

DERIVATION OF A TEST STATISTIC FOR EMPHYSEMA QUANTIFICATION

Gonzalo Vegas-Sanchez-Ferrero, *George Washko*, *Farbod N. Rahaghi*
Maria J. Ledesma-Carbayo, *R. San José Estépar*

ABSTRACT

Density masking is the de-facto quantitative imaging phenotype for emphysema that is widely used by the clinical community. Density masking defines the burden of emphysema by a fixed threshold, usually between -910 HU and -950 HU, that has been experimentally validated with histology. In this work, we formalized emphysema quantification by means of statistical inference. We show that a non-central Gamma is a good approximation for the local distribution of image intensities for normal and emphysema tissue. We then propose a test statistic in terms of the sample mean of a truncated non-central Gamma random variable. Our results show that this approach is well-suited for the detection of emphysema and superior to standard density masking. The statistical method was tested in a dataset of 1337 samples obtained from 9 different scanner models in subjects with COPD. Results showed an increase of 17% when compared to the density masking approach, and an overall accuracy of 94.09%.

Index Terms— Emphysema quantification, statistical test, non-central Gamma, truncated random variable.

1. INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is projected to be the 3rd leading cause of death worldwide by 2020 [1, 2]. Thus, an accurate diagnosis, quantification, and assessment of the progression of COPD is of paramount importance. For this purpose, density masking is currently the method of choice for noninvasive assessment of the anatomical changes caused by emphysema [3], one of the pathological components of COPD.

Pulmonary emphysema is defined as an abnormal and permanent enlargement of the airspaces distal to the terminal bronchioles, accompanied by destruction of their walls, without obvious fibrosis [4]. Emphysema is anatomically

described at the level of the secondary pulmonary lobule, the smallest functional unit of the lung that has a nominal size of 25 mm³. The characterization of the different stages and types of emphysema follows different patterns according to the enlargement of airspaces, leading to lower intensity levels in the image within the secondary pulmonary lobule. Density masking provides a detection by a deterministic thresholding of the intensity levels in the CT image (in Hounsfield units) within the lung region.

Although several attempts have been made to include spatial texture, structure and location [5, 6, 7], density masking has become the de-facto phenotype for clinical studies [8]. Several thresholds between -910 HU and -950 HU have been proposed based on the correlation with pathological samples of diseases [9].

A fixed threshold has the caveat that is dependent on image formation parameters like reconstruction kernels, slice thickness and dose [10]. Different experimental methods have been proposed to compensate these effects [11, 12, 13]. More recently, a Hidden Markov models have been proposed to adapt to the local statistics of the signal while exploiting spatial coherence [14, 15]. In all these cases, weak assumptions are made about the underlying distribution of intensities without any formal formulation. Furthermore, density masking does not employ the notion that emphysema is defined within the secondary lobe unit.

In this work, we proposed a new formulation for emphysema quantification based on statistical inference within the secondary pulmonary lobule. Our approach is based on a non-central Gamma distribution assumption for local intensities that we empirically validate. Then, emphysema quantification is cast in terms of a hypothesis testing problem based on a novel test statistic. The proposed test statistic is inspired by the density masking philosophy, that has shown good histological correlations, and includes assumptions about the partial volume effect between airspaces within the acinus. The test statistic follows a truncated Gamma distribution whose PDF and CDF are analytically derived in here. These derivations allow us to propose a constructive method for the determination of the threshold to accept/reject the test at a given significance level.

2. STATISTICAL CHARACTERIZATION OF EMPHYSEMA IN SECONDARY PULMONARY LOBE

We propose to characterize statistically the behavior of normal and emphysematous tissue with a family of parametric distributions that include the qualitative behavior commonly shown in CT scanners with a clear skewness to the higher attenuation levels, its origin is located in -1000 HU and approximates 0 with a derivative of higher order than 1. In Fig. 1 the histogram of normal pulmonary tissue is depicted clearly showing these characteristics. This qualitative behavior can be described by means of the non-central Gamma (nc- Γ) distribution, whose PDF is defined as follows:

$$f_X(x; \alpha, \beta, \delta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} (x - \delta)^{\alpha-1} e^{-\frac{x-\delta}{\beta}}; \quad x \geq \delta \quad (1)$$

where δ is the location parameter, α accounts for the shape and β for the scale and $\Gamma(\cdot)$ is the Euler's Gamma function.

