

Lost in Retraining: Roaming the Parameter Space of Exponential Families Under Closed-Loop Learning

Fariba Jangjoo

Kavli Institute for Systems Neuroscience, NTNU, Trondheim, Norway

Matteo Marsili

ICTP- International Centre for Theoretical Physics, Trieste, Italy

Yasser Roudi

Department of Mathematics, King's College London, London, UK

(Dated: July 10, 2025)

Abstract

Closed-loop learning is the process of repeatedly estimating a model from data generated from the model itself. It is receiving great attention due to the possibility that large neural network models may, in the future, be primarily trained with data generated by artificial neural networks themselves. We study this process for models that belong to exponential families, deriving equations of motions that govern the dynamics of the parameters. We show that maximum likelihood estimation of the parameters endows sufficient statistics with the martingale property and that as a result the process converges to absorbing states that amplify initial biases present in the data. However, we show that this outcome may be prevented if the data contains at least one data point generated from a ground truth model, by relying on maximum *a posteriori* estimation or by introducing regularisation.

Introduction— This paper focuses on the behaviour of statistical models and learning machines subject to *closed-loop* learning: the process in which models are repeatedly trained on data mainly generated by models themselves. When does this process converge and if so to what? These questions have raised considerable recent interest in the context of Large Language models (LLM) and other state of the art neural networks [1–8]. The success of LLMs relies on the availability of huge amounts of high-quality human-generated data. Yet the supply of such data has been argued to be running out [9]. At the same time, more and more AI-generated data is populating the Internet, thereby polluting the datasets with which models will be trained. It has been argued that this process of “data-autophagy” is pathologic with a severity that depends on the fraction of synthetic data used in training [1, 2, 6]. A possible outcome of this is *model collapse* [4] – a phenomenon observed also in Generative Adversarial Networks [10], diffusion models [11] and Variational Autoencoders [4] and LLM [5] – whereby the model generates only a limited variety of outputs, failing to capture the full diversity of the training data. Marchi *et al.* [3] find that closed-loop learning can also lead to the opposite outcome, where models generate random outputs. As we shall argue in the conclusions, close-loop learning is also of relevance for a wide variety of other phenomena, ranging from the evolution of cultures to rational herding [12, 13].

Here, we study closed-loop learning within a simple yet general setting, that of Exponential Family (EF). Distributions in EF take the form (see e.g. [14] sec. 3.4)

$$f(s|\boldsymbol{\theta}) = \frac{1}{Z(\boldsymbol{\theta})} e^{\sum_a \theta_a \phi_a(s)}, \quad Z(\boldsymbol{\theta}) = \sum_s e^{\sum_a \theta_a \phi_a(s)}, \quad (1)$$

where the random variable s takes value in \mathcal{S} , $\boldsymbol{\theta} \in \mathbb{R}^D$ is a vector of D parameters with components θ_a . Each $\phi_a(s)$, $a = 1, \dots, D$, is a bounded functions of s , and Z is the partition function. We aim to understand how the parameters evolve if they are fitted to data generated by the model with the previous set of parameters, including potentially data generated by an external source. As we show, this simplified setting offers a transparent and fully controllable rationalization of a set of phenomena that occur when retraining models on model-generated data, with the potential to unearth general mechanisms that apply to a broader domain.

We first study the dynamics of the model parameters follow in closed-loop learning when they are estimated using Maximum Likelihood (ML) - a widely used approach- and from data exclusively drawn from the model itself. In the case of ML, the relevant variables are

the sufficient statistics of the model [15] which, as we show, acquire a martingale property. This property in turn will be shown to lead to model collapse. We then discuss how this collapse may be avoided. To do this, we consider the more general case in which parameter learning may go beyond ML and may involve a prior distribution/regularization (as in MAP inference). We also consider the possibility that external data (e.g. from a ground truth distribution taken to be human generated data in LLMs) may also be supplied alongside self-generated data during repeated training. For this general setting, we derive the Fokker-Planck equations governing the stochastic dynamics of the model parameters during closed-loop learning. Finding the stationary state distribution of the Fokker-Planck equations, we find that both the presence of priors/regulations and external data may prevent model collapse. In particular, we show that even adding one single datapoint from outside may be sufficient for avoiding model collapse. Surprisingly, these happen even in the limit where, at each retraining iteration, the amount of available self-generated data approaches infinity. This is surprising because, in standard (non-closed-loop) learning, the differences between MAP estimation and maximum likelihood, as well as the effects of priors and regularization, become negligible.

Closed-loop learning in EF^- . Consider the following dynamics in the space of parameters $\boldsymbol{\theta} \in \mathbb{R}^d$. The dynamics starts at $\boldsymbol{\theta}_0$ which we take to be the result of fitting the model to some data. From the initial parameter vector, at iteration $t \geq 0$,

- 1: draw M independent samples, $\hat{s}^1(t), \dots, \hat{s}^M(t)$, from Eq. (1) with $\boldsymbol{\theta} = \boldsymbol{\theta}_t$ and m samples $\hat{s}^{M+1}(t), \dots, \hat{s}^{M+m}(t)$ generated from a distribution $q(s)$. Denoting these samples as $\hat{\mathbf{s}}(t) \equiv (\hat{s}^1(t), \dots, \hat{s}^{M+m}(t))$, their likelihood is

$$p(\hat{\mathbf{s}}(t)|\boldsymbol{\theta}_t) = \prod_{i=1}^M f(\hat{s}^i(t)|\boldsymbol{\theta}_t) \prod_{j=1}^m q(\hat{s}^j(t)).$$

- 2: Given $\hat{\mathbf{s}}(t)$ find the parameters

$$\boldsymbol{\theta}_{t+1} \equiv \hat{\boldsymbol{\theta}}(\hat{\mathbf{s}}(t)) = \arg \max_{\boldsymbol{\theta}} [p(\hat{\mathbf{s}}(t)|\boldsymbol{\theta}) p_0^u(\boldsymbol{\theta})]. \quad (2)$$

loop: Set $t \rightarrow t + 1$ and repeat.

When $u = 0$ the parameter are inferred by Maximum Likelihood (ML) whereas for $u = 1$ inference relies on maximum *a-posteriori* (MAP) inference with a prior $p_0(\boldsymbol{\theta})$. This includes

regularization schemes, e.g., Lasso [16], that can be represented by an appropriate prior. With $m = 0$ the dynamics is one in which data are generated uniquely from the model itself whereas with $m > 0$ new data generated from a distribution $q(\mathbf{s})$ are also included. The distribution q could also be the model generating the original data that determined $\boldsymbol{\theta}_0$ and reflect a *ground truth*

The steps above yield a Markov dynamics on the parameters: at each step the new data is generated consistently with the data just learned. In what follows, we first study this dynamics for $u = m = 0$, that is ML and completely self-generated data. We then study the general case.

Closed-loop learning under ML ($m = u = 0$)–. The components $\phi_a(s)$ in Eq. (1) are the sufficient statistics of the model. This means that for $m = u = 0$ given a sample $\hat{\mathbf{s}}$, the ML estimates, $\hat{\boldsymbol{\theta}}(\hat{\mathbf{s}})$, depend solely on the empirical averages

$$\bar{\phi}_a(\hat{\mathbf{s}}) = \frac{1}{M} \sum_{l=1}^M \phi_a(\hat{s}^l). \quad (3)$$

and are the solutions of the following set of equations

$$\langle \bar{\phi}_a(\mathbf{s}) \rangle_{\hat{\boldsymbol{\theta}}} = \bar{\phi}_a(\hat{\mathbf{s}}), \quad a = 1, \dots, d. \quad (4)$$

where $\mathbf{s} = (s^1, \dots, s^M)$, for $s^l \in \mathcal{S}$ and $\langle \dots \rangle_{\boldsymbol{\theta}}$ is expectation with respect to $p(\mathbf{s}|\boldsymbol{\theta}) = \prod_{l=1}^M f(s^l|\boldsymbol{\theta})$.

In the closed-loop dynamical process described above, $\hat{\mathbf{s}}(t)$ determines $\bar{\phi}_a(\hat{\mathbf{s}}(t))$ and thus, via Eq. (4), determines $\boldsymbol{\theta}_t = \hat{\boldsymbol{\theta}}(\hat{\mathbf{s}}(t))$ and the distribution $p(\cdot|\boldsymbol{\theta}_t)$. Consequently, given $\hat{\mathbf{s}}(t)$ and defining $\bar{\phi}_{a,t} \equiv \bar{\phi}_a(\hat{\mathbf{s}}(t))$, we can write

$$\langle \bar{\phi}_{a,t+1} | \bar{\phi}_{a,t} \rangle_{\boldsymbol{\theta}_t} = \bar{\phi}_{a,t}. \quad (5)$$

where $\langle \dots | \bar{\phi}_{a,t} \rangle_{\boldsymbol{\theta}_t}$ is the average over all possible samples $\hat{\mathbf{s}}(t+1)$ generated at fixed $\boldsymbol{\theta}_t$, which as described above is determined by $\bar{\phi}_{a,t}$.

