UNIVERSALITY AND CONVERGENCE OF GENERATIVE FLOWS

Anonymous authors
Paper under double-blind review

ABSTRACT

Generative flows sample from an unnormalized target by training a flow to be balanced, and the training loss is the signal a practitioner watches. We ask what that signal is worth: whether a small loss certifies an accurate sampler, whether the loss can be driven to zero, and how fast gradient descent does so. The loss decides the first. Losses that compare the two sides of the balance by their difference bound, in total variation, the error of the sampler the flow implies, with explicit constants that do not involve the policy; flow-matching losses that compare them through a ratio admit no such bound, already on a single cycle, whenever their generator is continuous at balance. On graphs, the backward policy decides the other two. Once it is frozen, balance becomes invariance under the backward chain, so that existence is free on finite graphs, and one constant — the norm of that chain's Green operator, which plays the role of an inverse spectral gap — sets a floor under both the curvature of the loss around the balanced flow and the rate at which training converges near it. The mechanism is that gradient descent diffuses the flow along the backward policy. For the squared-logarithm generator of detailed and trajectory balance, training the balance loss on states converges globally on every finite path-connected graph, from every positive initialization. The constant can be infinite while backward trajectories are short on average, and exact flow matching can then fail. The bounds and rates are tested by exact computation on enumerable state spaces, and every theorem carries a certification status computed from a Lean 4 development.

1 INTRODUCTION

1.1 GENERATIVE FLOWS AND THE COST OF MIXING

A generative flow or GFlowNet (Bengio et al., 2021; 2023) was originally defined on a directed acyclic graph as a positive weighting of its edges, called a flow. It defines a sampling process for an unnormalized target distribution on the vertices: it starts from some initial distribution on the vertices and, at each vertex, chooses between following an edge and stopping (which we see as following an edge toward a sink state, weighted by the reward), with probabilities proportional to the flow. The fundamental result of GFlowNets states that when the flow is balanced — at every state, what arrives or starts equals what leaves or stops — the stopped state is an exact sample of the target (Bengio et al., 2021; 2023). Training enforces balance through a loss: flow matching, detailed balance, trajectory balance and their variants (Bengio et al., 2021; 2023; Whitammer et al., 2022; Pan et al., 2023; da Silva et al., 2024; Hu et al., 2025). Built on directed acyclic graphs, the framework now covers continuous state spaces (Li et al., 2023; Lahlou et al., 2023), state spaces with cycles (Brunswic et al., 2024; 2025a; Morozov et al., 2025) and multiple agents (Brunswic et al., 2025b), and the models it fine-tunes are moving the same way. Generative flows fine-tune autoregressive models, whose generation explores a tree (Yu et al., 2025; Takase et al., 2024; Ho et al., 2024; Bartoldson et al., 2025; Wang et al., 2025); diffusion language models unmask tokens in any order and some revise them (Sahoo et al., 2024; Nie et al., 2025; von Rütte et al., 2025), so a state is reached along many paths and, where tokens are revised, along cycles.

In their seminal work, Bengio et al. (2021) explain that GFlowNets are expected to escape the cost of mixing that burdens Markov chain Monte Carlo (Levin & Peres, 2017) **during inference**. A common informal belief is that the choice of the initial distribution, and of the frozen backward

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

policy used in the detailed-balance and trajectory-balance losses, matters for the convergence of training. Whoever trains a generative flow watches a loss and needs three things from it: that a small value certifies an accurate sampler, that the value can be driven to zero, and that gradient descent drives it there in reasonable time. Answers so far are partial: da Silva et al. (2025b;a) give guarantees on finite acyclic graphs, Brunswic et al. (2025a) bound the total-variation sampling error of any flow by its flow-matching defect and prove L^1-universality of ergodic flows, and Morozov et al. (2025) show that freezing the backward policy makes the detailed-balance loss stable in a weaker sense. These last two works converge on the same conclusion: frozen backward policies control universality and stability. The present work combines these ideas with the Markov-chain viewpoint of Deleu & Bengio (2023) into an encompassing framework that formalizes the intuition that frozen backward policies matter for training convergence; we provide convergence bounds controlling the sampling error by losses and describe the training dynamics in terms of a frozen backward policy. In fact, all convergence bounds or universality results proved in the present work depend on mixing coefficients, train/invariant distribution ratio in crucial ways: **GFlowNets do not escape the cost of mixing during training**.

Contributions.

1. *Certificates and their limit* (Section 3). We bound the sampling error by the stable and KL-weak-FM losses, with constants that do not involve the policy. We show that ratio-based flow matching with a generator continuous at 1 admits no such bound, already on one cycle.
2. *The frozen chain* (Section 2). We freeze the backward policy and wrap the graph around, so that balance is the invariance of one Markov chain. The norm $\widehat{B}$ of that chain's Green operator is the constant of the paper.
3. *Existence* (Section 4). We prove that on a finite path-connected graph a positive backward policy admits, for each target positive on the stopping states, a balanced flow sampling it exactly, with no mixing hypothesis.
4. *Dynamics* (Section 5 and Appendix C). We show that gradient descent diffuses the flow along the frozen backward policy, and derive from this the local rate. We show that freezing is weaker than stability.
5. *The boundary* (Appendix A). We give a counterexample, the doubling graph, on which the choice of the frozen backward policy alone decides whether a balanced flow exists and whether the bounds constants are finite.
6. *A certified paper* (Section 1.3). The paper is fully formalized in Lean 4. Every theorem statement and badge is generated from the Lean certificate, and the whole development is unrolled in an appendix that opens on its proof-structure graph.

With a fixed backward policy on a directed acyclic graph, learning a generative flow is entropy-regularized reinforcement learning (Tiapkin et al., 2024), and entropy-regularized policy gradient converges globally at linear rates in tabular Markov decision processes (Mei et al., 2020; Cen et al., 2022); the analysis here works in the space of flows, allows cycles and names the constant that sets the rate. Gritsaev et al. (2025) train the backward policy this paper freezes.

1.2 Generative flows: framework and notation

We work in the measurable theory of generative flows of Brunswic et al. (2024), which admits cycles, in the reward-free form of Brunswic et al. (2025a) and Brunswic et al. (2025b). One set of definitions then covers graphs, where the background measure is the counting measure, and continuous state spaces, where it is the Lebesgue or Riemannian volume; it is also the setting of the Lean development, whose sampler results hold on an arbitrary measurable space.

Measures and kernels. The state space $\mathcal{S}$ is a measurable space with a finite background measure ν_{B}. Flows are finite measures on $\mathcal{S}$, and a lower-case letter denotes the density of a flow against ν_{B}. A Markov kernel π moves a measure μ to $\mu\pi$, and $\mu \otimes \pi$ is the measure on pairs (x, y) with x drawn from μ and y from $\pi(x)$. For an unnormalized μ, $x \sim \mu$ says that the law of x is $\mu/\mu(\mathcal{S})$, and $\mathrm{TV}(x \,\|\, \mu)$ is the total-variation distance from that law to $\mu/\mu(\mathcal{S})$, at most one. The target κ is a finite measure on $\mathcal{S}$, known through its density in reinforcement learning and through samples

in imitation learning (Brunswic et al., 2025a); the training measure ν_{T} is the measure under which losses are evaluated.

Flows, sampler, policies. A generative flow is a *star forward policy* $\pi^*_{\rightarrow}$, a Markov kernel on $\mathcal{S}$ that never stops, and a *star outflow* $F^*_{\rightarrow}$, the mass each state passes on. The action spaces of Brunswic et al. (2025b), bundled with the state an action leaves, are composed away: $\pi^*_{\rightarrow}$ is all the sampler sees. Given an initial flow F_{init} and a terminal flow F_{term}, the sampler draws $s_1 \sim F_{\mathrm{init}}$; at step t it stops with probability $dF_{\mathrm{term}}/d(F_{\mathrm{term}} + F^*_{\rightarrow})$ at s_t, and otherwise moves to $s_{t+1} \sim \pi^*_{\rightarrow}(s_t)$; its output is the stopped state s_τ. Adjoin a source s_0 and a sink s_f. The edge flow $F^*_{\rightarrow} \otimes \pi^*_{\rightarrow}$, extended by F_{init} out of the source and by F_{term} into the sink, has outflow $F_{\rightarrow} := F^*_{\rightarrow} + F_{\mathrm{term}}$ and inflow $F_{\leftarrow} := F_{\mathrm{init}} + F^*_{\leftarrow}$, where $F^*_{\leftarrow} := F^*_{\rightarrow}\pi^*_{\rightarrow}$ is the star inflow. Normalized by its origin, the edge flow gives the forward policy $\pi_{\rightarrow}$, which is the sampler's step, stopping included. Normalized by its destination, it gives the *backward policy* $\pi_{\leftarrow}$, which says where the mass arriving at a state came from: on a graph, $\pi_{\leftarrow}(s \to s')$ is the flow along $s' \to s$ divided by $F_{\leftarrow}(s)$, the source counting as a predecessor.

Balance and the sampling theorem. The flow is *balanced*, or flow-matching, when $F_{\leftarrow} = F_{\rightarrow}$ as measures, that is $F_{\mathrm{init}} + F^*_{\leftarrow} = F^*_{\rightarrow} + F_{\mathrm{term}}$: at every state, what arrives or starts equals what leaves or stops. The sampling theorem formalizes the intuition that balance makes the sampler exact. A balanced flow with $F_{\mathrm{term}} \neq 0$ has $F_{\mathrm{init}}(\mathcal{S}) = F_{\mathrm{term}}(\mathcal{S})$, its sampler stops after at most $1 + F^*_{\rightarrow}(\mathcal{S})/F_{\mathrm{init}}(\mathcal{S})$ steps in expectation, and $s_\tau \sim F_{\mathrm{term}}$ (Bengio et al., 2021; Brunswic et al., 2024); the statement is certified on an arbitrary measurable space (Theorem 1).

Theorem 1. *certified* [L1] *Let $\pi^*_{\rightarrow}$ be a Markov kernel and F_{init}, $F_{\mathrm{term}} \neq 0$, $F^*_{\rightarrow}$ finite measures on $\mathcal{S}$ with $F_{\mathrm{init}} + F^*_{\rightarrow}\pi^*_{\rightarrow} = F_{\mathrm{term}} + F^*_{\rightarrow}$. The sampler from F_{init} along $\pi^*_{\rightarrow}$, stopped at the rate F_{term} prescribes, stops almost surely, and*

$$F_{\mathrm{init}}(\mathcal{S}) = F_{\mathrm{term}}(\mathcal{S}), \qquad \mathbb{E}(\tau) \leq 1 + \frac{F^*_{\rightarrow}(\mathcal{S})}{F_{\mathrm{init}}(\mathcal{S})}, \qquad s_\tau \sim \frac{F_{\mathrm{term}}}{F_{\mathrm{term}}(\mathcal{S})}$$

(Bengio et al., 2021; Brunswic et al., 2024).

A 0-*flow* is a flow balanced with $F_{\mathrm{init}} = F_{\mathrm{term}} = 0$, its star outflow invariant under its star policy: the measurable form of a circulation around a cycle (Brunswic et al., 2024). Added to a balanced flow, a 0-flow leaves it balanced between the same ends. The sampler therefore still draws F_{term}, while the bound on $\mathbb{E}(\tau)$ grows with the circulation.

Forcing balance. A flow that is not balanced is balanced once its ends are corrected (Brunswic et al., 2025a). With $E := F_{\mathrm{init}} + F^*_{\leftarrow} - F^*_{\rightarrow}$, let $\delta F_{\mathrm{init}} := (E - F_{\mathrm{term}})^-$ and $\widehat{F}_{\mathrm{term}} := F_{\mathrm{term}} + (E - F_{\mathrm{term}})^+$; the flow is balanced from $F_{\mathrm{init}} + \delta F_{\mathrm{init}}$ to $\widehat{F}_{\mathrm{term}}$. The correction at the start is not tractable and the one at the end is, so the sampler is run from F_{init} and stopped by $\widehat{F}_{\mathrm{term}}$. It stops almost surely, and its output lies within total variation $\delta F_{\mathrm{init}}(\mathcal{S})/\widehat{F}_{\mathrm{term}}(\mathcal{S})$ of $\widehat{F}_{\mathrm{term}}$ whenever $F_{\mathrm{init}} \neq 0$ (Theorem 2).

Theorem 2. *certified* [L2] *Let $(\pi^*_{\rightarrow}, F^*_{\rightarrow})$ be a generative flow and $F_{\mathrm{init}} \neq 0$, F_{term} finite measures on $\mathcal{S}$, not necessarily balanced; let δF_{term} and δF_{init} be the positive and negative parts of $F_{\mathrm{init}} + F^*_{\rightarrow}\pi^*_{\rightarrow} - F^*_{\rightarrow} - F_{\mathrm{term}}$, and $\widehat{F}_{\mathrm{term}} := F_{\mathrm{term}} + \delta F_{\mathrm{term}}$. The sampler from F_{init} stopped by $\widehat{F}_{\mathrm{term}}$ stops almost surely, and*

$$\widehat{F}_{\mathrm{term}}(\mathcal{S}) = F_{\mathrm{init}}(\mathcal{S}) + \delta F_{\mathrm{init}}(\mathcal{S}), \qquad \mathrm{TV}\Big(s_\tau \,\Big\|\, \frac{\widehat{F}_{\mathrm{term}}}{\widehat{F}_{\mathrm{term}}(\mathcal{S})}\Big) \leq \frac{\delta F_{\mathrm{init}}(\mathcal{S})}{\widehat{F}_{\mathrm{term}}(\mathcal{S})}$$

(Brunswic et al., 2025a).

This theorem guards a caveat: the flow a practitioner trains is never exactly balanced, and it bounds what that costs. At inference the target is not queried: $F_{\mathrm{term}} = 0$, the correction is $\delta F_{\mathrm{init}} = E^-$, of density δf_{init}, and the sampler stops by the *implied terminal flow* $\widehat{\kappa} := E^+$, of density $\widehat{f}_{\mathrm{term}}$. In the terms of Brunswic et al. (2025b), E is the virtual reward and a flow with $E \geq 0$ is weakly flow-matching; its forced sampler is exact.

Losses and their generators. Training moves $\pi^*_{\rightarrow}$ and $F^*_{\rightarrow}$, and sometimes F_{init}, to drive a loss to zero. A balance loss integrates, under a training measure, a function of a comparison between the two sides of the balance, and that function is its *generator*. *Difference losses* compare the two sides by their difference. The stable loss of Brunswic et al. (2024) and the KL-weak-FM loss of Brunswic et al. (2025a) are

$$\mathcal{L}^{\mathrm{stable}}_{\mathrm{FM},q} := \left\| f_{\mathrm{init}} + f^*_{\leftarrow} - f^*_{\rightarrow} - f_{\mathrm{term}} \right\|_{L^q(\nu_{\mathrm{T}})} \qquad \text{with } F_{\mathrm{term}} = \kappa, \tag{1}$$

$$\mathcal{L}^{q}_{\mathrm{KL-wFM}} := b \left\| \delta f_{\mathrm{init}} \right\|_{L^q(\nu_{\mathrm{T}})} - \mathbb{E}_{s\sim\kappa} \log \widehat{f}_{\mathrm{term}}(s) \qquad \text{with } F_{\mathrm{term}} = 0. \tag{2}$$

The stable loss needs the target's density; its q-th power is generated by $|x|^q$, one of the functions of the difference with which Brunswic et al. (2024) generate their stable losses. The KL-weak-FM loss, of weight b, needs only samples of a normalized target: its first term is generated by the negative part of the difference alone, which is weak flow matching, and its cross-entropy term, taken under κ rather than ν_{T}, fits the implied terminal flow to κ. *Ratio losses* compare the two sides by their ratio, through a generator $g : (0,\infty) \to [0,\infty)$ that vanishes only at 1 (Bengio et al., 2021; 2023; Whitammer et al., 2022; Brunswic et al., 2024). Flow matching (FM) compares them on states, detailed balance (DB) on edges and trajectory balance (TB) along trajectories:

$$\mathcal{L}_{\mathrm{FM},g} := \int_{\mathcal{S}} g\Big(\frac{dF_{\leftarrow}}{dF_{\rightarrow}}\Big)\, d\nu_{\mathrm{T}} \quad \text{with } F_{\mathrm{term}} = \kappa, \tag{3}$$

$$\mathcal{L}_{\mathrm{DB},g} := \int g\big(r(x,y)\big)\, d\nu^{\mathrm{edge}}_{\mathrm{T}}(x,y), \quad r(x,y) := \frac{F_{\rightarrow}(dy)\,\pi_{\leftarrow}(y \to dx)}{F_{\rightarrow}(dx)\,\pi_{\rightarrow}(x \to dy)}, \tag{4}$$

$$\mathcal{L}_{\mathrm{TB},g} := \int g\Big(\prod_{k=1}^{n} r(x_{k-1},x_k)\Big)\, d\nu^{\mathrm{traj}}_{\mathrm{T}}(x_0,\dots,x_n). \tag{5}$$

The DB ratio r divides the flow the backward policy sends back along an edge by the flow the forward policy sends along it, $F_{\rightarrow}$ being read at the source and at the sink as the masses of F_{init} and of F_{term}. In practice DB and TB are generated by $(\log x)^2$. The results below use four properties of a generator: continuity at 1; curvature $g''(1) > 0$ at balance; *strict unimodality*, meaning that g' has the sign of $x - 1$, as for $(\log x)^2$ and $(x-1)^2$; and, at the other extreme, a derivative that vanishes on bands around 1.

Stability. Brunswic et al. (2024) call a loss *stable* when adding a 0-flow never lowers it, and the word is used here in that sense only. A 0-flow adds the same measure to both sides of the balance. A difference does not see it; a ratio is pulled towards one. The difference losses (1) and (2) are therefore unchanged by a 0-flow, while Brunswic et al. (2024) show that the ratio losses (3)–(5) are unstable: a circulation lowers them without changing what the sampler draws.

The frozen backward policy. On a graph, the backward policy can be held fixed while the flow is trained (Morozov et al., 2025). The flow is then determined by the mass passing through each state. Close the graph into a loop, sink back to source (Deleu & Bengio, 2023). The frozen backward policy then defines a Markov chain, the backward chain, with an invariant measure λ, and P moves densities one step along it on $L^2(\lambda)$. With Π the projection onto constants, the *Green operator* is $S := (I - P + \Pi)^{-1} - \Pi$, and $\widehat{B} := \|S\|$ is the constant of the paper: the best $\widehat{B}$ with $\|h - \Pi h\| \le \widehat{B}\,\|(I-P)h\|$. The mixing sum $B_p := \sum_n \|P^n - \Pi\|$ on L^p appears only where universality on a general state space needs it.

1.3 Overview

Figure 1 draws the argument; an arrow follows the order in which the story reaches a result. The sampling theorem answers the first question of Section 1.1: forcing balance measures any flow against it and reduces certification to one norm of the balance defect, which difference losses control and ratio losses cannot.

Freezing the backward policy answers the other two, once the graph is made a chain. The framework of Section 1.2 admits cycles, and the results on the frozen chain are stated for a Markov chain with an invariant measure, the ergodic presentation of generative flows (Brunswic et al., 2025a) that this paper adopts. The acyclic graphs of practice reduce to it by the loop trick of Deleu & Bengio (2023):

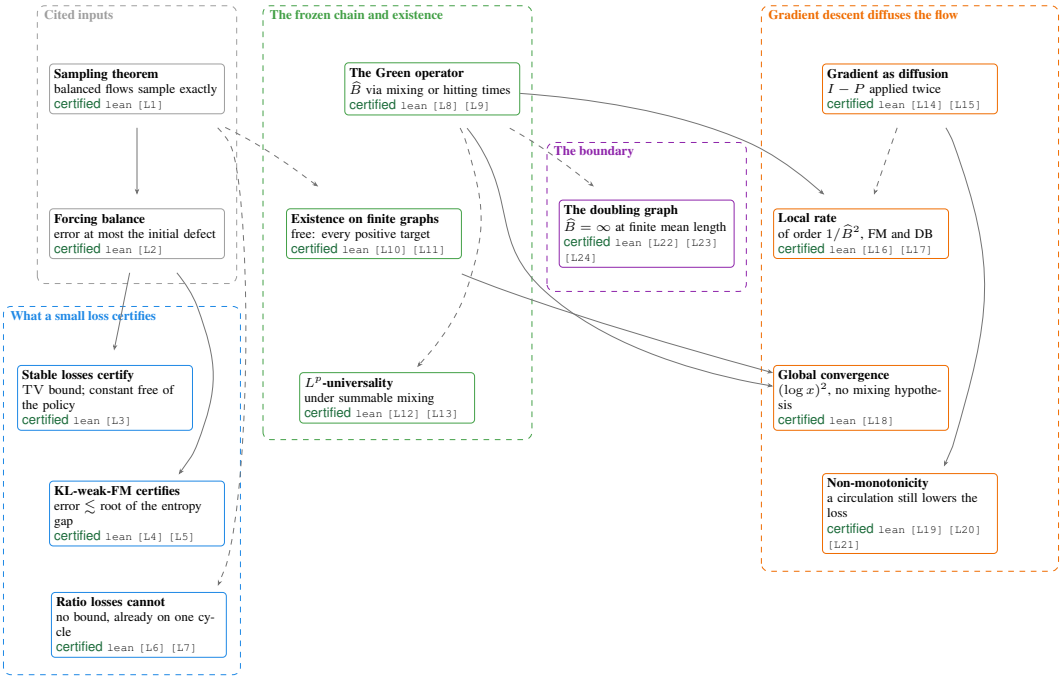


Figure 1: The results walked through in Section 1.3. Each badge is computed from the dependency closure of the Lean declarations it links to, never typed. An arrow follows the argument; it is solid when the Lean proof of its target runs through a module that proves its source, and dashed when the link is in the argument only. The map is trimmed to the results that carry the story; the full proof structure is drawn in the appendix.

wrap the graph around by an edge from the sink back to the source. On a finite path-connected graph the sampler's trajectories, read backward, become the excursions of one irreducible chain, the balanced flow becomes its invariant measure, and cutting the wrap returns the acyclic flow, so the reduction loses nothing; what it adds, a chain that is never reversible and may be periodic, is why the constant is the norm $\widehat{B}$ of the chain's Green operator and not a spectral gap (Section 2). Balance is then the invariance of that chain, existence is free on finite graphs, and $\widehat{B}$ is the cost of mixing that training pays: gradient descent, a diffusion along the backward policy, pays it at the local rate and, through hitting times, at the global one. The doubling graph closes the story where $\widehat{B}$ is infinite.

