

Article

Optimization Models for Reaction Networks: Information Divergence, Quadratic Programming and Kirchhoff's Laws

Julio Michael Stern ^{1,*} and Fabio Nakano ²

¹ IME-USP, Institute of Mathematics and Statistics of the University of São Paulo;
Rua do Matão, 1010, Cidade Universitária, 05508-090, São Paulo, Brazil

² EACH-USP, School of Arts, Sciences and Humanities of the University of São Paulo;
Av. Arlindo Béttio, 1000, Ermelino Matarazzo, 03828-000, São Paulo, Brazil.

* Author to whom correspondence should be addressed; E-Mail: jstern@ime.usp.br;
Tel./Fax: +55-11-3819-3922.

Received: 4 November 2013; in revised form: 19 February 2014 / Accepted: 21 February 2014 /
Published: 5 March 2014

Abstract: This article presents a simple derivation of optimization models for reaction networks leading to a generalized form of the mass-action law, and compares the formal structure of Minimum Information Divergence, Quadratic Programming and Kirchhoff type network models. These optimization models are used in related articles to develop and illustrate the operation of ontology alignment algorithms and to discuss closely connected issues concerning the epistemological and statistical significance of sharp or precise hypotheses in empirical science.

Keywords: affinity; conservation laws; entropy and quadratic optimization; mass-action law; ontology alignment; stoichiometry; variational principles

Our acceptance of an ontology is, I think, similar in principle to our acceptance of a scientific theory, say a system of physics: We adopt, at least insofar as we are reasonable, the simplest conceptual scheme into which the disordered fragments of raw experience can be fitted and arranged... To whatever extent the adoption of any system of scientific theory may be said to be a matter of language, the same -but no more- may be said of the adoption of an ontology. Willard van Orman Quine, [1]; On What There Is.

1. Introduction

In 1999, [2] presented the FBST-The Full Bayesian Significance Test for Precise Hypotheses. The FBST is based on the e -value, $ev(H|X)$, the epistemic value of hypothesis H given the observed data X . $ev(H|X)$ is a possibilistic measure obtained by integrating the posterior probability measure of the underlying Bayesian statistical model at focal sets defined by the surprise function, $s_n(\theta) = p_n(\theta)/r(\theta)$. The logical and operational properties of the FBST differ significantly from the well-known Bayes Factors, the standard Bayesian choice for accessing non-sharp hypotheses. However, the FBST approach can easily handle important sharp hypotheses problems using a direct approach and standard priors, while Bayes Factors require convoluted constructions, see [3,4].

In several Bayesian meetings, FBST-related presentations motivated heated discussions concerning the relevance of sharp or precise hypotheses in the practice of empirical science. In articles [5–7] we address this issue, showing the importance of sharp hypotheses in the definition of scientific ontologies. This article tries to give a very simple and direct derivation of some formalisms for reaction networks that are based primarily on entropy optimization. Our conclusions about these formalisms and how they relate to each other are important for a rigorous discussion of the arguments given in articles [5–7], and is also required for future research.

This article presents a formal but simple derivation, using only very basic and standard results of convex optimization theory, of a Minimum Information Divergence optimization model for chemical (or abstract) reaction networks leading to a generalized form of mass-action law. Information Divergence is a (non-symmetric) measure for the difference between two probability distributions. In Statistics this measure is known as the Kullback-Leibler “distance”; the denomination Relative Entropy is used in Physics and Engineering. Moreover, using again only very basic and standard tools of convex analysis and optimization theory, we derive Quadratic Programming optimization models for close to equilibrium conditions. Furthermore, we examine some interpretations for these models based on analogies to Kirchhoff (electrical) networks.

Each one of these alternative formulations could be used as the core of an axiomatic system for a theory on reaction networks. This article focus on the structural similarities between this alternative formulations, for these similarities can be used to link up the testable (sharp statistical) hypotheses of the corresponding theories that, in turn, are the key for developing ontology alignment algorithms, as discussed in [7]. For interesting applications and concrete examples of chemical networks see [8], whose notation we follow closely, and the references cited therein.

