

深層ニューラルネットワーク と 非平衡臨界現象

- ディープラーニングに潜む普遍的な法則の探求 -

東京大学 知の物理学研究センター

玉井 敬一

KT, T. Okubo, T. V.T. Duy, N. Natori and S. Todo. arXiv: 2307.02284

上智大学

第3回「学習物理」領域セミナー

(兼 第55回「ディープラーニングと物理学」セミナー)

2023年11月9日

自己紹介

• 玉井 敬一 博士 (理学)

学歴・職歴

2011年4月 - 2013年3月: 東京大学理学部物理学科

指導教員: 宮下 精二 教授 (理論)、酒井 広文 准教授 (実験)

2013年4月 - 2018年3月: 東京大学大学院理学系研究科物理学専攻

指導教員: 佐野 雅己 教授

2018年4月 - 2020年3月: 東京大学物性研究所 (CCMS)

受入所員: 川島 直輝 教授

2020年4月 - 2022年5月: 産業技術総合研究所 (CD-FMat)

受入所員: 三宅 隆 材料インフォマティクスチーム長

2022年6月 -: 東京大学大学院理学系研究科知の物理学研究センター

受入教員: 藤堂 眞治 教授

ディープラーニングの 応用研究の躍進

- 深層学習の応用は質・量ともに急激に発展中

- AlphaGo, MidJourney, ChatGPT, Bing AI Chat, ...

- モデル規模も急激に成長

- Transformer などでは大きなモデルを使うほど性能が上がる？
(Neural network scaling laws)

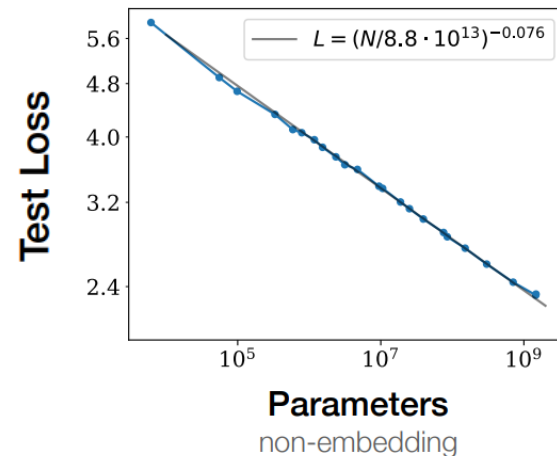
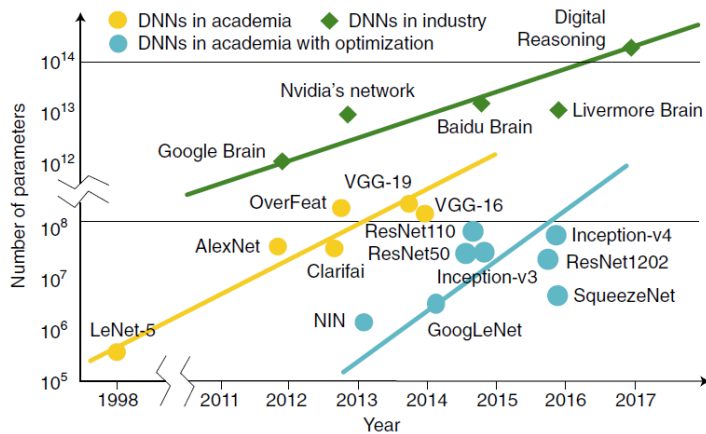


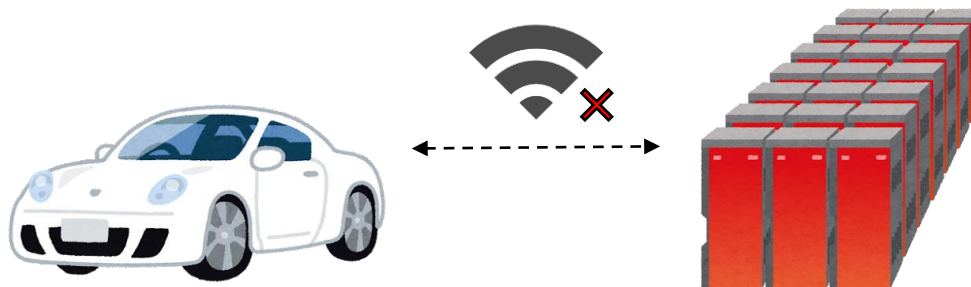
Fig. from: X.-w. Xu *et al.* Nat. Electron. **1**, 216, (2018);
J. Kaplan *et al.* arXiv: 2001.08361

ディープラーニングのエネルギー問題

- 深層学習のより広範な社会実装を目指していく上で、エネルギー問題がボトルネックになり得る
 - 大規模モデルは地球環境への負荷も大きい
 - そもそも大規模計算資源にアクセスできないケースもある
 - 少なくとも、モデルから予測・生成物を得る際のエネルギー消費は抑えたい

Consumption	CO ₂ e (lbs)
Air travel, 1 passenger, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000

Training one model (GPU)	CO ₂ e (lbs)
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468
Transformer (big)	192
w/ neural architecture search	626,155



ディープラーニングの エネルギー問題

- 深層学習のより広範な社会実装を目指していく上で、エネルギー問題がボトルネックになり得る
 - 大規模モデルは地球環境への負荷も大きい
 - そもそも大規模計算資源にアクセスできないケースもある
 - 少なくとも、モデルから予測・生成物を得る際のエネルギー消費は抑えたい
- 自然界から学べるところはないか？
 - 人間の脳が日常的に消費するエネルギー： $\sim 20\text{ W}$
(高性能 GPU の消費電力： $\gtrsim 600\text{ W}$)
 - ! 「人間の知的活動が現在の AI よりエネルギー効率が良いか」
自体は微妙な問題

