Proceedings of the 2007 IEEE Symposium on Foundations of Computational Intelligence (FOCI 2007)

Statistical Cryptography using a Fisher-Schrödinger Model

R. C. Venkatesan
Systems Research Corporation
Aundh, Pune 411007, India
ravi@systemsresearchcorp.com

Abstract

A principled procedure to infer a hierarchy of statistical distributions possessing ill-conditioned eigenstructures, from incomplete constraints, is presented. The inference process of the pdf's employs the Fisher information as the measure of uncertainty, and, utilizes a semi-supervised learning paradigm based on a measurement-response model. The principle underlying the learning paradigm involves providing a quantum mechanical connotation to statistical processes. The inferred pdf's constitute a statistical host that facilitates the encryption/decryption of covert information (code). A systematic strategy to encrypt/decrypt code via unitary projections into the null spaces of the ill-conditioned eigenstructures, is presented. Numerical simulations exemplify the efficacy of the model.

1. Introduction

This paper accomplishes a two-fold objective. First, a systematic methodology to infer from incomplete constraints, a hierarchy of statistical distributions corresponding to the *multiple energy states* of a time independent Schrödinger-like equation (TIS-lE), is presented. By definition, the case of incomplete constraints corresponds to scenarios where the number of constraints (physical observables) is less than the dimension of the distribution. The inference procedure employs a semi-supervised learning paradigm, based on a measurement-response model that utilizes the Fisher information (FI) as the measure of uncertainty.

The time independent Schrödinger equation (TISE) is a fundamental equation of physics, that describes the behavior of a particle in the presence of an external potential [1]

$$\underbrace{-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2}+V(x)\psi(x)}_{H^{QM}\psi(x)}=E\psi(x). \tag{1}$$

Here, $\psi(x)$ is the wave function, E is the total energy eigenvalue, $V(x)$ is the external potential, and, H^{QM} is the TISE Hamiltonian. The constants $\hbar$ and m are the Planck constant and the particle mass, respectively. In time independent scenarios, Lagrangians containing the FI as the measure of uncertainty, yield on variational extremization an equation similar to the TISE, i.e. the TIS-lE [2] [1]. *The TIS-lE provides a quantum mechanical connotation to a statistical process*.

Next, a self-consistent strategy to project covert information into the *null spaces* of ill-conditioned eigenstructures possessed by the inferred host statistical distributions corresponding to the *multiple energy states* of the TIS-lE, is described. The strategy of unitary projection of covert information into the *null spaces* of the ill conditioned eigenstructures of a hierarchy of statistical distributions, has been recently studied for host probability density functions (*pdf*'s, hereafter) inferred using the maximum entropy (MaxEnt) principle [3].

The selective projection of covert information into a hierarchy of statistical distributions implies that the dimension of the covert information is greater than that of any single host distribution. This selective projection endows the code[2] with multiple layers of security, without altering the host statistical distributions. The present paper accomplishes the task of achieving both symmetric and asymmetric cryptography [4] via a judicious amalgamation of statistical inference using an information theoretic semi-supervised learning paradigm, quantum mechanics, and, the theory of unitary projections.

In summary, the semi-supervised learning paradigm is utilized to infer the statistical hosts possessing ill-conditioned eigenstructures. The code is then projected into the *null spaces* of these ill-conditioned eigenstructures. Another example of the use of learning theory in cryptosystems, albeit within a different context, is described in [5].

[1] This property is the *raison d'etré* for the phrase "Fisher-Schrödinger model"

[2] The terms covert information and code are used interchangeably.

1.1 The TIS-lE

Consider a measured random variable $y = (y_1, ..., y_N)$, parameterized by $\theta = (\theta_1, ..., \theta_N)$ (the "true" value). A *fluctuation*, i.e. a random variable $x = (x_1, ..., x_N)$, defined by $x = y - \theta$ is introduced. For translational (or shift) invariant families of distributions, $p(y|\theta) = p(y-\theta) = p(x)$. The particular form of the FI that is chosen is the trace of the FI matrix for independent and identically distributed (*iid*, hereafter) data [3]. This is referred to as the Fisher channel capacity (FCC) [2]. The FCC is: $I^{FCC} = \int dyp(y|\theta)\left(\frac{\partial \ln p(y|\theta)}{\partial\theta}\right)^2 = \int dxp(x)\left(\frac{\partial \ln p(x)}{\partial x}\right)^2$ under translational invariance [2,6]. The probability amplitude (wave function) relates to the *pdf* as $\psi(x) = \sqrt{p(x)}$. The FCC acquires the compact form $I^{FCC} = 4\sum_{n=1}^{N}\int dx_n\left(\frac{\partial\psi(x_n)}{\partial x_n}\right)^2 = 4\int dx\left(\frac{\partial\psi(x)}{\partial x}\right)^2$. The form of the FCC is essential to the formulation of a variational principle. Within the framework of a measurement-response model, this implies that the observer who initiates the measurements, collects the response in the form of *iid* data. In many practical scenarios, the response of a system to measurements is not obliged to be *iid*. The presence of correlations contribute to off-diagonal elements in the FI matrix formed by the observer. These correlations may be mitigated, thereby eliminating the off-diagonal elements of the FI matrix, by performing ICA (or an equivalent procedure) as a pre-processing stage.

Consider a Lagrangian of the form

$$L^{FCC} = 4\int dx\left(\frac{\partial\psi(x)}{\partial x}\right)^2 + \underbrace{+\int dx\sum_{i=1}^{M}\lambda_i\Theta_i(x)\psi^2(x) - \lambda_o\int dx\psi^2(x)}_{-J[x]}, \quad (2)$$

