Equivalence of the Ashkin-Teller and the four-state Potts-glass models of neural networks

D. Bollé*,† and P. Kozłowski†,‡
Instituut voor Theoretische Fysica, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium
(Received 10 April 2001; published 26 November 2001)

We show that for a particular choice of the coupling parameters the Ashkin-Teller spin-glass neural network model with the Hebb learning rule and one condensed pattern yields the same thermodynamic properties as the four-state anisotropic Potts-glass neural network model. This equivalence is not seen at the level of the Hamiltonians.

DOI: 10.1103/PhysRevE.64.067102 PACS number(s): 02.50.−r, 05.50.+q, 64.60.Cn, 84.35.+i

It is well known that the classical Ashkin-Teller (AT) model is a generalization of the Ising, the four-state clock and the four-state Potts models. This can be easily seen already at the level of the Hamiltonian, especially when one rewrites the Hamiltonian of the AT model [1] using two Ising spins located at each site of the lattice interacting via two- and four-spin couplings [2].

For spin-glass systems similar observations can be made (see Ref. [3] and references therein). The AT spin-glass Hamiltonian contains as particular limits, for certain bond realizations, both the four-state clock spin glass and the four-state Potts-glass Hamiltonians.

Concerning neural network models, the situation is more complicated. It is straightforward to see at the level of the Hamiltonian that for two and one of the coupling strengths, respectively, taken to be zero, the AT neural network model [4,5] is equivalent to the Hopfield model [6] and the four-state clock neural network model, respectively, [7]. On the contrary, the possible relation with the four-state Potts neural network models existing in the literature [8,9] is, at first sight, unclear. However, since we discovered in the study of the thermodynamic and retrieval properties of the AT neural network [4,5] for equal coupling strengths some resemblance to the properties of the Potts-glass neural network [8,10], we expect that a relation with the latter does exist. To investigate this relation is the purpose of this paper.

The AT neural network with the Hebb learning rule is described by the following infinite-range Hamiltonian:

$$\mathcal{H}_{AT}=-\frac{1}{2N}\sum_{\mu=1}^{p}\sum_{(i,j)=1}^{N}[J_1\xi_i^\mu\xi_j^\mu s_i s_j+J_2\eta_i^\mu\eta_j^\mu\sigma_i\sigma_j + J_3\xi_i^\mu\eta_i^\mu\xi_j^\mu\eta_j^\mu s_i\sigma_i s_j\sigma_j], \tag{1}$$

with the two types of Ising neurons $s_i,\sigma_i, i=1,\ldots,N$ describing the state of the network. In this model storage and retrieval of the patterns $\{\xi_i^\mu\},\{\eta_i^\mu\}, \mu=1,\ldots,p$ are studied. The patterns are randomly chosen configurations of the network. Based upon our observations mentioned above we take equal coupling strengths $J_1=J_2=J_3=1$ in the sequel.

Two Potts-glass neural networks with Hebb learning have been studied in the literature. Considering four Potts states the first network [8] is described by the Hamiltonian

$$\mathcal{H}_K=-\frac{1}{2N}\sum_{\mu=1}^{p}\sum_{(i,j)=1}^{N}\frac{1}{16}(\mathbf{u}_i\cdot\boldsymbol{\psi}_i^\mu)(\mathbf{u}_j\cdot\boldsymbol{\psi}_j^\mu), \tag{2}$$

while the second one is given by [9]

$$\mathcal{H}_{VZ}=-\frac{1}{2N}\sum_{\mu=1}^{p}\sum_{(i,j)=1}^{N}\frac{1}{16}(\mathbf{u}_i\cdot\mathbf{u}_j)(\boldsymbol{\psi}_i^\mu\cdot\boldsymbol{\psi}_j^\mu), \tag{3}$$

where $\mathbf{u}_i$ and $\boldsymbol{\psi}_i$ are state and pattern vectors taken from the set of four-dimensional vectors $\mathbf{v}=\{\mathbf{v}^{(l)}\}$ with components $v_k^{(l)}=4\delta_{kl}-1$ for $l,k=1,2,3,4$. The main difference between the two models is that in the first, anisotropic model precisely one specific Potts state is favored at each site, while in the second, isotropic model the fact whether or not two neurons are in the same state is important.

