Appendix B

Probability Density Functions After Averaging Under Null Conditions

When combining K distributions in an averaging process, K convolutions are required. By using the characteristic function of a distribution, a convolutional process can be simplified to that of multiplication [103 ch5.5].

The Fourier transform of a random variable x with a pdf $f(x)$, is known as the characteristic function, $\phi_x(\omega')$, where ω' is used to distinguish it from ω associated with the original Fourier transform of the time domain signal. $\phi_x(\omega')$ is defined by:

$$\phi_x(\omega') = E[e^{j\omega' x}] = \int_{-\infty}^{\infty} e^{j\omega' x} f(x)dx$$

There is a one to one correspondence between a pdf and its characteristic function. Information contained in one function is equivalent to the information contained in the other function. Taking K independent random variables $x_1, x_{2,} \ldots\ldots, x_K$ having probability densities $f_1(x_1), f_2(x_2), \ldots\ldots, f_K(x_K)$ and associated characteristic functions $\phi_1(\omega'), \phi_2(\omega'), \ldots\ldots, \phi_K(\omega')$ and considering the sum of a number of random variables, the overall characteristic function will be given by:

$$\phi_z(\omega') = \int e^{j\omega' x_1} f_1(x_1)dx_1 \cdot \int e^{j\omega' x_2} f_2(x_2)dx_2 \cdot \ldots\ldots \int e^{j\omega' x_K} f_K(x_K)dx_K$$

$$= \phi_1(\omega') \cdot \phi_2(\omega') \cdot \ldots\ldots \phi_K(\omega')$$

Therefore, the characteristic function of the sum of a number of independent random variables is given by the product of all the individual characteristic functions. To obtain the pdf from $\phi_z(\omega')$ the inverse Fourier transform is applied:

$$f(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_z(\omega') e^{-j\omega' z} d\omega$$

B.1 'Log of Mean' pdf

The distribution of the power spectrum of a Fourier coefficient for a random variable q and an original noise power of σ_{ft}^2 was derived in chapter 6.5 and given by:

$$f_Q(q) = \frac{1}{2\sigma_{ft}^2} \mathrm{e}^{\left(\frac{-q}{2\sigma_{ft}^2}\right)} \qquad q \geq 0 \tag{B.1}$$

The characteristic function of the pdf described by (B.1) is found to be:

$$\phi_q(\omega') = \int_{-\infty}^{\infty} e^{j\omega' q_1} f_Q(q) dq$$

$$\phi_q(\omega') = \left(\frac{\sqrt{1 + j\sigma_{ft}^2 \omega'}}{\sqrt{1 - j\sigma_{ft}^2 \omega'} \cdot \sqrt{1 + j\sigma_{ft}^4 \omega'^2}} \right)$$

The averaging process requires the sum of a number of samples *K* produced from identical distributions and the result divided by *K*. The characteristic function $\phi_q(\omega')$ has to be multiplied by itself K times to produce a new characteristic function for the sum of random variables:

$$\phi_{q_K}(\omega') = \left(\frac{\sqrt{1 + j\sigma_{ft}^2 \omega'}}{\sqrt{1 - j\sigma_{ft}^2 \omega'} \cdot \sqrt{1 + j\sigma_{ft}^4 \omega'^2}} \right)^K \tag{B.2}$$

To obtain the final pdf of the average, (B.2) is first inverse Fourier transformed and a further transformation of the pdf is then required since to find the average the sum is divided by the number of samples *K*. The generalised result is given by:

$$f_{A_K}(q) = \frac{e^{-\frac{Kq}{\sigma_{ft}^2}} \cdot \sigma_{ft}^{-2K} \cdot q^{-1+K} \cdot K^K}{2 \cdot \Gamma(K)} \tag{B.3}$$

where $\Gamma(K) = (K-1)!$ for integer K

Equation (B.3) is plotted for a range of K values, representing a source of random noise having a fixed original variance of 1 and is shown in fig B.1. The narrower distributions correspond to the larger number of averaged samples. The non-averaged exponential distribution (B.1) is also shown by the green trace.