Sampling theorem: a balanced flow samples its terminal flow exactly. *Forcing balance*: a flow that is not balanced, stopped by its implied terminal flow, samples with an error set by its initial defect; both certificates start there. See Section 1.2).

What a small loss certifies (Section 3). *Stable losses certify* the forced sampler, with a constant free of the policy. *KL-weak-FM certifies* it from samples of the target, to leading order by the square root of the loss's excess over the target's self-entropy. *Ratio losses cannot*: flow matching with a generator continuous at 1 is blind to circulations, and already on one cycle, for every target but one, no function of the loss that vanishes with it bounds the sampling error.

The frozen chain and existence. *The Green operator* (Section 2): its norm $\widehat{B}$ is the constant of the paper, bounded through mixing or through the hitting times of backward trajectories. *Existence on finite graphs* (Section 4): on a finite path-connected graph, for each target positive on the stopping states, a backward policy positive on the edges has a balanced flow that samples it exactly. *L^p-universality* (Section 4): on a general state space a fixed forward policy with summable mixing has flows universal in L^p, exactly so in L^∞ — the one place where the mixing sum, and not $\widehat{B}$, is the price.

Gradient descent diffuses the flow (Section 5). *Gradient as diffusion*: near balance the gradient of a balance loss applies $I - P$ twice, once through its adjoint. *Local rate*: training on the flow-matching

and detailed-balance losses therefore converges near balance at a rate of order $1/\widehat{B}^2$, linearized on any state space and in full on a finite one. *Global convergence*: the $(\log x)^2$ balance loss on states converges from every positive initialization on a finite path-connected graph, at a rate read off backward trajectories and with no mixing hypothesis. *Non-monotonicity*: freezing buys curvature, not stability; a circulation can still lower these losses, and for strictly unimodal generators gaining circulation is how training reaches balance.

The boundary (Appendix A). *The doubling graph*: its backward trajectories have finite mean length, the quantity through which Morozov et al. (2025) control training, yet $\widehat{B}$ is infinite in every L^p and exact flow matching with square-integrable outflows fails for some pair; a suitable finite truncation restores it, at a constant growing at least like the square root of the truncation's length.

2 THE FROZEN CHAIN AND ITS CONSTANT

Freezing the backward policy turns training into a question about one Markov chain. On the loop closure the backward chain is irreducible, its invariant measure λ is unique up to scale and positive everywhere, and a flow is balanced exactly when its state mass is a multiple of λ: the flow-matching loss measures departure from invariance through $I - P$, and the detailed-balance loss measures the same departure for the chain lifted to edges. The target enters as the chain's step out of the sink, and λ has a trajectory reading: it is the visit ratio, the expected number of visits to a state by one backward trajectory drawn from the target row, divided by two plus that trajectory's expected length (Theorem 20).

The constant that prices the rest is the norm $\widehat{B}$ of the Green operator, and three readings of it are certified. Where the chain mixes summably, the Green operator is the sum of the centred powers of P, and $\widehat{B}$ is at most the mixing sum:

Theorem 3. certified [L8] *Let T be a Markov kernel on a standard Borel space with $\lambda T = \lambda$, $\lambda \neq 0$ finite, P its density action on $L^2(\lambda)$, Π the λ-mean projection, and $B_2 := \sum_{n\geq 0} \|P^n - \Pi\| < \infty$. Then $S := \sum_{n\geq 0}(P^n - \Pi)$ converges in operator norm, and*

$$S(I-P) = I - \Pi, \qquad \|h - \Pi h\| \leq B_2 \,\|(I-P)h\| \quad (h \in L^2(\lambda)).$$

If λ is not a multiple of a Dirac mass, $\|I - \Pi\| = 1$ and $B_2 \geq 1$.

For a reversible chain $\widehat{B}$ is exactly the reciprocal spectral gap (Levin & Peres, 2017). No loop closure of a marked graph is reversible, since the chain leaves the sink only through the wrap while it enters it from every terminating state, so for a backward policy this is a reading and not an identity. On a finite path-connected graph, whether or not the chain mixes, hitting times bound the constant: with σ_* the longest expected time for a backward trajectory to reach the source, $\bar{\sigma}$ the mean length of one drawn from the target row and $N_{\min}$ the fewest expected visits to a state, $\widehat{B} \leq \sigma_* \sqrt{(2+\bar{\sigma})/N_{\min}}$ (Theorem 21), with no aperiodicity hypothesis. The two constants differ in general, and the difference matters where the state space is most common: on a leveled graph, the state space of an autoregressive model, the loop closure is periodic, its mixing coefficients do not decay and the mixing sum is infinite (Theorem 35), while $\widehat{B}$ stays finite. This is why the constant of the paper is the norm of the Green operator and not the mixing sum, and why the rates of Section 5 hold on autoregressive state spaces.

3 WHAT A SMALL LOSS CERTIFIES

Section 1.2 reduced certification to one norm: the forced sampler is within $\delta F_{\mathrm{init}}(\mathcal{S})/\widehat{\kappa}(\mathcal{S})$ of $\widehat{\kappa}$, and both that defect and the distance from $\widehat{\kappa}$ to κ are integrals of $E - \kappa$. A loss certifies when it controls the $L^1(\nu_{\mathrm{B}})$ norm of that defect. Difference losses do.

Theorem 4. certified [L3] *Let $q \in [1,\infty]$, ν_{B} finite with $\nu_{\mathrm{B}} \leq M\nu_{\mathrm{T}}$, $(\pi^*_{\rightarrow}, f^*_{\rightarrow})$ a generative flow with star inflow $\ll \nu_{\mathrm{B}}$, $f_{\mathrm{init}}, k \geq 0$ integrable with positive masses and $f_{\mathrm{init}} + f^*_{\leftarrow} - f^*_{\rightarrow} - k \in L^q(\nu_{\mathrm{T}})$, $\kappa = k\,\nu_{\mathrm{B}}$, and $C_0 := \nu_{\mathrm{B}}(\mathcal{S})^{1-1/q} M^{1/q}$. The sampler from F_{init} stopped by $\widehat{\kappa} = E^+$ stops almost surely with a law $\ll \nu_{\mathrm{B}}$, and*

$$\mathrm{TV}\Big(s_\tau \,\Big\|\, \frac{\kappa}{\kappa(\mathcal{S})}\Big) \leq \frac{2C_0}{\kappa(\mathcal{S})}\,\mathcal{L}^{\mathrm{stable}}_{\mathrm{FM},q}, \qquad \big|F_{\mathrm{init}}(\mathcal{S}) - \kappa(\mathcal{S})\big| \leq C_0\,\mathcal{L}^{\mathrm{stable}}_{\mathrm{FM},q}.$$

When the target is known through samples, the KL-weak-FM loss certifies as well, and the mechanism is more delicate. Its cross-entropy term is taken against the implied terminal flow, which is not normalized: a mass defect lowers the cross-entropy through the logarithm of $\widehat{\kappa}(\mathcal{S})$, and the penalty $b\|\delta f_{\mathrm{init}}\|_{L^q(\nu_{\mathrm{T}})}$ must charge for it. Hölder's inequality bounds the defect by $\|\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}}\|_{L^{q^*}(\nu_{\mathrm{T}})}$ times the penalty's norm, so the penalty wins once b exceeds one plus that norm, and the loss is then floored by the target's self-entropy.

Theorem 5. *certified* [L4] *Let $q \in [1,\infty]$ with conjugate q^*, $\nu_{\mathrm{B}} \ll \nu_{\mathrm{T}}$ σ-finite with $\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}} \in L^{q^*}(\nu_{\mathrm{T}})$, $b \geq 1 + \|\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}}\|_{L^{q^*}(\nu_{\mathrm{T}})}$, $F_{\mathrm{init}} = f_{\mathrm{init}}\,\nu_{\mathrm{B}}$ and $\kappa = k\,\nu_{\mathrm{B}}$ probability measures with $\mathcal{H} := -\int k \log k \, d\nu_{\mathrm{B}}$ finite, and $(\pi^*_{\rightarrow}, f^*_{\rightarrow})$ a generative flow with star inflow $\ll \nu_{\mathrm{B}}$. The sampler from F_{init} stopped by $\widehat{\kappa} = E^+$ stops almost surely with a law $\ll \nu_{\mathrm{B}}$, $\mathcal{L}^q_{\mathrm{KL-wFM}} \geq \mathcal{H}$, and if $\mathcal{L}^q_{\mathrm{KL-wFM}} < \infty$,*

$$\mathrm{TV}(s_\tau \,\|\, \kappa) \leq \tfrac{1}{\sqrt{2}}\sqrt{\mathcal{L}^q_{\mathrm{KL-wFM}} - \mathcal{H}} + \min\bigl(1, \|\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}}\|_{L^{q^*}(\nu_{\mathrm{T}})}(\mathcal{L}^q_{\mathrm{KL-wFM}} - \mathcal{H})\bigr).$$

As the loss approaches the floor the linear term is dominated and the error is of order the square root of the excess (Theorem 15). The floor is attained at exact matching, which strong universality supplies; reaching it may require a large 0-flow component, since some regions of the state space take many moves to reach.

Ratio losses do not certify. A 0-flow pulls every ratio towards one without changing what the sampler draws (Section 1.2), and on a cycle this is not a loss of tightness but the absence of any bound. For every generator continuous at 1, every training measure and every target but one, no function of the flow-matching loss that vanishes with it bounds the sampling error on a single cycle (Theorem 24); and over the cycles of each size, at every loss level, the worst error is the largest possible:

Theorem 6. *certified* [L6] *Let g be continuous at 1 with $g(1) = 0$, $\varepsilon > 0$, and $\mathcal{C}_N$ the cycle graph on $N \geq 2$ internal states. Then*

$$\inf_{\nu_{\mathrm{T}}} \sup\bigl\{\mathrm{TV}(s_\tau \,\|\, \kappa) \;:\; \mathcal{L}_{g,\nu_{\mathrm{T}}}(F) \leq \varepsilon\bigr\} = 1,$$

the supremum over probability targets κ on the internal states and non-negative edge-carried flows F with a sampler, the loss taken against κ.

The statement is for the flow-matching loss on states. The same circulation lowers the detailed- and trajectory-balance losses (Brunswic et al., 2024); whether those can certify is not settled here, since a circulation does not move the terminal ratios that detailed balance also sees. On finite directed acyclic graphs the picture is different: model-free bounds exist there (da Silva et al., 2025b), but none is free of the state space — on a star graph with uniform backward policy and uniform training the residual is arbitrarily small while the sampling error stays bounded away from zero — so their constants grow with the graph (Section B.4).

4 What the backward policy can reach

The stable bound is empty unless the loss can be made small, so the question is whether the balance equation of a fixed policy has a solution. On a finite graph, with the backward policy frozen, it always does: the balanced flow is the invariant measure of the loop closure, and cutting the wrap edge and reading the sink row as the target gives a generative flow with exact balance at every internal state (Theorem 22). Its sampler draws the target exactly:

Theorem 7. *certified* [L10] *In the setting of Theorem 22, let G have an internal state. The sampler of e stops almost surely, with*

$$\mathbb{E}(\tau) \leq 1 + \frac{F^*_{\rightarrow}(V)}{F_{\mathrm{init}}(V)}, \qquad s_\tau \sim \frac{\pi_{\leftarrow}(s_f \to \cdot)}{1 - \pi_{\leftarrow}(s_f \to s_0)} \text{ on the internal states.}$$

Nothing beyond positivity of the policy on the edges is asked: no aperiodicity and no mixing. A target vanishing on some terminating states is handled on the subgraph they span, and a target charging an edge from the source straight to the sink loses that mass to the normalization. Existence is free on graphs; what is priced is speed.

On a general state space, as for ergodic generative flows (Brunswic et al., 2025a), one fixes the forward policy instead and leaves the outflow free. Universality then does need mixing:

Theorem 8. *certified [L12] Let $p \in [1, \infty)$ and $\pi^*_{\to}$ a Markov kernel on $\mathcal{S}$ with $\nu_{\mathrm{B}}\pi^*_{\to} = \nu_{\mathrm{B}}$, ν_{B} finite, whose density action $P_\star$ is bounded on $L^p(\nu_{\mathrm{B}})$ with*

$$\sum_{n\geq 0} \|P^n_\star - \Pi\|_{L^p(\nu_{\mathrm{B}})} < \infty.$$

The generative flows with forward policy $\pi^*_{\to}$ and free outflow $f^*_{\to} \geq 0$ are weakly L^p-universal: any two densities in $L^p(\nu_{\mathrm{B}})$ of equal mass are joined by flows of arbitrarily small L^p residual.

At $p = \infty$ the join is exact, with an outflow the mixing sum controls (Theorem 18). The condition is sufficient and not necessary; it is the one place where the mixing sum rather than the Green operator enters; and the boundedness of the density action is a requirement on the parameterization that the definitions do not supply. This is the L^p strengthening of the L^1-universality of ergodic flows, stated for the forward policy; on graphs the backward policy is frozen and the strengthening is not needed. Existence is not speed: Section 5 prices the solution, and Appendix A shows a graph on which the mean backward length is finite, exact flow matching with square-integrable outflows fails for some pair, and the price is infinite.

5 Gradient descent diffuses the flow along the backward policy

The gradient of a ratio loss has a closed form — the reversal of the kernel minus the balance ratio, applied to g' of the ratio (Appendix C) — and near balance it linearizes:

Theorem 9. *certified [L14] Let T be a Markov kernel on a standard Borel space with $\lambda T = \lambda$, $\lambda \neq 0$ finite, P its density action on $L^2(\lambda)$, $A = P - I$, $\nu = w\lambda$ with $0 \leq w \in L^\infty(\lambda)$, M_w the multiplication by w, and $g \in C^3[1-a, 1+a]$, $0 < a < 1$, with $g(1) = g'(1) = 0 < g''(1)$. There is an explicit K such that, for $0 < \varepsilon \leq a/4$ and $|h| \leq \varepsilon$, the flow $(1+h)\lambda$ has ratio $(1+Ph)/(1+h)$, and the gradient of $\mathcal{L}_{g,\nu}$ at it has an $L^2(\lambda)$ density $D(h)$ with*

$$\|D(h) - Hh\| \leq K\varepsilon \, \|Ah\|, \qquad H := g''(1)\, A^\dagger M_w A;$$

H is the only homogeneous H' with $\|D(h) - H'h\| = o(\|h\|)$ as $\|h\|_\infty \to 0$; and if T is reversible and $w \equiv 1$, then $H = g''(1)(I-P)^2$, and $P\varphi = \beta\varphi$ implies $H\varphi = g''(1)(1-\beta)^2\varphi$.

Training diffuses the flow along the frozen backward policy, once through P and once through its adjoint. Detailed balance is the same loss on the chain lifted to edges; trajectory balance is not. Reading this as the square of a heat flow needs a reversible kernel, which a loop closure never is.

The curvature transverse to the balanced flow is therefore at least $g''(1)w_{\min}/\widehat{B}^2$: linearized training contracts at that rate on any state space, and on a finite one the full dynamics inherits it from an explicit neighbourhood:

Theorem 10. *certified [L16] Let T be a Markov kernel on a finite state space with invariant probability $\lambda > 0$, P its density action, $\nu = w\lambda$ with $w \geq w_{\min} > 0$, $g \in C^3[1-a, 1+a]$ with $g'(1) = 0 < g''(1)$, $\widehat{B} \geq 1$ with $\|h - \Pi h\| \leq \widehat{B}\,\|(I-P)h\|$ on $L^2(\lambda)$, and $\varrho := g''(1)\, w_{\min}/\widehat{B}^2$. There are explicit $\varepsilon_0, \gamma_0 > 0$, functions of g, a, $w_{\min}$, $\|w\|_\infty$, $\widehat{B}$ and $\min\lambda$, such that for $\|h_0\|_{L^2(\lambda)} \leq \varepsilon_0$: (a) the gradient flow $(1+h_t)\lambda$ of $\mathcal{L}_{g,\nu}$ from $(1+h_0)\lambda$ exists, is unique, stays positive, and for some balanced $c_\infty\lambda$*

$$\|1 + h_t - c_\infty\|_{L^2(\lambda)} \leq 2e^{-\varrho t/2}\, \|h_0 - \Pi h_0\|_{L^2(\lambda)};$$

(b) for $0 < \gamma \leq \gamma_0$ the gradient descent h_k from h_0 is well defined and

$$\|h_{k+1} - \Pi h_{k+1}\| \leq (1 - \gamma\varrho/4)\, \|h_k - \Pi h_k\|.$$

Detailed balance obeys the same with $1 + \widehat{B}$ in place of $\widehat{B}$; for trajectory balance the development gives a step cap and no rate (Appendix C).

Far from balance, for the generator $(\log x)^2$, the balance loss on states converges on every finite path-connected graph from every positive initialization, in a Łojasiewicz phase and then an exponential

one at the local rate with the hitting-time constant of Section 2 in place of $\widehat{B}$ — read off the backward trajectories the practitioner already samples, with no mixing hypothesis (Theorem 62, Appendix C). That rate is a guarantee, not a prediction.

Freezing buys curvature, not stability: a circulation still lowers the flow-matching and detailed-balance losses (Theorem 63). This is the mechanism rather than a defect: at a fixed training measure, training moves on a sphere of $L^2(\lambda)$ whose mass is largest at the balanced flow, so for a strictly unimodal generator gaining circulation is how it gets there, while a generator whose derivative vanishes on bands around 1 freezes flows away from balance (Appendix C).

6 Conclusion and limitations

The constant that prices training is the norm of the Green operator of the frozen backward chain, and the doubling graph is the reminder that it is not the mean backward length: the mean can be finite where the constant is infinite, and on a truncation the constant grows with the size while the mean does not. The quantity to estimate from the backward trajectories a practitioner already samples is $\widehat{B}$ — their exit times and visit counts bound it, and for a policy reversible before the loop closure a single eigenvalue reads it — and not the length of those trajectories.

Reproducibility statement

The certificates are reproduced from the Lean development alone, which is the supplementary material. Every result this paper marks certified is a theorem of that development, and `make check` in it recompiles every proof, failing on any unproved goal and on any declaration that depends on an axiom beyond `propext`, `Classical.choice` and `Quot.sound`. Appendix E names the Lean and Mathlib releases it is pinned to and says how to build it, Appendix F locates each cited declaration by file and line, and Appendix B reads the development back into English, starting from its proof-structure graph.

AI use statement

We used generative AI tools (Claude, through Claude Code). The mathematical ideas, the counterexamples included, are ours, with no AI involvement. The proofs were drafted with AI assistance, and the Lean formalization, statements and proofs alike, was written entirely by AI under our review. The experiments were designed by us and implemented and run by AI. The paper was drafted with AI assistance, grounded in the Lean formalization and the experiment results, and AI wrote the tools that build its statements, badges, numbers, figures and appendix.

The paper is built so that AI output is checked rather than trusted. Every theorem statement is generated from its Lean declaration rather than transcribed, and its badge is computed from that declaration's dependency closure: a statement marked certified is checked by the Lean kernel, with no axiom beyond Lean's standard classical ones. Every number is read from a recorded experiment run, and every reference was checked against arXiv, OpenAlex and the publisher's own record. We reviewed all AI-assisted work, and we take responsibility for the final content of this work, including text, claims and artifacts produced with the aid of generative AI.

References

Brian R. Bartoldson, Siddarth Venkatraman, James Diffenderfer, Moksh Jain, Tal Ben-Nun, Seanie Lee, Minsu Kim, Johan Obando-Ceron, Yoshua Bengio, and Bhavya Kailkhura. Trajectory balance with asynchrony: Decoupling exploration and learning for fast, scalable LLM post-training. In *Advances in Neural Information Processing Systems*, volume 38, 2025. arXiv:2503.18929.

Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. *Advances in Neural Information Processing Systems*, 34:27381–27394, 2021.

Yoshua Bengio, Salem Lahlou, Tristan Deleu, Edward J. Hu, Mo Tiwari, and Emmanuel Bengio. GFlowNet foundations. *Journal of Machine Learning Research*, 24(210):1–55, 2023.

Leo Maxime Brunswic, Yinchuan Li, Yushun Xu, Yijun Feng, Shangling Jui, and Lizhuang Ma. A theory of non-acyclic generative flow networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 11124–11131, 2024. doi: 10.1609/aaai.v38i10.28989.

Leo Maxime Brunswic, Mateo Clémente, Rui Heng Yang, Adam Sigal, Amir Rasouli, and Yinchuan Li. Ergodic generative flows. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 5649–5668, 2025a.

Leo Maxime Brunswic, Haozhi Wang, Shuang Luo, Jianye Hao, Amir Rasouli, and Yinchuan Li. A theory of multi-agent generative flow networks. *arXiv preprint arXiv:2509.20408*, 2025b. NeurIPS 2025 Workshop on Structured Probabilistic Inference & Generative Modeling.

Shicong Cen, Chen Cheng, Yuxin Chen, Yuting Wei, and Yuejie Chi. Fast global convergence of natural policy gradient methods with entropy regularization. *Operations Research*, 70(4):2563–2578, 2022. doi: 10.1287/opre.2021.2151.

Tiago da Silva, Eliezer de Souza da Silva, and Diego Mesquita. On divergence measures for training GFlowNets. In *Advances in Neural Information Processing Systems*, volume 37, pp. 75883–75913, 2024.

Tiago da Silva, Rodrigo Barreto Alves, Eliezer de Souza da Silva, Amauri H. Souza, Vikas Garg, Samuel Kaski, and Diego Mesquita. When do GFlowNets learn the right distribution? In *The Thirteenth International Conference on Learning Representations*, 2025a.

Tiago da Silva, Amauri H. Souza, Omar Rivasplata, Vikas Garg, Samuel Kaski, and Diego Mesquita. Generalization and distributed learning of GFlowNets. In *The Thirteenth International Conference on Learning Representations*, 2025b.

Tristan Deleu and Yoshua Bengio. Generative flow networks: a Markov chain perspective. *arXiv preprint arXiv:2307.01422*, 2023.

Timofei Gritsaev, Nikita Morozov, Sergey Samsonov, and Daniil Tiapkin. Optimizing backward policies in GFlowNets via trajectory likelihood maximization. In *The Thirteenth International Conference on Learning Representations*, 2025.

Matthew Ho, Vincent Zhu, Xiaoyin Chen, Moksh Jain, Nikolay Malkin, and Edwin Zhang. Proof flow: Preliminary study on generative flow network language model tuning for formal reasoning. *arXiv preprint arXiv:2410.13224*, 2024. NeurIPS 2024 Workshop on System-2 Reasoning at Scale.