$M < N$ (incomplete constraints), where the Lagrange multiplier (LM) λ_o corresponds to the probability density function (*pdf*, hereafter) normalization condition $\int dx\psi^2(x) = 1$. The LM's $\lambda_i; i = 1, ..., M$ correspond to actual (physical) constraints of the form $\int dx\Theta_i(x)\psi^2(x) = \langle\Theta_i(x)\rangle = d_i$. Here, $\Theta_i(x)$ are operators, and, d_i are the constraints (physical observables). This work considers constraints of the geometric moment type: $\Theta_i(x) = x^i; i = 0, ..., M$. Here, (2) resembles the usual MaxEnt Lagrangian with the FCC replacing the Shannon entropy. In (2), the FCC is ascribed the role akin to the kinetic energy. The constraint terms manifest the potential energy. Variational extremization of (2) yields the minimum Fisher information (MFI) principle [7]

$$\underbrace{-\frac{d^2\psi(x)}{dx^2} + \frac{1}{4}\sum_{i=1}^{M}\lambda_i\Theta_i(x)\psi(x)}_{H^{FI}\psi(x)} = \frac{\lambda_o}{4}\psi(x), \quad (3)$$

where H^{FI} is an empirical Hamiltonian operator. Here, (3) is referred to as a TIS-lE. Note that the probability amplitudes are taken as being real quantities. This assumption is tenable since the model is spatially one dimensional in the continuum, and is time independent. Comparing the TIS-lE with the TISE immediately reveals that the constraint terms $V(x) = \frac{1}{4}\sum_{i=1}^{M}\lambda_i\Theta_i(x) = \sum_{i=1}^{M}\tilde{\lambda}_i x^i$, constitute an empirical pseudo-potential. The normalization LM and the total energy eigenvalue relate as $\lambda_o = 4E$. Further, the constants in the TISE relate as $\hbar^2/2m = 1$.

Solution of the TISE as an eigenvalue problem yields a number of *energy states* characterized by distinct values of E. These comprise the equilibrium state (*zero-energy state*) characterized by a Maxwellian distribution, and, higher energy *excited states* (non-equilibrium states). The wave functions are a superposition of Hermite-Gauss solutions. *By virtue of its similarity to the TISE, the TIS-lE "inherits" these energy states within an information theoretic context*. This feature permits the projection of covert information into *multiple energy states* of the TIS-lE, for an empirical pseudo-potential that approximates a TISE physical potential.

Employing the TIS-lE to infer *pdf*'s from incomplete constraints, requires an accurate evaluation of the LM's. This is accomplished in this paper through a semi-supervised learning paradigm, that iteratively couples the solution of (3) with the minimization of a Lagrangian that manifests a measurement-response model.

This procedure represents, in certain aspects, an extension of the optimization procedure employed to achieve *quantum clustering* using the TISE [8]. In the case of *quantum clustering*, the TISE probability amplitude/wave function is approximated by a non-parametric estimator (Parzen windows), and, the potential $V(x)$ is determined via a steepest descent in Hilbert space. In contrast, the semi-supervised paradigm presented in this paper achieves reconstruction of *pdf*'s (the inverse problem of statistics) without any *a-priori* strictures placed on the probability amplitudes of (3). The Fisher-Schrödinger model has been employed within a statistical setting in a number of studies ranging from quantum statistics to fuzzy clustering [9]. Within the context of securing covert information, the above features endow the statistical enryption/decryption strategy with a fundamental physical connotation.

[3]The FI matrix for *iid* data has vanishing off-diagonal elements (e.g., Appendix B in [2])

1.2 The Dirac notation

This paper utilizes the Dirac *bra-ket* notation [10] to describe linear algebraic operations in a compact form. By definition, a *ket* $|\bullet\rangle$ denotes a column vector, and, a *bra* $\langle\bullet|$ denotes a row vector. The scalar inner product and the projection operators are described by $\langle\bullet \mid \bullet\rangle$, and the outer product $|\bullet\rangle\langle\bullet|$, respectively. The expectation evaluated at the ε^{th} *energy state* is $\langle\bullet\rangle_{\varepsilon}$.

2 Semi-supervised Learning Paradigm

2.1 Theory

The task of density estimation involves the iterative determination of the LM's and probability amplitudes of the TIS-lE (3). In the MaxEnt and MFI theories, the observer is external to the system. The present work reconstructs the host *pdf*'s using a semi-supervised learning paradigm, by incorporating a *participatory observer*. This is accomplished by positioning the *participatory observer* in a *measurement space* characterized by the amplitude $\psi^{\varepsilon}(x)$, performing unbiased measurements [2, 11] on a given physical system (data).

The *system space*, inhabited by the physical system subject to measurements, is characterized by an amplitude $\phi^{\varepsilon}(\tilde{x})$. Here, x and $\tilde{x}$ are the conjugate basis coordinates of the *measurement space* and *system space*, respectively. Herein, the mutually conjugate spaces are taken to be the Cartesian coordinate and the linear momentum. Setting $\tilde{x}=\mu$, the commutation relation is $[x,\mu]=i\hbar$, respectively [1]. The group for the basis change is $G=-i\hbar\frac{d}{dx}$. The corresponding Hermitian unitary operator is $U=e^{-ai\hbar\frac{d}{dx}}$, where a is the group parameter of infinitesimal transformations. Within the present scenario, $U\psi^{\varepsilon}[x]=\varphi^{\varepsilon}[\mu]$. On the basis of the above discussions, it is easily proven that $I_{\varepsilon}^{FCC}[x]=4\int dx\left(\frac{d\psi(x)}{dx}\right)^2=\frac{4}{\hbar^2}\left\langle\mu^2\right\rangle_{\varepsilon}=4\left\langle\frac{\mu^2}{2m}\right\rangle_{\varepsilon}=I_{\varepsilon}^{FCC}[\mu]\,;\hbar^2/_{2m}=1$. In the non-relativistic limit, the kinetic energy of a particle is $T=\frac{\mu^2}{2m}$ [1]. For TIS-lE polynomial pseudo-potentials of the form $V(x)=\sum_{i=1}^{M}\lambda_i^{\varepsilon}x^i$, the quantum mechanical virial theorem [9, 12, 13] yields

$$I_{\varepsilon}^{FCC}=2\left\langle x\frac{dV(x)}{dx}\right\rangle_{\varepsilon}=2\sum_{i=1}^{M}i\lambda_i^{\varepsilon}d_i^{\varepsilon} \qquad (4)$$

The unitary relation between the amplitudes in conjugate spaces results in the potential energy term in (2), $J_{\varepsilon}[x]$*, being manifested as an empirical representation of the FCC*. Specifically, $I_{\varepsilon}^{FCC}[x]=J_{\varepsilon}[x]$. Each measurement (or set of measurements) initiated by the observer at a specific juncture, perturbs the amplitude of the *system space* as $\delta\varphi^{\varepsilon}(\mu)$. For mutually conjugate spaces related by a unitary transform, this results in a perturbation $\delta\psi^{\varepsilon}(x)=\delta\varphi^{\varepsilon}(\mu)$ of the *measurement space*. It is at this juncture that the observer constructs the FCC for *iid* data. Consequently, $\delta I_{\varepsilon}^{FCC}[x]=\delta J_{\varepsilon}[x]$ [2]. Such models are known as measurement-response models [14].