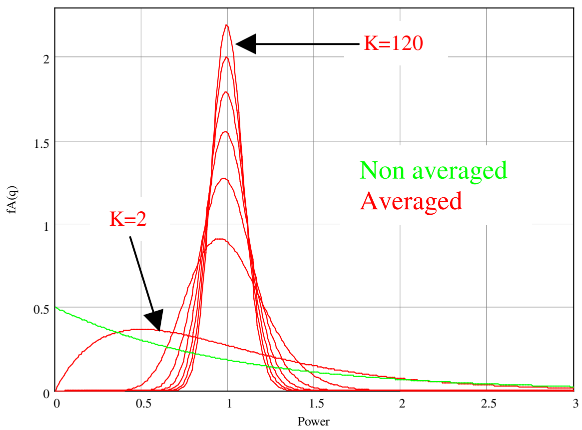

Fig B.1: consecutive arithmetic averages of the distribution of a power spectral coefficient. Green trace shows non-averaged distribution described by (B.1).

The averaged result is then converted logarithmically and this requires a further transformation of the pdf given by (B.4):

$$f_{LM_K}(q) = \frac{e^{K\left(\frac{e^{-q}}{\sigma_{ft}^2}+q\right)} \cdot \sigma_{ft}^{-2K} \cdot K^K}{\Gamma(K)} \qquad \text{(B.4)}$$

This is plotted for a range of K values for a source of random noise, having a fixed original variance of 1, and is shown in fig B.2. The narrower distributions correspond to the larger number of averaged samples. The non-averaged distribution where K equals 1 is shown by the green trace.

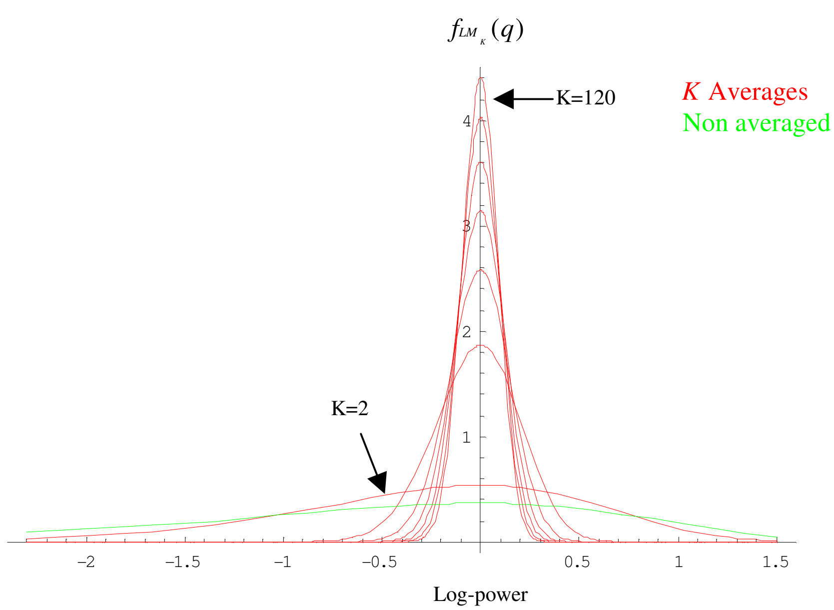

Fig B.2: Logarithmically transformed distributions described by (B.4).

The distributions are asymmetrical and remain so even for high values of K having a skewness value of -1. The mean value of $f_{LM_K}(q)$ is given by:

$$\overline{LM}_K = \int_{-\infty}^{\infty} q \cdot \frac{e^{K\left(\frac{e^{-q}}{\sigma_{ft}^2}+q\right)} \cdot \sigma_{ft}^{-2K} \cdot K^K}{\Gamma(K)} dq$$

The results are plotted over a range of K values and are shown in fig B.3. For a large K the mean converges to a value of $\ln(\sigma^2)$ and is the value of the mode, which remains constant for changing K values. Finding the mean square value and subtracting the square of the mean, (B.5), gives the variance of a distribution and this has been plotted for a range of K values shown in fig B.4.

$$V_{LM_K} = \overline{LM_K^{\ 2}} - \left(\overline{LM}_K\right)^2$$

$$= \int_{-\infty}^{\infty} q^2 \cdot \frac{e^{K\left(\frac{e^{-q}}{\sigma_{ft}^2}+q\right)} \cdot \sigma_{ft}^{-2K} \cdot K^K}{\Gamma(K)} dq - \left(\int_{-\infty}^{\infty} q \cdot \frac{e^{K\left(\frac{e^{-q}}{\sigma_{ft}^2}+q\right)} \cdot \sigma_{ft}^{-2K} \cdot K^K}{\Gamma(K)} dq\right)^2 \qquad \text{(B.5)}$$

The variance is independent of the original variance but dependent on the number of averages and is seen to be approximately inversely proportional to K, converging to zero as the number of averages increases to infinity.