Two models can be equivalent at the level of the Hamiltonian or at the level of the free energy. It is clear that for the Hamiltonians of the models we have introduced above there are one state and p pattern variables associated with each of the N sites of the network. Thus the Hamiltonians can be written in the form

$$\mathcal{H}_{mod}=-\frac{1}{2N}\sum_{\mu=1}^{p}\sum_{(i,j)}H_{mod}(C_{ij}^\mu), \tag{4}$$

where mod denotes AT, K or VZ. The energy of the interaction between two sites is a sum over patterns of $H_{mod}(C_{ij}^\mu)$ and depends on the state-pattern configuration C_{ij}^μ of sites i and j. Hence, it is enough to compare the $H_{mod}(C_{ij}^\mu)$. In the case of four state models we are considering here, all possible values of $H_{mod}(C_{ij}^\mu)$ can be written in the form of a

*Email address: desire.bolle@fys.kuleuven.ac.be
†Also at Interdisciplinair Centrum voor Neurale Netwerken, Katholieke Universiteit Leuven, Belgium.
‡Email address: piotr.kozlowski@fys.kuleuven.ac.be

16×16 matrix (16 state-pattern configurations for a given site). For the sake of easy comparison we write down these matrices explicitly

$$H_K=\begin{pmatrix}
9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 1 & \bar{3}\\
9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9
\end{pmatrix}, \tag{5}$$

$$H_{VZ}=\begin{pmatrix}
9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1\\
\bar{3} & 9 & \bar{3} & \bar{3} & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1\\
\bar{3} & \bar{3} & 9 & \bar{3} & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1\\
\bar{3} & \bar{3} & \bar{3} & 9 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1\\
1 & \bar{3} & 1 & 1 & \bar{3} & 9 & \bar{3} & \bar{3} & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1\\
1 & 1 & \bar{3} & 1 & \bar{3} & \bar{3} & 9 & \bar{3} & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1\\
1 & 1 & 1 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 9 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 1 & 1 & 1\\
1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & \bar{3} & 9 & \bar{3} & \bar{3} & 1 & \bar{3} & 1 & 1\\
1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & \bar{3} & \bar{3} & 9 & \bar{3} & 1 & 1 & \bar{3} & 1\\
1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9 & 1 & 1 & 1 & \bar{3}\\
\bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & 9 & \bar{3} & \bar{3} & \bar{3}\\
1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & \bar{3} & 9 & \bar{3} & \bar{3}\\
1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & \bar{3} & \bar{3} & 9 & \bar{3}\\
1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & 1 & 1 & 1 & \bar{3} & \bar{3} & \bar{3} & \bar{3} & 9
\end{pmatrix}, \tag{6}$$

$$H_{AT}=\begin{pmatrix}
3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3\\
\bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1}\\
\bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1}\\
\bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1}\\
\bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1}\\
3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3\\
\bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1}\\
\bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1}\\
\bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1}\\
\bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1}\\
3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3\\
\bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1}\\
\bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1}\\
\bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1}\\
\bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & 3 & \bar{1}\\
3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3 & \bar{1} & \bar{1} & \bar{1} & \bar{1} & 3
\end{pmatrix}, \tag{7}$$

where we have used the standard notation $\overline{c}=-c$.

One can see that H_{AT} differs from H_K and H_{VZ} by the number of energy levels and by their position in the matrix. We have not found a transformation of the Hamiltonians removing this difference. Therefore, we conclude that at the level of the Hamiltonian the AT neural network is not equivalent to any of the two four-state Potts models.

