and Restricted Integer Compositions and Integer Partitions. *Journal of Mathematical Modelling and Algorithms*, 9(1):53–97, 2010.
- [Pat49] P. B. Patnaik. The Non-Central χ^2 - and F -Distribution and their Applications. *Biometrika*, 36(1-2):202–232, 1949.
- [Pea82] J. Pearl. Reverend Bayes on inference engines: A distributed hierarchical approach. In *Second AAAI National Conference on Artificial Intelligence*, pp. 133–136. Pittsburgh, PA, USA, August 1982.
- [Per07] C. Perreault and M.-F. Auclair-Fortier. Speckle Simulation Based on B-Mode Echographic Image Acquisition Model. In *Fourth Canadian Conference on Computer and Robot Vision (CRV)*, pp. 379–386. Montreal, QC, Canada, May 2007.
- [Pop34] K. R. Popper. *The Logic of Scientific Discovery*. Hutchinson, London, UK, 1934.
- [Pou08] C. Poupon, A. Roche, J. Dubois, J. F. Mangin, and F. Poupon. Real-Time MR Diffusion Tensor and Q-ball Imaging using Kalman Filtering. *Medical Image Analysis*, 12(5):527–534, 2008.
- [Pra03] R. W. Prager, A. H. Gee, G. M. Treece, and L. H. Berman. Decompression and Speckle Detection for Ultrasound Images using the Homodyned K-Distribution. *Pattern Recognition Letters*, 24(4-5):705–713, 2003.
- [Ray80] L. Rayleigh. On the Resultant of a Large Number of Vibrations of the Same Pitch and of Arbitrary Phase. *Philosophical Magazine*, 10(60):73–78, 1880.
- [Ric45] S. O. Rice. Mathematical Analysis of Random Noise. *Bell System Technical Journal*, 24(1):46–156, 1945.
- [Rol88] T. J. Rolfe. Generations of Permutations with Non-Unique Elements. *SIGNAL Newsletter*, 23(2):24–28, 1988.
- [Roy62] J. Roy and M. L. Tiku. A Laguerre Series Approximation to the Sampling Distribution of the Variance. *Sankhyā, The Indian Journal of Statistics, Ser. A*, 24(2):181–184, 1962.
- [Sah11] R. K. Saha and M. C. Kolios. Effects of Cell Spatial Organization and Size Distribution on Ultrasound Backscattering. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 58(10):2118–2131, 2011.
- [Sal05] R. Salvador, A. Peña, D. K. Menon, T. A. Carpenter, J. D. Pickard, and E. T. Bullmore. Formal Characterization and Extension of the Linearized Diffusion Tensor Model. *Human Brain Mapping*, 24(2):144–155, 2005.
- [San03] J. M. Sanches and J. S. Marques. Compensation of Log-Compressed Images for 3-D Ultrasound. *Ultrasound in Medicine and Biology*, 29(2):239–253, 2003.
- [Sch78] G. Schwarz. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464, 1978.

- [Sch97] R. E. Schapire. Using Output Codes to Boost Multiclass Learning Problems. In *Fourteenth International Conference on Machine Learning (ICML)*, pp. 313–321. Nashville, TN, USA, July 1997.
- [Sch08] M. D. Schneider, L. Knox, S. Habib, K. Heitmann, D. Higdon, and C. Nakhleh. Simulations and Cosmological Inference: A Statistical Model for Power Spectra Means and Covariances. *Physical Review D*, 78(6):1–13, 2008.
- [Sea08] J. Seabra and J. Sanches. Modeling Log-Compressed Ultrasound Images for Radio Frequency Signal Recovery. In *30th IEEE Annual Conference of the EMB Society (EMBC)*, pp. 426–429. Vancouver, BC, Canada, August 2008.
- [Sea10a] J. Seabra and J. Sanches. On Estimating De-Speckled and Speckle Components from B-Mode Ultrasound Images. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 284–287. Rotterdam, The Netherlands, April 2010.
- [Sea10b] J. Seabra, J. Sanches, F. Ciompi, and P. Radeva. Ultrasonographic Plaque Characterization using a Rayleigh Mixture Model. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 1–4. Rotterdam, The Netherlands, April 2010.
