A New Uncertainty Framework for Stochastic Signal Processing

Rishabh Singh, Student Member, IEEE and Jose C. Principe, Life Fellow, IEEE

Abstract—The field of signal processing and information theory have evolved with the goal of developing formulations to extract intrinsic information from limited amount of data. When one considers the modeling of unpredictably varying processes and complex dynamical signals with a large number of unknowns (for instance, in the field of finance, NLP, communications, etc.), there is a need for algorithms to have increased sensitivity with short spans of data while maintaining stochastic generalization ability. This naturally calls for an increased focus on localized stochastic representation. Most metrics developed so far for characterizing signals envision data from entropic and probabilistic points of view that lack sensitivity towards varying signal dynamics. We hypothesize that models that work with the intrinsic uncertainties associated with local data induced metric spaces would be significantly more sensitive towards signal characterization. To this end, we develop a new framework for stochastic signal processing that is based on decomposing the local metric space of the signal in a Gaussian Reproducing Kernel Hilbert Space (RKHS). A major advantage of our framework is that we are able to implement this decomposition on a sample-by-sample basis. The key aspects of our framework are the following: 1) We use a data defined metric related to Parzen density estimation for quantifying the local structure of data in the Gaussian RKHS. 2) We use a quantum description of this metric which consequently introduces uncertainty in the structure of the local kernel space. Since the RKHS has been well established and known for providing universal data fitting capabilities, we submit that local quantifiers of the kernel space data projection could significantly improve acquisition of signal information.

Index Terms—Signal processing, uncertainty, stochastic, decomposition, metric space, dynamical signals, local quantifiers, kernel space, quantum, RKHS.

I. INTRODUCTION

A. Information Theory: Historical Perspective

INFORMATION theory (IT) provides a critical basis for the quantification of usefulness of data based on limited data occurrence or knowledge. The concept of entropy has been at the forefront of this theory ever since its inception in the domain of statistics by Shannon [1]. At its origins, Gibbs entropy has been a well-known metric for describing micro-states in a thermodynamical system. Hence Shannons formulation of this concept as a metric associated with behavior of data opened the doors to a completely new realization of information theory based on the uncertainty formulations in physics. Subsequent to its development, Jaynes formulated a method in the context of statistical mechanics to estimate the probability distribution of data as the distribution which maximizes the entropy of the system of data (assumed as micro-states) based on certain macro-state constraints which are typically chosen as an ensemble property of data for instance, mean [2]. This work has been widely celebrated in the field of density estimation [3], [4]. It also created the foundations of physics based interpretation of information theory. Further developments in information theory and stochastic signal processing have been closely intertwined with entropy and several other derived metrics and generalized forms of entropy [5], [6]. A significant related work is that of Renyi in 1956 where he proposes a generalized formulation of the Shannons entropic measure [7]. Renyis entropy has the following form:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left( \sum_{i=1}^{n} p_i^\alpha \right)$$  (1)

Here, $X$ is a discrete random variable with probabilities $p_i$ denoting the probability of $X$ taking the value $i$.

One of the difficulties of applying IT to machine learning is that normally the data statistics are not known a-priori, therefore the IT descriptors of entropy, mutual information and divergence must be estimated directly from data [8]. Perhaps the simplest of the approaches is to employ statistical methodologies of density estimation [9] and plug-in the estimators in the IT formulas. We will review first some alternatives.

B. Maximum Entropy Method of Moments

An important extension to Jaynes’ work is the maximum entropy method of moments which is useful for estimation of density functions in systems with undersized samples [10]. The main idea here is to construct an approximate density function of a particular system of samples by imposing constraints unique to the system (typically chosen as the moment means) on the various moments associated with data. It consists of a system of two equations from which one can obtain the Lagrangian multipliers (or the weights associated with various data moments) and the probability densities using numerical techniques such as the Newton-Raphson method. Determination of Lagrangian weights quantify the uncertainty associated with the different moments of data. Several application extensions and analysis of this method have been introduced over the past few decades [11], [12], [13]. This method of maximizing entropy with respect to the various moments of data is simple and elegant with very few assumptions in its formulation. However, the solution of the different Lagrangian multipliers and density values are not guaranteed since the optimization is based on numerical techniques. The possibility of spurious results increase when large number of moments are used (large n value). Furthermore, moment correlations are
being ignored which leads to more difficulties in solving the optimization problem.

C. Point Estimate Methods

A highly relevant class of methods involved in the formulation and modeling of data statistics are the point estimation methods (PEM) which make it possible to evaluate statistical moments of data using approximate realizations of their density only at specific set of points in the input space. Introduced by Hahn, Shapiro and Cox [13], [15], PEM methods, at their onset, usually involved Taylor series expansions of random variables about their means which involves computation of higher order derivatives of functions. This can, however, become very cumbersome computationally. One of the popular earlier works in point estimate methods was that of Rosenblueth where he develops a method for estimating densities of correlated variables by assuming the joint distribution of the variable to be concentrated within a set of $2^n$ hyper-quadrants of the space associated with the variable [16], [17]. This method becomes impractical for large numbers of quadrants. Li introduced an improved formulation based on Rosenblueth’s method that only needs the first two statistical moments to compute PEM with almost the same accuracy as that of Rosenblueth’s method [18]. A thorough mathematical analysis and evaluation of various spatial point processes is provided in [19]. A more recent work by Decreusefond and Flint introduces a generalized formulation for point processes with the only constraint that the processes consist of Papanegelou intensities [20].

Although the statistical underpinning of IT is widely accepted, Crutchfield presents a rather unique non-probabilistic perspective of information theory where he regards the information space as a metric space (instead of a probabilistic one) created by information sources [21]. He stresses more the topological structure (geometry) created by information and develops an information theoretic distance measure in such a space that more closely follows the metric properties of distance as compared to other measures (such as mutual information, Kullback information gain, etc.). This serves as an inspiration for us to treat the kernel space projections of data samples as a metric space instead of a probabilistic one in our framework as we previously have done [3].

D. Advances in Uncertainty Quantification Methods

The field of uncertainty quantification aims to evaluate the effect of intrinsic uncertainties present in the input signal on the model’s performance. Earlier methods of uncertainty quantification were based on large number of Monte Carlo simulations of a model to evaluate its variability [22]. This is obviously a very expensive computational procedure though reliable in terms of accuracy. To mitigate the high computational expense, surrogate models (or meta-models) were introduced as new response surfaces to give fast evaluations of changes in input statistics [23], [24]. Two such surrogate models that have gained wide popularity are Gaussian process regression (kriging) and polynomial chaos expansion [25], [26], [27]. Both of these algorithms can be implemented in a non-intrusive manner which means that the model being evaluated can be treated as a black box. The idea of polynomial chaos (PC) algorithms was first introduced by Wiener in 1938 when he suggested that Hermite polynomials and homogenous chaos had an important influence in integration theory related to Brownian motion [28]. However, only recently have PC based surrogate models become widely popular [29].

