GSI & CICT as coupled resource for multi-scale system biology and biomedical engineering modeling

Rodolfo A. Fiorini

***Abstract*—** To grasp a more reliable representation of reality and to get more resilient and antifragile system development techniques, researchers and scientists need two intelligently articulated hands: both stochastic and combinatorial approaches synergically articulated by natural coupling. The Geometric Science of Information (GSI) coupled to Computational Information Conservation Theory (CICT) can offer an effective framework to develop more competitive reality modeling. The first attempt to identify basic principles to get stronger modeling solution for scientific application has been developing at Politecnico di Milano University since the 1990s. This paper is a relevant contribute towards arbitrary multi-scale systems biology, biomedical engineering and computer science modeling, to show how GSI and CICT can offer stronger and more effective system modeling solutions for more reliable and resilient simulation.

***Keywords*—** arbitrary-scale system modeling, biomedical engineering, CICT, GSI.

I. Introduction

IN biology and biomedical research many questions are too complex to describe, let alone solve, in a practicable length of time. Even Stochastic vs. Combinatorially Optimized Noise generation ambiguity emphasises the information double-bind (IDB) problem in current most advanced instrumentation systems, just at the inner core of human knowledge extraction by experimentation in Science [1]. In fact, even the most sophisticated instrumentation system is completely unable to reliably discriminate so called "random noise" (RN) from any combinatorially optimized encoded message, called "deterministic noise" (DN) [2]. Epistemic and aleatory uncertainties are fixed neither in space nor in time. What is aleatory uncertainty in one model can be epistemic uncertainty in another model, at least in part. And what appears to be aleatory uncertainty at the present time may be cast, at least in part, into epistemic uncertainty at a later date [3]. Paradoxically if you don’t know the underlying hidden generating process for the folded information you can’t tell the difference between an information-rich message and a random jumble of letters [2]. The observer, having incomplete information about any generating process, and no reliable theory about what the data correspond to, will always make inference mistakes, but these mistakes have a certain pattern [3]. Statistical and applied probabilistic theory is the core of classic scientific knowledge; it is the logic of "Science 1.0"; it is the traditional instrument of risk-taking. Unfortunately, the "probabilistic veil" can be very opaque computationally in arbitrary multi-scale modeling, and misplaced precision leads to information dissipation and confusion [4]. To develop resilient and antifragile application, we need stronger biological and physical system correlates; in other words, we need asymptoctic exact global solution panoramas combined to deep local solution computational precision with information conservation and vice-versa. Can we achieve this goal?

Rodolfo A. Fiorini is with the Department of Electronics, Information and Bioengineering (DEIB), Politecnico di Milano University, Milano, 32 Piazza Leonardo da Vinci, MI 20133, ITALY (phone: +39 022399 3350; fax: +39 022399 3360; e-mail: rodolfo.fiorini@polimi.it).

II. The Root of the Problem

To find an innovative solution, we just need to remember the Relativity’s father inspiration quote: "We cannot solve our problems with the same thinking we used when we created them." Every approach that uses analytical function applies a top-down (TD) point-of-view (POV) implicitly. These functions belong to the domain of Infinitesimal Calculus (IC). In a multi-scale modeling framework, from a system computational pespective, all approaches that use a TD scale-free POV allow for starting from a global panorama of system parameter exact analytic solution families (macroscale infinite precision quantification). Then, going through system mesoscale, more and more shallow computational precision solution to real specific needs is reached, to arrive to system microscale where system parameter quantification is dominated by noise (quantification uncertainty). In other words, from global to local POV overall system information precision is not conserved, as misplaced precision leads to information dissipation and confusion [3,4] (see Fig. 1). In fact, usually further analysis and validation (by probabilistic and stochastic methods) is necessary to get localized computational solution of any practical value, in real application. A local discrete solution is worked out and computationally approximated as the last step in their line of reasoning, that started from an overall continuous system approach (from continuum to discreteness ≡ TD POV). Unfortunately, the IC methods are NOT applicable to discrete variable. To deal with discrete variables, we need the Finite Differences Calculus (FDC). FDC deals especially with discrete functions, but it may be applied to continuous function too. As a matter of fact, it can deal with both discrete and continuous categories conveniently. In other words, if we want to achieve an overall system information conservation approach, we have to look for a convenient bottom-up (BU) scale-relative POV (from discreteness to continuum view ≡ BU POV) to start from first, and NOT the other way around! Then, a TD POV can be applied, if needed (Fig.1). Current human made application and system can be quite fragile to unexpected perturbation because Statistics can fool you, unfortunately. Deep epistemic limitations reside in some parts of the areas covered in risk analysis and decision making applied to real problems [4]. We need tools able to manage ontological uncertainty more effectively [5],[6].

