

Different phases and phase transition points in hidden Markov models

Ruben Poghosyan¹, Vardan Bardakhchyan^{2,*} and David B. Saakian^{2,†}

¹*School of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

²*A.I. Alikhanyan National Science Laboratory (Yerevan Physics Institute) Foundation, 2 Alikhanian Brothers St., Yerevan 375036, Armenia*



(Received 14 November 2025; accepted 7 May 2026; published 14 July 2026)

We are looking for different phases in hidden Markov models (HMMs). By analyzing the stationary solution of the master equation, we discover different situations with localized or delocalized distributions for the stationary distributions of the master equation. We hypothesize that these phases are related to the relaxation dynamics of the HMM ensemble. We have derived precise phase transition points. We test whether the behavior of the HMMs and their associated numerical algorithms change near the transition points. We find that the accuracy of the Viterbi and Baum-Welch algorithms sometimes is worse near the phase transition, which contradicts the edge of chaos hypothesis.

DOI: [10.1103/vjy9-4r1f](https://doi.org/10.1103/vjy9-4r1f)

I. INTRODUCTION

Hidden Markov models (HMMs) [1] are a key tool in applied mathematics and data science, from communication theory [2] and speech recognition [3–7] to financial mathematics [8]. They are still widely used, even after the boom of deep neural networks in data science [9]. The main difference from conventional Markov models is that HMMs use a hidden state. It considers the entropy of a long sequence of observed signals describing a hidden Markov process. The latter belongs to the field of statistical physics with frozen disorder [10]. In Ref. [11], following the idea first introduced in Ref. [10], we formulated a problem of computing entropy similar to the problem of computing the product of random matrices [12] in the case of correlated random matrices [13]. This problem is one of the tools of interdisciplinary research and has recently been widely applied in biological evolution [14–20]. Recently, we investigated the entropy of hidden Markov models [21] using a functional equation for computing entropy (the maximum Lyapunov index). In Ref. [22], we deduced that there exists an infinite series of certain phase transitions in the product of correlated random matrices. Later, in Ref. [23], we investigated a two-qubit model, considering a similar functional equation.

Here, we analyze phase transitions in HMMs and the accuracy of the Viterbi and Baum-Welch (BW) algorithms near these phase transitions. BW is an expectation-maximization algorithm applied to HMMs and used to estimate model parameters. In general, it requires that all probabilities (of the initial state, transition, and observation) be unknown. These

parameters are estimated based on the observed sequence of outcomes. In contrast, in the Viterbi algorithm, all parameters must be known, and the goal is to estimate the most probable path of hidden states, given the observed sequence.

Since the scope of HMMs is much broader than that of correlated random matrix products, our result should have broader applicability.

The mathematical structure of an HMM can be represented as follows. There are Markov transition probabilities, where q_{ij} denotes the probability of a hidden state transitioning from state j to state i during one period. The probability conditions for two hidden states ($i, j \in \{1, 2\}$) are $q_{11} + q_{21} = 1$, $q_{22} + q_{12} = 1$. For d hidden states, we have a matrix of order $d \times d$ with the constraints $\sum_{j=1}^d q_{ji} = 1$.

We denote the observation probabilities by p_{xi} : the probability of observing state x when the hidden process is in state i .

We define a matrix T^x with elements $T_{ij}^x = p_{xi}q_{ji}$ (the probability that the hidden process moved from i to j and we observed x).

We have the following probability of observing the sequence $x_n \dots x_1$.

$$P = \langle \text{unit} | T \langle st \rangle$$

$$T = T^{x_n} \dots T^{x_2} T^{x_1}, \quad (1)$$

where $|st\rangle$ is the initial vector and $|\text{unit}\rangle = (1, \dots, 1)$ (a unit vector of length d). In what follows, we restrict ourselves to two hidden and two observed outcomes. The coordinates of the vector z are the probabilities of states 1 and 2. Then, (z_n^1, z_n^2) are the probabilities that the state will have values 1 and 2 for given observations $x_n \dots x_0$.

We define a sequence of vectors $|\mathbf{z}_n\rangle = (z_n^1, z_n^2)$.

We initially choose the vector $|st\rangle \equiv (z_0^1, z_0^2)$.