Rui Hu, Yifan Zhang, Zhuoran Li, and Longbo Huang. Beyond squared error: Exploring loss design for enhanced training of generative flow networks. In *The Thirteenth International Conference on Learning Representations*, 2025.

Salem Lahlou, Tristan Deleu, Pablo Lemos, Dinghuai Zhang, Alexandra Volokhova, Alex Hernández-García, Léna Néhale Ezzine, Yoshua Bengio, and Nikolay Malkin. A theory of continuous generative flow networks. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 18269–18300, 2023.

David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, 2nd edition, 2017.

Yinchuan Li, Shuang Luo, Haozhi Wang, and Jianye Hao. CFlowNets: Continuous control with generative flow networks. In *The Eleventh International Conference on Learning Representations*, 2023.

Jincheng Mei, Chenjun Xiao, Csaba Szepesvári, and Dale Schuurmans. On the global convergence rates of softmax policy gradient methods. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6820–6829, 2020.

Nikita Morozov, Ian Maksimov, Daniil Tiapkin, and Sergey Samsonov. Revisiting non-acyclic GFlowNets in discrete environments. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 44887–44910, 2025.

Shen Nie, Fengqi Zhu, Zebin You, Xiaolu Zhang, Jingyang Ou, Jun Hu, Jun Zhou, Yankai Lin, Ji-Rong Wen, and Chongxuan Li. Large language diffusion models. *arXiv preprint arXiv:2502.09992*, 2025.

Ling Pan, Nikolay Malkin, Dinghuai Zhang, and Yoshua Bengio. Better training of GFlowNets with local credit and incomplete trajectories. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 26878–26890, 2023.

Subham Sahoo, Marianne Arriola, Yair Schiff, Aaron Gokaslan, Edgar Marroquin, Justin Chiu, Alexander Rush, and Volodymyr Kuleshov. Simple and effective masked diffusion language models. *Advances in Neural Information Processing Systems*, 37:130136–130184, 2024.

Ryoichi Takase, Masaya Tsunokake, Yuta Tsuchiya, and Shota Inuzuka. GFlowNet fine-tuning for diverse correct solutions in mathematical reasoning tasks. *arXiv preprint arXiv:2410.20147*, 2024.

Daniil Tiapkin, Nikita Morozov, Alexey Naumov, and Dmitry Vetrov. Generative flow networks as entropy-regularized RL. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 4213–4221, 2024.

Dimitri von Rütte, Janis Fluri, Yuhui Ding, Antonio Orvieto, Bernhard Schölkopf, and Thomas Hofmann. Generalized interpolating discrete diffusion. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 61810–61843, 2025.

Yejing Wang, Shengyu Zhou, Jinyu Lu, Qidong Liu, Xinhang Li, Wenlin Zhang, Feng Li, Pengjie Wang, Chuan Yu, Jian Xu, Bo Zheng, and Xiangyu Zhao. GFlowGR: Fine-tuning generative recommendation frameworks with generative flow networks. *arXiv preprint arXiv:2506.16114*, 2025.

Esmeralda S. Whitammer, Moksh Jain, Emmanuel Bengio, Chen Sun, and Yoshua Bengio. Trajectory balance: Improved credit assignment in GFlowNets. In *Advances in Neural Information Processing Systems*, volume 35, pp. 5955–5967, 2022.

Fangxu Yu, Lai Jiang, Haoqiang Kang, Shibo Hao, and Lianhui Qin. Flow of reasoning: Training LLMs for divergent reasoning with minimal examples. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 73115–73145, 2025.

A WHERE THE CONSTANT IS INFINITE: THE DOUBLING GRAPH

Freezing a backward policy on the graph of Figure 2 is the example that separates the constant of this paper from the quantity through which Morozov et al. (2025) control training, the expected length of a backward trajectory, and it separates them in two ways. A bad policy fails outright: the uniform backward policy, which is the member $\varepsilon \equiv 1/2$ of the family, and every member with $s < 1$ leave the backward chain transient — no invariant probability, hence no balanced flow, and nothing for the expected backward length to rest on (Theorems 39 and 40). The subtle case is $s = 1$ with $c < 1$, and it is not fully settled. There the chain is positive recurrent and the mean backward length is $\bar{\jmath}/(1-c)$ for a target row of mean $\bar{\jmath}$ — the condition of Morozov et al. (2025) holds — and yet the invariant measure decays only polynomially, at the Cramér root of the doubling probability, the mixing sum is infinite, the Green operator is unbounded in every L^p, and exact flow matching with square-integrable outflows fails for some pair (Theorems 46, 41, 50 and 55). Truncating the ladder past the target's support restores exact flow matching and a finite constant, but one that grows at least like the

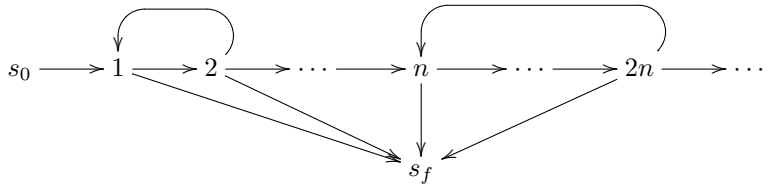


Figure 2: The doubling graph: a source s_0, a ladder of rungs and a sink s_f, with forward edges from the source to the first rung, from each rung to the next, from each even rung to its half, and from every rung to the sink. Its backward chain therefore steps down one rung or jumps to twice its rung, doubling with probability $\varepsilon(j)$ at rung j; the family studied is $\varepsilon_{c,s}(j) = c/(j+1)^s$, and geometric damping $\varepsilon(j) = a\rho^j$ is the contrast.

square root of the truncation's length while the mean length stays put (Theorem 58); whether weak universality holds on the infinite chain, and what happens at the boundary $c = 1/\ln 2$, is open. The obstruction belongs to the policy and not to the graph: a geometrically damped doubling probability keeps the Green operator bounded (Theorem 54). Appendix B.5 carries the analysis, statement by statement.

B THE LEAN DEVELOPMENT, UNROLLED

B.1 THE SAMPLER, THE CERTIFICATES AND THE CONSTANT

This appendix is an account of a Lean 4 development, checked against Mathlib `v4.31.0`. Every statement below is a theorem of that development; every proof sketch names the results that the corresponding proof term actually invokes, and nothing else; and each statement is followed by a note on how the Lean statement reads the paper's. The development contains no unproved goal, and each result stated here has been verified to depend on no axiom beyond Lean's own `propext`, `Classical.choice` and `Quot.sound`. Figure 3 is the whole of it; the chapters that follow unroll it, one namespace each, in the order of the main text.

This chapter. The general layer: a generative flow on an arbitrary measurable space, its sampler and the corrected sampler that forces balance; the two difference losses and what they bound; the Green operator of an invariant Markov kernel; universality for a fixed forward policy; and the first variation of a balance loss. Nothing here assumes a graph.

B.1.1 THE SAMPLER

Theorem 11. *Let $\pi^*_\to$ be a Markov kernel and F_{init}, $F_{\mathrm{term}} \neq 0$, $F^*_\to$ finite measures on $\mathcal{S}$ with $F_{\mathrm{init}} + F^*_\to \pi^*_\to = F_{\mathrm{term}} + F^*_\to$. The sampler from F_{init} along $\pi^*_\to$, stopped at the rate F_{term} prescribes, stops almost surely, and*

$$F_{\mathrm{init}}(\mathcal{S}) = F_{\mathrm{term}}(\mathcal{S}), \qquad \mathbb{E}(\tau) \leq 1 + \frac{F^*_\to(\mathcal{S})}{F_{\mathrm{init}}(\mathcal{S})}, \qquad s_\tau \sim \frac{F_{\mathrm{term}}}{F_{\mathrm{term}}(\mathcal{S})}$$

(Bengio et al., 2021; Brunswic et al., 2024).

Proof. Balance makes the total masses agree `SamplingGeneral.MFlow.mass_eq`. Summing the mass not yet stopped at each time against the outflow gives the bound on the expected length `SamplingGeneral.MFlow.expectedTau_le`, which forces that mass to vanish `SamplingGeneral.MFlow.tailProb_tendsto_zero`; the stopped law, the setwise limit of the laws stopped by each time, is then F_{term} normalized `SamplingGeneral.MFlow.sampleLaw_eq`. □

The sampling theorem of the main text as it is stated there, imported and not claimed: on an arbitrary measurable space rather than a Polish one, and with $F^*_\to$ any finite measure rather than a density against ν_{B}, both weakenings of the hypotheses. The expected length is a sum of tail probabilities in

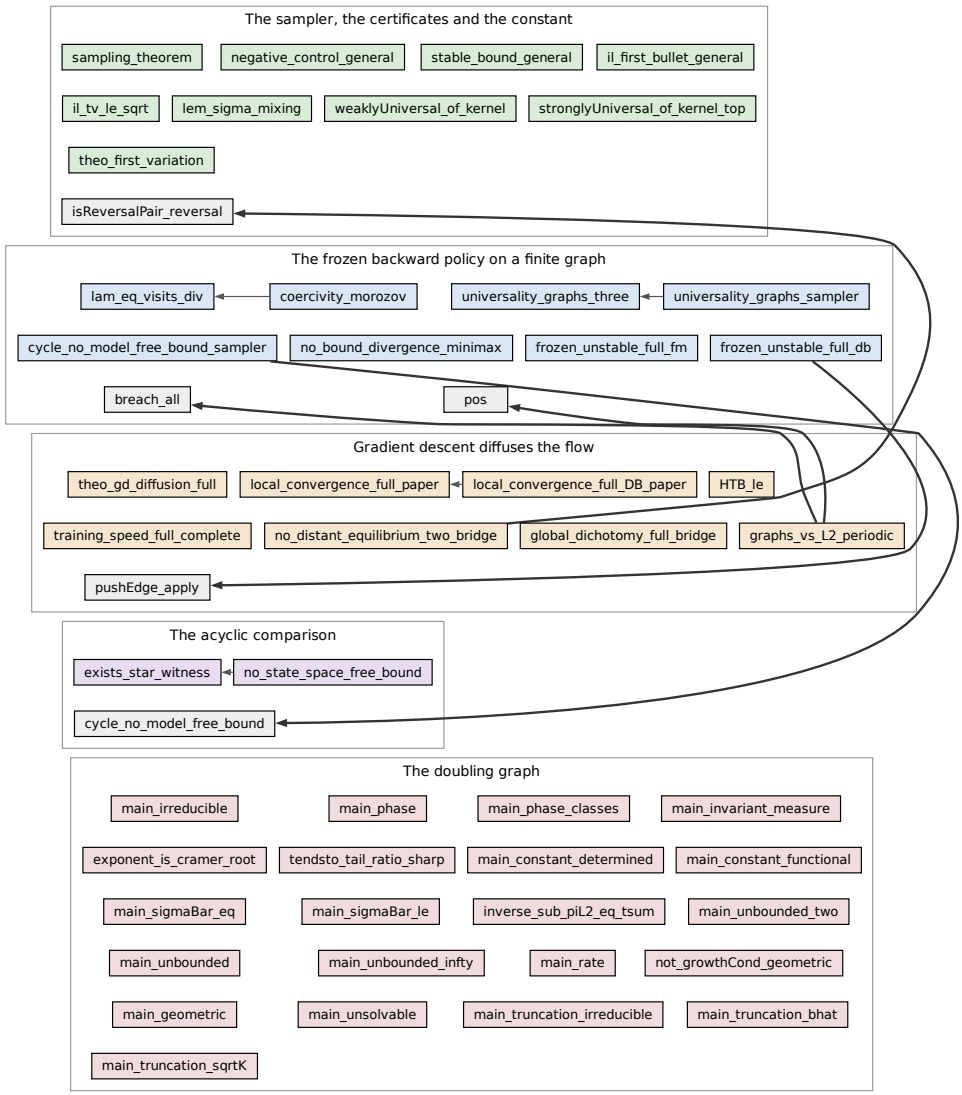


Figure 3: The proof structure of the development at the level of its headlines, read off the compiled environment: one box per result stated in this appendix, grouped by chapter, and an arrow from a result to another when the proof term of the first invokes the second. Heavier arrows leave their chapter — the places where one part of the development stands on another — and the grey boxes are the results they pass through. The graph with every named result the proofs invoke is generated alongside this one and is not printed.

$[0, \infty]$, and '$s_\tau \sim F_{\mathrm{term}}$' is read as an equality of measures, the law of the stopped state being F_{term} normalized.

Theorem 12. *Let $(\pi^*_\rightarrow, F^*_\rightarrow)$ be a generative flow and $F_{\mathrm{init}} \neq 0$, F_{term} finite measures on $\mathcal{S}$, not necessarily balanced; let δF_{term} and δF_{init} be the positive and negative parts of $F_{\mathrm{init}} + F^*_\rightarrow \pi^*_\rightarrow - F^*_\rightarrow - F_{\mathrm{term}}$, and $\widehat{F}_{\mathrm{term}} := F_{\mathrm{term}} + \delta F_{\mathrm{term}}$. The sampler from F_{init} stopped by $\widehat{F}_{\mathrm{term}}$ stops almost surely, and*

$$\widehat{F}_{\mathrm{term}}(\mathcal{S}) = F_{\mathrm{init}}(\mathcal{S}) + \delta F_{\mathrm{init}}(\mathcal{S}), \qquad \mathrm{TV}\Big(s_\tau \,\Big\|\, \frac{\widehat{F}_{\mathrm{term}}}{\widehat{F}_{\mathrm{term}}(\mathcal{S})}\Big) \leq \frac{\delta F_{\mathrm{init}}(\mathcal{S})}{\widehat{F}_{\mathrm{term}}(\mathcal{S})}$$

(Brunswic et al., 2025a).

Proof. The corrected flow is balanced from $F_{\mathrm{init}} + \delta F_{\mathrm{init}}$ to $\widehat{F}_{\mathrm{term}}$ `SamplingGeneral.MFlow.hat_matched`, so the sampling theorem applies to it. The samplers from F_{init}, from δF_{init} and from their sum share one kernel, so their laws are linear in the initial mass at every time, and the distance between the law from F_{init} and the exact law is at most the share of δF_{init} in the initial mass `SamplingGeneral.MFlow.negative_control_tv`. □

The theorem of Brunswic et al. (2025a) as the main text quotes it, imported, on a measurable space. Total variation is the half-convention distance between densities against any common σ-finite dominating measure, and the sampler is that of the sampling theorem run to the corrected terminal flow. The main text applies it at inference with $F_{\mathrm{term}} = 0$: the defect is then $E = F_{\mathrm{init}} + F^*_\leftarrow - F^*_\rightarrow$ itself, $\delta F_{\mathrm{init}} = E^-$ is the correction of the initial flow, and $\widehat{F}_{\mathrm{term}} = E^+ = \widehat{\kappa}$ is the implied terminal flow.

B.1.2 What a small loss certifies

Theorem 13. *Let $q \in [1, \infty]$, ν_{B} finite with $\nu_{\mathrm{B}} \leq M\nu_{\mathrm{T}}$, $(\pi^*_\rightarrow, f^*_\rightarrow)$ a generative flow with star inflow $\ll \nu_{\mathrm{B}}$, $f_{\mathrm{init}}, k \geq 0$ integrable with positive masses and $f_{\mathrm{init}} + f^*_\leftarrow - f^*_\rightarrow - k \in L^q(\nu_{\mathrm{T}})$, $\kappa = k\,\nu_{\mathrm{B}}$, and $C_0 := \nu_{\mathrm{B}}(\mathcal{S})^{1-1/q} M^{1/q}$. The sampler from F_{init} stopped by $\widehat{\kappa} = E^+$ stops almost surely with a law $\ll \nu_{\mathrm{B}}$, and*

$$\mathrm{TV}\Big(s_\tau \,\Big\|\, \frac{\kappa}{\kappa(\mathcal{S})}\Big) \leq \frac{2C_0}{\kappa(\mathcal{S})}\,\mathcal{L}^{\mathrm{stable}}_{\mathrm{FM},q}, \qquad \big|F_{\mathrm{init}}(\mathcal{S}) - \kappa(\mathcal{S})\big| \leq C_0\,\mathcal{L}^{\mathrm{stable}}_{\mathrm{FM},q}.$$

Proof. The negative-control theorem at inference bounds the distance from s_τ to $\widehat{\kappa}$ normalized by the share of E^- in the initial mass `SamplingGeneral.inference_negative_control`. That share, the distance from $\widehat{\kappa}$ to κ and the mass defect are all integrals of the defect, hence bounded by its $L^1(\nu_{\mathrm{B}})$ norm, which Hölder's inequality against $\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}}$ turns into C_0 times its $L^q(\nu_{\mathrm{T}})$ norm; the mass of $\widehat{\kappa}$ is kept away from zero by a floor read off the kernel `RLBound.mass_floor_kernel`, and the assembly is `SamplingGeneral.stable_bound_general_hNC`. □

Item (1) of the stable bound of the main text, both displays, at every $q \in [1, \infty]$ and on a general measurable space; the $L^q(\nu_{\mathrm{T}})$ norm of the defect is $\mathcal{L}^{\mathrm{stable}}_{\mathrm{FM},q}$. Total variation is the half-convention distance between ν_{B}-densities, and s_τ is the output of the sampler of the sampling theorem run at inference to $\widehat{\kappa}$, the corrected terminal flow of the negative-control theorem at $F_{\mathrm{term}} = 0$. The constant is the printed one: it depends on ν_{B}, on how ν_{T} covers it and on the target's mass, and not on the policy. At $q = 1$ the factor $\nu_{\mathrm{B}}(\mathcal{S})^{1-1/q}$ is one. The clauses on driving the loss to zero over a universal family are separate statements.

Theorem 14. *Let $q \in [1, \infty]$ with conjugate q^*, $\nu_{\mathrm{B}} \ll \nu_{\mathrm{T}}$ σ-finite with $\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}} \in L^{q^*}(\nu_{\mathrm{T}})$, $b \geq 1 + \|\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}}\|_{L^{q^*}(\nu_{\mathrm{T}})}$, $F_{\mathrm{init}} = f_{\mathrm{init}}\,\nu_{\mathrm{B}}$ and $\kappa = k\,\nu_{\mathrm{B}}$ probability measures with $\mathcal{H} := -\int k \log k \, d\nu_{\mathrm{B}}$ finite, and $(\pi^*_\rightarrow, f^*_\rightarrow)$ a generative flow with star inflow $\ll \nu_{\mathrm{B}}$. The sampler from F_{init} stopped by $\widehat{\kappa} = E^+$ stops almost surely with a law $\ll \nu_{\mathrm{B}}$, $\mathcal{L}^q_{\mathrm{KL-wFM}} \geq \mathcal{H}$, and if $\mathcal{L}^q_{\mathrm{KL-wFM}} < \infty$,*

$$\mathrm{TV}(s_\tau \,\|\, \kappa) \leq \tfrac{1}{\sqrt{2}}\sqrt{\mathcal{L}^q_{\mathrm{KL-wFM}} - \mathcal{H}} + \min\big(1, \|\frac{d\nu_{\mathrm{B}}}{d\nu_{\mathrm{T}}}\|_{L^{q^*}(\nu_{\mathrm{T}})}(\mathcal{L}^q_{\mathrm{KL-wFM}} - \mathcal{H})\big).$$

Proof. The distance from s_τ to κ splits at $\widehat{\kappa}$ normalized. The first leg is the negative-control bound at inference `SamplingGeneral.inference_negative_control`, the share δ of E^- in the initial mass; the second is Pinsker's inequality against the Kullback–Leibler divergence of κ from $\widehat{\kappa}$ normalized. The cross-entropy term of the loss decomposes into that divergence, the self-entropy and $-\log(1+\delta)$, and Hölder's inequality bounds δ by $(b-1)$ times the penalty, so the penalty absorbs the normalization; `SamplingGeneral.il_first_bullet_general_hNC` assembles the floor and the bound. □

The first bullet of the KL-weak-FM bound of the main text, on a general measurable space, the loss taking the value $+\infty$ where the paper's does. The floor at the self-entropy is unconditional; the bound is the explicit form behind 'of order the square root of the excess', which the next statement extracts. The second bullet — that the floor is attained on a family dominating the target — is a separate statement.

Theorem 15. *Let $N \geq 0$ and $0 \leq L - H \leq 1$. Then*

$$t \leq \tfrac{1}{\sqrt{2}}\sqrt{L-H} + \min\bigl(1, N(L-H)\bigr) \implies t \leq \bigl(\tfrac{1}{\sqrt{2}} + N\bigr)\sqrt{L-H}.$$

Proof. For $0 \leq x \leq 1$, $\min(1, Nx) \leq Nx \leq N\sqrt{x}$. □

The main text's 'to leading order, the square root of the excess': an arithmetic step, stated for real numbers so that it composes with the preceding bound wherever the excess is at most one.

B.1.3 The Green operator

Theorem 16. *Let T be a Markov kernel on a standard Borel space with $\lambda T = \lambda$, $\lambda \neq 0$ finite, P its density action on $L^2(\lambda)$, Π the λ-mean projection, and $B_2 := \sum_{n\geq 0} \|P^n - \Pi\| < \infty$. Then $S := \sum_{n\geq 0}(P^n - \Pi)$ converges in operator norm, and*

$$S(I-P) = I - \Pi, \qquad \|h - \Pi h\| \leq B_2 \,\|(I-P)h\| \quad (h \in L^2(\lambda)).$$

If λ is not a multiple of a Dirac mass, $\|I - \Pi\| = 1$ and $B_2 \geq 1$.

Proof. Since P preserves λ-integrals and fixes the constants, $\Pi P = P\Pi = \Pi$, so $(P^n - \Pi)(I - P) = P^n - P^{n+1}$ and the partial sums telescope to $I - P^{N+1} \to I - \Pi$; convergence, the resolvent identity and the coercivity are `sigma_mixing`. A finite measure on a standard Borel space that is not a multiple of a Dirac mass charges a set of measure strictly between 0 and $\lambda(\mathcal{S})$ `exists_measurableSet_of_ne_smul_dirac`; the centred indicator of that set is non-zero, so $I - \Pi$ is a non-zero orthogonal projection, $\|I - \Pi\| = 1$ `sigma_mixing_beta_zero_of_set` and $B_2 \geq 1$ `sigma_mixing_one_le_B_of_set`. □

The reading of the Green operator through mixing: when the mixing sum of the backward chain is finite, the S here is the Green operator $(I - P + \Pi)^{-1} - \Pi$ of the main text and $\widehat{B} = \|S\| \leq B_2$. The density action, the mean projection and the boundedness of P are constructed, not assumed; summability of the coefficients is the one hypothesis. On a finite space with at least two states and $\lambda > 0$ everywhere the last clause applies, which is the case the graph chapters use.

