Applying the Histogram Multi-Fractal formalism for extracting sub-LUN skew

D Gueorguiev Jan 3rd , 2013

The distribution of IOs along a LUN may have multi-fractal nature. The goal of this report is to explore the Histogram method and its feasibility to obtain a closed form expression estimating the Sub-LUN skew for various workload mixtures.

By obtaining the *multifractal singular spectrum* and *the multifractal singular exponents* of the generalized LUN workload we may be able to approximate the sub-LUN skew with a parametrized closed form expression. Here the parameters may be LUN size, write percent and IO size.

The steps in this process are as follows:

1. Write a code which implements the WTMM to obtain singular spectrum and singular exponents of the fractal signal.
2. Verify the computation process against a fractal distribution with known analytical expression for the singular spectrum and singular exponents
3. Write a module which parses btp/trace workload and transforms it into a Devil staircase signal which can be supplied as an input into the multifractal analysis code
4. Produce singular spectra for various LUN workloads with different LUN size, write percent, IO size, and interval duration
5. Create a closed form expression fitting the produced spectra for various values of the LUN size, write percent , IO size and trace duration.
6. Handle traces with significant number of sequential IOs. We hope that such workloads can be modeled as non-everywhere singular fractal signals i.e. fractal signals which have non-fractal component. Those signals can be analyzed relatively easily by WTMM formalism (see [1]).

What is the multifractal singular spectrum of LUN IO workload

The multifractal singular spectrum represents the degree of clustering of the IOs along the LUN. The singular exponent identifies the density of the clusters

Figure 1: Simplified LUN IO layout. L – the LUN size; B – the block size

Generalized Cantor Set

To illustrate the WTMM method and its application we will consider Generalized Cantor sets as a signal source. Starting with a line of unit length subdivided into *b* sectors we discard sectors from it. For each of the present sectors we apply the same step recursively. Then on the -th iteration we will have a set of sectors each of length interspersed with holes between them. With each sector it is associated a non-zero measure obtained in the following way – at the first iteration the -th sector is assigned a value , . At the second iteration this sector is subdivided into sectors the measure of which is formed as , and so on. Thus on the -th iteration the measure of a sector takes the form , . The holes can be represented as sectors with zero measures. Then the fractal set can be uniquely characterized by its generating vector giving the measure distribution on the first iteration where can be zero in which case it indicates a hole at that location.

Figure 2: Generalized Cantor set with generating vector [p1, 0, p2, p3, 0, p4, 0, p5] . Here b = 8 and k = 3.

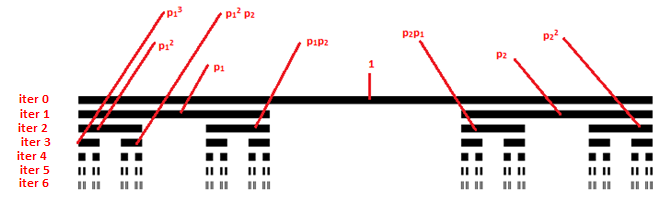


Figure 3: Six iterations of Cantor set with generating vector [p1 0 p2]. On the sixth iteration the fractal object is subdivided into number of sectors with non-zero measure.

Example 1 – generalized Cantor set with weights p1=0.6 and p2=0.4, b = 3, 8 iterations, N = 6561 sectors,

The smallest feature is ymin = 6.5536 x 10-4 and the largest feature is ymax = 0.016796 .