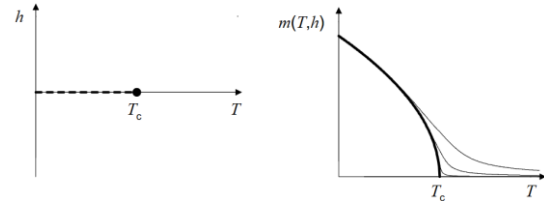
本講演のテーゼ：

**非平衡臨界現象を「横の糸」にすると
自然知能と人工知能の共通/相違点を
見通しよく整理できる（かも）**

臨界現象のスケーリング則： 平衡相転移の場合

- 二次相転移の転移点近傍で物理量が冪乗則を示す

$$m \sim (T_c - T)^\beta, \quad \xi \sim |T - T_c|^{-\nu}, \dots$$



- 系の詳細に依存しない普遍的な臨界指数 ➡ 普遍クラス
 - 非普遍的パラメータ込みでリスケールすれば
様々な系の振る舞いが1本の曲線上に乗る

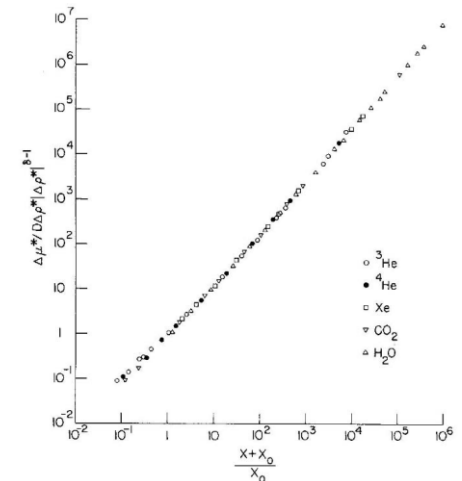
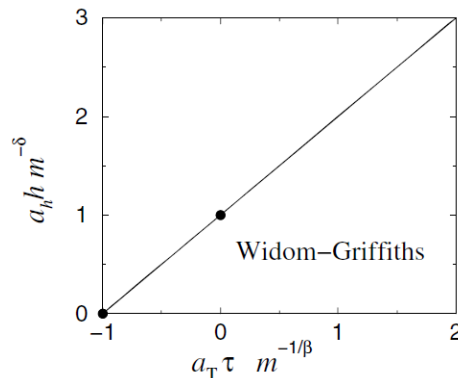
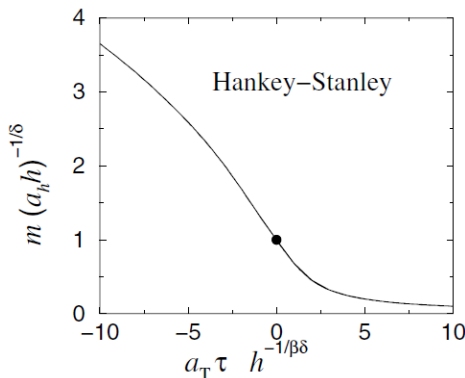
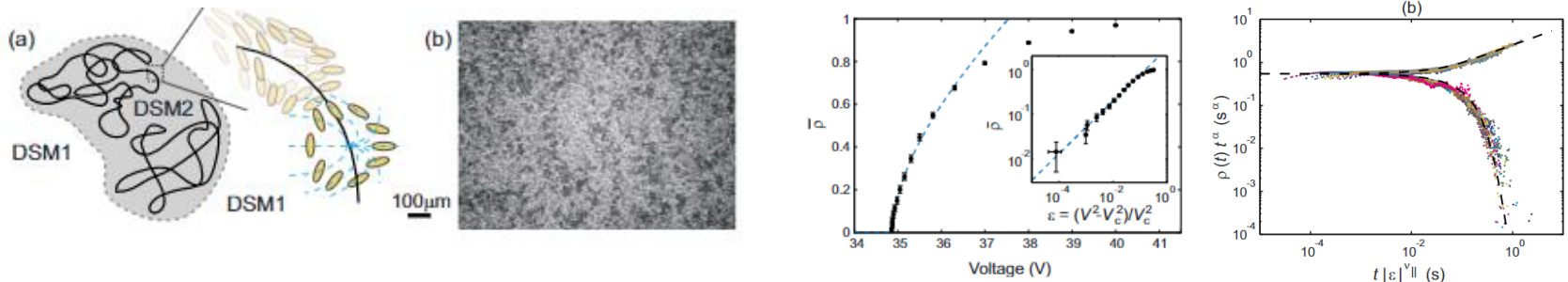


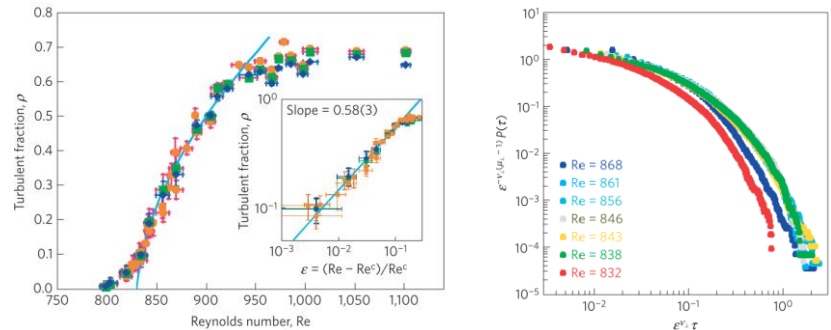
Fig. from S. Lubeck. Int. J. Mod. Phys. B **18**, 3977 (2004);
J. V. Sengers and J. M. H. Levelt Sengers in *Progress in liquid Physics* (1978, John Wiley&Sons)

実験で見る非平衡臨界現象

- スケーリング則が非平衡系の実験でも見られ始めた
液晶電気対流の欠陥乱流転移



開放せん断流における乱流遷移



K. A. Takeuchi *et al.* Phys. Rev. Lett. **99**, 234503 (2007); Phys. Rev. E **80** 051116 (2009)

M. Sano and KT. Nat. Phys. **12**, 249 (2016); K. Kohyama *et al.* (incl. KT) Proc. TSFP-12 (2022)

See also: G. Lemoult *et al.* Nat. Phys. **12**, 254 (2016); L. Klotz *et al.* Phys. Rev. Lett. **128**, 014502 (2022)

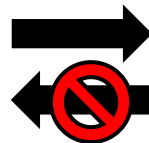
代表的な非平衡臨界現象 としての吸収状態転移

- ふたつの系の共通点

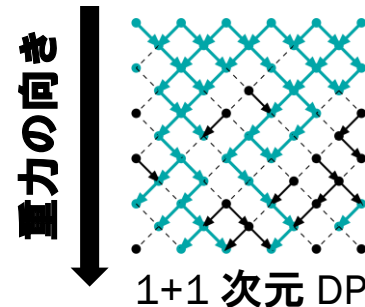
- 定性的に異なるふたつの状態の空間的共存
- 一方の状態がもう一方の状態より好まれやすい

液晶電気対流：トポロジカル欠陥による弾性エネルギー損
層流-乱流遷移：層流の線型安定性

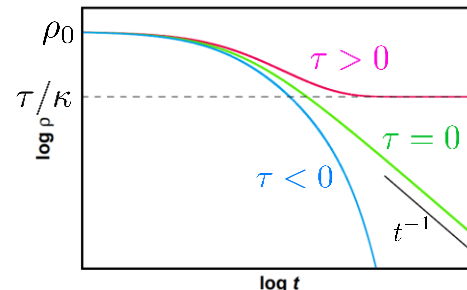
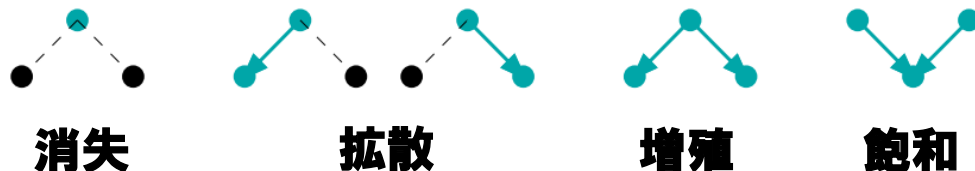
- **吸収状態**：一度到達すると自発的に出ていけない状態