Given a vector $|\mathbf{z}_n\rangle$ at some round n , we define the probability of choosing x at the next round as

$$P_x(|\mathbf{z}_n\rangle) = \frac{\langle \text{unit} | T^x |\mathbf{z}_n\rangle}{\sum_y \langle \text{unit} | T^y |\mathbf{z}_n\rangle} \quad (2)$$

*Contact author: vardanbardakhchyan@gmail.com

†Contact author: saakian@yerphi.am

In Sec. II we investigate the functional equation associated with a HMM, and in Sec. III we are looking for the phase transitions on the basis of the steady-state solution of the functional equation. We discovered three phases in the HMM and derived exact phase transition points. In Sec. IV we are looking at high dimensions and suggest an alternative expression for the phase transition points, which exactly coincide with the rigorous results of Sec. III for the $d = 2$ case. In the Appendix we work with the Viterbi and BW algorithms and confirm the significant impact of transition points on numerical algorithms: the accuracy of both algorithms are becoming worse near the phase transition points.

II. THE MASTER EQUATION IN POLAR COORDINATES

We can look at the polar angle,

$$|\mathbf{z}_n\rangle = [\cos(\alpha_n), \sin(\alpha_n)]h_n. \quad (3)$$

The meaning of the vector $\mathbf{z}_n$ is that it gives a relative probability of having up and down Markovian states in case of observed sequence of x 's.

We have matrices $T^1 = A$, $T^2 = B$.

$$A = \begin{pmatrix} p_{11}q_{11} & p_{11}q_{12} \\ p_{12}q_{21} & p_{12}q_{22} \end{pmatrix},$$

$$B = \begin{pmatrix} p_{21}q_{11} & p_{21}q_{12} \\ p_{22}q_{21} & p_{22}q_{22} \end{pmatrix}. \quad (4)$$

If the norm of the vector z_n grows exponentially with the coefficient λ ,

$$\langle \mathbf{z}_n | \mathbf{z}_n \rangle \sim \exp[-2\lambda n], \quad (5)$$

then λ is the mean entropy of the sequence of x 's [11].

Acting to the vector with the angle α , the matrix A moves it to the point $f^{-1}(\alpha)$,

$$f_1^{-1}(\alpha) = \arccos \left(\frac{[A_{11} \cos(\alpha) + A_{21} \sin(\alpha)]}{\sqrt{[A_{11} \cos(\alpha) + A_{21} \sin(\alpha)]^2 + [A_{12} \cos(\alpha) + A_{22} \sin(\alpha)]^2}} \right),$$

$$f_2^{-1}(\alpha) = \arccos \left(\frac{[B_{11} \cos(\alpha) + B_{21} \sin(\alpha)]}{\sqrt{[B_{11} \cos(\alpha) + B_{21} \sin(\alpha)]^2 + [B_{12} \cos(\alpha) + B_{22} \sin(\alpha)]^2}} \right). \quad (6)$$

Then, we define the probability of choosing the type $x = 1$ of the observation matrix at the next round,

$$\pi_1(\alpha) = \frac{(A_{11} + A_{21}) \cos(\alpha) + (A_{21} + A_{22}) \sin(\alpha)}{(A_{11} + A_{21} + B_{11} + B_{21}) \cos(\alpha) + (A_{21} + B_{22} + B_{21} + A_{22}) \sin(\alpha)} \quad (7)$$

and $\pi_2(\alpha) = 1 - \pi_1(\alpha)$. We have the following equation for the steady-state distribution:

$$\begin{aligned} \rho_1(\alpha, n+1) &= \pi_1[f_1(\alpha)]\{\rho_1[f_1(\alpha), n] \\ &\quad + \rho_2[f_1(\alpha), n]\}f'_1(\alpha) \\ \hat{\rho}_2(h, n+1) &= \pi_2[f_1(\alpha)]\{\rho_1[f_2(\alpha), n] \\ &\quad + \rho_2[f_2(\alpha), n]\}f'_2(\alpha), \end{aligned} \quad (8)$$

and while we will not concentrate on it, still it is useful to state that the Lyapunov index λ here is defined as

$$\begin{aligned} \lambda &= \int d\alpha \left[\rho_1(\alpha)g_1(\alpha) + \int d\alpha \rho_2(\alpha)g_2(\alpha) \right] \\ g_1(\alpha) &= \frac{1}{2} \ln\{[A_{11} \cos(\alpha) + A_{21} \sin(\alpha)]^2 \\ &\quad + [A_{12} \cos(\alpha) + A_{22} \sin(\alpha)]^2\} \\ g_2(\alpha) &= \frac{1}{2} \ln\{[B_{11} \cos(\alpha) + B_{21} \sin(\alpha)]^2 \\ &\quad + [B_{12} \cos(\alpha) + B_{22} \sin(\alpha)]^2\}. \end{aligned} \quad (9)$$

In Eq. (9) ρ_1 is the probability distribution via α for the first choice of the matrix and ρ_2 for the second choice of the matrix.