Incorporation of a *participatory observer* results in a *zero-sum game* [15] of information acquisition played between the observer and the system under observation. The observer seeks to maximize her/his information about the system. Simultaneously, the *system space* is inhabited by a *demon*, remniscent to the Maxwell *demon*, who seeks to minimize this information transfer. This *zero-sum game* between the observer and the *demon* is hereafter referred to as the *Fisher game*.

Game theoretic studies in MaxEnt and MFI follow the traditional pattern of having the *arbiter*, who assigns strategies to the players, residing external to the system. *In this paper, the probe measurements initiated by the observer constitute the arbiter, and, the probability amplitudes manifest the strategies.* A future publication treats the game theoretic aspects of the semi-supervised learning paradigm, within the ambit of the *bounded rationality theory* [16].

The incomplete constraints $d_i^{\varepsilon}, i=1,\ldots,M$ are evaluated as the moments of the Cartesian coordinates at each *energy state* ε, by solving the TISE (1) as an eigenvalue problem on a lattice, for an *a-priori* specified physical potential. *The incomplete constraints represent the only manifestation of the target values of the amplitudes/pdf's, made available to the designer at the commencement of the inference procedure*.

The host *pdf* inference is solved by an iterative optimization process *In this paper, the host pdf is independently inferred for each energy level* ε. The optimization process couples the solution of the TIS-lE (3), with the steepest descent minimization of an empirical quantity known as the *residue*. The *residue* represents the discrepancy between the value of $J_{\varepsilon}[x]$ in (2) evaluated at an intermediate *iteration level* l for a specific *energy state*, and, the value of the exact (target) FCC expressed at the same *iteration level* l.

In (4), the LM's λ_i^{ε} are target values of the optimization process, which are unknown at the commencement of the inference procedure. Here, (4) is made consistent with the iteration process by specifying the relation between the target values of the LM's, and, the LM's at some intermediate iteration level l as

$$2\sum_{i=1}^{M}i\lambda_i^{\varepsilon}\cong\sum_{i=1}^{M}\lambda_{i,l}^{\varepsilon}\left\langle\psi_l^{\varepsilon}\mid\psi_l^{\varepsilon}\right\rangle_{\varepsilon} \qquad (5)$$

Here, (5) is critical to the optimization process since it infuses a representation of the target response state into the iterative procedure. The final values of the expectation of the amplitudes satisfy $\left\langle \psi^{\varepsilon}_{l=final} \middle| \psi^{\varepsilon}_{l=final} \right\rangle_{\varepsilon} = 1$. Note that the expectation $\langle \psi^{\varepsilon}_{l} \mid \psi^{\varepsilon}_{l} \rangle_{\varepsilon}$ is not assumed to be unity. Combining (4) and (5) allows the target value of the FCC to be manifested at some intermediate iteration level l. At the l^{th} iteration level, the term $J_{\varepsilon}[x]$ in (2) is

$$J\left(\lambda^{\varepsilon}_{i,l}\right) = \lambda^{\varepsilon}_{o,l} \langle \psi^{\varepsilon}_{l} \mid \psi^{\varepsilon}_{l} \rangle_{\varepsilon} - \sum_{i=1}^{M} \lambda^{\varepsilon}_{i,l} \langle \psi^{\varepsilon}_{l} | \Theta(x) | \psi^{\varepsilon}_{l} \rangle_{\varepsilon} \tag{6}$$

The *residue* at the l^{th} iteration level, re-scaled with respect to $\langle \psi^{\varepsilon}_{l} \mid \psi^{\varepsilon}_{l} \rangle_{\varepsilon}$, is

$$\begin{aligned} &\frac{R_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}\right)}{\langle \psi^{\varepsilon}_{l} | \psi^{\varepsilon}_{l} \rangle_{\varepsilon}} = \tilde{R}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}\right) \\ &\cong -\lambda^{\varepsilon}_{o,l} + \sum_{i=1}^{M} \lambda^{\varepsilon}_{i,l} \left\{ \frac{\langle \psi^{\varepsilon}_{l} | \Theta(x) | \psi^{\varepsilon}_{l} \rangle_{\varepsilon}}{\langle \psi^{\varepsilon}_{l} | \psi^{\varepsilon}_{l} \rangle_{\varepsilon}} + d^{\varepsilon}_{i} \right\} \end{aligned} \tag{7}$$

Here, (7) is $\tilde{R}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}\right) \cong \tilde{I}^{FCC}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}, d^{\varepsilon}_{i}\right) - \Im_{\varepsilon}\left(\lambda^{\varepsilon}_{o,l}, \lambda^{\varepsilon}_{i,l}\right); i = 1, ..., M$, where the re-scaled target FCC and the *potential energy* are $\tilde{I}^{FCC}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}, d^{\varepsilon}_{i}\right)$, and, $\Im_{\varepsilon}\left(\lambda^{\varepsilon}_{o,l}, \lambda^{\varepsilon}_{i,l}\right)$, respectively. A steepest descent procedure $\frac{\partial \tilde{R}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}\right)}{\partial \lambda^{\varepsilon}_{i,l}} \to 0$ along the gradient of the LM's yields "optimal" values of the LM's

$$\frac{\partial \left[\tilde{I}^{FCC}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}, d^{\varepsilon}_{i}\right) - \Im_{\varepsilon}\left(\lambda^{\varepsilon}_{o,l}, \lambda^{\varepsilon}_{i,l}\right)\right]}{\partial \lambda^{\varepsilon}_{i,l}} \to 0; i = 1, ..., M \tag{8}$$

The optimization procedure is carried out till the target values are achieved. The steepest descent procedure requires the analytical values of $\frac{\partial \tilde{R}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}\right)}{\partial \lambda^{\varepsilon}_{i,l}}$, and thus, $\frac{\partial \lambda^{\varepsilon}_{o,l}}{\partial \lambda^{\varepsilon}_{i,l}}; i = 1, ..., M$. Left multiplying (3) for the *energy state* ε and iteration level l by ψ^{ε}_{l} and integrating, yields $\frac{\partial \lambda^{\varepsilon}_{o,l}}{\partial \lambda^{\varepsilon}_{i,l}} = \frac{d^{\varepsilon}_{i,l}}{\langle \psi^{\varepsilon}_{l} | \psi^{\varepsilon}_{l} \rangle_{\varepsilon}}; i = 1, ..., M$. The theory of the semi-supervised learning paradigm based on the *Fisher game* is summarized by the pseudo-code in Algorithm 1.