In order to find out whether a possible equivalence exists on the level of the free energies we start from the model with Hamiltonian $\mathcal{H}_K$ because the matrix H_K has the same global symmetry as H_{AT}. We note that the symmetry of 4×4 blocks in H_K and H_{AT} is different. It is a consequence of the fact that $\mathcal{H}_{AT}$ is invariant under inversion of all the spins, while $\mathcal{H}_K$ is not invariant under any permutation of the state variables. The Hamiltonian $\mathcal{H}_{VZ}$ on the contrary is completely invariant under any permutation of those variables.

As remarked in Ref. [8] H_K can be rewritten using two different types of Ising spins

$$\begin{aligned}H_K(C_{ij}^{\mu})=&(s_i\xi_i^{\mu}+\sigma_i\eta_i^{\mu}+s_i\xi_i^{\mu}\sigma_i\eta_i^{\mu})\\ &\times(s_j\xi_j^{\mu}+\sigma_j\eta_j^{\mu}+s_j\xi_j^{\mu}\sigma_j\eta_j^{\mu}).\end{aligned} \tag{8}$$

Applying the usual replica method [11] to calculate the quenched average over an arbitrary number of patterns, chosen to be independent identically distributed random variables taking the values $+1$ and -1 with equal probability, the free energy density can be written in the thermodynamic limit $N\to\infty$ in the form $f=\lim_{n\to 0}=\phi_n/n$ with ϕ_n the replicated free energy. For the model at hand, assuming at first that there is only one pattern condensed, say $\mu=1$, we get

$$\phi_{n,K}=\frac{1}{2}\sum_{a=1}^{n} m_a'^2+\frac{9}{2}\alpha'\beta'\sum_{a<b} r'_{ab}q'_{ab}+\frac{9\alpha'}{4\beta'}\mathrm{Tr}\ln\Lambda$$
$$-\frac{1}{\beta'}\ln\left\langle\left\langle\sum_{\{s,\sigma\}}\exp\left\{\beta'\sum_a m'_a b_a+\frac{9}{2}\alpha'\beta'^2\sum_{a<b} r'_{ab}b_{ab}\right\}\right\rangle\right\rangle, \tag{9}$$

where we have dropped the index 1 and where

$$b_a=s^a\xi+\sigma^a\eta+s^a\sigma^a\xi\eta,$$
$$b_{ab}=s^as^b+\sigma^a\sigma^b+s^a\sigma^as^b\sigma^b,$$
$$\Lambda_{ab}=(1-3\beta')\delta_{ab}-\beta' q'_{ab},\qquad a,b=1,\ldots,n.$$

The brackets $\langle\langle\cdots\rangle\rangle$ indicate the average over the condensed pattern. As usual β' is the inverse temperature, α' the capacity defined as the number of patterns per number of couplings per spin, i.e., $\alpha'=2p/9N$, $\Sigma_{\{s,\sigma\}}$ denotes the sum over all configurations at one site and $\Sigma_{a<b}$ denotes the sum over pairs of different replicas $a<b$. Finally, the set of order parameters is given by

$$m'^{\mu}_a=\left\langle\left\langle\frac{1}{N}\sum_{i=1}^{N}\langle s_i^a\rangle\xi_i^\mu+\langle\sigma_i^a\rangle\eta_i^\mu+\langle s_i^a\sigma_i^a\rangle\xi_i^\mu\eta_i^\mu\right\rangle\right\rangle,$$
$$q'_{ab}=\left\langle\left\langle\frac{1}{N}\sum_{i=1}^{N}\langle s_i^a\rangle\langle s_i^b\rangle+\langle\sigma_i^a\rangle\langle\sigma_i^b\rangle+\langle s_i^a\sigma_i^a\rangle\langle s_i^b\sigma_i^b\rangle\right\rangle\right\rangle,$$
$$r'_{ab}=\frac{2}{9\alpha'}\sum_{\mu>1}^{p}\langle\langle m'^{\mu}_a m'^{\mu}_b\rangle\rangle$$

where $\langle\cdots\rangle$ denotes the thermal average and the brackets $\langle\langle\cdots\rangle\rangle$ now indicate the average over all patterns.