The minimum divergence optimization problems studied in this article are formally very similar to some important variational problems discussed in [5]. That article examines, using formal and rigorous methods of analysis, two important questions: (1) The relation between conserved quantities and invariance properties of a theory; (2) The ontological status of such invariant entities according to an epistemological framework based on cognitive constructivism and supported by the e -value. For more information on this Bayesian epistemological framework, see [2,5–7,9–15]. Formal and interpretative similarities between the variational problems discussed at [5] and the present article should make it easy to compare the two papers’ analyses and align its conclusions.

2. Stoichiometry and Affinity

The reaction processes discussed in this article obey two kinds of conservation laws, namely, mass conservation, concerning chemical compounds or abstract simple elements, and energy conservation, concerning the potentials driving the system's dynamics. This and the following sections present a model for reaction networks within the general structure of the variational problems studied in [5] and, in so doing, make it possible to highlight important relations between conservation laws and invariance properties of the model. Furthermore, the following sections investigate similarities in the formal structure of alternative models for reaction networks.

Whenever possible, matrices are represented by upper-case Latin letters, vectors by lower-case, and scalars by Greek letters. The superscripts -1 , t and $-t$ indicate the matrix inverse, transpose and inverse-transpose. The Kronecker delta, δ_i^k , takes the value 1 if $i = k$ and the value 0 otherwise. The o-operators, \odot , \otimes , *etc.*, indicate the point-wise matrix version of the corresponding scalar arithmetic operator.

In a reaction network, mass conservation is described by stoichiometric linear equations in the form $Sv = dx/dt$. The *stoichiometric matrix*, S , is an $m \times n$ sparse and integer matrix; each of its m rows represents an individual compound; each of its n columns represents an individual reaction; and each non-zero element, $S_{i,j}$, represents the stoichiometric coefficient of compound i in reaction j . Typically, these quantities exhibit several interdependence relations, implying $\text{rank}(S) < m < n$; for more details, see Section 3.1. Conventionally, in a given reaction, reactants and products have, respectively, negative and positive values. The system's reaction rates are given by the n -dimensional *flux vector*, v . Mass conservation constraints are expressed by the flux balance equation, $Sv = dx/dt$, relating flux to the differential increase in the m -dimensional vector of chemical concentrations, x . For a system at equilibrium or steady state, $dx/dt = 0$. The condition $Sv = 0$ for a reaction network is analogous to Kirchhoff's current law in an electrical network, as explained in Section 4.

Open systems can be modeled adding a mass exchange with the external environment in the form of an influx or outflux, $b = -S_e v_e$. The extended steady state conservation equation then becomes $Sv = b$. The special case of a network having only elementary reversible reactions can be highlighted splitting the internal flux vector and stoichiometric matrix in its forward (f) and reverse (r) parts, namely, $v^t = [v_f^t, v_r^t]^t$ and $S = [S_f, S_r]$, with $S_r = -S_f$. For such elementary reversible networks the steady state balance equation takes the form $S_f v_f - S_r v_r = b$. The examples studied at [7] correspond to this case. Therefore, without loss of generality, we retain this special form in the following analyses.

Energy conservation is described by stoichiometric linear equations in the form $\Delta u = S^t u$, where the m -dimensional vector u represents the *chemical potential* of each component in the reaction network, and the n -dimensional vector Δu represents *change in chemical potential* for the system's internal reactions. The scalar product of drop in chemical potential and flux, $-\Delta u^t v$, represents a power dissipation rate. In a complex system, $m < n$, the energy conservation equations imply an analogous of Kirchhoff's voltage law, stating that the change in chemical potential around any stoichiometrically balanced loop is zero, see Section 4. The absolute value of weighted partial sums over products and reactants defining Δu are called in chemistry the forward and reverse chemical *affinity* of each reaction.

The *mass-action* law for reaction networks as above states that $\rho \log(v_f \oslash v_r) = -\Delta u$, where $\rho = RT$, the product of gas constant and absolute temperature. The mass-action law is valid for ideal gas reactants under (near) equilibrium, constant temperature, pressure, *etc.* and homogeneous (ergodic, good mixing) conditions. Moreover, the system must be non-degenerate in the sense that it complies with regularity conditions carefully described in Sections 3 and 4. Furthermore, the ideal conditions stated above should be supplemented with auxiliary conditions concerning molecular kinetics, see Sections 12.8 of [16], Sections 4.1 and 5.3 of [17], [18], Sections 3.6 and 3.7 of [19], Chapter 4 of [20], and Section 14.4 of [21]; see also [22,23]. Nevertheless, it is well known that mass-action flow relations can provide good practical approximations for reaction networks even if some of these strict conditions are somewhat relaxed. For an historical perspective, see Section 3.3.3 of [24] and Chapter 15 of [25].