By definition, Eq. (5) means that $\bar{\phi}_{a,t}$ is a martingale [17]. As shown in Feller [18] (see also Appendix A), a Markov chain $\bar{\phi}_{a,t}$ that is also a martingale in a bounded state space converges to absorbing states. In other words, the fluctuations of $\bar{\phi}_{a,t}$ vanish, i.e. $\phi_a(s)$ attains the same value for all data points $\hat{\mathbf{s}}(t)$, and

$$\lim_{t \rightarrow \infty} f(s|\boldsymbol{\theta}_t) = c \prod_a \delta(\phi_a(s) - \bar{\phi}_{a,\infty}) \quad (6)$$

with c a normalizing constant. This is what is meant by model collapse: while the parameters may be initialized at any value, closed-loop learning prevents the model to represent anything but few possibilities. Below we illustrate this for three example models.

Consider an Ising model defined over spin configurations $s = (\sigma_1, \dots, \sigma_N)$ with $\sigma_i = \pm 1$ as

$$f(s|\theta_1, \theta_2) = \frac{1}{Z} \exp \left[\theta_1 \sum_i \sigma_i + \frac{\theta_2}{N} \sum_{i < j} \sigma_i \sigma_j \right]. \quad (7)$$

With $\theta_2 = 0$ (no interaction) only $\phi_1(s) = \sum_i \sigma_i$ is relevant and $\bar{\phi}_1(\hat{s})$ converges to $\pm N$, with (see Appendix A)

$$\lim_{t \rightarrow \infty} P\{\bar{\phi}_{1,t} = N | \bar{\phi}_{1,0}\} = \frac{\bar{\phi}_{1,0} + N}{2N}; \quad (8)$$

Correspondingly, $\theta_{1,t}$ diverges to $\pm \infty$ and the distribution $f(s|\theta_1, \theta_2 = 0)$ converges to a fully polarized state where in the sample \hat{s} , either $\sigma_i^l = +1$ or $\sigma_i^l = -1$ for all l . Fig. 1(a) shows examples of this dynamics. The empirical cumulative distributions of t_{max} , the time it takes for the dynamics to converge, shown in Fig. 1(b), suggests that $t_{max} \sim NM$. Fig. 1(c) shows simulations results confirming Eq. (8).

If instead $\theta_2 \neq 0$, any $\bar{\phi}_{1,\infty} = N - 2k$ with $k = 0, 1, 2, \dots, N$, can be an absorbing state. For $|\bar{\phi}_{1,\infty}| < N$, this occurs because both θ_1 and θ_2 diverge, i.e. $\theta_2 \rightarrow -\infty$ with a finite ratio $\theta_1/\theta_2 = -\bar{\phi}_{1,\infty}/N$ see SI. In this case, $f(s|\theta_1, \theta_2)$ converges to a uniform distribution on all configurations s with the same value of $\phi_1(s)$. Fig. 1(d) shows that for $\theta_{2,t=0} = 0$, the limiting value of $\bar{\phi}_{1,\infty}$ is not far from the initial value $\bar{\phi}_{1,0}$.

To further illustrate model collapse in closed-loop learning, we next consider a spin model where θ_s is an independent parameter for each of the $d = 2^n$ spin configurations $s = (\sigma_1, \dots, \sigma_N)$. The corresponding operator in Eq. (1) is thus $\phi_s(s') = \delta_{s,s'}$. When θ_s are drawn from a Gaussian distribution, this model coincides with the *Random Energy Model* [19]. Yet here we focus on the closed-loop learning setting with $m = u = 0$, where $\boldsymbol{\theta}_{t+1}$ are the ML estimates inferred from a M samples $\hat{s}(t)$ drawn i.i.d. from the model with parameters $\boldsymbol{\theta}_t$. For $M \ll 2^N$ this is an over-parametrized model. The average of ϕ_s over the data $\hat{s}(t)$ is $\bar{\phi}_{s,t} = k_{s,t}/M$, where $k_{s,t}$ is the number of times configuration s occurs in the data $\hat{s}(t)$, and $f(s|\hat{\boldsymbol{\theta}}(\hat{s}(t))) = \frac{k_{s,t}}{M}$. A state s which is not observed at time t , i.e. if $k_{s,t} = 0$, will thus never be observed forever after. The number of states with $k_{s,t} > 0$ is thus a non-increasing function of time. The asymptotic state will be one where $k_{s^*,t} = M$ for one state and $k_{s,t} = 0$ for all $s \neq s^*$, which correspond to the two absorbing states $\bar{\phi}_s^+ = M$ and $\bar{\phi}_s^- = 0$. Furthermore, it

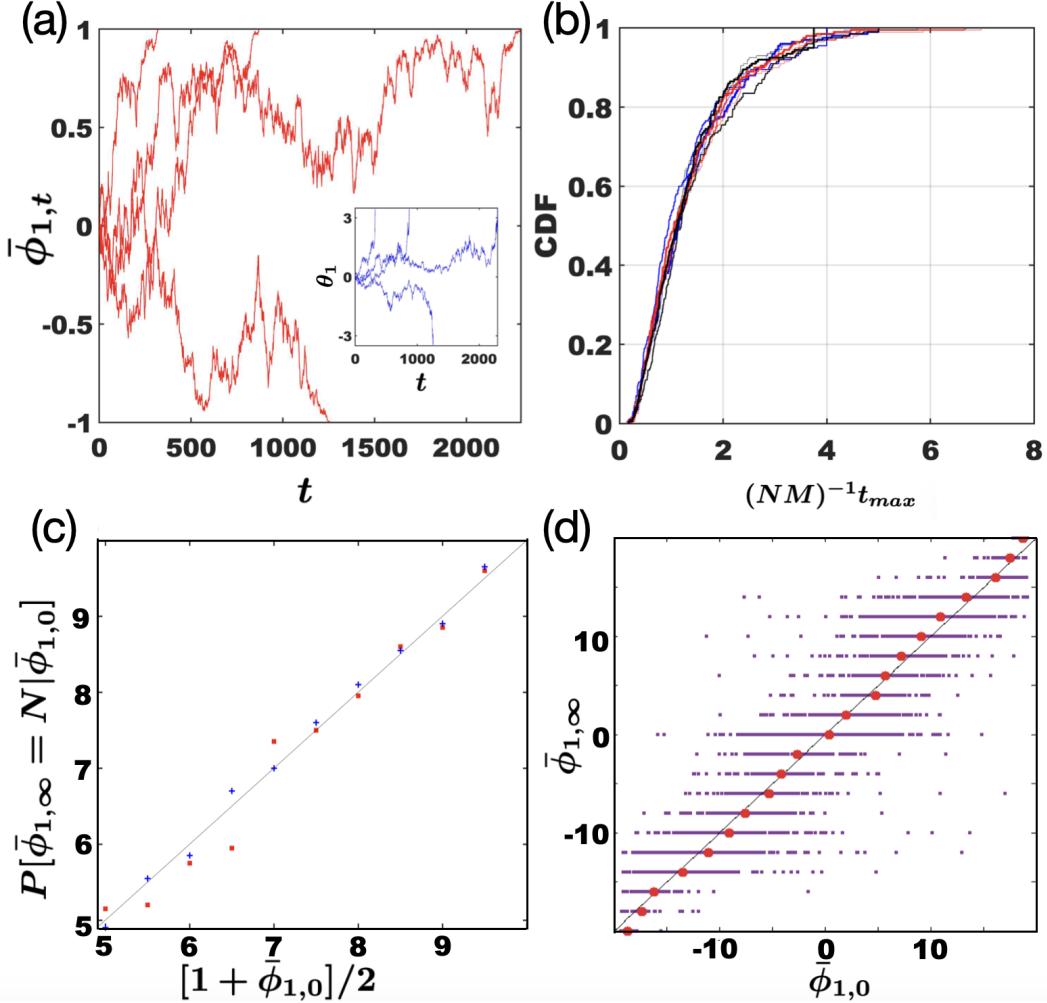


FIG. 1. Closed-loop dynamics and model collapse for $m = u = 0$ in the Ising model (Eq. (7)). (a), (b) and (c) refer to the case $\theta_2 = 0$ whereas in (d), $\theta_2 \neq 0$. (a) Example trajectories of $\phi_{1,t}$, starting at $\phi_{1,0} = 0$; $N = 20$ and $M = 50$. Corresponding trajectories for θ_1 shown in the inset. (b) Cumulative distribution for the convergence time to fully polarized states normalized by NM for $N = 5, 10, 20$ (blue, red, black lines), and $M = 50, 100, 200$ (thinnest to thickest lines). (c) Probability of convergence to all spins up, as a function of $(1 + \bar{\phi}_{1,0})/2$ (see Eq. (8)); each point is calculated from 200 simulations with the same initial $\bar{\phi}_{1,0}$, with $N = 10$; $M = 50$ (red squares) and $M = 100$ (blue circles). (d) The final values of $\bar{\phi}_1$ vs. initial values for $\theta_2 \neq 0$, for 100 different simulations with $N = 20$ and $M = 100$. θ_2 is set to zero at $t = 0$ but updated afterwards. Full circles are averages of $\bar{\phi}_{1,0}$ for the same value of $\bar{\phi}_{1,\infty}$.

can be shown that (see Appendix A)

$$\lim_{t \rightarrow \infty} P\{k_{s,t} = M | k_{s,0}\} = \frac{k_{s,0}}{M}. \quad (9)$$

Summarizing, closed-loop learning in EF based on ML and not data from outside ($m = 0$) leads to a reduction of the diversity of the data. How severe this reduction is depends on the model and the initial conditions.

