STATISTICAL STUDY OF HARD X-RAY SOLAR FLARES

Deborah Leddon, B.A., B.S., M.S.

Dissertation Prepared for the Degree of
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS
December 2001

APPROVED:
Paolo Grigolini, Major Professor
Bruce West, Minor Professor
William Deering, Committee Member
Michael Fanelli, Committee Member
Jim Roberts, Committee Member
Sam Matteson, Chair of the Department of Physics
C. Neal Tate, Dean of the Robert B. Toulouse School of Graduate Studies

Leddon, Deborah L. A Statistical Study of Hard X-Ray Solar Flares. Doctor in Philosophy (Physics), December 2001, 4 tables, 21 figures, 78 references, 30 titles. The results of a statistical study of hard x-ray solar flares are presented in this dissertation. Two methods of analysis were used, the Diffusion Entropy (DE) method coupled with an analysis of the data distributions and the Rescaled Range (R/S) Method, sometimes referred to as "Hurst's method". Chapter one provides an introduction to hard x-ray flares within the context of the solar environment and a summary of the statistical paradigms solar astronomers currently work under. Chapter two presents the theory behind the DE and R/S methods. Chapter three presents the results of the two analysis methodologies: most notably important evidence of the conflicting results of the R/S and DE methods, evidence of a Lévy statistical signature for the underlying dynamics of the hard x-ray flaring process and a possible separate memory signature for the waiting times. In addition, the stationary and nonstationary characteristics of the waiting times and peak intensities, are revealed. Chapter four provides a concise summary and discussion of the results.

ACKNOWLEDGEMENTS

I am most grateful to my dissertation advisor, Paolo Grigolini, for his guidance and support throughout the time of my dissertation research. I am thankful to Dr. Bruce West, Dr. Duncan Weathers, Dr. Don Kobe, Dr. Sam Matteson, and Dr. William Deering for their continued faith and encouragement of my efforts over the years which at times must have seemed a lost cause. Without their ongoing support, I could not have finished my research.

To Carol Bowden, Physics Department secretary, who has guided me through the nether worlds of university rules and red tape, who has sheparded countless numbers of physics students to their degrees and beyond, and who has generally bailed me out, God bless you and thank you.

To Ronnie Trietsch, Physics Department technician, who has rescued me a countless number of times when I locked myself out of my office and my filing cabinet, and who coordinated the cleanup effort after a pidgeon got locked in my office over a weekend; many, many thanks.

I am grateful to Kay Littler who provided much needed instructional support and materials, greatly easing my work load in preparing for class instruction and to Dr. Atul Sharma, pediatrician and scientist extraordinaire, for Mathematica and Tex help and for inspiring me overall.

To the multi-talented folks in the Mathematica MathGroup for their help not only with Mathematica code, but with computational techniques and software in general, many thanks for your time and consideration. To the fine folks in the College of Arts and Sciences computer support, the UNT Science Library and the Interlibrary Loan Office, thanks so much for your timely efforts in support of my research.

I am grateful to Mauro Bologna, Roberto Failla, Massi Ignaccolo, Heather Appleby, Gayla Scully, Reiko Lukic, and my coworkers in the astronomy group for making my graduate student life more bearable and often, fun.

To Markus Aschwanden who answered uncountable questions and endured more requests for data information and preprints than I care to admit to and in so doing managed to teach me a bit of solar physics, thank you from the bottom of my heart.

To Massi Ignaccolo whose brilliant grasp of the most subtle and beautiful aspects

of statistical physics amazes me, and whose dedication and hard work continues to inspire me, thank you so much from my heart and my head!

To my husband, James, thank you for your loving support and encouragement, and for understanding what is generally not understandable; the graduate research experience.

To my greatest treasures in life, my sons, Jim and Sam, blessings and Todah Ma'od for letting me see that *the eyes of a child are truly, the eyes of discovery.*

And finally, to God I give thanks and praise for allowing me to experience the wonders of his universe. To be a subject of his infinite wisdom, his love, and his creation, I often wonder:

מה אנוש כי תזכרנו ובן אדם כי תפקדנו

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 The Solar Environment

In order to understand solar flares and their hard x-ray components, hard x-ray flares (also known as hard x-ray bursts), it is useful to examine the environment in which they occur. The sun itself is a complex dynamical system consisting of multiple layers or zones whose processes are often unobservable or must be inferred through whatever observations we can obtain. How these zones are related and by what mechanisms they generate and influence flaring activity is not completely understood.

Figure [1.1] illustrates the layered sun system. Energetic photons generated by nuclear burning in the core travel through the overlapping radiation zone, a region in which the pressure, density and temperature is inversely related to the increasing radius. At approximately 0.74 to 0.75 solar radii [1], all of the energy is absorbed by the plasma medium of the convection zone with the result that columns of hot gas rise, cool, and then descend creating a characteristic pattern of convection cells. This zone is characterized by a complex hierarchy of convection cells whose large sizes at the base of the zone decrease to approximately 1000 km across at the top of the zone. Above the convection zone lies the photosphere where the gas density is too low to sustain convective type motion. The photosphere is seen as the visible surface of the sun since most of the visible light we receive is emitted from this zone. The photosphere lies at the bottom of the *Solar Atmosphere* which is comprised of the chromosphere, the transition zone, and the corona. It is at the base of the convection zone where magnetic fields are believed to be generated [1]. The generation of magnetic fields within the turbulent conducting medium of the interior, their subsequent rise into the photosphere and their resultant dynamical behavior is what generates a wide range of phenomena known as *Solar Transients* [2] or *Solar Activity* [3]. The spatial and temporal characteristics of the transient phenomena is a product of the interplay between the nonlocalized dynamo, solar rotation, and the photospheric and turbulent

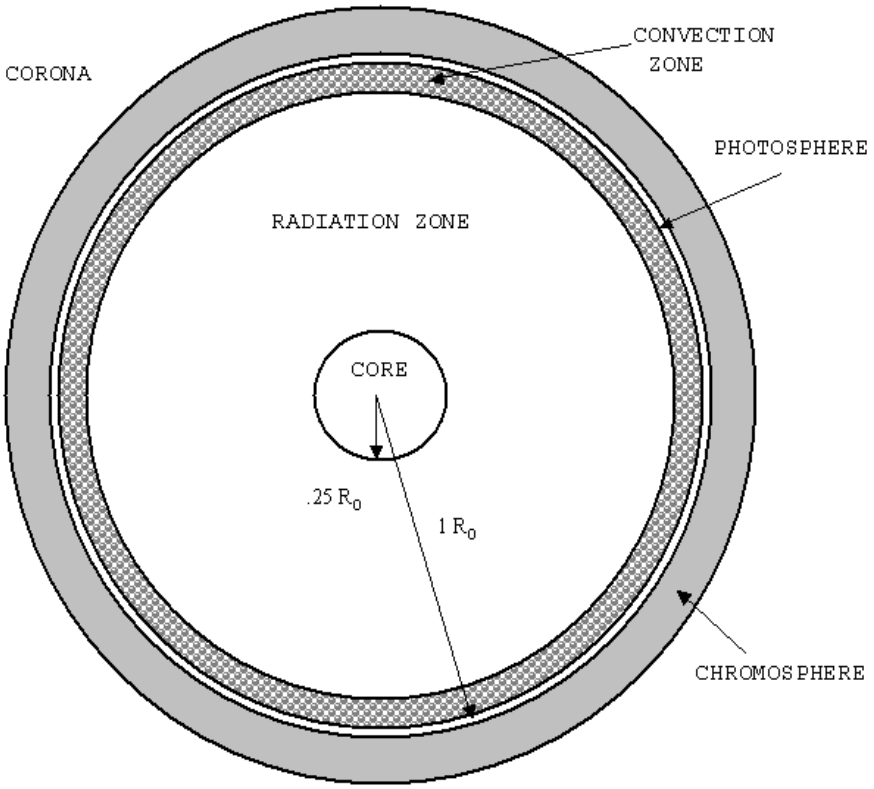

Figure 1.1: Illustration of the layered regions of the internal sun. Magnetic fields are believed to originate at the base of the convection zone.

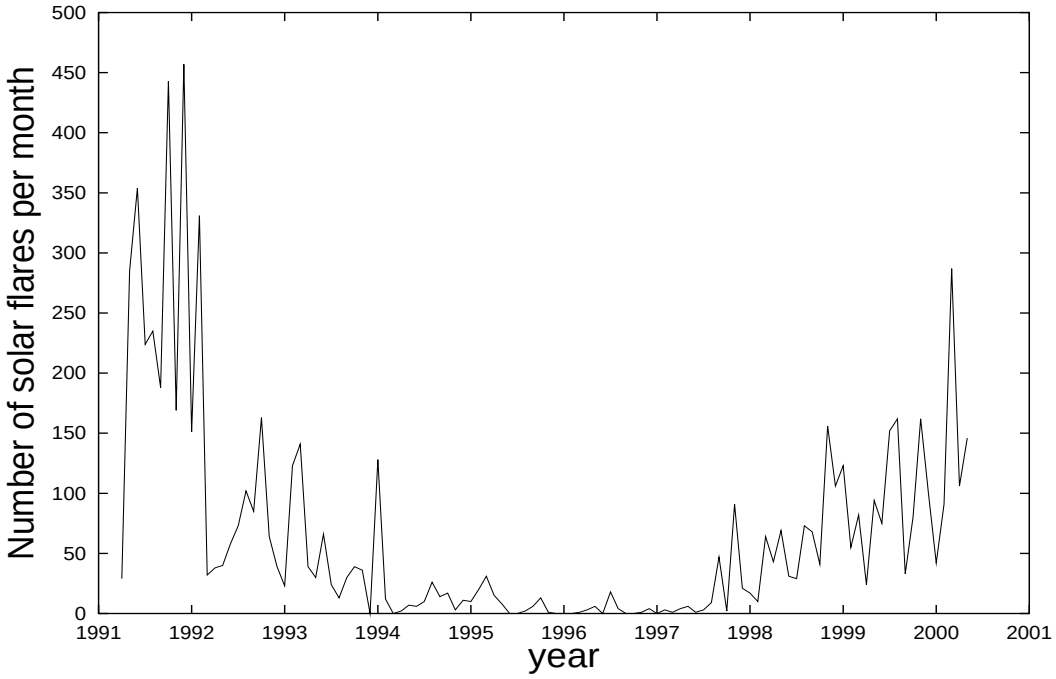

Figure 1.2: Illustration of flaring rate per month. Courtesy: N.Scafetta.

motions of the sun's outer atmosphere. Flares are just one example of the incredible variety of transient active phenomena which include coronal mass ejections [CMEs], spikes, ejecta, and more [6]. These activity complexes are governed by a system-wide regime of actions and responses propagating through both temporal and spatial scales in such a way as to produce an overall cyclic behavior known as the *Solar Cycle* [7]. With respect to flaring events, it is known that rates of flaring and the magnitude of their peak intensities directly correlate with the solar cycle [8]. Figures [1.2] and [1.3] illustrate this phenomenon from the perspective of the hard x-ray (HXR) data examined in this study. The flaring rate per month was calculated from the HXR flaring events ranging from April, 1991 to May, 2000. The rate of flaring seen in Figure [1.2] directly correlates with solar cycle 22 whose maximum occurred in 1989 and minimum in 1996. The rate increases again for cycle 23 whose maximum was predicted to occur around April-May 2000 [9]. Tracking of the solar cycle is done in a number of ways, irradiance and 10.7 cm flux values, and most notably through the

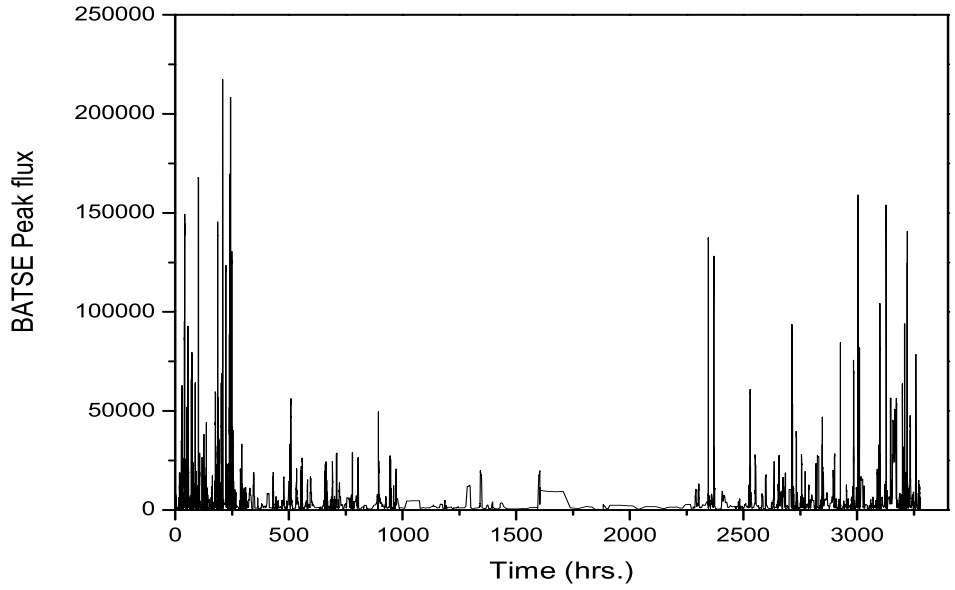

Figure 1.3: Hard X-Ray Peak Intensities (Two point smoothed). Time in hours from April 1991 to May 2000. Data represents the cyclical behavior of the solar total irradiance output starting with the downward phase from the solar cycle 22 maximum of 1989.

recording of sunspot numbers. Predictions of many research groups generally make good approximations to the empirical record through the use of dynamo theory [10] and past sunspot or sunspot group number data, [11, 12].

Flaring peak intensities (number of photons s^{-1} per detector area), the maximum number of HXR photons emitted at a certain time during the flaring event known as the *peak time*, also increase in response to the cyclic behavior of the cycle as seen in Figure [1.3]. Both of these observed effects are just a small part of the many fascinating aspects of the solar cycle itself providing strong evidence of the interconnectedness of both long and short term memory signatures inherent within the solar dynamical system.