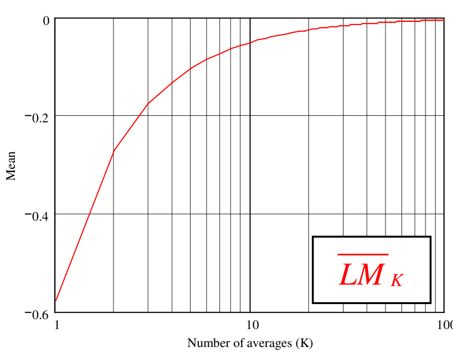

Fig B.3: Averaged log of mean value converges to the mode of its non-averaged distribution as K increases.

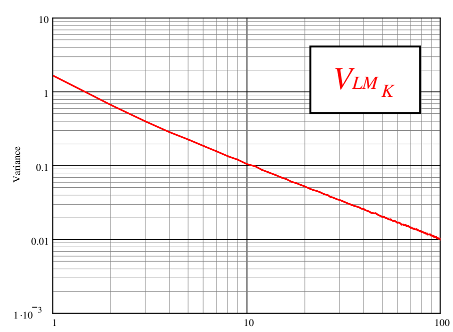

Fig B.4: Variance of the averaged log of mean value is approximately inversely proportional to K.

B.2 'Mean of Log' pdf

The distribution of the logarithm for a Fourier coefficient of the power spectrum for a random variable r and original spectral noise power σ^2 was derived in chapter 5.5 and given by:

$$f_R(r) = \frac{e^r}{2\sigma_{ft}^{2\,2}} e^{\left(\frac{-e^r}{2\sigma_{ft}^2}\right)} \qquad -\infty \geq r \leq \infty \tag{B.6}$$

The characteristic function of the pdf described by (B.6) is found to be:

$$\phi_r(\omega') = \int e^{j\omega r_1} f_R(r)dr = \left(\frac{1}{\sigma_{ft}^2}\right)^{j\omega'} \cdot \Gamma(1 - j\omega')$$

where Γ is the Gamma function [179].

The averaging process requires the sum of a number of samples K, produced from identical distributions. The characteristic function $\phi_r(\omega')$ has to be multiplied by itself K times to produce a new characteristic function:

$$\phi_{r_K}(\omega') = \sigma_{ft}^{-2j\omega' K} \cdot \Gamma^K(1 - j\omega') \tag{B.7}$$

To obtain the final pdf of the average, (B.7) is first inverse Fourier transformed and a further transformation of the pdf is then required due to dividing the sum by the number of samples K. This integral proves difficult and a generalised symbolic result is not readily obtained. However, the distribution $f_{ML_K}(r)$ can be calculated for given values of parameters and this has been plotted for a range of K values having a fixed original variance of 1, and is shown in fig B.5. The narrower distributions correspond to the larger number of averaged samples. The non averaged distribution where K is 1 is shown by the green trace.