III. THE CRITICAL POINT

First, we should identify the range of α values $[\alpha_1, \alpha_2]$, where the solution of Eq. (8) is defined. We define α_1, α_2 as

the angles of the eigenvectors, related to the maximal eigenvalues:

$$\alpha_1 = f_1(\alpha_1), \quad \alpha_2 = f_2(\alpha_2). \quad (10)$$

The range of α where the functions ρ_1, ρ_2 are defined and can be nonzero is $[\alpha_1, \alpha_2]$. For the higher dimensions, with more than two choices of the matrices T_x , the definition of the domain of parameters for the steady-state solution of the master equation is a nontrivial problem.

For the case of Eq. (8), we approximate

$$\begin{aligned} f_1(\alpha) &\approx \alpha_1 + k_1(\alpha - \alpha_1), \\ f_2(\alpha) &\approx \alpha_2 + k_2(\alpha - \alpha_2). \end{aligned} \quad (11)$$

We look at *Ansätze*

$$\begin{aligned} \rho_1(\alpha) &= a_1(\alpha - \alpha_1)^{-1+n_1}, \\ \rho_2(\alpha) &= a_2(\alpha - \alpha_2)^{-1+n_2}. \end{aligned} \quad (12)$$

Let us define n_1, n_2 . Looking at the solutions with the scaling near the corresponding critical points

$$\rho_1(\alpha) \sim (\alpha - \alpha_1)^{n_1}, \quad \rho_2(\alpha) \sim (\alpha - \alpha_2)^{n_2}, \quad (13)$$

we obtain from Eq. (5)

$$\begin{aligned} 1 &= \pi_1(\alpha_1)(k_1)^{n_1}, \\ 1 &= \pi_2(\alpha_2)(k_2)^{n_2}. \end{aligned} \quad (14)$$

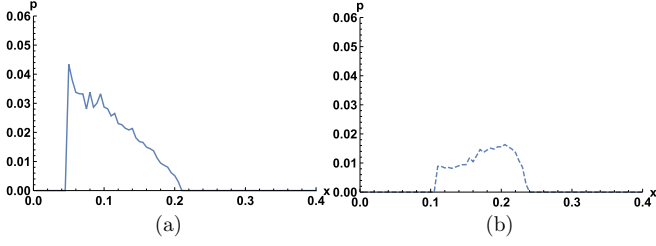


FIG. 1. The steady-state distribution vs $x = \alpha/(2\pi)$. $q_{11} = 0.9$, $q_{22} = 0.9$, $p_{11} = 0.8$, $p_{22} = 0.4$. We have done 1 000 000 sampling with 50 iterations for each. (a) ρ_1 and (b) ρ_2 .

Thus,

$$n_1 = -\frac{\ln[\pi_1(\alpha_1)]}{\ln(k_1)}, \quad n_2 = -\frac{\ln[\pi_2(\alpha_2)]}{\ln(k_2)}. \quad (15)$$

The critical cases are when $n_1 = 1$.

$$\pi_1(\alpha_1)k_1 = 1 \quad (16)$$

For the Fig. 1 case, we take $q_{11} = 0.9$, $q_{22} = 0.9$, $p_{11} = 0.80$, $p_{22} = 0.40$, $\alpha_1 = 0.2866$, $\alpha_2 = 1.289$, $\beta_1 = 0.20517$, $\beta_2 = 0.10601$.

Next, $n_1 = 0.812$, $n_2 = 1.666$, thus we have a singular behavior for ρ_1 at α_1 and smooth behavior for ρ_2 at α_2 . We identify three phases:

- (1) We have singularities in α_1, α_2 (Fig. 2).
- (2) We have one singularities either at α_1 or at α_2 (Fig. 1).
- (3) We have no singularities (Fig. 3).