The statistical characterization of the secondary pulmonary lobe assumes that there are two different kinds of tissues: normal tissue (NT) accounting for the description of conventional parenchyma, and panlobular emphysema (PL), which implies the complete destruction of the entire acinus in the secondary pulmonary lobule causing an increase of airspaces and their volume. We will consider that any degree of detectable emphysema with CT scans is a combination of the statistical response of normal tissue and airspaces and, thus, it can be described by a combination of these two cases.

Empirical validation. In order to see the performance of the model for describing both cases, a set of 148 regions of Panlobular emphysematous tissue and 370 regions of Normal Tissue were fitted. The regions were manually selected by a professional pulmonologist in patches of size 24.18×24.18 mm². These ROIS were acquired from 16 sites hosting CT scanners by 3 vendors and a total of 9 specific models in subjects with COPD. This is a representative sample of a general population and a wide variate of scanners. In Fig. 1 we show the fitting of the proposed model for the whole set of regions. Note that the statistical behavior is described by the model for each of the two kinds of tissue considered. The parameters obtained for each of the tissues are: $\Theta_0 = \{\alpha_0, \beta_0, \delta_0\} = \{20.95, 2.18, -1050\}$ for NT tissue, and $\Theta_1 = \{\alpha_1, \beta_1, \delta_1\} = \{11.81, 2.58, -1050\}$ ¹.

Emphysema detection. The characterization of emphysema by a statistical behavior resulting from a combination of NT and PS tissues lead to the definition of the most likely threshold to distinguish between both kinds of tissues. This can be obtained by simply equaling the classification error for NT type to the classification error for PL type as $F_X(t|\Theta_1) = 1 - F_X(t|\Theta_0)$.

¹Note that the location parameter δ is set to -1050 HU because noisy response of the attenuation may cause more negative attenuation values than -1000 HU.

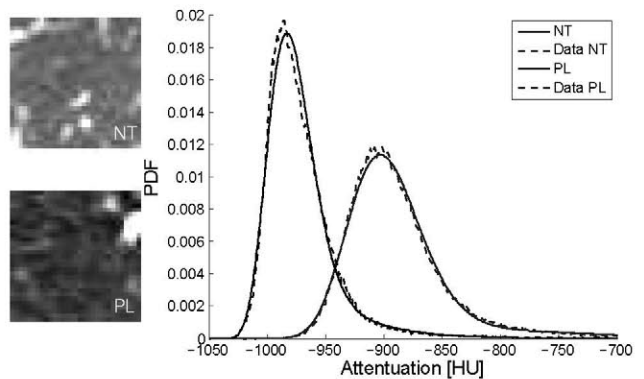


Fig. 1. Fitting of the proposed model to the two normal pulmonary tissue (NT) and panlobular tissue (PL).

So, for the analyzed dataset, this threshold is located at $t = -940.74$ HU. Note that this value is in the range normally used in density masking for $t \in [-950, -910]$ [3].

So far, the definition of this threshold by considering the attenuation response of airspaces and normal tissue differs little from the perspective of the density masking. However, the combination of the probabilistic model proposed for normal and emphysematous tissue with this optimal threshold will allow us to propose a statistical test that can evaluate those pixels below this threshold to increase the classification performance within an inferential and hypothesis testing framework. The next section is devoted to the derivation of the statistical test.

3. STATISTICAL TEST FOR EMPHYSEMA

3.1. Definition of the statistic

For the development of the statistical test we consider the same hypothesis assumed in the mask density method: *emphysema implies a decrease of the attenuation levels due to the destruction of lung tissue*. With this hypothesis, we want to define a statistical test that accounts for samples below the threshold that differentiates normal tissue and destruction of tissue, so the statistical test will consider samples below the threshold between normal and PL tissue calculated in the previous section. Additionally, the partial volume effect of combinations of airspaces and normal parenchyma will affect the attenuation levels of samples below this threshold. This fact allows us to define the mean of truncated samples as the ideal statistic for testing emphysema, where the truncation is set to the threshold t , i.e. from a set of samples $\mathcal{X} = \{x_i\}_{i=1}^N$ we consider $\mathcal{Y} = \{y \in \mathcal{X} : y < t\}$ and the statistic:

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n y_i, \quad \text{with } y_i \in \mathcal{Y} \quad (2)$$

where n is the number of samples in \mathcal{Y} . So, if we assume that samples \mathcal{X} come from normal parenchyma, the resulting samples would follow a *truncated nc- Γ distribution* with

PDF:

$$f_Y(y; \Theta_0) = \frac{(y - \delta_0)^{\alpha_0 - 1} e^{-\frac{y - \delta_0}{\beta_0}}}{\Gamma(\alpha_0) \beta_0^{\alpha_0} F_X(t, \Theta_0)}, \quad (3)$$

where $y \in [0, t]$, $\Theta_0 = \{\alpha_0, \beta_0, \delta_0\}$ and $F_X(\cdot, \Theta_0)$ is the cumulative distribution function (CDF) of the nc- Γ distribution describing the original random variable X already introduced.