2.2 Physical interpretations

The *Fisher game* constitutes a self-consistent information theoretic optimization procedure, with a quantum mechanical connotation. The above theory contains three interesting observations. First, the commencement of each *iteration loop* corresponds to the juncture at which the observer initiates measurements. Next, as the iterative process advances, the FCC approaches a minimum. This corresponds to an increase in the uncertainty at the location of the observer.

Finally, at the termination of the l^{th} iteration loop, the condition (8) that yields the "optimal" LM's is the statement of a *contract* between the *demon* and the observer, whereby, the *demon* makes the last move in the iterative Fisher game. This implies that the *participatory observer* acquires a state of maximum uncertainty (minimum Fisher information). Such a *contract* is the underlying basis for determining the "optimal" LM's, corresponding to amplitudes that decrease the FCC at the termination of each iteration level.

Scenarios of such type cannot be modeled within the framework of traditional game theory [17], thus, justifying the use of the *bounded rationality theory* to study the game theoretic aspects of the *Fisher game*. A future publication studies the information landscape and its relation to the *Fisher game*.

Algorithm 1 Inverse Problem of Statistics

PROCEDURE FOR EACH ENERGY STATE ε
INITIALIZATION
1.Solve TISE (1) for known physical potential $V(x)$ as an eigenvalue problem in the event space $[a, b]$. Obtain incomplete constraints $d^{\varepsilon}_{i}; i = 1, ..., M$.
2. Solve TIS-lE (3) for arbitrary Lagrange multipliers $\lambda^{\varepsilon}_{i,arbitrary}; i = 1, ..., M$. Obtain amplitudes $\psi^{\varepsilon}_{i,l=1}$.
3. Input tolerance parameter $\delta_i; i = 1, ..., M$
ALGORITHM FOR $l^{th} \geq 1$ **ITERATION LOOP**
1. Obtain ψ^{ε}_{l} by solving the TIS-lE (3) as an eigenvalue problem, using the optimized LM's $\lambda^{\varepsilon}_{i,l-1}$ obtained from the $(l-1)^{th}$ iteration loop.
2. Minimize the re-scaled *residue* $\tilde{R}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}\right)$ (7), to obtain optimized LM's that correspond to ψ^{ε}_{l}
$\lambda^{\varepsilon}_{i,l} \leftarrow \frac{\partial \left[\tilde{I}^{FCC}_{\varepsilon}\left(\lambda^{\varepsilon}_{i,l}, d^{\varepsilon}_{i}\right) - \Im_{\varepsilon}\left(\lambda^{\varepsilon}_{o,l}, \lambda^{\varepsilon}_{i,l}\right)\right]}{\partial \lambda^{\varepsilon}_{i,l}}$
3. Solve TIS-lE (3) as an eigenvalue problem with LM's $\lambda^{\varepsilon}_{i,l}$. Obtain moments $d^{\varepsilon}_{i,l}; i = 1, ..., M$.
IF
$\left| d^{\varepsilon}_{i} - d^{\varepsilon}_{i,l} \right| > \delta_i; i = 1, ..., M$
$l = l + 1$
RETURN
ELSE IF
$\left| d^{\varepsilon}_{i} - d^{\varepsilon}_{i,l} \right| \leq \delta_i; i = 1, ..., M$
END IF

2.3 Numerical results

The asymmetric harmonic oscillator (AHO) potential $V(x) = x + \frac{x^2}{2}$ is chosen as TISE physical potential. The TISE with the AHO potential is solved as an eigenvalue problem for 201 data points within the event space $[a = -1, b = 1]$ for different *energy states*. Boundary conditions on the amplitudes, $\psi^{\varepsilon}(a) = \psi^{\varepsilon}(b) = 0$, are enforced. An empirical pseudo-potential of the form $V_{\varepsilon}(x) = \tilde{\lambda}^{\varepsilon}_{1} x + \tilde{\lambda}^{\varepsilon}_{2}$, that approximates the TISE AHO physical potential is specified. Here, $\tilde{\lambda}_i^\varepsilon = \frac{\lambda_i^\varepsilon}{4}$. In this case, $M = 2, N = 201$.

The values of the incomplete constraints are $d_1^{\varepsilon=0,1} = (-0.0344, 0.0097,), and, d_2^{\varepsilon=0,1} = (0.1302, 0.2809)$. The final values of the LM's are $\tilde{\lambda}_1^{\varepsilon=0,1} = (1.0892, 1.1121)$, and, $\tilde{\lambda}_2^{\varepsilon=0,1} = (0.6635, 0.7175)$, respectively. The inferred total energy eigenvalues are $E_{\text{inf}}^{\varepsilon=0,1} = (2.5333, 10.0769)$. The corresponding TISE total energy eigenvalues are $E_{exact}^{\varepsilon=0,1} = (2.5152, 10.0147)$. Fig. 1 depicts the inferred *pdf*'s overlaid upon the TISE solution. Here, the Maxwellian distribution corresponds to $\varepsilon = 0$, and, double peaked *pdf* corresponds to the first *excited state* $\varepsilon = 1$. Note that the inferred *pdf*'s almost exactly coincide with the TISE solutions.