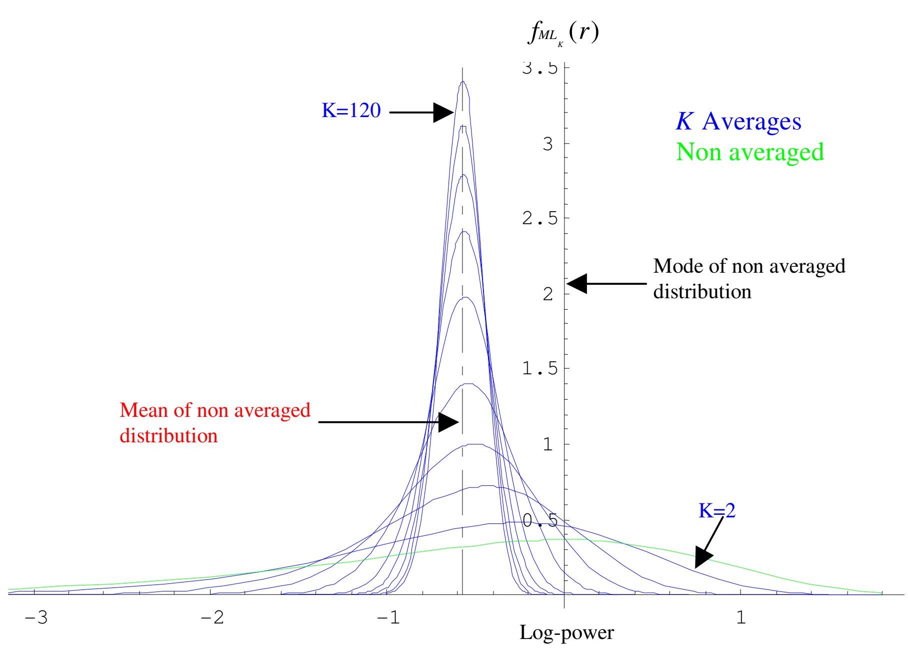

Fig B.5: pdf results from averaging a number of distributions.

As predicted by the central limit theorem as K increases the distribution converges to that of Gaussian and the mode of the averaged distributions can be seen to converge to the mean value of the non averaged distribution. It is shown in appendix A.3 that the variance of the non-averaged distribution is $\pi^2/6$, the variance of the averaged distribution will now be calculated. Even though it was not possible to invert the characteristic function, the higher order moments may be found by differentiating the characteristic function n times to obtain the n^{th} algebraic moment [103 ch5.5]. Taking into account the division by the number of samples, the n^{th} moment of K averages is given by:

$$\frac{d^n}{d\omega'^n}\left[\frac{\phi_{r_K}(\omega')}{K^n}\right]\Bigg|_{\omega'=0} = \frac{d^n}{d\omega'^n}\left[\frac{\sigma_{ft}^{-2j\omega' K}\cdot\Gamma^K(1-j\omega')}{K^n}\right]\Bigg|_{\omega'=0}$$

Subtraction of the square of the mean from the mean square values results in the variance:

$$V_{ML_K} = \frac{d^2}{d\omega'^2}\left[\frac{\phi_{r_K}(\omega')}{K^2}\right]\Bigg|_{\omega'=0} - \left(\frac{d}{d\omega'}\left[\frac{\phi_{r_K}(\omega')}{K}\right]\Bigg|_{\omega'=0}\right)^2 = \frac{\pi^2}{6K}$$

The variance is independent of the original variance but dependent on the number of averages and is seen to be inversely proportional to K, converging to zero as the number of averages increases to infinity. Figure B.6 shows this change in variance for changing average numbers compared with the results of the variance for the 'log of mean'.

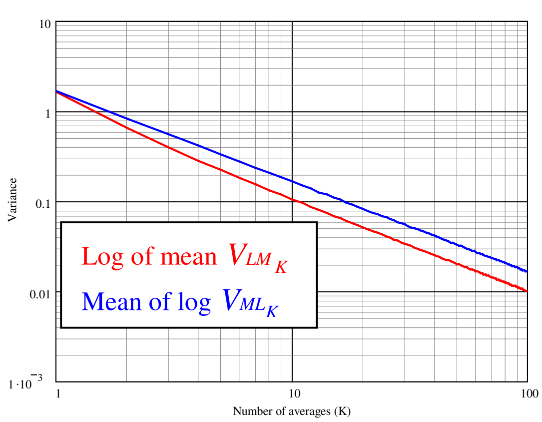

Fig B.6: Change in variance for a change in the number of averages, shown for both log of mean and mean of log methods.