IV. THE IDENTIFICATION OF THE PHASE TRANSITION IN HIGHER DIMENSIONS

A. Three-matrices case

Consider the case when there are $d > 2$ observed values of the state of HMM. For the $d = 3$ case we have

$$\begin{aligned} A &= \begin{pmatrix} p_{11}q_{11} & p_{11}q_{12} & p_{11}q_{13} \\ p_{12}q_{21} & p_{12}q_{22} & p_{12}q_{23} \\ p_{13}q_{31} & p_{13}q_{32} & p_{13}q_{33} \end{pmatrix}, \\ B &= \begin{pmatrix} p_{21}q_{11} & p_{21}q_{12} & p_{21}q_{13} \\ p_{22}q_{21} & p_{22}q_{22} & p_{22}q_{23} \\ p_{23}q_{31} & p_{23}q_{32} & p_{23}q_{33} \end{pmatrix}, \\ C &= \begin{pmatrix} p_{31}q_{11} & p_{31}q_{12} & p_{31}q_{13} \\ p_{32}q_{21} & p_{32}q_{22} & p_{32}q_{23} \\ p_{33}q_{31} & p_{33}q_{32} & p_{33}q_{33} \end{pmatrix}. \end{aligned} \quad (17)$$

We again should introduce polar coordinates, looking the vectors with a unit norm

$$|\mathbf{z}\rangle = (v_1, v_2, \dots, v_d)h. \quad (18)$$

In principle we can repeat the derivations of the previous section, introducing the scaling indices and looking for the singularities, but it is easier to use an alternative method.

B. Simple formula for the phase transition point

Let us look at matrix A, whose maximal eigenvector (i.e., eigenvector corresponding to maximal eigenvalue) is $|\mathbf{v}\rangle$, denoted by $|\gamma\rangle$ the next eigenvector, denoting the eigenvalues as

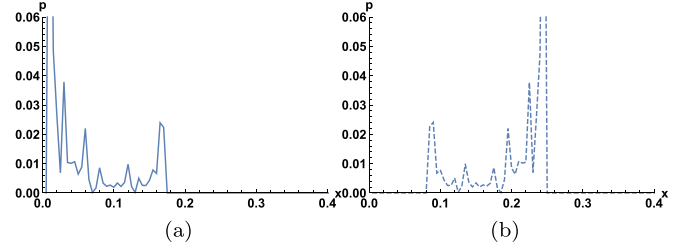


FIG. 2. The steady-state distribution vs $x = \alpha/(2\pi)$. $q_{11} = 0.9$, $q_{22} = 0.9$, $p_{11} = 0.8$, $p_{22} = 0.8$. We have done 1 000 000 sampling with 50 iterations for each. (a) ρ_1 and (b) ρ_2 .

λ_1, λ_2 . If we take a vector $|\mathbf{v}\rangle + \delta|\gamma\rangle$, then multiply by A, we get $\sim|\mathbf{v}\rangle + \delta\frac{\lambda_2}{\lambda_1}|\gamma\rangle$, thus the distance to the $\mathbf{v}$ decreased λ_1/λ_2 times. We have that there is

$$\frac{\lambda_2}{\lambda_1} \quad (19)$$

time attraction. Thus, for the $d = 2$ case we identify $f'_1(\alpha) = \lambda_1/\lambda_2$.

α must be taken as a maximal eigenvalue of the matrix. Next, we should define the probability of choosing in the next round again matrix A by Eq. (2) as π . If the eigenvector of the maximal eigenvalue of A is $|\mathbf{z}\rangle$, then define the probability π in Eqs. (20) and (21) as

$$\pi = \frac{\langle \text{unit} | T_1 | \mathbf{z} \rangle}{\sum_y \langle \text{unit} | T_y | \mathbf{z} \rangle} \quad (20)$$

The exact critical point corresponds to the limit

$$\pi \left| \frac{\lambda_1}{\lambda_2} \right| = 1. \quad (21)$$

When

$$\pi \left| \frac{\lambda_1}{\lambda_2} \right| > 1, \quad (22)$$

the steady-state distribution should be focused on an eigenvector related to the maximal eigenvalue, see Fig. 4.

We verified that the last equation is identical to Eq. (16) for the simplest case $d = 2$.

Equations (20) and (21) together give the critical point.

In the Appendix we work with these algorithms near the critical points and explore how algorithms (BW and Viterbi) change when approaching the critical point.