3.2. Derivation of probabilistic distribution of the statistic

In order to define a statistical test for hypotheses involving the sample mean of truncated nc- Γ variates, the PDF of the sample mean should be estimated. Many efficient methods can be applied for the estimation of the sample mean distribution of nc- Γ variates. In this work we choose the Edgeworth expansion which approximates the distribution around the standard Gaussian distribution by means of combinations of Hermite polynomials whose coefficients depend on the moments of the random variable. This approach has been demonstrated to be an asymptotic expansion of the PDF and has good convergence properties [16].

In our case, let $Z_n = \frac{\bar{Y}_n - E\{\bar{Y}_n\}}{\sqrt{\text{Var}\{\bar{Y}_n\}}}$ be a variate with CDF $F_n(\cdot)$ where Y_1, \dots, Y_n are independent and identically distributed (IID) random variables distributed as in eq. (3). The Central Limit Theorem states that $F_n(z) \rightarrow \Phi(z)$ for every z where $\Phi(z)$ is the CDF of a standard Normal $\mathcal{N}(0, 1)$. The bound of this approximation is provided by the Berry-Essen theorem as $|F_n(z) - \Phi(z)| = O\left(\frac{1}{\sqrt{n}}\right)$ uniformly in z when \bar{Y}_n has three finite moments [17]. When this approximation is modified by suitable constants and polynomials, $q_r(\cdot)$, a better bound of the approximation can be derived:

$$F_n(z) = \Phi(z) + \sum_{r=1}^k \frac{q_r(z)}{n^{r/2}} + o(n^{-r/2}) \quad (4)$$

uniformly in z , which is the so called Edgeworth expansion for \bar{Y}_n [17] and order k . The coefficients in eq. (4) depend on the moments of \bar{Y}_n and, thus, the Edgeworth expansion is especially convenient for our formulation since the central moments with respect to δ_0 can be calculated as follows:

First, the r -th central moment, $\mu_r = E\{(Y - \delta_0)^r\}$, of a truncated nc- Γ variate is:

$$\begin{aligned} \mu_r &= \int_0^t \frac{(y - \delta_0)^{r + \alpha_0 - 1} e^{-\frac{y - \delta_0}{\beta_0}}}{\Gamma(\alpha_0) \beta_0^{\alpha_0} F_X(t, \Theta_0)} dy \\ &= \frac{\Gamma(\alpha + r) \beta^\alpha F_X(t, \alpha_0 + r, \beta_0, \alpha_0)}{\Gamma(\alpha) F_X(t, \Theta_0)}. \end{aligned} \quad (5)$$

Note that these moments converge to the ones of the conventional nc- Γ as $t \rightarrow \infty$.

Now the moments of the sample mean with respect δ_0 , $m_r = E\left\{\left(\frac{1}{n} \sum_{i=1}^n (y_i - \delta_0)\right)^r\right\}$, can be exactly calculated by

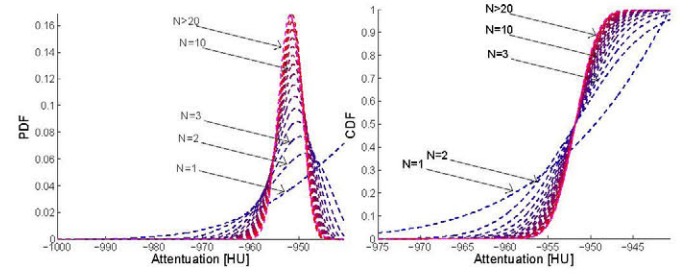


Fig. 2. Probability distributions of sample mean of truncated nc- Γ for an increasing number of samples $N = n$.

means of the *theorem of moments of sample variates* [18] as:

$$m_r = \frac{1}{n^r} \sum_{s=1}^r \binom{r}{s} \sum_{\mathbf{p} \in \mathcal{P}_s} \sigma(\mathbf{p}) \sum_{j=1}^s \prod_{l=1}^s \mu E_{j,l} \quad (6)$$

where \mathcal{P}_s is the set of solutions of the integer partition problem for n into s terms, $\sigma(\mathbf{p})$ is a multinomial coefficient, and $E_{j,l}$ accounts for the permutations without repetition of s elements with multiplicities the elements of \mathbf{p} (see [18] for more details).