B.1.4 Universality on a general state space

Theorem 17. *Let $p \in [1, \infty)$ and $\pi^*_{\to}$ a Markov kernel on $\mathcal{S}$ with $\nu_{\mathrm{B}}\pi^*_{\to} = \nu_{\mathrm{B}}$, ν_{B} finite, whose density action $P_\star$ is bounded on $L^p(\nu_{\mathrm{B}})$ with*

$$\sum_{n\geq 0} \|P_\star^n - \Pi\|_{L^p(\nu_{\mathrm{B}})} < \infty.$$

The generative flows with forward policy $\pi^*_{\to}$ and free outflow $f^*_{\to} \geq 0$ are weakly L^p-universal: any two densities in $L^p(\nu_{\mathrm{B}})$ of equal mass are joined by flows of arbitrarily small L^p residual.

Proof. The density action fixes the constants `densityActionCLM_constOne` and preserves ν_{B}-integrals `integral_densityActionCLM`, so the mean projection commutes with it and the abstract mixing hypothesis holds; the abstract theorem then solves the residual equation to any accuracy by truncating the Neumann series of the Green operator and clipping the outflow at a level that goes to infinity `weaklyUniversal_of_massPreserving`. □

The weak half of the L^p-universality theorem of the main text, from its own hypotheses: invariance, boundedness of the density action — a requirement on the parameterization, not a theorem — and summable mixing. Nothing is hypothesised about an operator: $P_\star$ is built from the kernel, absolute continuity of $(f\nu_{\mathrm{B}})\pi^*_{\to}$ follows from invariance, and $\Pi P_\star = P_\star \Pi = \Pi$ is derived. The condition is sufficient, not necessary.

Theorem 18. *Let $\pi^*_{\to}$ be a Markov kernel on $\mathcal{S}$ with $\nu_{\mathrm{B}}\pi^*_{\to} = \nu_{\mathrm{B}}$, ν_{B} finite, whose density action $P_\star$ is bounded on $L^\infty(\nu_{\mathrm{B}})$ with $\sum_{n\geq 0} \|P_\star^n - \Pi\|_{L^\infty(\nu_{\mathrm{B}})} < \infty$. The generative flows with forward policy $\pi^*_{\to}$ and free outflow $f^*_{\to} \geq 0$ are strongly universal: any two densities in $L^\infty(\nu_{\mathrm{B}})$ of equal mass are joined by a flow of zero residual.*

Proof. As before, the density action fixes the constants `densityActionCLM_constOne` and preserves integrals `integral_densityActionCLM`, which gives the abstract mixing hypothesis `mixing_of_massPreserving`; at $p = \infty$ the Green operator's image of the mismatch is bounded, so shifting it by its supremum norm gives a non-negative outflow that solves the residual equation exactly `stronglyUniversal_top`. □

The strong half of the L^p-universality theorem, at $p = \infty$ only, which is where the main text claims it. The outflow bill the main text quotes — at most twice the mixing sum times the mismatch — is read off the construction and is not a clause of this statement; transfer of exactness to other exponents holds per pair of bounded densities and is not stated.

B.1.5 THE FIRST VARIATION

Theorem 19. *Let T be a Markov kernel on $\mathcal{S}$, μ a finite measure with ratio $r = d(\mu T)/d\mu \in [a, b]$, $0 < a \leq b$, $\nu \ll \mu$ finite with $d\nu/d\mu \in L^\infty(\mu)$, and $g \in C^1(0, \infty)$. For every bounded u, with $v := d((u\mu)T)/d\mu$,*

$$\frac{d}{ds}\Big|_{s=0} \mathcal{L}_{g,\nu}\big((1 + su)\mu\big) = \int g'(r)\,(v - ru)\,d\nu.$$

If moreover $\mathcal{S}$ is standard Borel and $\lambda T = \lambda$ for a finite $\lambda \neq 0$ equivalent to μ, the gradient has λ-density $T\psi - r\psi$, with $\psi := g'(r)\,d\nu/d\mu$.

Proof. Differentiating under the integral along the perturbed flow $\mu + s\,u\mu$ gives the directional derivative `General.FirstVariation.hasDerivAt_loss_perturb`; with an invariant measure equivalent to μ the reversal exists, and the integrand is rewritten against it as an inner product with the stated density `General.FirstVariation.first_variation_general`. □

The first-variation theorem as the main text states it. 'Differentiable' is certified in the directional sense — along every essentially bounded direction — which is the sense the full theorem uses; the uniform sense needs the locally Lipschitz derivative of the full form. No invariant measure and no standard Borel hypothesis enter the first clause.

B.2 THE FROZEN BACKWARD POLICY ON A FINITE GRAPH

This chapter. A finite path-connected marked graph — a source s_0, a sink s_f and internal states between them — with a backward policy positive on its edges, closed into a loop by the wrap edge $s_f \to s_0$. The objects are the invariant probability λ of the loop-closed backward chain, the flows it generates, and the sampler of the generative flow obtained by cutting the wrap. The chain is handled through its transition matrix and the linear identities it satisfies; where a statement speaks of backward trajectories, their expected lengths and visit counts enter as the solutions of the corresponding linear systems (`BackwardPolicy.IsGreen`, `BackwardPolicy.IsHitExp`), not as random variables.

B.2.1 THE LOOP CLOSURE AND ITS INVARIANT MEASURE

Theorem 20. *Let G be a finite path-connected marked graph, $\pi_{\leftarrow}$ a backward policy positive on its edges, λ an invariant probability of the loop closure, $N(x)$ the expected number of visits to x of a backward trajectory drawn from the target row and stopped at the source, and $\bar{\sigma}$ its expected length. Then*

$$\lambda(x) = \frac{N(x)}{2 + \bar{\sigma}} \qquad (x \in V).$$

Proof. The visit counts form an invariant vector of the backward chain of total mass $2 + \bar{\sigma}$ `BackwardPolicy.IsGreen.visits_invariant`; an invariant vector of an irreducible chain is a multiple of λ `BackwardPolicy.eq_smul_lam_of_invariant`, and the mass fixes the multiple. □

Kac's formula on a marked graph with the visit counts kept: the invariant measure of the backward chain is the visit ratio, which is how the main text reads it. Visits and expected lengths are the solutions of the renewal and Poisson systems of the chain, not expectations over a constructed process; $N(s_0) = N(s_f) = 1$ by the counting convention.

Theorem 21. *Let G be a finite path-connected marked graph, $\pi_{\leftarrow}$ a backward policy positive on its edges, λ an invariant probability of the loop closure, P its density action on $L^2(\lambda)$, and N, $\bar{\sigma}$ and σ_* the visit counts, the mean backward length and the longest expected time to reach the source. For every h,*

$$\|h - \Pi h\|_{L^2(\lambda)} \le \sigma_* \sqrt{\frac{2 + \bar{\sigma}}{\min N}}\, \|(I - P)h\|_{L^2(\lambda)}.$$

Proof. The invariant probability is positive everywhere `BackwardPolicy.IsInvProb.pos`; a coercivity bound with constant $\sigma_*/\sqrt{\lambda_{\min}}$ follows from the Poisson system for the hitting times `BackwardPolicy.coercivity_lamMin`; and $\lambda_{\min} = \min N/(2 + \bar{\sigma})$ by the visit ratio `BackwardPolicy.lam_eq_visits_div`. □

The hitting-time bound on the Green operator: $\widehat{B} \le \widehat{B}_\sigma$ on every finite path-connected graph, with no aperiodicity and no mixing hypothesis. It is the constant through which the global clock of the main text is read off backward trajectories; it is a guarantee, not an estimate of $\widehat{B}$.

B.2.2 EXISTENCE

Theorem 22. *Let G be a finite path-connected marked graph, $\pi_{\leftarrow}$ a backward policy positive on its edges, λ an invariant probability of the loop closure, $Z > 0$, $c := Z/\lambda(s_0)$, and $e(u \to v) := c\,\lambda(v)\,\widehat{\pi_{\leftarrow}}(v \to u)$ with the wrap edge cut. Then e is a generative flow on G, balanced at every internal state, with*

$$e(s_f \to s_0) = Z, \qquad F_{\text{init}}(V) = Z, \qquad F_{\text{term}} = Z\,\pi_{\leftarrow}(s_f \to \cdot), \qquad F^*_{\to} \ge 0,$$

and $F_{\text{term}} > 0$ at every state with an edge to the sink.

Proof. Positivity of λ `BackwardPolicy.IsInvProb.pos` makes c well defined. The wrap edge carries $c\lambda(s_0) = Z$ `BackwardPolicy.edgeFlow_wrap`, and the initial flow, the source row, sums to the same `BackwardPolicy.sum_initFlow`; the terminal flow is the sink row scaled by Z `BackwardPolicy.termFlow_eq_target`, positive wherever the policy charges the edge `BackwardPolicy.termFlow_pos_of_edge`; the star outflow $c\lambda(u) - e(u \to s_f)$ is non-negative `BackwardPolicy.outflowStar_nonneg`; and invariance of λ under the backward chain is exactly the balance constraint at every internal state `BackwardPolicy.fmDefect_eq_zero`. □

Item (3) of the graph universality theorem: existence is free, under positivity of the policy on the edges alone, with no aperiodicity and no mixing hypothesis. Every target positive on the terminating states is reached exactly by freezing the sink row of the backward policy to it; a target vanishing on some terminating states is handled through the subgraph they span. That the invariant probability is unique and positive — items (1) and (2) — is established separately in the development.

Theorem 23. *In the setting of Theorem 22, let G have an internal state. The sampler of e stops almost surely, with*

$$\mathbb{E}(\tau) \le 1 + \frac{F^*_{\to}(V)}{F_{\text{init}}(V)}, \qquad s_\tau \sim \frac{\pi_{\leftarrow}(s_f \to \cdot)}{1 - \pi_{\leftarrow}(s_f \to s_0)} \text{ on the internal states.}$$

Proof. The flow of `BackwardPolicy.universality_graphs_three` is balanced, so the sampling theorem applies to its sampler `SamplerWiring.edgeFlow_sampler`: it stops almost surely and emits the terminal flow normalized. The terminal mass on the internal states is $Z(1 - \pi_{\leftarrow}(s_f \to s_0))$ `SamplerWiring.sum_internal_termFlow`, which is positive because the direct edge cannot carry the whole sink row when an internal state exists `SamplerWiring.pb_snk_src_lt_one`. □

The sampler clause of item (3): the target is sampled exactly, with or without an edge from the source straight to the sink. '$s_\tau \sim \kappa$' is read as the law of s_τ being κ normalized on the states where the sampler can stop; a target charging the direct edge loses that mass to the normalization.

B.2.3 Ratio losses on a cycle

Theorem 24. *Let $\mathcal{C}_N$ be the cycle graph on $N \ge 2$ internal states, $g \ge 0$ on $(0, \infty)$, continuous at 1, with $g(1) = 0$, $\nu_{\mathrm{T}} \ge 0$ on the states, and κ a probability on the internal states other than the point mass at the second one. No Φ with $\Phi(L) \to 0$ as $L \to 0^+$ satisfies*

$$\mathrm{TV}(s_\tau \,\|\, \kappa) \le \Phi\big(\mathcal{L}_{g,\nu_{\mathrm{T}}}(F)\big)$$

for every non-negative edge-carried flow F on $\mathcal{C}_N$ that is balanced with positive denominators and positive terminal mass.

Proof. The circulation of size k on the cycle is balanced and its sampler emits the second state `SamplerWiring.cycle_sampler`, so its error against κ is $1 - \kappa(x_2) > 0$ for every k while its loss tends to 0 with k; a Φ vanishing at 0^+ cannot dominate a positive constant along that sequence `Remarks.cycle_no_model_free_bound`. □

On one cycle, whatever the training measure and for every target but one, no function of a ratio loss that vanishes with it bounds the sampling error: a circulation drives every ratio to one while the sampler keeps emitting the second state. This is the statement for the flow-matching loss on states; that the same circulation lowers the detailed- and trajectory-balance forms is due to Brunswic et al. (2024), and whether those can certify is not settled here.

Theorem 25. *Let g be continuous at 1 with $g(1) = 0$, $\varepsilon > 0$, and $\mathcal{C}_N$ the cycle graph on $N \ge 2$ internal states. Then*

$$\inf_{\nu_{\mathrm{T}}} \sup\big\{\mathrm{TV}(s_\tau \,\|\, \kappa) \;:\; \mathcal{L}_{g,\nu_{\mathrm{T}}}(F) \le \varepsilon\big\} = 1,$$

the supremum over probability targets κ on the internal states and non-negative edge-carried flows F with a sampler, the loss taken against κ.

Proof. For every $\delta \in (0, 1)$ a circulation of loss at most ε samples with error $1 - \delta/(N-1)$ against a suitable target, and no error exceeds 1; the supremum is therefore 1 for every weight `SamplerWiring.sSup_tvSet`, and so is the infimum over weights. □

The minimax identity of the main text on the cycle of each size: at any loss level, whatever the training measure, some target is sampled with error arbitrarily close to the largest possible. Balance is not asked of the flows in the supremum, only a loss below ε; the value 1 is a supremum, approached and not attained.

B.2.4 FREEZING IS NOT STABILITY

Theorem 26. *Let G be a finite path-connected marked graph, $\pi_{\leftarrow}$ a backward policy positive on its edges, λ an invariant probability of the loop closure, F^* its balanced edge flow, $\gamma \neq 0$ a non-negative circulation with no inflow at the source or the sink, g with $g(1) = 0 < g(x)$ for $x \neq 1$, $\nu_{\mathrm{T}} > 0$ on the states, and $\varepsilon > 0$ with $2\varepsilon\gamma \leq F^*$. With $F := F^* - \varepsilon\gamma$ and $F_0 := \varepsilon\gamma$,*

$$0 \leq F_0 \leq F, \qquad \mathcal{L}_{g,\nu_{\mathrm{T}}}(F + F_0) = 0 < \mathcal{L}_{g,\nu_{\mathrm{T}}}(F).$$

Proof. The invariant probability is positive `BackwardPolicy.IsInvProb.pos`, so the balanced edge flow charges every edge and ε is admissible: $F_0 \leq F$ `BackwardPolicy.epsCirc_le_perturbedFlow`. The inflow of $F + F_0 = F^*$ `BackwardPolicy.perturbedFlow_add_epsCirc` is λ itself `BackwardPolicy.edgeInflow_edgeFlow_one`, whose balance ratios are all one `BackwardPolicy.stateRatio_eq_one_iff`, so the loss vanishes `BackwardPolicy.fmLoss_eq_zero`. Removing the circulation leaves an inflow that is not invariant `BackwardPolicy.not_invariant_perturbedFlow` — it stays positive `BackwardPolicy.edgeInflow_perturbedFlow_ge` with positive ratios `BackwardPolicy.stateRatio_pos` — so some ratio differs from one and the loss is positive `BackwardPolicy.fmLoss_pos`. □

The flow-matching half of 'freezing does not restore stability': the stability of Brunswic et al. (2024), which forbids a 0-flow from improving a flow, fails for every admissible generator and every positive training measure. The circulation is carried as the properties the proof uses — non-negative, non-zero, no inflow at the marks — of which the indicator of a directed cycle avoiding both marks is an instance. The loss is the state-space loss read on the inflow of an edge flow.

Theorem 27. *Let G, $\pi_{\leftarrow}$, λ, γ, g, ε, F and F_0 be as in Theorem 26, and $\widehat{\nu}$ positive on the edges of the loop closure and zero off them. Then $0 \leq F_0 \leq F$ and*

$$\mathcal{L}_{\mathrm{DB},g,\widehat{\nu}}(F + F_0) = 0 < \mathcal{L}_{\mathrm{DB},g,\widehat{\nu}}(F).$$

Proof. Adding F_0 back gives F^* `BackwardPolicy.perturbedFlow_add_epsCirc`, whose edge ratios are all one `BackwardPolicy.dbRatio_edgeFlow_eq_one`, so the DB loss vanishes `BackwardPolicy.dbLoss_eq_zero_of`. The perturbed flow is positive on the edges of the loop closure `BackwardPolicy.perturbedFlow_pos_of_hatEdge` and zero off them `BackwardPolicy.perturbedFlow_eq_zero_of_not_hatEdge`, its edge ratios are positive `BackwardPolicy.dbRatio_perturbedFlow_pos`, and its push through the lift `Balance.pushEdge_apply` differs from itself somewhere `BackwardPolicy.pushEdge_perturbedFlow_ne`, so some ratio differs from one and the loss is positive `BackwardPolicy.dbLoss_pos_of`; admissibility is `BackwardPolicy.epsCirc_le_perturbedFlow`. □

The detailed-balance half. The DB loss is read on edge flows through the edge lift of the backward chain, as the lift proposition of the development defines it; the same circulation lowers it.

B.3 GRADIENT DESCENT DIFFUSES THE FLOW

This chapter. The dynamics of a balance loss under a frozen kernel: the linearization of its gradient near balance on a standard Borel space, the exponential clock from an explicit neighbourhood on a finite state space for the flow-matching and detailed-balance losses, the global two-phase clock of the squared-logarithm generator on a finite graph, and the equilibria a generator allows. The gradient flow is constructed in each statement that speaks of it, in the class C^1; it is never hypothesised.

B.3.1 THE GRADIENT AS A DIFFUSION

Theorem 28. *Let T be a Markov kernel on a standard Borel space with $\lambda T = \lambda$, $\lambda \neq 0$ finite, P its density action on $L^2(\lambda)$, $A = P - I$, $\nu = w\lambda$ with $0 \leq w \in L^\infty(\lambda)$, M_w the multiplication by w, and $g \in C^3[1-a, 1+a]$, $0 < a < 1$, with $g(1) = g'(1) = 0 < g''(1)$. There is an explicit K such*

that, for $0 < \varepsilon \le a/4$ and $|h| \le \varepsilon$, the flow $(1+h)\lambda$ has ratio $(1+Ph)/(1+h)$, and the gradient of $\mathcal{L}_{g,\nu}$ at it has an $L^2(\lambda)$ density $D(h)$ with

$$\|D(h) - Hh\| \le K\varepsilon \, \|Ah\|, \qquad H := g''(1)\, A^\dagger M_w A;$$

H is the only homogeneous H' with $\|D(h) - H'h\| = o(\|h\|)$ as $\|h\|_\infty \to 0$; and if T is reversible and $w \equiv 1$, then $H = g''(1)(I-P)^2$, and $P\varphi = \beta\varphi$ implies $H\varphi = g''(1)(1-\beta)^2\varphi$.

Proof. The ratio of $(1+h)\lambda$ is $(1+Ph)/(1+h)$ `GdDiffusionGeneral.ratio_flowOf`, within explicit bounds of one `GdDiffusionGeneral.gd_ratio_bounds`; the gradient exists with density $P^\dagger\psi - r\psi$ `GdDiffusionGeneral.gd_gradient`, which expands to second order `GdDiffusionGeneral.gd_expansion` and vanishes at $h = 0$ `GdDiffusionGeneral.gradD_zero`; the linear part is unique `GdDiffusionGeneral.linearization_unique`. For a reversible kernel $A^\dagger = A$, so $H = g''(1)A^2$ `GdDiffusionGeneral.linHess_reversible`; its eigenmodes are those of P `GdDiffusionGeneral.linHess_eigen` and decay accordingly `GdDiffusionGeneral.eigenmode_decay`. □

The full form of 'gradient descent as backward diffusion', on a standard Borel space, which covers the Polish setting of the main text; the constant K is the printed one, and the gradient is that of the first-variation theorem read against λ. The heat-flow reading needs reversibility, and no loop closure of a marked graph is reversible, so on graphs it is a reading and not an identity; for the edge lift of the detailed-balance loss it is false.

B.3.2 THE CLOCK NEAR BALANCE

Theorem 29. *Let T be a Markov kernel on a finite state space with invariant probability $\lambda > 0$, P its density action, $\nu = w\lambda$ with $w \ge w_{\min} > 0$, $g \in C^3[1-a, 1+a]$ with $g'(1) = 0 < g''(1)$, $\widehat{B} \ge 1$ with $\|h - \Pi h\| \le \widehat{B}\,\|(I-P)h\|$ on $L^2(\lambda)$, and $\varrho := g''(1)\, w_{\min}/\widehat{B}^2$. There are explicit $\varepsilon_0, \gamma_0 > 0$, functions of g, a, $w_{\min}$, $\|w\|_\infty$, $\widehat{B}$ and $\min\lambda$, such that for $\|h_0\|_{L^2(\lambda)} \le \varepsilon_0$: (a) the gradient flow $(1+h_t)\lambda$ of $\mathcal{L}_{g,\nu}$ from $(1+h_0)\lambda$ exists, is unique, stays positive, and for some balanced $c_\infty\lambda$*

$$\|1 + h_t - c_\infty\|_{L^2(\lambda)} \le 2e^{-\varrho t/2}\, \|h_0 - \Pi h_0\|_{L^2(\lambda)};$$

(b) for $0 < \gamma \le \gamma_0$ the gradient descent h_k from h_0 is well defined and

$$\|h_{k+1} - \Pi h_{k+1}\| \le (1 - \gamma\varrho/4)\, \|h_k - \Pi h_k\|.$$

Proof. $C_\infty = (\min\lambda)^{-1/2} \ge 1$ `one_le_Cinf` controls the passage from $L^2(\lambda)$ to the uniform norm; within the neighbourhood the loss is differentiable along the flow with the gradient as its field `hasDerivAt_loss_ipL2_local`; the constants are those of the printed formulas `constW_bounds`; existence, uniqueness, positivity and the exponential bound are `local_convergence_full_exists`, and the descent clause is `local_convergence_gd_C3On`. □

The local exponential convergence theorem of the main text for the flow-matching loss, as printed. $\widehat{B}$ is any coercivity constant, so the norm of the Green operator gives the best rate and the hitting-time bound a guaranteed one. The flow is constructed, in the class C^1, and its field is the $L^2(\lambda)$ gradient at every time; $\min\lambda$ and $\|w\|_\infty$ are the actual extrema, and the constants are the printed formulas.