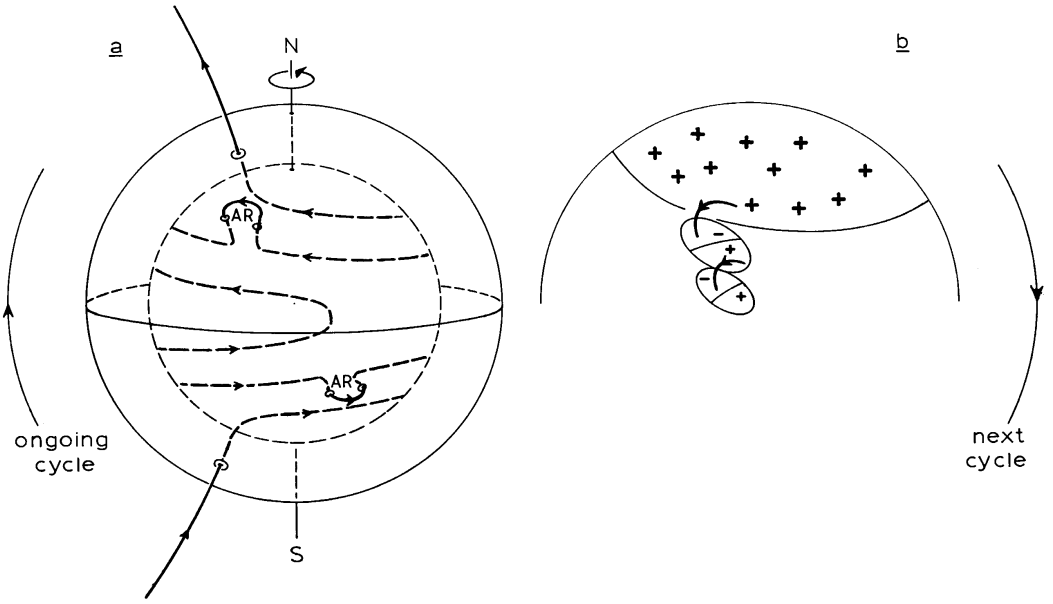

Figure 1.4: Effects of Solar Rotation on Poloidal Magnetic Field. Part (a) depicts the nearly toroidal magnetic field configuration produced by rotation stretching of the field in the E-W direction. Ω loops seen in active regions (AR). Part (b) illustrates reconnecting loops whose poloidal component dirction matches that of the next cycle. Courtesy: C. Zwaan, 1996, via C. Scrijver, Lockheed Martin Solar and Astrophysical Lab.

1.2 Hard X-Ray Flares

Solar Flares are rapid, high energy release events occuring within the active regions of the Sun's atmosphere and are characterized by strong magnetic fields. Flaring geometry is not unique in that there exists a variety of flaring configurations which may include point sources [1], and single and multi-ribbon loops [1, 13]. The origin of these magnetic loops in which flares occur and their associated active regions has been explained, in part, by $\alpha - \omega$ dynamo models [4, 5]. The effects of solar rotation seen in Figure [1.4] yield significant changes in both the overall and localized spatial geometries of the rising magnetic fields. If we assume the initial field to be poloidal, subsequent differential rotation *("ω effect" of the $\alpha - \omega$ model)* induces a complex

pattern of toroidal movements in which 'kinks' or loops can form. Bipolar active regions are then formed by the breakthrough of "Ω" shaped loops [6]. The Coriolis force due to solar rotation induces the *"α effect"*. This is a twisting of the rising Ω loop until it orients itself to the initial poloidal direction. The complex character of these fields may initially be present as they emerge from the underlying convection zone, it may evolve through stages of flux emergence, or may form through continuing interactions with either the ambient chaotic medium or collisions between separate magnetic bipoles [3]. The onset of a flaring event is believed to be due to an instability or set of instabilities either within the magnetic field itself or as a result of the interaction between the magnetic field and the surrounding solar atmosphere. Thermal and nonthermal radiation is emitted from the flare in phases which encompass emissions throughout the entire range of the electromagnetic spectrum. The *pre-flare phase*, may contain any one or more of a range of phenomenon often classified as *precursors* [14]. These events may include phenomena such as pre-flare heating visible in the microwave, UV, and x-ray spectrums, filament eruptions and surges, and occasionally the onset of coronal mass ejections (CMEs) [6]. The initial release of stored magnetic energy due to a triggering instability is believed to be responsible for this range of phenomena [14]. The second phase, known as the *impulsive* or *flash phase* [14] includes the sudden release of energy seen as spikes in the radio, HUV, and hard X-ray (HXR) spectrum in the form of non-thermal radiation. Since HXR events most notably occur during impulsive flaring events [1], we can look at a generalized picture of an impulsive flare as seen in Figure[1.4] in order to see what is happening. According to the *Standard Model* [3], which applies to two-ribbon flares, the equilibrium state of stored magnetic energy is disrupted resulting in an ejection of chromospheric and/or coronal material. Reconnection in the acceleration region initiates the release of stored magnetic energy. Local heating of the surrounding plasma causes particle aceleration in both the upwards and downwards directions. Nonthermal electrons with energies, $E \geq 20keV$ accelerate in the downward direction experiencing Coulombic collisions with charged particles in the denser chromospheric region resulting in the emission of HXR bremsstrahlung radiation. Coulombic collisions, by definition, happen when an electron is deflected in the Coulomb field of the target charged particle. The energy

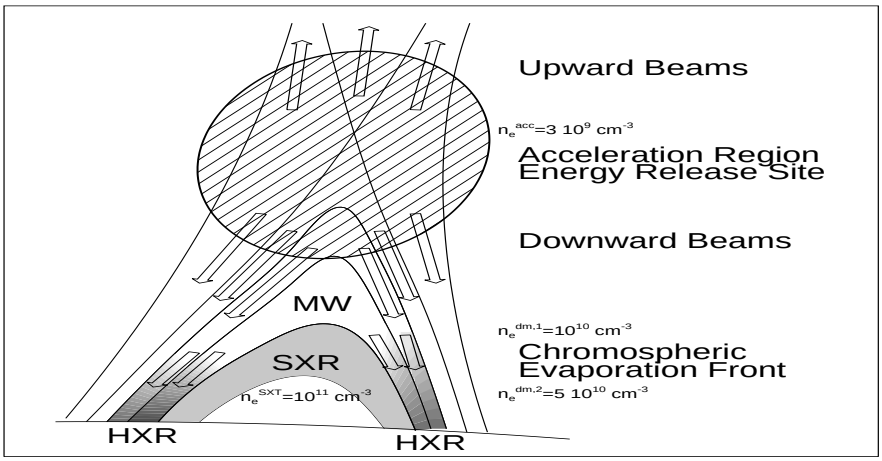

Figure 1.5: Illustration of a flaring event. Note the increase of electron density (n_e) with decreasing height. It is in this denser region of the chromosphere and close to the footprints of the flare loop that HXR emission can occur due to collisions. (Courtesy: M.J. Aschwanden, Lockheed Martin Solar and Astrophysical Labs, Palo Alto, CA)

spectrum of HXR emission contains energies in the range 10 keV to 1MeV. The lower end of the energy range; $E \approx 1 - 10$ keV corresponds to HXR emission while the higher energy range end; 1 - 100 MeV corresponds to Gamma emission. Heating of the plasma also produces an overpressure resulting in the evaporation of plasma upwards eventually filling the entire loop. This is believed to be the mechanism responsible for the production of soft x-ray (SXR) emission seen in the figure.

The power law signature of these bursts,

$$\frac{dJ\,(E)}{dE} = CE^{-\gamma} \tag{1.1}$$

where $J\,(E)$ is the photon flux (photons $\mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{keV}^{-1}$) and γ, denotes the spectral hardness index, generally accounts for in a simplistic way, the explanation of the event as an impulsive one. Hence, the collection of these emission spikes, the inverse power law signature and the preponderence of non-thermal radiation points to particle acceleration at the reconnection site of the flaring event as the most likely explanation for the source of the hard x-ray emission. The *main phase* [3] of the event follows with the release of the bulk of the radiation which is largely thermal and can be attributed to plasma heating.

While it can be seen that x-ray bursts or flares are just part of the overall flaring event itself, their occurance during the impulsive phase is often seen as important evidence of the dynamical triggering and initial energy release processes. A number of studies have addressed this aspect; Masuda et.al. [18], noting several types of hard x-ray sources and their possible relationship to impulsive phase reconnection, and Petrosian and Donaght [19] noting the spectral characteristics of hard x-ray emission with respect to turbulent scattering and electron acceleration. Observations by Harra-Murnion et al. [20], noted the overlaying of soft x-ray loops with hard x-ray emission suggesting magnetic reconnection.

1.3 Statistics of Solar Transients

It is reasonable then that investigators turn to statistical methods in hopes of making sense of flaring behavior. Up until recently, these methods have generally been based on the assumption that fluctuations leading to flare generation and energy release are independent, random processes that can be characterized by the finite moments of distributions subject to the restrictions of the Central Limit Theorem [21]. Within the framework of ordinary Brownian motion (OBM), two of these distributions, the Gaussian and it's related distribution, the Poisson, are most often used in the analysis of flare time series. A major implication of this use, however, is that the frequency distribution exponents and hence the scaling of the time series being analyzed is often measured incorrectly or is obscured.

It is now generally accepted that most physical phenomena do not evolve according to well defined inventories of physical processes but rather exhibit characteristics of systems sensitive not only to the interactions of internal fluctuations but to environmental ones as well. The composite traits of randomness and order displayed by these systems are evident in the invariant scaling of the system's statistical behavior in both the temporal and spatial scales and by moments which may be non-finite. In addition, the ubiquity of the inverse power law, which is noted not only in flare observations but elsewhere in the observation and analysis of other physical phenomena is an example of this universal behavior [22, 23]. Examples of other nonlinear systems which scale in such a way include stockmarket indices [24], heart interbeat intervals[25], teen birth behavior [26], DNA sequences [27], and more.

Past efforts to relate frequency distributions to the underlying dynamics of the flaring process have involved not only the assumptions of gaussian and/or Poisson statistics but also the fitting of these distributions to data. In the case of waiting time distributions, results differ. For example, the waiting time distribution of hard x-ray flare events recorded for 8 years by the ICE/ISEE3 spacecraft and examined by Wheatland et al [28] as a Poisson process, evidenced correlations between events, particularly an overclustering with respect to short waiting times. Normally this is taken as proof of sympathetic flaring; i.e.; where one flaring event triggers another,

however, the authors were unable to distinguish between this result and the possiblity that individual flares might be comprised of several bursts events. Biesecker [29], using using one year of x-ray data from the BATSE catalog found a waiting time distribution based on a time varying poisson in which the mean flaring rate was not finite. Using waiting time distributions fitted on 20 years of GOES x-ray flaring data, Boffetta et al [30] determined that the inverse power law signature was indicative of MHD turbulence and in contradiction with the time dependent Poisson statistics of Wheatland [31] for the same set of data. Aschwanden et al [32] implemented the logistic function, as the form of the dissipated energy rate, an assumption, which does in part address the obvious stochastic nature through the evolution of cascading bifurcations. However, fluctuations were again Poisson distributed with frequency distributions seen as inverse power law or exponential in form. The interpretation of results such as inverse power law forms has often been seen as evidence that the underlying dynamics is manifested as a self organized state which has exceeded it's critical limits of stability.

Within the last few years, efforts have been made to better describe the obvious deviations from ordinary brownian motion and occasionally fractional brownian motion in the analysis of solar time series in general. Use of the rescaled range analysis technique to uncover long term memory signatures in solar activity began with the analysis of average monthly sunspot numbers yielding H values of 0.9 indicating a high level of memory [33]. Other studies, based on Carbon 14 solar proxy data [34] and solar doppler rotation data [35] yielded Hurst values indicating a high degree of persistance. Recently, a Hurst value of 0.74 was determined for the Hydrogen alpha flaring index, Q, a measure of flaring activity which correlated well with the measure of sunspot activity [36]. More recently, Lepreti et. al, [37] found that the waiting time distribution of hard x-ray flares retrieved from the Geosynchronous Orbiting Satellite System (GOES) was inconsistent with Poisson statistics and was better characterized by the Lévy symmetrical long-time distribution.

The investigation of the HXY emission seen during the impulsive phase of a flare is the subject of this dissertation. In this study, the waiting times and peak intensities of HXY events from 1991 to 2000, extracted from the solar flare catalog

list produced by the Burst and transient source experiment (BATSE) onboard the Compton Gamma Ray observatory (CGRO) satellite, have been examined using the the method of rescaled range analysis (R/S) and the method of diffusion entropy (DE) analysis. The purpose of this study was a general one is which the scaling and the underlying statistics of the diffusing processes are examined. The results, which include the presence of a dual memory signature in the waiting times, persistant memory in the intensities, detection of Lévy scaling, and the identification of stationary/nonstationary regimes in the data are believed to be the first evidence of its kind. We shall also see that the DE analysis and the joint use of waiting times and intensities distributions bypasses the ambiguities of the rescaled range method (R/S analysis) which is based on the assumption that the process is characterized by Gaussian statistics.

CHAPTER 2

THEORY

2.1 Rescaled Range Statistics.

Typically, time series are modeled with Gaussian or Poisson statistics which suppress the fluctuations of the underlying dynamics of the system. Realization of the inadequacy of these descriptions can be noted through the inspection of the central moments such as the mean and the variance. For complex systems exhibiting inverse power law behavior, the absence of a finite time and/or spatial scale implies the self-similarity property for the measured parameter, L;

$$L(\alpha r) = \beta L(r) \tag{2.1}$$

The simplest solution to Eq.[2.1] is of the form

$$L(r) = Ar^{mu}, \tag{2.2}$$

where A and μ are assumed constant. Hence, for systems which scale in such a way, it is obvious that the central moments will no longer maintain a finite value;

$$< L(r)^n >= \int_{-\infty}^{\infty} P_L(r) L^n dr, \tag{2.3}$$

where

$$P_L(r) \sim r^{-\mu} \tag{2.4}$$

and n is the index for the moments, e.g.; n = 1 (first moment) and n=2 (second moment or variance). What Eq.[2.1] says is that fluctuations at small scales lead to larger fluctuations at larger scales. These larger fluctuations tend to increase the variance in such a way that the mean and/or the variance no longer assumes finite values. This scaling invariance is the result of the propagation of a systems memory

as it evolves through time. The knowledge that a system's memory will produce measurable effects that can be analyzed statistically, implies greater power in attempting to decipher the underlying dynamical process. Since we know the moments of a distribution are non-finite and are a function of the spatial and/or temporal scaling, it has often been assumed that an examination of the scale dependency of these moments may yield important information. The R/S method sometimes known as the 'Hurst analysis', is an example of a second moment technique yielding an exponent which is a measure of the power law scaling. The originator of this technique, H.E. Hurst, developed the method in order to determine the optimal reservoir size in the design of the Aswan Dam in the Nile river basin. He noticed in over a thousand years of Nile river height data, that there was a persistency effect, in other words, the river volumes were not randomly distributed in time but evidenced periods of low volume flows and high volume flows [38]. His conclusion was that some sort of memory effect was in place between the river flow dynamics and the outside environment. Mandelbrot [39, 40] extended the rescaled range technique by incorporating the Hurst exponent as a scaling parameter under the assumption of Fractional Brownian motion (FBM). In this paradigm, data points are considered to be random gaussian fluctuations but with an inverse power law signature to account for scaling and a parameter known as 'H', the Hurst exponent, which can achieve values in the range [0,1]. H can reveal three different regimes of scaling. In the 'persistant' regime, $0.5 < H \leq 1.0$, data fluctuations will be positively correlated, meaning that a fluctuation increment in one direction is likely to be followed by another fluctuation increment in the same direction. For the 'anti-persistant' regime, $0 < H < 0.5$, the opposite effect is achieved; fluctuation increments in one direction are likely to be followed by increments in the opposite direction. Ordinary Brownian motion is recovered for $H = 0.5$, a random walk situation in which no memory effects can be present. The method examines the accumulated departures of the mean fluctuations (in this study, the waiting time and peak intensities data points) within windowed partitions of the data which are then normalized by the standard deviation. Any long term memory effects are evidenced in these accumulated mean departures.