Finally, both the PDF and CDF of the typified sample mean Z_n of truncated nc- Γ variates is derived by the Edgeworth expansion as:

$$F_n(z) = \Phi(z) - \Phi(z) \sum_{r=1}^k \frac{h_r(z)}{n^{r/2}}; f_n(z) = \phi(z) \left(1 + \sum_{r=1}^k \frac{\bar{h}_r(z)}{n^{r/2}} \right) \quad (7)$$

where Φ and ϕ are the CDF and PDF of a standard Normal $\mathcal{N}(0, 1)$ respectively. The polynomials h_r, \bar{h}_r can be easily obtained as linear combinations of Hermite Polynomials [17].

From eqs. (7) one can easily obtain both the CDF and PDF of \bar{Y}_n by simply applying the transformation $\bar{Y}_n = Z_n \sqrt{m_2 - m_1^2} + m_1$.

In Fig. 2 we show the PDFs and CDFs of the proposed statistic for an increasing number of samples calculated with an Edgeworth of order 5. Note that the distributions converge to a Gaussian distribution with mean $\mu_1 = \frac{\Gamma(\alpha+r)\beta^\alpha F_X(t, \alpha_0+1, \beta_0)}{\Gamma(\alpha) F_X(t, \Theta_0)} = -951.82$ HUs, which is interestingly related to the common measure used in the standard density mask usually set to -950 HU.

Note also that for a number of samples $N > 20$, due to the Central Limit Theorem, both the PDF and CDFs can be effectively approximated by a Normal distribution with mean m_1 and variance $\frac{1}{n}(m_2 - m_1^2)$.

3.3. Significance levels for one-tailed test

With the calculation of the PDF and CDF of the sample mean, we can develop the statistical test for an arbitrary statistical significance. We state the null hypothesis H_0 as “*The truncated mean of samples below the threshold t come from normal parenchyma*”, whereas the alternative hypothesis, H_1 is “*The truncated mean of samples below the threshold come*

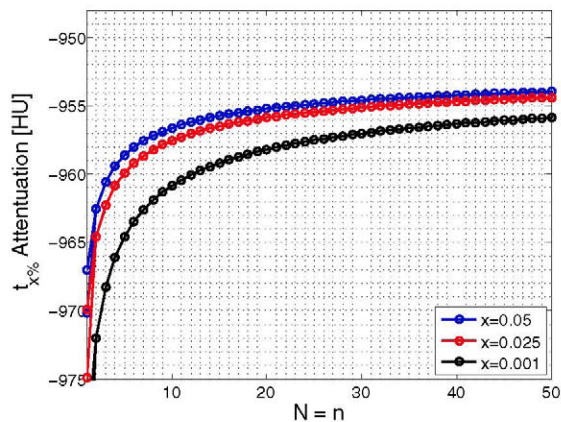


Fig. 3. Acceptance thresholds for an increasing number of samples and significance levels 0.05, 0.025, 0.001.

from emphysema”. In this context, the threshold accounting for accepting/rejecting the test can be directly calculated from the CDF obtained by the Edgeworth expansion for the desired significance. In Fig. 3 the acceptance threshold for different significance values ($t_{0.05,n}$, $t_{0.025,n}$ and $t_{0.001,n}$) is shown for an increasing number of samples.

The threshold was exactly calculated for $N \leq 20$, while the Gaussian approximation is used for $N > 20$ since the sample distribution quickly converges to the Gaussian distribution for higher number of samples.

4. PERFORMANCE TEST

In this section we test the performance of the hypothesis test by comparing the conventional threshold of -950 HU to the proposed test. For this purpose, we consider a dataset of 1337 lung regions of interest (ROIs) of size 24.18×24.18 mm² segmented and classified by a pulmonologist as normal parenchyma (NT) and emphysematous tissue including paraseptal (PS), panlobular (PL), emphysema and centrilobular for three different severity levels (CL1-3) from mild to severe distributed as: 370 NT, 184 PS, 148 PL, 170 CL1, 287 CL2, 178 CL3. These ROIs were also acquired from 16 sites hosting CT scanners by 3 vendors and a total of 9 specific models in subjects with COPD. The classification follows a leave-one-out methodology as explained in sections 2 and 3.