Theorem 30. *Let $\pi_\leftarrow$ be a backward policy on a finite state space with invariant probability $\lambda > 0$, E its edge set, K_2 its edge lift, and $\lambda_2(s,s') := \pi_\leftarrow(s' \to s)\,\lambda(s')$. Under the hypotheses of Theorem 29 on g and a, its conclusions hold for the detailed-balance loss, that is, for (K_2, λ_2) in place of (T, λ), with $w \ge w_{\min} > 0$ on E and $\widehat{B} \ge 1$ a coercivity constant of K_2 on $L^2(\lambda_2)$.*

Proof. The edge lift is a Markov kernel on E `edgeKernelE_isMarkovOn` leaving λ_2 invariant `edgeMeasureE_isInvariant`, positive `edgeMeasureE_pos` and of total mass one `edgeMeasureE_total`; the flow-matching theorem on (E, K_2, λ_2) `local_convergence_full_paper` is the statement. □

The detailed-balance theorem is the flow-matching theorem for the edge-lifted chain, whose Markov property, invariance, positivity and total mass are proved from the backward policy; that the balance loss on E is the DB loss is the lift identity of the development. A coercivity constant of K_2 is one more than one of the chain on states, which is the main text's '$1 + \widehat{B}$ for DB'.

Theorem 31. *Let $\pi_{\leftarrow}$ be a backward policy on a finite state space with invariant probability $\lambda > 0$, K_2 and λ_2 its edge lift and edge measure, $g = (\log x)^2$, $H_{\mathrm{DB}} = g''(1)\, A^{\dagger} A$ and $H_{\mathrm{TB}} = \ell\, g''(1)\, A^{\dagger} \Theta_{\ell} A$ the Hessians at balance of the detailed- and trajectory-balance losses on the edges, Θ_{ℓ} the window operator of length ℓ, and $B' \geq \sum_{m=1}^{N} \|K_2^m - \Pi_2\|_{L^2(\lambda_2)}$ for every N. For every h on the edges,*

$$\langle h, H_{\mathrm{TB}} h \rangle \leq \ell\,(1 + 2B')\,\langle h, H_{\mathrm{DB}} h \rangle.$$

Proof. The window operator is self-adjoint and positive, and its quadratic form is at most $1 + 2B'$ times the identity's by the mixing bound on the powers of the lift `TBHessian.Theta_le`. □

The trajectory-balance Hessian is controlled by the detailed-balance one up to a factor growing with the trajectory length: a step cap for TB follows from the DB one, and no rate is claimed for TB anywhere in this paper. The bound is vacuous where the lifted mixing sum is infinite, exactly as the printed one.

B.3.3 The clock far from balance

Theorem 32. *Let G be a finite path-connected marked graph, $\pi_{\leftarrow}$ a backward policy positive on its edges, λ the invariant probability of the loop closure, N, $\bar{\sigma}$ and σ_* the visit counts, the mean backward length and the longest expected exit time, $g = (\log x)^2$, $\nu = w\lambda$ with $w \geq w_{\min} > 0$, $u_0 > 0$, and*

$$\varrho_{\sigma} := \frac{2\, w_{\min} \min N}{\sigma_*^2\,(2 + \bar{\sigma})}.$$

Then $\lambda = N/(2 + \bar{\sigma})$, and there are explicit $\kappa, \gamma_* > 0$ such that: (a) the gradient flow μ_t of $\mathcal{L}_{g,\nu}$ from $u_0\lambda$ exists, is the unique positive solution, and

$$\mathcal{L}(\mu_t) \leq \left(\mathcal{L}(\mu_0)^{-1} + \kappa^2 t\right)^{-1}, \qquad \mu_t \longrightarrow \|u_0\|_{L^2(\lambda)}\, \lambda,$$

exponentially at the rate $\varrho_{\sigma}/(2m^2)$ after an explicit time, m the mass then; (b) for $0 < \gamma \leq \gamma_*$ the gradient descent u_k from u_0 is bounded below, converges to a balanced flow, satisfies $\mathcal{L}(u_k) \leq \left(\mathcal{L}(u_0)^{-1} + k\gamma\kappa^2/4\right)^{-1}$, and after an explicit step k_0

$$\|u_{k+1} - \Pi u_{k+1}\| \leq \Big(1 - \frac{\gamma \varrho_{\sigma}}{4m^2}\Big)\|u_k - \Pi u_k\|, \qquad m \text{ the mass at } k_0.$$

The constants see $(G, \pi_{\leftarrow})$ only through σ_*, $\bar{\sigma}$ and $\min N$; γ_* also through the least positive transition probability and $|V|$.

Proof. The flow and its two phases, the Łojasiewicz decay and the exponential tail, are `training_speed_full_paper`; the descent, with the step cap and the entry time computed from u_0, is `DiscreteGlobal.training_speed_gd_exact`. □

The principal declaration of the global clock: the whole theorem in one statement, the flow constructed rather than hypothesised, the quantifiers of the descent clause from u_0 alone, and $\mathcal{L}(\mu_0)^{-1} := +\infty$ at a balanced start. The rate ϱ_{σ} is the local rate with the hitting-time constant $\widehat{B}_{\sigma}$ in place of $\widehat{B}$; it is a guarantee, and the main text says where it is loose. The dependence sentence is read off the explicit formulas.

B.3.4 Equilibria and the generator

Theorem 33. *Let T be a Markov kernel on a standard Borel space $\mathcal{S}$ with $\lambda T = \lambda$, λ finite, μ a flow equivalent to λ meeting the first-variation hypotheses for ν, $u = d\mu/d\lambda$, and g differentiable*

on $(0,\infty)$*. Then* $\mathcal{L}_{g,\nu}(c\mu) = \mathcal{L}_{g,\nu}(\mu)$ *for* $c > 0$*, and every* λ*-density* D *of the gradient satisfies* $\langle D, u\rangle_{L^2(\lambda)} = 0$*. If* T *is ergodic for* $\lambda \neq 0$*,*

$$\mu(\mathcal{S}) \leq \sqrt{\lambda(\mathcal{S})}\, \|u\|_{L^2(\lambda)},$$

with equality iff μ *is balanced.*

Proof. The λ-reversal exists on a standard Borel space `IsInvariant.isReversalPair_reversal`, so the gradient has a density, unique in $L^2(\lambda)$ `General.Bridge.isGradDensity_unique`; scale invariance, orthogonality and the Cauchy–Schwarz bound with its equality case are `General.no_distant_equilibrium_two_general`. □

The static half of 'the inflation is the algorithm': training moves on a sphere of $L^2(\lambda)$, the mass is bounded on that sphere, and the bound is attained exactly at the balanced flow. That the gradient never lowers the mass for a strictly unimodal generator is the first clause of the next statement. The sphere is invariant only at a fixed training measure.

Theorem 34. *Let* T *be a Markov kernel on* $\mathcal{S}$ *with* $\lambda T = \lambda$*,* λ *finite, and* $\mu \ll \lambda$ *a finite flow of ratio* r*. (1) If* g *is a strictly unimodal generator and* ν*, with* $d\nu/d\mu > 0$*, meets the first-variation hypotheses, every* λ*-density* D *of the gradient of* $\mathcal{L}_{g,\nu}$ *at* μ *satisfies*

$$\int D\, d\lambda = \int g'(r)\,(1-r)\,\frac{d\nu}{d\mu}\, d\lambda \;\leq\; 0,$$

with equality iff μ *is balanced; and* μ *is balanced if it is a critical point of* $\mathcal{L}_{g,\nu}$*. (2) If* g' *is continuous on* $(0,\infty)$ *and vanishes on two bands, one on each side of* 1*, that contain the values of* r*, then* μ *is a critical point of* $\mathcal{L}_{g,\nu}$ *for every finite* $\nu \ll \mu$*.*

Proof. Item (1) is the mass identity for the gradient density, whose sign is that of $g'(r)(1-r)$ `General.Bridge.no_distant_equilibrium_one_bridge`; item (2) reads the derivative along every direction as an integral of $g'(r)$ against a measure carried by the bands, where it vanishes `General.Bridge.nonlinear_freezing_one_critical`. □

The general halves of the global dichotomy: for a strictly unimodal generator, such as $(\log x)^2$, the gradient never lowers the mass and the only critical points are balanced; for a generator whose derivative vanishes on bands, flows a fixed distance from balance are frozen. The convergence clause of item (1) on a finite space, and that some such generators freeze an open set already on two states, are separate statements of the development.

B.3.5 Periodic chains

Theorem 35. *Let* G *be a finite path-connected marked graph,* $\pi_{\leftarrow}$ *a backward policy positive on its edges,* λ *an invariant probability of the loop closure,* P *its density action, and* $d \geq 2$ *a divisor of every return time of the backward chain to some state. Then*

$$\|P^{dk} - \Pi\|_{L^2(\lambda)} \geq 1 \quad (k \geq 0), \qquad \sum_{n \geq 0} \|P^n - \Pi\|_{L^2(\lambda)} = +\infty.$$

Proof. Positivity of λ `IsInvProb.pos` and the reach of every state from every other along the loop `BackwardPolicy.breach_all` carry periodicity from one state to the chain; at multiples of the period the powers of P keep a centred function of norm one at distance one from its mean, so the coefficients do not vanish `RemarksA.mixing_fails_of_periodicAt`. □

Why the main text reads the constant through the Green operator and not through the mixing sum: on every leveled graph — an autoregressive state space among them — the loop closure is periodic and the mixing sum diverges, while $\widehat{B}$, the norm of the Green operator, stays finite on any finite path-connected graph by the hitting-time bound.

B.4 THE ACYCLIC COMPARISON

This chapter. The comparison with the acyclic guarantees of da Silva et al. (2025b) on finite directed acyclic graphs, where a model-free bound on the sampling error does exist. The statements concern the star graph with a uniform backward policy and uniform training: its residual can be made arbitrarily small while the sampling error stays at one eighth, so no bound that depends on the residual alone holds uniformly over the state space.

B.4.1 NO BOUND FREE OF THE STATE SPACE

Theorem 36. *For every $\varepsilon > 0$ there is K such that the star graph with terminals x^*, x' and K fillers, target $\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2K}, \dots)$, forward policy $p_F = (\frac{3}{8}, \frac{1}{8}, \frac{1}{2K}, \dots)$ and uniform training distribution has $\chi^2 = 0$ between the training and the uniform distributions, and*

$$M' \leq 1, \qquad 0 < \mathcal{R} \leq \varepsilon, \qquad \mathrm{TV}(p_F, \pi) = \tfrac{1}{8},$$

$\mathcal{R}$ its residual.

Proof. The ratio bound $M' \leq 1$ `star_Mprime`, the residual's closed form and its bound `star_energy`, and the total variation, which does not depend on K `star_tv`. □

Item (ii) of the acyclic comparison: an explicit instance in which the residual of da Silva et al. (2025b) is arbitrarily small while the sampling error is one eighth, so the constants of their bound cannot be free of the state space. $K = \lceil 1/\varepsilon \rceil + 1$ suffices, the residual being $\big((\log \frac{3}{2})^2 + (\log 2)^2\big)/(K+2) \leq 1/(K+2)$.

Theorem 37. *No F with $F(\mathcal{R}) \to 0$ as $\mathcal{R} \to 0^+$ satisfies*

$$\mathrm{TV}(p_T, \pi) \leq F(\mathcal{R})$$

over all finite instances with $\chi^2(q_{E,T} \| p_{E,T}) = 0$ and $M' \leq 1$, where p_T is the sampled distribution, π the target and $\mathcal{R}$ the residual.

Proof. Along the star witnesses the residual tends to 0 while the total variation stays at one eighth `exists_star_witness`, and F of the residual would have to tend to 0. □

The 'in particular' of the acyclic comparison: model-free bounds exist on finite directed acyclic graphs, but none is uniform over the state space. The quantification over instances is over types in the lowest universe, where the star witnesses live.

B.5 THE DOUBLING GRAPH

The object. The graph is the doubling graph of Figure 2: a source s_0, a ladder $1, 2, 3, \dots$, a sink s_f, and the forward edges $s_0 \to 1$, $j \to j+1$, $2j \to j$, $j \to s_f$. The predecessors of a ladder state j are $j-1$ and $2j$, so the backward chain either *decrements* or *doubles*, and the loop closure adds the wrap $s_f \to s_0$. The backward policy is $\pi_{\leftarrow}(j \to 2j) = \varepsilon(j)$ and $\pi_{\leftarrow}(j \to j-1) = 1 - \varepsilon(j)$, and the family studied throughout is

$$\varepsilon_{c,s}(j) \;:=\; \frac{c}{(j+1)^s}, \qquad s \geq 0, \quad 0 < c < 2^s, \tag{6}$$

with geometric damping $\varepsilon(j) = a\rho^j$, $0 < \rho < 1$, as the contrast. The target row $\widehat{\pi_{\leftarrow}}(s_f \to \cdot)$ is a probability supported in $\{1, \dots, d\}$ of mean $\bar{\jmath}$. A *truncation at K* deletes every doubling edge $j \to 2j$ with $2j > K$, leaving a chain on $K+2$ states. The operator this chapter calls the *diffusion operator*, $S = (\mathrm{Id} - P_\star + \Pi)^{-1} - \Pi$ on $L^2(\lambda)$, is the Green operator of the main text, and $\widehat{B} = \|S\|$; on the loop closure $P_\star$ is the density action of the backward chain.

Phases. Whether the loop closure has an invariant probability at all depends on the parameters, and the table is complete but for one point: transient for $s < 1$ and for $s = 1$, $c > 1/\ln 2$; null recurrent for $s = 1$, $1 \leq c < 1/\ln 2$; positive recurrent for $s = 1$, $c < 1$ and for $s > 1$; and at $s = 1$, $c = 1/\ln 2$ only the absence of positive recurrence is established (Theorems 38, 39 and 40). The

uniform backward policy is the member $c = 1/2$, $s = 0$, and is transient. Everything below is about the positive recurrent rows, and the finer results about the row $s = 1$, $c < 1$.

Polynomial decay. On that row the invariant measure decays like a power, $\lambda_m \sim C\, m^{-p_*}$ with $p_* > 1$ the root of the Cramér equation $c\,(2^p - 1) = p$, and with an explicit rate of approach (Theorem 41); no other exponent is possible, since any polynomial decay of a cut-balanced sequence forces its exponent to be a Cramér root (Theorem 42); the constant C is determined by the first d rungs and is a linear functional of them (Theorems 44 and 45). The consequence that drives the chapter is the tail ratio: the mass beyond a rung, divided by the mass at the rung, grows linearly, $L(m)/\lambda_m = m\,(1 + o(1))/(p_* - 1)$ (Theorem 43).

Summability. A bounded Green operator would keep that ratio bounded — it forces the tail to decay geometrically — so the linear growth makes the Rayleigh quotient of the centred tail indicators vanish, at a rate at least $\sqrt{m}$ along the cuts (Theorem 52): there is no bounded inverse in any L^p, the series $\sum_n (P_\star^n - \Pi)$ does not converge, and the mixing sum $\sum_n \widehat{\beta}_n$ is infinite, on the whole positive recurrent part of the family (Theorems 50, 51 and 49). The two constants of the main text are ordered one way only: where the mixing sum converges, the Green operator is its sum and $\widehat{B} \le \sum_n \widehat{\beta}_n$ (Theorem 48), which on a finite truncation is a strict inequality in general and on the infinite chain says nothing.

Geometric damping. The unboundedness theorem rests on a growth condition on the doubling probability that the polynomial family meets and geometric damping does not (Theorem 53); and under geometric damping the picture reverses: the loop closure is positive recurrent, the tail is comparable to the mass at the cut, and the Green operator is bounded (Theorem 54). The obstruction of this chapter is therefore a property of the backward policy and not of the graph. No numeric bound on $\widehat{B}$ and no summable mixing is claimed for the geometric family.

The boundary against the condition of Morozov et al. The quantity through which Morozov et al. control training is the expected length of a backward trajectory, $\bar{\sigma}$. On the row $s = 1$, $c < 1$ it is finite, $\bar{\sigma} = \bar{\jmath}/(1 - c)$, uniformly over the truncations (Theorems 46 and 47), while $\widehat{B}$ is infinite; on the truncations both are finite, and $\widehat{B}_K$ grows at least like $\sqrt{K}$ while $\bar{\sigma}$ stays put (Theorems 56, 57 and 58). Exact flow matching with square-integrable outflows fails for some pair on the infinite chain (Theorem 55); whether weak universality holds there is open. On the transient rows there is no invariant probability, no balanced flow, and nothing for the condition to rest on.

What is not modelled. The development builds no stochastic process. It works with the invariant measure through the balance identity across a cut, with the expected hitting time through the supersolution that dominates it, with the recurrence classes through the least solutions of the corresponding systems, and with the descent through a recursion on the state rather than a chain of random variables. Everything below is therefore a statement about those objects, and §B.6 says where that costs something.

B.5.1 The chain, and when it settles

Theorem 38. *The loop closure is irreducible, and every invariant measure of it is positive at every state.*

Proof. The graph half is a pair of explicit paths `reach_all_none`. Every state reaches the ladder state 1: from a ladder state the decrements walk down to the source, the wrap carries it to the sink, and the target row re-enters the ladder. Conversely 1 reaches every state, by climbing: from n the doubling edge to $2n$ followed by $n - 1$ decrements lands on $n + 1$, and $2n \ge n + 1$. Concatenating the two gives reachability between any pair.

The measure half is general and needs only that `PreStat.pos_of_irreducible`: an invariant measure is positive somewhere, and invariance transports that positivity along any path, so irreducibility spreads it over the whole state space. □

Irreducibility and positivity are the standing hypotheses under which the rest of this appendix speaks of *the* invariant probability and of the operators built from it. The main text's universality and

convergence theorems are stated for chains carrying such a measure; what follows establishes, for each member of the family, whether one exists.

Theorem 39. *Let $\varepsilon = \varepsilon_{c,s}$ with $c > 0$, in the standing range. The loop closure has an invariant probability if $s > 1$, or $s = 1$ and $c < 1$; it has none if $s < 1$, or $s = 1$ and $c \geq 1$.*

Proof. Three cases, each a drift computation against a Lyapunov function, and no more.

For $s < 1$ the drift coefficient $c(m+1)^{1-s}$ grows without bound, which forces the tail of any candidate invariant sequence to grow too fast to be summable `isEmpty_stat_of_family_lt_one`. For $s = 1$ the coefficient is the constant c, and $c \geq 1$ is exactly the threshold past which the same tail argument applies `isEmpty_stat_of_family_ge_one`. For $s > 1$ the drift is eventually negative and a Foster argument on the ladder height produces an invariant probability `exists_stat_of_family_gt_one`; the remaining regime $s = 1$, $c < 1$ is the same argument at the explicit threshold.

Each conclusion is drawn from the cut balance directly, without constructing the chain as a stochastic process. □

This is the phase diagram that decides which members of the family the main text's theorems can even be stated about. In the two regimes with no invariant probability there is no λ, hence no Π, no mixing coefficients and no diffusion constant: the hypotheses of those theorems fail there for want of the objects they quantify over, which is a different kind of failure from the one this appendix is about.

Theorem 40. *Let $\varepsilon = \varepsilon_{c,s}$ with $c > 0$ in the standing range. The loop closure is irreducible, and its backward chain is: transient for $s < 1$; positive recurrent for $s = 1$ and $c < 1$; null recurrent for $s = 1$ and $1 \leq c < 1/\ln 2$; positive recurrent at no state for $s = 1$ and $c = 1/\ln 2$; transient for $s = 1$ and $c > 1/\ln 2$; and positive recurrent for $s > 1$.*

Proof. Irreducibility is the reach of every state from every other along the loop `reach_all_none`. Each row is its own theorem, from the drift of the chain: $s > 1$ `phase_gt_one` and $s < 1$ `phase_lt_one`; at $s = 1$, the row $c < 1$ `phase_one_lt`, and the rows $c \geq 1$, where positive recurrence fails `phase_one_ge`, split into the null-recurrent band below $1/\ln 2$ `phase_one_null` and the transient tail above it `phase_one_transient`. □

The labels of the phase table. Transience, null recurrence and positive recurrence are read off the least solutions of the chain's linear systems — the return probabilities and the expected return times — since no stochastic process is constructed; on a countable irreducible chain these are the textbook classes. The case $c = 1/\ln 2$ is only excluded from positive recurrence and is not resolved further. For the main text: on a transient row there is no invariant probability, hence no balanced flow, and the expected backward length of Morozov et al. has nothing to rest on; the uniform backward policy is the member $c = 1/2$, $s = 0$.

B.5.2 Polynomial decay: the invariant measure

Theorem 41. *Let $\varepsilon = \varepsilon_{c,1}$ with $0 < c < 1$, λ an invariant probability of the loop closure, and $L(m) := \sum_{j \geq m} \lambda_j$. There are a root $p > 1$ of ψ_c, $C > 0$, $\vartheta \in (0,1)$ and $c_6 > 0$ such that*

$$\lambda_m \, m^p \to C, \qquad (p-1)\, m^{p-1} L(m) \to C, \qquad |\lambda_m m^p - C| \leq c_6 \, m^{-\vartheta} \quad (m \geq 1).$$

Proof. The Cramér root and the parameters it fixes are packaged by `Decay.ofC`. Given them, the invariant probability yields a positive sequence satisfying the cut balance above d, and the whole asymptotic follows from that alone `Decay.sharp_of_cutBal`: a contraction estimate on dyadic blocks shows the rescaled profile $\lambda_m m^p$ is Cauchy, its limit lies between the two constants of the two-sided decay bound and is therefore positive, and balancing the block index against the level gives the exponent ϑ. The statement about the tail is the same limit summed `Decay.tendsto_tail_sharp`: comparing $\sum_{j \geq m} j^{-p}$ with its integral turns a limit for the profile into one for the tail. □

The invariant measure of this graph decays like a power, not geometrically, and the exponent is the Cramér root of the doubling probability. That is the mechanism behind everything that follows: a power tail is heavy enough that the centred tail indicators below have vanishing Rayleigh quotient, which is what breaks the mixing hypothesis of the main text.

Theorem 42. *Let $c > 0$ and let (λ_j) satisfy the cut balance $\lambda_m\big(1 - \frac{c}{m+1}\big) = \sum_{j \in W(m)} \lambda_j \, \frac{c}{j+1}$ at every $m > d$, $W(m)$ the window of rungs whose doubling lands past the cut at m. If $\lambda_j \, j^p \to A$ for some $A > 0$ and $p \neq 0$, then p solves the Cramér equation $c\,(2^p - 1) = p$.*

Proof. Dividing the cut balance by λ_m and letting $m \to \infty$, the rescaled window sums converge to $(2^p - 1)/p$ `tendsto_windowV`, and the limit of the balance is the Cramér equation. □

No polynomial decay exponent other than a Cramér root is possible: the exponent is forced by the flux balance across a cut alone, before any existence or sharpness result. Only $s = 1$ is covered, the shape $\varepsilon(j) = c/(j+1)$ being what the cut balance encodes; for $c < 1$ the equation has exactly one root $p_* > 1$, which is the exponent of the invariant measure.