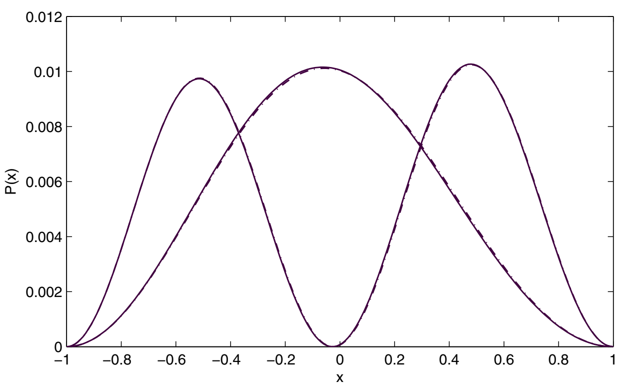

Figure 1. Inferred TIS-IE *pdf* (dash-dots) and exact TISE *pdf* (solid lines)

3 Projection Strategy

Consider M constraints $d_1^\varepsilon, ..., d_M^\varepsilon$. In a discrete setting, these are expectation values of a random variable $x_{i,n}; n = 1, ..., N$:

$$d_i^\varepsilon = \sum_{n=1}^{N} p_n^\varepsilon x_{i,n}; i = 1, ..., M. \tag{9}$$

The *pdf* $|p^\varepsilon\rangle \in \Re^N$ is a *ket*, where $|n\rangle ; n = 1, ..., N$ is the standard basis in $\Re^N$, is expressed as $|p^\varepsilon\rangle = \sum_{n=1}^{N} |n\rangle \langle n \mid p^\varepsilon\rangle = \sum_{n=1}^{N} p_n^\varepsilon |n\rangle$. The *ket* of observable's is expressed as $|d^\varepsilon\rangle \in \Re^{M+1}$ with components $d_1^\varepsilon, ..., d_M^\varepsilon, 1$, and, an operator $A : \Re^N \to \Re^{M+1}$ given by $A = \sum_{n=1}^{N} |x_n\rangle \langle n|$. Defining vectors $|x_n\rangle \in \Re^{M+1}; n = 1, ..., N$ as the expansion $|x_n\rangle = \sum_{i=1}^{M+1} |i\rangle \langle i \mid x_n\rangle = \sum_{i=1}^{M+1} x_{i,n} |i\rangle$, where i is a basis vector in $\Re^{M+1}$, (9) acquires the compact form

$$|d^\varepsilon\rangle = A |p^\varepsilon\rangle ; A : \Re^N \to \Re^{M+1}. \tag{10}$$

The physical significance of the constraint operator A in (10) is as follows. Inference of the *pdf* and the TIS-IE pseudo-potential in (3) from physical observables is achieved by specifying $V(x) = \sum_{i=1}^{M} \lambda_i x^i$. In a discrete setting, $x_n^i \to x_{i,n}; i = 1, ..., M; n = 1, ..., N$. The $x_{i,n}$ constitute the elements of the M rows and N columns of the operator A, and represent the spatial elements of the TIS-IE pseudo-potential in matrix form. The unity element in $|d^\varepsilon\rangle \in \Re^{M+1}$ enforces the normalization constraint of the probability density $|p^\varepsilon\rangle$.

The operator A is independent of the host pdf, and thus, the energy state. This may be mitigated by defining

$$\tilde{A}^\varepsilon = A + k^\varepsilon |d^\varepsilon\rangle \langle I| \tag{11}$$

Here, $\langle I|$ is a $1 \times N$ *bra*, and, $k^\varepsilon \neq -1$ is a constant parameter introduced to adjust the condition number of $\tilde{A}^\varepsilon$, and hence its sensitivity to perturbations. In (11), dependence upon the host *pdf* is "injected" into the operator $\tilde{A}^\varepsilon$ by the incorporation of $|d^\varepsilon\rangle$. Specifically, each element of the *ket* $|d^\varepsilon\rangle$ is defined by $\sum_{n=1}^{N} p_n^\varepsilon x_n^i; i = 1, ..., M$.

Thus, (10) becomes $\tilde{A}^\varepsilon |p^\varepsilon\rangle = |d^\varepsilon\rangle + \langle k^\varepsilon |d^\varepsilon\rangle \langle I|\rangle_\varepsilon$. Expanding $\langle k |d^\varepsilon\rangle \langle I|\rangle_\varepsilon = k |d^\varepsilon\rangle \langle I| \ p^\varepsilon\rangle$, and evoking the *pdf* normalization, $\langle I| \ p^\varepsilon\rangle = 1$, yields

$$\left|\tilde{d^\varepsilon}\right\rangle = (k^\varepsilon + 1) |d^\varepsilon\rangle = \tilde{A}^\varepsilon |p^\varepsilon\rangle ; \tilde{A}^\varepsilon : \Re^N \to \Re^{M+1}. \tag{12}$$

The operator $\tilde{A}^\varepsilon$ is ill-conditioned and rectangular. Thus, (12) becomes:

$$|p^\varepsilon\rangle = \left(\tilde{A}^\varepsilon\right)^{-1} \left|\tilde{d^\varepsilon}\right\rangle + \left|p^{\varepsilon'}\right\rangle, \tag{13}$$

where, $\left(\tilde{A}^\varepsilon\right)^{-1}$ is the pseudo-inverse [18] of $\tilde{A}^\varepsilon$, and lies in $range\left(\tilde{A}^\varepsilon\right)$. All necessary data dependent information resides in $\left(\tilde{A}^\varepsilon\right)^{-1} \left|\tilde{d^\varepsilon}\right\rangle$.

The *null space* term in (13) is of particular importance since the code is embedded into it via unitary projections. Here, $\left|p^{\varepsilon'}\right\rangle \in null\left(\tilde{A}^\varepsilon\right)$ is explicitly data independent. However, it is critically dependent on the solution methodology employed to solve (13). The operator $G^\varepsilon = \tilde{A}^{\varepsilon\dagger}\tilde{A}^\varepsilon$ is introduced. Here, $\tilde{A}^{\varepsilon\dagger}$ is the conjugate transpose of $\tilde{A}^\varepsilon$. Projection of the covert information into $null\left(G^\varepsilon\right)$ instead of $null\left(\tilde{A}^\varepsilon\right)$, leads to increased instability of the eigenstructure, which is exploited to increase the security of the covert information [3, 9].

Given the operator $\tilde{A}^{\varepsilon}$ and the probability vector $|p^{\varepsilon}\rangle$, whose inference is described in Section 2, the normalized eigenvectors corresponding to the eigenvalues in the *null space* of G^{ε} having value zero (*zero eigenvalues*) are defined as $|\eta_n^{\varepsilon}\rangle\,; n = 1, ..., N-(M+1)$. Here, $|\eta_n^{\varepsilon}\rangle$, defined as the basis of $null\,(G^{\varepsilon})$, are evaluated using SVD [18]. To introduce cryptographic keys (cryptographic primitives), an operator $\tilde{G}^{\varepsilon}$ is formed by perturbing select elements of G^{ε} by $\delta G_{i,j}^{\varepsilon}$. Here, $\delta G_{i,j}^{\varepsilon}$ is a perturbation to the element inhabiting the i^{th} row and j^{th} column of the operator G^{ε}. In *symmetric* cryptography, only a single element of G^{ε} is perturbed. The security of the code may be ensured by adopting an *asymmetric* cryptographic strategy. Here, more than one element of G^{ε} is perturbed.