For a data set containing the fluctuations; ξ_i, we denote the mean value

$$< \xi >_T = \frac{1}{T} \sum_{i=1}^{T} \xi_i \tag{2.5}$$

in order to define a collection of departures from the mean value for windows of the data,

$$X(t,T) = \sum_{u=1}^{t} (\xi_u - < \xi >_T). \tag{2.6}$$

Next, the range of differences, $R(T)$ between the maximum and minimum values of X(t) is defined as

$$R(T) = \max_{1 \le t \le T} X(t,T) - \min_{1 \le t \le T} X(t,T). \tag{2.7}$$

Finally, we divide the $R(T)$ by the standard deviation, $S(T)$, for each partition

$$S(T) = \left[\frac{1}{T} \sum_{i=1}^{T} (\xi_i - < \xi >_T)^2 \right]^{(1/2)}. \tag{2.8}$$

Note, that since different values of $R(t,T)$ for each interval start at different times, t, the division by the standard deviation normalizes and rescales the range of fluctuations within the same time period. The rescaled range relationship is then defined as

$$R(T)/S(T) = aT^H \tag{2.9}$$

In and of itself, the Hurst method can clearly indicate deviations from ordinary brownian motion and persistency effects within data, however, it cannot be used to assess whether or not the data is stationary or nonstationary or provide a measure of the correct scaling in the case where the underlying distribution is non-gaussian. Let us examine what is meant by the term 'correct scaling'. The Hurst exponent is based on the deviations from mean values divided by a standard deviation for partitions of a data set. Hence, this scaling is a measure of the changes in the first and second moments as the diffusion process evolves. In addition, these moments are Gaussian, subject to the restrictions of the Central Limit Theorem (CLT) [41] which states that

the distribution of the sum of N independent and identically distributed variables with a finite variance converges to a Gaussian as $N \rightarrow \infty$. Therefore we are not looking at the scaling of the distribution spreading, even under the Gaussian assumption. This limits the viability of our analysis in that a statistical signature applicable to a particular set of dynamical behaviors cannot be detected for a stationary data set. In fact, it is now known that the Gaussian assumption is no longer adequate to describe many types of phenomena which evidence intermittency or repeated bursts within bursts types of behaviors [42, 43, 44]. To see this, we may look intuitively at the diffusion process where for a time series of data points, ξ_i, we define a brownian trajectory,

$$x\left(N\right) = \sum_{i=1}^{N} \xi_i. \tag{2.10}$$

The continuous form of Eq.[2.10]

$$\dot{x} = \xi \tag{2.11}$$

allows us to obtain through integration and the definition of the second moment,

$$< x^2\left(t\right) > = < \xi^2 > \int_0^t dt' \int_0^{t''} dt'' \Phi_\xi \left(\mid t' - t'' \mid \right) \tag{2.12}$$

where Φ_ξ is the correlation function. We assume that $\Phi\left(\xi\right)$ is stationary and in the asymptotic time limit, approaches the value,

$$\lim_{t \rightarrow \infty} \Phi_\xi\left(t\right) \sim \frac{1}{t^\beta} \tag{2.13}$$

where $0 < \beta < 1$. It has been proven [45] that the asymptotic correlation function, Eq. [2.13], allows Eq. [2.12] to yield;

$$< x_t^2 > \sim t^{2H} \tag{2.14}$$

where $H = 1 - \beta/2$. Now suppose we consider the case where it is unclear as to

whether moments of the data exist or not, the assumption is therefore, of a nonfinite first and /or second moment,

$$< x(t) >=< x^2(t) >= \infty \tag{2.15}$$

which leads us to consider distributions which may not contain finite moments and are known for their scaling invariance. A class of distributions fitting these two conditions and subject to the Generalized Central Limit Theorem (GCLT) [46, 47] is known as the Stable distributions, one of which, the Lévy distribution [48], appears quite ubiquitous in the description of phenomenological distributions [49]. The Lévy process rescales in the same manner as the gaussian;

$$P_L(x,t) = \frac{1}{t^\delta} F\left(\frac{x}{t^\delta}\right) \tag{2.16}$$

where $\delta = 1/(1+\beta) = 1/\alpha$ and β denotes the characteristic exponent which has the range of values, $1 < \beta < 2$. Note, that for $\beta = 2$, the Gaussian is obtained. The parameter α, is known as the Lévy characteristic exponent. Comparing the δ scaling parameter with the expression of hurst scaling parameter, $H = 1 - \beta/2$, it is quite obvious that the two are distinctively different.

2.2 Diffusion Entropy

At this point, one may ask, if FBM is clearly not sufficient to describe non-gaussian systems with long time memory correlations, then what statistical paradigm can give a better description? The answer is: no general statistical picture exists which can fully describe enhanced diffusion. Since second moment techniques are clearly inadequate in detecting the correct scaling behavior of a time series in which these moments may not exist, it is useful to consider another method which would address the evolution of the systems statistics in a different way. Here, we examine two basic considerations with regards to the behavior of the unknown underlying distributions; first, the spreading of the distribution in question as the system evolves and secondly, consideration of statistical distributions known for their scaling invariance. Let us

first examine the time evolution of a distribution, $p(x,t)$. If we start from an initial time and position, t_0 and x_0, where the system can be described as a Dirac delta function;

$$p(x_0,t_0) = \delta_{Dirac} \tag{2.17}$$

then allow the system to evolve through a spreading of the probability in such a manner that it eventually achieves the generalized inverse power law form;

$$p(x,t) = \frac{1}{t^{\delta}} F\left(\frac{x}{t^{\delta}}\right) \tag{2.18}$$

where delta is defined as the scaling parameter. At this point, we want to establish a connection between the initial probability, $p(x,t)$ and its value at a later time. Assuming that a rescaling exists between the final and intial distributions, we can examine the spreading of the distribution through the entropic relationship;

$$S(t) = -\int_{-\infty}^{\infty} p(x,t) \ln p(x,t)\, dx \tag{2.19}$$

where $S(t)$ is the familiar form of the Shannon entropy. We make the simplifying assumption that the statistics of the process defined by $F\left(\frac{x}{t^{\delta}}\right)$ remains unchanged. Substituting Eq. [2.18] into Eq. [2.19] results in a linearized form of the entropy,

$$S(\omega) = A + \delta(\omega), \tag{2.20}$$

where

$$A \equiv -\int_{-\infty}^{\infty} dy F(y) \ln[F(y)] \tag{2.21}$$

for

$$y \equiv \frac{x}{t^{\delta(t)}} \tag{2.22}$$

and

$$\omega \equiv \ln(t). \tag{2.23}$$

Equation [2.20] forms the basis of the methodology known as *Diffusion Entropy* (DE), an analysis technique developed by Paolo Grigolini and group [50] for the purpose of analyzing the scaling properties of stochastic time series. The method has now been successfully applied to a number of phenomena which include teen birth behavior [50], DNA sequences [51], hard x-ray solar flares waiting times [52] and human heart beat data [53].

The application of the method proceeds with a given time series of M fluctuating variables (ie. data points), labeled ξ_i. For most time series data, the finite length of the data presents serious problems in that a lack of significant data points may not evidence a reliable statistical signature. We can circumvent this problem by creating many trajectories of the data such that the length of the data is significantly increased. To do this, we first of all, select a range of integers, l, where $1 \leq l \leq M$. We refer to l as "time" so that for any given time, l, we can obtain $M - l + 1$ sub-sequences defined as

$$\xi_i^s = \xi_{i+s} \tag{2.24}$$

where $s = 0, ..., M-1$ For any one of these subsequences, a diffusion trajectory labeled with the index, s, can be defined as the position;

$$x^{(s)}(l) = \sum_{i=1}^{l} \xi_i^{(s)} = \sum_{i=1}^{l} \xi_{i+s} \tag{2.25}$$

In keeping within the context of a continuous time random walk (CTRW) [54] scheme we may assume each position as corresponding with the position of a random walker. We can then imagine each position as one which a brownian particle attains at time, l, after having made l jumps. The jumps made at each ith step have the intensity, $\mid \xi_i^{(s)} \mid$, and will occur in either one continuous direction or in a random backwards or forwards direction according to a coin tossing prescription. These three coin toss prescriptions or better known; walking rules, which will be employed in this analysis, will be discussed shortly.

In order to evaluate the entropy, one must define a probability in which the bin size is a partition of the x-axis into cells of size, $\varepsilon(l)$. We define the probability, $p_{i(l)}$,

that a particle can be found in the ith cell at time, l, as

$$p_{i(l)} = \frac{N_i(l)}{M - l + 1}. \tag{2.26}$$

Hence, the entropy of the diffusion process at time, l, and dictated by Eq. [2.20] is

$$S_d(l) = -\sum_i p_i(l) \ln [p_i(l)]. \tag{2.27}$$

It is important to mention that the assumptions behind the diffusion entropy method defined by Equations [2.18] and [2.19] in which the scaling parameter, δ, is normally a function of time, would not hold true during a transition process. As a consequence, the scaling detected by the DE method must be made in the asymptotic time limit as the scaling parameter $\delta(t)$ approaches a finite value, δ and after the transition has been completed. The methodology, itself, however, can be used to shed light into the regime of transition considered to be deeply connected to the foundations of statistical mechanics itself. According to Kinchin [55], the realization of canonical equilibrium is based fundamentally on the central limit theorem. A process resulting from the sum of N independent variables can be described by a gaussian distribution provided that N is large enough and that a finite second moment exists. The evolution of a physical process to large enough values of N, subject to the predictions of the central limit theorem, sees a transition from the microscopic to the macroscopic regime, where thermodynamics would apply. If the distribution of the variables does not have a finite second moment, then the central limit theorem no longer applies. At this point the distribution of the variables must be specified by an invariant distribution whose second moment does not exist and is thus subject to the predictions of the generalized central limit theorem. Let us examine intuitively how this transition from dynamics to thermodynamics obtains by examing the results of applying the Shannon entropy to a diffusion process. We allow the parameter, N, to assume the role of "time", t, where the microscopic regime corresponds to the fluctuations of ξ_i and the macroscopic regime corresponds to the fluctuations of the diffusion coordinate, $x(t)$. The time evolution of $\delta(t)$ to its time independent value corresponds to the transition from

dynamics to thermodynamics. This transition regime becomes much longer for the case of anomalous diffusion than for the case of ordinary diffusion. Hence, δ becomes a true measure of the scaling only at relatively large times, thereby setting limitations to the detection of scaling. In dealing with actual time series data, the finite size of the data produces saturation effects in the long time limit. As a consequence, the region where true scaling can be detected is an intermediate region of time preceded by an initial transition region and followed by the asymptotic saturation region. Due to the statistical limitations of our data, especially with respect to the AJM and SJM rules whose transition to the thermodynamical region of a finite δ parameter, can be quite slow, we employ the use of artificial sequences in order to obtain the proper scaling. This method overcomes the accuracy and detection limitations of the real data by generating artificial sequences with the same statistics as the real data and then searching for the distribution exponent, μ, that best represents the most accurate fitting with the DE results derived from the real data. The application and results of this method will be described in Chapter 3 where the results of the analysis are presented.

As mentioned earlier in this section, the scaling detected by the DE method is dependent on the random walk prescription one adopts for the diffusion trajectories. In this study, the application of the DE method utilized 3 random walk rules as described below. Note, that for the first two walking prescriptions, the data sequence must first be converted into a dichotomous one in which the data points are represented by zeroes and ones.

Asymmetric Jump Model (AJM rule): Walk according to the dichotomous value of the datapoint, ξ_i, then make a jump of one always in the same direction.

Symmetric Jump Model (SJM rule): Walk according to the dichotomous value of the data point, then make a jump of one in either a forward or backward direction according to a random coin toss prescription.

Flight Rule: Walk at fixed increments of time, then make a jump whose length corresponds to the value of ξ_i, always in the same direction.

The data points for the intensities data sets utilized in this study, correspond to the maximum intensity value of the HXY flares detected at time, t_i. The second set of

datapoints, τ_i, are the waiting times calculated as the difference between subsequent nearest neighbor maximum event times, t_i. Hence according to the AJM and SJM rules, we may look at the intensities and waiting times as corresponding to a random walker making instantaneous jumps at the occurance of a flare. The flight rule implies that the random walker makes jumps in one direction equal to values of the actual intensities and waiting times of the data.