3. Relative Entropy Minimization

Reaction networks are modeled as a dynamical process under invariance constraints. In this context, it is natural to look for a fundamental variational principle, see [5]. This goal requires an optimization problem, constrained by mass conservation stoichiometric equations, that renders optimal fluxes in a reaction network that are compatible with an underlying mass-action law. The following minimum Information Divergence (Relative Entropy) optimization problem achieves this goal:

$$\min_v \varphi(v, q) = v_f^t \log\left(\frac{v_f}{q_f}\right) + v_r^t \log\left(\frac{v_r}{q_r}\right)$$

for forward and reverse fluxes $v_f > 0$ and $v_r > 0$ such that $S_f v_f - S_r v_r = b$, and fixed *reference* fluxes q_f and q_r . This relative entropy optimization problem is a simplification, but also a slight generalization, of the version presented in Theorem 1 of [8]. More general formulations are considered in Chapter VIII of [26]. Section 3.6 of [21] and [27] give a unified view of the minimum relative entropy formalism, see also [26,28–33]. Conceptually, one could work with normalized fluxes, $v \geq 0 \mid \mathbf{1}^t v = 1$, in order to highlight probabilistic or informational interpretations of these problems.

Convexity properties of the objective function φ , on its vector argument v , can be asserted from the gradient vector and positive definite (and diagonal) Hessian matrix given by the following derivatives:

$$d(v, q)_i = \frac{\partial \varphi}{\partial v_i} = 1 + \log\left(\frac{v_i}{q_i}\right), \quad D(v, q)_{i,k} = \frac{\partial^2 \varphi}{\partial v_i \partial v_k} = \delta_i^k \frac{1}{v_i}$$

Using m -dimensional Lagrange multipliers, y , we obtain the following Karush-Kuhn-Tucker viability and optimality conditions, see [34,35],

$$\begin{aligned} v_f > 0, \quad v_r > 0, \quad S_f v_f - S_r v_r &= b \\ S_f^t y = d(v_f, q_f) &= \log(v_f) + 1 - \log(q_f) \\ -S_r^t y = d(v_r, q_r) &= \log(v_r) + 1 - \log(q_r) \end{aligned}$$

Subtracting the last two equations, we have

$$\log(v_f \oslash v_r) = 2S_f^t y + \log(q_f \oslash q_r) = 2S_f^t y + c$$

In the special case of $q_f = q_r$, the same forward and reverse reference fluxes, we have $c = 0$. Hence, taking $u = -2\rho y$, we get flux relations compatible with standard mass-action law.

3.1. Constraint Qualification and Rank Condition

Before ending this section, we make some comments on the delicate issue of constraint qualification and rank condition. If the constraint matrix does not have full row-rank, that is, if $\text{rank}(S) < m < n$, then, even if the primal solution is uniquely defined, the corresponding dual vector (Lagrange multipliers) is not unique, see ([36] p. 323), and ([37] p. 54). Under appropriate normalization and regularity conditions, rank deficiency does not prevent the convergence of well-known relaxation algorithms for entropy optimization. For a general overview see [38–40]; for the celebrated Bregman method, see [5,41], and Theorem 3.1 at ([37] p. 61); for the Multiplicative Algebraic Reconstruction Technique (MART), see Theorem 3.3 at ([37] p. 69) and [36]. Closely related Proximal Point Algorithms can handle more general optimization problems under similar rank deficient conditions, see Proposition 4.1 at ([42] pp. 233–235) and [43].

In contrast, more traditional optimization algorithms often require an independence constraint qualification. Hence, employment of such algorithms requires pre-processing the original problem by manipulation techniques capable of reducing it to a full rank version. Furthermore, even when employing proximal point methods, convergence times for rank deficient problems can be significantly altered by such pre-processing, see [36, p. 337]. Regularization methods are yet another important alternative used to overcome the typical difficulties of dealing with optimization problems with dependent constraints, see [44] and ([39] pp. 49,50,130).