Theorem 43. *At $s = 1$ and $0 < c < 1$, let p_* be the Cramér root, (λ_j) a summable positive sequence with $\lambda_j \, j^{p_*} \to C > 0$, and $L(m) = \sum_{j \geq m} \lambda_j$. Then*

$$\frac{(p_* - 1)\, L(m)}{m\, \lambda_m} \;\to\; 1 :$$

the tail ratio grows linearly, $L(m)/\lambda_m = m\,(1 + o(1))/(p_* - 1)$.

Proof. The tail asymptotic $(p_* - 1)\, m^{p_* - 1} L(m) \to C$ `Decay.tendsto_tail_sharp`, divided by $\lambda_m m^{p_*} \to C$. □

The linear growth of the tail ratio is the mechanism that separates polynomial from geometric decay. A bounded Green operator keeps $L(m)/\lambda_m$ bounded — as it is under geometric damping, where the tail is comparable to the mass at the cut — and here it is not; that is what makes the Rayleigh quotient of the centred tail indicators vanish. Stated for any sequence with the sharp asymptotic, which the invariant measure has.

Theorem 44. *Let $\varepsilon = \varepsilon_{c,1}$ with $0 < c < 1$, and λ, λ' invariant probabilities of the loop closure with $\lambda_j = \lambda'_j$ for $1 \leq j \leq d$. Then $\lambda_m = \lambda'_m$ for every $m \geq 1$.*

Proof. At $s = 1$ the cut balance expresses λ_m, for every $m > d$, as a positive combination of the λ_j with $\lceil m/2 \rceil \leq j \leq m - 1$ — indices all smaller than m. Induction on m therefore determines the whole sequence from $\lambda_1, \ldots, \lambda_d$, which is the uniqueness statement `Decay.cutBal_unique`; the Cramér root that the surrounding parameters need is supplied by `Decay.ofC`. □

The boundary data of the target row therefore fix the whole invariant measure, and with it the constant of the asymptotic below. This is what makes that constant a property of the policy and the target row rather than a free parameter.

Theorem 45. *Let $\varepsilon = \varepsilon_{c,1}$ with $0 < c < 1$. There are a root $p > 1$ of ψ_c and $\nu_1, \ldots, \nu_d \geq 0$ with $\sum_j \nu_j > 0$ and $\nu_j = 0$ for $2j \leq d$ such that every invariant probability λ of the loop closure satisfies*

$$\lambda_m \, m^p \;\longrightarrow\; \sum_{j=1}^{d} \nu_j \, \lambda_j .$$

Proof. The sequences satisfying the cut balance above d form a real vector space, and restriction to the first d coordinates is an isomorphism onto $\mathbb{R}^d$ — that is the content of the determinacy above. Each map $\lambda \mapsto \lambda_m m^p$ is linear in those coordinates, so their limit is additive and positively homogeneous on the cone of positive sequences and extends to a linear functional on the whole space; its coefficients are the ν_j `Decay.constant_functional`. Positivity of $\sum_j \nu_j$ comes from the two-sided decay bound, and $\nu_j = 0$ for $2j \leq d$ because no window above d ever reads those indices. The root and the parameters it determines are again `Decay.ofC`. □

The constant of the tail asymptotic is thus a fixed non-negative linear functional of the boundary data, the same functional for every invariant probability. The vanishing of ν_j for $2j \le d$ says that the low half of the target row contributes nothing: only the states the doubling edge can reach from above carry weight.

B.5.3 THE BACKWARD LENGTH

Theorem 46. *Let $\varepsilon = \varepsilon_{c,1}$ with $0 < c < 1$, and $\bar\sigma^{(n)}$ the truncated expected backward-trajectory lengths of the loop closure. Then*

$$\bar\sigma := \sup_n \bar\sigma^{(n)} = \frac{\bar\jmath}{1-c}.$$

Proof. The linear function $h(x) = x/(1-c)$ is an exact solution of the one-step equation at $s = 1$: the substitution $\varepsilon(j) = c/(j+1)$ collapses the average of the state after one step to $c + j - 1$, and $1 + h(c + j - 1) = h(j)$. A non-negative solution of that equation dominates the truncated hitting-time expectations by induction, and the constant drift $c - 1$ makes the inequality an equality in the limit `sbar_iSup_eq`. Averaging over the target row, whose mean is $\bar\jmath$ and on which h is linear, gives the stated value. □

This is the quantity Morozov et al. (2025) require to be finite, and on this graph it is finite throughout the positive recurrent regime at $s = 1$. It is the first half of the separation this appendix is for: the condition of the main text that is repaired here.

Theorem 47. *Let $\varepsilon = \varepsilon_{c,1}$ with $0 < c < 1$, and $\bar\sigma^{(n)}$ the truncated expected backward-trajectory lengths of the loop closure or of any truncation. Then $\bar\sigma^{(n)} \le \bar\jmath/(1-c)$ for every n.*

Proof. The same linear supersolution as above serves here, and the one place the truncation differs is harmless: at a state whose doubling edge has been cut the decrement carries all the mass, and the one-step equation becomes an inequality in the right direction, because $1 \le 1/(1-c)$. The induction and the target-row average then run unchanged `sigmaBar_le`. □

The bound is uniform in the truncation, so the finiteness that repairs the condition of the main text is not an artefact of the infinite chain: it survives every finite approximation of it, which is what the measurements of a truncated system can be compared against.

B.5.4 SUMMABILITY AND THE GREEN OPERATOR

Theorem 48. *Let λ^K be an invariant probability of the truncation at some $K \ge d$, $P_\star$ the backward chain's density action on $L^2(\lambda^K)$, Π the mean projection, and $\widehat\beta_n = \|P_\star^n - \Pi\|$. If $\sum_n \widehat\beta_n < \infty$, then*

$$S := (\mathrm{Id} - P_\star + \Pi)^{-1} - \Pi = \sum_{n\ge 0}(P_\star^n - \Pi) \qquad \text{and} \qquad \widehat B = \|S\| \le \sum_{n\ge 0}\widehat\beta_n.$$

Proof. Summability makes the partial sums of $\sum(P_\star^n - \Pi)$ Cauchy in operator norm; their limit satisfies the resolvent identities, so it is the inverse minus Π `Stat.sub_piL2_eq_tsum`, and the triangle inequality bounds its norm by the sum. □

The two constants of the main text, ordered one way: where the mixing sum converges, the Green operator is its sum and its norm is at most the sum. The statement is vacuous exactly where $\sum \widehat\beta_n = +\infty$, which on the polynomial family is everywhere; on a finite truncation both constants are finite, and they differ in general, the reverse inequality holding nowhere in general.

Theorem 49. *Let $\varepsilon = \varepsilon_{c,s}$ with $c > 0$, $s \ge 0$, and λ an invariant probability of the loop closure. On $L^2(\lambda)$ no bounded S satisfies $S(\mathrm{Id} - P_\star) = \mathrm{Id} - \Pi$, the series $\sum_{n\ge 0}(P_\star^n - \Pi)$ does not converge in operator norm, and*

$$\sum_{n\ge 0}\|P_\star^n - \Pi\|_{L^2(\lambda)} = +\infty.$$

Proof. All three are the same fact at $p = 2$. The family satisfies the growth condition `GrowthCond`, discharged as before `growthCond_of_family`; the vanishing infimum of the Rayleigh quotient then excludes a bounded left inverse `Stat.no_diffusionOp`. The other two follow from that one: a norm-convergent series would furnish such an inverse, since its sum U satisfies $(\mathrm{Id} - P_\star)U = U(\mathrm{Id} - P_\star) = \mathrm{Id} - \Pi$ by telescoping `Stat.not_tendsto_partialSum`; and a summable $\sum_n \widehat{\beta}_n$ would make that series absolutely convergent, the norms being the $\widehat{\beta}_n$ `Stat.not_summable_betaHat`. □

This is the sentence that separates the two conditions. The summable-mixing hypothesis of the main text's L^2 universality theorem, and of its frozen detailed-balance theorem in the flow-matching instance — where the constant is exactly the state-space mixing sum $\sum_n \widehat{\beta}_n$ — is repaired by no positive recurrent member of this family, while the backward length above is finite throughout. The divergence is a statement about the flow-matching instance; the detailed-balance instance, whose coefficients are the edge-lifted ones, is not covered.

Theorem 50. *Let $\varepsilon = \varepsilon_{c,s}$ with $c > 0$, $s \geq 0$, λ an invariant probability of the loop closure, f_m the centred tail indicators, and $1 \leq p < \infty$. For every $t > 0$ some $m > d$ satisfies*

$$\|(\mathrm{Id} - P_\star) f_m\|_{L^p(\lambda)}^p \leq t \, \|f_m\|_{L^p(\lambda)}^p,$$

and no constant B satisfies $\|f\|_{L^p(\lambda)} \leq B \, \|(\mathrm{Id} - P_\star) f\|_{L^p(\lambda)}$ for every bounded f with $\sum_x \lambda_x f(x) = 0$.

Proof. The family satisfies the growth condition $\sum_{i<D} \log(1/\varepsilon(2^i)) = o(2^D)$ `GrowthCond`, because for $\varepsilon_{c,s}$ that sum is $O(D^2)$ `growthCond_of_family`.

Given the growth condition, the ratio $L(m)/\lambda_m$ is unbounded: were it bounded by Q, the tail would decay geometrically at rate $e^{-1/Q}$, while iterating the doubling inequality along the powers of two forces $L(2^D) \geq \lambda_1 \prod_{i<D} \varepsilon(2^i)$, and the growth condition makes the second bound incompatible with the first. Choosing cuts along that unbounded ratio makes the centred tail indicator's defect small against its own mass `Stat.exists_small_mass_ratio`, which is the first clause; the second is that same infimum read as the non-existence of a bound `Stat.no_bounded_inverse`. □

The second clause is the failure of a bounded diffusion operator at every finite exponent. The main text's convergence bounds are stated with a constant of exactly this shape; on this graph no such constant exists, at any p, at any positive recurrent member of the family. The first clause is the sharper fact behind it: the infimum of the Rayleigh quotient is not merely unattained but zero.

Theorem 51. *Let $\varepsilon = \varepsilon_{c,s}$ with $c > 0$, and λ an invariant probability of the loop closure. No constant B satisfies*

$$\sup |f| \;\leq\; B \, \sup \big| (\mathrm{Id} - P_\star) f \big|$$

for every bounded f with $\sum_x \lambda_x f(x) = 0$.

Proof. The witnesses are clipped ramps: r_N is the ladder height cut off at N. Its flow-matching defect is bounded independently of N — at a state below the clip the defect is $1 - \varepsilon(j)(j+1)$, at a state above it the defect vanishes, and at the sink it is $-\bar{\jmath}$ — while centring leaves the defect unchanged and leaves the function attaining at least $N/2$. So M/D grows without bound `Stat.no_bounded_inverse_infty`.

That the defect is bounded uses $\gamma = \sup_j (j+1)\varepsilon(j) < \infty$, which for this family holds exactly when $s \geq 1$, with $\gamma = c\, 2^{1-s}$ `gamma_family`. Below $s = 1$ there is no invariant probability to quantify over `isEmpty_stat_of_family_lt_one`, and the statement holds for that reason instead. □

The failure therefore reaches the uniform norm as well, where the main text's stable bounds are most often read. No hypothesis on s is needed: above $s = 1$ the clipped ramps do the work, and below it there is no invariant probability at all, so the claim is vacuous exactly where the objects it quantifies over do not exist.

Theorem 52. *Let $\varepsilon = \varepsilon_{c,1}$ with $0 < c < 1$, λ an invariant probability of the loop closure, and f_m the centred tail indicators. There are a root $p > 1$ of ψ_c, $0 < c_1 \le c_2$ and $m_2 > d$ such that*

$$c_1\, m\, \big\|(\mathrm{Id} - P_\star) f_m\big\|^2_{L^2(\lambda)} \;\le\; 4c_2(p-1)\, \|f_m\|^2_{L^2(\lambda)} \qquad (m \ge m_2).$$

Proof. The two-sided decay bound $c_1 j^{-p} \le \lambda_j \le c_2 j^{-p}$ holds unconditionally at this level `Decay.decay_two_sided_of_cutBal`, run at the level m_0 at which the descent-weight window is controlled `Decay.m0`, with the root supplied by `Decay.ofC`. Summing that bound over $j \ge m$ gives $L(m)/\lambda_m \ge c_1 m/(c_2(p-1))$, so the per-cut inequality — the defect of f_m is carried by the cut and the window alone, and the cut balance collapses the window — reads as the display once the mass below the cut is bounded away from zero. The cut m_2 past which that last condition holds exists because $L(m) \to 0$ `Stat.exists_rayleigh_family`. □

Unboundedness above says only that the Rayleigh quotient has infimum zero; this says how fast, along an explicit family of trial functions. Squared and rearranged it is $\|f_m\| \ge c_7\sqrt{m}\,\|(\mathrm{Id} - P_\star) f_m\|$. It is not a lower bound on the diffusion constant of the infinite chain, which is $+\infty$ there; it is a rate along the f_m, and it is the truncation below that converts a rate into a bound on a finite constant.

B.5.5 Geometric damping: a finite constant

Theorem 53. *For a geometrically damped doubling probability $\varepsilon(j) = a\rho^j$ with $a > 0$ and $0 < \rho < 1$, the growth condition $\sum_{i<D} \log\big(1/\varepsilon(2^i)\big) = o(2^D)$ of the unboundedness theorem fails: the left side equals $D\log(1/a) + (2^D - 1)\log(1/\rho)$, which is of order 2^D.*

Proof. The identity $\sum_{i<D} \log\big(1/(a\rho^{2^i})\big) = D\log(1/a) + (2^D - 1)\log(1/\rho)$ `geometric_sum`, whose right side is not $o(2^D)$ since $\log(1/\rho) > 0$. □

The unboundedness theorem indicts the polynomial family and not the graph: its hypothesis is a growth condition on the doubling probability, and geometric damping does not meet it. What geometric damping gives instead is the next statement.

Theorem 54. *For a geometrically damped doubling probability $\varepsilon(j) = a\rho^j$ on the rungs, with $a > 0$ and $0 < \rho < 1$: the loop closure has an invariant probability, and at every invariant probability λ the operator $\mathrm{Id} - P_\star + \Pi$ is invertible on $L^2(\lambda)$, its inverse R satisfying*

$$(\mathrm{Id} - P_\star)(R - \Pi) = (R - \Pi)(\mathrm{Id} - P_\star) = \mathrm{Id} - \Pi, \qquad \Pi\,(R - \Pi) = (R - \Pi)\,\Pi = 0;$$

so the Green operator $S = R - \Pi$ exists, and $\widehat{B} = \|S\| < \infty$.

Proof. The drift of the chain at rung m is $(m+1)\,a\rho^m$ times a bounded factor, which tends to 0, so the loop closure is positive recurrent `geometric_exists_stat`. Geometric damping satisfies the ratio condition $\varepsilon(i+1) \le \theta\,(1 - \varepsilon(i+1))\,\varepsilon(i)$ past some rung with $\theta = (1+\rho)/2$; under that condition the tail is comparable to the mass at the cut, $L(k) \le B\lambda_k$ `Stat.tail_bound_of_ratio`, a Poincaré inequality follows, and the open mapping theorem turns the coercivity into a bounded inverse `Stat.exists_diffusionOp_of_ratio`. □

Exponential damping restores what the polynomial family loses: a bounded Green operator, hence a finite coercivity constant and the local rate of the main text, on the very same graph. The obstruction is therefore a property of the backward policy and not of the graph. Only existence is stated: no numeric bound on $\widehat{B}$ is claimed, and nothing is said about the mixing sum $\sum \widehat{\beta}_n$, whose convergence is neither used nor asserted. Behind the boundedness is a geometric tail — the mass beyond a rung is comparable to the mass at the rung — which is the opposite of the linear tail ratio of the polynomial family.

B.5.6 FLOW MATCHING

Theorem 55. *Let $\varepsilon = \varepsilon_{c,s}$ with $c > 0$, $s \geq 0$, and λ an invariant probability of the loop closure. Then*

$$(\mathrm{Id} - P_\star)L^2(\lambda) \subsetneq \overline{(\mathrm{Id} - P_\star)L^2(\lambda)} = \ker \Pi,$$

and there are probability densities $f_{\mathrm{init}}, f_{\mathrm{term}} \in L^2(\lambda)$ for which $(\mathrm{Id} - P_\star)f = f_{\mathrm{init}} - f_{\mathrm{term}}$ has no solution $f \in L^2(\lambda)$.

Proof. Write A for the restriction of $\mathrm{Id} - P_\star$ to $\ker \Pi$. It is injective, its kernel being the mean-zero fixed points of $P_\star$, which irreducibility reduces to the constants and hence to zero; its adjoint inside $\ker \Pi$ is the restriction of $\mathrm{Id} - P$, injective for the same reason, so the range of A is dense. It is not surjective: a surjective bounded injection of a Hilbert space is boundedly invertible by the open mapping theorem, and that contradicts the vanishing infimum established above. Picking $\psi \in \ker \Pi$ outside the range and splitting it into positive and negative parts produces the two densities `doubling_unsolvable`. The growth condition the argument needs is discharged for the family as before `growthCond_of_family`. □

What this refutes is *exact* L^2 flow matching on the infinite chain: a pair of L^2 densities the loop closure cannot balance at all. It does not refute the weak universality of the main text, whose conclusion is a defect infimum of zero — density of the range is exactly what the first clause asserts, and whether the weak form fails here is not decided.

B.5.7 THE TRUNCATION

Theorem 56. *Let $K \geq d$ be even. The truncation at K is irreducible on its $K + 2$ states, and every invariant measure of it is positive at each of them.*

Proof. Only the climbing half differs from the infinite chain, because the doubling edges above $K/2$ have been cut. Let 2^r be the largest power of two at most K: the edges $2^i \to 2^{i+1}$ all survive for $i < r$, so the ladder state 1 reaches 2^r and, by decrements, every state below it. Above 2^r an even n is reached by doubling from $n/2 \leq K/2 < 2^r$, and an odd n is not the cap — which is even — so $n + 1$ is even and in range and one decrement finishes `reach_all_some`. Positivity is the same general fact as on the loop closure `PreStat.pos_of_irreducible`. □

The truncation is what a numerical experiment can actually run, and this is what makes it a legitimate finite approximation: the same qualitative structure as the infinite chain, on a finite state space where the diffusion constant is finite and can be measured.

Theorem 57. *Let $K \geq d$ be even. The truncation at K has exactly one invariant probability λ^K, and there is a bounded S on $L^2(\lambda^K)$ with*

$$S(\mathrm{Id} - P_\star) = (\mathrm{Id} - P_\star)S = \mathrm{Id} - \Pi, \qquad \Pi S = S\Pi = 0;$$

so $\widehat{B}_K := \|S\| < \infty$.

Proof. Existence and uniqueness of the invariant probability are those of a finite irreducible chain `exists_stat`, `stat_unique`. On a finite state space $\mathrm{Id} - P_\star + \Pi$ is then injective — a vector it kills has zero mean and is fixed by $P_\star$, hence is a constant of zero mean — and injective is invertible in finite dimension; writing R for the inverse and $S = R - \Pi$, the relations $\Pi P_\star = P_\star \Pi = \Pi$ and $\Pi^2 = \Pi$ give the four identities `Stat.exists_bhat`. Boundedness is automatic, the space being finite-dimensional. □

So the object that fails to exist on the infinite chain does exist on every truncation of it. Nothing diverges at a finite cap; what the next result shows is that the constant grows without bound as the cap does, which is how an infinite constant is visible to a finite computation.

Theorem 58. *Let $\varepsilon = \varepsilon_{c,1}$ with $0 < c < 1$, and λ^K an invariant probability of the truncation at K. There are $c_8 > 0$ and K_0 such that, for $K \geq K_0$, every bounded S on $L^2(\lambda^K)$ with $S(\mathrm{Id} - P_\star) = \mathrm{Id} - \Pi$ satisfies*

$$\|S\|^2 \geq c_8 K.$$

Proof. The argument is the per-cut inequality applied at a cut of order K. The decay bound transfers to the truncation with constants that do not depend on the cap `Stat.decayK`: extending λ^K above the cap gives a positive sequence satisfying the cut balance below $K/2$, so the block estimate applies at the level m_0 `Decay.m0`, and the ratio of the two constants is controlled by chaining the cut-balance inequalities across one block. The descent weight at that level is two-sided `Decay.descOne_two_sided` with the escape constant `Decay.c5`, and the Cramér root is again `Decay.ofC`.

Taking the cut at $\lfloor K/4 \rfloor$ then makes $L_K(m)/\lambda_m^K$ of order K, while the mass below the cut is bounded away from zero uniformly in K — on a finite irreducible chain the mass at the source is the reciprocal of the expected return time, which is $2 + \bar{\sigma}_K \le 2 + \bar{\jmath}/(1-c)$. Feeding both into the per-cut inequality gives the displayed bound `Stat.sqrtK`. □

This is the quantitative form of the separation, and the one a measurement can be compared against: the diffusion constant of the truncation is finite but grows at least like $\sqrt{K}$, so it diverges as the truncation is lifted, while the backward length stays at $\bar{\jmath}/(1-c)$ throughout. The bound holds for *every* left inverse, so it is a property of the chain and not of a particular construction.

B.6 What is not established here

Three claims that would naturally accompany the results above are absent, and in each case for the same reason: they are properties of the backward chain as a stochastic process, and no such process is constructed.

- *Transience and null recurrence as properties of a process.* The labels of the phase table are classes of the least solutions of the chain's linear systems — the return probabilities and the expected return times — and not of a constructed process; on a countable irreducible chain the two readings agree, but the second is not what is stated. The case $s = 1$, $c = 1/\ln 2$ is only excluded from positive recurrence at every state and is not resolved further.
- *An infinite expected hitting time.* At $s = 1$ with $1 \le c < 2$ the absence of an invariant probability is established; the accompanying statement that $\mathbb{E}(\sigma \mid X_0 = j) = +\infty$ at every ladder state is not.
- *The trajectory reading of the diffusion operator.* The operator excluded at $p = 2$ is characterised by the resolvent identities, not as the sum $\sum_{n \ge 0} \big(\mathbb{E}(\theta(X_n) \mid X_0 = x) - \lambda(\theta)\big)$ along backward trajectories. That the two agree needs the chain.