The scaling corresponding to the AJM rule was studied by Grigolini et al., [56] and yields the scaling relationship between δ and the distribution exponent, μ as follows:

$$\delta_{AJM} = \begin{cases} \mu - 1, & 1 < \mu < 2 \\ 1/(\mu - 1), & 2 < \mu < 3 \\ 0.5, & \mu > 3. \end{cases} \tag{2.28}$$

The scaling of the SJM studied by Shlesinger [57] yields the prescription:

$$\delta_{SJM} = \begin{cases} 0.5(\mu - 1), & 1 < \mu < 2 \\ 0.5, & \mu > 2. \end{cases} \tag{2.29}$$

The Flight rule prescription is:

$$\delta_{flight} = 1/(\mu - 1), \qquad 1 < \mu < 3. \tag{2.30}$$

Figure [2.1] is a depiction of the scaling rules. It clearly demonstrates that on occasion, the detection of the scaling often requires the the joint application of more than one rule. For example, a detection of $\delta = 0.9$ in the AJM framework would yield 2 different values of μ; $\mu = 2.1$ or 1.9. Should a joint application of the SJM rule yield $\delta = 0.5$, then obviously the correct scaling corresponds to $\mu = 2.1$, else for $\delta \neq 0.5$, the scaling belongs to the region below $\mu = 2.0$. In addition, it is clear that the AJM and SJM rules reflect the phase transition character of the $\mu = 2$ limit while the flight rule does not. The scaling according to the flight rule for $\mu = 2$ yields values of $\delta > 1$ which means that the diffusion process falls into the superballistic region. It will be shown later that this rule provides a useful assessment of the scaling in that the fast

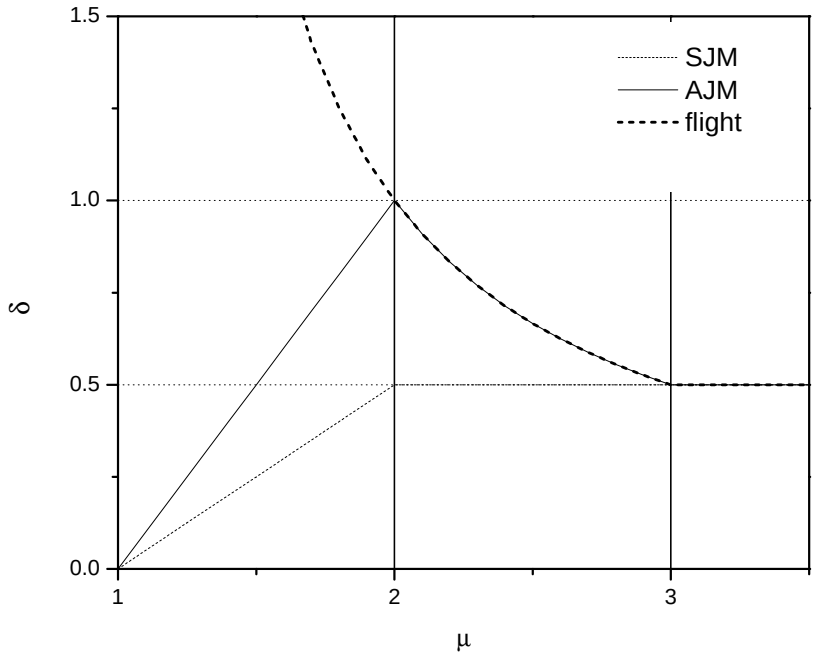

Figure 2.1: Illustration of the relationship between the scaling parameter, δ, and the distribution exponent, μ.

transition to the thermodynamical regime provides a large enough region for fitting to a finite value of δ .

At this point it would be useful to examine the regions of the μ parameter as depicted in Figure [2.1] and how they reflect the memory of a diffusion process. In the region, $\mu \geq 3.0$, the statistics of ordinary Brownian motion applies where the diffusing process is characterized by random Gaussian statistics. In this region, the first and second moments are finite. In the region, $2.0 < \mu < 3.0$, the three random walk rules dictate the type of statistics. The δ scaling parameter assumes a range of values; $0 < \delta \leq 1.0$, where $\delta = 0.5$ corresponds to a random Gaussian signature. How can we define memory in this region? If we look at the predictions of the Manneville map [58] which is the dynamical generator of fluctuations leading to Lévy processes, we can examine how the effects of moving from a laminar to a chaotic region can dictate the value of μ. The Manneville map is defined as:

$$x_{n+1} = x_n + x_n^z \, (mod 1) \tag{2.31}$$

where $z \geq 1.0$ and

$$d^z + d = 1. \tag{2.32}$$

The Manneville map dictates the course of what is known as *sporadic randomness* [59] in which a particle moving in the laminar region can escape into the chaotic region. From the chaotic region it is injected back into the laminar region again. This model has been used in a number of studies to describe the intermittent characteristics of turbulence [60, 61].

The distribution of times, $\psi(t)$, spent in the laminar region [62] is defined as:

$$\psi(t) = d^{z-1} \left[1 + d^{z-1} (z-1) \, t\right]^{-\mu}, \tag{2.33}$$

for $\mu = z/(z-1)$. This means that the region, $1.5 \leq z \leq 2.0$, corresponds to the case where the first moment is finite and the second moment diverges. In addition, if the system, represented as a Lévy walker, is considered to make a jump in this region, then it becomes equivalent to a Lévy flight [63]. The equivalent μ range,

$2 \leq \mu \leq 3$, also contains the finite first moment and the diverging second moment. The divergence of the second moment corresponds to the broadening of a distribution by increasing values of it's standard deviation as the diffusion process evolves. In this region, the Lévy distribution mimics this behavior by the heaviness of it's tails. Hence, the amount of information or memory can now be located further from the center of the distribution which is reflected in the range of μ values in this region.

At the value, $\mu = 2.0$, we move from the stationary to the non-stationary regime. In this region, both the first and the second moments do not exist in which case, the predictions of the GCLT states that the distribution of a sum of stable distributions (in our case, the family of Lévy stable distributions) is a stable distribution (or a Lévy stable distribution). The Lévy distribution is defined by it's characteristic function (Fourier transform):

$$P(x) = \exp\left[imx - \gamma \mid x \mid^{\alpha} (1 + i\beta x/ \mid x \mid) \tan(\alpha\pi/2)\right], \tag{2.34}$$

where $0 \leq \alpha \leq 2$. For $\alpha = 2.0$, the Gaussian distribution results, while for $\alpha < 2.0$, $P(x)$ assumes an asymptotic inverse power form,

$$P(x) \sim Ax^{-(1+\alpha)}, \tag{2.35}$$

where the second moment is infinite. The lack of both moments indicate an even stronger presence of memory where the distribution is now noninvariant.

It is important to remind ourselves that the random walk rules imply that there is no correlation between the ξ_i data points. Furthermore, these rules assume that the asymptotic limit of the distribution of the intensities and waiting times, is an inverse power law with no truncation. The truncation of the inverse power law at large distances is another serious issue worth commenting on. Laherrere and Sornette [64], suggest that the stretched exponential representation may have a theoretical generalization more suitable than the power law distribution by itself. In the intermediate time region the stretched exponential approximates the inverse power law form, however, both functions become distinguishable from one another in the long time region

which is affected by poor statistics. However, the studies of Mantegna and Stanley [65] and Floriani et al [66] show that a truncation of the power law form of ψ at large times yields an ultra-slow convergence to normal diffusion with effects that are beyond the range of observation of an entropic analysis such as DE. We shall observe that the application of the AJM and SJM rules yields a slow transition to the scaling regime which, due to the statistical limitations of both data sets turns out to be a relatively short region in the transition between the thermodynamical and saturation regimes. The adoption of the Flight rule yields a faster convergence to the thermodynamic regime and consequently makes it possible to achieve a direct evaluation of δ. In all cases, however, the physical consequences of the inverse power law truncation are still beyond the range of observation.

CHAPTER 3

Analysis and Results

In this chapter we will look at the results of the Rescaled range method described in Chapter 2 as applied to the HXY data. It is important to keep in mind that the application of the two methods spans three statistical paradigms; ordinary brownian motion, fractional brownian motion, and the as yet, unformulated statistical picture that is likely to encompass both. The data set consists of 7212 hard x-ray components of 7212 flaring events ocurring between April 1991 and May 2000. The data was extracted from the solar flare catalog of the Burst and Transient Source experiment (BATSE) onboard the Compton Gammay Ray observatory (CGRO) satellite. An event consists of peak flux values (photons s^{-1} per detector area of 2000 cm^2) and their corresponding times. The analysis was conducted on two parameters extracted from the data set, the waiting times and intensities (peak flux values). The waiting times are defined as the laminar lengths between two nearest neighbor flare times, $\tau_i = t_{i+1} - t_i$, corresponding to successive flaring events. The distribution of these waiting times is denoted as $\psi(\tau)$. The intensities data points denoted as x_i and corresponding to the maximum amount of HXY photon flux emitted during the event are represented in distribution form as $M(x)$. We will also utilize the integral of the distribution data, denoted as $\Psi(\tau)$ for the waiting times and $\Psi_M(x)$ for the intensities.

3.1 Rescaled Range Analysis and Results

A rescaled range analysis was performed on the waiting times and intensities through non-overlapping windows of the data. Based on Eq. [2.9], a linear regression of the $\log(R/S)$ values of the lags or partitions of the waiting times and the intensities was applied to determine the H value from the slopes. These windowed values are based on whole number divisors of the data length. Figures [3.1] and [3.2] illustrate the results. It should be noted that the regression was not performed through the entire

range of H values. As seen in Figure [3.1], waiting times data fluctuations due to solar cycle periodicity and possible saturation effects start to occur around a lag of T = 600. This would have resulted in an H value higher than 0.79. For the intensities data, the same reasoning applied; the fit performed through the first 18 data points yielded H=0.59 and would have resulted in a higher value if the saturation and periodicity range had been included in the fit. From our brief discussion of Fractional Brownian motion in Chapter 2, the assumption that the correlation function exists, dictates that correlations for the waiting times are of a much longer range nature than for the peak intensities, e.g.; the actual number of peak photons emitted in the HXR range. Recent observations by Moon et al. [67] have verified the persistancy characteristics for the waiting times of HXR flares occuring within the same active region. Since it was also found that *homologous flares* (flares ocurring within the same active region) outnumber interdependent *sympathetic flares* (flares occuring in different active regions), the correlations the Hurst method is seeing is from within a data set comprised of high memory flare groupings interspersed with lower memory interdependent flares. It is also been noted that the relationship between peak flaring intensities and waiting times are not significantly correlated (Moon and Crosby) [67, 68] and that autocorrelation results for HXR emissions within a single flaring event indicate little to no correlations (Aschwanden) [69]. However, in looking at the change in peak intensities over the time of a solar cycle as seen in Figure [1.3], it is obvious that flaring intensities are highly correlated with the solar cycle. Hence, flux emission changes due to the global memory of the solar cycle are not being seen by the Hurst method. In summary, it is clear that the data does depart from the statistics of ordinary Brownian motion. Both values fall into the persistent region where the probability of positive fluctuations recurring is high. The Hurst value, H= 0.79, for the waiting times falls much closer to 1.0, the ballistic Gaussian limit, than does the intensities values of H=0.59 which is much closer to the random Gaussian regime. In essence, this means from the FBM standpoint, that the waiting times are more highly correlated and have a much stronger memory signature as opposed to the intensities.

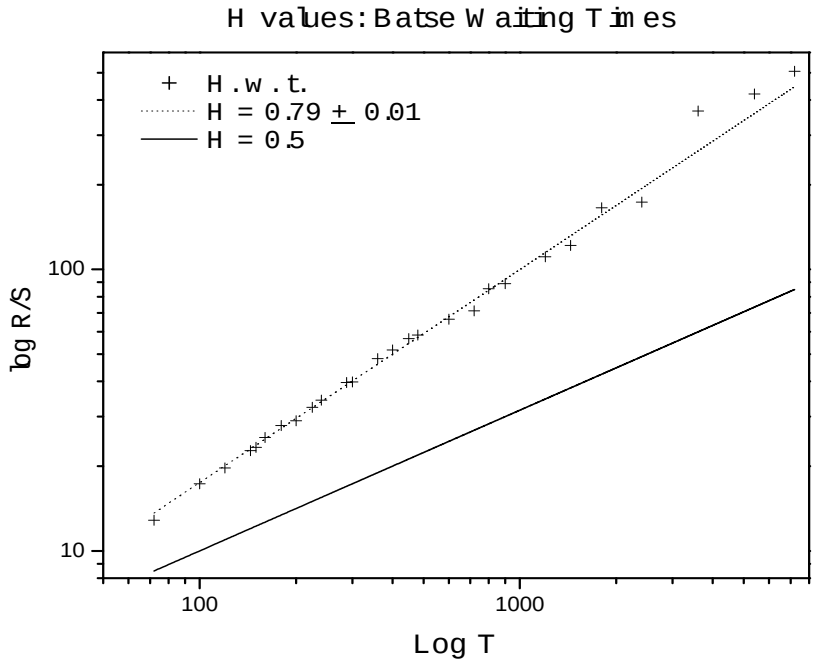

Figure 3.1: Rescaled Range Analysis results. Note the persistancy characteristics of the waiting times. A linear least squares fit was performed from the 2nd through the 19th data points. This was done in order to eliminate the influence of the fluctuations in the data in the larger windows. These particular fluctuations appear to be due to the effects of solar rotation and in part, possible saturation.

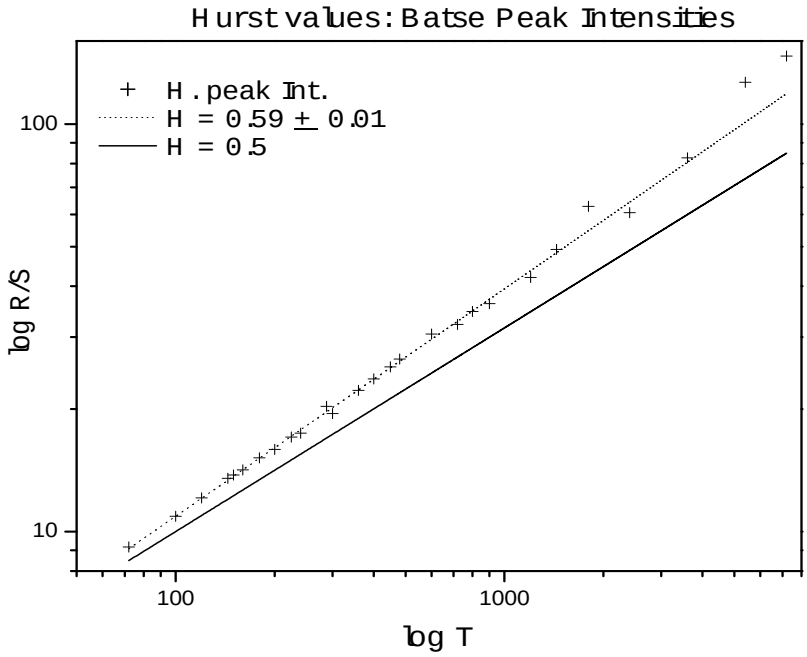

Figure 3.2: Rescaled Range Analysis results. Note the persistancy characteristics of the intensities. A linear least squares fit was performed through the first 18 data points. This was done in order to eliminate the influence of the fluctuations in the data in the larger windows. These particular fluctuations appear to be due to the effects of solar rotation and in part, possible saturation.

3.2 Diffusion Entropy Results

In consideration of the case that the statistics of HXY flares may not fit the FBM picture, the DE algorithm was applied to the data sets. The most important assumption in this case is that the distribution is a normalized inverse power law of the form:

$$\phi(y) = \frac{A}{(B+y)^{\mu}}, \tag{3.1}$$

where

$$A = (\mu - 1)\, B^{(\mu-1)} \tag{3.2}$$

and y represents the intensities and waiting times parameters, x_i and τ_i respectively. Note that the most utilized form of the inverse power law,

$$\phi(y) = \frac{A}{y^{\mu}} \tag{3.3}$$

is recovered for the asymptotic condition. This generalized inverse power law form has been used in a number of studies [56, 70] in order to make use of its integrable properties within the CTRW scheme. Our motivation for examing the flaring distributions in terms of this form is two-fold; it is a more realistic representation of the dual parameter appearance of these distributions (for example, Figure [3.3]) and to stay within the CTRW scheme.