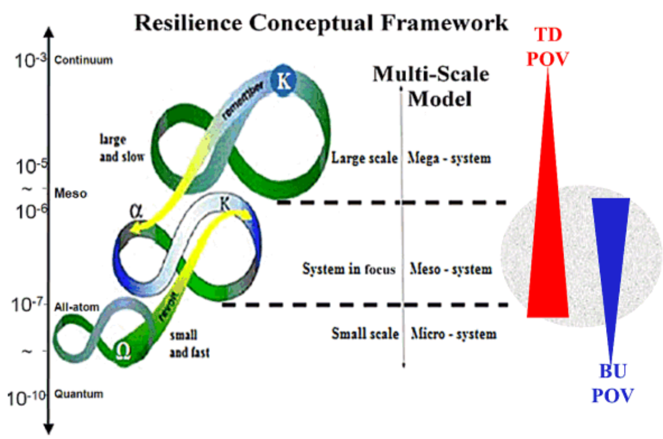

Fig. 1 Top-Down (TD) and Bottom-Up (BU) scale-relative Point-of-View (POV) in a Multi-scale Modeling Framework (see text).

III. Mathematical Theories For System Modeling

In the past, many attempts to arrive to a continuum-discrete articulated mathematical formulation have been proposed, all of them with big operational compromises. The most recent ones find their roots at the beginning of the 1980s, even if their publication may record a later public release [7]-[10]. The basic framework of statistics has been virtually unchanged since English statistician Ronald Fisher (1890-1962), Polish mathematician Jerzy Neyman (1894-1981) and British statistician Egon Pearson (1895-1980) introduced it starting in the second half of the 1920s, till the first half of the 1940s. In 1945, by considering the space of probability distributions, Indian-born mathematician and statistician Calyampudi Radhakrishna Rao (1920-) suggested that families of statistical distributions with continuous parameters may be regarded as Riemannian manifolds with parameters playing the role of coordinates. He used Fisher information matrix in defining the metric, so it was called Fisher–Rao metric [11]. This allowed the use of differential geometrical methods in the analysis of estimation, testing and other inference problems. The Riemannian geometry of statistical models was then studied as a mathematical curiosity for some years, with an emphasis in the geodesic distances associated with the Levi-Civita connection for this metric. In 1972, Russian mathematician Nikolai Chentsov proved that the Fisher information matrix is the only invariant Riemannian metric for statistical manifolds (up to some scalar factor), defined on the tangent space, that is decreasing under Markov morphisms (published in his Russian monograph and translated into English in 1982 by the AMS [12]). Because Markov morphisms represent coarse graining or randomization, it means that the Fisher information is the only Riemannian metric possessing the attractive property that distinguishability of probability distributions becomes more difficult when they are observed through a noisy channel. Since the 1980s the application of geometrical concepts to statistical theory and practice has been producing a number of different approaches which can be considered foundational to the modern Geometric Science of Information (GSI) [13]. This differential geometrization of Statistics [13] links with information theory through entropy functions, which appear as special cases of divergences. Its natural setting as part of probability theory in general, rendered this theory what is known today as the field of Information Geometry (IG). A greater amount of attention was devoted to the IG subject after American statistician B. Efron (1938-) introduced the concept of statistical curvature (called "e-curvature"), pointing out its importance to statistical inference, as well as implicitly using a new affine connection, which would be known as the "exponential connection" [14]. Every simply-connected Riemann surface can be given one of three geometries (Euclidean, spherical, or hyperbolic). Hyperbolic Geometry (HG) is the most prevalent geometry in this picture and also the most complicated.