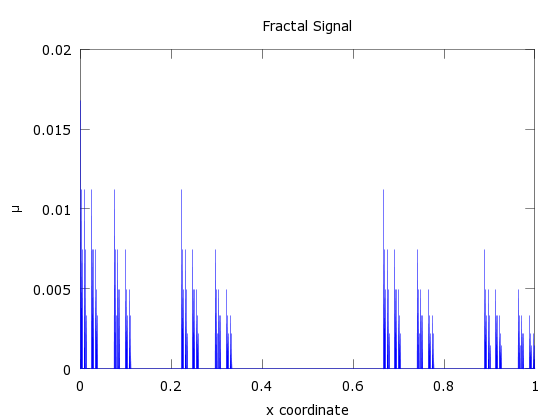
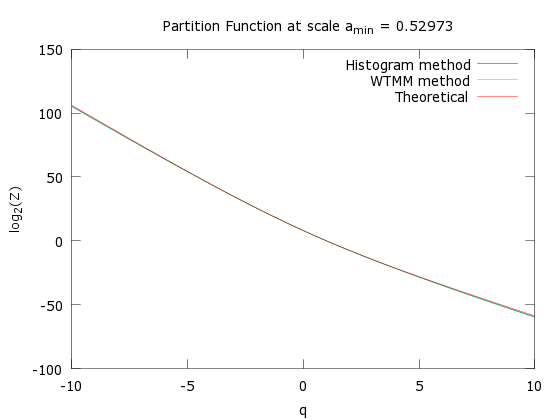
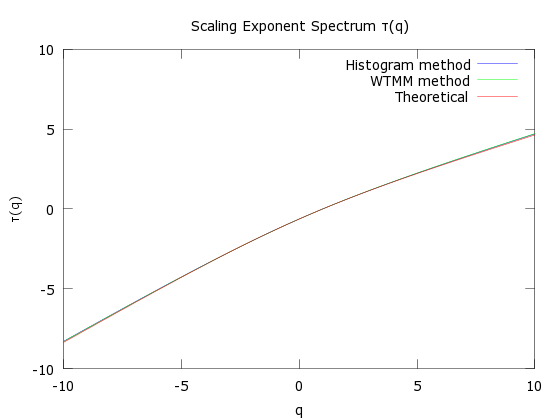
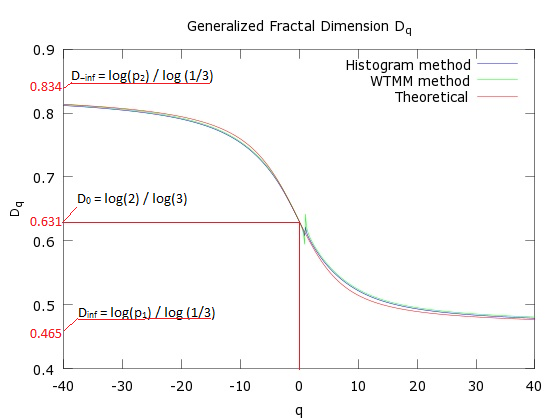
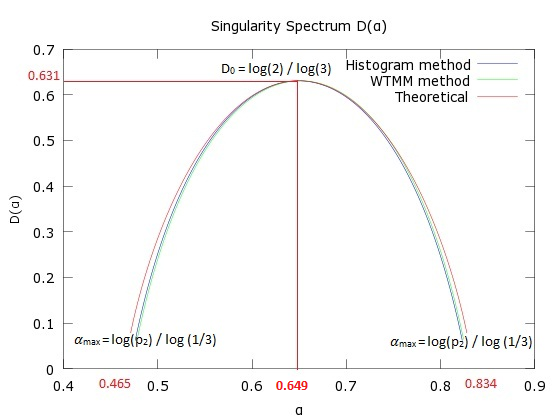


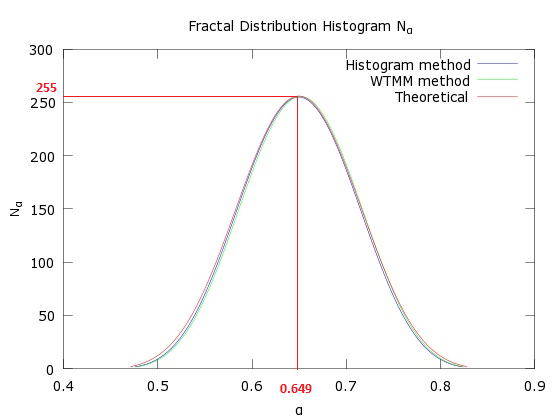
Figure 5: The fractal signal. To each sector of the fractal object it is prescribed a measure value which has the form where *n1*+*n2*=8