First, we will focus on obtaining a value for the distribution exponent, μ, by direct evaluation of the distribution and it's integral. The use of artificial sequences was limited to our study of the waiting times in order to demonstrate a method for accurately determining the distribution exponent, μ. This was important since the initial value of μ determined from the waiting time distribution lay close to the border of transition ($\mu = 2.0$) from the stationary to the nonstationary regime.

The application of the DE method proceeds as follows; first we convert the waiting times and intensities data point values to a dichotomous representation of sequences

of zeroes and ones. This conversion takes place according to the prescription,

$$\beta_j = Int\left[\frac{\xi_i}{B}\right] + 1 \tag{3.4}$$

where $Int\,[z]$, denotes the integer value of the argument, z. The choice of B values can be made up to the largest data point value within the data set, however, preliminary trials with different values of B proved there problems associated with excessively small or large values of B. Small values of B certainly increase the computational time with a possible effect of spurious statistics whereas values that are too large would produce earlier statistical saturation and unphysical subdiffusion effects which again would present an innacurate picture of the underlying dynamics. To avoid this problem, a number of preliminary trials were conducted in order to narrow down the most stable values of $\delta\,(t)$, which would point to the proper thermodynamic fit region on the entropy curves. We then proceed to create trajectories in which the number of zeroes corresponding to the value of β_j are placed in a group followed by a number whose value corresponds to either a $+1$ for the AJM rule (always in the same direction) or a random ± 1 according to the SJM rule (forward or backward direction). The flight rule consists of an incremental sum of the entire sequence of groups, by which the DE algorithm is then applied.

3.2.1 DE Analysis of The Waiting Times.

We examine the waiting time distribution, $\psi\,(\tau)$, and it's integral, $\Psi\,(\tau)$, defined as;

$$\Psi\,(\tau) = \int_{\tau}^{\infty} \psi\,(\tau)\,d\tau. \tag{3.5}$$

The evaluation of the probability of finding time distances no larger than a given τ_i, allows us to obtain a better fit. This is due to the fact that integration has the effect of smoothing the data points within the distribution curve. The fitting procedure utilizes the nonlinear least squares Levenberg-Marquardt algorithm [71]. The fitting was implemented with the generalized inverse power law equation [3.1] for

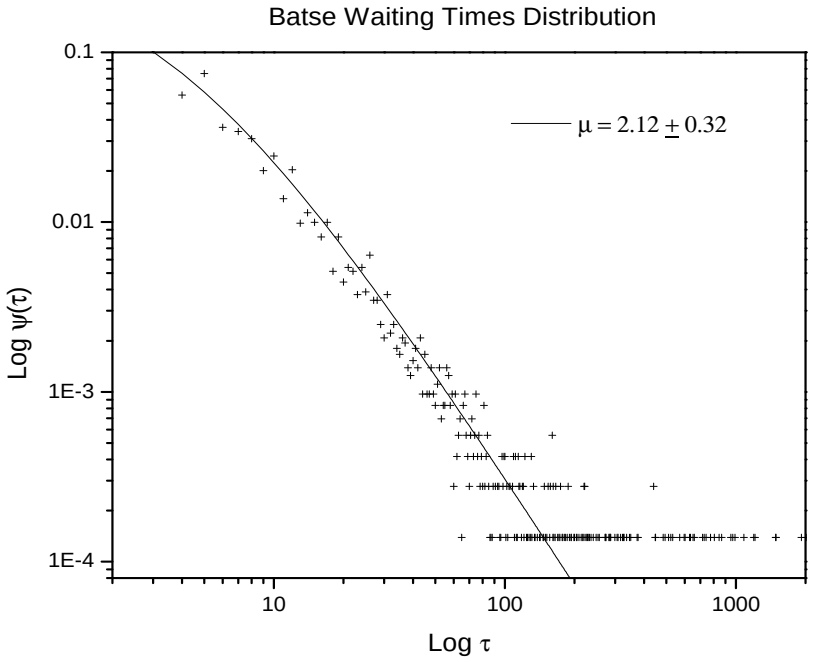

Figure 3.3: Waiting times distribution illustration. Note the truncation of the distribution due to statistical scatter.

$\phi(y) \equiv \psi(\tau)$ and it's integral form;

$$\Psi(\tau) = \frac{B}{(T+\tau)^{\mu}}. \tag{3.6}$$

Figure [3.3] illustrates the result of the distribution fit. There was a significant error of 15 per cent in fitting to the overall distribution such that $\mu = 2.12 \pm 0.32$. This implies that μ lies in the interval, [1.8, 2.44] containing the transition value of $\mu = 2.0$, a result that would prevent us from assessing whether the waiting times were stationary or not. The fit to the integral of the distribution, $\Psi(\tau)$, is illustrated in Figure [3.4]. The smoothed curve is more readily fitted, the result being, $\mu = 2.144 \pm 0.05$. This establishes the distribution as a stationary one within the range, [2.094,2.194]. It can now be noted that the uncertainty interval associated with $\Psi(\tau)$ is contained within

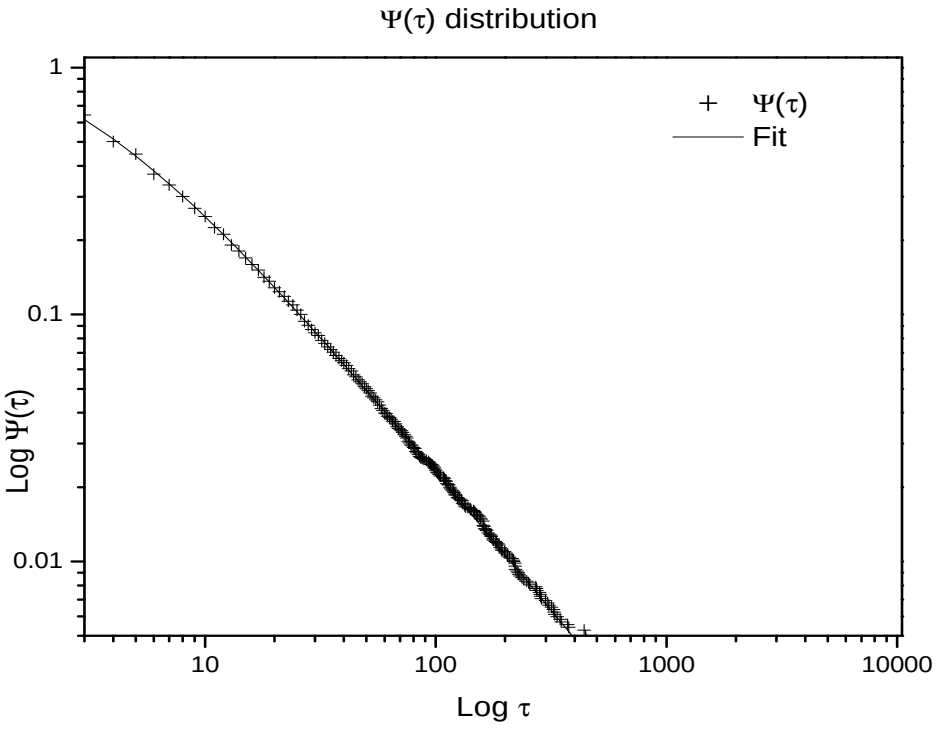

Figure 3.4: Integral of the waiting time distribution. The fit line corresponds to $\mu = 2.144 \pm 0.05$.

$\psi(\tau)$	$\Psi(\tau)$
$1.8 < \mu < 2.144$	$2.094 < \mu < 2.194$
$T = 8787$	$T = 8422 \pm 500$
$A = 3.1 \times 10^4$	$A = 30657 \pm 16590$

Table 3.1: Summary of parameters used in the artificial sequences verification of the final distribution exponent, $\mu = 2.138$

the uncertainty interval for $\psi(\tau)$. Table [3.1] provides a summary of our results for the ranges of μ and the fitting parameters, T and B. We now proceed to narrow down the value of μ through the use of artificial sequences and the DE method. First, we assume the distribution, $\psi(\tau)$, of artificial data to have the form:

$$\psi(\tau) = \frac{A}{(T+\tau)^{\mu}}, \tag{3.7}$$

where A is determined by the normalization condition:

$$\frac{1}{A} = \int_{45}^{\infty} \frac{1}{(T+\tau)\,\mu} d\tau. \tag{3.8}$$

and the minimum waiting time value of 45 seconds allows us to establish the integration from 45 to ∞. We then proceed by first constructing two different sets of artificial sequences defined by the normalized distribution, where (1) μ is held fixed while T is allowed to change in the interval, [7922, 8922], and (2) T is held fixed at 8422, while μ is allowed to vary in the interval, [2.094, 2.194]. The DE results of these sequences based on the AJM and SJM rules are then compared to the DE results of the shuffled waiting times based on the AJM and SJM rules. The comparison to the shuffled waiting times is more physically appropriate since the artificial sequences are based on uncorrelated waiting times, themselves. The results are shown in Figures [3.5] and [3.6]. Both figures indicate that the error strips corresponding to the changing T values are significantly smaller than the μ error strips by a factor of 5. This allows us to assess that T is the stable parameter defining our artificial sequences, since any fluctuations within the fitting interval results in very little change on the

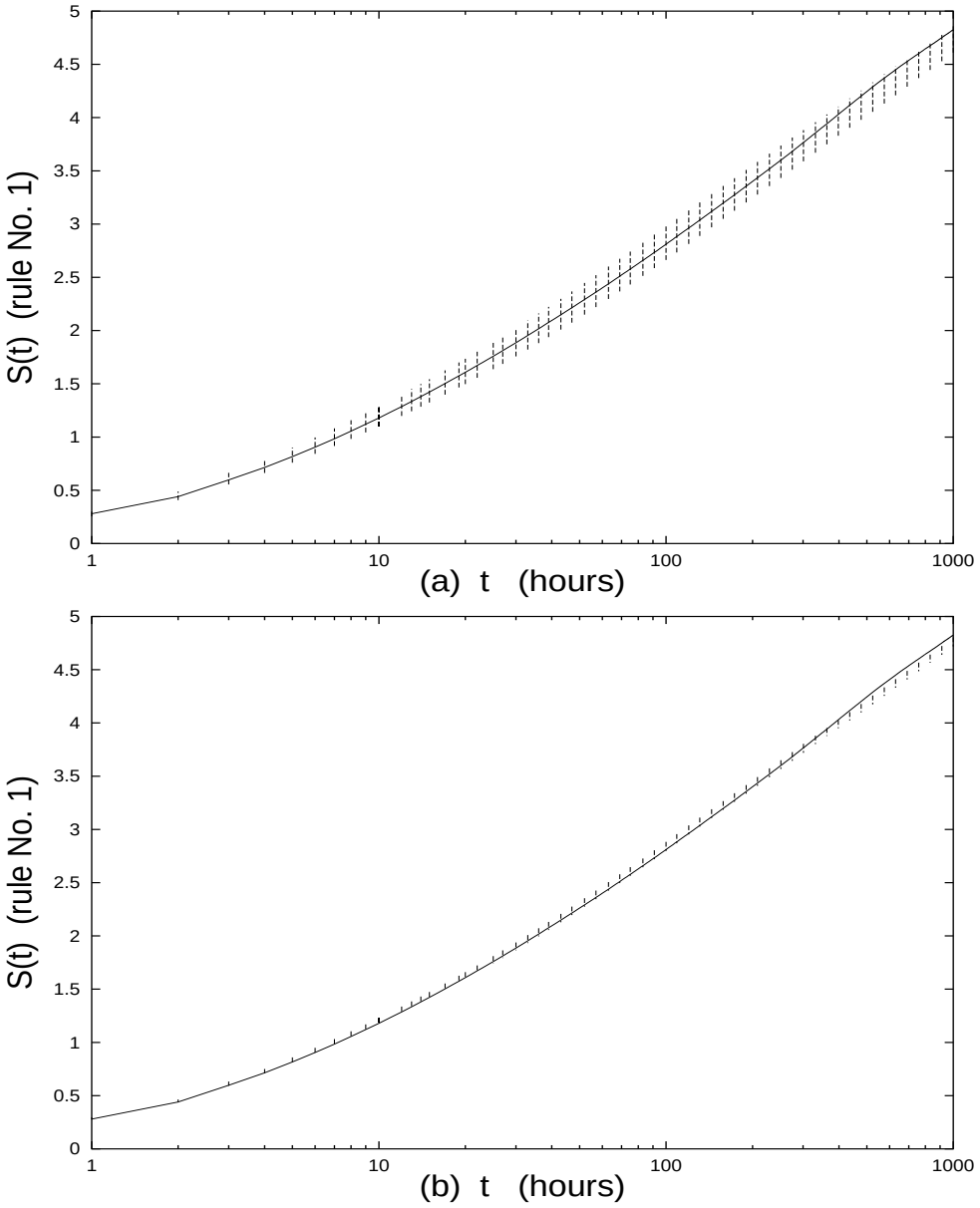

Figure 3.5: Ilustration of results of DE based on the AJM rule (on ordinate axis denoted as rule no. 1) for shuffled waiting times data (solid line) and artificial sequences. (a) Vertical bars depict changes of artificial sequences DE curves from waiting times DE for T = 8422 and μ changing in the interval [2.094,2.194]. (b) Vertical bars depict changes of DE of artificial sequences with $\mu = 2.144$ and T changing in the interval [7922,8922] in comparison with DE of shuffled waiting times