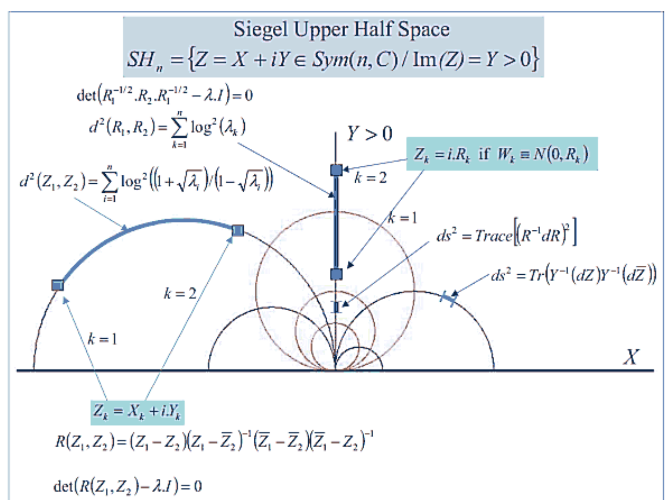

Fig.2 The Poincaré upper-half plane (PUHP) for 2-D problems, and the Siegel upper-half space (SUHS) for 3D problems (PUHP rotational symmetry along Y axis) [13].

In the two dimension hyperbolic case, the Beltrami-Klein model can be related to the Poincaré disk model (due to Riemann) and to the Poincaré upper-half plane model (PHUP). PHUP is named after French mathematician Henri Poincaré (1854–1912), but originated with Italian mathematician Eugenio Beltrami (1835–1899), who used it, along with the

Klein model and the Poincaré disk model, to show that hyperbolic geometry was equiconsistent with Euclidean geometry. The Poincaré disk model and the PHUP model are isomorphic under a conformal mapping. HG can describe a projective relativistic geometry. In the case of multivariate Gaussian probability distribution functions (pdfs), we can consider to replace the PUHP model by the Siegel Upper-Half Space (SUHS), as it is shown in Fig. 2. IG is applicable to convex analysis, even when it is not connected with probability distributions. This widens the applicability of IG to convex analysis, machine learning, computer vision, Tsallis entropy, economics, game theory, etc. The first attempt to identify basic principles, to synergically articulate CICT by natural coupling to GSI and IG, for scientific research and application, has been developing at Politecnico di Milano University since the 1990s. In 2013, the basic principles on CICT, from discrete system parameter and generator, appeared in literature and a brief introduction to CICT appeared in 2014 [1]. Traditional Number Theory and modern Numeric Analysis use LTR (left-to-right) mono-directional interpretation for ***Q*** Arithmetic single numeric group generator, so information entropy generation cannot be avoided in contemporary computational algorithm and application. On the contrary, according to CICT, it is quite simple to show information conservation and RTL (right-to-left) generator reversibility, by using basic considerations only. Eventually, CICT defines an arbitrary-scaling discrete Riemannian manifold uniquely, under HG metric, that, for arbitrary finite point accuracy level W going to infinity, under the criterion of scale relativity invariance, is isomorphic (even better, homeomorphic) to classic IG Riemannian manifold (exact solution theoretically).