A polynomial chaos model involves expression of a stochastic process with finite second order moments $Y(x, t, \theta)$ in terms of orthogonal basis functions (given by Wiener-Askey scheme [30]):

$$Y(x, t, \theta) = \sum_{i=0}^{\infty} \alpha_i(x, t) \phi_i(\zeta(\theta))$$

where $\phi_i$ represents the basis functions (or stochastic modes) and $\zeta$ represents second order random variables of the stochastic process parameterized by $\theta$. $\alpha_i$ represent the PC coefficients or weights of the stochastic modes. The choice of the family of basis functions depends on the distribution of the random variables ($\zeta$) and is given by the Wiener-Askey scheme. Any chosen basis function family, nevertheless, satisfies the following orthogonality conditions:

$$\langle \phi_r(\zeta), \phi_q(\zeta) \rangle = \int \phi_r(\zeta) \phi_q(\zeta) P(\zeta) d\zeta = \delta_{rq} ||\phi_r||^2$$

where $r$ and $q$ represent the basis function orders.

The main objective in PC expansion is to determine the coefficients ($\alpha$ values in (2)) corresponding to the different modes. This can be done using intrusive or non-intrusive methods. A popular example of an intrusive method is Galerkin projection which involves solving a system of deterministic equation where the coefficients are considered as the unknowns [31]. Galerkin projection first involves generalizing (2) in the following manner:

$$Y(x, t, \theta) = L \left( x, t; \sum_{r=0}^{K} \alpha_r(t) \phi_r(\zeta(\theta)) \right)$$

Projecting (4) into the different polynomial basis functions $\phi_j$, we obtain:

$$\langle Y(x, t, \theta), \phi_q(\zeta(\theta)) \rangle = \left< L \left( x, t; \sum_{r=0}^{K} \alpha_r(t) \phi_r(\zeta(\theta)) \right), \phi_q(\zeta(\theta)) \right>$$

This is a system of $K + 1$ deterministic equations. These can be solved using integration of the unknown with respect to the particular random variable over a suitably sized stochastic space.

An example of a non-intrusive method to solve the PC coefficients is linear regression [32]. Here the coefficients are computed using an overdetermined least squares problem of the following form:

$$\phi(\zeta_k) \alpha(t) = Y(t, \zeta_k)$$

A popular class of ‘surprise’ quantification algorithms over the past few decades have been those related to Bayesian
approaches [33, 34, 35]. Bretthorst, in his work, shows how the entropy method of moments can be formulated in terms of Bayesian probability theory [36]. Through this approach, Lagrange multipliers are expressed in terms of their marginalized posterior distributions with respect to the number of multipliers and data. Markov Chain Monte Carlo based techniques are then used in obtaining the solutions. While this approach finds solutions that are more optimal than the method of moments, there is a significant computational cost involved with Bayesian approaches and Monte Carlo methods that cannot be ignored. A very popular algorithm in this class that gained significant attention over the past decade is that based on a Bayesian definition of surprise by Itti and Baldi [37, 38]. The authors formulate the definition of surprise by using the concept of relative entropy. They evaluate the expected Kullback-Leibler (KL) divergence between the prior and posterior distributions of the model after obtaining new data. The formulation is given below:

$$S(D, \mathcal{M}) = KL(P(M|D), P(M)) = \int_{\mathcal{M}} P(M|D) \log \frac{P(M|D)}{P(M)} dM$$

Here, $M$ is the set of models in a model space given by $\mathcal{M}$. D represents the observed data. We can infer from the formulation that if an observed data leads to a significant change in the posterior distribution of the model, it will subsequently lead to high KL divergence between the prior and posterior thereby causing a high surprise quantity.

E. Quantum Modeling of Stock Markets

The idea of applying quantum theory to data analysis has been pervasive in the field of econophysics [39-45]. The main motivation for doing so comes from treating highly dynamical signals with a large number of unknown generating functions (such as stock data) as a mixture of different quantum eigenstates. The earliest such attempts of using physics to describe data can be traced back to 1933 when Frisch attempted to use concepts of classical physics to model finance related dynamics [40]. One of the first crucial attempts in linking quantum physics with finance was made with the quantum field theory interpretation of financial markets where critical concepts from quantum mechanics such as path integrals and differential manifolds were intensively used in the study of the dynamics of financial markets [41, 42]. Recent work in this field involves non-classical oscillator models of the Chinese stock markets [42, 43, 44, 45].

F. Proposed Framework and Contributions

We propose a framework for characterizing a real valued time series or signal in terms of the moment decomposition of its point-wise stochasticity. Here, stochasticity refers to the uncertainty associated with the local structure of data functionals, instead of employing a probability measure. In principle our methodology can also be linked to Gaussian processes theory [25], except that we do not operate in the data space with its intrinsic probability law, but we prefer to concentrate on the Reproducing Kernel Hilbert Space (RKHS) equivalent functional representation to quantify the local kernel structure of the data. Hence our framework is more closely associated with point estimate methods. The hallmark of our framework of signal decomposition lies in its ability to perform sample by sample decomposition of stochasticity, which is not possible to implement using current techniques. This is done by using a functional called the information potential field which is motivated from Parzen density estimation of data using Gaussian kernel windows [8]. Hence, due the universal approximation property of the Gaussian RKHS, we are able to provide an inclusive local representation of the data, without making assumptions on the type of distribution or stochastic process associated with the data. This enables our framework to exhibit sensitivity towards local signal characteristics while simultaneously providing a generalized representation over all past samples. Since the information potential field is based on the average of the pairwise distances of a point from all available samples, it becomes feasible to treat the associated functional in RKHS as a force field (where the samples are assumed to be identical information particles) and hence formulate a physical description of the information potential field based on the time independent Schrodingers equation [46]. This quantum description of the information potential field has significant consequence in our framework since it provides a quantification of uncertainty associated with the local RKHS structure. A crucial feature of this quantum description is that the wave function is defined in the RKHS where the basis functions keep getting updated with every sample. The time-independent Schrodingers equation hence represents the local structure of the RKHS (which formally corresponds to a pointwise estimate of the PDF in a probability interpretation [47]) as a combined representation of updated standing waves at any time instance. Since the temporal dependencies of the signal are embedded in the wave function along all even order moments (see section III), the Schrodingers equation, despite being time-independent, provides an efficient spatiotemporal characterization of a time series (or signal). As is evident from the literature of entropy and stochastic decomposition, uncertainties associated with data consist of a mixture of contributions from different moments of data or its stochastic representation [13, 16, 26]. Hence, we utilize this quantum description of the signals local RKHS structure to extract higher order intrinsic modes (eigenstates) associated with the its uncertainty and their corresponding eigenvalues. We do this by projecting the first order mode of the wave function into successively higher orders of Hermite polynomial space at every sample. This is similar to formulations associated with polynomial chaos expansion in the field of uncertainty quantification and it also follows the principles of the well-known solution of the quantum harmonic oscillator. The projections, in our case however, are carried out on a sample-by-sample basis. Thereafter, we compute the corresponding information potentials associated with the various extracted modes. The computed information potentials at the different modes summarize the signal uncertainties (in the RKHS) at these modes. The overall framework is depicted in fig. 1 with the associated formulations and reasonings described in detail.
in section III. A summary of the main steps involved in our sample based framework is as follows:

- Quantum description of the local RKHS space using an information potential field created by the selected kernel (and hence a description of uncertainty). We restrict the analysis to the Gaussian kernel in this paper.
- Extraction of different modes of the wave-function associated with the quantum description of the kernel space using Hermite polynomial embeddings.
- Evaluation of the Laplacians over the wave-function modes.
- Computation of the information potential along each mode using the computed Laplacians.

Apart from the ability to operate on a sample-by-sample basis, our framework for uncertainty decomposition of signals presents several advantages over related methods summarized before. Due to the way the Lagrangians are computed, a fundamental limitation of the maximum entropy method of moments is that the moment correlations embedded in the data are ignored. This limits the analysis of the PDF to a very restricted space close to the space of available samples. Our framework, on the other hand, expands the space of analysis much further away from the space of samples by evaluating moments of the wave-function, which in itself, already contains information related to all even order temporal dependencies of the data (i.e. pairwise differences between the point under consideration and all the other samples that have occurred over time). Moreover, we show in section III that the quantum description of the information potential field in our framework involves computation of the Laplacian which takes into account both temporal and inter-modal local dynamics. Furthermore, Hermitian expansions of the wave-function provide a systematic relationship between the successive moments of the wave-function, i.e. of the different states of information potential field. Hence, our framework maximizes information gain and also provides a more stable basis for evaluating as many moments as needed unlike the method of moments where moment expansions depend on how feasible the optimization problem is for finding the Lagrangians. We show in the formulation of our framework in section III that the eigenvalues associated with the different modes are empirically defined as a result of simply imposing a constraint on the information potential field corresponding to any mode to be always positive. This greatly simplifies the eigenvalue determination process and hence provides an advantage to our framework when compared to polynomial chaos based methods where eigenvalue determination process is an optimization problem and hence faces convergence and computational cost issues depending on the type of numerical technique used. Our proposed framework also has several advantages over the stock modeling methods in the field of econometrics summarized before. All of the aforementioned work on quantum modeling are restricted to only specific properties pertaining to the class of stock market data being considered thereby not providing a general framework. In many of these models (43), for instance), authors make specific assumptions on the stochasticity of data (type of motion, drifts, etc) based on the fluctuations of particular stocks. We make no assumptions on the stochastic nature of data. Additionally, unlike some of the stock models (for instance, in 43) where authors use time-dependent formulations of the stock data PDF, we use a time-independent Schrodinger's equation since our local functional, defining the wave equation, implicitly contains all temporal dependencies of the data. This use of a time-independent formulation greatly simplifies our framework.

II. BACKGROUND INFORMATION

A. Derivation of Information Potential Field (IPF)

Renyis entropy of order alpha ($\alpha \neq 1$, $\alpha > 0$) for a continuous random variable $x$ is given by:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \int p(x)\alpha dx$$

(8)

The case of $\alpha = 2$, Renyi's quadratic entropy, is particular important because it leads to an efficient nonparametric family of estimators using the Parzen window method [48]. In this case, Renyi quadratic entropy becomes just the log of the expected value of the probability density function, i.e.

$$H_2(X) = -\log \int p(x)^2 dx = -\log V(X)$$

(9)

Let us call the argument of the logarithm $V(X)$, the information potential (IP) of the data set, which is a number that is nothing but the mean value of the PDF. Let us assume that we use a Gaussian window for the Parzen estimation, with bandwidth (or kernel size) $\sigma$. One can readily estimate directly from experimental data $x_i$, $i = 1, ..., N$ the information potential,

$$V(X) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_{\sigma\sqrt{\sigma}}(x_i - x_j)$$

(10)

i.e., the IP is a number obtained by the double sum of the Gaussian functions centered at differences of samples with a larger kernel size. There is a physical interpretation of $V(X)$ if we think the samples as particles in a potential field, hence the name information potential. Let us define a function over the data space $V(x)$ as

$$V(x) = \frac{1}{N} \sum_{i=1}^{N} G(x - x_i)$$

(11)

which we will call the information potential field (IPF). $V(x)$ is a continuous function obtained by the sum of Gaussian bumps centered on the samples, which is the estimated PDF $\hat{p}(x)$ obtained by the Parzen window, but it can also be interpreted as a potential field similar to gravity, over the space of the samples. In fact, if we attribute unitary mass to all our samples, we can say that the IPF is always positive and regions of space with more samples will have a larger IP, while regions of the space with few samples will have less IP. Here, the shape of the Parzen window will determine the gravity, instead of the inverse law of physics. The information potential $V(X)$ in
Renys entropy is nothing but the total potential field created by the samples in the data set, i.e. \( V(x) = \frac{1}{N} \sum_{j=1}^{N} V(x_j) \).