- [Sea11] J. C. Seabra, F. Ciompi, O. Pujol, J. Mauri, P. Radeva, and J. Sanches. Rayleigh Mixture Model for Plaque Characterization in Intravascular Ultrasound. *IEEE Transactions on Biomedical Engineering*, 58(5):1314–1324, 2011.
- [Sed84] R. Sedgewick. *Algorithms*. Addison-Wesley Longman Publishing Co. Inc., Boston, MA, USA, 1984.
- [Sha93] P. M. Shankar, J. M. Reid, H. Ortega, C. W. Piccoli, and B. B. Goldberg. Use of Non-Rayleigh Statistics for the Identification of Tumors in Ultrasonic B-Scans of the Breast. *IEEE Transactions on Medical Imaging*, 12(4):687–692, 1993.
- [Sha00] P. M. Shankar. A General Statistical Model for Ultrasonic Backscattering from Tissues. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 47(3):727–736, 2000.
- [Sha01] P. M. Shankar. Ultrasonic Tissue Characterization using a Generalized Nakagami Model. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 48(6):1716–1720, 2001.
- [Sha03] P. M. Shankar, F. Forsberg, and L. Lown. Statistical Modeling of Atherosclerotic Plaque in Carotid B Mode Images – A Feasibility Study. *Ultrasound in Medicine and Biology*, 29(9):1305–1309, 2003.
- [Sij04] J. Sijbers and A. J. den Dekker. Maximum Likelihood Estimation of Signal Amplitude and Noise Variance Form MR Data. *Magnetic Resonance in Medicine*, 51(3):586–594, 2004.
- [Spa04] L. G. Spagnoli, A. Mauriello, G. Sangiorgi, S. Fratoni, E. Bonanno, R. S. Schwartz, D. G. Piegras, R. Pistolese, A. Ippoliti, and D. R. J. Holmes. Extracranial Thrombotically Active Carotid Plaque as a Risk Factor for Ischemic Stroke. *The Journal of the American Medical Association*, 292(15):1845–1852, 2004.
- [Spe02] T. P. Speed. John W. Tukey’s Contributions to Analysis of Variance. *Annals of Statistics*, 30(6):1649–1665, 2002.

- [Sta65] E. W. Stacy and G. A. Mihram. Parameter Estimation for a Generalized Gamma Distribution. *Technometrics*, 7(3):349–358, 1965.
- [Sto62] F. Stockmal. Algorithm 95: Generation of Partitions in Part-Count Form. *Communications of the ACM*, 5(6):344, 1962.
- [Sza04] T. Szabo. *Diagnostic Ultrasound Imaging: Inside Out*. Elsevier, Burlington, MA, USA, 2004.
- [Tak10] A. Taki, H. Hetterich, A. Roodaki, S. K. Setarehdan, G. Unal, J. Rieber, N. Navab, and A. Konig. A New Approach for Improving Coronary Plaque Component Analysis Based on Intravascular Ultrasound Images. *Ultrasound in Medicine and Biology*, 36(8):1245–1258, 2010.
- [Tan77] W. Y. Tan and S. P. Wong. On the Roy-Tiku Approximation to the Distribution of Sample Variances from Non-normal Universes. *Journal of the American Statistical Association*, 72(360):875–880, 1977.
- [Tao06] Z. Tao, H. D. Tagare, and J. D. Beaty. Evaluation of Four Probability Distribution Models for Speckle in Clinical Cardiac Ultrasound Images. *IEEE Transactions on Medical Imaging*, 25(11):1483–1491, 2006.
- [Thi03] J. M. Thijssen. Ultrasonic Speckle Formation, Analysis and Processing Applied to Tissue Characterization. *Pattern Recognition Letters*, 24(4-5):659–675, 2003.
- [Tik65] M. L. Tiku. Laguerre Series Forms of Non-Central χ^2 and F Distributions. *Biometrika*, 52(3-4):415–427, 1965.