IV. Operative Considerations

To better understand the CICT fundamental relationship that tie together numeric body information of divergent and convergent monotonic power series in any base (in this case decimal, with no loss of generality) with D ending by digit 9 is given by the following CICT fundamental correspondence equation [2]:

$$\frac{1}{D} = \sum_{k=0}^{\infty} \frac{1}{10^W}\left(\frac{\overline{D}}{10^W}\right)^k \Rightarrow \ldots \Leftarrow \; Div\left(\frac{1}{D}\right) = \sum_{k=0}^{\infty} (D+1)^k \qquad (01)$$

where $\overline{D}$ is the additive 10^W complement of D, i.e. $\overline{D} = (10^W - D)$, W is the word representation precision length of the denominator D and "*Div*" means "Divergence of". Further generalizations related to D ending by digit 1, 3 and 7 are straightforward [15]. Furthermore, When $\overline{D} > D$ the formal power series on the left of (01) can be rescaled mod D, to give multiple convergence paths to $1/D$, but with different "convergence speeds." The total number of allowed convergent paths, as monotonic power series, is given by the corresponding Q_L value, at the considered accuracy level L [2]. So, increasing the level of representation accuracy, the total number of allowed convergent paths to $1/D$, as monotonic power series (as allowed conservative paths), increases accordingly and can be counted exactly, and so on, till maximum machine word length and beyond, like discrete quantum paths denser and denser to one another, towards a never ending "blending quantum continuum," by a TD perspective [2].

V. Computational Example

CICT can be considered a natural framework for arbitrary-scale Systems Biology, Biomedical Engineering and Computer Science Modeling, in the current landscape of modern GSI. We present an example that takes advantage from scale related self-similarity modeling. Looking back to geometric series one can see a remarkable correspondence to the self-similarity concept. If we formally scale the following series S:

$$S = \sum_{k=0}^{\infty} q^k = 1 + q^1 + q^2 + q^3 + \ldots \qquad (02)$$

with the factor q, we obtain:

$$qS = q\sum_{k=0}^{\infty} q^k = q^1 + q^2 + q^3 + q^4 \ldots, \qquad (03)$$

therefore

$$S = \sum_{k=0}^{\infty} q^k = 1 + q\sum_{k=0}^{\infty} q^k \quad . \qquad (04)$$

This is the well-known "self-similarity" property of geometric series. The value of the sum S is $\bar{0}1.\bar{0}$ plus the scaled down version of the whole series. Naturally, self similarity only holds for the limit, but not for any finite stage. For example, let us suppose finite stage $k = 2$:

$S_2 = 1 + q + q^2$, then:

$$S_3 = 1 + qS_2 = 1 + q + q^2 + q^3 \neq S_2 \quad . \qquad (05)$$

As an example, for the sake of simplicity, let us use decimal base representation system, with no loss of generality, and let us consider $q = 3/10$ as LTR elementary generator. Then, we obtain the following convergent series:

$$S = \sum_{k=0}^{\infty} (3/10)^k = 1 + (3/10^1) + (9/10^2) + (27/10^3) + \ldots = \\ = 10/7 \qquad (06)$$

and

$$(3/10)\, S = \sum_{k=0}^{\infty} (3/10)^k = (3/10^1) + (9/10^2) + (27/10^3) + \ldots , \qquad (07)$$

therefore

$$\sum_{k=0}^{\infty} (3/10)^k = 1 + (3/10)\sum_{k=0}^{\infty} (3/10)^k \qquad (08)$$

as stated previously.

For finite stage, as we already stated, let us suppose $k = 2$, then:

$$S_2 = 1 + 3/10 + 9/100 = 139/100 \tag{09}$$

and

$$1 + (3/10)S_2 = 1 + (3/10^1) + (9/10^2) + (27/10^3) = \\ = 1417/1000 \neq S_2 \ . \tag{10}$$

Apparently by finite step we lost self-similarity, the fundamental property of geometric series, and this fact may seem, at first sight, a strong limitation to proceed further. But, self-similarity is still there, just a little less manifest. We can turn an apparent limitation into a striking computational advantage. In fact, it is possible to conceive an evolutive self-similar arithmetic correspondence (called complementary series or co-series) to original geometric series, step by step, which can act as a continuous connection from finite geometric power increment to its asymptotic limit to conserve characteristic computational information in a coherent way. To compute the corresponding LTR evolutive complementary arithmetic co-series (additive complement series), we introduce the fundamental concept of "coherent correspondence". Therefore, given any single term of original geometric series S_k, as s_k with an operational representation N_k/D_k , its "coherent correspondent term" m_c for correspondent complementary co-series M_1 is given, in this case, by $(D_k - N_k)/(D_k)^2$. So, the correspondent LTR complemenary co-series M_1 of our example is given by:

$$M_1 = \sum_{c=0}^{\infty} m^c = m_0 + m_1 + m_2 + m_3 + \ldots = \\ = 0 + (7/10^2) + (91/10^4) + (973/10^6) + \ldots = \\ = 7\{(0/10^0) + (1/10^2) + (13/10^4) + (139/10^6) + \ldots\} = \\ = 7\{S_0/(10^2) + S_1/(10^3) + S_2/(10^4) + \ldots\} = \\ = 7\{10/873\} = 7\{10/(9*97)\} \ . \tag{11}$$

As a matter of fact, M_1 is just the first co-series of a countable family M_r of complementary co-series to original series S, at different accuracy level r, given by:

$$M_r = D \sum_{c=0}^{\infty} S_c / (10)^{c+r+1} \quad r = 1, 2, \ldots, \infty \tag{12}$$

where D is the reduced denominator of the limit of original series S and S_c is its finite stage sum at stage c. So, in other words, each M_r is an irreducible co-domain, at different accuracy level r, for the original domain S. It is immediate to verify the following co-series limiting values:

$$M_2 = 7\{10^2/98703\} = 7\{10^2/(99*997)\} \ , \\ M_3 = 7\{10^3/9987003\} = 7\{10^3(999*9997)\} \ , \\ \ldots \\ M_7 = 7\{10^7/999999870000003\} \\ = 7\{10^7/(9999999*99999997)\} \ , \tag{13}$$

and so on. So, in other words, each M_r represents an irreducible co-domain, at accuracy level r, for the original domain S, with $q = 3/10$ as LTR elementary generator, in this case. Then, co-domain multiscale evolutive structured information, synthesized by its limiting value, can be used for deterministic noise source coherent tuning or checking for the presence of such specific generator in system "background noise." In fact, their numeric limiting values, by elementary arithmetic long division algorithm, supply us with cyclic remainder sequences perfectly tuned to deterministic source generators. By this kind of operational flexibility, a machine can generate autonomously, either on-the-fly or in advance stored in a-priori knowledge-base, combinatorially optimized exponential cyclic sequences (OECS) [2] to check for the presence of suspect "deterministic noise sources" in its probing field and then acting accordingly to compensate and to obtain a virtually homogeneous and uniform machine experimental reference domain. Following this line of thought, it is possible to overcome the dreadful ambiguity and limitations of the traditional Shannon entropy concept [2], and to solve the IDB problem successfully.

VI. Conclusion

CICT new awareness of a rational hyperbolic framework of encoded heterogeneous hyperbolic structures (reciprocal space), underlying the familiar Euclidean surface representation system (direct space) can open the way to holographic information geometry. This formulation has the great merit of maintaining close contact between the mathematical description and the physical phenomenon described, showing how to obtain a purely algebraic formulation of information and physical laws relating directly elementary information generators to experimental measurements. In fact, traditional elementary arithmetic long division remainder sequences can be interpreted as combinatorically optimized exponential cyclic sequences (OECS) for hyperbolic geometric structures, as points on a discrete Riemannian manifold, under HG metric, indistinguishable from traditional random noise sources by classical Shannon entropy, and current most advanced instrumentation approach. CICT defines an arbitrary-scaling discrete Riemannian manifold uniquely, under hyperbolic geometry (HG) metric, that, for arbitrary finite point accuracy level W going to infinity under scale relativity invariance, is isomorphic (even better, homeomorphic) to classic Riemannian manifold (exact solution theoretically). In other words, HG can describe a projective relativistic geometry directly hardwired into elementary arithmetic long division remainder sequences, offering many competitive computational advantages over traditional Euclidean approach only. More generally, CICT is a natural framework for arbitrary multi-scale computer science and systems biology modeling in the current landscape of the modern Geometric Science of Information (GSI). Specifically, high reliability organization (HRO), mission critical project (MCP) system, very low technological risk (VLTR) and crisis management (CM) system will be highly benefitted mostly by these new techniques. The present paper is a relevant contribute towards arbitrary multi-scale systems biology, biomedical engineering and computer science modeling, to show how GSI and CICT as coupled resource can offer stronger and more effective system modeling solutions for more reliable, effective and powerful complex system simulation.