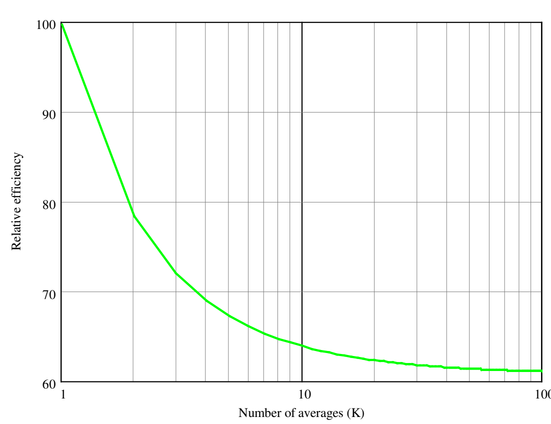

Fig B.7: Relative efficiency of the 'mean of log' compared to the 'log of mean'.

The relative efficiency of the 'mean of log' compared to the 'log of mean', given by the ratio V_{LM_K}/V_{ML_K} is lower than the optimum 'log of mean' as shown as a percentage by fig B.7.

B.3 pdf for the RDT of the 'Log of Mean'

In appendix B.1, a generalised result for the averaged distribution of the log of the mean method $f_{LM_K}(q)$ was derived (B.4). The requirement is to establish the probability density function of the difference between two random variables. The same method as that used for the calculation of the non-averaged RDT distribution, described in chapter 5.5 will be used to derive the averaged RDT distribution. The method involves placing one random variable in terms of the other, allowing the integration of the product of the two separate distributions with respect to one random variable only:

$$f_{LMR_K}(q) = \int_{-\infty}^{\infty} \left[\frac{e^{K\left(\frac{e^{-q}}{\sigma_{\omega 1}^2}+q\right)} \cdot \sigma_{\omega 1}^{-2K} \cdot K^K}{\Gamma(K)} \cdot \frac{e^{K\left(\frac{e^{a+q}}{\sigma_{\omega 2}^2}+a+q\right)} \cdot \sigma_{\omega 2}^{-2K} \cdot K^K}{\Gamma(K)} \right] da$$

when $\sigma_{\omega 1}^2 = \sigma_{\omega 2}^2 = \sigma^2$ this simplifies to:

$$f_{LMR_K}(q) = \int_{-\infty}^{\infty} \left[\frac{e^{-\frac{K(e^{-q}+e^{a+q}-\sigma^2(a+2q))}{\sigma^2}} \cdot \sigma^{-2K} \cdot K^{2K}}{\Gamma^2(K)} \right] da \qquad \text{(B.8)}$$

It is found that (B.8) has no closed form solution, but the distribution for individual K values can be found by numerical integration. Fig B.8 shows the distributions for K values of 1 to 10 produced from (B.8). Making $\sigma_{\omega 1}^2$ equal to $\sigma_{\omega 2}^2$, the mean value of all the distributions will be zero. As the number of averages increase, the distributions become narrower.

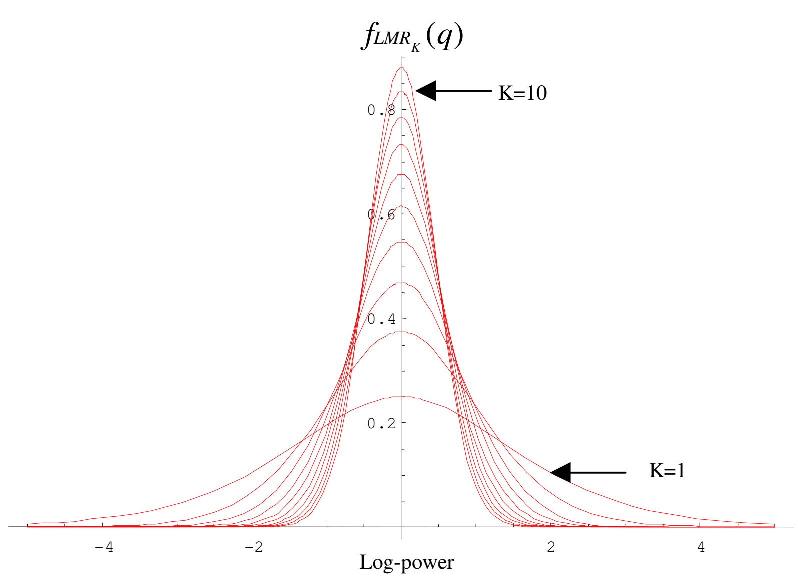

Fig B.8: Distribution of the RDT under null conditions for the log of the mean method, shown for K values of 1 to 10.