B. Quantum Information Potential field (QIPF)

The idea of a potential field (aka probability density) over the space of the samples can be readily extended with quantum theory concepts. The Schrodinger stationary (time independent) equation for a particle in the presence of a potential field can be written as

\[
\frac{\hbar^2}{2m} \nabla^2 \psi(x) + \psi(x)(E - V_s(x)) = 0 \tag{12}
\]

where \( \hbar \) is the Planck’s constant, \( m \) the mass of the particle and the wave function \( \psi \) determines the spatial probability of the particle with \( \psi(x) = |\psi(x)|^2 \). \( V_s(x) \) is the potential energy as a function of position, \( E \) corresponds to the allowable energy state of the particle and \( \psi \) becomes the corresponding eigenvector. For a set of information particles with a Gaussian kernel, the wave-function for one dimensional information particle becomes,

\[
\psi(x) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} G_s(x - x_i)} \tag{13}
\]

We can also rescale \( V_s(x) \) such that there is a single free parameter \( \sigma \) in (12) to yield

\[
H \psi(x) = \left( -\frac{\sigma^2}{2} \nabla^2 + V_s(x) \right) \psi(x) = E \psi(x) \tag{14}
\]

Solving for \( V_s(x) \), we obtain:

\[
V_s(x) = E + \frac{\sigma^2}{2} \frac{\nabla^2 \psi(x)}{\psi(x)} \tag{15}
\]

which we called the quantum information potential field (QIPF) denoted by \( V_s(x) \). This can also be simplified as:

\[
V_s(x) = E - \frac{1}{2} \left( \frac{1}{2\sigma^2 \psi(x)} \right) \sum_i (x - x_i)^2 e^{-(x - x_i)^2/2\sigma^2} \tag{16}
\]

To determine the value of \( V_s(x) \) uniquely we can require that \( \min V_s(x) = 0 \), which makes \( E = -\min \frac{\sigma^2}{2} \frac{\nabla^2 \psi(x)}{\psi(x)} \)

where \( 0 \leq E \leq 1/2 \). Note that \( \psi(x) \) is the eigenfunction of \( H \) and \( E \) is the lowest eigenvalue of the operator, which corresponds to the ground state. Given the data set, we expect \( V_s(x) \) to increase quadratically outside the data region and to exhibit local minima associated with the locations of highest sample density (clusters). This can be interpreted as clustering since the potential function affects the data distribution function \( \psi(x) \) to its minima, while the Laplacian drives it away, producing a complicated potential function in the space. We should remark that in this framework \( E \) sets the scale at which the minima are observed. This derivation can be easily extended to multidimensional data.

We can see that \( V_s(x) \) is also a potential function that differs from \( V(x) \) in (11) because it is associated with a quantum description of the IPF. The two fields are similar to each other for Gaussian kernels since the derivative of the Gaussian is a Gaussian, but it presents a big advantage because now \( V_s(x) \) can be operated independently as a counterpart of the samples, describing the duality between particles (samples) and waves (functionals).

III. PROPOSED APPROACH: EXTRACTING MODES OF UNCERTAINTY IN THE RKHS

As one can notice, when we derive information potential as a special case of Parzen’s density estimator, it becomes a localized PDF estimator. However, from a deterministic sense, information potential field is nothing but a local functional based on the averaged pairwise distance of a point from all the other points in the Gaussian kernel space. From this perspective, its quantum description introduces uncertainty in the functional space associated with the sample location being considered at any point of time: One can consider that at an information particle location, we have no uncertainty (the field collapses to a delta function), but at any point in the RKHS that has no sample, there is an uncertainty that is quantified by the wave function. Our framework attempts to utilize these concepts to extract the different eigenstates of the QIPF associated with the signal. The main motivation for doing so comes from assuming the signal was created by a quantum physical system that is governed by a large number of unknown forces. Consequently, the extracted eigenstates associated with the quantum framework of the signal would provide a characterization of the intrinsic governing forces acting on the system that produced the signal.

As is well known in the field of quantum physics, the time-independent Schrödinger’s equation describing the dynamics of the quantum harmonic oscillator can be formulated as:

\[
\frac{d^2 \psi}{dx^2} + \left( \frac{2mE}{\hbar^2} - \frac{m^2 \omega^2}{\hbar^2} x^2 \right) \psi(x) = 0 \tag{17}
\]

where the different terms have their usual meaning.

It is widely known that (17) can be solved using the power series method by introducing a dimensionless variable as an expression of \( x \) given by \( y = \sqrt{\frac{m}{\hbar}} x \). This yields wave functions of different modes, \( \psi_n(x) \), that are in fact, consecutive projections of \( y \) in the space of Hermite polynomials. These solutions are given as:

\[
E_0 = \frac{\hbar \omega}{2}, \quad \psi_0 = \alpha_0 e^{-\frac{y^2}{2}} \\
E_1 = \frac{3\hbar \omega}{2}, \quad \psi_1 = \alpha_0 (2y) e^{-\frac{y^2}{2}} \\
E_2 = \frac{5\hbar \omega}{2}, \quad \psi_2 = \alpha_0 (4y^2 - 2) e^{-\frac{y^2}{2}} \\
\ldots
\]

Hence, the solution to the Schrödinger’s equation for the harmonic oscillator yields infinite eigenfunctions (denoted by \( \psi_0, \psi_1, \psi_2 \ldots \)) that are embedded in the orthogonal space of Hermite polynomials. Their corresponding eigenvalues are denoted by \( E_0, E_1, E_2 \ldots \).
One can obtain similar Hermitian embeddings by projecting the wave-function of the QIPF (13) into the orthogonal spaces of consecutive Hermite polynomials. Our conjecture is that by doing so, we are obtaining the approximate intrinsic modes associated with the local RKHS structure (or the local PDF, from a probabilistic perspective) of the signal as defined by the corresponding information potential field. Hence, by exploiting the Hermitian relationship between the different higher order modes of the QIPF, it becomes possible to construct its eigenfunctions. This is conceptually very similar to working with the moment expansions of the characteristic function in statistics, but without imposing a preset family of basis functions (the complex exponentials), here one uses a data centric decomposition. The generating function of the Hermite polynomial family is given by:

$$H_n(x) = \frac{(-1)^n}{\sqrt{n!}} e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$$  \hspace{1cm} (19)

Upon inserting the QIPF wave function in (19), we obtain the following generating function for the wave-functions at different eigenstates:

$$H_n(\psi(x)) = (-1)^n e^{\psi(x)^2} \frac{d^n}{dx^n} e^{-\psi(x)^2}$$  \hspace{1cm} (20)

The recurrence property of Hermite polynomials allows us to analytically compute successive orders of eigenfunctions without having to use the generating function given by (19). This saves significant computational time when evaluating higher order eigenfunctions of the signal’s QIPF in real time by computing the next order of polynomial expansion from the current order and the previous order using the following relation:

$$H_{n+1}(y) = 2yH_n(y) - 2nH_{n-1}(y)$$  \hspace{1cm} (21)

Since the Hermite polynomial expansion is being done in the Gaussian kernel space (which is an even function), we only consider even order Hermite expansions [49]. Thus, on expanding (20) for different even values of n (eigenstates), we get the following series of eigenfunctions in the QIPF:

$$\psi^0 = 1$$

$$\psi^2 = 4\psi(x)^2 - 2$$

$$\psi^4 = 16\psi(x)^4 - 48\psi(x)^2 + 12$$

$$\psi^6 = 64\psi(x)^6 - 480\psi(x)^4 + 720\psi(x)^2 - 120$$

$$\cdots$$

The family of functions thus obtained in (22) form a new space of orthogonal basis functions within the RKHS totally defined by the time series. While the Gaussian kernel framework provides us with a with a universal tool for signal representation, the subsequent Hermitian projections (quantified as wave functions) decomposes the local RKHS structure of the signal in terms of its various underlying moments thereby quantifying uncertainties along those moments.