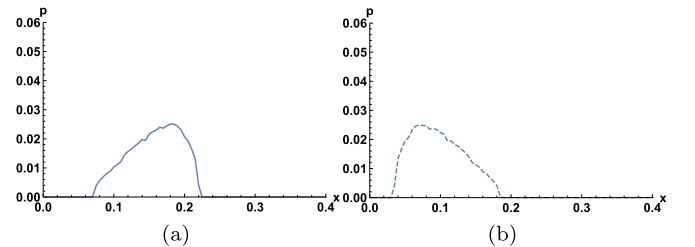


FIG. 3. The steady-state distribution vs $x = \alpha/(2\pi)$. $q_{11} = 0.9$, $q_{22} = 0.9$, $p_{11} = 0.4$, $p_{22} = 0.4$. We have done 1 000 000 sampling with 50 iterations for each. (a) ρ_1 and (b) ρ_2 .

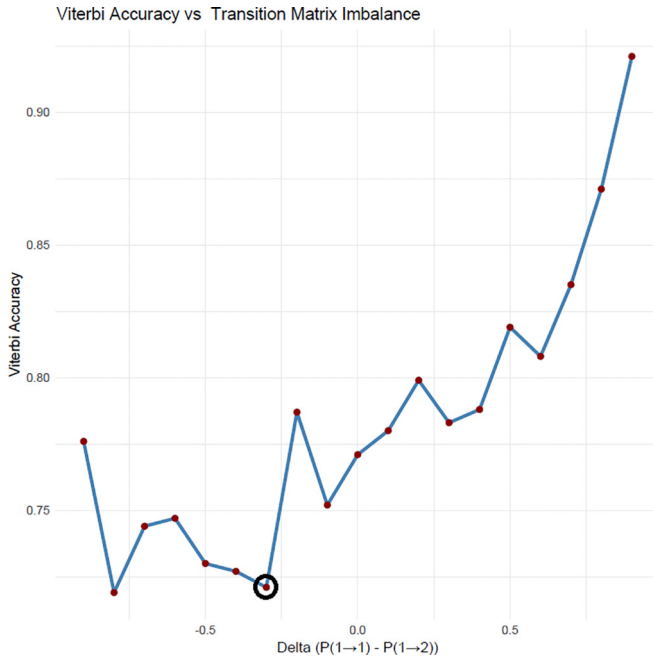


FIG. 4. The accuracy of the Viterbi algorithm vs the parameters of the model. The circled point corresponds to the phase transition point. It is seen that the algorithm shows the lowest accuracy.

V. CONCLUSIONS

Hidden Markov models continue to play a fundamental role in data analysis and are considered singularity-free analytical objects. In this paper, we examined long sequences of observed values and discovered three phases in their distribution. First, we derived the phase transition point by carefully analyzing the steady state of the master functional equation for the simplest case of hidden Markov models, deriving singularity indices. Later, we found an equivalent version of the transition-point equation, which can easily be generalized to a general version of the hidden Markov model. The phase transition is related to the well-known transition to a transition phase in evolutionary models [15–19,24].

There are pure processes with relaxation periods and some random mixing between pure dynamic processes. If the mixing rate is less than the relaxation rate, then the system has

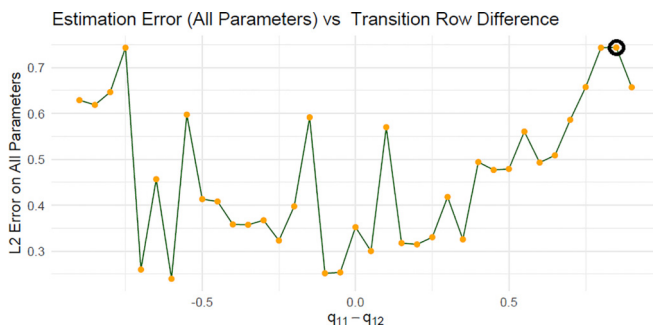


FIG. 5. The accuracy of the Baum-Welch algorithm vs the parameters. The circled point corresponds to the phase transition point. It is seen that the Baum-Welch algorithm shows the lowest accuracy (highest estimation error).

sufficient time to relax; otherwise, it is entirely in a transition phase between different pure states.

Why is the transition point important? There was great hope that living systems exist on the brink of chaos, which is often identified with a phase transition point [25,26]. Therefore, we might have hoped that working with HMMs near the critical point, defined by (16) and (20), would improve the efficiency of HMM algorithms. However, according to our numerical calculations, this turned out to be incorrect, and in some cases, the opposite is even true. See the appendix and Figs. 4–5

ACKNOWLEDGMENT

This work was supported by HESC of Armenia, Grant No. 24FP-1D009.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX: ACCURACY OF THE VITERBI AND BAUM-WELCH ALGORITHMS

HMMs are generally learned using two powerful algorithms—specifically, the Viterbi algorithm [9] and the Baum-Welch algorithm [3]. The former is used to estimate the most probable path (most probable sequence) of hidden states. The latter estimates the HMM parameters.