The test was performed to all the ROIs with a statistical significance 0.05, while the conventional approach assumed a threshold for -950 HU considering that the sample has emphysema 5% of the samples in the ROI are below this threshold. The classification performance is shown in table 1 where the superiority of the proposed approach is clear, showing a good balance between sensitivity (97.64%) and specificity (85.93%) compared to the conventional approach. This supposes an increase of 7% of the accuracy and an increase of 16% in the specificity, confirming the suitability of the probabilistic model of emphysema/parenchyma and the statistic proposed.

Actual	Predicted		Actual	Predicted	
	Normal	Emph.		Normal	Emph.
Normal	94.05%	5.89%	Normal	89.19%	14.48%
Emph.	5.95%	94.11%	Emph.	10.81%	85.52%
Acc.: 94.09%			Acc.: 86.54%		
Sen.: 97.64%			Sen.: 95.39%		
Spe.: 85.93%			Spe.: 70.21%		
(a) Statistical Test.			(b) Threshold at -950 HU.		

Table 1. Confusion Matrices of the proposed statistical test at significance 0.05 (a), the density masking (b) and global accuracy (Acc.), Sensitivity (Sen.) and Specificity (Spe.).

5. CONCLUSION

In this paper we propose a statistical test to evaluate the existence of emphysematous tissue in the secondary pulmonary lobe. This is achieved by proposing a family of statistical distributions (the nc- Γ distribution) to describe the attenuation of conventional tissue and airspaces caused by panlobular emphysema. This family of distributions accurately describes the statistical behavior of both kinds of tissues and leads to the definition of an optimal threshold for considering samples in the secondary pulmonary lobe as potential emphysema.

A statistical test is then proposed by defining a statistic that accounts for the average attenuation of samples under the optimal threshold that differentiates between normal tissue and panlobular tissue. This statistic is modeled as the sample mean of truncated nc- Γ distributions. The PDF and CDF of the statistic are derived and the statistic threshold is obtained for different confidence intervals.

Interestingly, the proposed statistic shows an asymptotic convergence to -951.8 HU, which is very close to the one proposed for density masking based on histological analysis of lung lobes[9]. This threshold is optimal in terms of the statistic for large regions ($n \rightarrow \infty$). However, when a smaller region of the lung is considered, our model suggests that lower thresholds (≤ -950) should be used to maintain the same significance level.

The classification was evaluated in a dataset from a multicenter study with 1,337 samples from different CT scanner models and vendors. We chose this dataset to provide a general evaluation of the statistical test. The accuracy reached was 94.09% with a sensitivity 97.64% and Specificity 85.93%. The statistical test supposes an increment of 7.5% in total accuracy and 16% in specificity when compared to the density masking approach.

The model proposed in this study can be combined to other approaches such as Hidden Markov models to take advantage of the local statistics with spatial coherence. For this purpose, the family of distributions purposes may serve to pave the way to well grounded potentials of the Random Field definition. Our future work will focus in this direction. Alternatively, this work opens the door to propose additional test statistics that consider the severity of disease in terms of the percentage of secondary lobule that is affected.