We can normalize the Hermitian expansions such that they satisfy the following relation:

$$\int_{-\infty}^{\infty} e^{-y^2} |H(y)|^2 dy = 1$$  \hspace{1cm} (23)

We now evaluate the QIPF of samples at the different eigenfunction projections of the wave equation using (15) which is generalized here for all orders of expansion:

$$V_{si}^k(x) = E_i^k + \frac{\sigma^2/2\nabla^2 \psi_i^k(x)}{\psi_i^k(x)}$$  \hspace{1cm} (24)

Here, $V_{si}^k$ represents the $k^{th}$ mode of the Schrodinger’s information potential at sample $i$, $E_i^k$ is the corresponding eigenvalue of the $k^{th}$ mode at sample $i$ and $\psi_i^k(x)$ is the $k^{th}$ mode wave-function value at sample $i$. The Laplacian
leading to:

\[ \min V \]

empirically by asserting the requirement that function evaluations of information potential are determined extracted modes. The energy values the local dynamics of the signal along time and across the advantage to our framework since it efficiently characterizes

operator used in the second term in (24) provides a critical advantage to our framework since it efficiently characterizes the local dynamics of the signal along time and across the extracted modes. The energy values \( E_i \) at the different eigenfunction evaluations of information potential are determined empirically by asserting the requirement that \( \min V \) at the different energy levels or modes. Due to the simplicity of the constraint, it may be argued that the eigenvalues in our case do not provide exact measures of the contribution of each mode. However, we show in our experimental evaluations in section IV that they are capable of uniquely characterizing the dynamics of different signals. The proposed framework is summarized in fig. 1 and Algorithm 1.

**Algorithm 1** Quantum Decomposition of Local PDF

**Input:**

\( x \): Signal
\( \sigma \): Kernel Width
\( m \): Number of Quantum Modes

**Initialization:**

\( \psi \): Wave Function
\( \psi^2, \psi^4, ... , \psi^m \): Wave Function Hermitian Embeddings
\( V^2, V^4, ... , V^m \): QIPF Modes
\( E^2, E^4, ... , E^m \): Eigenvalue of Each Mode

**Computations:**

for \( i = 1 \) to length(\( x \)) do

\[ \psi = 0 \]

for \( j = 1 \) to \( i \) do

\[ \psi \leftarrow \psi + e^{-\frac{(x_i-x_j)^2}{2\sigma^2}} \]
end for

\[ \psi_i \leftarrow \sqrt{\text{mean}(\psi)} \]

\[ [\psi^2, \psi^4, ... , \psi^m] \leftarrow \text{HermiteProjections}(\psi_i) \]

\[ [\nabla^2 \psi^2, \nabla^2 \psi^4, ... , \nabla^2 \psi^m] \leftarrow \text{Laplacians} \]

for each mode \( k \) do

\[ E^k_i = -\min_{q=1...m} \frac{\sigma^2/2\nabla^2 \psi^k_q(x)}{\psi^k_q(x)} \]

\[ V^k_{s(i)} = E^k_i + \frac{\sigma^2/2\nabla^2 \psi^k}{\psi^k} \]
end for

end for

IV. Simulation Results and Analysis

In order to study the trends of the extracted QIPF modes in the sample space and analyze their response to various kernel widths, we first perform a spatial analysis of the different QIPF modes. We implement our framework on 500 samples of Lorenz series (which is a chaotic dynamical signal) to extract the first 6 even order modes of the QIPF. All simulations were conducted using MATLAB R2018b on a 2.3 Ghz intel i5 processor machine. The generating functions of the Lorenz series consist of the following three mutually coupled differential equations (with \( \sigma, \rho \) and \( \beta \) as system parameters):

\[ \frac{dx}{dy} = \sigma(y-x) \]

\[ \frac{dy}{dt} = x(\rho-z) - y \]

\[ \frac{dz}{dt} = xy - \beta z \]

(26)

The Lorenz series is generated with the parameters set as \( \sigma = 10 \), \( \rho = 28 \) and \( \beta = 8/3 \). The initial conditions were set as \( x_1 = 1 \), \( y_1 = 1 \) and \( z_1 = 1.05 \). The signal was also normalized to zero mean and unit variance and hence roughly varied between the amplitude levels of -2 and 2. We follow the steps shown in fig. 1 and algorithm 1 and first compute the fundamental wave function of the Schrödinger’s information potential field (QIPF) using all of the samples of the signal at different locations in the space of the samples. This is followed by computation of the Hermite polynomial embeddings of the fundamental mode wave function. We use 6 successive even order Hermite polynomial expansions for our analysis. We compute the QIPF mode corresponding to each of the extracted modes of the wave function using (24) and (25), where \( k \) represents the mode number and \( i \) represents the sample number. Fig. 2 shows the plots of these modes extracted at different kernel widths with respect to the different locations in the sample space. Several observations can be made from the plots. Firstly, one can observe that the different QIPF modes peak in a mutually exclusive manner and spread out upon increasing the kernel width. The mutual exclusivity can be attributed to the orthogonal nature of Hermite polynomial expansions. Secondly, we observe that the QIPF modes peak away from the mean (especially in regions outside of the dynamic range of the signal) and hence emphasize the more uncertain regions of the sample space. The higher order states successively emphasize the more distant regions in an increasing order of the state number. The trends of the different QIPF states in the region within the dynamic range of the signal are more sensitive to the kernel width. For extremely low kernel widths, as can be seen from fig. 2(a), the first few QIPF modes get completely diminished and the higher order modes dominate the signal’s representation in all regions of sample space. Since the local kernel space of the signal modeled using extremely low kernel widths will be very uncertain, it is expected for higher order modes to dominate the signal’s representation in such cases. On increasing the kernel width (fig. 2(b)), some lower order modes (modes 2 and
3) begin to dominate the signal’s representation in the region around the mean. For moderate kernel widths (figs. 2(c) and 2(d)), the first mode of the QIFP dominates the immediate region around the mean. For larger kernel widths (figs. 2(e) and 2(f)) which exceed the dynamic range of the signal, we can see that high order modes begin to emerge in the region around the mean. Since very large kernel widths imply equally likely outcomes (or increased uncertainty), it again makes sense for the higher order modes to dominate the signal’s representation around the mean region. This behavior is remarkable similar to physical systems. Suppose that we have a drum, with a skin membrane. If we increase the tension of the membrane, and we hit the membrane the drum will vibrate for a long time. In our potential field the stiffness is controlled by the kernel size. If the kernel size is large the QIFP becomes stiffer so the energy in the higher models will increase. If one decreases the kernel size, the membrane becomes more elastic and there are many local modes that decay much faster. One remarkable variable that we have not yet fully identified is the wavelength of the QIFP propagation, but its dependency on the kernel size is clear.