吸収状態転移の臨界現象



- 確率数理モデルを使った理論研究が大きく進展
 - モデルの例：Directed percolation (DP), Contact process
- 描像：消失過程と増殖過程のせめぎあい + 拡散・飽和

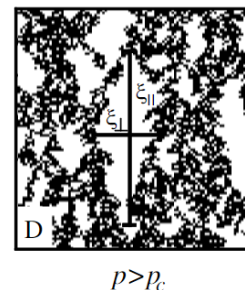
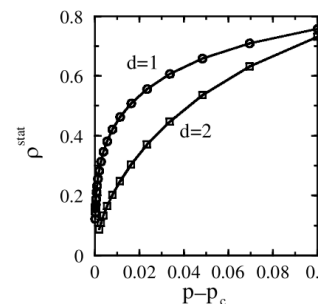


➡ 平均場理論 $\frac{d\rho}{dt} = \tau\rho - \kappa\rho^2$ (ρ : ● の割合)

- 空間次元が 3 以下では平均場理論と異なる指数の冪乗則
 - 相関長 ξ_{\perp} と相関時間 ξ_{\parallel} も同様

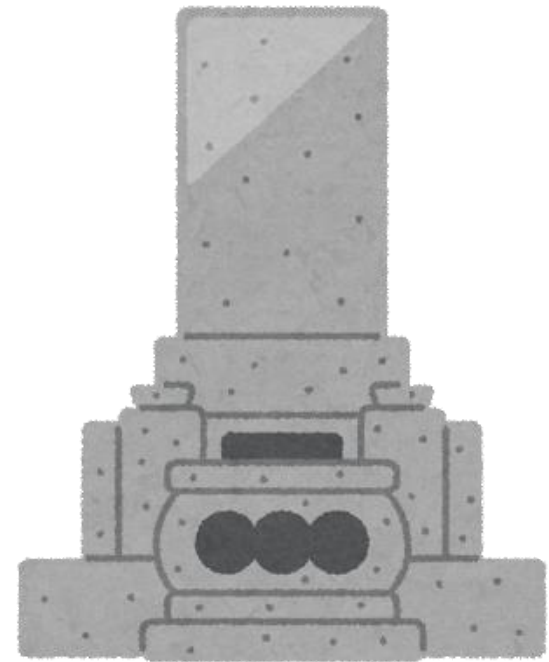
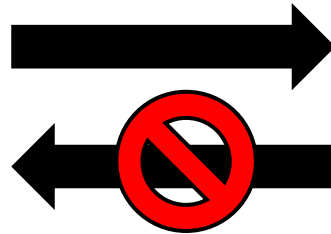
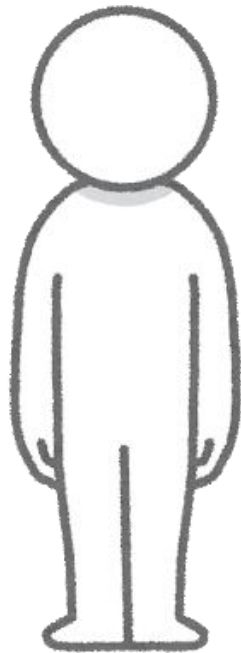
$$\rho \sim \tau^{\beta}, \quad \xi_{\perp} \sim |\tau|^{-\nu_{\perp}}, \quad \xi_{\parallel} \sim |\tau|^{-\nu_{\parallel}}$$

(平均場理論では $\beta_{\text{MF}} = 1, \nu_{\parallel \text{MF}} = 1$)

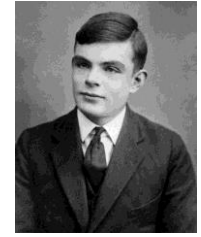


神経系ダイナミクスと 吸収状態転移の関係？

- ひとつは誰もが既に馴染みがあるはず！



Turing の慧眼：神経発火と雪崩のアナロジー



Alan Turing

- 「脳がある種の臨界状態にあるのでは？」という指摘が 1950 年の頃からなされていた

quiescence, like a piano string struck by a hammer. Another simile would be an atomic pile of less than critical size: an injected idea is to correspond to a neutron entering the pile from without. Each such neutron will cause a certain disturbance which eventually dies away. If, however, the size of the pile is sufficiently increased, the disturbance caused by such an incoming neutron will very likely go on and on increasing until the whole pile is destroyed. Is there a corresponding phenomenon for minds, and is there one for machines? There does seem to be one for the human mind. The majority of them seem to be 'sub-critical', i.e. to correspond in this analogy to piles of sub-critical size. An idea presented to such a mind will on average give rise to less than one idea in reply. A smallish proportion are super-critical. An idea presented to such a mind may give rise to a whole 'theory' consisting of secondary, tertiary and more remote ideas. Animals minds seem to be very definitely sub-critical. Adhering to this analogy we ask, 'Can a machine be made to be super-critical?'

⋮

神経系ダイナミクスと 吸収状態転移の関係？

- **神経雪崩**：ニューロンのバーストの連鎖現象
 - 一度雪崩が収まると、再度雪崩を引き起こすイベントが発生するまでそのまま（吸収状態）
- **臨界現象のような冪的振る舞いがパラメータの細かい調整なしで現れる（自己組織化臨界）**