Problem reduction to full rank condition may be achieved detecting, by simple inspection, and subsequently pruning an appropriate number of linearly dependent constraints. Of course, such a naïve approach is only feasible for small applications. General purpose reduction techniques can be implemented using (sparse) matrix factorization algorithms. For example, see [45,46] for the incomplete (sparse) LU or QR factorizations. These algorithms are part of standard libraries for computational linear algebra, allowing straightforward implementation. For specific examples and manipulation techniques specially devised for network applications, see [47] and ([48] Chapter 3).

4. Quadratic Programming and Kirchhoff’s Laws

The algebraic form of the objective function’s gradient and Hessian renders the following second order Taylor series approximation around v^* , the (primal) optimum solution of the original minimum divergence problem:

$$\min_v \xi(v, q) = \frac{v^t D(v^*, q)v}{2} + d(v^*, q)v$$

for $v > 0$, such that $Sv = b$

This positive definite and diagonal *quadratic programming* problem generalizes the minimum power dissipation variational formulation for resistive electrical circuits. In this case, the stoichiometric matrix is simplified to a node-edge incidence matrix where each column, $S_{\bullet,j}$, has only two non-zero elements, namely, -1 and $+1$ at the j -th (directed) arch’s node of origin and destination in a digraph representing the topology of the electrical circuit. A dual optimal solution, y^* , represents node’s potentials (voltages), vector $\text{diag}(D)$ represents arch’s resistances, and vector b represents external power (current) sources.

Lagrange optimality condition for this quadratic programming problem, see Chapter 7 of [49] or Section D.3 of [50], imply the *complementary slackness* condition: $y^t(Sv - b) = 0$. In the electro/chemical network literature this transversality condition is known as *Tellegen's theorem*, an expression that encodes several important conservation laws, including Kirchhoff's voltage and current laws, see [51].

More general forms of the divergence function, $\varphi(v, q)$, can be used to model alternative interaction kinetics, see [52] and Chapter VIII of [26]. However, as long as they retain some basic projection properties, the underlying optimization problem will be well behaved, see [36,37,39,42,44]. Moreover, as long as the divergence's Hessian remains positive definite, network flow interpretations may be useful to understand the reaction system, see [53–58]. Matrix transformations to diagonal, block-diagonal or other structured forms may facilitate further interpretations in specialized applications.

5. Further Research and Final Remarks

Each one of the alternative formulations for reaction networks presented in this article, namely, minimum divergence, quadratic programming or Kirchhoff's networks, could be used as the core of an axiomatic system for a scientific discipline on reaction networks, entailing different theoretical perspectives and, consequently, having distinct commitments (or being committed in different degrees) to entities in their respective ontologies. For example, entropy may be a central concept in a good ontology for one of these models, while it may be regarded as an auxiliary construct in another alternative. Our opening quotation points to fascinating implications of choosing between these alternatives, an issue to be fully explored in [7] and subsequent articles, using the models developed in this article as basic examples.

Information about the structural similarities between alternative models for reaction networks can be used to identify or match the "same" chemical concept or entity as it occurs in different models. Such a systematic entity correspondence mapping is known in computer science as an *ontology alignment*. Diachronic ontology alignments can be used to trace the development of a given concept in the historical evolution of a scientific discipline. Synchronic ontology alignments can be used to compile (approximate) translation dictionaries for contemporary disciplines and their respective languages. Following articles discuss how structural similarities can be used to suggest validation criteria or (partial) objective functions for ontology alignment algorithms.

For a review of the historical evolution of conservation laws in chemistry related to affinity and stoichiometry, see [7,24,25], and the references cited therein. The additive algebraic structure of the entropy optimization problem constraints, together with the differential decoupling properties of its objective function, can be used to render first order approximations for the laws governing chemical reaction fluxes that are analogous to Kirchhoff's laws governing electrical networks, as demonstrated in Section 4. The simple additive algebraic structure of Kirchhoff's laws for network flows replicates the additive structure for chemical affinities that Louis-Bernard Guyton de Morveau could foresee way ahead of his time, see [24,25,59,60]. In [7] and following articles of this research group, these models are further compared, and also used to illustrate the operation of ontology alignment algorithms.