The order parameters m'_a, q'_{ab}, and r'_{ab}, and the inverse temperature β' can be rescaled in such a way that the resulting replicated free energy density Eq. (9) satisfies $\phi_{n,K}=3\phi_{n,AT}$, with $\phi_{n,AT}$ the replicated free energy density for the AT neural network. Hereby we have taken into account that for the AT neural network model with equal coupling strengths and one condensed pattern, the nine order parameters $(m_a^\nu,q_{ab}^\nu,r_{ab}^\nu)$ with $\nu=1,2,3$ referring to ξ, η, and $\xi\eta$ reduce to three, i.e., (m_a,q_{ab},r_{ab}), where a reference to a specific type of pattern is now irrelevant. This is due to the fact that for this AT model only states satisfying $m_a^\nu=m_a,q_{ab}^\nu=q_{ab},r_{ab}^\nu=r_{ab},\nu=1,2,3$, i.e., so-called simple states, minimize the free energy. For the replica symmetric anzatz this property of the simple states has been shown in Ref. [5] to be related to taking the quenched average over just one condensed pattern. Since patterns do not carry replica indices, we assume that it is also valid in the fully replicated case.

The proper rescaling is the following:

$$m'_a=3m_a,\qquad q'_{ab}=3q_{ab},\qquad r'_{ab}=3r_{ab},$$
$$\beta'=\frac{1}{3}\beta,\qquad \alpha'=\alpha. \tag{10}$$

Next, assuming more than one condensed pattern, the order parameters $\mathbf{m}_a^\nu$ get a vector character in μ and stable states for which the $\mathbf{m}_a^\nu$ are different for different ν occur. This no longer allows for a reduction of the order parameters. We remark that these states have a bigger replica symmetric free energy than the one for the simple states and, hence, they play a minor role in the thermodynamics of the model. Nevertheless, they do destroy the thermodynamic equivalence with the Potts model.

In brief, we conclude that the AT neural network with equal coupling strengths and one condensed pattern is thermodynamically equivalent to the four-state anisotropic Potts model studied in Ref. [8], in spite of the different Hamiltonians. In fact, the AT Hamiltonian (1) does not contain three-spin interaction terms present in Eq. (8). We have demonstrated this thermodynamic equivalence by rewriting the Hamiltonian of the Potts model using two different types of Ising variables and calculating the replicated free energy.

These results clarify the resemblance found before [4,5,12] of the thermodynamic properties of the AT and the Potts neural networks. Furthermore, they imply that the four-state Potts model described by $\mathcal{H}_{VZ}$ is thermodynamically equivalent to the AT neural network model with one condensed pattern only in the limit of low loading, i.e., for $\alpha=0$, or at zero temperature assuming replica symmetry, where we know that the fixed-point equations for the two four-state Potts models are the same, as shown in Ref. [9].

[1] J. Ashkin and E. Teller, Phys. Rev. **64**, 178 (1943).
[2] C. Fan, Phys. Lett. **39A**, 136 (1972).
[3] F.D. Nobre and D. Sherrington, J. Phys. A **26**, 4539 (1993).
[4] D. Bollé and P. Kozłowski, J. Phys. A **31**, 6319 (1998).
[5] D. Bollé and P. Kozłowski, J. Phys. A **32**, 8577 (1999).
[6] J.J. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982); **81**, 3088 (1984).
[7] J. Cook, J. Phys. A **22**, 2057 (1989).
[8] I. Kanter, Phys. Rev. A **37**, 2739 (1988).
[9] H. Vogt and A. Zippelius, J. Phys. A **25**, 2209 (1992).
[10] D. Bollé, P. Dupont, and J. Huyghebaert, Phys. Rev. A **45**, 4194 (1992).
[11] M. Mézard, G. Parisi, and M.A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[12] D. Bollé and P. Kozłowski, Phys. Rev. E **64**, 011915 (2001).