Technical result

(Dated: October 18, 2025)

Can a critical system obey ETH? a few key ingredients in achieving bounded error: (1) exponential decay may bound the higher-order term. The exponential function after integration keeps the same - maybe this is the key?? (2) for the double integral, the change of variable - t and t-s is key.

I. SET UP

Consider a quantum many-body Hamiltonian H_{tot}. We divide the system into two regions: a *system* region and a *bath* region, such that $H_{\mathrm{tot}} \in \mathcal{H}^{\mathrm{sys}} \otimes \mathcal{H}^{\mathrm{bath}}$. The original Hamiltonian can be formally written as

$$H_{\mathrm{tot}} = H_{\mathrm{sys}} + H_{\mathrm{bath}} + H_{\mathrm{SB}}, \tag{1}$$

though here H_{sys} and H_{bath} are uniform and identical just acting on different regions. Formally let us write $H_{\mathrm{SB}} = \sum_a S_a \otimes B_a$. Our goal is to construct an effective model that reproduces the original dynamical behaviour of H_{tot}. The central idea is to replace the explicit dynamics of the bath by *random operators*, which are generated from suitably defined random bath operators.

The key requirement is that this effective description accurately captures the dynamics within a specified region of interest. That is, in a region within $\mathcal{B}_{\mathrm{sys}}$ with length ℓ_0, the local obserbable dynamics error is bounded

$$\left|\langle O_{\ell_0}(t)\rangle_H - \langle O_{\ell_0}(t)\rangle_{\mathrm{rand}}\right| \tag{2}$$

by some time-independent quantities.

Now, formally, denote the new random Hamiltonian as that can be decomposed as

$$H_{\mathrm{rand}} = H_{\mathrm{sys}} + H_{\mathrm{rand,\ bath}} + H_{\mathrm{rand,\ int}}, \tag{3}$$

where H_{sys} acts on the $\mathcal{B}_{\mathrm{sys}}$ and $H_{\mathrm{rand,\ bath}}$ are some to-be-selected random operators acting on the $\mathcal{B}_{\mathrm{bath}}$.

To make progress and get a rough idea of the complexity of local observable dynamics, let us first make a few strong assumptions about the Hamiltonian.

(S1) Locality. The Hamiltonian is finite-range or exponentially decaying:

$$\|h_i\| \le J e^{-|i-j|/r_0}. \tag{4}$$

This ensures a Lieb–Robinson velocity v_{LR} such that

$$\|[A_x(t), B_y]\| \le c\, e^{-(|x-y|-v_{\mathrm{LR}}|t|)/r_0} \tag{5}$$

for any local operators A_x, B_y.

(S2) Rapid decay of spatio-temporal correlations. For any local bath operators B_x, B_y,

$$\left|\langle B_x(t) B_y(0)\rangle - \langle B_x(t)\rangle\langle B_y(0)\rangle\right| \le C_0\, e^{-|x-y|/\xi_x} e^{-|t|/\tau_c}, \tag{6}$$

where ξ_x and τ_c are the spatial and temporal correlation lengths.

A. Circuit construction for a truncated random bath and error analysis

1. Setting and notation

Let the full lattice be partitioned into a system region $\mathcal{B}_{\mathrm{sys}}$ and a bath region. Fix a truncation radius R and denote by $\mathcal{R}(R)$ the bath sites within graph distance R of $\mathcal{B}_{\mathrm{sys}}$. We construct a *truncated random bath* acting only on $\mathcal{R}(R)$. Let S_a denote local system operators (supported on $\mathcal{B}_{\mathrm{sys}}$) and B_x denote local bath operators at site x.

For the exact model define

$$H_{\mathrm{tot}} = H_{\mathrm{sys}} + H_{\mathrm{bath}} + H_{SB}, \qquad H_{SB} = \sum_{a\in\mathcal{B}_{\mathrm{sys}}} S_a \otimes B_a.$$