The general case and continuous time dynamics–. Let us now consider the general case with an arbitrary integer value of m and $u \in \{0, 1\}$.

Between two consecutive iterations, we expect that the difference $\boldsymbol{\theta}_{t+1} - \boldsymbol{\theta}_t$ to be small, i.e. of the same order of the error on the ML estimator $\delta \hat{\boldsymbol{\theta}} \propto M^{-1/2}$. So for large M , we expect the evolution of $\boldsymbol{\theta}_t$ to be well approximated by a continuous time dynamics. We thus consider the transition matrix between t and $t + 1$

$$p(\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}' | \boldsymbol{\theta}_t = \boldsymbol{\theta}) = \sum_{\hat{\boldsymbol{s}}} p(\hat{\boldsymbol{s}} | \boldsymbol{\theta}) \delta(\boldsymbol{\theta}' - \boldsymbol{\theta}_{t+1}(\hat{\boldsymbol{s}})) \quad (10)$$

where $\boldsymbol{\theta}_{t+1}(\hat{\boldsymbol{s}})$ is given by Eq. (2) and the probability of $\hat{\boldsymbol{s}}$ may depend both on $\boldsymbol{\theta} = \boldsymbol{\theta}_t$ and on q , as detailed in step1 of the dynamics described above. In the limit $M \rightarrow \infty$ the transition matrix is dominated by a sharp maximum close to the ML estimator $\hat{\boldsymbol{\theta}}$. Expanding around this point (see Appendix B), to the leading order in M , $\boldsymbol{\theta}_{t+1}$ is well approximated by a Gaussian variable with covariance $J^{-1}(\boldsymbol{\theta}_t)/M$, where

$$J_{a,b}(\boldsymbol{\theta}) = \frac{\partial^2}{\partial \theta_a \partial \theta_b} \log Z = \langle \phi_a \phi_b \rangle_{\boldsymbol{\theta}} - \langle \phi_a \rangle_{\boldsymbol{\theta}} \langle \phi_b \rangle_{\boldsymbol{\theta}} \quad (11)$$

is the Fisher Information Matrix (FIM). The Gaussian has a mean $\boldsymbol{\theta}_t + A(\boldsymbol{\theta}_t)/M$ where

$$A_a(\boldsymbol{\theta}) = - \sum_b J_{a,b}^{-1} \frac{\partial U}{\partial \theta_b} \quad (12)$$

$$U(\boldsymbol{\theta}) = \frac{1}{2} \log \det J - u \log p_0(\boldsymbol{\theta}) + m D_{KL}(q | \boldsymbol{\theta}). \quad (13)$$

Here $D_{KL}(q | \boldsymbol{\theta}) = \sum_s q(s) \log \frac{q(s)}{f(s|\boldsymbol{\theta})}$ is the Kullback-Leibler divergence between $q(s)$ and $f(s|\boldsymbol{\theta})$. Note that in the ML case ($u = m = 0$) the drift term (Eq 12) is non-zero, as a consequence of the fact that ML estimators are, in general, biased [20].

In the rescaled time $\tau = t/M$, the vector of parameters satisfies the (Ito) stochastic differential equation [21] (see Appendix B)

$$d\boldsymbol{\theta}_{\tau} = A(\boldsymbol{\theta}_{\tau}) d\tau + dW_{\tau} \quad (14)$$

$$\langle dW_{\tau} dW_{\tau'} \rangle = J^{-1}(\boldsymbol{\theta}_{\tau}) \delta(\tau - \tau') d\tau. \quad (15)$$

This dynamics is fully consistent with the model collapse discussed in the previous section for $m = u = 0$. Indeed Eqs. (14)-(15) can be used to show that the evolution of the sufficient statistics $\varphi_a \equiv \langle \phi_a(\mathbf{s}) \rangle_{\boldsymbol{\theta}}$ is a purely diffusive dynamics (see Appendix B), eventually hitting and getting absorbed in the boundaries.

The continuous time dynamics indeed shows that closed-loop learning leads to a contraction of the configuration space. For example, for an Ising model with $\theta_2 = 0$, the dynamics of the magnetization $\mu = \langle \bar{\phi}_1 \rangle_{\boldsymbol{\theta}} / N = \langle \sum_i \sigma_i \rangle_{\boldsymbol{\theta}} / N$ per spin is given by $d\mu = \sqrt{1 - \mu^2} dW$, where dW is the differential of a Wiener process. This dynamics clearly has two absorbing states at $\mu = \pm 1$. Indeed $\langle \mu^2 \rangle_{\theta_\tau} = 1 - (1 - \langle \mu^2 \rangle_{\theta_0}) e^{-\tau}$. More generally, one can consider the evolution of the entropy

$$S(\boldsymbol{\varphi}) = \log Z - \sum_a \theta_a \varphi_a, \quad (16)$$

as a function of the conjugate variables $\boldsymbol{\varphi}$ to show that the configuration space, indeed, shrinks. To see this, note that $-S(\boldsymbol{\varphi})$ is the Legendre transform of $\log Z$, in the conjugate coordinates $\boldsymbol{\varphi}$, and hence $\frac{\partial S}{\partial \varphi_a} = -\theta_a$ and $\frac{\partial^2 S}{\partial \varphi_b \partial \varphi_c} = -J_{b,c}^{-1}$ [22]. The differential of the entropy is thus $dS = -\sum_a \theta_a d\varphi_a - D d\tau$. Consequently, the expected value of the entropy changes with time τ as

$$\langle S \rangle_{\boldsymbol{\varphi}_\tau} = \langle S \rangle_{\boldsymbol{\varphi}_0} - \tau D \quad (17)$$

diverges to $-\infty$ as $\tau \rightarrow \infty$. Importantly, Eq. (17) shows that the entropy reduces at a faster rate for higher dimensional models.

In the presence of priors or samples drawn from $q(s)$ ($u = 1$ or $m > 0$), the asymptotic behavior of the dynamics can be studied by deriving the Fokker-Planck equation associated to Eq. (14) using standard techniques (see Appendix B). This shows that the dynamics converges to a stationary state

$$p_{st}(\boldsymbol{\theta}) = c p_0^{2u}(\boldsymbol{\theta}) e^{-2mD_{\text{KL}}(q|\boldsymbol{\theta})}. \quad (18)$$

assuming the normalization constant c is finite. Model collapse can be identified when it is not. In particular, consistent with the previous section, this happens when $u = m = 0$. However, even the addition of one external data point ($m = 1$) or turning on the prior ($u = 1$) may change this. For example, for the Ising model in Eq. (7) with $\theta_2 = 0$ under ML ($u = 0$) and with $m = 1$, model collapse will be avoided, as confirmed numerically in Fig. 2a. In Fig. 2, we also numerically confirm the analytical expression in Eq. (18) for MAP