Nothing above depends on these. They are named so that a reader who expects the vocabulary of recurrence classification knows where it has been replaced by the weaker statement that is actually proved.

B.7 The development

Each result above is a declaration of a Lean 4 development checked against Mathlib `v4.31.0`. Table 1 gives the correspondence. Every one of them has been verified to rest on no axiom beyond Lean's own `propext`, `Classical.choice` and `Quot.sound`, and the development contains no unproved goal.

C The diffusion picture in full

Section 5 kept the two statements that carry the story: the gradient of a balance loss linearizes to the backward policy's diffusion applied twice, and near balance training converges at a rate the Green operator sets. The other results of the dynamics group of Figure 1 are stated here in full, each with its computed badge; how each Lean statement reads the paper's, and the sketch of its proof, are in Sections B.1, B.2 and B.3.

C.1 THE GRADIENT

Theorem 59. *certified [L15] Let T be a Markov kernel on $\mathcal{S}$, μ a finite measure with ratio $r = d(\mu T)/d\mu \in [a, b]$, $0 < a \le b$, $\nu \ll \mu$ finite with $d\nu/d\mu \in L^\infty(\mu)$, and $g \in C^1(0, \infty)$. For every bounded u, with $v := d((u\mu)T)/d\mu$,*

$$\frac{d}{ds}\Big|_{s=0} \mathcal{L}_{g,\nu}\big((1 + su)\mu\big) = \int g'(r)\,(v - ru)\,d\nu.$$

If moreover $\mathcal{S}$ is standard Borel and $\lambda T = \lambda$ for a finite $\lambda \neq 0$ equivalent to μ, the gradient has λ-density $T\psi - r\psi$, with $\psi := g'(r)\,d\nu/d\mu$.

Differentiability is certified in the directional sense, along every essentially bounded direction, which is the sense the full theorem uses; the uniform sense needs a locally Lipschitz derivative. The reversal of the kernel exchanges the backward policy with the forward policy of the balanced flow, so the gradient descent of Section 5 moves the flow along the policy that training holds fixed.

C.2 THE CLOCK NEAR BALANCE, FOR DETAILED AND TRAJECTORY BALANCE

The detailed-balance loss is the flow-matching loss of the chain lifted to edges, whose Markov property, invariance and positivity follow from the backward policy; a coercivity constant of the lift is one more than one of the chain on states, which is the $1 + \widehat{B}$ of Section 5.

Theorem 60. *certified [L17] Let $\pi_{\leftarrow}$ be a backward policy on a finite state space with invariant probability $\lambda > 0$, E its edge set, K_2 its edge lift, and $\lambda_2(s, s') := \pi_{\leftarrow}(s' \to s)\,\lambda(s')$. Under the hypotheses of Theorem 29 on g and a, its conclusions hold for the detailed-balance loss, that is, for (K_2, λ_2) in place of (T, λ), with $w \ge w_{\min} > 0$ on E and $\widehat{B} \ge 1$ a coercivity constant of K_2 on $L^2(\lambda_2)$.*

For trajectory balance the development proves a comparison of Hessians and nothing more: a step cap for detailed balance, divided by the factor below, stays valid for trajectory balance, and no rate is claimed. The bound is vacuous where the lifted mixing sum is infinite, as the printed one is.

Theorem 61. *certified [L25] Let $\pi_{\leftarrow}$ be a backward policy on a finite state space with invariant probability $\lambda > 0$, K_2 and λ_2 its edge lift and edge measure, $g = (\log x)^2$, $H_{\mathrm{DB}} = g''(1)\,A^\dagger A$ and $H_{\mathrm{TB}} = \ell\, g''(1)\,A^\dagger \Theta_\ell A$ the Hessians at balance of the detailed- and trajectory-balance losses on the edges, Θ_ℓ the window operator of length ℓ, and $B' \ge \sum_{m=1}^{N} \|K_2^m - \Pi_2\|_{L^2(\lambda_2)}$ for every N. For every h on the edges,*

$$\langle h, H_{\mathrm{TB}} h\rangle \le \ell\,(1 + 2B')\,\langle h, H_{\mathrm{DB}} h\rangle.$$

C.3 THE CLOCK FAR FROM BALANCE

Theorem 62. *certified [L18] Let G be a finite path-connected marked graph, $\pi_{\leftarrow}$ a backward policy positive on its edges, λ the invariant probability of the loop closure, N, $\bar\sigma$ and σ_* the visit counts, the mean backward length and the longest expected exit time, $g = (\log x)^2$, $\nu = w\lambda$ with $w \ge w_{\min} > 0$, $u_0 > 0$, and*

$$\varrho_\sigma := \frac{2\, w_{\min} \min N}{\sigma_*^2\,(2 + \bar\sigma)}.$$

Then $\lambda = N/(2 + \bar\sigma)$, and there are explicit $\kappa, \gamma_* > 0$ such that: (a) the gradient flow μ_t of $\mathcal{L}_{g,\nu}$ from $u_0\lambda$ exists, is the unique positive solution, and

$$\mathcal{L}(\mu_t) \le \big(\mathcal{L}(\mu_0)^{-1} + \kappa^2 t\big)^{-1}, \qquad \mu_t \longrightarrow \|u_0\|_{L^2(\lambda)}\,\lambda,$$

exponentially at the rate $\varrho_\sigma/(2m^2)$ after an explicit time, m the mass then; (b) for $0 < \gamma \le \gamma_*$ the gradient descent u_k from u_0 is bounded below, converges to a balanced flow, satisfies $\mathcal{L}(u_k) \le \big(\mathcal{L}(u_0)^{-1} + k\gamma\kappa^2/4\big)^{-1}$, and after an explicit step k_0

$$\|u_{k+1} - \Pi u_{k+1}\| \le \Big(1 - \frac{\gamma\varrho_\sigma}{4m^2}\Big)\|u_k - \Pi u_k\|, \qquad m \text{ the mass at } k_0.$$

The constants see $(G, \pi_{\leftarrow})$ only through σ_*, $\bar\sigma$ and $\min N$; γ_* also through the least positive transition probability and $|V|$.

The rate ϱ_σ is the local rate of Section 5 with the hitting-time constant of Section 2 in place of $\widehat{B}$, and the crossover time and the step cap are explicit in the same quantities. The rate is a guarantee and not a prediction, since the exact constant can be far below the hitting-time one. The gradient flow is constructed, in the class C^1, and is unique among positive solutions; the descent's step is chosen from the initialization.

C.4 Freezing is not stability

Brunswic et al. (2024) call a loss stable when adding a 0-flow never lowers it, and show that the ratio losses are not. Freezing the backward policy does not change that:

Theorem 63. *certified* [L19] *Let G be a finite path-connected marked graph, $\pi_{\leftarrow}$ a backward policy positive on its edges, λ an invariant probability of the loop closure, F^* its balanced edge flow, $\gamma \neq 0$ a non-negative circulation with no inflow at the source or the sink, g with $g(1) = 0 < g(x)$ for $x \neq 1$, $\nu_{\mathrm{T}} > 0$ on the states, and $\varepsilon > 0$ with $2\varepsilon\gamma \leq F^*$. With $F := F^* - \varepsilon\gamma$ and $F_0 := \varepsilon\gamma$,*

$$0 \leq F_0 \leq F, \qquad \mathcal{L}_{g,\nu_{\mathrm{T}}}(F + F_0) = 0 < \mathcal{L}_{g,\nu_{\mathrm{T}}}(F).$$

Theorem 64. *certified* [L20] *Let G, $\pi_{\leftarrow}$, λ, γ, g, ε, F and F_0 be as in Theorem 26, and $\widehat{\nu}$ positive on the edges of the loop closure and zero off them. Then $0 \leq F_0 \leq F$ and*

$$\mathcal{L}_{\mathrm{DB},g,\widehat{\nu}}(F + F_0) = 0 < \mathcal{L}_{\mathrm{DB},g,\widehat{\nu}}(F).$$

The circulation is carried as the properties the proofs use — non-negative, non-zero, no inflow at the source or the sink — of which the indicator of a directed cycle avoiding both marks is an instance; the loss is the state-space loss read on the inflow of an edge flow, and the same pair witnesses both halves. The obstruction is the mechanism of training. At a fixed training measure the loss is invariant under rescaling the flow, and the gradient is orthogonal to the radial direction:

Theorem 65. *certified* [L26] *Let T be a Markov kernel on a standard Borel space $\mathcal{S}$ with $\lambda T = \lambda$, λ finite, μ a flow equivalent to λ meeting the first-variation hypotheses for ν, $u = d\mu/d\lambda$, and g differentiable on $(0, \infty)$. Then $\mathcal{L}_{g,\nu}(c\mu) = \mathcal{L}_{g,\nu}(\mu)$ for $c > 0$, and every λ-density D of the gradient satisfies $\langle D, u\rangle_{L^2(\lambda)} = 0$. If T is ergodic for $\lambda \neq 0$,*

$$\mu(\mathcal{S}) \leq \sqrt{\lambda(\mathcal{S})}\, \|u\|_{L^2(\lambda)},$$

with equality iff μ is balanced.

Training therefore moves on a sphere of $L^2(\lambda)$, and on that sphere the mass is largest exactly at the balanced flow. Whether the gradient climbs towards it is decided by the generator:

Theorem 66. *certified* [L21] *Let T be a Markov kernel on $\mathcal{S}$ with $\lambda T = \lambda$, λ finite, and $\mu \ll \lambda$ a finite flow of ratio r. (1) If g is a strictly unimodal generator and ν, with $d\nu/d\mu > 0$, meets the first-variation hypotheses, every λ-density D of the gradient of $\mathcal{L}_{g,\nu}$ at μ satisfies*

$$\int D\, d\lambda = \int g'(r)\,(1 - r)\,\frac{d\nu}{d\mu}\, d\lambda \;\leq\; 0,$$

with equality iff μ is balanced; and μ is balanced if it is a critical point of $\mathcal{L}_{g,\nu}$. (2) If g' is continuous on $(0, \infty)$ and vanishes on two bands, one on each side of 1, that contain the values of r, then μ is a critical point of $\mathcal{L}_{g,\nu}$ for every finite $\nu \ll \mu$.

For a strictly unimodal generator — $(\log x)^2$ and $(x-1)^2$ among them — the gradient never lowers the mass and every critical point is balanced, so gaining circulation is how training reaches balance; on a finite state space the convergence is global, as the preceding subsection states for $(\log x)^2$. For a generator whose derivative vanishes on bands on either side of 1, every flow whose ratios lie in those bands is a critical point, for every training measure, and some such generators freeze an open set of flows a fixed distance from balance already on the two-state chain with the fastest mixing there is. The freezing is invisible to the linearization, and its mechanism is the vanishing derivative, not the curvature.

result	declaration	named results	declarations
The sampler, the certificates and the constant			
11	`SamplingGeneral.MFlow.sampling_theorem`	6	55
12	`SamplingGeneral.MFlow.negative_control_general`	8	78
13	`SamplingGeneral.stable_bound_general`	39	128
14	`SamplingGeneral.il_first_bullet_general`	37	137
15	`ILBoundFull.il_tv_le_sqrt`	0	0
16	`lem_sigma_mixing`	38	97
17	`weaklyUniversal_of_kernel`	31	76
18	`stronglyUniversal_of_kernel_top`	30	69
19	`General.FirstVariation.theo_first_variation`	20	64
The frozen backward policy on a finite graph			
20	`BackwardPolicy.lam_eq_visits_div`	15	60
21	`BackwardPolicy.coercivity_morozov`	24	104
22	`BackwardPolicy.universality_graphs_three`	25	61
23	`SamplerWiring.universality_graphs_sampler`	46	128
24	`SamplerWiring.cycle_no_model_free_bound_sampler`	35	149
25	`SamplerWiring.no_bound_divergence_minimax`	40	160
26	`BackwardPolicy.frozen_unstable_full_fm`	31	75
27	`BackwardPolicy.frozen_unstable_full_db`	45	99
Gradient descent diffuses the flow			
28	`GdDiffusionGeneral.theo_gd_diffusion_full`	56	155
29	`local_convergence_full_paper`	134	242
30	`local_convergence_full_DB_paper`	149	267
31	`TBHessian.HTB_le`	44	113
32	`DiscreteGlobal.training_speed_full_complete`	303	560
33	`General.Bridge.no_distant_equilibrium_two_bridge`	33	111
34	`General.Bridge.global_dichotomy_full_bridge`	31	114
35	`RemarksA.graphs_vs_L2_periodic`	17	85
The acyclic comparison			
36	`exists_star_witness`	3	28
37	`no_state_space_free_bound`	4	32
The doubling graph			
38	`main_irreducible`	3	67
39	`main_phase`	25	162
40	`main_phase_classes`	66	371
41	`main_invariant_measure`	98	366
42	`exponent_is_cramer_root`	2	17
43	`Decay.tendsto_tail_ratio_sharp`	7	34
44	`main_constant_determined`	5	98
45	`main_constant_functional`	104	377
46	`main_sigmaBar_eq`	10	83
47	`main_sigmaBar_le`	4	39
48	`Stat.inverse_sub_piL2_eq_tsum`	53	194
49	`main_unbounded_two`	60	267
50	`main_unbounded`	16	128
51	`main_unbounded_infty`	11	116
52	`main_rate`	71	302
53	`not_growthCond_geometric`	2	4
54	`main_geometric`	58	271
55	`main_unsolvable`	44	203
56	`main_truncation_irreducible`	4	69
57	`main_truncation_bhat`	37	183
58	`main_truncation_sqrtK`	96	393

Table 1: The results of this appendix as declarations of the Lean development, in the namespaces `GFNBounds.Core`, `GFNBounds.Graph`, `GFNBounds.Balance`, `GFNBounds.Silva` and `GFNBounds.Doubling`. The last two columns count what each proof rests on, unfolded to the bottom of the development: the declarations it reaches, and how many of those are results the development names rather than steps internal to one.

D EXPERIMENTAL GROUNDING OF THE LEAN THEOREMS

The Lean certificate settles whether each statement of this paper holds; no measurement can add to that. What a measurement can add is of another kind: the size of a constant a theorem leaves unspecified, a prediction the theorem makes and an experiment confirms, the place where a certified bound becomes vacuous or its hypothesis fails, or a view of the mechanism a proof constructs. Table 2 names the six forms this takes, and fixes how the body may phrase each; a *check*, which evaluates a certified inequality on instances, is the weakest of them, since a violation would indict the code or the reading of a Lean statement as the paper's, never the proof.

The statements are grouped as in the map of certified results, each with its certification badge and its Lean declarations. Every number below is computed from the experiment results at the pinned commit of the experiment sub-repository, never typed; a measurement that has been run but is not in the pin is listed as pending, without its numbers, and a statement no measurement bears on is listed as such. Discrepancies between a result file and the lab record that are not yet resolved are flagged on the entries they affect and stated at the end.

Table 2: The six forms a measurement takes when it bears on a certified statement.

kind	what it adds to a certified statement	how the body may phrase it
CHECK	the certified inequality or identity, evaluated on instances; a violation would indict the code or the reading of the Lean statement as the paper's, never the proof	holds on every instance, with the count
SIZE	a number for a constant the theorem leaves unspecified: an order, a tightness, a slack	measured at, with its range
PREDICTION	the theorem's quantity predicts an independent measurement	predicts to within
LIMIT	where the statement is vacuous or conservative, or where its hypothesis fails	valid but conservative by
BEYOND	the statement's conclusion tested outside its hypotheses	evidence, not proof
MECHANISM	the proof's construction, made visible	a figure, no count

Of the 26 boxes of the reading map, 17 carry a measurement read from the pinned results, 0 carry only measurements not yet in the pin, and 9 carry none; the boxes without a measurement are listed as such.

IMPORTED

Sampling theorem certified lean [L1]

- CHECK (exp19a) On 130 balanced flows computed exactly, the expected stopping time equals $1 + F_{\mathrm{out}}(\mathcal{S})/F_{\mathrm{init}}(\mathcal{S})$: the ratio of the two differs from one by at most $7.7 \cdot 10^{-11}$: the bound is attained.

Forcing balance (Brunswic et al.) certified lean [L2] No measurement bears on this statement yet.

THE FROZEN BACKWARD POLICY AND ITS CONSTANT

The λ-reversal certified lean [L27] No measurement bears on this statement yet.

$\widehat{B}$ through mixing certified lean [L8] [L28]

- LIMIT (exp3) On the 17 chains of the rate study where the mixing sum B is finite, it exceeds $\widehat{B}$ by a factor in 1.01–65.3: a valid bound, loosest on the transposition graphs.

$\widehat{B}$ through hitting times certified lean [L29] [L9] [L30]

- CHECK (exp3, exp21a) The occupation identity $\lambda(x) = N(x)/(2 + \bar{\sigma})$ holds to $5.4 \cdot 10^{-15}$ on the 20 chains of the rate study, and on 20 cyclic sorting graphs the return time to s_0 exceeds $\bar{\sigma}$ by 2, its value at s_0, to within $5.5 \cdot 10^{-9}$.

- LIMIT (`exp3`) The hitting-time bound $\widehat{B}_\sigma$ holds on every chain and exceeds $\widehat{B}$ by a factor in 13.6–$3.14 \cdot 10^4$; on product-replacement graphs it degrades with the state space while $\widehat{B}$ stays in 3.8–5.6.

Periodic chains certified lean [L31]

- CHECK (`exp3`) On the 3 leveled graphs the mixing sum is infinite, yet $\widehat{B}$ is at most 1.46, and gradient descent converges there at a rate of at least 0.944, against at most 0.583 on every other chain of the study (Figure 7).

DB is a lifted balance loss certified lean [L32] No measurement bears on this statement yet.

What a small loss certifies

Stable losses certify certified lean [L3]

- CHECK (`exp1`, `exp2`, `exp10`, `exp11`) Over 25,668 (flow, ν_T, q) rows of perturbed tabular flows, on permutation, hypergrid, Cayley, product-replacement and torus state spaces of 6 to 386,880 states, the certificate is violated 0 times.
- SIZE (`exp2`, `exp10`) open: D5 At $q = 1$ under a uniform ν_T the median of $C\mathcal{L}/\mathrm{TV}$ lies in 6–9.08 across state spaces of up to 720 states, and is 6.61 on product-replacement graphs of up to 386,880 states: the bound carries no factor of the state space.
- LIMIT (`exp1`) open: D5 The constant is set by the coverage of ν_B by ν_T: with the visit measure of the flow as ν_T instead of the uniform one, the median of $C\mathcal{L}/\mathrm{TV}$ at $q = 1$ rises from 6.7 to 5,155.
- SIZE (`exp20`) open: D2, D5 On the doubling ladder, where every mixing constant is infinite, the certificate is violated 0 times in 9,216 rows, and where it is informative the median of $C\mathcal{L}/\mathrm{TV}$ at $q = 1$ lies in 6–6.03 across truncations.
- LIMIT (`exp12`) open: D5 Along the training of a parameterized flow the certificate holds at all 24 checkpoints, but that flow never learns (its TV stays in 0.966–0.974) and the certificate is vacuous throughout: its smallest value at $q = 1$ is 9.44.
- CHECK (`exp23`) Along 576 training runs on the acyclic and the wrapping hypergrid, with $\nu_B = \nu_T$ on-policy or uniform, $q \in \{1, 2\}$, exact or sampled losses, every learning rate of the sweep, diverged runs included, and state spaces of 64 to 1,024 states, each configuration at its longest budget, the certificate $C\mathcal{L}$ with $C = 2/\kappa(\mathcal{S})$ bounds the certified sampler's TV at each of the 281,134 evaluations of those runs: it is violated 0 times.
- SIZE (`exp23`) At the end of training at $q = 1$, each configuration taken at its chosen learning rate and as the median over seeds, the certificate exceeds the true TV by a factor of 3.2–6.2: once training has converged, the certificate of a trained flow is informative, and conservative by a small constant factor.
- LIMIT (`exp23`) With $\nu_B = \nu_T$ the constant is the same for on-policy and uniform training, and at $q = 1$ so is the loss: exact training certifies $\mathrm{TV} \le \varepsilon$ at the same step in both regimes, the settling steps differing by 0, with ε at 0.05. At $q = 2$ the on-policy loss $\sum_x D(x)^2/\rho(x)$ weighs rarely visited states more. On the wrapping grid, where the least-visited state keeps at least 0.61 of its uniform share, on-policy training certifies up to 2.4 times sooner; on the acyclic grid, where that share falls to 0.034, it certifies up to 6.7 times later.
- LIMIT (`exp23`) Writing the same loss in masses, ν_B the counting measure, while training on on-policy data multiplies the certificate at the end of exact training by up to 92.7: the constant then carries $\|d\nu_B/d\nu_T\|_\infty$, the inverse of the least visitation, which is the coverage clause of the certificate at work during training.
- LIMIT (`exp23`) With minibatches of sampled states the true error settles below ε sooner on-policy at $q = 2$, by up to 25 times, yet 0 of the 16 sampled configurations on grids larger than 64 states certify $\mathrm{TV} \le \varepsilon$ within 100,000 steps, in either regime: the certificate needs the defect small on every state, including those a minibatch rarely draws.

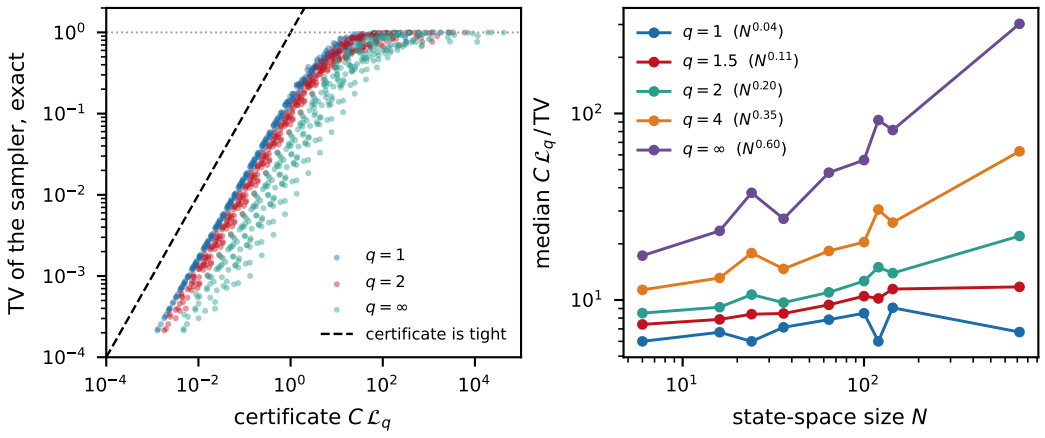


Figure 4: open: D5 Left: the exact total variation of the sampler against the certificate $C\mathcal{L}_q$, one point per perturbed flow under a uniform ν_T; every point lies below the diagonal. Right: the median of their ratio against the size of the state space, with its fitted exponent, flat at $q = 1$.