We analyze how the different regions and local dynamics of a signal are represented by our framework in real time by doing a sample by sample implementation. As an illustrative example to this end, we implement our framework on 500 samples of Lorenz series generated and normalized in the same way as before. Furthermore, the signal is scaled to half its original amplitude level. This scaled signal is used for constructing the QIFP. The points of evaluation of the QIFP, however, are sample locations of the unscaled signal as shown in fig. 3(a). This is done to evaluate the QIFP at regions outside of the dynamical range of the signal (in addition to the inside ones). It should be noted that the points of evaluation inside the dynamical range of the signal do not generally overlap with the scaled signal. They do, however, roughly have the same dynamical structure at the corresponding points. We evaluate the QIFP at each point of evaluation by extracting the QIFP modes and taking their average. At each point of evaluation, only the past samples of the scaled signal are utilized for the computations at that point. Fig. 3(b) shows the average of first modes.

![Fig. 2: Analysis of peak locations of the different modes in the input space as a function of kernel width](image)

![Fig. 3: Comparison of conventional information potential and the average value of QIFP modes evaluated at every sample of a section of Lorenz series](image)
5 extracted modes of the QIPF at each point of evaluation and fig. 3(c) shows the same for the first 10 modes of the QIPF. For visual clarity, only values from 200th to 500th points of evaluation are shown for both cases. We can observe that in both cases, there is a significant difference between the sensitivity of the QIPF and that of the conventional form of information potential. They also follow opposite trends with respect to the points of evaluation since the conventional (classical) IPF follows the PDF of the signal whereas the QIPF follows the uncertainty. The QIPF can also be seen to drastically increase at points of evaluation that are outside of the dynamical range of the scaled signal. It tends to be the lowest at high sample density regions. Upon comparing fig. 3(b) and fig. 3(c), one can notice that increasing the number of modes of the QIPF leads to a more detailed evaluation of how uniquely uncertainty gets quantified at every point even within the dynamical signal range. It is interesting to observe that there are more jumps/peaks in the uncertainty within the dynamical signal range when the number of modes is increased, especially at points where the signal dynamics visibly change.

To demonstrate the effectiveness of the proposed framework in characterizing a signal in terms of its various intrinsic dynamical modes, we use a pedagogical example of how a sine wave function differs from Lorenz series with respect to their QIPF modes. Since our framework in based on the decomposition of local kernel space structure in terms of even order oscillator harmonics, we expect the sine wave, which consists of a single oscillator generating function, to get encoded in much fewer quantum information potential modes when compared to Lorenz series which is a chaotic dynamical system. We generated a sine wave having a frequency of 100 Hz and was sampled at a rate of 8000 samples per second to mimic a continuous signal for a total time of 0.16 seconds. The Lorenz series was generated in the same way as before (without amplitude scaling). Both signals were normalized to zero mean and unit variance. We compute the QIPF modes using (24) and (25) and extract 18 successive even order modes of the QIPF to encode the information associated with the sine wave and the Lorenz series. The kernel width used for doing so for both signals was fixed to 0.3 which is sufficiently small for more emphasis of the local regions of the sample space. Fig. 4 shows the signals (left column) and the corresponding histogram plots (right column) of the number of times the value of each QIPF mode dominated over the others throughout the durations of the signals. As can be seen in fig. 4, there are only two dominant modes in case of the sine wave (modes 2 and 3). The dominant modes in the case of Lorenz series, on the other hand, are significantly more spread out indicating a more complex dynamical structure of the signal.

We extend this analysis to different frequencies and sampling rates of the sine wave signal (figs. 5 and 6) as well as different parameters and initial conditions of the Lorenz series (fig. 7). It is interesting to note in fig. 5 that the effect of increasing the signal’s frequency \(f_0\) has no significant effect on the distribution of QIPF modes that dominate over the signal (which is still limited to modes 2 and 3). There is, however, a slight reduction in the third mode’s proportion when the signal is aliased as can be seen in fig. 5(c). Fig. 6 shows the same analysis done on sine waves generated using a mixture of different frequency components. As can be seen
Fig. 7: Left: Lorenz series of different parameters or initial conditions. Right: Dominance frequency of QIPF states.

From the corresponding histogram plots, the distribution of QIPF modes that are dominant begins to spread out more towards the higher modes when the number of frequency components in the signal increases. On the other hand, the modal distribution on the Lorenz series, in fig. 7, can be seen to be significantly more sensitive to changing generating system parameters and initial conditions. Since chaotic dynamical systems have a high sensitivity and long-term dependency towards system parameters and initial conditions, the increased sensitivity observed towards these factors is expected. This is indicative of the framework’s increased sensitivity towards the signal’s generating systems in spite of all computations being done locally.

We also compute the QIPF energy levels of the first 18 even order modes of the information potential extracted from various signals. The formulation for computing the energy levels associated with various states is given by (25), which is the minimum value of the Laplacians at each mode required to constraint the information potential at that mode to be positive. Fig. 8(a) shows the normalized energy levels associated with the first 18 even order states of sine wave signals of different frequencies. All of these signals had the same sampling rate (8000 Hz) and the samples in the first 0.05 seconds of each signal were considered for the computations. The eigenvalues shown here for each mode were computed using the Laplacians over all of the samples simultaneously (unlike sample-by-sample approach used in the first experiment). Fig 8(b) similarly shows the normalized QIPF energy levels associated with the first 18 even order modes of the information potential of the Lorenz series signals generated using different parameters or initial conditions. 500 samples for each Lorenz series signal were used for the evaluations. It can be seen from both subfigures in fig. 8 that the energy levels increase consistently with the state number. It is interesting to note here that the slope of the energy level vs state number curve generally flattens out rather quickly (within first 4-5 modes) for the sine wave signals as compared to that associated with the Lorenz series signals. This suggests (and is also expected) that higher order states do not contribute significantly in the representation of the sine wave dynamics as they do in the representation of Lorenz series signals. This is also consistent with the trends observed in the first set of results. For increased frequency sine wave, however, more variance is observed in the energy levels of higher order states, perhaps because of the finite sampling frequency. For both the classes of signals in fig. 8, one can observe that the first few modes of the QIPF are quite similar. Discriminative properties particular to specific signals (generated by changing parameters, initial conditions or fundamental frequency components) start to become increasingly significant in higher order QIPF states.