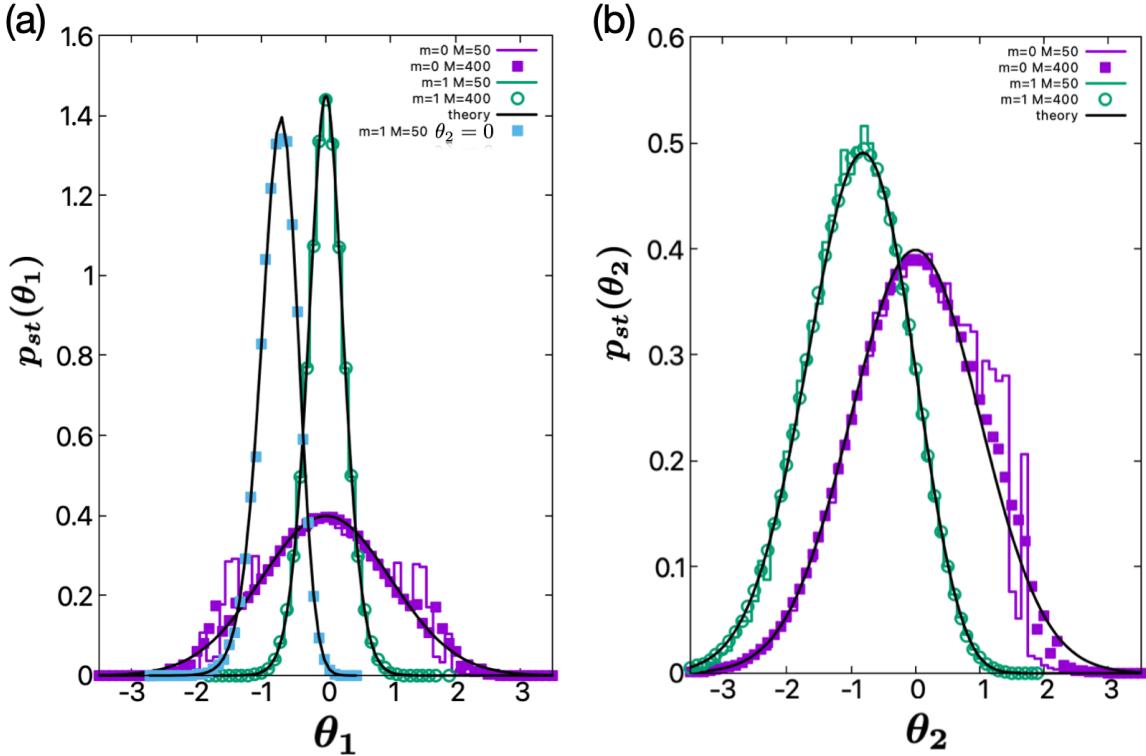


FIG. 2. Numerical simulation confirming Eq. (18) and prevention of model collapse for the Ising model. (a) The distribution of the field θ_1 . (b) The distribution of the coupling θ_2 . We consider three cases all with $N = 10$: *i*) $\theta_2 = 0$ and $M = 50$, with one data point with $\sum_i \sigma_i = -6$ being added ($m = 1$) while $u = 0$ (ML); *ii*) $\theta_2 \neq 0$, $m = 0$ and $M = 50$ and 400 , with Gaussian priors on both θ_1 and θ_2 with zero mean and variance 2 (which is equivalent to $L2$ regularization with parameter $1/4$); *iii*) $\theta_2 \neq 0$, $M = 50$ and 400 , with the same priors and the addition of one data point ($m = 1$) with $\sum_i \sigma_i = 0$. In all cases model collapse is avoided as predicted by Eq. (18) with a good agreement (see full black lines). Data is obtained from simulations of 10^7 iterations of the dynamics.

$(u \neq 0)$ for the Ising model when $\theta_2 \neq 0$, $m = 0$ and $m = 1$, showing that model collapse is again avoided.

In general, whether, or to what degree, using MAP ($u = 1$) or adding external data ($m > 0$) helps avoiding model collapse depends on the details of the model, the choice of the prior and the distribution q . To further demonstrate this, it is instructive to consider another example besides the Ising model discussed above. In Appendix C we consider a Poisson distribution with mean ϑ . In terms of $\theta = \log \vartheta$, the model takes the form in

Eq. (1). Assuming an exponential prior $p_0(\vartheta) = \lambda e^{-\lambda\vartheta}$ and $q(s)$ a Poisson distribution with mean ϑ_0 , one can analytically derive the stochastic differential equation that ϑ follows, which turns out to be a Cox–Ingersoll–Ross process [23]. With $m = 0$, the stationary distribution does not exist and the system converges to the absorbing state $\vartheta_\infty = 0$ even with $\lambda > 0$. This can, however, be prevented even with a small amount of data from outside ($m > 0$) for which the dynamics converges to a stationary state $p_{st}(\vartheta)$ which is a Gamma distribution with shape parameter $2m\vartheta_0$ which coincides with Eq. (18) for θ . These analytical results for the Poisson model are compared and confirmed with simulations in Appendix Fig. 3, where we again show simulation examples where adding only one external data point prevents model collapse, even in the absence of priors or regularization.

Discussion– We studied closed-loop learning for models in the exponential families, analytically deriving the equations governing the stochastic dynamics of the parameters. We found that in this setting, as a result of the martingale property in Eq. (5), model collapse occurs when inference is based on the maximization of the likelihood on data is entirely generated by the model itself. Remarkably, we showed that the introduction of even one single data point generated from a ground truth distribution, or adopting MAP estimates or introducing regularisation can avoid model collapse. In classical statistical theory, the difference between MAP ($u = 1$) and ML ($u = 0$) estimates vanishes for $M \rightarrow \infty$. Likewise, polluting a sample with a finite number m of data points drawn from a different distribution has no effect on the asymptotic behavior of statistical estimates. Our results, however, show that these conclusions do not hold in closed-loop learning.

These results may help shed light in the current debate on closed-loop learning in large artificial network models [1–8]. Although here the martingale property and its consequences were derived and studied for the exponential family, it may hold for other models as well, with similar consequences; see e.g. [1]. Yet, it is important to acknowledge that these models are largely over-parametrized and operate in a regime that is very different from that of classical statistics [24] assumed here. In spite of this, our analysis aligns with observations reported in more complex cases of much recent interest in AI [1–8]. This suggests that our analysis can provide a starting point for addressing issues such as implicit biases arising from algorithms, architectures, and/or initial conditions in artificial neural networks [25–27], providing intuition and testable prediction within an analytically controllable framework. In particular, our results suggest that while ML amplifies biases in the initial conditions,

MAP or the introduction of regularization emphasizes biases implicit in algorithms and/or architectures. Our analysis shows that FIM is what determines drift and diffusion in closed-loop learning. Studying the FIM in deep networks using techniques such as those of [28, 29] may thus help understanding the precise dynamics of closed-loop learning in such networks. We also note that deriving the continuous time dynamics here relied on the fact that the FIM of EF is invertible. Generalizing our analysis to singular models where this does not hold [30] is another interesting avenue of further research.

Dynamics similar to closed-loop learning is present in many domains including word-of-mouth communication, repeated model based imitation [31–33] and rational herding [13], or cultural transmission across generations [12]. Pre-literacy cultural evolution has been based on one generation learning from the previous one, leading to a corpus of “data” of, e.g., mythological nature that shares remarkable similarities across distant cultures [12]. The debate remains open as to whether these similarities stem from shared ancestry and historical migrations [12], from innate archetypal structures or from universal cognitive mechanisms and biases [34]. In social contexts, individuals learn from the observed actions of others, and their resulting behavior is observed by others. This mechanism potentially leads to “informational cascades” that have been argued to determine many collective phenomena [13]. In prenatal development, the brain develops to a large extent in the absence of external stimuli, driven by its own spontaneous activity [35, 36]. What is common in these phenomena is that learning takes place using data generated by another learnt model. Although the process discussed here is very simple and stylized, it offers a platform for studying these phenomena and for engineering models that implement specific functions or that feature pre-assigned structures of asymptotic states.

Acknowledgments—. We are grateful for interesting discussions with Manfred Opper, Fernando Villegas and Claudio Arezzo. We also acknowledge the use of LLM at different stages of this research.

-
- [1] S. Alemohammad *et al.*, Self-consuming generative models go MAD, in *International Conference on Learning Representations (ICLR)*, 2024.
 - [2] Q. Bertrand, A. J. Bose, A. Duplessis, M. Jiralerspong, and G. Gidel, On the stability of

- iterative retraining of generative models on their own data, in *International Conference on Learning Representations (ICLR)*, 2024.
- [3] M. Marchi, S. Soatto, P. Chaudhari, and P. Tabuada, Heat death of generative models in closed-loop learning, in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, pp. 1524–1530, 2024.
 - [4] I. Shumailov *et al.*, *Nature* **631**, 755 (2024).
 - [5] Y. Guo, G. Shang, M. Vazirgiannis, and C. Clavel, The curious decline of linguistic diversity: Training language models on synthetic text, in *Findings of the Association for Computational Linguistics: NAACL 2024*, edited by K. Duh, H. Gomez, and S. Bethard, pp. 3589–3604, 2024.
 - [6] M. Briesch, D. Sobania, and F. Rothlauf, Large language models suffer from their own output: An analysis of the self-consuming training loop, 2024, 2311.16822.
 - [7] E. Dohmatob, Y. Feng, and J. Kempe, Model collapse demystified: The case of regression, in *Advances in Neural Information Processing Systems*, edited by A. Globerson *et al.*, volume 37, pp. 46979–47013, Curran Associates, Inc., 2024.
 - [8] H. A. Vu, G. Reeves, and E. Wenger, arXiv preprint arXiv:2505.21677 (2025).
 - [9] P. Villalobos *et al.*, arXiv preprint arXiv:2211.04325 **3** (2024).
 - [10] H. Thanh-Tung and T. Tran, Catastrophic forgetting and mode collapse in gans, in *2020 international joint conference on neural networks (ijcnn)*, pp. 1–10, IEEE, 2020.
 - [11] G. Biroli, T. Bonnaire, V. De Bortoli, and M. Mézard, *Nature Communications* **15**, 9957 (2024).
 - [12] E. M. Witzel, *The origins of the world's mythologies* (Oxford University Press, 2012).
 - [13] S. Bikhchandani, D. Hirshleifer, and I. Welch, *Journal of political Economy* **100**, 992 (1992).
 - [14] C. M. Bishop and H. Bishop, *Deep learning: Foundations and concepts* (Springer Nature, 2023).
 - [15] G. Casella and R. Berger, *Statistical inference* (CRC press, 2024).
 - [16] R. Tibshirani, *Journal of the Royal Statistical Society Series B: Statistical Methodology* **58**, 267 (1996).
 - [17] É. Roldán *et al.*, *Advances in Physics* **72**, 1 (2023).
 - [18] W. Feller, *An introduction to probability theory and its applications, Vol. 1* (John Wiley & Sons, New York, 1968).