The extreme sensitivity to perturbations of G^{ε} causes the eigenstructure of $\tilde{G}^{\varepsilon} = G^{\varepsilon} + \delta G_{i,j}^{\varepsilon}$ to substantially differ from that of G^{ε}, even for infinitesimal perturbations. The values of $|\tilde{\eta}_n^{\varepsilon}\rangle\,; n = 1, ..., N-(M+1)$, the basis of $null(\tilde{G}^{\varepsilon})$, are evaluated using SVD. The unitary operators of decryption U_{dec}^{ε} (without perturbations) and $\tilde{U}_{dec}^{\varepsilon}$ (with perturbations), $U_{dec}^{\varepsilon}, \tilde{U}_{dec}^{\varepsilon} : \Re^N \rightarrow \Re^{N-(M+1)}$, and, the corresponding encryption operators $U_{enc}^{\varepsilon}, \tilde{U}_{enc}^{\varepsilon} : \Re^{N-(M+1)} \rightarrow \Re^N$ for the *energy state* ε are

$$\begin{aligned} U_{dec}^{\varepsilon} &= \sum_{n=1}^{N-M-1} |n\rangle \langle \eta_n^{\varepsilon}|; \\ \tilde{U}_{dec}^{\varepsilon} &= \sum_{n=1}^{N-M-1} |n\rangle \langle \tilde{\eta}_n^{\varepsilon}|, \\ &and, \\ U_{enc}^{\varepsilon} &= U_{dec}^{\varepsilon\dagger} = \sum_{n=1}^{N-M-1} |\eta_n^{\varepsilon}\rangle \langle n|; \\ \tilde{U}_{enc}^{\varepsilon} &= \tilde{U}_{dec}^{\varepsilon\dagger} = \sum_{n=1}^{N-M-1} |\tilde{\eta}_n^{\varepsilon}\rangle \langle n|, \end{aligned} \tag{14}$$

respectively.

3.1 Encryption

Given a code $|q^{\varepsilon}\rangle \in \Re^{N-(M+1)}$ to be encrypted in the energy state ε , the $N-(M+1)$ components are given by $\langle n \mid q^{\varepsilon}\rangle = q_n^{\varepsilon}; n = 1, ..., N-(M+1)$. The *pdf* of the embedded code is:

$$|p_c^{\varepsilon}\rangle = \tilde{U}_{enc}^{\varepsilon}\,|q^{\varepsilon}\rangle = \sum_{n=1}^{N-M-1} |\tilde{\eta}_n^{\varepsilon}\rangle \langle n \mid q^{\varepsilon}\rangle. \tag{15}$$

The total *pdf* comprising the host *pdf* and the *pdf* of the code is

$$|\tilde{p}^{\varepsilon}\rangle = |p^{\varepsilon}\rangle + |p_c^{\varepsilon}\rangle\,. \tag{16}$$

Note that since $|p_c^{\varepsilon}\rangle \in null\left(\tilde{A}^{\varepsilon}\right)$, $\tilde{A}^{\varepsilon}\,|p_c^{\varepsilon}\rangle = 0$.

3.2 Transmission

Information may be transferred from the encrypter to the decrypter in two separate manners , via a *public channel*. The first mode is to transmit the constraint operators $\tilde{A}^{\varepsilon}$ and the total *pdf*'s $|\tilde{p}^{\varepsilon}\rangle$. An alternate mode is to transmit the LM's obtained on solving the *Fisher game*(Section 2), and, the total *pdf*'s $|\tilde{p}^{\varepsilon}\rangle$. Owing to the large dimensions of the constraint operators $\tilde{A}^{\varepsilon}$, the latter transmission strategy is more attractive.

The values of parameters k^{ε} for each *energy state*, and, the cryptography key/keys are transmitted through a *secure/covert channel*. The cryptographic primitives are labeled in order to identify the elements of the operator G^{ε} that are perturbed. In the case of *asymmetric* cryptography, some of the keys may be declared public, while keeping the remainder private [4]. *Asymmetric* cryptography provides greatly enhanced security to the covert information, and, provides protection against attacks, such as *plaintext attacks* [3, 4, 9].

3.3 Decryption

The decrypter and encrypter have an *a-priori* "agreement" concerning the nature of TIS-lE pseudo-potential, and, the number of *energy states*. The legitimate receiver recovers the key/keys $\delta G_{i,j}^{\varepsilon}$ and the parameter k^{ε} from the *covert channel*. The operators $\tilde{A}^{\varepsilon}$, G^{ε} , and, $\tilde{G}^{\varepsilon}$ are constructed. The host *pdf* may be recovered in two distinct manners, depending upon the transmission strategy employed. Note that both methods to reconstruct the host *pdf* require the total *pdf* $\tilde{p}^{\varepsilon}$ to be provided by the encrypter. First, the scaled incomplete constraints, defined in (12), are obtained by solving $\langle i|\,\tilde{A}^{\varepsilon}\,|\tilde{p}^{\varepsilon}\rangle = \left|\tilde{d}^{\varepsilon}\right\rangle$. Here, i is a basis vector in $\Re^{M+1}$. This procedure is possible because $|p_c^{\varepsilon}\rangle \in null\left(\tilde{A}^{\varepsilon}\right)$. Thus, $\tilde{A}^{\varepsilon}\,|p_c^{\varepsilon}\rangle = 0$. The host *pdf* are then computed for each *energy state* by solving the *Fisher game*, using the re-scaled set of incomplete constraints. Alternatively, the host *pdf* may be obtained by solving the TIS-lE (3) as an eigenvalue problem, given the values of the LM's $\lambda_i^{\varepsilon}\,; i = 1, ..., M$, and, the event space (Section 2). Both methods allow the reconstructed host *pdf*'s to be obtained with a high degree of precision. The code *pdf* is recovered using

$$|p_{rc}^{\varepsilon}\rangle = \left|\tilde{p}^{\varepsilon}\right\rangle - |p^{\varepsilon}\rangle\,. \tag{17}$$