In order to illustrate the sensitivity of the proposed frame-
work on the changing dynamical behavior and uncertainties associated with the signal, we implement our framework on a sample by sample basis and test it on 1100 samples of Lorenz series where the last 500 samples are corrupted using heteroscedastic white Gaussian noise, where the variance (and consequently the SNR) randomly changes after every 100 samples (shown in table 1). The first 600 samples are uncorrupted. We use a window of past 100 samples to perform computations at each sample. The signal under consideration is shown in fig. 9 (top). The noise is added from the 600th sample onwards. For reference, the corresponding original (without decomposition) information potential values of the samples are shown in fig. 9 (bottom). We compute the first 25 quantum states of the QIPF using the same formulations depicted by fig. 1 at every sample using a moderate kernel width of 0.8 (and with the IP being computed using the window of 100 past samples)

The values of the different states of the QIPF at every sample is expressed as a heat-map in fig. 10. For clarity of analysis, all values in the heat-map matrix are normalized with respect to all past values only in its corresponding row (i.e. all past values of the particular state). The first two sets of noise, as can be seen from table 1, have low variances. From the heat-map, we can observe how the different states of the QIPF react to the inclusion of noise. The first few states (1 to 5) show very little response to all sets of noise in general. As we move on to the middle order states (6 to 11), we observe that the response to higher variance noise increases. The response to the low variance noise in these states continues to remain low. Higher order states (state 12 onwards) show increased sensitivity to all noise variances. It is interesting to observe here that the noise sets get encoded quite differently among the high order states. Although all of the higher order states generally show increased sensitivity to noise, the sensitivity of each state is significantly different from the others for different noise variances. Fig. 11 shows the expected values of the different sets of QIPF states (1-7, 8-15 and 18-25). The sensitivity trends of the different sets of QIPF states are apparent here with the expected values of the first 7 states showing very little change with the addition of noise. The expected value of the middle order states (8 to 15) can be seen to be significantly more sensitive to high variance noise form the 800th sample onwards. Expected value of the higher order states (18 to 25) shows increased sensitivity towards all noise variances as is evident from the relatively quick responses to all sets of noise seen in fig. 11.

We use this approach to quantify the sensitivity of our framework towards statistical changes in the signal so as to compare our framework with entropy based surprise quantification methods [37], [38]. We perform the quantification and analysis over 5000 samples of Mackey Glass chaotic dynamical series. Mackey Glass chaotic dynamical system is governed by the following non-linear delay differential equation:

\[
\frac{dx}{dt} = \alpha \frac{x(t - \tau)}{1 + x(t - \tau)^n} - \beta x \quad \alpha, \beta and n > 0 \quad (27)
\]
where \( \tau \) is the delay and \( \alpha, n \) and \( \beta \) are other parameters.

The 5000 samples were generated by setting the delay parameter \( (\tau) \) as 30 and other parameters as \( \alpha = 0.2, \beta = 0.1 \). Heteroscedastic noise, with the variance changing after every 500 samples, was added to the last 2500 samples of the generated Mackey Glass signal. We implement our framework on the signal by extracting the first 10 even order modes of the generated Mackey Glass signal. We implement our framework at each sample and independent of other frameworks before comparing the sensitivity of our framework with that of the classical IP states are more sensitive at lower kernel widths. On increasing the kernel width, the lower order set changes in signal statistics than other models. Upon analyzing how the sensitivity of our framework changes with respect to kernel width, one can notice that the higher order set of IP states are more sensitive at lower kernel widths. On increasing the kernel width, the lower order IP states start becoming more sensitive as compared to other states. This is expected since increasing the kernel width starts to spread out the higher order states to beyond the dynamical range of the signal within which changes take place. However, there is also a compromise here because the best sensitivity values (among the state groups) decrease with increase in the kernel width. There is also a compromise here because the best sensitivity values (among the state groups) decrease with increase in the kernel width. Hence, using this measure, we evaluate the change in a particular state group with respect to the change in noise from one sample interval, \( R \), to the next, \( R+1 \) (similar to the previous analysis). The sample intervals are non-overlapping and have a length of 500 samples which is the same as the interval length in the heteroscedastic noise where the variance is constant. We then quantify the change detection performance of the three groups of states by measuring their sensitivity with respect to the change in noise variance. The sensitivity is measured by evaluating the change in Euclidean norm of the state values from one interval of the noise corrupted samples to the next where the variance of the noise changes. The formulation of sensitivity \( (\zeta) \) is given as follows:

\[
\zeta = E_i \left( \left| \frac{\left| V_{R}^{K} \right| - \left| V_{R+1}^{K} \right|}{D_{R} - D_{R+1}} \right| \right)
\]  

(28)

Here \( R \) is the sample interval and \( K \) is 1-3, 4-6 or 7-9 depending on which state group we consider. \( D_{R} \) is the decibel measure of noise in the sample interval \( R \). Hence, using this measure, we evaluate the change in a particular state group with respect to the change in noise from one sample interval, \( R \), to the next, \( R+1 \) (similar to the previous analysis). The sample intervals are non-overlapping and have a length of 500 samples which is the same as the interval length in the heteroscedastic noise where the variance is constant. We compare the sensitivity of our framework with that of the Bayesian surprise model \( (7) \) which measures the amount of surprise associated with each data sample by evaluating the KL divergence between the prior and posterior model distributions. It is reformulated here as below:

\[
S(D, M) = \int_{M} P(M|D) \log \frac{1}{P(M|D)} dM - \int_{\mathcal{M}} P(M|D) \log \frac{1}{P(M|D)} dM
\]  

(29)