It has been found by calculation for the individual distributions shown, that the variance is approximately inversely proportional to K and that the variance of the averaged distribution for any given number of averages is twice that of the distribution before the RDT has been applied. Due to the symmetry of the distribution the skew is zero.

B.4 pdf for the RDT of the 'Mean of Log'

In appendix B.2 a closed form solution of a generalised result for the averaged distribution of the mean of the log method was not available, but a generalised result for the characteristic function $\phi_{r_K}(\omega')$ was derived (B.7). The RDT subtracts an estimate of the log of the power spectral density at one frequency from the log of the power spectral density at another frequency. Assuming the terms are independent, then the characteristic function of the RDT will be the characteristic function of one estimate times the complex conjugate of the other characteristic function:

$$\phi_{rRDT_K^{\#}}(\omega') = \sigma_1^{-2j\omega' K} \cdot \Gamma^K\left(1 - j\omega'\right) \cdot \sigma_2^{-2j\omega' K} \cdot \Gamma^K\left(1 + j\omega'\right)$$

when $\sigma_1^2 = \sigma_2^2$ this simplifies to:

$$\phi_{rRDT_K^{\#}}(\omega') = \pi^K [\omega' \cdot \operatorname{cosec}(\pi\omega')]^K$$

The result has to be scaled to accommodate the division by the number of averages:

$$\phi_{rRDT_K}(\omega') = [\frac{\pi\omega'}{K} \cdot \operatorname{cosec}(\frac{\pi\omega'}{K})]^K \tag{B.9}$$

The inverse Fourier transform of (B.9) for $K > 1$ creates analytical difficulties in the calculation of the distribution, but it can be calculated for given values of parameters, and fig B.9 shows the distributions for K values of 1 to 10. Due to $\sigma_1^2 = \sigma_2^2$ the mean values of all the distributions are zero. As the number of averages increase, the distributions become narrower.

It is found by calculation for the individual distributions shown, that the variance is inversely proportional to K and from chapter 5.5 the variance of the non averaged RDT distribution is $\pi^2/3$ therefore the variance of the averaged distribution for any given number of averages is twice that of the distribution before the RDT has been applied (B.10):

$$V_{MLR_K} = \frac{\pi^2}{3K} \tag{B.10}$$

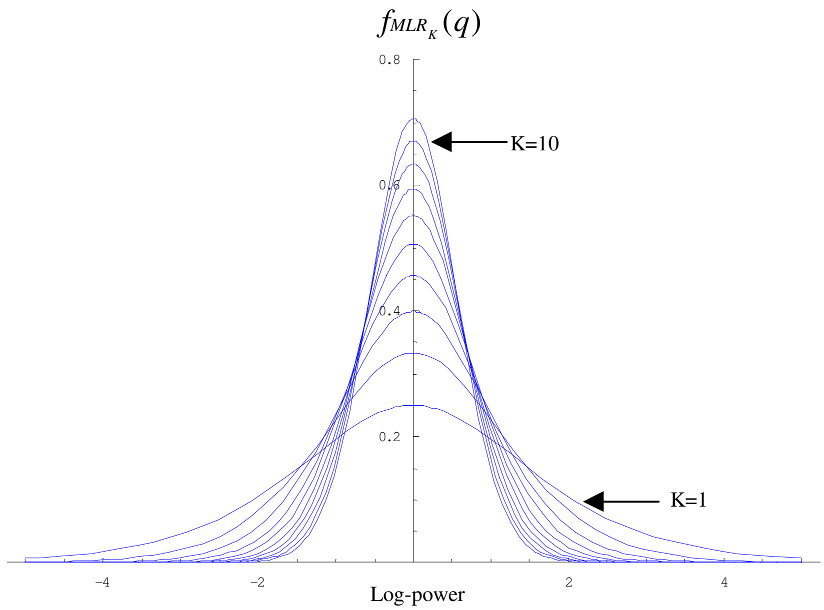

Fig B.9: Distribution of the RDT under null conditions for the mean of the log method, shown for K values of 1 to 10.