- [19] B. Derrida, Physical Review B **24**, 2613 (1981).
- [20] D. R. Cox and D. V. Hinkley, *Theoretical statistics* (CRC Press, 1979).
- [21] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, Springer Series in Synergetics, 3rd ed. (Springer, Berlin, Heidelberg, 2009).
- [22] S.-i. Amari, *Information geometry and its applications* volume 194 (Springer, 2016).
- [23] J. C. Cox *et al.*, Econometrica **53**, 385 (1985).
- [24] P. L. Bartlett, A. Montanari, and A. Rakhlin, Acta numerica **30**, 87 (2021).
- [25] N. Rahaman *et al.*, On the spectral bias of neural networks, in *International conference on machine learning*, pp. 5301–5310, PMLR, 2019.
- [26] E. Francazi, A. Lucchi, and M. Baity-Jesi, arXiv preprint arXiv:2306.00809 (2023).
- [27] A. Jain, R. Nobahari, A. Baratin, and S. Sarao Mannelli, Advances in Neural Information Processing Systems **37**, 24435 (2024).
- [28] T. Hayase and R. Karakida, The spectrum of fisher information of deep networks achieving dynamical isometry, in *International Conference on Artificial Intelligence and Statistics*, pp. 334–342, PMLR, 2021.
- [29] W. Cao, A. Dytso, M. Fauß, H. V. Poor, and G. Feng, arXiv preprint arXiv:2005.03622 (2020).
- [30] S. Watanabe, The Journal of Machine Learning Research **14**, 867 (2013).
- [31] C. L. Baker, R. Saxe, and J. B. Tenenbaum, Cognition **113**, 329 (2009).
- [32] A. Najar, E. Bonnet, B. Bahrami, and S. Palminteri, PLoS biology **18**, e3001028 (2020).
- [33] C. J. Charpentier, K. Iigaya, and J. P. O'Doherty, Neuron **106**, 687 (2020).
- [34] P. Boyer, *Religion Explained: The Evolutionary Origins of Religious Thought* (Basic Books, New York, 2001).
- [35] F. J. Martini, T. Guillamón-Vivancos, V. Moreno-Juan, M. Valdeolmillos, and G. López-Bendito, Neuron **109**, 2519 (2021).
- [36] R. O. Wong, Annual review of neuroscience **22**, 29 (1999).

Appendix A: Absorption probabilities in martingales

In the dynamical process described in the main text, $\hat{\mathbf{s}}(t)$ determines $\bar{\phi}_a(\hat{\mathbf{s}}(t))$ and thus determines $\hat{\boldsymbol{\theta}}_t$ and the distribution $p(\cdot|\hat{\boldsymbol{\theta}}_t)$. Consequently, given $\hat{\mathbf{s}}(t)$ and thus $\bar{\phi}_{a,t} \equiv \bar{\phi}_a(\hat{\mathbf{s}}(t))$,

we have

$$\langle \bar{\phi}_{a,t+1} | \bar{\phi}_{a,t} \rangle_{\theta_t} = \bar{\phi}_{a,t}. \quad (\text{A1})$$

implying that $\bar{\phi}_{a,t}$ is a martingale. Since the sequence of samples $\hat{s}(0), \hat{s}(1), \dots$ forms a Markov chain, so does $\bar{\phi}_{a,0}, \bar{\phi}_{a,1}, \dots$. Under the assumption that $\phi(s)$ is bounded, the states

$$\phi_a^+ = \max_s \phi_a(s), \quad \phi_a^- = \min_s \phi_a(s) \quad (\text{A2})$$

are absorbing states, for any Markov chain that is also a martingale. In other words, $P\{\bar{\phi}_{a,t+1} | \bar{\phi}_{a,t} = \phi_a^\pm\} = 1$ if $\bar{\phi}_{a,t+1} = \phi_a^\pm$ and $P\{\bar{\phi}_{a,t+1} | \bar{\phi}_{a,t} = \phi_a^\pm\} = 0$ otherwise. In addition, if there are no other absorbing states, the asymptotic probability to be absorbed at ϕ_a^\pm starting from $\bar{\phi}_{a,0}$ is given by

$$\lim_{t \rightarrow \infty} P\{\bar{\phi}_{a,t} = \phi_a^+ | \bar{\phi}_{a,0}\} = \frac{\bar{\phi}_{a,0} - \phi_a^-}{\phi_a^+ - \phi_a^-} \quad (\text{A3})$$

The proof of these statements is straightforward; we recall the argument in Feller [18] (p. 399). For a single variable ϕ taking discrete values in $[\phi_-, \phi_+]$, let $p_{\phi, \phi'} = P(\phi_{t+1} = \phi' | \phi_t = \phi)$ be the transition probability matrix. The martingale property implies

$$\langle \phi_{t+1} | \phi_t \rangle = \sum_{\phi'} p_{\phi_t, \phi'} \phi' = \phi_t.$$

For $\phi_t = \phi_-$ this necessarily implies that $p_{\phi_-, \phi'} = 0$ for all $\phi' > \phi_t$ and $p_{\phi_-, \phi_-} = 1$, so ϕ_- is an absorbing state. The same is true for $\phi_t = \phi_+$.

Let $p_{\phi, \phi'}^{(t)}$ be the ϕ, ϕ' matrix element of the t^{th} power of the transition matrix. If there is no other absorbing state, then $p_{\phi, \phi'}^{(t)} \rightarrow 0$ for $t \rightarrow \infty$, for all $\phi' \neq \phi_\pm$. The iteration of the martingale property infinite times then implies that

$$\lim_{t \rightarrow \infty} \left[p_{\phi_0, \phi_-}^{(t)} \phi_- + p_{\phi_0, \phi_+}^{(t)} \phi_+ \right] = \phi_0$$

Eq. (A3) follows from the fact that $p_{\phi_0, \phi_-}^{(t)} = 1 - p_{\phi_0, \phi_+}^{(t)}$.

Absorption probabilities in Ising model. In the case of the Ising model, $\bar{\phi}_1(\hat{s}) = M^{-1} \sum_{l,j=1}^{M,N} \hat{\sigma}_j^l$. From Eq. (A2) above, we have $\phi_1^\pm = \pm N$ and thus using Eq. (A3) above, we retrieve Eq. (8) in the main text for the case of $\theta_2 = 0$. If, instead $\theta_2 \neq 0$, we can write

$$f(s | \theta_1, \theta_2) = \frac{1}{Z'} \exp \left[\frac{\theta_2}{2N} (\phi_1(s) - \phi^*)^2 \right], \quad \phi^* = -N \frac{\theta_1}{\theta_2} \quad (\text{A4})$$

with Z' a normalization constant. Consequently, θ_1 and θ_2 are both coupled to the same order parameter $\phi_1 = \sum_i \sigma_i$. For a given N and $\phi^* = N - 2k$ with $k = 0, 1, 2, \dots, N$, $f(s|\theta_1, \theta_2)$ in Eq. (A4) the distribution of ϕ_1 becomes more and more concentrated around ϕ^* , as $\theta_2 \rightarrow -\infty$ (with $\theta_1 = -\theta_2\phi^*/N \rightarrow \infty$ also diverging). This limit is realized as soon as, at some time step t , $\phi_1(\hat{s}^l(t)) = \phi^*$ for all data points $\hat{s}(l)(t)$, so that $\bar{\phi}_{1,t} = \phi^*$. For all $t' > t$ we will also have $\bar{\phi}_{1,t'} = \phi^*$ and thus ϕ^* is an absorbing state for $\bar{\phi}_{1,t}$. As a result, for θ_1 and θ_2 diverging with a finite ratio $\theta_1/\theta_2 = -\phi^*/N$, the system thus has $N + 1$ absorbing states, corresponding to the $N + 1$ different values that ϕ^* can take.