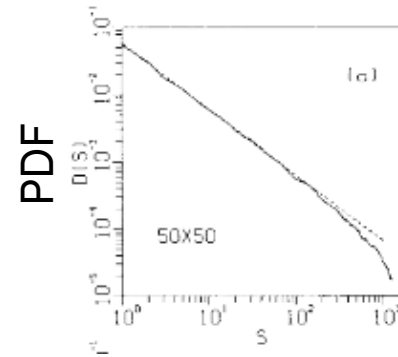
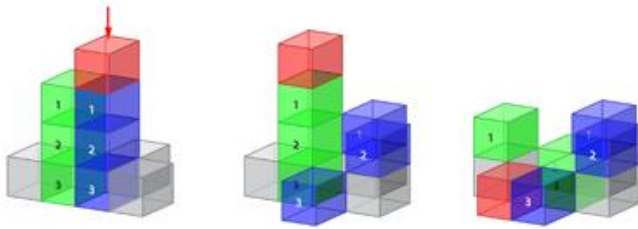
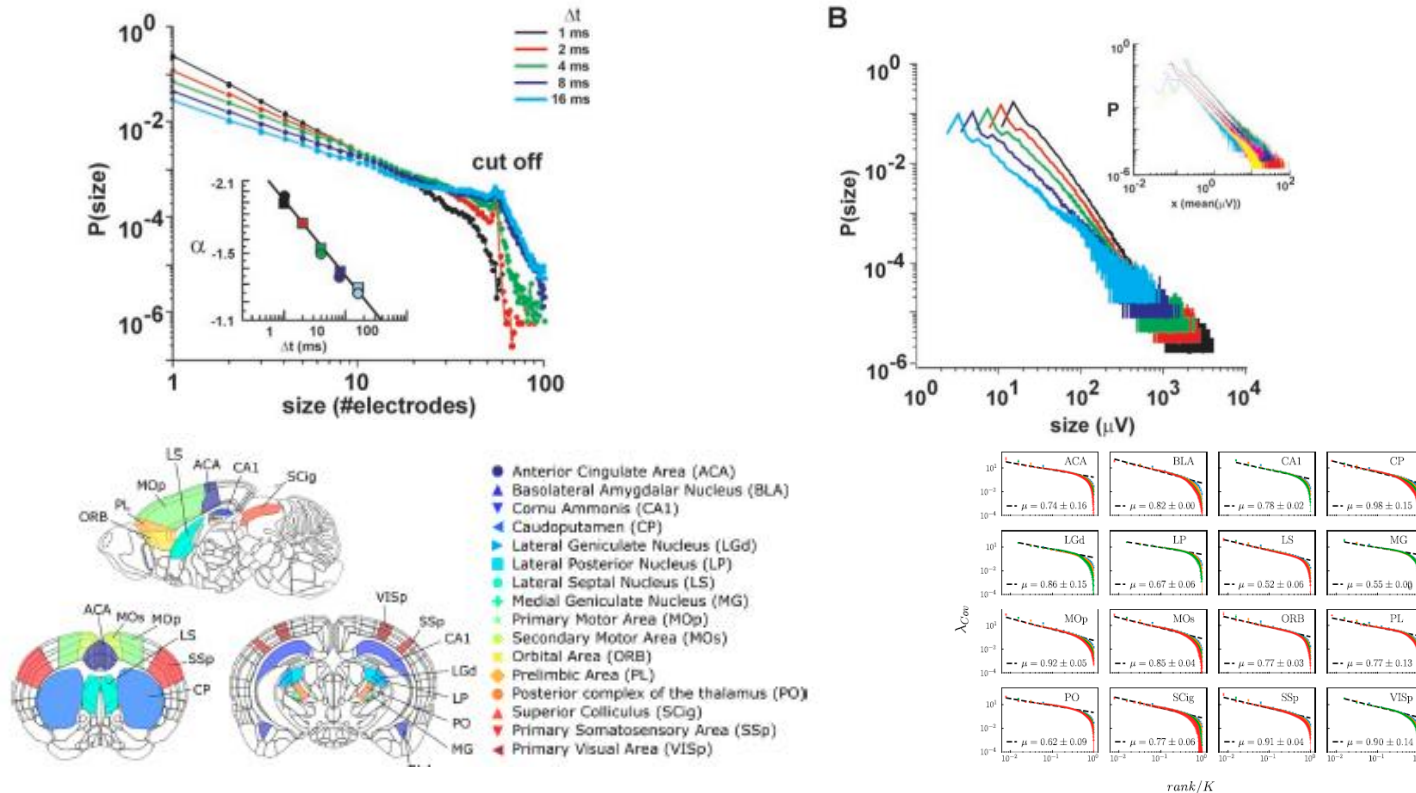


Fig. by Alan Stonebraker ([Stonebraker Design | Scientific illustration and more... \(alanstonebraker.com\)](http://www.alanstonebraker.com))

P. Bak, C. Tang & K. Wiesenfeld. Phys. Rev. Lett. **59**, 381 (1987); Phys. Rev. A **38**, 364 (1988)

神経系ダイナミクスと吸収状態転移の関係？

- 動物の脳でも（少なくとも擬似的に）自己組織化臨界を実現しているという実験証拠が集まってきている



J. M. Beggs & D. Plenz. *J. Neurosci.* **23**(35), 11167 (2003)

G. B. Morales et al. *PNAS* **120**, e2208998120 (2023)

深層ニューラルネットワークではどうなの？

深層ニューラルネットワークの「忘却転移」？

- 初期化した全結合深層パーセプトロンの信号順伝播

- x : 入力、 $z^{(l)}$: preactivation, $W^{(l)}$: 重み行列、 b : バイアス

- n : 隠れ層のニューロン数 (「幅」)、 σ_w, σ_b : hyperparameter, h : 活性化関数

$$z_i^{(1)} = \frac{\sigma_w}{\sqrt{n}} \sum_j W_{ij}^{(1)} x_j + \sigma_b b_i^{(1)} \quad \text{with } W_{ij}^{(l)}, b_i^{(l)} \sim \mathcal{N}(0, 1)$$

$$z_i^{(l+1)} = \frac{\sigma_w}{\sqrt{n}} \sum_j W_{ij}^{(l+1)} h(z_j^{(l)}) + \sigma_b b_i^{(l+1)}$$

Q. 異なる入力信号 x_1, x_2 を同じ初期化ニューラルネットに入れたとき、信号が伝播するにつれ差異はどうか？

「相関係数」 (を 1 から引いたもの) で定量化

$$\rho^{(l)}[\sigma_w; n] := 1 - \frac{\sum_i (z_{1;i}^{(l)} - Z_1^{(l)})(z_{2;i}^{(l)} - Z_2^{(l)})}{\sqrt{\sum_i (z_{1;i}^{(l)} - Z_1^{(l)})^2 \sum_i (z_{2;i}^{(l)} - Z_2^{(l)})^2}}, \quad Z_j^{(l)} := \frac{1}{n} \sum_i z_{j;i}^{(l)}$$

深層ニューラルネットワークの「忘却転移」？

- 初期化した全結合深層パーセプトロンの信号順伝播

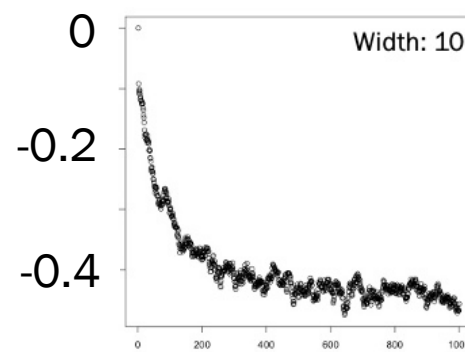
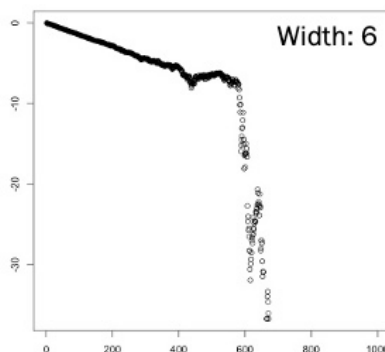
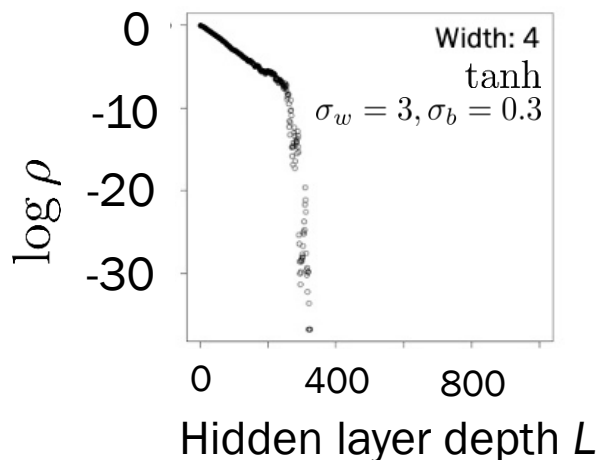
- x : 入力、 $z^{(l)}$: preactivation, $W^{(l)}$: 重み行列、 b : バイアス

