Advanced Mean Field Methods: Theory and Practice by M. Opper and D. Saad (Eds.). The MIT Press, Cambridge, Massachusetts, London, England, 2001,
Neural Information Processing Series, 273 pages, hardbound, ISBN 0-262-15054-9.
Reviewed by Jacek Mańdziuk
Published in *IEEE Transactions on Neural Networks*, 13(3), 2002, pp. 785-786

Several problems in statistical physics, information sciences or neural computing require probabilistic modeling of multivariable systems composed of a large number of variables. Typically, the exact calculation of such models is computationally infeasible, hence there is a strong need for efficient, approximate methods in this area. One type of such methods well known to the statistical physics community are the Mean Field Methods (MFMs) in which the value of each random variable is approximated by the so-called effective field. Consequently, in the simple MFM approach the true (yet generally intractable) probability distribution of random variables is approximated by a factorized distribution, which then can be approached by variational optimization methods. More advanced MFMs are based on the TAP (Thouless, Anderson, Palmer) method which incorporates several non-trivial dependencies between variables neglected in simple MFM approximations.

MFMs were originally introduced and developed for calculations of the spin-glass models in quantum mechanics or, equivalently, for the Hopfield-type models in neural networks. Recently, they have been applied extensively in the rapidly growing field of probabilistic graphical models. These models are "*a marriage between probability theory and graph theory. They provide a natural tool for dealing with two problems that occur throughout applied mathematics and engineering – uncertainty and complexity – and in particular they are playing an increasingly important role in the design and analysis of machine learning algorithms…*" -- *Michael Jordan.*

One of the recent events related to Mean Field Methods was the NIPS 1999 Workshop on *Advanced Mean Field Methods* organized by Manfred Opper and David Saad. The book discussed here is a collection of articles presented at the Workshop, with a few other field-related papers. The book has 17 chapters, each of which presenting a separate piece of research and allowing for reading in isolation. Chapter 1, written by the Editors, outlines the book's scope and summarizes its content chapter by chapter. Chapters 2 and 3 provide a concise and comprehensive introduction to the main MFMs from the statistical physics point of view.

The next five chapters are generally devoted to TAP-type approaches. Chapters 4, 5 and 7 offer generalizations of the classical TAP method by deriving TAP equations under less restricted assumptions. In Chapter 4, the approach based on a Taylor series expansion of the marginal probabilities is proposed and applied to several arbitrary probability distributions, for example those arising in stochastic neural networks with asymmetric couplings or in sigmoid belief networks. In Chapter 5, the approach developed for highly coupled systems with intensive connectivity is experimentally tested against the classical TAP formulation in case of extensively connected systems, such as the Hopfield associative memories. Chapter 7 presents a novel TAP-type approach to models defined by quadratic interactions, which does not require specific assumptions on the randomness of couplings (typical for the TAP formulation). Chapters 6 and 8 are devoted to applications of TAP-like methods: respectively, to the problem of decoding corrupted codewords that were encoded by the sparse parity-check error-correcting codes and to the average case performance analysis of the stochastic batch mode learning algorithms for one layer perceptron.

In Chapter 9, saddle-point methods are presented as an alternative to TAP and variational methods for the problem of inference with Bayesian belief networks. Chapter 10 serves as a tutorial on modern variational methods, with emphasis put on presenting the ways in which the inference and estimation problems can be transformed into a suitable variational form. Chapter 11 provides a comprehensive description of the theory and practice of using variational methods for approximating inference and learning in the context of graphical models, with a special focus on Bayesian learning in probabilistic graphical models with hidden variables. In the next chapter, the Bayesian inference problem with hidden variables is considered from a different angle, with a certain recursive variational approach proposed and tested on a toy neural network model and a simple hidden Markov model.

Chapter 13 introduces a new approach to Bayesian inference problem in densely connected directed graphical models. The method for a class of directed, loopy models provides an approximate implementation of the Belief Propagation (BP) technique based on Fourier integral representation. In Chapter 14, the max-product algorithm for large, loopy probabilistic graphical models is analyzed. In the context of codes on graphs it is theoretically shown that if BP messages are properly attenuated then, assuming that the algorithm converges, the maximum *a posteriori* configuration of variables is reached. Chapter 15 aims at comparing the Mean Field (MF) approximation and the BP method in the context of inference approximation problem in Markov Random Fields. Based on several low-level vision examples, it is shown that BP typically outperforms a simple MF method – mainly due to a superior optimization technique which efficiently avoids local minima.

The last two chapters analyse Mean Field approximations from the information geometry point of view. In Chapter 16, properties of the naive MF approximation and those of the TAP approach are studied in the simple spin models, such as the Sherrington-Kirkpatrick model, or the Boltzmann machine. Chapter 17 presents a unified information-geometrical interpretation of the two MF approximation approaches – naive and perturbative – in the Boltzmann machine framework. Likewise, the information-geometrical approaches to the variational Bayes method and to variational approximation used in the EM algorithm are presented in this chapter.

In summary, the book provides an up-to-date overview of the theory and practice of the advanced Mean Field Methods. The authors' background spans a wide range of disciplines, from theoretical, applied, statistical and medical physics, mathematical statistic, applied engineering and computer science to neural computation and neuroscience. This diversity makes the book a valuable source of knowledge on recent advances in MFMs, as seen from different angles. Several authors aim at bridging the above disciplines by presenting the interrelation between different approaches. Further investigation of these interrelations may ultimately lead to the development of qualitatively new methods.

The book deserves strong recommendation to anyone interested in Mean Field Methods, albeit with a word of warning that, in order to fully appreciate its content, some background in statistical mathematics and physics may be necessary.

Warszawa, 10/02/2002