Absorption probabilities in REM. Similar to the Ising model discussed in the previous section, the Random Energy model (REM) is also defined over configurations of N spins $s = (\sigma_1, \dots, \sigma_N)$ with $\sigma_i = \pm 1$. It assigns a different parameter – the energy $E(s)$ – to each of the 2^n spin configuration, and the probability of each state is proportional to $\exp[-E(s)]$. In the notation of Eq. (1) of the main text, for each configuration s , we have a corresponding θ_s and ϕ_s and

$$\phi_s(s') = \delta_{s,s'} \quad (\text{A5})$$

The model is an over-parametrized model in the sense that there is corresponding parameter for each spin configuration.

Closed loop learning in this case is straightforward. For \hat{s} comprised of M data points, let us denote by k_s the number of times configuration s has been observed. Since $\sum_s k_s = M$, we have $\bar{\phi}_s(\hat{s}) = k_s$. Starting from a sample where state s is observed $k_{s,0}$ times, with $\sum_s k_{s,0} = M$, at iteration t , let $k_{s,t}$ be the count of state s . Maximum likelihood returns the estimate $\exp[\hat{\theta}_s^t]/Z = \mu_t(s) = \frac{k_{s,t}}{M}$. From this a new sample $\hat{s}^{(t+1)}$ is drawn. It is easy to see that if a state s fails to be observed at time t , i.e. if $k_{s,t} = 0$, then $\hat{\theta}_{s,t} \rightarrow -\infty$, which means that state s will never be observed forever after. Then the number of states with $k_{s,t} > 0$ is a non-increasing function of time. The asymptotic state will be one where $k_{s^*,t} = M$ for one state and $k_{s,t} = 0$ for all $s \neq s^*$. Alternatively, by noting that for each s , $\bar{\phi}_s^+ = M$ and $\bar{\phi}_s^- = 0$, and that $\bar{\phi}_{s,t} = k_{s,t}$, Eq. (A3) above can then be written as

$$\lim_{t \rightarrow \infty} P\{k_{s,t} = M | k_{s,0}\} = \frac{k_{s,0}}{M}. \quad (\text{A6})$$

which is Eq. (9) of the main text. Eq. (A3) clearly applies to the reduced process $\tilde{s} = 1$ for state s and $\tilde{s} = 0$ for any other state $s' \neq s$.

Appendix B: The continuous time limit

Let us begin by deriving an expression for the transition matrix

$$p(\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}' | \boldsymbol{\theta}_t = \boldsymbol{\theta}) = \sum_{\hat{s}} f(\hat{s}|\boldsymbol{\theta}) \delta(\boldsymbol{\theta}' - \hat{\boldsymbol{\theta}}(\hat{s})). \quad (\text{B1})$$

In order to do so, we first write, assuming $M \gg 1$,

$$p(\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}' | \boldsymbol{\theta}_t = \boldsymbol{\theta}) = \sum_{\hat{s}} \frac{f(\hat{s}|\boldsymbol{\theta})}{f(\hat{s}|\hat{\boldsymbol{\theta}})} f(\hat{s}|\hat{\boldsymbol{\theta}}) \delta(\boldsymbol{\theta}' - \hat{\boldsymbol{\theta}}) \quad (\text{B2})$$

$$\cong \sum_{\hat{s}} e^{-\frac{M}{2}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\hat{J}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})} f(\hat{s}|\hat{\boldsymbol{\theta}}) \delta(\boldsymbol{\theta}' - \hat{\boldsymbol{\theta}}) \quad (\text{B3})$$

$$= e^{-\frac{M}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}')J'(\boldsymbol{\theta}-\boldsymbol{\theta}')} \sum_{\hat{s}} f(\hat{s}|\hat{\boldsymbol{\theta}}) \delta(\boldsymbol{\theta}' - \hat{\boldsymbol{\theta}}) \quad (\text{B4})$$

here $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\hat{s})$ is a function of \hat{s} . In the first line, the term $\frac{f(\hat{s}|\boldsymbol{\theta})}{f(\hat{s}|\hat{\boldsymbol{\theta}})}$ is non-zero only for $\boldsymbol{\theta} \approx \hat{\boldsymbol{\theta}}$, so we expand the (log of the) numerator around $\hat{\boldsymbol{\theta}}$ and retain only terms up to 2nd order (with an error which is of order $1/M$). In the second line,

$$J_{a,b}(\boldsymbol{\theta}) = \frac{\partial^2}{\partial \theta_a \partial \theta_b} \log Z = \langle \phi_a \phi_b \rangle_{\boldsymbol{\theta}} - \langle \phi_a \rangle_{\boldsymbol{\theta}} \langle \phi_b \rangle_{\boldsymbol{\theta}} \quad (\text{B5})$$

is the Fisher Information matrix and we use the shorthands $\hat{J} = J(\hat{\boldsymbol{\theta}})$ and $J' = J(\boldsymbol{\theta}')$. In Eq. (B4) above we substitute $\hat{\boldsymbol{\theta}}$ with $\boldsymbol{\theta}'$, exploiting the presence of the δ -function. For the last term, we use the identity

$$\int d\boldsymbol{\theta} g(\boldsymbol{\theta}) = \sum_{\hat{s}} \int d\boldsymbol{\theta} g(\boldsymbol{\theta}) f(\hat{s}|\boldsymbol{\theta}) \cong \left(\frac{2\pi}{M} \right)^{D/2} \sum_{\hat{s}} \frac{f(\hat{s}|\hat{\boldsymbol{\theta}})g(\hat{\boldsymbol{\theta}})}{\sqrt{\det J(\hat{\boldsymbol{\theta}})}} \quad (\text{B6})$$

where $g(\boldsymbol{\theta})$ is a generic function. This identity follows by estimating the integral in the middle expression by the saddle point method for $M \gg 1$. Using $g(\boldsymbol{\theta}) = \left(\frac{2\pi}{M} \right)^{-d/2} \sqrt{\det J(\boldsymbol{\theta})} \delta(\boldsymbol{\theta}-\boldsymbol{\theta}')$ in Eq. (B4), we find

$$p(\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}' | \boldsymbol{\theta}_t = \boldsymbol{\theta}) \cong \left(\frac{M}{2\pi} \right)^{-D/2} \sqrt{\det J'} e^{-\frac{M}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}')J'(\boldsymbol{\theta}-\boldsymbol{\theta}')} \quad (\text{B7})$$

$$\cong \left(\frac{M}{2\pi} \right)^{-D/2} \sqrt{\det J} e^{-\frac{M}{2}(\boldsymbol{\theta}'-\boldsymbol{\theta})J(\boldsymbol{\theta}'-\boldsymbol{\theta}) - \frac{1}{2}(\boldsymbol{\theta}'-\boldsymbol{\theta})\alpha} \quad (\text{B8})$$

where we expanded the term $\sqrt{\det J'} \cong \sqrt{\det J} e^{\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}')\alpha}$ with

$$\alpha_a(\boldsymbol{\theta}) = \frac{\partial}{\partial \theta_a} \log \det J(\boldsymbol{\theta}),$$

and substituted J' with J in the quadratic term at the exponent, neglecting sub-leading corrections. This is a Gaussian distribution for $\boldsymbol{\theta}'$ with mean and variance

$$\langle \theta'_a \rangle_{\boldsymbol{\theta}} = \theta_a - \frac{1}{2M} \sum_b J_{a,b}^{-1} \frac{\partial}{\partial \theta_b} \log \det J(\boldsymbol{\theta}), \quad \langle (\theta'_a - \langle \theta'_a \rangle_{\boldsymbol{\theta}})(\theta'_b - \langle \theta'_b \rangle_{\boldsymbol{\theta}}) \rangle_{\boldsymbol{\theta}} = J_{a,b}^{-1}. \quad (\text{B9})$$

In the continuous time limit, this leads to the stochastic differential equation

$$d\theta_a = -\frac{1}{2} \sum_b J_{a,b}^{-1} \frac{\partial}{\partial \theta_b} \log \det J(\boldsymbol{\theta}) d\tau + dW_a, \quad \langle dW_{a,\tau} dW_{a,\tau'} \rangle_{\boldsymbol{\theta}_\tau} = J_{a,a}^{-1}(\boldsymbol{\theta}_\tau) \delta(\tau - \tau') d\tau. \quad (\text{B10})$$