- n : 隠れ層のニューロン数（「幅」）、 σ_w, σ_b : hyperparameter, h : 活性化関数

$$z_i^{(1)} = \frac{\sigma_w}{\sqrt{n}} \sum_j W_{ij}^{(1)} x_j + \sigma_b b_i^{(1)}$$

with $W_{ij}^{(l)}, b_i^{(l)} \sim \mathcal{N}(0, 1)$

$$z_i^{(l+1)} = \frac{\sigma_w}{\sqrt{n}} \sum_j W_{ij}^{(l+1)} h(z_j^{(l)}) + \sigma_b b_i^{(l+1)}$$



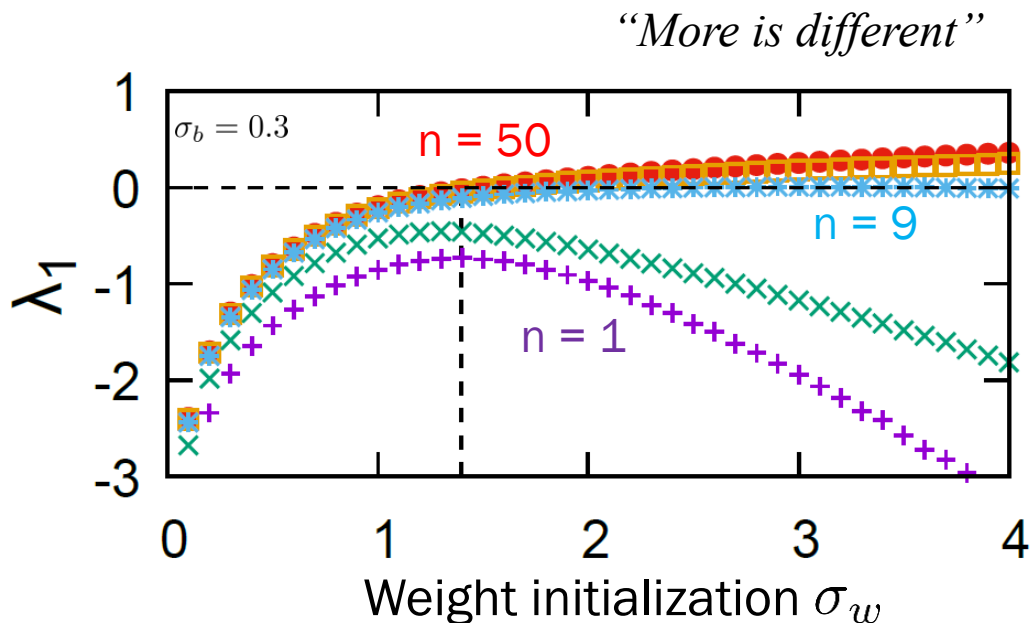
ニューロン多体現象としての 「忘却転移」

- 忘却解 ($z_1^{(l)} = z_2^{(l)}$) の微小摂動に対する安定性を定量化
 - 幅が狭いと任意の hyperparameter で安定
- ↕
- 幅が広いと不安定な hyperparameter 領域が現れる

「最大 Lyapunov 指数」
(用語の濫用を含む)

$$\lambda_1 := \lim_{l \rightarrow \infty} \frac{1}{l} \log \frac{\|J^{(l)}(z^{(l)}) \cdots J^{(1)}(z^{(1)}) \mathbf{u}_0\|}{\|\mathbf{u}_0\|}$$

$$J^{(l)}(z) = \begin{pmatrix} J_{11}^{(l)}(z) & \cdots & J_{1n}^{(l)}(z) \\ \vdots & \ddots & \vdots \\ J_{n1}^{(l)}(z) & \cdots & J_{nn}^{(l)}(z) \end{pmatrix} \quad \text{with} \quad J_{ij}^{(l)}(z) := W_{ij}^{(l)} h'(z_j).$$



幅無限極限での 「忘却転移」の記述

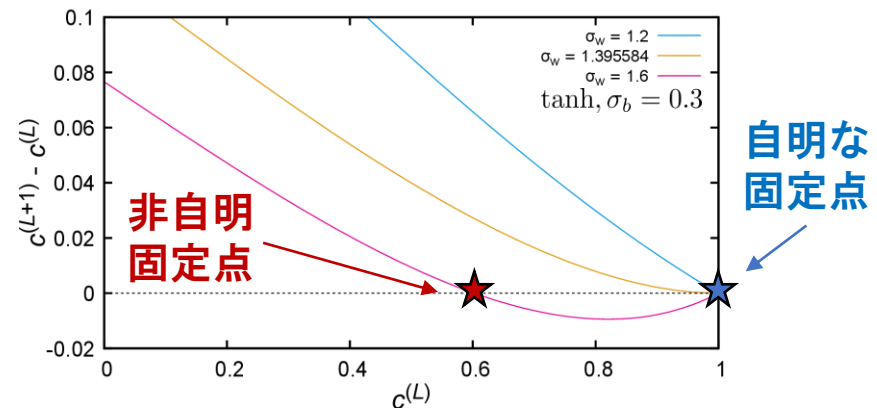
$$z_i^{(l+1)} = \frac{\sigma_w}{\sqrt{n}} \sum_j W_{ij}^{(l+1)} h(z_j^{(l)}) + \sigma_b b_i^{(l+1)}$$

- 完全に忘却した（異なる入力信号 x_1, x_2 を与えたにもかかわらず対応する preactivation $z_1^{(l)}, z_2^{(l)}$ が全く同じ）状態 (ordered state) は吸収状態
- 幅 $n \rightarrow \infty$ の極限（『熱力学極限』）が扱いやすい
 - Preactivation pair $(z_1^{(l)}, z_2^{(l)})$ が多変量ガウス分布に従う
 - Preactivation の分散 $q^{(l)}$ ・ 相関係数 $c^{(l)}$ の漸化式が書き下せる

$$q^{(l+1)} = \sigma_w^2 \int dz \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} h^2(\sqrt{q^{(l)}} z) + \sigma_b^2$$

$$c^{(l+1)} = \frac{1}{\sqrt{q_1^{(l+1)} q_2^{(l+1)}}} \left[\sigma_w^2 \int dz_1 \int dz_2 \frac{1}{\sqrt{(2\pi)^2}} e^{-\frac{z_1^2 + z_2^2}{2}} h(u_1^{(l)}) h(u_2^{(l)}) + \sigma_b^2 \right]$$