The graph or hyper-graphs representation and the underlying matroid algebraic structure of network models reveals very interesting properties from the cognitive constructivism perspective, for these properties directly relate to the decoupling concept studied in [61] and Chapter 3 of [50], see also [62–65]. We hope to further explore these properties in following articles.

Finally, an anonymous referee suggests possible extensions of the approach presented in this article for systems of reactions with different time scales in networks under high frequency observation, conditions likely to induce non-neglectable relaxation times. Following our analogies to electrical circuits, this suggests the incorporation of some form of “capacitive” or “inductive” elements to the reaction network model, see [66]. We intend to explore this interesting suggestion in future research.

Acknowledgments

The authors are grateful for the support of IME-USP and EACH-USP, the Institute of Mathematics and Statistics and the School of Arts, Sciences and Humanities of the University of São Paulo; FAPESP-the State of São Paulo Research Foundation (grant CEPID 2013/07375-0); and CNPq-the Brazilian National Counsel of Technological and Scientific Development (grant PQ 301206/2011-2). Finally, the authors are grateful for advice of anonymous referees and comments received from participants of MaxEnt 2013-the 33rd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, held on December 15–12 at Canberra, Australia, and from participants of EBEB 2014-the 12th Brazilian Meeting on Bayesian Statistics, held on March 10–14 at Atibaia, Brazil.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Quine, W.O. On What There Is. Review of Metaphysics In *From a Logical Point of View*; Harvard University Press: Cambridge, USA, 1961; Volume 2, pp.21–38.
2. Pereira, C.A.B.; Stern, J.M. Evidence and credibility: Full bayesian significance test for precise hypotheses. *Entropy J.* **1999**, *1*, 69–80.
3. Diniz, M.; Pereira, C.A.B.; Stern, J.M. Unit roots: Bayesian significance test. *Commun. Stat.-Theory Methods* **2011**, *40*, 4200–4213.
4. Diniz, M.; Pereira, C.A.B.; Stern, J.M. Cointegration: Bayesian significance test. *Commun. Stat.-Theory Methods* **2012**, *41*, 3562–3574.
5. Stern, J.M. Symmetry, invariance and ontology in physics and statistics. *Symmetry* **2011**, *3*, 611–635.
6. Stern, J.M.; Pereira, C.A.B. Bayesian epistemic values: Focus on surprise, measure probability! *Logic J. IGPL* **2013**, doi:10.1093/jigpal/jzt023.
7. Stern, J.M. Jacob’s ladder and scientific ontologies. *Cybernet. Hum. Know.* **2013**, in press.