For comparisons, we use the Parzen density (using Gaussian kernel windows) as the model distribution \( P(\hat{M}) \) in the Bayesian surprise framework since the information potential, which is used by our framework as the local kernel space quantifier, is ultimately a metric derived from Parzen density estimation (as has been shown in section II). The model distribution in the Bayesian surprise framework is updated with every sample. However, the space over which integral computations are done at each sample (model space \( \mathcal{M} \) in (29)) is fixed to be a regularly spaced set of values lying in the dynamic range of the generated Mackey Glass series. The sensitivity here is measured by replacing \( V_{R}^{K} \) and \( V_{R+1}^{K} \) in (28) with the two integral computations seen in (29). Here \( P(M|D) \) is the set of updated distributions with respect to the data in the next interval of samples. Hence sensitivity related to KL divergence evaluated here measures the differences in statistics of the neighboring interval of samples where the noise variances are different from each other. The other metric we compare our framework with is the sensitivity associated with simply the entropy differences between the different sample intervals where the noise variance changes. The sensitivity values for the different frameworks evaluated at various kernel widths are shown in table II. The values shown are the average results of 10 simulation runs for each framework. For each framework, normalization (to zero mean and unit variance) was done with respect to evaluated values at each sample and independent of other frameworks before evaluating the sensitivity. It can be seen here that for all kernel widths above 0.2, our framework has higher sensitivity to changes in signal statistics than other models. Upon analyzing how the sensitivity of our framework changes with respect to kernel width, one can notice that the higher order set of IP states are more sensitive at lower kernel widths. On increasing the kernel width, the lower order IP states start becoming more sensitive as compared to other states. This is expected since increasing the kernel width starts to spread out the higher order states to beyond the dynamical range of the signal within which changes take place. However, there is also a compromise here because the best sensitivity values (among the state groups) decrease with increase in the kernel width. Our framework is the most sensitive at moderate kernel widths of 0.5 and 0.6 for this case.

### Table II: Sensitivity values of several frameworks at different kernel widths

| Framework          | 0.2         | 0.4         | 0.5         | 0.6         | 0.8         | 1           |
|--------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| States 1-3         | 1.1860      | 1.0996      | 0.9047      | 0.5578      | 2.5123      | 1.7937      |
| QIPF States 4-6    | 1.4799      | 0.8292      | 2.8727      | 4.4312      | 1.7927      | 1.6161      |
| States 7-10        | 1.4956      | 3.7847      | 3.7879      | 1.4028      | 1.2349      | 1.2133      |
| Bayesian Surprise   | 1.8301      | 1.6548      | 1.6681      | 1.6255      | 1.2913      | 0.5803      |
| Entropy Difference | 1.9174      | 1.7482      | 1.5896      | 1.4160      | 1.0191      | 0.7853      |
| Classical IP       | 0.3174      | 0.2942      | 0.2644      | 0.2451      | 0.2343      | 0.2400      |
V. CONCLUSION

In this paper, we have introduced a new framework for stochastic signal processing that utilizes the information potential as the local kernel space quantifier and is able to process signals stochastically on a sample by sample basis. We exploit the quantum physical description of the information potential to extract its various dynamical modes in a principled manner based on the analogical behavior of a quantum harmonic oscillator. We have highlighted several key advantages of our framework and have showed both theoretically and experimentally how it is sensitive to the local dynamics of the signal while also providing a generalized global representation. We have also shown experimentally that our framework is more sensitive to changing intrinsic signal dynamics when compared to other entropy based methods. For future work, we intend to apply this framework in the context of adaptive filtering and predictive algorithms. We also intend to utilize this framework as an RKHS based uncertainty quantifier to analyze the performance and robustness of deep learning models.

REFERENCES

[1] C.E. Shannon and W. Weaver, “The Mathematical Theory of Communication,” University of Illinois Press, Urbana, 1949.
[2] E.T. Jaynes, “Information Theory and Statistical Mechanics,” Physical Review, Vol. 106, No. 4, 1957.
[3] George R. Terrell, “The Maximal Smoothing Principle in Density Estimation,” Journal of the American Statistical Association, 85:410, 470-477, DOI: 10.1080/01621459.1990.10476223
[4] Hang K. Ryu, “Maximum entropy estimation of density and regression functions,” Journal of Econometrics, Volume 56, Issue 3, 1993, Pages 397-440, ISSN 0304-0076
[5] C. Tsallis, “Possible generalization of Boltzmann-Gibbs statistics,” Journal of Statistical Physics, 52 (1988), p. 479.
[6] S. Kullback and R.A. Leibler, “On information and sufficiency,” Annals of Mathematical Statistics 31 (1960): 79-86.
[7] A. Renyi, “On measures of entropy and information,” Proc. 4th Berkeley Symp. Math. Stat. Prob., 1960, vol. 1, page 547, Berkeley, 1961. University of California Press.
[8] J. Mller and R. P. Waagepetersen, “Modern Statistics for Spatial Point Processes,” AIP Conference Proceedings 1553, 3 (2013).
[9] L. Itt and B. Pál, “Bayesian surprise attracts human attention,” Advances in neural information processing systems (19) MIT Press; Cambridge, MA, 2006. pp. 547554. NIPS*2005
[10] L. Itt and B. Pál, “A Principled Approach to Detecting Surprising Events in Video,” Proceedings in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005, pp. 631637.
[11] R. Frisch, “Propagation problems and impulse problems in dynamic economics,” “Economic Essays in Honor of Gustav Cassel, George Allen and Unwin, London, 1933, pp. 171205.
[12] K. Ilinski, “Physics of Finance,” Wiley, New York, 2001.
[13] Belal E. Baaquie, “Quantum Finance,” Cambridge University Press, Cambridge, 2004.
[14] Ali Ataullah, Ian Davidson and Mark Tippett, “A wave function for stock market returns,” Physica A, 388 (2009) 455-461.
[15] Chao Zhang and Lu. Huang, “A quantum model for the stock market,” Physica A: Statistical Mechanics and its Applications, 389. 10.1016/j.physa.2010.09.008 (2010).
[16] Xiangyi Meng, J.W. Zhang and Hong Guo, “Quantum Brownian motion model for the stock market,” Physica A: Statistical Mechanics and its Applications, 452. 281-288. 10.1016/j.physa.2016.02.026 (2016).
[17] K. Ahn, M. Y. Choi, B. Dai, S. Sohn and B. Yang, “Modeling stock return distributions with a quantum harmonic oscillator,” EPL, 120 (3) 38003 (2017) DOI: 10.1209/0295-5075/120/38003
[18] E. Schrödinger, “An Undulatory Theory of the Mechanics of Atoms and Molecules,” Physical Review, 28 (6): 10491070 (1926).
[19] Kenji Fukumizu, Franciscus R. Bach, and Michael I. Jordan, “Dimensionality Reduction for Supervised Learning with Reproducing Kernel Hilbert Spaces,” J. Mach. Learn. Res., 73-99, Res. 5 (December 2004).
[20] E. Parzen, “On the estimation of a probability density function and mode,” Ann Statist, 33:10651076 (1962)
[21] L. C. Andrews, “Special Functions of Mathematics for Engineers,” 2nd ed., SPIE Optical Engineering Press, Bellingham, WA; Oxford University Press, Oxford (1998)[formerly published as 2nd ed. by McGraw-Hill, New York (1992)].