KL-weak-FM certifies certified lean [L4] [L5]

- SIZE (exp13) On 3,472 flows the loss never falls below the self-entropy (0 exceptions). Where it is finite, $\mathrm{TV}/\sqrt{\mathcal{L} - \mathcal{H}}$ is at most 0.323 inside the proved scope and 0.612 beyond it, both under $1/\sqrt{2}$: the square-root term alone accounts for the error.
- CHECK (exp14, exp15, exp16) Along the training of flows on the checkerboard and on knots, at 126 checkpoints with a finite loss, the certified bound $\frac{1}{\sqrt{2}}\sqrt{\mathcal{L} - \mathcal{H}} + \min(1, N(\mathcal{L} - \mathcal{H}))$, taken at $N = 1$, is violated 0 times; where it is informative the error is at most 66% of it, and early in training the linear term is needed, $\mathrm{TV}/\sqrt{\mathcal{L} - \mathcal{H}}$ reaching 1.000.
- LIMIT (exp13) The loss is unforgiving of a missing support: it is infinite on 1,900 of the 3,472 perturbed flows, wherever the target has mass and the flow stops nowhere.

Ratio losses cannot certify certified lean [L6] [L7] [L33]

- MECHANISM (exp4) open: D5 On the eight-cycle C_8, adding the circulation $k\gamma$ for k up to $1 \cdot 10^6$ leaves the sampling error at 0.98571 for every k (spread 0.0), while the $(\log x)^2$ loss falls to $2.2 \cdot 10^{-13}$; the stable loss stays at 0.2464 throughout (spread $1.1 \cdot 10^{-15}$). Its bound is vacuous on this instance (5.91), so the contrast is in shape, not in a usable certificate.

The acyclic comparison certified lean [L34] [L35] No measurement bears on this statement yet.

WHAT CAN BE REACHED

Universality on finite graphs certified lean [L36] [L37] [L11] [L10] [L38]

- CHECK (exp20, exp21a) The construction runs on graphs with cycles: the balanced flow closes flow matching to $3.7 \cdot 10^{-11}$ on 25 cyclic sorting graphs, and the universality certificate closes it to $6.0 \cdot 10^{-8}$ on all 9,216 doubling-ladder rows.

L^p-universality certified lean [L12] [L13]

- BEYOND (exp10) On product-replacement graphs property (T) makes the hypothesis a theorem, with a spectral gap uniform over groups: across the ten groups carried at every k, the inverse spectral gap of the uniform walk, which is its $\widehat{B}$ since the walk is symmetric,

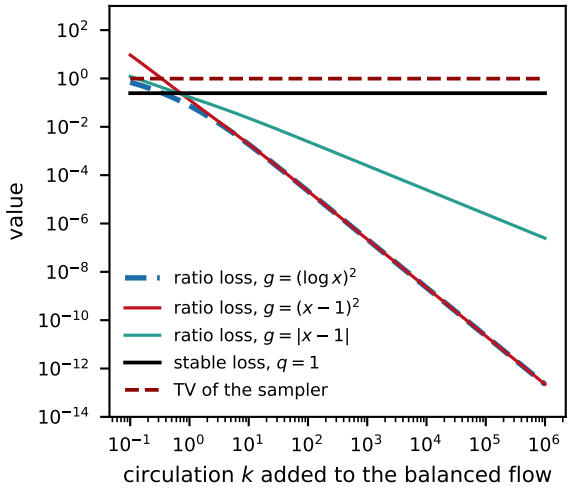


Figure 5: On the eight-cycle C_8, adding a circulation to the balanced flow drives the ratio losses to zero while neither the stable loss nor the sampling error moves.

lies in 3.78–14.8 at $k = 3$, where (T) fails, and in 5.39–12.9 and 6.6–12.7 at $k = 4$ and $k = 5$, on state spaces of up to 386,880 states.

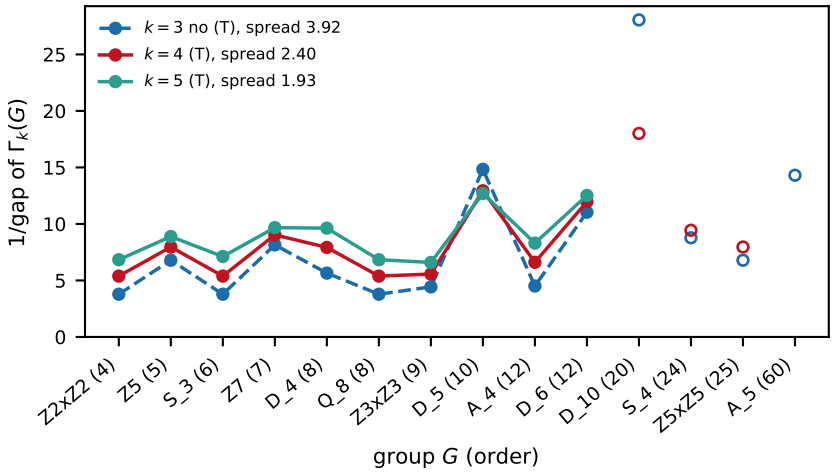


Figure 6: The inverse spectral gap of the uniform product-replacement walk, which is its $\widehat{B}$ since the walk is symmetric, against the group order: lines join the ten groups carried at every k, open markers are larger groups. The spread across groups narrows as k passes the threshold of property (T).

Gradient descent diffuses the flow

First variation certified lean [L15] No measurement bears on this statement yet.

The gradient is a diffusion certified lean [L14]

- PREDICTION (exp3) The exact gradient flow, with gradients taken by automatic differentiation rather than from the first-variation formula, relaxes at the rate of the linearized diffusion on all 20 chains; see the local rate.

Curvature and the local rate certified lean [L39] [L40] [L16] [L17]

- PREDICTION (exp3) On 20 leveled, Cayley and product-replacement chains, the measured rate over the predicted $g''(1)w_{\min}/\widehat{B}^2$ lies in 1.003–1.102: the prediction is never optimistic (Figure 7).

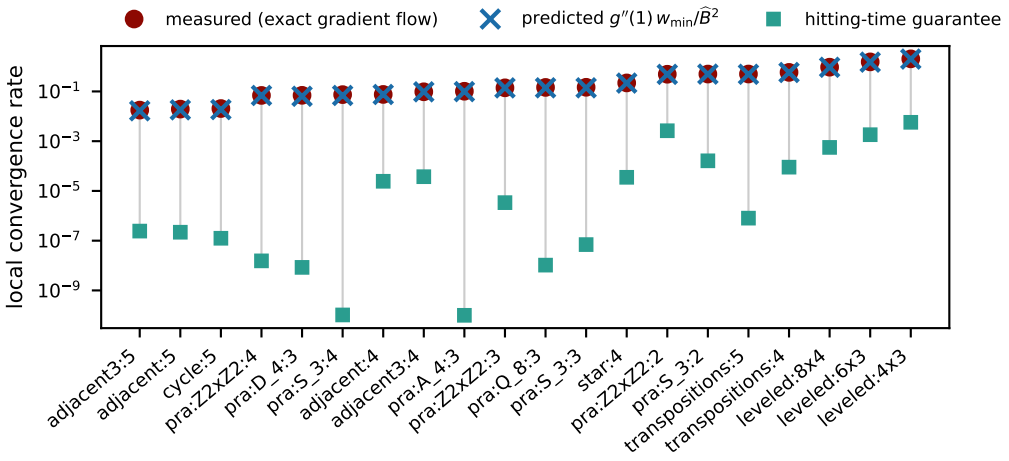


Figure 7: The local rate of the exact gradient flow on each of the 20 chains of the rate study, against the rate predicted from the Green operator and the rate the hitting-time bound guarantees.

TB: a step cap, no rate certified lean [L25] No measurement bears on this statement yet.

Global convergence for $(\log x)^2$ certified lean [L18] [L41]

- LIMIT (exp3) The exponential rate the theorem guarantees holds on all 20 chains and is conservative by a factor in 190–$1.0 \cdot 10^9$, most on a product-replacement chain (Figure 7). The global two-phase dynamics has not been measured.

Freezing is not stability certified lean [L19] [L20] No measurement bears on this statement yet.

Mass ascent and freezing certified lean [L26] [L42] No measurement bears on this statement yet.

The global dichotomy certified lean [L21] No measurement bears on this statement yet.

THE DOUBLING GRAPH: WHERE THE CONSTANT IS INFINITE

Phases and backward length certified lean [L43] [L44] [L45] [L46]

- CHECK (exp20) At the largest truncation the mean backward length is 6.00000 at $c = 1/4$ and 8.99991 at $c = 1/2$, against the closed forms $\bar{\jmath}/(1-c)$ of 6 and 9.

Cramér decay of λ certified lean [L47] [L48] [L49] [L50]

- CHECK (exp20) The decay exponent fitted to λ at the largest truncation differs from the Cramér root p^* by at most $1.3 \cdot 10^{-3}$ over 8 values of c, negative roots included.

$\widehat{B} = \infty$ at finite mean length certified lean [L22] [L51] [L52] [L53] [L54]

- MECHANISM (exp20) At $c = 1/4$ the mean backward length is 6.0000, yet $\widehat{B}_K$ grows from 17.2 to 1,048 as the truncation grows from 64 to 4,096 states, with log-log slope 0.989; every c tested shows the same doubling.

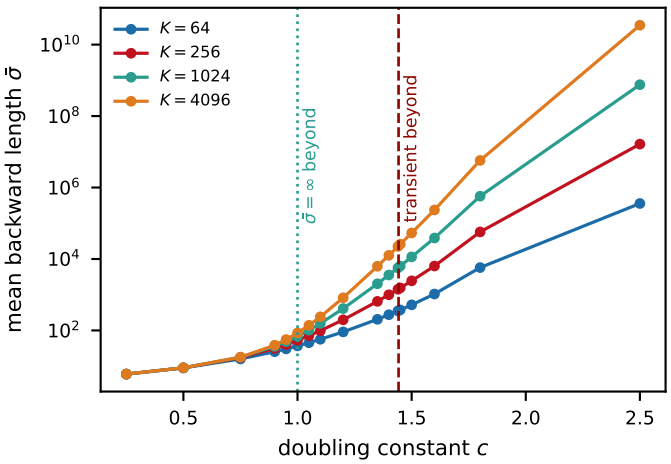


Figure 8: The mean backward length against the doubling constant c at $s = 1$, one curve per truncation K: the curves coincide for $c < 1$, where the length is finite, and fan out beyond, where the truncation hides a divergence; the chain is recurrent up to $c = 1/\ln 2$ and transient past it.

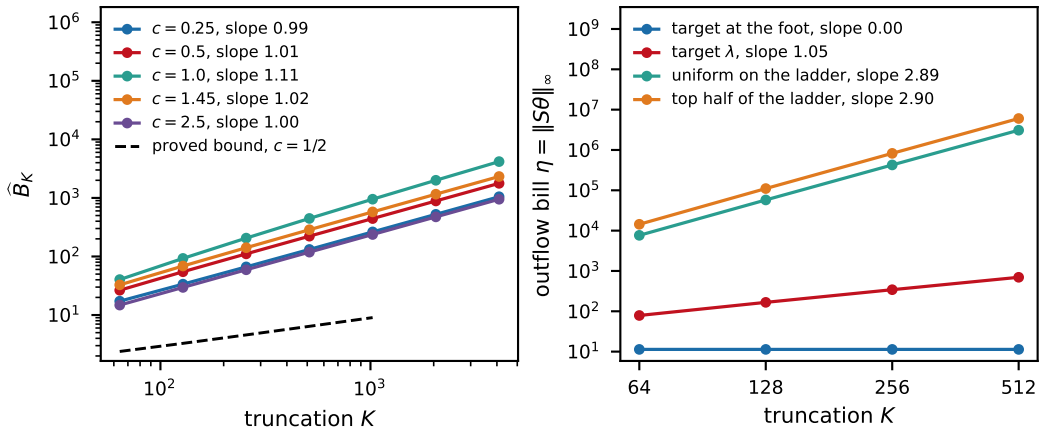


Figure 9: Left: $\widehat{B}_K$ against the truncation, one curve per doubling constant c, with the proved lower bound. Right: the outflow bill of the universality certificate at $c = 1/2$ from a Dirac initial flow: flat for a target at the foot of the ladder, growing polynomially once the target reaches into it.

Exact L^2 flow matching fails certified lean [L23]

- LIMIT (exp20) The failure cannot be observed at finite size, where every truncation is universal; its shadow is the outflow bill of the certificate, which at $c = 1/2$ from a Dirac initial flow stays at 11.35 at every truncation (spread $4.6 \cdot 10^{-13}$) for a target at the foot of the ladder and grows from $1.44 \cdot 10^4$ to $6.03 \cdot 10^6$ for one loaded at its top (Figure 9, right).

The price at finite size certified lean [L55] [L56] [L24]

- SIZE (exp20) The proof gives growth at least $\sqrt{K}$; the measured log-log slope of $\widehat{B}_K$ lies in 0.989–1.11 across c, and the per-cut lower bound of the proof holds in 40 of 40 cells (Figure 9).

OPEN DISCREPANCIES

D5 The committed tightness ratios use the constant $C_0(1 + 2/\kappa(\mathcal{S}))$ of the experiment library, not the certified $2C_0/\kappa(\mathcal{S})$; with $\kappa(\mathcal{S}) = 1$ they overstate the certified tightness by a factor of 1.5, and are corrected when the re-run results reach the pin.

D2 On the doubling ladder the tightness is flat only where the bound is informative; over all rows at $q = 1$ its median across truncations spans 15.6–102.

D1 The committed rates on the torus lattices do not match the corrected table of the experiment register, which derives the backward policy of the flow rather than taking it uniform; on one chain the measured rate exceeds the predicted one by a factor of 1.79, against at most 1.102 on every other chain. No torus rate is quoted until this is settled.

D3 The register compares the inverse spectral gap with $\widehat{B}$ for policies reversible before the loop closure on a set of chains that is not in the pinned results; the reading that the two nearly agree is therefore not quoted here.

E PROVENANCE

The Lean development that certifies this document is its supplementary material, which unpacks to `supplementary/lean/`. Every mark `[Ln]` in the text refers to an entry of Appendix F, which names the Lean declaration and gives its file and line in that material. The development was checked against `lean4` v4.31.0 and Mathlib `v4.31.0`; it contains 5,721 declarations and 0 unproved goals, and `make check` in that directory compiles it and replays its audits, as its `README.md` describes.

F LEAN REFERENCES

Each mark `[Ln]` in this document refers to the entry below with the same key: the full name of a Lean declaration and, beneath it, the file and line where the declaration begins (its docstring). Paths are relative to the Lean project of the supplementary material, the directory that holds its `lakefile.toml`.

[L1] GFNBounds.Core.SamplingGeneral.MFlow.sampling_theorem
GFNBounds/Core/SamplingGeneral.lean:637

[L2] GFNBounds.Core.SamplingGeneral.MFlow.negative_control_general
GFNBounds/Core/SamplingGeneralBounds.lean:361

[L3] GFNBounds.Core.SamplingGeneral.stable_bound_general
GFNBounds/Core/SamplingGeneralBounds.lean:603

[L4] GFNBounds.Core.SamplingGeneral.il_first_bullet_general
GFNBounds/Core/SamplingGeneralBounds.lean:704

[L5] GFNBounds.Core.ILBoundFull.il_tv_le_sqrt
GFNBounds/Core/ILBoundFull.lean:725

[L6] GFNBounds.Graph.SamplerWiring.no_bound_divergence_minimax
GFNBounds/Graph/SamplerWiring.lean:493
[L7] GFNBounds.Graph.SamplerWiring.cycle_no_model_free_bound_sampler
GFNBounds/Graph/SamplerWiring.lean:528
[L8] GFNBounds.Core.lem_sigma_mixing
GFNBounds/Core/SigmaMixing.lean:505
[L9] GFNBounds.Graph.BackwardPolicy.coercivity_morozov
GFNBounds/Graph/Morozov.lean:1027
[L10] GFNBounds.Graph.SamplerWiring.universality_graphs_sampler
GFNBounds/Graph/SamplerWiring.lean:749
[L11] GFNBounds.Graph.BackwardPolicy.universality_graphs_three
GFNBounds/Graph/Universality.lean:520
[L12] GFNBounds.Core.weaklyUniversal_of_kernel
GFNBounds/Core/Kernel.lean:589
[L13] GFNBounds.Core.stronglyUniversal_of_kernel_top
GFNBounds/Core/StrongUniversality.lean:199
[L14] GFNBounds.Balance.GdDiffusionGeneral.theo_gd_diffusion_full
GFNBounds/Balance/GdDiffusionGeneral.lean:1253
[L15] GFNBounds.Core.General.FirstVariation.theo_first_variation
GFNBounds/Core/FirstVariationGeneral.lean:1555
[L16] GFNBounds.Balance.local_convergence_full_paper
GFNBounds/Balance/LocalConvergenceClose.lean:151
[L17] GFNBounds.Balance.local_convergence_full_DB_paper
GFNBounds/Balance/LocalConvergenceClose.lean:241
[L18] GFNBounds.Balance.DiscreteGlobal.training_speed_full_complete
GFNBounds/Balance/DiscreteGlobal.lean:2643
[L19] GFNBounds.Graph.BackwardPolicy.frozen_unstable_full_fm
GFNBounds/Graph/FrozenUnstable.lean:543
[L20] GFNBounds.Graph.BackwardPolicy.frozen_unstable_full_db
GFNBounds/Graph/FrozenUnstableDB.lean:412
[L21] GFNBounds.Balance.General.Bridge.global_dichotomy_full_bridge
GFNBounds/Balance/FreezingGradientBridge.lean:351
[L22] GFNBounds.Doubling.main_unbounded
GFNBounds/Doubling/Main.lean:188
[L23] GFNBounds.Doubling.main_unsolvable
GFNBounds/Doubling/Main.lean:274
[L24] GFNBounds.Doubling.main_truncation_sqrtK
GFNBounds/Doubling/Main.lean:316
[L25] GFNBounds.Balance.TBHessian.HTB_le
GFNBounds/Balance/TBHessian.lean:963

```
[L26] GFNBounds.Balance.General.Bridge.no_distant_equilibrium_two_bridge
      GFNBounds/Balance/FreezingGradientBridge.lean:449
[L27] GFNBounds.Core.General.adjoint_general
      GFNBounds/Core/AdjointGeneral.lean:598
[L28] GFNBounds.Core.Mixing.coercivity
      GFNBounds/Core/Mixing.lean:206
[L29] GFNBounds.Graph.BackwardPolicy.morozov_rate
      GFNBounds/Graph/Morozov.lean:1067
[L30] GFNBounds.Graph.BackwardPolicy.lam_eq_visits_div
      GFNBounds/Graph/Morozov.lean:800
[L31] GFNBounds.Balance.RemarksA.graphs_vs_L2_periodic
      GFNBounds/Balance/RemarksA.lean:1395
[L32] GFNBounds.Balance.LiftGeneral.lift_coercivity_general
      GFNBounds/Balance/LiftGeneralMixing.lean:535
[L33] GFNBounds.Graph.SamplerWiring.cycle_counterexample_full
      GFNBounds/Graph/SamplerWiring.lean:347
[L34] GFNBounds.Silva.no_state_space_free_bound
      GFNBounds/Silva/NoUniform.lean:459
[L35] GFNBounds.Silva.exists_star_witness
      GFNBounds/Silva/NoUniform.lean:425
[L36] GFNBounds.Graph.BackwardPolicy.universality_graphs_one
      GFNBounds/Graph/Setting.lean:590
[L37] GFNBounds.Graph.BackwardPolicy.universality_graphs_two
      GFNBounds/Graph/Universality.lean:296
[L38] GFNBounds.Graph.UniversalityClosing.
universality_graphs_strongly_universal
      GFNBounds/Graph/UniversalityClosing.lean:579
[L39] GFNBounds.Balance.StableFrozenGeneral.hessG_coercive
      GFNBounds/Balance/StableFrozenGeneral.lean:406
[L40] GFNBounds.Balance.StableFrozenGeneral.db_stable_frozen_general
      GFNBounds/Balance/StableFrozenGeneral.lean:489
[L41] GFNBounds.Balance.training_speed_full_paper
      GFNBounds/Balance/TrainingSpeedAssembled.lean:1079
[L42] GFNBounds.Balance.General.Bridge.nonlinear_freezing_three_bridge
      GFNBounds/Balance/FreezingGradientBridge.lean:534
[L43] GFNBounds.Doubling.main_irreducible
      GFNBounds/Doubling/Main.lean:108
[L44] GFNBounds.Doubling.main_phase
      GFNBounds/Doubling/Main.lean:114
[L45] GFNBounds.Doubling.main_sigmaBar_eq
      GFNBounds/Doubling/Main.lean:236
```

```
[L46] GFNBounds.Doubling.main_sigmaBar_le
      GFNBounds/Doubling/Main.lean:244
[L47] GFNBounds.Doubling.main_invariant_measure
      GFNBounds/Doubling/Main.lean:132
[L48] GFNBounds.Doubling.main_constant_determined
      GFNBounds/Doubling/Main.lean:154
[L49] GFNBounds.Doubling.main_constant_functional
      GFNBounds/Doubling/Main.lean:167
[L50] GFNBounds.Doubling.main_rate
      GFNBounds/Doubling/Main.lean:252
[L51] GFNBounds.Doubling.main_unbounded_infty
      GFNBounds/Doubling/Main.lean:205
[L52] GFNBounds.Doubling.main_unbounded_two
      GFNBounds/Doubling/Main.lean:223
[L53] GFNBounds.Doubling.General.lem_doubling_operator
      GFNBounds/Doubling/OperatorGeneral.lean:668
[L54] GFNBounds.Doubling.Stat.exists_diffusionOp
      GFNBounds/Doubling/OperatorL2.lean:157
[L55] GFNBounds.Doubling.main_truncation_irreducible
      GFNBounds/Doubling/Main.lean:292
[L56] GFNBounds.Doubling.main_truncation_bhat
      GFNBounds/Doubling/Main.lean:301
```