Zeitschrift: Helvetica Physica Acta
Band: 64 (1991)
Heft: 1

Artikel: Statistical mechanics of low autocorrelation skew-symmetric binary sequences

Autor: Groot, Claas de / Würtz, Diethelm

DOI: https://doi.org/10.5169/seals-116303

Helvetica Physica Acta
Vol. 64 (1991)

0018-0238/91/010086-06$1.50+0.20/0

Statistical Mechanics of Low Autocorrelation Skew-symmetric Binary Sequences

Claas de Groot[†] and *Diethelm Würtz*[†]
Interdisziplinäres Projektzentrum für Supercomputing
ETH-Zentrum, CH-8092 Zürich

(7. VI. 1990, revised 11. VII. 1990)

Abstract

Based on the ergodicity postulate of Golay and mean field like arguments of Bernasconi we investigate the statistical mechanics of low autocorrelation skew-symmetric binary sequences of finite length. We evaluate the density of states and thermodynamic quantities and compare them to results obtained from an exact enumeration algorithm. We find further confidence in the conjecture that the skew-symmetric sequence space provides good candidates for optimal autocorrelation sequences.

[†]Also at: Institut für Theoretische Physik, ETH-Hönggerberg, Zürich.

Introduction

Technical applications and theoretical interest led to the investigation of the properties of binary sequences with a given autocorrelation function [1]. In several communication engineering problems there is the need for sequences with a low autocorrelation with respect to this function. And this minimization problem has influenced the theoretical interest on the problem, because finding sequences with low autocorrelation turned out to be an extremely hard optimization problem. Golay investigated many properties of this autocorrelation function [1] and proposed his famous "ergodicity postulate". Motivated by his results Beenker [2], Bernasconi [3], and also we [4] used stochastic search algorithms (such as simulated annealing and evolution strategies) to find near optimal autocorrelation sequences. Since now more powerful computers are available "skew-symmetric" sequences can be enumerated allowing the investigation of the complete configuration space. A major step forward in the understanding of the numerical data and of the "structure" of low autocorrelation sequences was made by Bernasconi [3] from a physicist's point of view applying methods of statistical mechanics. His mean-field like arguments can also be extended to the skew-symmetric subspace. As a first approximation he argued that skew-symmetry is essentially breaking the sequence space into halves yielding reasonable results in the thermodynamic limit. Taking the structure of the skew-symmetric sequences explicitely into account one can derive analytic expressions for finite chains, which can directly be compared to numerical data. The results for the density of states, mean energy and specific heat are in good agreement with data from exact enumeration. From the analytical derivation of an optimal merit factor we find further confidence in the conjecture that the skew-symmetric sequence space provides good candidates for optimal autocorrelation sequences.

Problem Definition

The configuration space $\tilde{S}$ of the binary sequences is a N-dimensional Boolean hypercube $\tilde{S} = \{\pm 1\}^N$. The elements of the configuration space are the binary sequences $S \in \tilde{S}$, $S = (s_1, \ldots, s_N)$ of length N. We define the autocorrelation of these sequences as follows. Let

$$R_k = \sum_{i=1}^{N-k} s_i s_{i+k}, \qquad \text{and} \qquad E = \sum_{k=1}^{N-1} R_k^2. \tag{1}$$

We call E the *energy* of the sequence. Golay introduced the *merit factor* F of a sequence to be

$$F = \frac{N^2}{2E}. \tag{2}$$

This allows us to state the optimization problem. Find $S_o \in \tilde{S}$ such that $F_{max} = F(S_o) = \max_{S \in \tilde{S}} F(S)$. This can be alternatively formulated as a *minimum* for $E(S_o)$, from which we see that only sequences with low autocorrelation are of interest. The cost function E becomes equivalent to the energy of a one dimensional spin system with long range four spin interactions. This allows the application of the methods of statistical mechanics to the problem. Sequences of odd length $N = 2n - 1$ are called *skew-symmetric* if the s_i satisfy

$$s_{n-l} = (-1)^l s_{n+l}, \qquad l = 1, \ldots, n-1. \tag{3}$$

The skew-symmetric sequences are good candidates for large merit factors, because all R_k with k odd vanish as can be seen from the definition of skew-symmetry. The size of the phase space is greatly reduced so the skew-symmetric space can be exactly enumerated for quite large N and the approximate expressions of an statistical mechanics approach to the problem can be compared to these exact results [4].