This corresponds to a stochastic differential equation for the expected value of the sufficient statistics $\varphi_a \equiv \langle \phi_a(\mathbf{s}) \rangle_{\boldsymbol{\theta}} = \frac{\partial \log Z}{\partial \theta_a}$ that is readily obtained with Ito calculus

$$d\varphi_a = \sum_b \frac{\partial \varphi_a}{\partial \theta_b} d\theta_b + \frac{1}{2} \sum_{b,c} \langle dW_{b,\tau} dW_{c,\tau'} \rangle_{\boldsymbol{\theta}_\tau} \frac{\partial^2 \varphi_a}{\partial \theta_b \partial \theta_c} d\tau \quad (\text{B11})$$

$$= \sum_b J_{a,b} \left[-\frac{1}{2} \sum_c J_{b,c}^{-1} \frac{\partial \log \det \hat{J}}{\partial \theta_c} d\tau + dW_b \right] + \frac{1}{2} \sum_{b,c} J_{b,c}^{-1} \frac{\partial^2 \varphi_a}{\partial \theta_b \partial \theta_c} d\tau \quad (\text{B12})$$

$$= d\xi_{a,\tau} \quad (\text{B13})$$

where $d\xi_{a,\tau} = \sum_b J_{a,b} dW_{b,\tau}$ is a stochastic noise differential with covariance

$$\langle d\xi'_{a,\tau} d\xi'_{b,\tau'} \rangle_{\boldsymbol{\varphi}_\tau} = J_{a,b} \delta(\tau - \tau') d\tau \quad (\text{B14})$$

In the derivation of Eq. (B13) we used the identities

$$\frac{\partial}{\partial \theta_a} \log \det J(\boldsymbol{\theta}) = \sum_{b,c} J_{b,c}^{-1} \frac{\partial J_{b,c}}{\partial \theta_a}, \quad \frac{\partial^2 \varphi_a}{\partial \theta_b \partial \theta_c} = \frac{\partial^3 \log Z}{\partial \theta_a \partial \theta_b \partial \theta_c} = \frac{\partial J_{b,c}}{\partial \theta_a}. \quad (\text{B15})$$

Note that Eq. (B13) is drift-less, as it should be given the martingale property.

In the general case, the transition matrix is given by

$$p(\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}' | \boldsymbol{\theta}_t = \boldsymbol{\theta}) = \sum_{\hat{\mathbf{s}}, \hat{\mathbf{t}}} p(\hat{\mathbf{s}} | \boldsymbol{\theta}) q(\hat{\mathbf{t}}) \delta(\boldsymbol{\theta}' - \bar{\boldsymbol{\theta}}(\hat{\mathbf{s}}, \hat{\mathbf{t}})) \quad (\text{B15})$$

where $\bar{\boldsymbol{\theta}}$ is determined via

$$\bar{\boldsymbol{\theta}}(\hat{\mathbf{s}}, \hat{\mathbf{t}}) = \arg \max_{\boldsymbol{\vartheta}} \mathcal{G}(\boldsymbol{\vartheta}, \hat{\mathbf{s}}, \hat{\mathbf{t}}), \quad \mathcal{G}(\boldsymbol{\vartheta}, \hat{\mathbf{s}}, \hat{\mathbf{t}}) = \log f(\hat{\mathbf{s}} | \boldsymbol{\vartheta}) + \log f(\hat{\mathbf{t}} | \boldsymbol{\vartheta}) + u \log p_0(\boldsymbol{\vartheta}) \quad (\text{B16})$$

With $m = u = 0$ this problem reduces to the case where $\bar{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\hat{\mathbf{s}})$ is the maximum likelihood (ML) estimator and all samples are generated from the ML distribution at the previous step. With $m > 0$ this problem envisages the possibility that some samples $\hat{\mathbf{t}}$ are generated from

the “true” distribution $q(\cdot)$ while the others are generated from $\bar{\boldsymbol{\theta}}$ and with $u = 1$ we look at the case where $\bar{\boldsymbol{\theta}}$ is the maximum *a posteriori* (MAP) estimator, with prior p_0 .

With finite m and $M \gg 1$, the maximum is dominated by the ML estimator $\hat{\boldsymbol{\theta}}(\hat{\mathbf{s}})$, so we expand around it:

$$\mathcal{G}(\boldsymbol{\vartheta}) \simeq \log \left[f(\hat{\mathbf{s}}|\hat{\boldsymbol{\theta}})f(\hat{\mathbf{t}}|\hat{\boldsymbol{\theta}}) \right] + u \log p_0(\hat{\boldsymbol{\theta}}) - \frac{M}{2}(\boldsymbol{\vartheta} - \hat{\boldsymbol{\theta}})^T \hat{\mathbf{J}}(\boldsymbol{\vartheta} - \hat{\boldsymbol{\theta}}) + \boldsymbol{\alpha}^T(\boldsymbol{\vartheta} - \hat{\boldsymbol{\theta}}) + \dots , \quad (\text{B17})$$

where $\hat{\mathbf{J}} = \mathbf{J}(\hat{\boldsymbol{\theta}})$ and $\tilde{\boldsymbol{\alpha}} \in \mathbb{R}^d$ is a vector with components

$$\tilde{\alpha}_a = \frac{\partial}{\partial \vartheta_a} \log f(\hat{\mathbf{t}}|\boldsymbol{\vartheta}) \Big|_{\hat{\boldsymbol{\theta}}} + u \frac{\partial}{\partial \vartheta_a} \log p_0(\boldsymbol{\vartheta}) \Big|_{\hat{\boldsymbol{\theta}}} . \quad (\text{B18})$$

To leading order in M , the maximum is given by

$$\bar{\boldsymbol{\theta}}(\mathbf{s}, \mathbf{t}) = \hat{\boldsymbol{\theta}}(\mathbf{s}) + \frac{\hat{\mathbf{J}}^{-1} \tilde{\boldsymbol{\alpha}}}{M} . \quad (\text{B19})$$

Therefore, in the continuous time limit, the stochastic differential equation acquires an additional drift

$$\mathbf{J}^{-1} \langle \tilde{\boldsymbol{\alpha}} \rangle_{\boldsymbol{\theta}} = \mathbf{J}^{-1} \frac{\partial}{\partial \boldsymbol{\theta}} [u \log p_0(\boldsymbol{\theta}) - m D_{\text{KL}}(q|\boldsymbol{\theta})]$$

, so that

$$d\theta_a = - \sum_b J_{a,b}^{-1} \frac{\partial}{\partial \theta_b} \left[\frac{1}{2} \log \det \mathbf{J}(\boldsymbol{\theta}) - u \log p_0(\boldsymbol{\theta}) + m D(q|\boldsymbol{\theta}) \right] d\tau + dW_a , \quad (\text{B20})$$

$$\langle dW_{a,\tau} dW_{a,\tau'} \rangle_{\boldsymbol{\theta}} = J_{a,b}^{-1}(\boldsymbol{\theta}) \delta(\tau - \tau') d\tau \quad (\text{B21})$$

as in Eqs. (14)-(15) of the main text.

1. The stationary state

As discussed in Gardiner [21], the Fokker-Planck equation associated to the stochastic differential equation (14) in the Letter can be written as

$$\frac{\partial p}{\partial t} = \sum_a \frac{\partial F_a}{\partial \theta_a} \quad (\text{B22})$$

where the probability current \mathbf{F} is given by

$$F_a = -A_a p + \frac{1}{2} \sum_b \frac{\partial}{\partial \theta_b} J_{a,b}^{-1} p = - \left[A_a - \frac{1}{2} \sum_b \frac{\partial J_{a,b}^{-1}}{\partial \theta_b} \right] p + \frac{1}{2} \sum_b J_{a,b}^{-1} \frac{\partial p}{\partial \theta_b} \quad (\text{B23})$$

The zero current condition $\mathbf{F} = 0$ reads

$$2A_a - \sum_b \frac{\partial J_{a,b}^{-1}}{\partial \theta_b} = \sum_b J_{a,b}^{-1} \frac{\partial \log p}{\partial \theta_b} \quad (\text{B24})$$

or

$$\frac{\partial \log p}{\partial \theta_b} = \sum_a J_{b,a} \left[2A_a - \sum_c \frac{\partial J_{a,c}^{-1}}{\partial \theta_c} \right] \quad (\text{B25})$$

$$= -2 \frac{\partial U}{\partial \theta_b} - \sum_{a,c} J_{b,a} \frac{\partial J_{a,c}^{-1}}{\partial \theta_c} \quad (\text{B26})$$

$$= -2 \frac{\partial U}{\partial \theta_b} + \sum_{a,c} J_{a,c}^{-1} \frac{\partial J_{a,b}}{\partial \theta_c} \quad (\text{B27})$$

$$= -2 \frac{\partial U}{\partial \theta_b} + \sum_{a,c} \frac{\partial \log \det J}{\partial J_{a,c}} \frac{\partial J_{a,c}}{\partial \theta_b} \quad (\text{B28})$$

$$= \frac{\partial}{\partial \theta_b} [-2U + \log \det J] \quad (\text{B29})$$

where

$$U(\boldsymbol{\theta}) = \frac{1}{2} \log \det J - u \log p_0(\boldsymbol{\theta}) + m D_{\text{KL}}(q|\boldsymbol{\theta})$$

is the same as Eq. (13) in the main text and we used the identities

$$\sum_{a,c} J_{b,a} \frac{\partial J_{a,c}^{-1}}{\partial \theta_c} = \frac{\partial}{\partial \theta_c} \left[\sum_{a,c} J_{b,a} J_{a,c}^{-1} \right] - \sum_{a,c} \frac{\partial J_{a,b}}{\partial \theta_c} J_{a,c}^{-1} = - \sum_{a,c} J_{a,c}^{-1} \frac{\partial J_{a,b}}{\partial \theta_c} \quad (\text{B30})$$

$$\frac{\partial J_{a,b}}{\partial \theta_c} = \frac{\partial^3 \log Z}{\partial \theta_a \partial \theta_b \partial \theta_c} = \frac{\partial J_{a,c}}{\partial \theta_b} \quad (\text{B31})$$

$$J_{a,c}^{-1} = \frac{\partial \log \det J}{\partial J_{a,c}} \quad (\text{B32})$$

where Z is the normalization constant in Eq. (1) of the main text. In Eq. (B30) we used the fact that the term in square brackets is the identity matrix and hence its derivative vanishes, and then we used identity (B31) to swap indices $c \leftrightarrow b$.