Both algorithms are based on maximum likelihood estimation (MLE). While the Viterbi algorithm uses pure maximum likelihood estimation, the Baum-Welch algorithm, due to the complexity of the problem, uses an iterative estimation algorithm (i.e., an expectation-maximization algorithm).

The algorithm is explicitly deduced in Ref. [27–29]. Any attempt to add additional information in the form of constraints can lead to a significant increase in the computational load and additional iterative cycles of the algorithm, as the authors demonstrate.

Since we are interested specifically in the case where we have additional information about the probabilities (i.e., additional relations), we avoid mathematical calculations for such cases.

Our goal is to determine whether the accuracy of the Viterbi and BW algorithms is related to the proximity of the true parameters to the phase transition boundaries. More precisely, we define accuracy in the Viterbi algorithm as the percentage of correct guesses in the simulated HMM and the total error calculated in the Euclidean norm (i.e., the sum of squares) in the BW algorithm. We then ask the following question: when the true HMM parameters are close to the phase transition boundary, does the accuracy of the Viterbi and BW algorithms increase or decrease? Or is there no such general relationship?

The precise method for calculating the error in the BW algorithm is presented in Ref. [29]. How does computational complexity increase dramatically when adding new variables?

Even in simple hidden Markov models with two hidden and two observed states, we need to calculate the eigenvalues of a five-by-five matrix, which is a highly nontrivial task (different for each sequence of realizations and increasingly complex with increasing length of the observed sequence).

The phase transition is calculated according to the following system of equations (for the case of two hidden and two observed states), which is essentially a combination of (20) and (22):

$$\begin{aligned} & \frac{(1-p_{11})x + (1-p_{12})y + \sqrt{[(1-p_{11})x + (1-p_{12})y]^2 + 4(1-p_{11})(1-p_{12})(1-x-y)}}{2} \\ & \frac{(1-p_{11})x + (1-p_{12})y - \sqrt{[(1-p_{11})x + (1-p_{12})y]^2 + 4(1-p_{11})(1-p_{12})(1-x-y)}}{2} \\ & = \frac{\frac{1}{2}}{\frac{(1-p_{11})x\pi_1 + (1-p_{11})(1-y)(1-\pi_1) + (1-p_{12})(1-x)\pi_1 + (1-p_{12})y(1-\pi_1)}{2}} \\ \pi_1 & = \frac{\frac{p_1x + p_2y + \sqrt{(p_1x + p_2y)^2 + 4p_1p_2(1-x-y)}}{2} - p_1(1-y)}{p_1x - p_1(1-y)}, \quad (\text{A1}) \end{aligned}$$

where x and y denote q_{11} and q_{22} , respectively.

We thus generate an HMM sequence of length $T = 1000$ with fixed initial emission probabilities and variable

transition probabilities in the first row. To do this, we first calculate the phase transition probabilities according to (A1). We then evaluate the accuracy of both the BW and Viterbi algorithms.

There is no obvious relationship between the accuracy of the Viterbi algorithm and the closeness to the phase transition point, though we included the graph where at the transition point the algorithm shows the least accuracy. This is the case in most but not all situations. So, the most which can be concluded is that generally the accuracy worsens close to transition point, but sometimes there is no crucial difference. We have never seen the opposite effect (i.e., the closeness to the transition point never improves the accuracy), where the phase transition point is defined as a point where the smallest eigenvalue equates to the probability of steady state times the maximal eigenvalue.

Again, though we included here the case where the accuracy of BW is the worst in the vicinity of the transition point, there is no obvious relationship. Nevertheless, in most cases the algorithm shows worse performance (only in terms of accuracy and not convergence speed) near the transition point. Sometimes, the closeness has no effect.