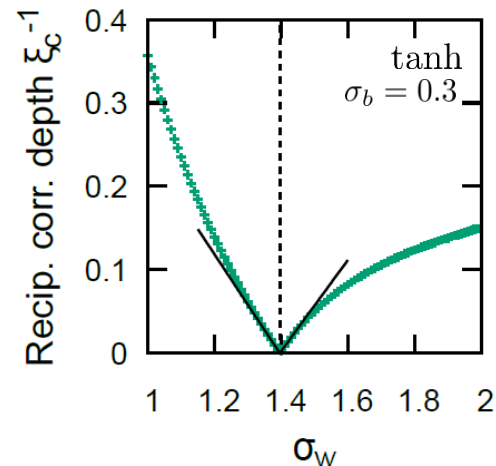
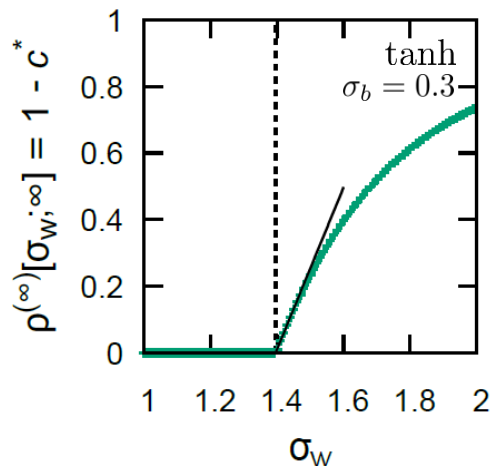
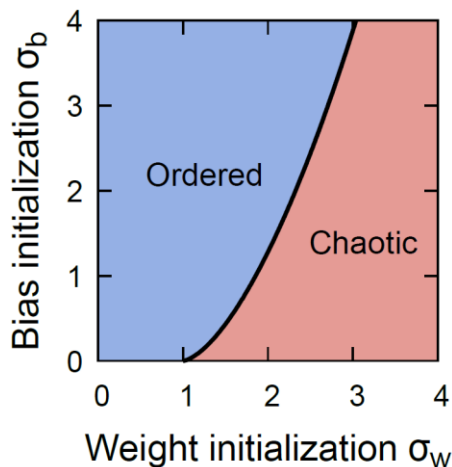
with $u_1^{(l)} := \sqrt{q_1^{(l)}} z_1, \quad u_2^{(l)} := \sqrt{q_2^{(l)}} (c^{(l)} z_1 + \sqrt{1 - c^{(l)2}} z_2)$



幅無限極限での 「忘却転移」の記述

$$z_j^{(l+1)} = \frac{\sigma_w}{\sqrt{n}} \sum_j W_{ij}^{(l+1)} h(z_j^{(l)}) + \sigma_b b_i^{(l+1)}$$

- 完全に忘却した（異なる入力信号 x_1, x_2 を与えたにもかかわらず対応する preactivation $z_1^{(l)}, z_2^{(l)}$ が全く同じ）状態 (ordered state) は吸収状態
- 幅 $n \rightarrow \infty$ の極限（『熱力学極限』）が扱いやすい
 - Preactivation pair $(z_1^{(l)}, z_2^{(l)})$ が多変量ガウス分布に従う
 - Preactivation の分散 $q^{(l)}$ ・相関係数 $c^{(l)}$ の漸化式が書き下せる



「忘却転移」のスケーリング: 全結合パーセプトロンの場合

- 臨界点近傍の秩序変数 $\rho^{(\infty)} = 1 - c^{(\infty)}$ や「相関深さ ξ_c 」に冪乗則が見られる

$$\rho \sim (\sigma_w - \sigma_{w;c})^1, \quad \xi_c \sim |\sigma_w - \sigma_{w;c}|^{-1}$$

- 指数が吸収状態転移の平均場理論と同じ！ (MF 普遍クラス)
- 現象論的スケーリングも有効

$$\rho^{(l)}[\sigma_w] \simeq l^{-1} f((\sigma_w - \sigma_{w;c})l)$$

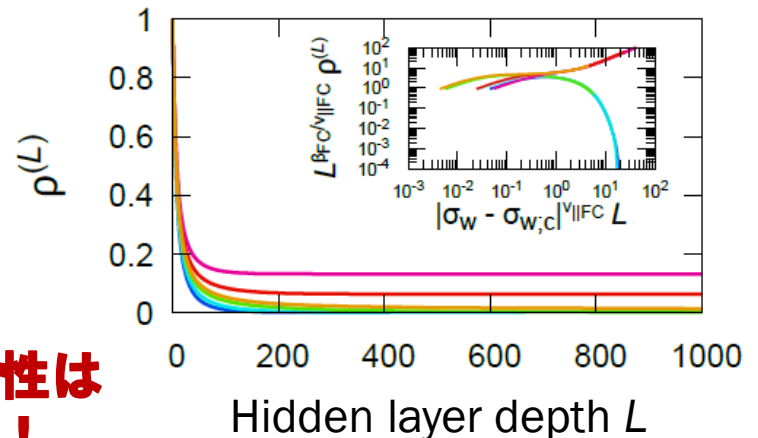
- 秩序変数の平均場理論的記述：

$$\frac{d\rho}{dl} = \gamma \cdot (\sigma_w - \sigma_{w;c})\rho - \kappa\rho^2$$

γ, κ : 非普遍的パラメータ

(活性化関数形、 σ_b に依存)

➡ **臨界点 (近傍) のネットワークの個性は
非普遍パラメータに押し込まれる！**



非普遍的パラメータ： 全結合パーセプトロンの場合

- 臨界点でのネットワークの振る舞いは preactivation の分散 q^* とパラメータ κ で特徴づけられる

(1) 臨界点の位置と preactivation 分散 q^*

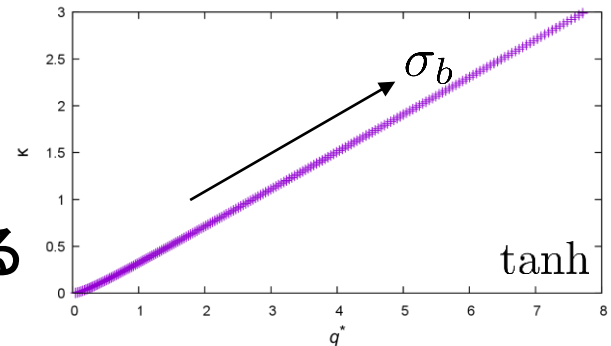
- 相関係数の自明な固定点の安定性条件と分散の固定点条件を連立して（数値的に）解くのが手っ取り早い
 - 純経験的手段なら有限サイズスケーリング

$$1 = \sigma_{w;c}^2 \int dz \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} h'^2(\sqrt{q^*}z) \quad q^* = \sigma_{w;c}^2 \int dz \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} h^2(\sqrt{q^*}z) + \sigma_b^2$$

(2) 非普遍的パラメータ κ

- (1) の結果を使えば一種の順問題
 - 秩序変数の冪減衰の係数の逆数でもある

$$\kappa = \frac{\int dz \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} h''^2(\sqrt{q^*}z)}{2}$$



“So What?”