Derivation of Thermodynamic Equations

In order to derive the statistical properties of the sequence space we write down the partition function Z of the problem in terms of the energy.

$$Z = \sum_{\{S\}} e^{-\beta E(S)} = 2^{(N+1)/2} \int_0^\infty dE\, \rho(E)\, e^{-\beta E}, \tag{4}$$

where $\rho(E)$ gives the density of states, and β denotes the inverse temperature $\beta = 1/T$. We now proceed by giving an approximate expression for the partition function which is based on the *ergodicity hypothesis* of Golay [1]. This hypothesis assumes that the R_k can be seen as independent random variables. If we take a close look at the skew-symmetric sequences we observe that we have $(N-k-1)/2$ terms of a sum which are at random ± 2, and a single term which is either $+1$ or -1 depending on k. If we put $y_k = R_k^2$ we get from this the continuous probability densities $p_k(y)$:

$$p_k(y) = \frac{1}{\sqrt{2\pi\sigma_k^2 y}} \exp\left(-\frac{y+1}{2\sigma_k^2}\right) \cosh\frac{\sqrt{y}}{\sigma_k^2}, \tag{5}$$

where $\sigma_k^2 = 2(N-k-1)$. The probability density $\rho(E)$ for the distribution of E is given by the convolution of these $p_k(y)$: $\rho(E) = p_2 * p_4 * \ldots * p_{N-1}$ with the Laplace transform $\tilde{\rho}(\beta) = \mathcal{L}\left[\rho(E)\right] = \int_0^\infty dE\, \rho(E)\, e^{-\beta E}$. If we compare this to the partition function we see that we can write

$$Z = 2^{(N+1)/2}\, \tilde{\rho}(\beta). \tag{6}$$

We have $\tilde{\rho}(\beta) = \prod_{k=2}^{N-1} \tilde{p}_k(\beta) = \prod_{k=2}^{N-1} e^{-\beta/(1+2\sigma_k^2\beta)}/(1+2\sigma_k^2\beta)^{1/2}$. If we insert σ_k^2 and renumber with respect to k we obtain

$$\tilde{\rho}(\beta) = \prod_{k=0}^{\frac{N-3}{2}} \frac{e^{-\beta/(1+8\beta k)}}{(1+8\beta k)^{1/2}}. \tag{7}$$

