

Statistical Recovery of the Hamiltonian

Vinit Kumar Singh^{1,*} and Jung Hoon Han^{2,†}

¹*Department of Physics, Indian Institute of Technology, Kharagpur 721302, India*

²*Department of Physics, Sungkyunkwan University, Suwon 16419, Korea*

(Dated: June 28, 2018)

Suppose we had a Hamiltonian

$$H = \sum_{ij} J_{ij} \sigma_i \sigma_j \quad (1)$$

with each $\sigma_i = \pm$ serving as a binary variable and the interaction J_{ij} extended over an arbitrary pair of sites ij of the finite $L \times L$ lattice. We are given equilibrium configurations for such Hamiltonian through, say, Monte Carlo simulation. Given such equilibrium configurations at one's disposal, is it possible to recover the original Hamiltonian from which those configuration came? In other words, can we accurately recover all J_{ij} from analysis of the configurations alone without any prior knowledge about the Hamiltonian? And can we use some machine learning algorithm in making such recovery?

This is in part inspired by the recent proposal[?] that a single eigenstate contains enough information to allow the faithful reconstruction of the original quantum Hamiltonian. I am asking for an analogue question in the classical equilibrium mechanics. One method that is guaranteed to work is to train the ML on the energy of each configuration. Since there are at most $O(N^2)$ number of parameters in any given Hamiltonian (N is the number of lattice sites), training set consisting of more than N^2 number of configurations will eventually get all the parameters of the Hamiltonian right.

This is not quite legal, though, since from the outset the assumption was that we do not know what the microscopic Hamiltonian was. We could instead use information gathered from the equilibrium configurations, because they were given to us from the outset. The way the information about the configuration is gathered in equilibrium statistical mechanics is by measuring the correlation functions

$$C_{ij}^{(O)} = \langle O_i O_j \rangle - \langle O_i \rangle \langle O_j \rangle. \quad (2)$$

Suppose the as-yet unknown Hamiltonian has the most general form

$$H = \sum_{1 \leq i, j \leq N} h_{ij} \sigma_i \sigma_j, \quad (3)$$

where N labels the total number of degrees of freedom. For the Ising model N also equals the number of sites in the lattice, and for the Heisenberg model N will be three times the lattice sites. A well-known theorem of equilibrium statistical mechanics says that

$$\langle H^2 \rangle - \langle H \rangle^2 = T^2 C(T) \quad (4)$$

where $C(T)$ refers to the specific heat. In terms of the general Hamiltonian (3), we get

$$\langle H^2 \rangle - \langle H \rangle^2 = \sum_{ij, kl} h_{ij} C_{ij, kl} h_{kl} \quad (5)$$

where the four-point correlation matrix $C_{ij, kl}$ would be

$$C_{ij, kl} = \langle \sigma_i \sigma_j \sigma_k \sigma_l \rangle - \langle \sigma_i \sigma_j \rangle \langle \sigma_k \sigma_l \rangle = C_{kl, ij}. \quad (6)$$

One can also treat $A = (ij)$ and $B = (kl)$ as one-dimensional array of size N^2 , and write the condition (5) as

$$\sum_{A, B} h_A C_{AB} h_B = \mathbf{h}^T \mathbf{C} \mathbf{h} = T^2 C(T). \quad (7)$$

The spectral decomposition of the Hermitian matrix $\mathbf{C} = \sum_{\alpha} \lambda_{\alpha} \mathbf{x}_{\alpha}^T \mathbf{x}_{\alpha}$ as

$$\sum_{\alpha} \lambda_{\alpha} (\mathbf{h}^T \mathbf{x}_{\alpha}) (\mathbf{x}_{\alpha}^T \mathbf{h}) = \sum_{\alpha} \lambda_{\alpha} (\mathbf{x}_{\alpha}^T \mathbf{h})^2 = T^2 C(T). \quad (8)$$

Positivity on the right-hand side is guaranteed. In order to reproduce 0 we must set all terms $\lambda_{\alpha} (\mathbf{x}_{\alpha}^T \mathbf{h})^2 = 0$ for every α , requiring either $\lambda_{\alpha} = 0$ or $\mathbf{x}_{\alpha}^T \mathbf{h} = 0$. What happens at $T = 0$ is that the correlation matrix \mathbf{C} vanishes identically.