- 全結合に興味を絞るのをひとまずよしとしても.....

1. 訓練過程でパラメータは初期値から離れるわけで、初期化時のネットワークの性質が分かったところでどうもしないのでは？
2. よしんばどうかするとして、各パラメータを知ってディープラーニング設計上の意味があるのか？

たかが初期化、されど初期化： NTK と Lazy Learning

- 勾配降下法による学習ダイナミクス

- 訓練データ： $\{(\mathbf{x}^{[1]}, y^{[1]}), \dots, (\mathbf{x}^{[D]}, y^{[D]})\}$
- 損失関数： $C = \frac{1}{D} \sum_j |y^{[j]} - g(\mathbf{x}^{[j]}; \boldsymbol{\theta})|^2$
- 時間発展

$$\text{パラメータ： } \frac{d\boldsymbol{\theta}}{dt} = -\alpha \nabla_{\boldsymbol{\theta}} C$$

ニューラルネットワーク出力：

$$\begin{aligned} \frac{\partial g(\mathbf{x}; \boldsymbol{\theta}(t))}{\partial t} &= \nabla_{\boldsymbol{\theta}} g(\mathbf{x}; \boldsymbol{\theta}(t)) \cdot \frac{d\boldsymbol{\theta}}{dt} \\ &= -\alpha \sum_m \sum_j \frac{\partial g}{\partial \theta_j}(\mathbf{x}; \boldsymbol{\theta}) \frac{\partial g}{\partial \theta_j}(\mathbf{x}^{[m]}; \boldsymbol{\theta}) \left[-\frac{2}{D} (y^{[m]} - g(\mathbf{x}^{[m]}; \boldsymbol{\theta})) \right] \\ &=: \Theta(\mathbf{x}, \mathbf{x}^{[m]}; \boldsymbol{\theta}) \end{aligned}$$

Neural Tangent Kernel (NTK)

簡単のため、出力は
1次元とする
 $\boldsymbol{\theta} (:= \text{vect}\{W^{(1)}, \mathbf{b}^{(1)}, \dots\})$
NNの学習パラメータ
 $g(\mathbf{x}; \boldsymbol{\theta})$: NNの出力
 α : 学習率

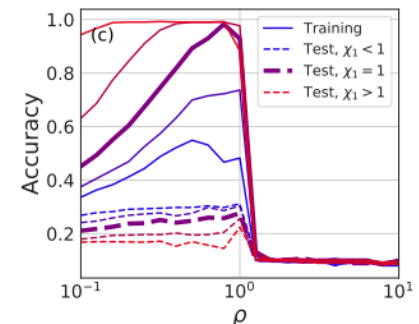
たかが初期化、されど初期化： NTK と Lazy Learning

$$\Theta(x_1, x_2) = \sum_j \frac{\partial g}{\partial \theta_j}(x; \theta) \frac{\partial g}{\partial \theta_j}(x^{[m]}; \theta)$$

無限幅の場合……

- 訓練過程でニューラルネットワークの各パラメータが初期値からほとんど動かない (lazy)
- NTK Θ が決定論的になり、訓練過程でも一定
- 学習率 α が $2(\lambda_{\max}(\Theta_{\text{train,train}}) + \lambda_{\min}(\Theta_{\text{train,train}}))^{-1}$ 以下なら、訓練過程が初期値まわりの線型化ダイナミクスに帰着
 - 経験的には、学習率はその閾値を超えたあたりから、性能が急激に悪くなる

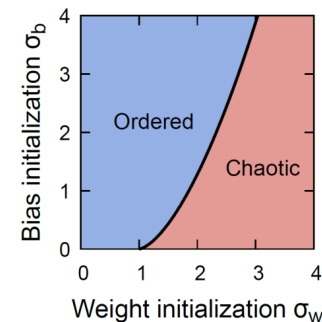
$$\Theta_{\text{train,train}} = \begin{pmatrix} \Theta(x^{[1]}, x^{[1]}) & \dots & \Theta(x^{[1]}, x^{[D]}) \\ \vdots & \ddots & \vdots \\ \Theta(x^{[D]}, x^{[1]}) & \dots & \Theta(x^{[D]}, x^{[D]}) \end{pmatrix}$$



See also: A. Jacot et al. NeurIPS2018, J. Lee et al. NeurIPS2019

Fig. from: L. Xiao et al. ICML2020

それぞれの相での学習ダイナミクス



- 無限幅全結合パーセプトロンの場合、NTK が具体的に解析可能

結果の要点

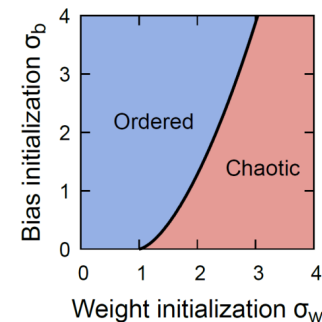
- **秩序相**では、 $\Theta_{\text{train,train}}$ の最大固有値は深さによらず、深層での preactivation 分散 q^* に比例。最小固有値が深さに関して指数関数的に小さくなる

➡ 条件数増大、訓練困難

- **カオス相**では、 $\Theta_{\text{train,train}}$ の最大固有値も最小固有値も深さに関して指数関数的に増大。 q^* に比例もする