The encrypted code is recovered by the operation

$$|q_r^{\varepsilon}\rangle = \tilde{U}_{dec}^{\varepsilon}\,|p_{rc}^{\varepsilon}\rangle = \sum_{n=1}^{N-M-1} |n\rangle \langle \tilde{\eta}_n^{\varepsilon} \mid p_{rc}^{\varepsilon}\rangle \tag{18}$$

The thresholds for the cryptographic keys is accomplished by the designer, who performs a simultaneous encryption/decryption without effecting perturbations to the operator G^{ε}. The host *pdf*'s are inferred from the *Fisher game*. The code $|q^{\varepsilon}\rangle$ having dimension $N-(M+1)$ is formed. The designer implements (15)-(18) for each *energy state* ε. The threshold for the cryptographic key/keys is $\delta^{\varepsilon}=\||q^{\varepsilon}\rangle-|q_r^{\varepsilon}\rangle\|$. Hardware independence is demonstrated by performing the encryption on an IBM RS-6000 workstation cluster, and, decryption on an IBM Thinkpad running MATLAB v 7.01. The encryption/decryption strategy is critically dependent upon the exact compatibility of the routines to calculate the basis $|\tilde{\eta}_n^{\varepsilon}\rangle$ and the eigenvalue solvers, available to the encrypter and decrypter.

4 Numerical Examples

The encryption/decryption strategy is tested using the *energy state* dependent model vis-á-vis an *energy state* independent model [9], for the case of *asymmetric* cryptography. These are characterized by the constraint operators $\tilde{A}^{\varepsilon}$ (described in (11) and (12)) for $k^{\varepsilon=0,1}=-0.1$, and, A (described in (10)), respectively. The *energy state* independent model corresponds to the ground state Maxwellian distribution.

A random number generator generates code in $[0,1]$. Two identical *kets* of the code having dimension $N-(M+1)=198$ are created for projection into the *null spaces* of the *energy state* dependent operators $\tilde{G}^{\varepsilon=0,1}$, respectively. This "emulates" the selective projection of a code comprising of a single *ket* of dimension 396, into the two *energy state* of $null(\tilde{G}^{\varepsilon}$. For the *energy state* independent operator $null(\tilde{G})$, only a single *ket* is projected. The cryptographic primitives are $\delta G_{1,3}=\delta\tilde{G}_{1,3}^{\varepsilon=0,1}=3.0e-013$ and $\delta G_{2,7}=\delta\tilde{G}_{2,7}^{\varepsilon=0,1}=7.0e-013$, respectively. All numerical examples in have a threshold for perturbations $\delta^{\varepsilon}\sim 2.0e-014$. The condition numbers, $cond(\bullet)$ of the constraint operators A and $\tilde{A}^{\varepsilon}$ provides a measure of the sensitivity to perturbations of the operators $G=A^{\dagger}A$ and G^{ε}.

Values of $cond(A)$, $cond(\tilde{A}^{\varepsilon=0})$, and, $cond(\tilde{A}^{\varepsilon=1})$ are 3.73048, 3.41742, and, 3.37888 , respectively. Going by conventional logic, the *energy state* independent model should afford greater security to the covert information, owing to the greater value of $cond(A)$, vis-á-vis $cond(\tilde{A}^{\varepsilon=0,1})$. Numerical simulations reveal a dichotomy in this regard.

A more relevant metric of the extreme sensitivity of $null(G^{\varepsilon})$ to perturbations, induced by the cryptographic keys $\delta G_{i,j}^{\varepsilon}$, is the distortion of the code *pdf* $\left|p_{c,unpert}^{\varepsilon}\right\rangle$. Here, $\left|p_{c,unpert}^{\varepsilon}\right\rangle$ is evaluated from (15), using η_n^{ε} (the unperturbed basis of $null(G^{\varepsilon})$). The distorted code *pdf* is $|\tilde{p}_c^{\varepsilon}\rangle$, which is calculated from (15) using $\tilde{\eta}_n^{\varepsilon}$ (the perturbed basis of $null\left(\tilde{G}^{\varepsilon}\right)$), as described in Section 3.1.

For the *energy state* dependent model, the *RMS error of encryption*is defined as: $RMS_{enc}^{\varepsilon}=\frac{\|err_{enc}^{\varepsilon}\|}{\sqrt{length(err_{enc}^{\varepsilon})}}$. Here, $\||err_{enc}^{\varepsilon}\rangle\|=\left\|\left(|p_c^{\varepsilon}\rangle-\left|p_{c,unpert}^{\varepsilon}\right\rangle\right)\right\|$, and, $length\left(|err_{enc}^{\varepsilon}\rangle\right)$ is the dimension of $\left(|p_c^{\varepsilon}\rangle-\left|p_{c,unpert}^{\varepsilon}\right\rangle\right)$.

A further quantitative metric of the degree of security of the encrypted code is the RMS error of reconstruction between the embedded code and the code reconstructed without the keys. For the *energy state* dependent model, this is: $RMS_{recon}^{\varepsilon}=\||err_{recon}^{\varepsilon}\rangle\|/length\left(|err_{recon}^{\varepsilon}\rangle\right)$; $\||err_{recon}^{\varepsilon}\rangle\|=\||q\rangle-|q_{r1}\rangle\|$. Here, $RMS_{recon}^{\varepsilon}$ provides a measure of the error of recovery of the code by an *unauthorized eavesdropper* who does not possess the keys $\delta G_{i,j}^{\varepsilon}$, but, possesses the total *pdf* and the code *pdf*. Such attackers are known as *semi-honest adversaries*, since no attempt is made to distort the information transmitted via the public channel. In this case, the reconstructed code becomes $|q_{r1}^{\varepsilon}\rangle=\sum_{n=1}^{N-M-1}|n\rangle\langle\eta_n^{\varepsilon}\mid p_{rc}^{\varepsilon}\rangle$.