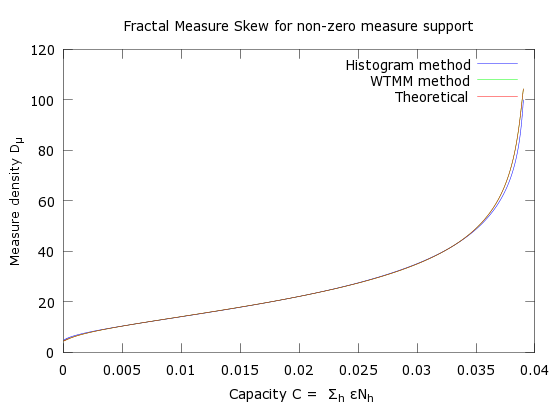






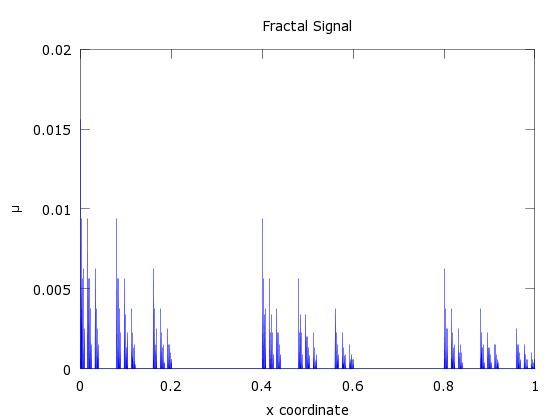


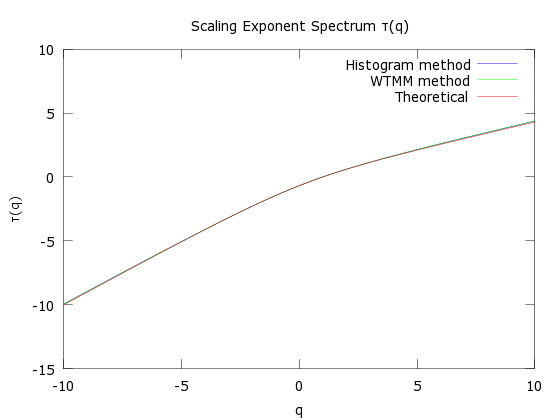
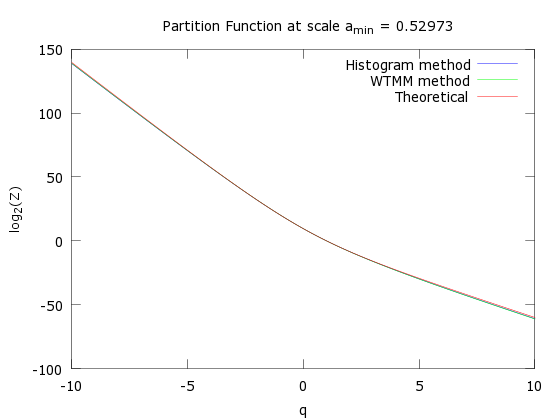


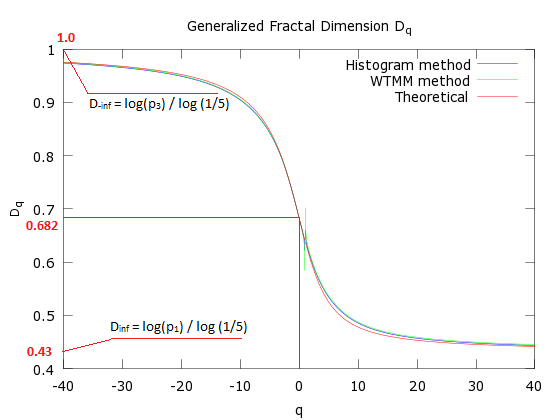


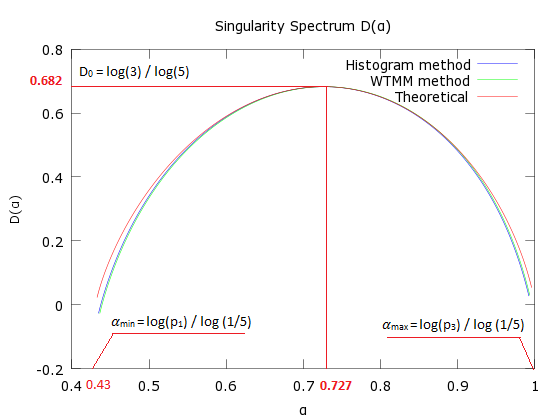
Example 2 – generalized Cantor set with weights p1=0.5, p2=0.3 and p3=0.2, base = 5, N = 15625 sectors, 6 iterations

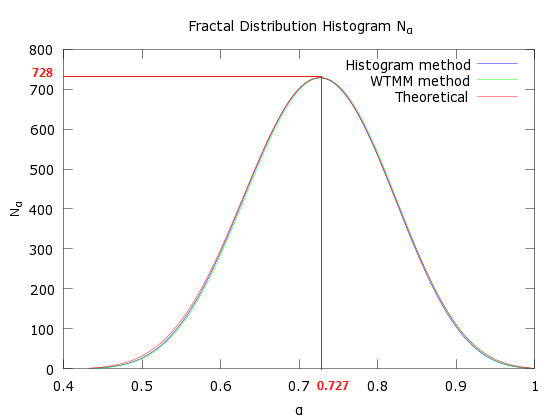
The smallest feature is ymin = 6.4 x 10-5 and the largest feature is ymax = 0.015625 .

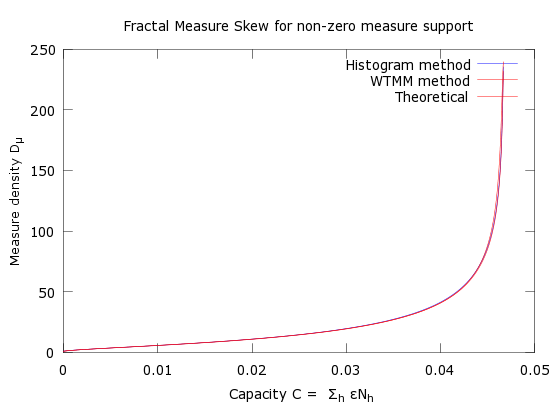












Example 3 – generalized Cantor set with weights p1=0.5, p2=0.3 and p3=0.2, base = 4, N = 4096 sectors, 6 iterations

The smallest feature is ymin = 3.2 x 10-5 and the largest feature is ymax = 0.015625 .

References

[1] The Multifractal formalism revisited with wavelets. J.F. Muzy, E. Bacry, A. Arneodo, *International Journal of Bifurcation and Chaos* Vol. 4, No. 2 (1994), pp 245-302

[2] Wavelet based multifractal formalism : applications to DNA sequences, satellite images of the cloud structure and stock market data in *The Science of Disasters*, 2002, eds A. Arneodo, B. Audit, N. Decoster, J.F. Muzy, C. Vaillant, *(Springer, Berlin)*, pp 26-102

[3] Direct determination of the f(𝛼) Singularity spectrum. , A. Chhabra, R. V. Jensen, *Physical Review Letters*, Vol. 62, Number 12, pp 1327-1330