➡ 訓練は可能。ただ、 $\Theta_{\text{train,test}}$ がデータセットにほとんど依存せず、汎化しない

それぞれの相での学習ダイナミクス



- 無限幅全結合パーセプトロンの場合、NTK が具体的に解析可能

結果の要点

Lazy learning において、
深層ニューラルネットワークの訓練・汎化には
初期化が臨界点に十分近いことが必要

学習率設計の際に q^* が重要な役割を果たす

深さに関して指数関数的に増大。 q^* に比例もする

- ➡ 訓練は可能。ただ、 $\Theta_{\text{train, test}}$ がデータセットにほとんど依存せず、汎化しない

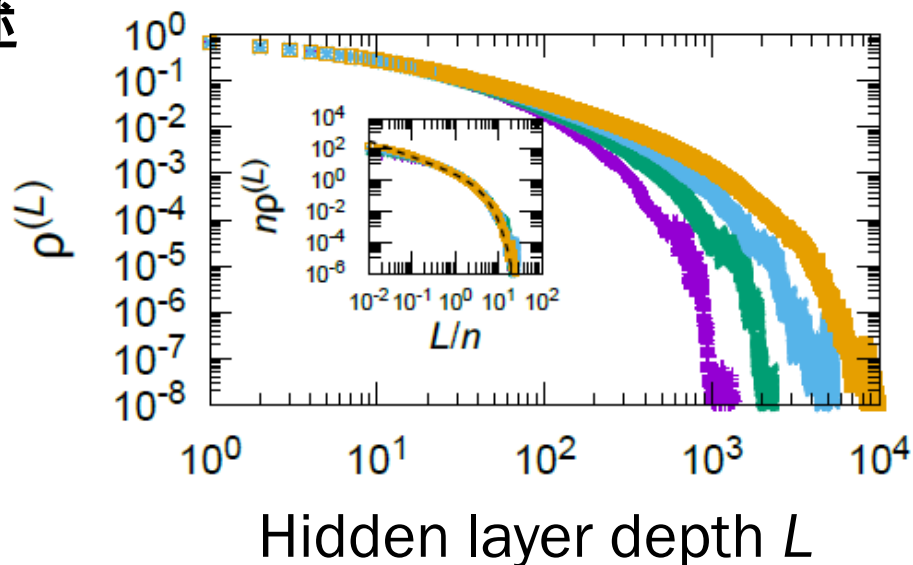
有限幅の場合は？

- 無限幅のように分散・共分散を ($O(n^{-1})$ の誤差内で) 追跡する理論構築も不可能ではない (Edgeworth 展開)
 - ニューロン間の非自明な相関を考慮するため煩雑
- 有限サイズスケーリングの考え方が有効
 - 臨界点直上の現象論的記述

$$\frac{d\rho}{dl} = -\frac{\lambda}{n}\rho - \kappa\rho^2$$

- 普遍スケーリング

$$\rho^{(l)}[\sigma_w; c; n] \simeq n^{-1} f(l/n)$$



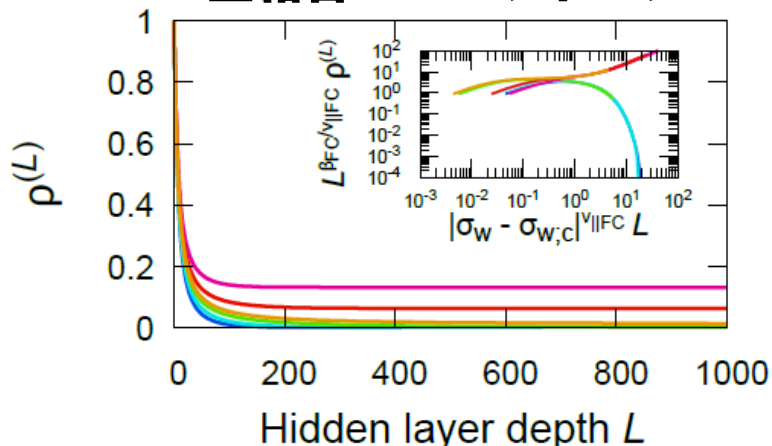
See also: S. Yaida. MSML2020 (Proceeding: PMLR **107**, 165 (2020));
K. T. Grosvenor and R. Jefferson. SciPost Phys. **12**, 081 (2022)

「忘却転移」のスケーリング: 畳み込み NN の場合

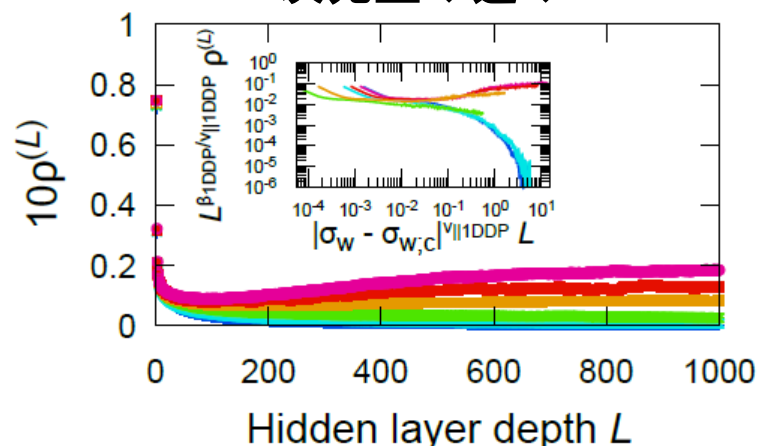
- 畳み込みニューラルネットワークに対しても同様のスケーリング解析を行う
- 「ニューロン間の『相互作用』が局所的」という、アーキテクチャの定性的な違いが普遍クラスに反映
 ➡ **Directed percolation (DP) 普遍クラス**

$$\beta_{1\text{DDP}} \sim 0.276, \quad \nu_{\parallel} \sim 1.734$$

全結合パーセプトロン



1次元畳み込み NN



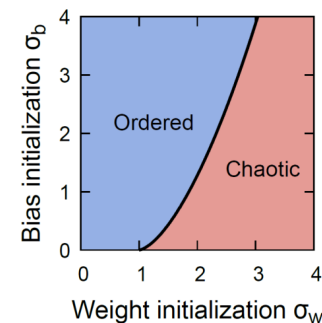
今後の展望・私見

- 実際の学習ではさらに、データ点の分布の仕方に由来するスケールが入るはず
 - 臨界点に初期化したからと言って闇雲に深くしても、性能は出ない。データ・アーキテクチャ・アルゴリズムのすべてを踏まえた設計吟味の必要性は lazy learning でも変わらなそう
 - Neural network scaling law を示すアーキテクチャでは、non-scalability の問題が打破されている？ Yes ならどのように？
- 少なくとも古典的なパーセプトロンや畳み込み NN は、自己組織化臨界を実現していない（普通の相転移）
 - よりモダンなディープラーニング技術（残差ブロック等）の効果は臨界現象の文脈で解釈できるか？

今後の展望・私見

- 経験的には、性能とモデル規模対性能比の間に一種のトレードオフがありそう
 - 熱機関のパワーと効率のトレードオフのように、統計力学で言えることは無いか？
- 深層学習の理論的理解を目指す場合でも、（当面は）数値計算やヒューリスティクスを駆使して進んでいくのが良さそう（物理屋が得意？）
 - Directed percolation は（現時点で）数学的には超難問（厳密解は空間 1 次元ですら未解決）

まとめ



「非平衡臨界現象」をキーワードに、自然知能、人工知能のダイナミクスを整理する試みを紹介

- 吸収状態転移が自然知能・人工知能を統一的に理解する足掛かりとなる可能性を示唆
- 神経系が（擬似的）自己組織化臨界現象を実現している可能性を示唆する理論・実験研究が増えている
- 古典的な人工深層ニューラルネットワークにおける信号伝搬ダイナミクスは吸収状態転移の臨界現象の枠組みで見通しよく整理できる。ただし自己組織化臨界現象は示さない