For the *energy state* independent model, the values of $RMS_{enc}=0.79869$ and $RMS_{recon}=0.81302$. The corresponding values for the *energy state* dependent model are $RMS_{enc}^{\varepsilon=0,1}=(0.83878,0.89651)$, and, $RMS_{recon}^{\varepsilon=0,1}=0.84596,0.88931$, respectively. The higher values of the RMS_{enc}^{ε} for the *energy state* dependent model explains the vastly enhanced degree of security it provides by demonstrating a greater value of $RMS_{recon}^{\varepsilon}$, despite the value of $cond(\tilde{A}^{\varepsilon})$ being less than $cond(A)$. Simulations results for select values for the case of the *energy state* independent and dependent models are described in Table 1 and Table 2, respectively. The reconstructed code *with* the keys is exactly similar to the original code. *On the other hand, the code reconstructed without the keys bears no resemblance to the original code*. The highly oscillatory nature of the total *pdf* (16) for the *energy state* dependent model depicted in Fig. 2, demonstrates the extreme instability of the statistical coding process.

5 Ongoing Work

The *Fisher game* has been extended to multi-dimensional and temporal cases. The model presented herein is in the process of being amalgamated with existing quantum key distribution protocols [19], to yield a hybrid statistical/quantum mechanical cryptosystem. Such a hybrid cryptosystem mitigates the current limitations of quantum channels to transmit large amounts of data. A *covert* quantum key distribution protocol may be utilized for the secure delivery of the cryptographic primitives (the $\delta G_{i,j}^{\varepsilon}$). Finally, the statistical encryption/decryption strategy has been modified to perform privacy protection in statistical databases. These results will be published elsewhere.

Table 1. Energy State Independent Model

$\|q\rangle = \|q_r\rangle$	$\|q_{r1}\rangle$
0.23813682639005	-0.00668168344388
0.69913526160795	0.20008072567388
0.27379424177629	-0.14186802540956
0.90226539453884	0.36853370671177

Table 2. Energy State Dependent Model

$\|q^{\varepsilon}\rangle = \|q_r^{\varepsilon}\rangle$	$\|q_{r1}^{\varepsilon}\rangle$
$\varepsilon = 0$ *Zero-energy/ground state*	
0.23813682639005	-0.26776249759842
0.69913526160795	0.77862610842042
0.27379424177629	-1.16636783859136
0.90226539453884	0.02881517541356
$\varepsilon = 1$ *First excited state*	
0.23813682639005	1.25161826270042
0.69913526160795	-3.255410114151938e-005
0.27379424177629	-0.11041660156776
0.90226539453884	0.61665920565232

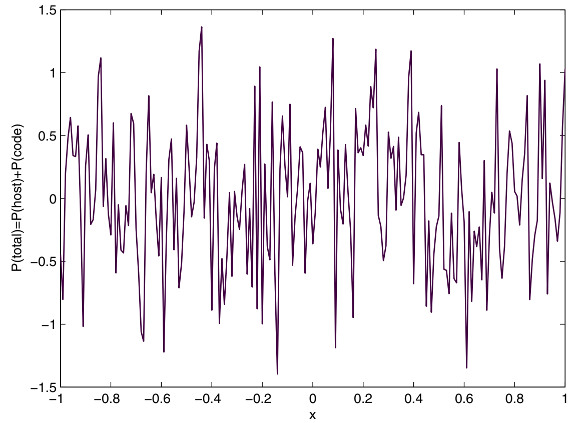

Figure 2. "Chaotic" Nature of the Total Pdf, Eq. (16)

Acknowledgements

This work was supported by *RAND-MSR contract CSM-DI & S-QIT-101107-2005*. Gratitude is expressed to B. R. Frieden, A. Plastino, and, B. Soffer for helpful discussions.

References

[1] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics*. 3rd edn, Pergamon Press, Oxford, 1977.
[2] B. R. Frieden. *Science from Fisher Information: A Unification*. Cambridge University Press, Cambridge. 2004.
[3] R. C. Venkatesan. Encryption of Covert Information into Multiple Statistical Distributions. *Phys. lett. A*, To Appear. Preprint available at at http://arxiv.org/abs/cond-mat/0607454, 2007.
[4] B. Schneier. *Applied Cryptography*. 2nd edn., John Wiley & Sons, New York, NY, 1996.
[5] R. L. Rivest. Cryptography and Machine Learning. *Proceedings ASIACRYPT '91*, 2006, pp.427-439, Springer 1993. Article available at http://theory.lcs.mit.edu/ rivest/Rivest-CryptographyAndMachineLearning.pdf.
[6] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, NY, 1991.
[7] P. J. Hüber. *Robust Statistics*. John Wiley & Sons, New York, NY, 1981.
[8] D. Horn and A. Gottlieb. Algorithm for Data Clustering in Pattern Recognition Problems based on Quantum Mechanics. *Phys. Rev. Lett.*, **88**, pp 18702(1)-18702(4), 2006.
[9] R. C. Venkatesan. Encryption of Covert Information through a Fisher Game. In *Exploratory Data Analysis Using Fisher Information*, Frieden, B.R. and Gatenby, R.A., (Eds.), Springer-Verlag, London, pp. 181-216, 2006.
[10] P. A. M. Dirac. *Principles of Quantum Mechanics (The International Series of Monographs on Physics)*. Oxford University Press, Oxford, 1982.
[11] H. L. van Trees. *Detection, Estimation, and Modulation Theory, Part I,*. Wiley, New York, NY, 1968.
[12] M. S. Rogalski and S. B. Palmer. *Quantum Physics*. Gordon & Breach Science Publications, Amsterdam, The Netherlands, 1998.
[13] A, Puente, M. Casas, and A. Plastino. Fisher Information and Semiclassical Methods. *Phys. Rev. A*, **59**(5), pp 3211-3217, 1999.
[14] L. Brillouin. *Science and Information Theory*. Academic Press, New York, NY, 1956.
[15] O. Morgenstern and J. von Neumann. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ, 1947.
[16] G. Gigerenzer and R. Selten. *Bounded Rationality*. The MIT Press, Cambridge, MA, 2002.
[17] David Wolpert. Information Theory - The Bridge Connecting Bounded Rational Game Theory and Statistical Physics. *Complex Engineering Systems*, Braha D. and Bar-Yam Y. (Eds.), Perseus Books, Boston, MA, 2004. Article available at http://arxiv.org/abs/cond-mat/0402508.
[18] G. H. Golub and C. F. van Loan. *Matrix Computations*. 3rd edn, Johns Hopkins University Press, Baltimore, MD, 1995.
[19] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.