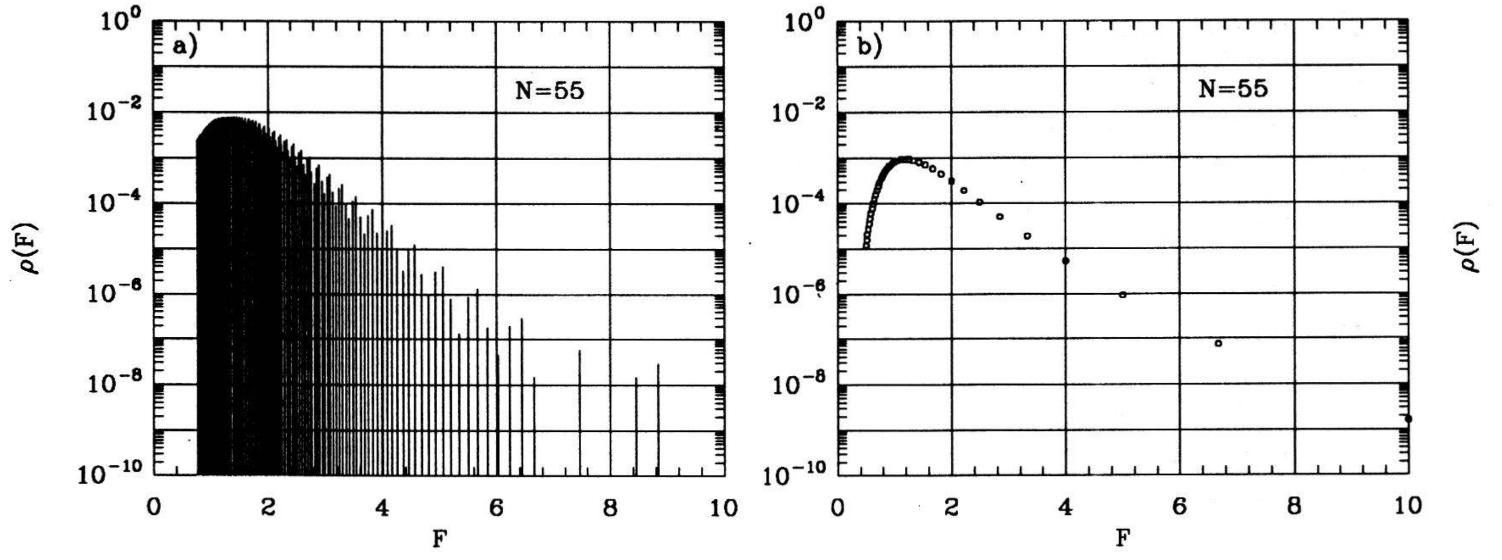

Fig. 1) a) Normalized distribution $\rho(F)$ of the merit factors F from complete enumeration of the skew-symmetric configuration space for $N = 55$, from ref. [4]. Note, the histogram is truncated for small values of F. b) In comparison the density of states as obtained from the numerical Laplace inversion of eq. (7).

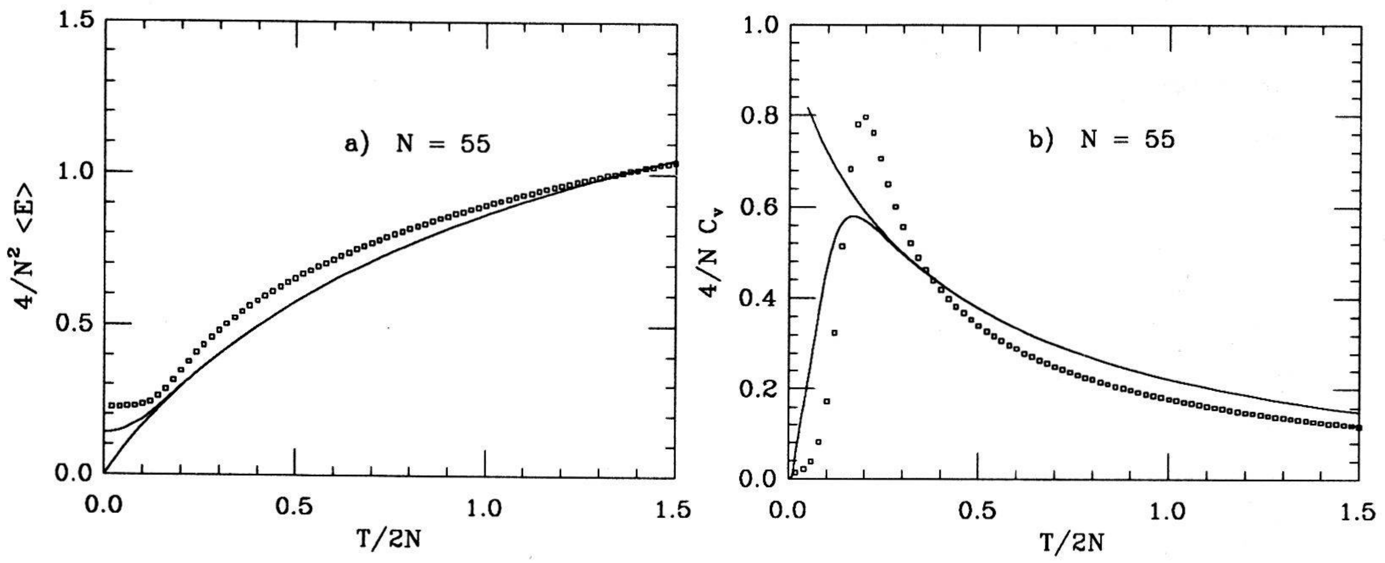

Fig. 2) Normalized mean energy (a) and specific heat (b) vs. normalized temperature obtained from exact enumeration of the skew-symmetric configuration space (square symbol). The curves represent the finite chain result. The finite value of the energy and the peak in the specific heat are due to the low energy cut off.

This is our main result from which we calculate the partition function to be

$$\ln Z(\beta) = \frac{N+1}{2}\ln 2 - \sum_{k=0}^{\frac{N-3}{2}} \left[\frac{1}{2}\ln(1+8\beta k) + \frac{\beta}{1+8\beta k}\right], \tag{8}$$

the mean energy $\langle E\rangle = -\partial \ln Z(\beta)/\partial\beta$, and finally the specific heat $C = -\beta^2\partial\langle E\rangle/\partial\beta$. In the case of the full sequence space the results for $N \to \infty$ were already given by Bernasconi. In comparison to (7) and (8) he obtained $\tilde{\rho}(\beta) = \prod_{k=1}^{N-1}(1+2\beta k)^{-1/2}$ which in the thermodynamic limit leads to $(-2/N)\ln Z(\beta) \approx (1+2\beta)\ln(1+2\beta)/2\beta - (1+2\ln 2)$.