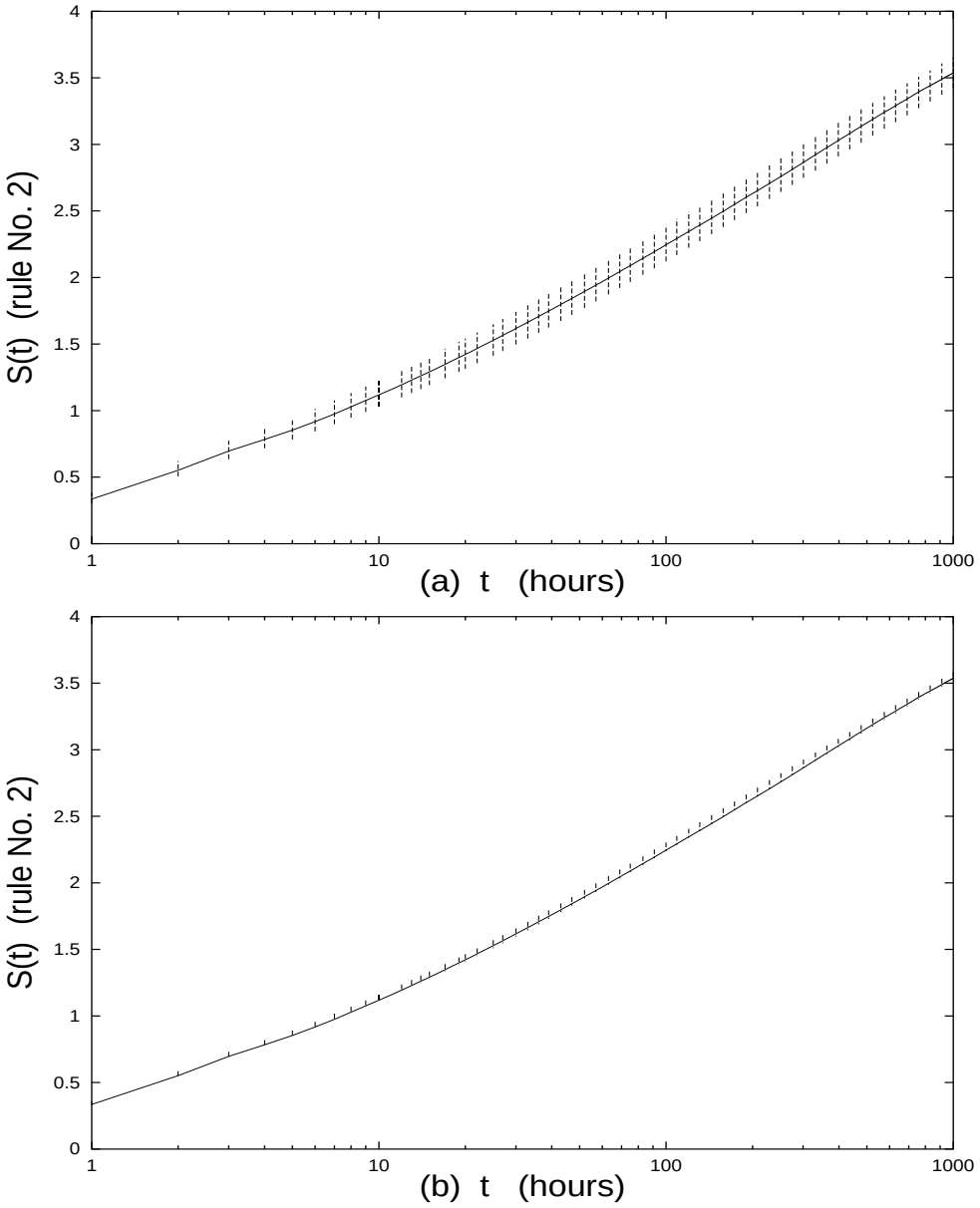

Figure 3.6: Illustration of results of DE based on the SJM rule (denoted as rule no. 2 on ordinate axis) for shuffled waiting times (solid line) and artificial sqeunces. (a) vertical bars indicate changes of artificial sequences DE curves where T = 8422 and μ changes in the interval [2.094,2.194]. (b) vertical bars depict changes of DE curves of artificial data where $\mu = 2.144$ and T changes in the interval [7922,8922].

outcome of the comparison between the DE results of the artificial sequences where $\mu = 2.144 \pm 0.05$ and the DE of the shuffled waiting times. Hence we can use artificial sequences of waiting times where T=8422 and μ assumes values in the interval, [2.094, 2.194] in order to find the proper μ value in this interval that allows the best comparison fit with the shuffled real waiting times. Figure [3.7] illustrates the final result where $\mu = 2.138 \pm 0.01$. The extremely close comparison between the DE results of the shuffled real waiting times and the DE results of the artificial waiting times is quite striking and therefore we should be confident that for $\mu = 2.138$, the following values for δ in the range $2.0 < \mu < 3.0$ according to the three rules are outlined in Table [3.2]. The results of the DE analysis based on the AJM rule for the shuffled

Rule	δ Parameter
AJM	$\delta = 0.879$
SJM	$\delta = 0.5$
Flight	$\delta = 0.879$

Table 3.2: Scaling parameters resulting from $\mu = 2.138$ as determined from the analysis of artificial sequences.

(randomized) and non-shuffled waiting time data is shown in Figure [3.8]. We can see that the slope of the shuffled waiting times lies closer to the $\delta = 0.874$ fit line as opposed to the nonshuffled waiting times which approach in only in the saturation region. The presence of these two distinct curves indicates that there exists two distinct forms of memory among the waiting times of the original data sequence. Figure [3.9] illustrates the result of applying the SJM rule. Again, there are two distinct curves. The entropy increase for both the shuffled and nonshuffled waiting times is much faster than the entropy based on the AJM rule. This is due to the fact that any kind of periodicity in the original data has less of an effect on a walker moving in a random backwards or forwards direction than it does for a walker moving in a biased direction. In addition, we may note that saturation effects show up at a later time for the same reason. Figure [3.10] illustrates the results of the DE analysis based on the flight rule. Despite the fact that the scaling relationship of Eq. [2.20] is based on a generalized inverse power law, if the difference between a distribution characterized

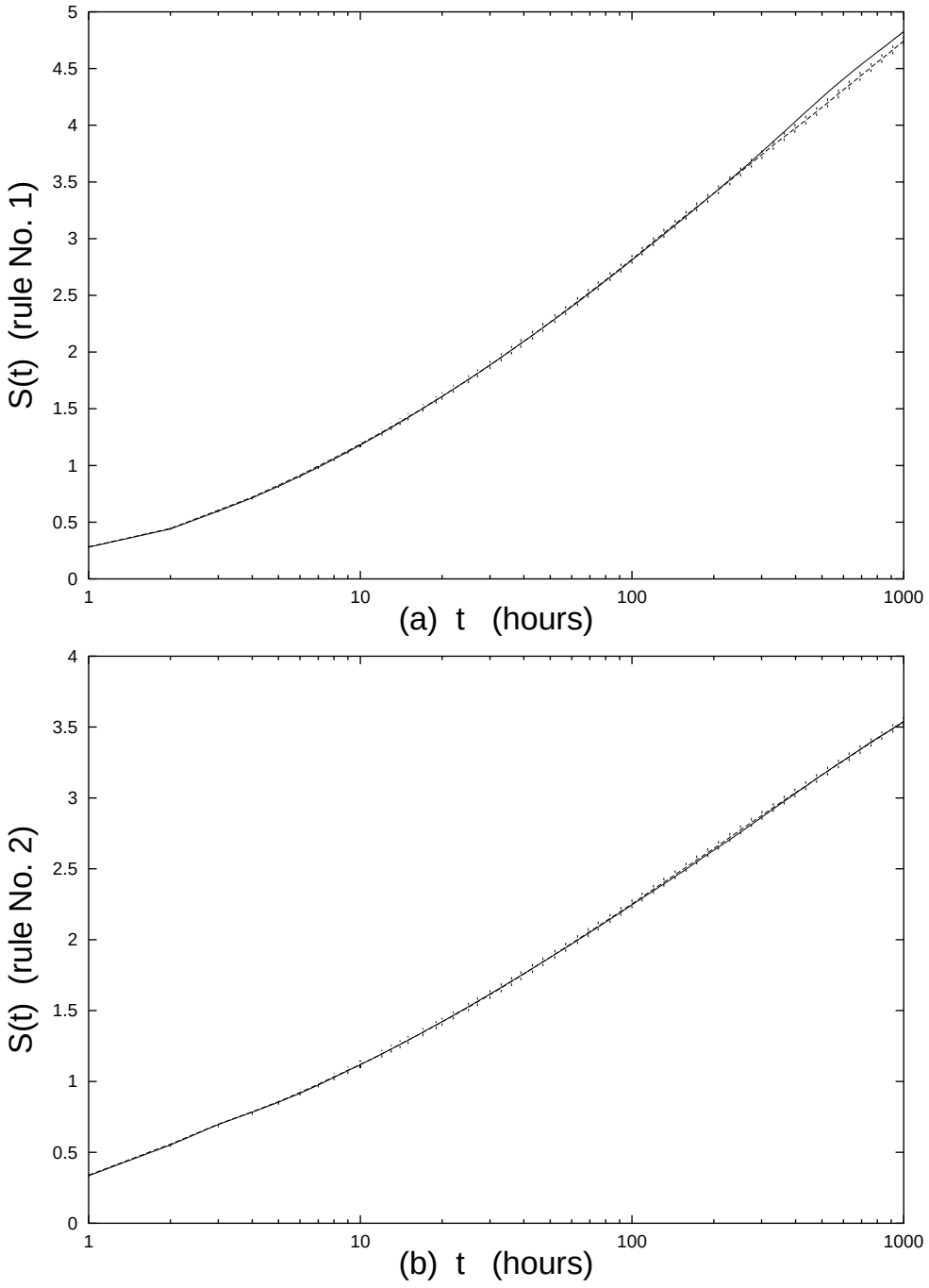

Figure 3.7: DE of artificial sequences based on the AJM rule (rule No. 1) and the SJM rule (rule No. 2). DE of shuffled waiting times denoted by solid lines. Dashed lines which denote artificial sequence data with $\mu = 2.138$ and T= 8422 almost coincide with the solid lines.

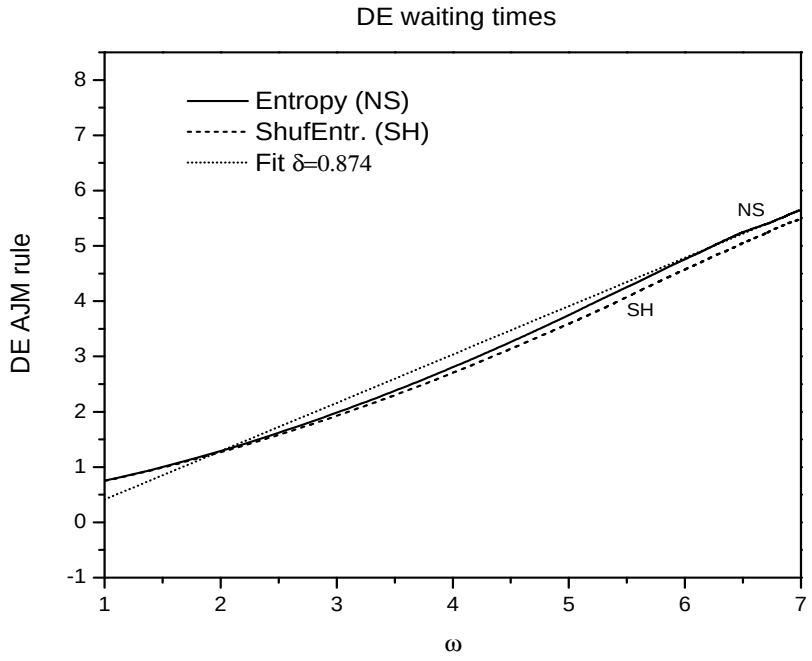

Figure 3.8: Rule AJM entropic analysis of Waiting times. The dashed line represents the scaling parameter value, $\delta = 0.874$ and $\mu = 2.144$ as determined by the analysis of the artificial data. The unshuffled entropy of the waiting times is denoted by the solid line; the shuffled waiting times DE is denoted by the short dashed line. Note how the entropy of the unshuffled data is superdiffusive with respect to the entropy of the shuffled data (dashed line).

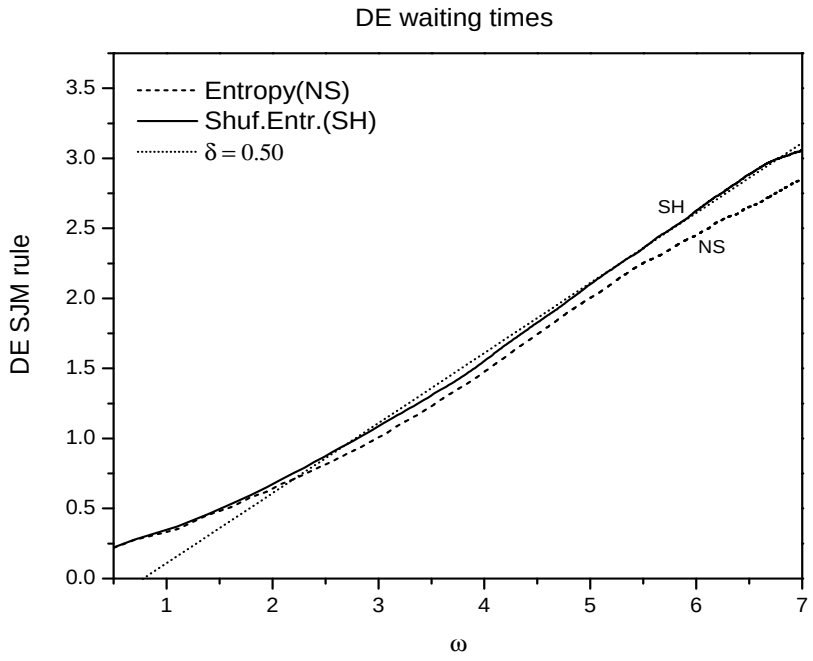

Figure 3.9: Rule SJM entropic analysis of waiting times. Note; in contrast with the AJM rule results, the entropy of the shuffled data is now superdiffusive with respect to the entropy of the unshuffled data (dashed line).

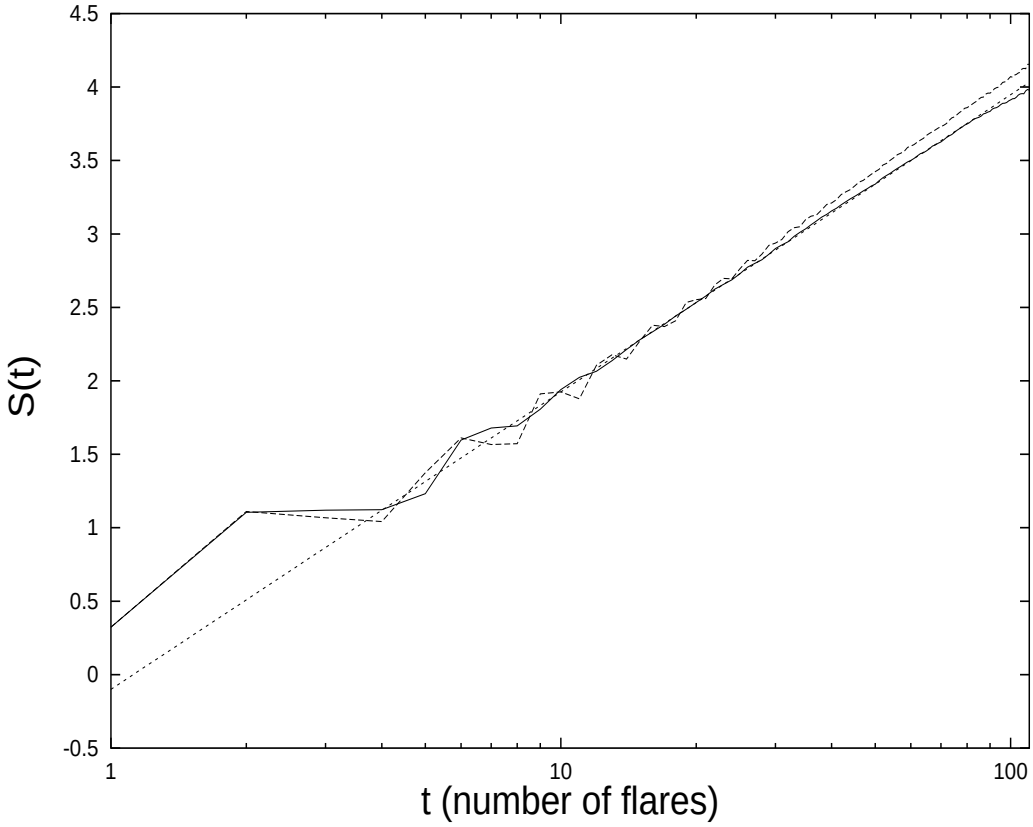

Figure 3.10: Illustration of results of flight rule on waiting times for $\delta = 0.879$ (fit - dotted line) and $\mu = 2.138$. The number of flares corresponds to the actual number of data points that are included in the moving window of the DE analysis. The dashed line refers to the unshuffled waiting times and the solid line refers to the shuffled waiting times.

by the inverse power law form is nearly equivalent to a Lévy distribution, then one could expect a very fast transition to the thermodynamical regime of the flight rule which is a Lévy signature. The figure illustrates this very fact, in fact the transition is so fast as to allow a large region on which we base our fit of $\delta = 0.879$. It should be noted that Lepreti et al [37] in their examination of the HXY GOES flares find a very accurate fit of the distribution to a symmetric long time Lévy distribution. Hence, this result is somewhat expected.