REFERENCES

[1] R.A. Fiorini, "Computational Information Conservation Theory: An Introduction," in *2014 Proceedings of the 8th International Conference on Applied Mathematics, Simulation, Modelling (ASM '14)*, N.E. Mastorakis, M. Demiralp, N. Mukhopadhyay, F. Mainardi, eds., Mathematics and Computers in Science and Engineering Series, No.34, NAUN Conferences, WSEAS Press, November 22-24, 2014, Florence, Italy, pp.385-394, ISSN: 2227-4588, ISBN: 978-960-474-398-8.

[2] R.A. Fiorini, "How Random is Your Tomographic Noise? A Number Theoretic Transform (NTT) Approach," *Fundamenta Infomaticae*, vol.135, no.1-2, pp.135-170, 2014.

[3] T.C. Hanks, C.A. Cornell, "Probabilistic Seismic Hazard Analysis: A Beginner's Guide," in *1994 Proceedings of the Fifth Symposium on Current Issues Related to Nuclear Power Plant Structures, Equipment and Piping*. Raleigh, N.C.: North Carolina State University, 1994, pp.I/1-1 to I/1-17.

[4] N.N. Taleb, *Silent Risk.* Preliminary Book Draft, January, 2015.

[5] R.A. Fiorini, "Stronger Physical and Biological Measurement Strategy for Biomedical and Wellbeing Application by CICT," in *2014 Proceedings of the 3rd International Conference on Health Science and Biomedical Systems (HSBS '14)*, N.E. Mastorakis, Atsushi Fukasawa, eds., Series: Recent Advances in Biology and Biomedicine Series No.6, NAUN Conferences, WSEAS Press, November 22-24, 2014, Florence, Italy, pp.36-45.

[6] R.A. Fiorini, "Stronger Quantum Decoherence Incomputability Modeling by CICT," in *2014 Proceedings of the 3rd International Conference on Applied and Computational Mathematics (ICACM '14)*, Nikos E. Mastorakis, Gen Qi Xu, Eds. NAUN Conferences, WSEAS Press, December 29-31, Geneva, Switzerland, 2014, pp.78-86.

[7] S. Hilger, "Ein Masskettenkalkül mit Anwendung auf Zentrumsmannigfaltigkeiten," PhD Thesis, Universität Würzburg, 1988.

[8] M. Bohner, A. Peterson, Eds, *Advances in Dynamic Equations on Time Scales.* Berlin: Birkhäuser, 2003.

[9] C. Auffray and L. Nottale, "Scale relativity theory and integrative systems biology: 1. Founding principles and scale laws," *Progr. Biophys. Mol. Bio.*, vol.97, no.1, pp.79-114, May 2008.

[10] L. Nottale and C. Auffray, "Scale relativity theory and integrative systems biology: 2. Macroscopic quantum-type mechanics," *Progr. Biophys. Mol. Bio.*, vol.97, no.1, pp.115-157, May 2008.

[11] C.R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, no.37, pp.81–89, 1945.

[12] N.N. Chentsov, Statistical Decision Rules and Optimal Inferences, *Transactions of Mathematics Monograph*, no.53, Lev J. Leifman, ed. Providence, RI, USA: AMS, 1982, (Published in Russian in 1972).

[13] F. Nielsen, (ed.), *Geometric Theory of Information.* Heidelberg: Springer, 2014.

[14] B. Efron, "The Geometry of Exponential Families," *The Annals of Statistics*, no.6, pp.362-366, 1978.

[15] R.A. Fiorini, "Computerized Tomography Noise Reduction by CICT Optimized Exponential Cyclic Sequences (OECS) Co-domain," *Fundamenta Informaticae*, 2015, in press.