Comparison to Exact Data

In ref. [4] we have performed the exact enumeration of the skew-symmetric configuration space for different chain lengths. As an example we present here the case $N = 55$. Our approximate equation for $\tilde{\rho}(\beta)$ can be Laplace transformed numerically to yield $\rho(F)$ of the skew-symmetric configuration space. Fig. 1a) shows the normalized histogram of the exact enumeration and Fig. 1b) compares to $\rho(F)$. These two figures show that the ergodicity hypothesis of Golay leads to a reasonable estimate of the configuration space especially with respect to sequences of low energy (corresponding to a large merit factor). From the exact enumeration of skew-symmetric autocorrelation sequences we have evaluated the mean energy and the specific heat, Fig. 2). The finite chain result for $N = 55$ is already close to the result in the thermodynamic limit for the whole temperature region considered. In the zero temperature limit the curves clearly show the failure of the mean field theory which arises from "ergodicity breaking". This ergodicity breaking can be investigated in some detail if we consider the lower bound of the partition function $Z > \exp(-\beta E_{min})$. In both cases, skew-symmetric and full sequence space, one obtains identical results in the thermodynamic limit $F_{max} \approx 12.32$. The dependence of F_{max} as a function of N is shown in Fig. 3) and again compared to data from exact enumeration [4]. We observe a slow convergence to the thermodynamic limit. Since F_{max} is independent of the reduction to the skew-symmetric sequence space we can consider this as further confidence in the conjecture that skew symmetric autocorrelation sequences are good candidates for optimal autocorrelation values. Furthermore we can now interpret E_{min} as a lower energy cut off in the density of states, i.e., we write for the corresponding F_{max}

$$\tilde{\rho}_{cut\ off}(F) = \begin{cases} \rho(F), & \text{if } F \leq F_{max}; \\ 0, & \text{otherwise.} \end{cases}$$

For finite chains this produces a nonzero ground state energy and a nondiverging specific heat with a peak similar to the results from exact enumeration, (see Fig. 2).

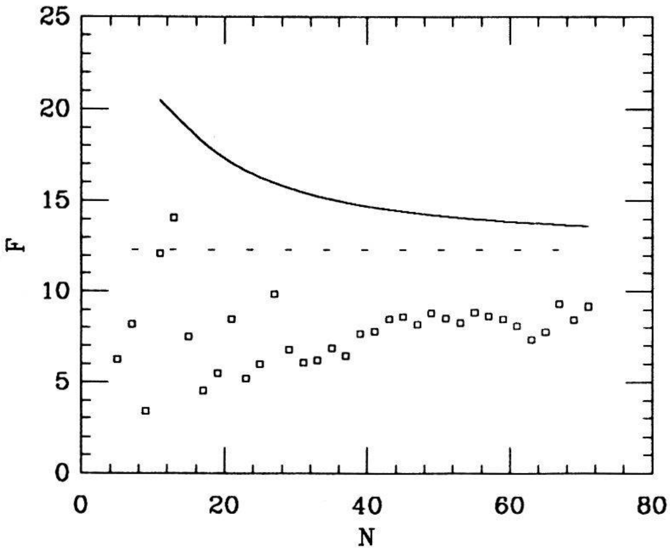

Fig. 3) Optimal merit factor of skew-symmetric sequences from exact enumeration (square symbol), compared to our approximation (upper curve). The dotted line represents the common value for $N \to \infty$.

Summary

We considered the statistical mechanics of low autocorrelation binary sequences in the skew-symmetric configuration space. We derived approximate expressions for the partition function, the mean energy, and the specific heat of the sequences which we compared to exact numerical results obtained by complete enumeration of the skew-symmetric configuration space.

Acknowledgement

One of us (CdG) acknowledges partial support from a grant given by the Swiss Bank Corporation.

References

[1] M.J.E. Golay, IEEE Trans. on Inf. Theory IT-18 (1972) 449, IT-23 (1977) 43 and IT-28 (1982) 543

[2] G.F.M. Beenker, T.A.C.M. Claasen, and P.W.C. Hermens, Philips J. of Research 40 (1985) 289

[3] J. Bernasconi, J. Physique 48 (1987) 559 and in: Proceedings of the Workshop on "Chaos and Complexity", Torino, World Scientific, Singapore, 1988

[4] C. de Groot, D. Würtz, and K.H. Hoffmann, IPS Research Report No. 89-09