In conclusion, it is useful to comment on the facts presented herein, especially

with regards to the data we used. Most notably, the presence of the two memory signatures in our waiting times DE results appears to be due to a different set of dynamical processes. The unshuffled data corresponds to the waiting times of flares whose dynamics are strongly dictated by the time evolution of the solar dynamo in conjunction with contributions from global and local plasma conditions. Shuffling or randomizing the waiting times data destroys the ordering of these waiting times such that we are left with a frequency distribution representing the occurances of the flares and not how they are ordered.

3.2.2 DE Analysis of the Intensities

Historical data on the distributions of HXR peak intensities indicated the high probability that our intensities data would not admit an invariant distribution. This can be seen in Table [3.3] where all of the distribution exponents fall into a range [1.6,2.0] with only one actually equal to the transition range value, $\mu = 2.0$. Having clearly demonstrated the joint use of the DE method, artificial sequences and distribution analysis in securing an exact value for μ, it was important at this point to establish the case for a nonstationary distribution for the intensities. The study of the intensities proceeded with the plotting of the $M(x)$ and the $\Psi_M(x)$ distributions as shown in Figures [3.11] and [3.12]. We utilized the same nonlinear Levenberg-Marquadt algorithm used in the fitting of the waiting times distributions. The value of $\mu = 1.9$ for $M(x)$ and $\mu = 1.869$ for $\Psi_M(\tau)$ pointed to the data distribution as being an invariant one, requiring the use of the DE analysis for verification. In this case, the statistics lie in a nonstationary region where the first and second moments for a distribution do not exist;

$$< x >=< x^2 >= \infty. \tag{3.9}$$

The moments for the underlying diffusion process however, are different, for example, for the SJM rule, $< x^2 > < \infty$. For a sequence of fluctuations whose distribution exponent lies in the range, $1 < \mu < 2$, the theory relating the scaling parameter δ, to

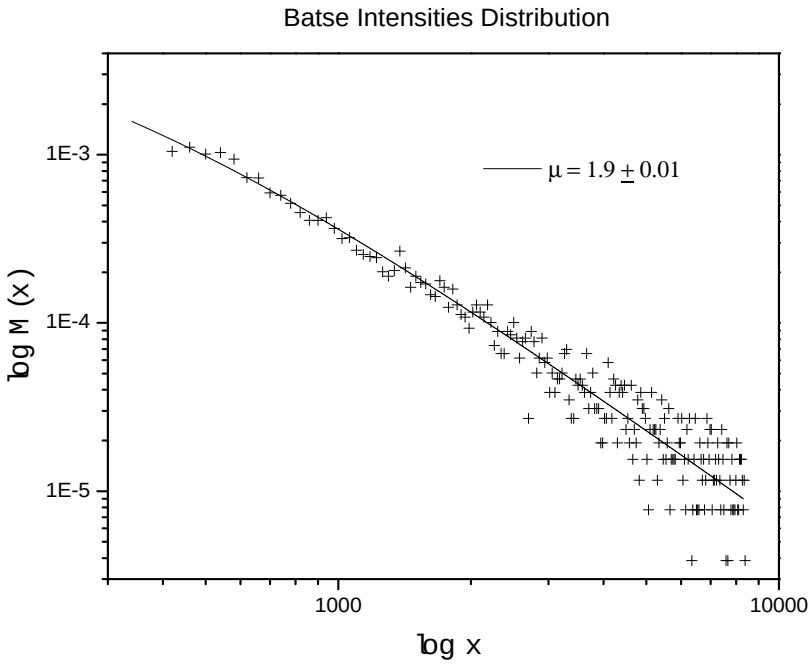

Figure 3.11: Illustration of distribution, $M(x)$. Note that the value of $\mu = 1.9$ points to the distribution as being a noninvariant one.

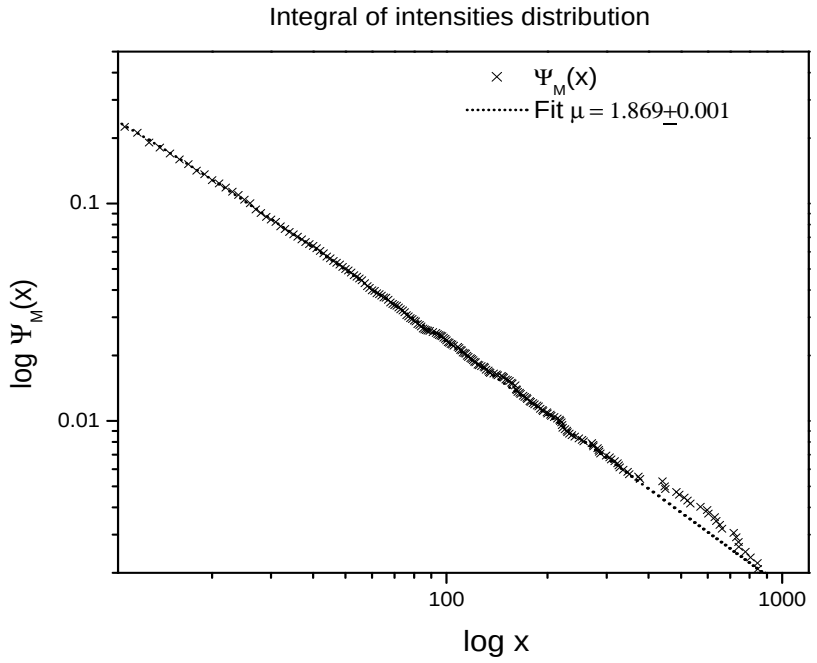

Figure 3.12: Integral of the waiting time distribution. The close fit to the integrated data distribution is readily seen and yields a value for μ slightly lower than $\mu = 1.9$ for the $M(x)$ distribution

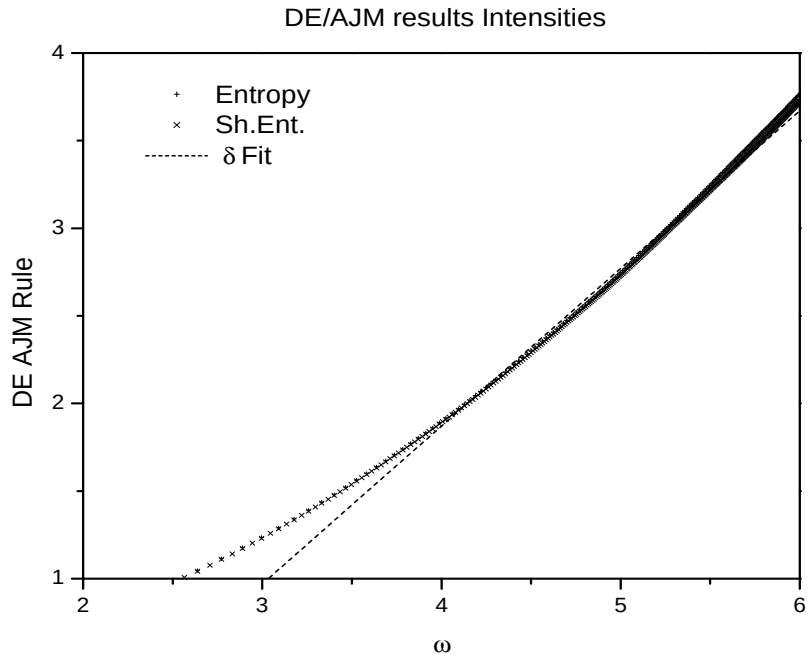

Figure 3.13: Illustration of results of AJM rule on intensities where $\omega = \ln x$ and $\delta = 0.868 \pm 0.004$

μ is based on the assumption of independent fluctuations. Hence, if there are no correlations between data points, then there are no correlations to destroy by shuffling. This means that the scaling parameter, δ, for both the shuffled and unshuffled data should be the same. This can be seen in figures [3.13], [3.14], and [3.15] for the AJM, SJM, and the Flight rules respectively.

We look first at Figure [3.13] which illustrates the result of DE algorithm based on the AJM assumption as applied to the intensities time series. The fit to the most linear region of δ corresponds to a value of $\delta = 0.868$, indicating that $\mu = 1.868$ or 2.13. As with the waiting times, the distribution of the data lies very close to the border of transition $\mu = 2.0$ between the stationary and nonstationary regimes. The application of the analysis based on the SJM rule yields the result, $\delta = 0.44$ which corresponds to a value of $\mu < 2.0$, confirming the nonstationary character

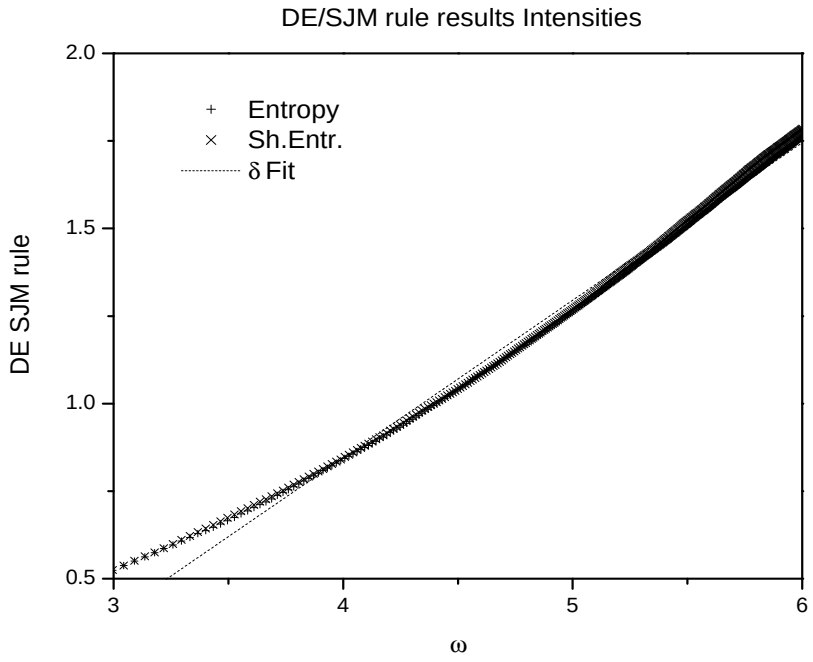

Figure 3.14: Illustration of results of SJM rule on intensities where $\omega = \ln x$ and $\delta = 0.44 \pm 0.004$

of the underlying distribution. This is illustrated in Figure [3.14]. Figure [3.15] illustrates the result of the DE analysis based on the Flight rule. The result itself is somewhat problematic in that despite the confirmation of the fast transition to scaling, an early saturation occurs, limiting our fit options to early transition. Two fits were performed before the saturation, one yielding a minimum scaling parameter value; $\delta = 1.177 \pm 0.014$ and the other yielding a maximum scaling parameter value; $\delta = 1.229 \pm 0.015$. A number of preliminary fits were done before obtaining these final values. The corresponding values for $\mu = 1.81$ and 1.85 are somewhat lower than our expected range of values, [1.879, 1.9], however, are not inconsistant. Table [3.4] illustrates just some of the many observed μ values for HXR intensities distributions, and the range of values for our results are not inconsistent with any of these values as well. Table [3.3] summarizes our results for the intensities. By inspection, we

DE Results on Intensities

δ	Corresponding μ	Analysis
———	1.9 ± 0.01	M(x) fit
———	1.869 ± 0.001	$\Psi_M(x)$ fit
0.868 ± 0.004	1.868 ± 0.004	DE/AJM fit
0.440 ± 0.004	1.880 ± 0.004	DE/SJM fit
$(1.229 \pm 0.015,$ $1.177 \pm 0.014)$	(1.814,1.850)	DE/Flight fits

Table 3.3: Illustration of results of joint use of DE and distribution analysis on the intensities.

notice that the lowest value of μ is 1.803, which corresponds to the uncertainty for the flight fit of $\delta = 1.229$. We do not assume the highest value of 1.9 for the $M(x)$ fit, however. In maintaining consistency we note that the integral of the distribution for the waiting times yielded a smaller error in fit than the fit for $\psi(\tau)$ which is the same case here for the intensities. We will assume the upper limit of $\Psi_M(x)$, $\mu = 1.870$. Hence, we note μ to lie within the range 1.803 to 1.870, below the transition value of $\mu = 2.0$. This is an important conclusion in that it says that the memory of the non-invariant distribution is higher than the memory of the waiting times, a result in conflict with the results of the Hurst analysis.