Suppose the heat capacity $C(T)$ is known along with the four-point correlation function, both as a function of temperature T . In a translationally invariant system we can write $h_{ij} = h_{i, i+r} \equiv h_r$ and $h_{kl} = h_{k, k+r'} \equiv h_{r'}$ with $r, r' \in L^2$ both spanning the lattice. Carrying out the partial sum over i, k gives

$$\begin{aligned} \sum_{i, r, k, r'} h_r \left(\sum_{i, k} C_{i, i+r; k, k+r'} \right) h_{r'} \\ = \sum_{r, r'} h_r C_{r, r'} h_{r'} = T^2 C(T). \end{aligned} \quad (9)$$

Having the complete knowledge of $C_{r, r'}$ and $C(T)$ for all temperatures, figuring out h_r appears to be a problem of least-square fitting. Since $C_{r, r'} = C_{r' r}$, one can further

apply spectral decomposition to the reduced four-point correlation function

$$C_{r,r'} = \sum_{\alpha} \lambda_{\alpha} [I^{(\alpha)}]_r [I^{(\alpha)}]_{r'} \quad (10)$$

and write

$$\begin{aligned} T^2 C(T) &= \sum_{\alpha} \lambda_{\alpha} \sum_r (h_r [I^{(\alpha)}]_r) (h_{r'} [I^{(\alpha)}]_{r'}) \\ &= \sum_{\alpha} \lambda_{\alpha} \left(\sum_r h_r [I^{(\alpha)}]_r \right)^2. \end{aligned} \quad (11)$$

To test the feasibility of this idea, we run MC calculation of the 1D Ising model and obtain the specific curve $C(T)$ and the reduced four-point correlation function $C_{r,r'}(T)$ over a wide range range. Then an error function is defined

$$\begin{aligned} F[\mathbf{h}] &= \sum_T [D(T)]^2, \\ D(T) &= \sum_{r,r'} h_r C_{r,r'}(T) h_{r'} - T^2 C(T). \end{aligned} \quad (12)$$

The gradient of the error function gives

$$\frac{\partial F}{\partial h_r} = 2 \sum_T D \frac{\partial D}{\partial \mathbf{h}} = 4 \sum_T D \sum_{r'} C_{r,r'} h_{r'}. \quad (13)$$

The gradient descent method moves the initial guess of the configuration \mathbf{h} to $\mathbf{h} - \epsilon \nabla_{\mathbf{h}} F$. To avoid redundancy, it is important to keep the range of interaction $0 \leq r \leq L/2$ for even lattice and $0 \leq r \leq [L/2] + 1$ for odd lattice.

The procedure goes as follows. (1) Work out the four-point correlation function $C_{ij,kl}$ and reduce it to $C_{r,r'}$ by summing over i, k . This will be done for every temperature T where the four-point and the specific heat data are available. (2) Do spectral decomposition of $C_{r,r'}$ and obtain the leading eigenvalues and eigenvectors. (3) The

desired Hamiltonian vector h_r is the one that satisfies Eq. (11) as accurately as possible. This can be done by least-square fitting or gradient descent method. This way, we recover the original Hamiltonian $H = \sum_{i,r} h_r \sigma_i \sigma_{i+r}$.

I. APPENDIX: ANALYTIC SOLUTION

Two- and four-point correlation functions of the 1D Ising model can be found by the transfer matrix method. For the two-point case we find

$$\langle \sigma_i \sigma_j \rangle = \frac{\lambda_+^{N-r} \lambda_-^r + \lambda_-^{N-r} \lambda_+^r}{\lambda_+^N + \lambda_-^N} \quad (14)$$

on a periodic chain of length N . The symbols are $\lambda_+ = \cosh K$, $\lambda_- = \sinh K$ ($K = J/T$), and $r = |j - i|$ is the separation. Four-point correlations are

$$\langle \sigma_i \sigma_{i+r} \sigma_j \sigma_{j+r'} \rangle = \frac{\lambda_+^{N-r-r'} \lambda_-^{r+r'} + \lambda_-^{N-r-r'} \lambda_+^{r+r'}}{\lambda_+^N + \lambda_-^N}, \quad (15)$$

assuming $i \leq i+r \leq j \leq j+r'$. Note that any product of four operators $\sigma_i \sigma_j \sigma_k \sigma_l$ can be re-arranged in ascending order $i \leq j \leq k \leq l$. For large N we can approximate

$$\begin{aligned} \langle \sigma_i \sigma_{i+r} \rangle &\approx (\tanh K)^r, \\ \langle \sigma_i \sigma_{i+r} \sigma_j \sigma_{j+r'} \rangle &\approx (\tanh K)^{r+r'}. \end{aligned} \quad (16)$$

Under this approximation we have $C_{ij;kl} = C_{ji;kl} = C_{ij;lk} = C_{ji;lk} = 0$ and $C_{kl;ij} = C_{lk;ij} = C_{kl;ji} = C_{lk;ji} = 0$ if $i \leq j \leq k \leq l$. Nonzero correlations are found for other sequences of sites. For instance when $i \leq k \leq j \leq l$, we get

$$C_{ij;kl} = x^{|k-i|+|l-j|} - x^{|j-i|+|l-k|}, \quad (17)$$

and so on.

* Electronic address: vinit Singh911@gmail.com

† Electronic address: hanjh@skku.edu

¹ Xiao-Liang Qi and Daniel Ranard, arXiv:1712.01850 (2017).