Therefore, the stationary state distribution is

$$p_{st}(\boldsymbol{\theta}) = c \det J e^{-2U(\boldsymbol{\theta})} \quad (\text{B33})$$

with c a normalization constant. The stationary state exists as long as p_{st} is normalizable, i.e. as long as c is finite. This yields Eq. (18) in the main text.

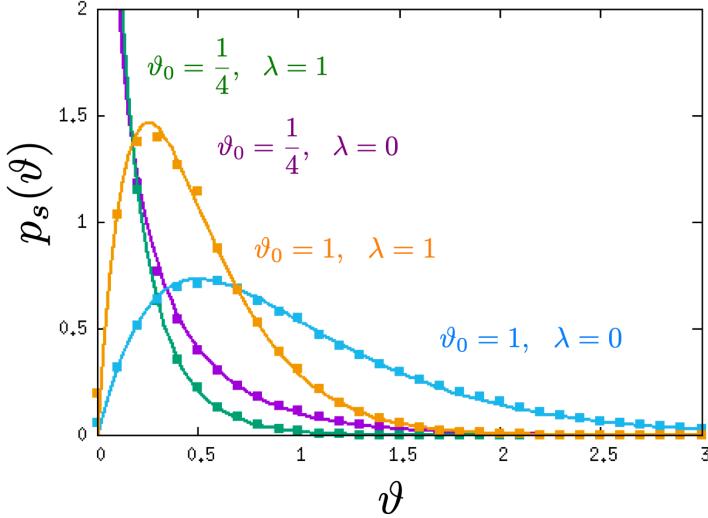


FIG. 3. Distribution of ϑ for the Poisson distribution with $m = 1$, $M = 200$ and different values of λ controlling the prior, and ϑ_0 the mean of the external data from the distribution q . Points refer to sampling of the closed-loop retraining process and full lines refer to the analytical solution Eq. (C2). As discussed in the text below, when $\lambda \rightarrow 0$, we have the case of $u = 0$ that is no regularization or priors, and thus with $m = 0$, model collapse occurs. The results in this figure shows how adding one data point prevents this.

Appendix C: The case of Poisson distribution

Let us consider the case of a Poisson distribution

$$p(s|\vartheta) = \frac{\vartheta^s}{s!} e^{-\vartheta}$$

with mean $\langle s \rangle_\vartheta = \vartheta \geq 0$ and prior $p_0(\vartheta) = \lambda e^{-\lambda\vartheta}$. Since the distribution of s depends on a single parameter we do not use boldface characters for parameters. We use notation ϑ instead of θ because $p(s|\vartheta)$ is an EF but it is not expressed in terms of the natural parameter $\theta = \log \vartheta$, which is also why we do not use the notation $f(s|\theta)$. Finally, to simplify the notation, we shall redefine $u\lambda = \lambda$, so that the case $u = 0$ is obtained in the limit $\lambda \rightarrow 0$.

With $p_0(\vartheta) = \lambda e^{-\lambda\vartheta}$ and $\hat{\mathbf{t}}$ being m i.i.d. samples from a Poisson distribution with mean ϑ_0 , the optimization of

$$\mathcal{G}(\boldsymbol{\vartheta}, \hat{\mathbf{s}}, \hat{\mathbf{t}}) = \log p(\hat{\mathbf{s}}|\boldsymbol{\vartheta}) + \log p(\hat{\mathbf{t}}|\boldsymbol{\vartheta}) + \log p_0(\boldsymbol{\vartheta})$$

yields

$$\frac{\partial \mathcal{G}}{\partial \vartheta} = \frac{S+T}{\vartheta} - M - m - \lambda \quad \Rightarrow \quad \bar{\vartheta} = \frac{S+T}{M+m+\lambda}, \quad S = \sum_{i=1}^M \hat{s}_i, \quad T = \sum_{i=1}^m \hat{t}_i.$$

Notice that $\bar{\vartheta} = 0$ is not an absorbing state when $m > 0$ because even if $S = T = 0$ at some t , so that $\vartheta_{t+1} = 0$, the m samples $\hat{\mathbf{t}}$ will again be drawn from a Poisson distribution with mean $\vartheta_0 > 0$. Conversely, for $m = 0$ as soon as $S = 0$ for some t , $\vartheta_{t'} = 0$ for all $t' > t$.

The conditional distribution of ϑ' given ϑ is given by

$$\begin{aligned} p(\vartheta_{t+1} = \vartheta' | \vartheta_t = \vartheta) &= \sum_{S=0}^{\infty} \frac{(M\vartheta)^S}{S!} e^{-M\vartheta} \sum_{T=0}^{\infty} \frac{(m\vartheta_0)^S}{S!} e^{-M\vartheta} \delta\left(\vartheta' - \frac{S+T}{M+m+\lambda}\right) \\ &= \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{ik\vartheta'} \exp[M\vartheta(1 - e^{-ik/(M+m+\lambda)})] \exp[m\vartheta_0(1 - e^{-ik/(M+m+\lambda)})] \\ &\cong \int_{-\infty}^{\infty} \frac{dk}{2\pi} \exp\left[ik\left(\vartheta' - \frac{M\vartheta + m\vartheta_0}{M+m+\lambda}\right) - \frac{M\vartheta + m\vartheta_0}{2(M+m+\lambda)^2} k^2 + O(k^3/M^2)\right] \\ &\cong \frac{M+m+\lambda}{\sqrt{2\pi(M\vartheta + m\vartheta_0)}} \exp\left[-\frac{1}{2} \frac{(\vartheta' - \langle \vartheta' \rangle_{\vartheta})^2}{\langle (\vartheta' - \langle \vartheta' \rangle_{\vartheta})^2 \rangle}\right] \end{aligned}$$

which is a Gaussian with mean and variance

$$\langle \vartheta' \rangle_{\vartheta} = \frac{M\vartheta + m\vartheta_0}{M+m+\lambda} = \vartheta + \frac{m+\lambda}{M+m+\lambda} \left(\frac{m\vartheta_0}{m+\lambda} - \vartheta \right), \quad \langle (\vartheta' - \langle \vartheta' \rangle_{\vartheta})^2 \rangle = \frac{M\vartheta + m\vartheta_0}{(M+m+\lambda)^2}$$

respectively. In the limit $M \rightarrow \infty$, in rescaled time, the parameter ϑ follows the stochastic differential equation

$$d\vartheta_{\tau} = (m+\lambda) \left[\frac{m\vartheta_0}{m+\lambda} - \vartheta_{\tau} \right] d\tau + \sqrt{\vartheta_{\tau}} dW \quad (\text{C1})$$

which is a Cox-Ingersoll-Ross model [23]. When $2m\vartheta_0 > 0$ the process admits a stationary distribution

$$p_{st}(\vartheta) = \frac{[2(m+\lambda)]^{2m\vartheta_0}}{\Gamma(2m\vartheta_0)} \vartheta^{2m\vartheta_0-1} e^{-2(m+\lambda)\vartheta} \quad (\text{C2})$$

which is given by a Gamma distribution.

The results of Eq. (C2) are in excellent agreement with numerical simulations, as shown in Fig. 3. Note that with $m = 0$ the distribution is not normalizable, i.e. $\vartheta = 0$ is an absorbing state for all $\lambda > 0$ when $m = 0$. Finally we note that when the Poisson distribution is expressed in terms of the canonical parameter $\theta = \log(s)_{\theta}$, Eq. (B33) leads to a stationary state

$$p_{st}(\theta) = \frac{[2(m+\lambda)]^{2m\vartheta_0}}{\Gamma(2m\vartheta_0)} e^{2m\vartheta_0\theta - 2(m+\lambda)e^{\theta}} \quad (\text{C3})$$

that coincides with Eq. (18) of the main text, because $D_{KL}(q|\theta) = \vartheta_0 \log \frac{\vartheta_0}{\vartheta} + \vartheta - \vartheta_0$ in this case.