- [Tit92] D. M. Titterington, A. F. Smith, and U. E. Makov. *Statistical Analysis of Finite Mixture Distributions*. Probability and Mathematical Statistics. Cambridge University Press, New York, NY, USA, 1992.
- [Tuk50] J. W. Tukey. Some Sampling Simplified. *Journal of the American Statistical Association*, 45(252):501–519, 1950.
- [Tuk56] J. W. Tukey. Keeping Moment-Like Sampling Computations Simple. *The Annals of Mathematical Statistics*, 27(1):37–54, 1956.
- [Tut88] T. A. Tuthill, R. H. Sperry, and K. J. Parker. Deviations from Rayleigh Statistics in Ultrasonic Speckle. *Ultrasonic Imaging*, 10(2):81–89, 1988.
- [TV09] A. Tristan-Vega and S. Aja-Fernandez. Design and Construction of a Realistic DWI Phantom for Filtering Performance Assessment. In G.-Z. Yang, D. Hawkes, D. Ruecker, A. Noble, and C. Taylor, eds., *12th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, vol. 5761 of *Lecture Notes in Computer Science*, pp. 951–958. London, UK, September 2009.
- [VSF10a] G. Vegas-Sanchez-Ferrero, S. Aja-Fernandez, M. Martin-Fernandez, A. F. Frangi, and C. Palencia. Probabilistic-Driven Oriented Speckle Reducing Anisotropic Diffusion with Application to Cardiac Ultrasonic Images. In T. Jiang, N. Navab, J. P. W. Pluim, and M. A. Viergever, eds., *13th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 6361 of *Lecture Notes in Computer Science*, pp. 518–525. Beijing, China, September 2010.
- [VSF10b] G. Vegas-Sanchez-Ferrero, D. Martin-Martinez, S. Aja-Fernandez, and C. Palencia. On the Influence of Interpolation on Probabilistic Models for Ultrasonic Images. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pp. 292–295. Rotterdam, The Netherlands, April 2010.

- [Wag83] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez. Statistics of Speckle in Ultrasound B-Scans. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):156–163, 1983.
- [Wag87] R. F. Wagner, M. F. Insana, and D. G. Brown. Statistical Properties of Radio-Frequency and Envelope-Detected Signals with Applications to Medical Ultrasound. *Journal of the Optical Society of America A*, 4(5):910–922, 1987.
- [Wah11] M. E.-S. Waheed, O. A. Mohamed, and M. E. A. El-aziz. Mixture of Generalized Gamma Density-Based Score Function for Fastica. *Mathematical Problems in Engineering*, pp. 1–14, 2011.
- [Wan04] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [WD07] N. Wiest-Daessle, S. Prima, P. Coupe, S. P. Morrissey, and C. Barillot. Non-Local Means Variants for Denoising of Diffusion-Weighted and Diffusion Tensor MRI. In N. Ayache, S. Ourselin, and A. Maeder, eds., *10th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 4792 of *Lecture Notes in Computer Science*, pp. 344–351. Brisbane, Australia, October 2007.
- [Web00] A. R. Webb. Gamma Mixture Models for Target Recognition. *Pattern Recognition*, 33(12):2045–2054, 2000.
- [Wil98] J. E. Wilhjelm, M.-L. Gronholdt, B. Wiebe, S. K. Jespersen, L. K. Hansen, and H. Sillesen. Quantitative Analysis of Ultrasound B-Mode Images of Carotid Atherosclerotic Plaque: Correlation with Visual Classification and Histological Examination. *IEEE Transactions on Medical Imaging*, 17(6):910–922, 1998.
- [Win87] D. R. Wingo. Computing Maximum-Likelihood Parameter Estimates of the Generalized Gamma Distribution by Numerical Root Isolation. *IEEE Transactions on Reliability*, 36(5):586–590, 1987.
- [Yu02] Y. Yu and S. T. Acton. Speckle Reducing Anisotropic Diffusion. *IEEE Transactions on Image Processing*, 11(11):1260–1270, 2002.
- [Zog98] A. Zoghbi and I. Stojmenovic. Fast Algorithms for Generating Integer Partitions. *International Journal of Computer Mathematics*, 70(2):319–332, 1998.