8. Fleming, R.M.T.; Maes, C.M.; Saunders, M.A.; Ye, Y.; Palsson, B.O. A variational principle for computing nonequilibrium fluxes and potentials in genome-scale biochemical networks. *J. Theor. Biol.* **2012**, *292*, 71–77.
9. Borges, W.; Stern, J.M. The rules of logic composition for the bayesian epistemic e -values. *Logic J. IGPL* **2007**, *15*, 401–420.
10. Madruga, M.R.; Esteves, L.; Wechsler, S. On the bayesianity of pereira-stern tests. *Test* **2001**, *10*, 291–299.
11. Madruga, M.R.; Pereira, C.A.B.; Stern, J.M. Bayesian evidence test for precise hypotheses. *J. Stat. Plan. Inference* **2003**, *117*, 185–198.
12. Pereira, C.A.B.; Wechsler, S.; Stern, J.M. Can a significance test be genuinely bayesian? *Bayesian Anal.* **2008**, *3*, 79–100.
13. Stern, J.M. Cognitive constructivism, eigen-solutions, and sharp statistical hypotheses. *Cybernet. Hum. Know.* **2007**, *14*, 9–36.
14. Stern, J.M. Language and the self-reference paradox. *Cybernet. Hum. Know.* **2007**, *14*, 71–92.
15. Stern, J.M. Constructive verification, empirical induction, and fallibilist deduction: A threefold contrast. *Information* **2011**, *2*, 635–650.
16. Callen, H.B. *Thermodynamics: An Introduction to the Physical Theories of Equilibrium Thermostatics and Irreversible Thermodynamics*; John Wiley: New York, NY, USA, 1960.
17. Callaghan, C.A. *Kinetics and Catalysis of the Water-Gas-Shift Reaction: A Microkinetic and Graph Theoretic Approach*. Ph.D. Thesis, Worcester Polytechnic Institute, 2006.
18. Gillespie, D. A rigorous derivation of the chemical master equation. *Phys. A* **1992**, *188*, 404–425.
19. Prigogine, I. *Introduction to the Thermodynamics of Irreversible Processes*, 2nd ed.; Interscience: New York, NY, USA, 1961.
20. Ross, J.; Berry, S.R. *Thermodynamics and Fluctuations far from Equilibrium*; Springer: New York, NY, USA, 2008.
21. Tribus, M. *Thermostatistics and Thermodynamics: An Introduction to Energy, Information and States of Matter, with Engineering Applications*; van Nostrand: Princeton, NJ, USA, 1961.
22. Gardiner, C. *Stochastic Methods: A Handbook for the Natural and Social Sciences*; Springer: New York, NY, USA, 2010.
23. Van Kanpen, N.G. *Stochastic Processes in Physics and Chemistry*; Elsevier: Amsterdam, The Netherlands, 2007.
24. Goupil, M. *Du Flou au Clair? Histoire de l’Affinité Chimique de Cardan à Prigogine* (in French); CTHS: Paris, France, 1991.
25. Muir, P. *A History of Chemical Theories and Laws*; John Wiley: New York, NY, USA, 1907;
26. Kapur, J.N.; Kesavan, H.K. *Entropy Optimization Principles with Applications*; Academic Press: Boston, MA, USA, 1992.
27. Tribus, M.; McIrvine, E.C. Energy and information. *Sci. Am.* **1971**, *224*, 178–184.
28. Jaynes, E.T. The minimum entropy production principle. *Ann. Rev. Phys. Chem.* **1980**, *31*, 579–601.

29. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: Cambridge, UK, 2003.
30. Kapur, J.N. *Maximum Entropy Models in Science and Engineering*; John Wiley: New Delhi, India, 1989.
31. Niven, R.K. Steady state of a dissipative flow-controlled system and the maximum entropy production principle. *Phys. Rev. E* **2009**, *80*, 1–15.
32. Niven, R.K. Minimization of a free-energy-like potential for non-equilibrium flow systems at steady state. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2010**, *365*, 1323–1331.
33. Niven, R.K. Maximum entropy analysis of steady-state flow systems (and extremum entropy production principles). *AIP Conf. Proc.* **2011**, *1443*, 270–281.
34. Luenberger, D.G. *Linear and Nonlinear Programming*; Addison-Wesley: Reading, UK, 1984.
35. Minoux, M.; Vajda, S. *Mathematical Programming*; John Wiley: Chichester, USA, 1986.
36. Elfving, T. On some methods for entropy maximization and matrix scaling. *Linear Algebra Appl.* **1980**, *34*, 321–339.
37. Fang, S.C.; Rajasekera, J.R.; Tsao, H.S.J. *Entropy Optimization and Mathematical Programming*; Kluwer: Dordrecht, The Netherlands, 1997.
38. Censor, Y. Row-action methods for huge and sparse systems and their applications. *SIAM Rev.* **1981**, *23*, 444–466.
39. Censor, Y.; Zenios S.A. *Parallel Optimization: Theory, Algorithms, and Applications*; Oxford University Press: New York, NY, USA, 1997.
40. Censor, Y.; de Pierro, A.; Elfving, T.; Herman, G.; Iusem, A. On iterative methods for linearly constrained entropy maximization. *Num. Anal. Math. Model. Banach Center Publ. Ser.* **1990**, *24*, 145–163.
41. Bregman, L.M. The relaxation method for finding the common point convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. Math. Phys.* **1967**, *7*, 200–217.
42. Bertsekas, D.P.; Tsitsiklis, J.N. *Parallel and Distributed Computation, Numerical Methods*; Prentice Hall: Englewood Cliffs, NJ, USA, 1989.
43. Garcia, M.V.P.; Humes, C.; Stern, J.M. Generalized line criterion for gauss seidel method. *J. Comput. Appl. Math.* **2002**, *22*, 91–97.
44. Iusem, A.N. *Métodos de Ponto Proximal em Otimização* (in Portuguese); IMPA: Rio de Janeiro, Brazil, 1995.
45. Golub, G.H.; van Loan, C.F. *Matrix Computations*; Johns Hopkins: Baltimore, MD, USA, 1989.
46. Stern, J.M. *Esparsidade, Estrutura, Estabilidade e Escalonamento em Álgebra Linear Computacional* (in Portuguese); UFPE, IX Escola de Computação: Recife, Brazil, 1994.
47. Steuer, R.; Junker, B.H. Computational models of metabolism: Stability and regulation in metabolic networks. *Adv. Chem. Phys.* **2009**, *142*, 105–252.
48. Heinrich, R.; Schuster, S. *The Regulation of Cellular Systems*; Chapman and Hall: New York, NY, USA, 1996.
49. Hadley, G. *Nonlinear and Dynamic Programming*; Addison-Wesley: New York, NY, USA, 1964.