6. REFERENCES

- [1] M.M. Ward, H.S. Javitz, W.M. Smith, and A. Bakst, "Direct medical cost of chronic obstructive pulmonary disease in the U.S.A.," *Respiratory Medicine*, vol. 94, no. 11, pp. 1123 – 1129, 2000.
- [2] Amund Gulsvik, Gunnar Boman, Ronald Dahl, Thorarinn Gislason, and Markku Nieminen, "The burden of obstructive lung disease in the nordic countries," *Respiratory Medicine*, vol. 100, Supplement 1, pp. S2 – S9, 2006.
- [3] Nestor L Müller, Catherine A Staples, Roberta R Miller, and Raja T Abboud, "Density mask. An objective method to quantitate emphysema using computed tomography.," *Chest*, vol. 94, no. 4, pp. 782–787, Oct. 1988.
- [4] Souza Junior AS Rubin AS Zanetti G. Hochegger B, Irion KL, "Normal variance in emphysema index measurements in 64 multidetector-row computed tomography," *J Appl Clin Med Phys.*, vol. 4, no. 14, pp. 4215, 2013.
- [5] Renuka Uppaluri, Theophano Mitsa, Milan Sonka, Eric A. Hoffman, and Geoffrey McLennan, "Quantification of pulmonary emphysema from lung computed tomography images. am j respir crit care med 156(1,," 1997.
- [6] M Mishima, T Hirai, H Itoh, Y Nakano, H Sakai, S Muro, K Nishimura, Y Oku, K Chin, M Ohi, T Nakamura, J H Bates, A M Alencar, and B Suki, "Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease.," *PNAS*, vol. 96, no. 16, pp. 8829–8834, Aug. 1999.
- [7] N. Kim E. J. Chae Y. M. Oh S. D. Lee Y. Lee Y. S. Park, J. B. Seo and S. H Kang, "Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: Comparison with density-based quantification and correlation with pulmonary function test," *Investigative radiology*, vol. 43, no. 6, pp. 395402, 2008.
- [8] Alvar Agusti, Peter MA Calverley, Bartolome Celli, Harvey O Coxson, Lisa D Edwards, David A Lomas, William MacNee, Bruce E Miller, Steve Rennard, Edwin K Silverman, Ruth Tal-Singer, Emiel Wouters, Julie C Yates, Jørgen Vestbo, and author firstName author lastName, "Characterisation of COPD heterogeneity in the ECLIPSE cohort," *Respiratory research*, vol. 11, no. 1, pp. 122, Sept. 2010.
- [9] P A Gevenois, P De Vuyst, V de Maertelaer, J Zanen, D Jacobovitz, M G Cosio, and J C Yernault, "Comparison of computed density and microscopic morphometry in pulmonary emphysema.," *American journal of respiratory and critical care medicine*, vol. 154, no. 1, pp. 187–192, July 1996.
- [10] David S Gierada, Andrew J Bierhals, Cliff K Choong, Seth T Bartel, Jon H Ritter, Nitin A Das, Cheng Hong, Thomas K Pilgram, Kyongtae T Bae, Bruce R Whiting, Jason C Woods, James C Hogg, Barbara A Lutey, Richard J Battafarano, Joel D Cooper, Bryan F Meyers, and G Alexander Patterson, "Effects of CT Section Thickness and Reconstruction Kernel on Emphysema Quantification: Relationship to the Magnitude of the CT Emphysema Index.," *Academic Radiology*, pp. –, Nov. 2009.
- [11] Seth T Bartel, Andrew J Bierhals, Thomas K Pilgram, Cheng Hong, Kenneth B Schechtman, Susan H Conradi, and David S Gierada, "Equating quantitative emphysema measurements on different CT image reconstructions.," *Medical physics*, vol. 38, no. 8, pp. 4894–4902, Aug. 2011.
- [12] Mario Ceresa, Gorka Bastarrika, Juan P de Torres, Luis M Montuenga, Javier J Zulueta, Carlos Ortiz-de Solorzano, and Arrate Muñoz-Barrutia, "Robust, standardized quantification of pulmonary emphysema in low dose CT exams.," *Academic Radiology*, vol. 18, no. 11, pp. 1382–1390, Nov. 2011.
- [13] A.M.R Schilham, B van Ginneken, H Gietema, and M Prokop, "Local noise weighted filtering for emphysema scoring of low-dose CT images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 4, pp. 451–463, 2006.
- [14] Y Hame, E D Angelini, E.A. Hoffman, R G Barr, and A F Laine, "Robust quantification of pulmonary emphysema with a Hidden Markov Measure Field model," in *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*. 2013, pp. 382–385, IEEE.
- [15] Y Hame, E D Angelini, E.A. Hoffman, R G Barr, and A F Laine, "Adaptive Quantification and Longitudinal Analysis of Pulmonary Emphysema With a Hidden Markov Measure Field Model," *IEEE transactions on medical imaging*, vol. 33, no. 7, pp. 1527–1540, July 2014.
- [16] S. Blinnikov and R. Moessner, "Expansions for nearly gaussian distributions," *Astron. Astrophys. Suppl. Ser.*, , no. 130, pp. 193–205, 1998.
- [17] A. DasGupta, *Asymptotic Theory of Statistics and Probability.*, Springer Texts in Statistics, Springer Science+Business Media, New York, 2008.
- [18] Gonzalo Vegas-Sánchez-Ferrero, Santiago Aja-Fernández, Marcos Martín-Fernández, and Cesar Palencia, "A direct calculation of moments of the sample variance," *Mathematics and Computers in Simulation*, vol. 82, no. 5, pp. 790–804, 2012.