-
- [1] L. E. Baum and T. Petrie, Statistical inference for probabilistic functions of finite state Markov chains, *Ann. Math. Stat.* **37**, 1554 (1966).
 - [2] Y. Ephraim and N. Merhav, Hidden Markov processes, *IEEE Trans. Inform. Theory* **48**, 1518 (2002).
 - [3] L. R. Rabiner, A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. IEEE*, **77**, 257 (1989).
 - [4] T. Starner and A. Pentland, Visual recognition of American sign language using hidden Markov models, Master's thesis, MIT, Program in Media Arts, 1995.
 - [5] R. M. Durbin, M. Richard, S. R. Eddy, A. Krogh, and G. Mitchison, *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, 1st ed. (Cambridge University, Cambridge, New York, 1998).
 - [6] A. Schliep, A. Schönhuth, and C. Steinhoff, Using hidden Markov models to analyze gene expression time course data, *Bioinformatics* **19**, i255 (2003).
 - [7] W. J. Ewens and G. R. Grant, *Statistical Methods in Bioinformatics: An Introduction* (Springer, New York, 2001).
 - [8] Edited by R. S. Mamon and R. J. Elliott, in *Hidden Markov Models in Finance: Further Developments and Applications* (Springer Nature, New York, 2014), Vol. 2.
 - [9] J. Schmidhuber, Deep learning in neural networks: An overview, *Neural Networks* **61**, 85 (2015).
 - [10] O. Zuk, I. Kanter, and E. Domany, The entropy of a binary hidden Markov process, *J. Stat. Phys.* **121**, 343 (2005).
 - [11] A. E. Allahverdyan, Entropy of hidden Markov processes via cycle expansion, *J. Stat. Phys.* **133**, 535 (2008).
 - [12] H. Furstenberg and H. Kesten, Products of random matrices, *Ann. Math. Statist.* **31**, 457 (1960).
 - [13] A. Crisanti, G. Paladin, and A. Vulpiani, *Products of Random Matrices in Statistical Physics* (Springer, Berlin, 1993).
 - [14] A. Mayer, T. Mora, O. Rivoire, and A. M. Walczak, Diversity of immune strategies explained by adaptation to pathogen statistics, *Proc. Natl. Acad. Sci. USA* **113**, 8630 (2016).
 - [15] E. Kussell and S. Leibler, Phenotypic diversity, population growth, and information in fluctuating environments, *Science* **309**, 2075 (2005).
 - [16] P. G. Hufton, Y. T. Lin, T. Galla, and A. J. McKane, Intrinsic noise in systems with switching environments, *Phys. Rev. E* **93**, 052119 (2016).
 - [17] A. Skanata and E. Kussell, Evolutionary phase transitions in random environments, *Phys. Rev. Lett.* **117**, 038104 (2016).
 - [18] K. Wienand, E. Frey, and M. Mobilia, Evolution of a fluctuating population in a randomly switching environment, *Phys. Rev. Lett.* **119**, 158301 (2017).
 - [19] O. Rivoire and S. Leibler, The value of information for populations in varying environments, *J. Stat. Phys.* **142**, 1124 (2011).
 - [20] O. Rivoire, Informations in models of evolutionary dynamics, *J. Stat. Phys.* **162**, 1324 (2016).
 - [21] D. B. Saakian, Exact solution of the hidden Markov processes, *Phys. Rev. E* **96**, 052112 (2017).
 - [22] H. Mineo, V. Suvorov, and D. B. Saakian, Investigation of the product of random matrices and related evolution models, *Mathematics* **11**, 3430 (2023).
 - [23] P. Pöpperl, I. V. Gornyi, D. B. Saakian, and O. M. Yevtushenko, Localization, fractality, and ergodicity in a monitored qubit, *Phys. Rev. Res.* **6**, 013313 (2024).
 - [24] M. Nilsson and N. Snoad, Error thresholds for quasispecies on dynamic fitness landscapes, *Phys. Rev. Lett.* **84**, 191 (2000).
 - [25] T. Mora and W. Bialek, Are biological systems poised at criticality? *J. Stat. Phys.* **144**, 268 (2011).

- [26] M. A. Muñoz, *Colloquium: Criticality and dynamical scaling in living systems*, [Rev. Mod. Phys. **90**, 031001 \(2018\)](#).
- [27] A. J. Viterbi, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, [IEEE Trans. Inf. Theory **13**, 260 \(1967\)](#).
- [28] D. Steeneck, Daniel, and E. L. Fredrik, Welch algorithm with limiting distribution constraints, [Oper. Res. Lett. **46**, 563 \(2018\)](#).
- [29] M. Jamshidian and R. I. Jennrich, Standard errors for EM estimation, [J. R. Stat. Soc. Series B: Stat. Methodol. **62**, 257 \(2000\)](#).