50. Stern, J.M. Cognitive Constructivism and the Epistemic Significance of Sharp Statistical Hypotheses. In Proceedings of the Tutorial Book for MaxEnt 2008, the 28th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Boracéia, São Paulo, Brazil, July 6-11, 2008.
51. Penfield, P.; Spence, R.; Duinker, S. *Tellegen's Theorem and Electrical Networks*; MIT Press: Cambridge, MA, USA, 1970.
52. Gorban, A.N.; Shahzad, M. The michaelis-menten-stueckelberg theorem. *Entropy* **2011**, *13*, 966–1019.
53. Bertsekas, D.P. Thevelin decomposition and large scale optimization. *JOTA* **1996**, *89*, 1–15.
54. Fishtik, I.; Callaghan, C.A.; Datta, R. Reaction route graphs I: Theory and algorithm. *J. Phys. Chem. B* **2004**, *108*, 5671–5682.
55. Fishtik, I.; Callaghan, C.A.; Datta, R. Wiring diagrams for complex reaction networks. *Ind. Eng. Chem. Res.* **2006**, *45*, 6468–6476.
56. Millar, W. Some general theorems for non-linear systems possessing resistance. *Philos. Mag.* **1951**, *7*, 1150–1160.
57. Peusner, L. *Studies in Network Thermo-Dynamics*; Elsevier: Amsterdam, The Netherlands, 1986.
58. Wiśniewski, S.; Staniszewski, B.; Szymanik, R. *Thermodynamics of Nonequilibrium Processes*; Reidel: Dordrecht, The Netherlands, 1976.
59. Morveau, L.B.G.de. Affinity. In *Supplement to the Encyclopaedia or Dictionary of Arts, Sciences and Miscellaneous Literature*; Thomas Dobson: Philadelphia, USA, 1803; pp. 391–405.
60. De Morveau, L.B.G.; Lavoisier, A.L.; Berthollet, C.L.; Fourcroy, A. *Méthode de Nomenclature Chimique*; Chez Cuchet: Paris, France, 1787.
61. Stern, J.M. Decoupling, sparsity, randomization, and objective bayesian inference. *Cybernet. Hum. Know.* **2008**, *15*, 49–68.
62. Bryant, V.; Perfect, H. *Independence Theory in Combinatorics: An Introductory Account with Applications to Graphs and Transversals*; Chapman and Hall: London, UK, 1980.
63. Recski, A. *Matroid Theory and its Applications in Electrical Network Theory and in Statics*; Akadémiai Kiadó: Budapest, Hungary, 1989.
64. Swamy, M.N.S.; Thulasiraman, K. *Graphs, Networks and Algorithms*; Wiley: New York, NY, USA, 1981.
65. Vágó, I. *Graph Theory: Applications to the Calculation of Electrical Networks*; Elsevier: Amsterdam, The Netherlands, 1985.
66. Thoma, J.; Mocellin, G. *Simulation with Entropy in Engineering Thermodynamics. Understanding Matter and Systems with Bondgraphs*; Springer: New York, NY, USA, 2006.