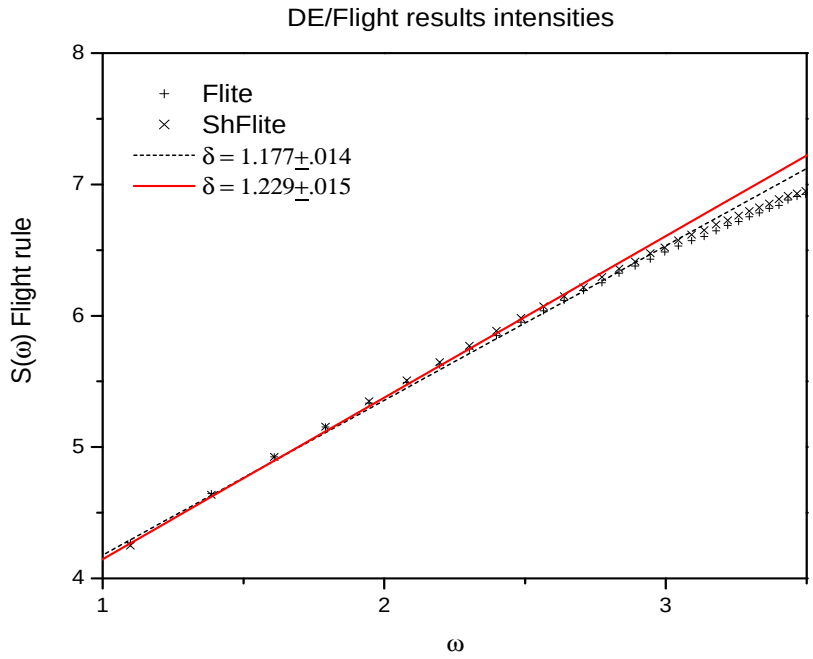

Figure 3.15: Illustration of results of flight rule on intensities. The two fit lines correspond to the extreme fit values obtained for the straight line region prior to saturation.

Observed HXR Peak Flux Distributions in Solar Flares

$-\mu$	Data Set	Reference
1.8	OSO VII, 20 keV; 1971-1972	Datlowe et al. 1974 [72]
1.9	OSO V; 1969-1971	Dennis 1985 [73]
1.70-1.86	HXRBS/SMM, > 30 keV;1980-1984	Bai 1993 [74]
1.74	HXRBS/SMM, > 30 keV;1980-1989	Kucera et al. 1997 [75]
1.75	ISEE3/ICE,>26 keV;1978-1986	Lee et al 1993[76]
1.86-2.0	ISEE 3/ICE,>26 keV;1978-1986	Bromund, McTiernan, Kane 1995[77]
1.6-1.74	BATSE/CGRO,>25 keV;1991-1992	Biesecker 1994 [29]
1.69-1.82	BATSE/CGRO,>50 keV;1991-1994	Aschwanden et al. 1995[78]

Table 3.4: Acronym Explanation: OSO: Orbiting Solar Observatory; HXRBS/SMM: Hard X-ray Burst Spectrometer/Solar Max.Mission; ISEE 3/ICE: International Cometary Explorer.

CHAPTER 4

Conclusions and Discussion

In this dissertation, the results of the application of two nonlinear methodologies on a nine year time series of hard x-ray flares are presented. The study commenced with the application of the Rescaled Range method on the waiting times and peak flux intensities data. The results yielded two values for the Hurst exponents, H= 0.59 for the intensities and H= 0.79 for the waiting times. According to the FBM paradigm, these results imply that the peak intensities are characterized by a weaker range of correlations as opposed to the waiting times whose higher H value implied considerably more memory. The results of the DE analysis on both the waiting times and the intensities indicated a different conclusion: that the intensities contained a higher memory signature than the waiting times by virtue of the fact that it's distribution exponent, which lay in the range [1.803,1.870], fell below the transition value of 2.0, and was significantly lower than the waiting time distribution exponent of $\mu = 2.138$. According to the random walk framework on which the DE method is based, the region, $1 < \mu < 2$ is characterized by at least several different realizations of how a dynamical process may experience diffusion. Despite the presence of a distribution which is nonstationary, the memory dictated by μ according to the prescription limitations of the Manneville map and the fact that this region is characterized by Lévy processes whose first and second moments are infinite, is stronger than it would be for a distribution whose exponent lay above the value of 2.0.

In addition we note that, the lack of the first two moments for a distribution in this region implies a propogation of memory through time and possibly spatial scales as well. As mentioned in Chapter 3, it is well known that the hard x-ray intensities values correlate well with the solar cycle, in that peak flux emissions, while not coupled strongly with the waiting times and hence their dynamical source, nevertheless must depend strongly on the density and temperature of the surrounding coronal medium. The bulk properties of this medium are modulated quite strongly by the

solar cycle. The theory of MHD processes has been used to describe the solar atmosphere along with its coupling to rising magnetic flux tubes and the underlying convective layer. The phenomena of intermittancy and turbulence which characterize many MHD processes, are known to be associated with statistical distributions described by the AJM, SJM and Flight rules. At this point we can conclude that the Hurst method was not able to see this more global memory largly due to its basis in the Gaussian assumption. In addition, a limitation may have been placed on the detection of scaling by the Hurst parameter due to the examination of the time series through non-overlapping windows. The DE method on the other hand, works through "moving windows" or better said, on all trajectories of the data under the assumption of a generalized inverse power law. This method of analysis better enables the accurate detection of the scaling from a global sense without the restrictions of a specified set of statistics. In addition, it is worthwhile to remind ourselves of the nature of the scaling parameters, themselves. The Hurst scaling parameter detects a scaling but not in the same way that the DE method does. The scaling parameter, δ, is directly related to the distribution exponent through the spreading of its probability density function through time. Hence, "true scaling" or a scaling that we could normally relate to the definition of a probability distribution is defined exactly by the entropic scaling parameter, δ. Thus, the type of scaling seen by the Hurst exponent, H, is definitively different and though not a second moment parameter, is related to second moment scaling; nonetheless.

Another significant finding is that the statistical signature of the underlying dynamics is best characterized by Lévy distributions. This fact has been noted earlier by Lepreti et al. [37] whose methodology involved the assumption that a distribution of hard x-ray waiting times was already Lévy prior to its fit with a symmetric Lévy power law. The detection of the presence of a stable distribution whose first two moments are infinite lends support to the intermittency interpretation of flaring occurances. It is not however, within the scope of this study to address the issue of SOC vs. intermittency interpretations. However, it is hoped that the solar physics community will benefit in general, from the better descriptive power of these distributions.

It was also found that the waiting times appear to be comprised of two different memory signatures, one relating to the solar cycle and the other whose origin is possibly due to localized or a combination of localized active region effects. In addition it is unclear as to how much as to how much of an effect the duty cycle of the detector plays in this. It remains the subject of future analysis efforts which will involve the DE examination of waiting times extracted from an interplanetary spacecraft.

BIBLIOGRAPHY

[1] M. B. Kallenrode, *Space Physics; An Introduction to Plasmas and Particles in the Heliosphere and Magnetospheres*, Springer-Verlag, Berlin., (1998).

[2] T. Lee, Ph.D. Dissertation, Stanford University, (1996).

[3] L. Golub, and J. Pasachoff, *The Solar Corona*, page 289, Cambridge University Press., (1997), page 282.

[4] E. N. Parker, Astrophys. J., **12**, 293, (1955).

[5] M.R.E. Proctor, and A.D.Gilbert, *Lectures on Solar and Planetary Dynamos*, eds. 1994, Cambridge Univ. Press., Cambridge UK.

[6] C. Schrijver, and C. Zwaan, *Solar and Stellar Magnetic Activity.* New York: Cambridge University Press. 2000, 204-205.

[7] L. Golub and J. Pasachoff, *Nearest Star: the Surprising Science of Our Sun*, Harvard Univ. Press, Cambridge MA. (2001).

[8] BATSE SOLAR Web page: *www.umbra.nascom.nasa.gov/batse/batse.years.html.*

[9] Hathaway, Wilson, and J. Reichmann, Geophys. Res., **104**, 22, 375, (1999).

[10] K. Schatten, P. Scherrer, L. Svalgaard, and J. Wilson, Geophys. Res. Letts., **5**, 411, (1978).

[11] Josleyn et al, EOS, Trans. Amer. Geophys. Union., **78**, 205, (1997).

[12] R. Thompson, Solar Phys., **148**, 383, (1993).

[13] M. J. Aschwanden, *Encyclopedia of Astronomy and Astrophysics*, Vol. 3, 2543, Inst. of Phys. Pub., New York,(2000).

[14] P. Martens, *Preflare Phase; Encyclopedia of Astr. and Astrophys.*, Inst. of Phys. Pub., New York (2000).

[15] L. Golub, and J. Pasachoff, *The Solar Corona*, page 282, Cambridge University Press., (1997).

[16] A. O.,Benz, Solar Phys., **104**, 99, (1986).

[17] A.L. Kiplinger, B. Dennis, A. Emslie, K. Frost, and L. Orwig, Astrophys. J. Lett.,**265**, L99, (1983).

[18] S. Masuda, T. Kosugi, T. Sakao, and J. Sato, Astrophysical and Space Science Library: Conference Proceedings; Observational Plasma Astrophysics: Five years of Yohkoh and beyond., **229**, 259, (1998).

[19] V. Petrosian, and T. Q. Donaght, Astrophys. J., **527**, 945, (1999).

[20] L. Harra-Murnion, Astr. and Astrophys., **337**, 911, (1998).

[21] L. Reichl, *A Modern Course in Statistical Physics, 2nd Ed.*, J Wiley and Sons, N.Y., (1998).

[22] J. Feder, *Fractals*, Plenum Press, N.Y., (1988).

[23] M. Schroeder, *Fractals, Chaos, and Power Laws, Minutes from an Infinite Paradise.*, W.H. Freeman Co., N.Y. (1991).

[24] R. Mantegna, and H. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge University Press, U.K., (2000).

[25] B. West, R. Zhang, A.W. Sanders, S. Miniyar, J.H. Zuckerman, and B.D. Levine, Physica A, **270**, 552, (1999).

[26] B. West, P. Hamilton, and D. West, NonLinear Dynamics, Psychology, and Life Sciences, **4**, No. 1, 87, (2000).

[27] P. Allegrini, M. Buiatti, P. Grigolini, and B. West, Phys. Rev. E, **57**, 4558, (1998).

[28] M. Wheatland, P. Sturrock, andJ. McTiernan, ApJ, **509**,448, (1998).

[29] D. Biesecker, Ph.D. Dissertation, University of New Hampshire, (1994).

[30] G. Boffetta, V. Carbone, P.Giuliani, P. Veltri, and A. Vulpiani, Phys. Rev. Lett., **83**, 4662, (1999).

[31] M. Wheatland, ApJ Lett., **536**, L109-L112, (2000).

[32] M.J. Aschwanden, B. Dennis, and A. Benz, ApJ, **497**, 972, (1998).

[33] B. Mandelbrot, *Fractal Geometry of Nature.*, W. H. Freeman Co., (1988).

[34] A. Ruzmaikin, J. Feynman, and P. Robinson, So. Phys., **149**, 395, (1994).

[35] R. Komm, So. Phys., **156**, 17, (1995).

[36] F. Lepreti, P. Fanello, F. Zaccaro, and V. Carbone, So. Phys., **197**, 149, (2000).

[37] F. Lepreti, V. Carbone, and P.Veltri, Astrophys. J. Lett., **555**, L133, (2001).

[38] H. Hurst, *Trans. Amer. Soc. Civ. Engrs.*, **116**, 770-808, (1951).

[39] B. Mandelbrot and J. Wallis, Water Resour. Res.,**4**, (1968).

[40] B. Mandelbrot and J. Van Ness, SIAM Rev.,**10**, 422, (1968).

[41] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. I, Wiley, N.Y., (1971).

[42] D. Leddon, Submitted to Phys. Rev. D, **cond-mat**, (2001).

[43] J. Klafter, M. Shlesinger, and G. Zumofen, Physics Today, 33, (Feb., 1996).

[44] S. Buldyrev, A. Goldberger, S. Havlin, C. Peng, M. Simons, and H. Stanley, Phys. Rev. E., **47**, 4514, (1993).

[45] R. Mannella, P. Grigolini, and B. West, Fractals, **2**, 81 (1994).

[46] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. II, Wiley, N.Y., (1966).

[47] C. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, John Wiley and Sons, INC, N.Y., (1995)

[48] P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris (1937).

[49] B. West and W. Deering, Physics Reports,**246**, 1, (1994).

[50] N. Scafetta, P. Hamilton, and P. Grigolini, Fractals, **9**, 193 (2001).

[51] N. Scafetta, V. Latora, and P. Grigolini, **cond-mat 0105041** (2001).

[52] P. Grigolini, D. Leddon and N. Scafetta, submitted Phys. Rev. E., **cond-mat** (2001).

[53] P. Hamilton, M. Ignaccolo, and P. Grigolini, in prep.

[54] G. W. Weiss, *Aspects and Applications of the Random Walk*, ed. H. Stanley and E. Guyon, North Holland Press, Amsterdam, (1994).

[55] A. I. Kinchin, *Mathematical Foundations of Statistical Mechanics*, Dover Publications, Inc., New York (1949)

[56] P. Grigolini, L. Palatella, and G. Raffaelli, in press, Fractals (2001).

[57] M. Shlesinger, J. Stat. Phys., **10**, 421 (1974).

[58] P. J. Manneville, J. Phys., **41**, 1235 (1980).

[59] M. Ignaccolo, P. Grigolini, and A. Rosa, Phys. Rev. E, **64**, 026210 (2001).

[60] M. Cosenza, and A. Parravano, Phys. Rev. E, **53**, 6032 (1996).

[61] F. Argoul, A. Arneodo, and P.Richetti, J. Phys., **49**, 767 (1988).

[62] M. Buiatti, P. Grigolini, and A. Montagnini, Phys. Rev. Lett., **82**, 3383 (1999).

[63] M. Bologna, P. Grigolini, and L. Palatella, Physica A, **268**, 214 (1999).

[64] J. Laherre and D. Sornette, Eur. Phys. J. B, **2**, 525 (1998).

[65] R. Mantegna and H. Stanley, Phys. Rev. Lett., **73**, 2946 (1994).

[66] E. Floriani, R. Mannella, and P. Grigolini, Phys. Rev. E, **52**, 5910 (1995).

[67] Y. Moon, G. Choe, H. Yun, and D. Park, AGU conf. presentation, (May, 2001).

[68] N. Crosby, N. Vilmer, N. Lund, and R. Sunyaev, Astr. Astrophys., **334**, 299 (1998).

[69] M. Aschwanden, M. Wills, H. Hudson, T. Kosugi, and R. Schwartz, ApJ, **468**, 398 (1996).

[70] M. Stefancich, L. Bonci, P. Grigolini, B. West, Phys. Rev. E, **57**, 6625 (1998).

[71] W. H. Press and Others, *Numerical Recipes in C,* Cambridge University Press, Cambridge (1992).

[72] D. Datlowe, M. Elcan, and H. Hudson, Solar Phys., **39**, 155 (1974).

[73] B. Dennis, Solar Phys., **100**, 465 (1985).

[74] T. Bai, ApJ, **404**, 805, (1993).

[75] T. Kucera, B. Dennis, R. Schwartz, and D. Shaw, ApJ, **475**, 338 (1997).

[76] T. Lee, V. Petrosian, and J. McTiernan, ApJ, **412**, 401, (1993).

[77] K. Bromund, J. McTiernan, and S. Kane, ApJ, **455**, 733 (1995).

[78] M. Aschwanden, A. Benz, B. Dennis, and R. Schwartz, **455**, 347 (1995).