We denote the exact bath Heisenberg evolution by $B_x(t) = U_{\mathrm{bath}}^{\dagger}(t) B_x U_{\mathrm{bath}}(t)$ where $U_{\mathrm{bath}}(t) = \exp(-\mathrm{i} H_{\mathrm{bath}} t)$. The target two-point correlator is

$$C_{xy}(t) \;=\; \langle B_x(t) B_y(0) \rangle_{\rho_{\mathrm{bath}}}.$$

We build a circuit on $\mathcal{R}(R)$ with timestep $\delta t > 0$. Let N denote the number of steps such that $t = N\delta t$. At step $n \in \{0, \dots, N-1\}$ apply a local unitary layer U_n on $\mathcal{R}(R)$ (a brickwork of nearest-neighbour two-site gates interleaved with single-site rotations). The ensemble of circuits is generated by taking the gate parameters at each step to be random classical variables with specified spatial and temporal correlations (not necessarily independent across steps). The circuit bath Heisenberg evolution is then

$$\widetilde{B}_x(n\delta t) \;=\; U_0^{\dagger} U_1^{\dagger} \cdots U_{n-1}^{\dagger}\, B_x\, U_{n-1} \cdots U_1 U_0,$$

and the ensemble-averaged correlator is

$$\widetilde{C}_{xy}(t) \;=\; \mathbb{E}_{\mathrm{ens}}\big[\widetilde{B}_x(t) \widetilde{B}_y(0)\big].$$

As assumed in S1 and S2, the exact bath satisfies the clustering assumption and that system evolution obeys a Lieb–Robinson bound with velocity v_{LR}. Norms of local operators are uniformly bounded: $\|S_a\| \le s$, $\|B_x\| \le b$, $\|O_{\ell_0}\| \le o$.

B. Circuit as a piecewise-constant Hamiltonian and splitting

To connect the circuit approximation error with δt we view each layer U_n as approximating the short-time evolution generated by a local bath Hamiltonian $H_{\mathrm{bath}}(t)$ that one would like to emulate on $\mathcal{R}$. Concretely one can choose a decomposition

$$H_{\mathrm{bath}}(t) = \sum_{\alpha} H_{\alpha}(t)$$

where each $H_\alpha(t)$ is a sum of commuting local terms grouped so that $\mathrm{supp}(H_\alpha)$ are small (single-site or two-site). A standard product formula of order r (Lie–Trotter for $r = 1$, symmetric Strang for $r = 2$, higher-order Suzuki for $r \ge 2$) gives an approximation

$$U_{\mathrm{bath}}(\delta t) = \mathrm{e}^{-\mathrm{i}\,\delta t H_{\mathrm{bath}}(t)} = V(\delta t) + E(\delta t),$$

where $V(\delta t)$ is the product of exponentials of the constituent pieces $\exp(-\mathrm{i}\,\delta t H_\alpha(t))$ in the chosen ordering, and the local error operator satisfies

$$\|E(\delta t)\| \le K_{\mathrm{loc}}\, \delta t^{\,r+1} \tag{7}$$

for some constant K_{loc} depending on nested commutators of the H_α (BCH expansion). When the circuit layers U_n are chosen to be the product formula $V(\delta t)$ (possibly with stochasticised parameters to match ensemble statistics) the same scaling holds for the circuit approximation error.

Let $U(t)$ denote the exact bath evolution on $\mathcal{R}$ restricted from the full bath (or the evolution under the true H_{bath} projected onto $\mathcal{R}$ for the time window of interest), and let $V_{\delta t}(t)$ denote the circuit (product formula) approximation accumulated over $N = t/\delta t$ steps. From the standard product formula error we have the global unitary error bound

$$\|U(t) - V_{\delta t}(t)\| \le K_1\, t\, \delta t^{\,r}, \tag{8}$$

for $t \le \tau_{\mathrm{cut}}$, where K_1 depends on K_{loc} and the decomposition; the factor t arises from summation of local errors over N steps.

For any fixed local bath operator B_x (with $\|B_x\| \le b$), the difference in Heisenberg evolution satisfies

$$\|B_x(t) - \widetilde{B}_x(t)\| = \|U^{\dagger}(t) B_x U(t) - V_{\delta t}^{\dagger}(t) B_x V_{\delta t}(t)\| \le 2b\, \|U(t) - V_{\delta t}(t)\|.$$

Using ((8)) we obtain the correlator error bound for $t \le \tau_{\mathrm{cut}}$:

$$\begin{aligned}
\big|C_{xy}(t) - \widetilde{C}_{xy}(t)\big| &= \big|\langle B_x(t) B_y(0)\rangle - \mathbb{E}_{\mathrm{ens}}[\widetilde{B}_x(t)\widetilde{B}_y(0)]\big| \\
&\le \big\|B_x(t) - \widetilde{B}_x(t)\big\|\, \|B_y\| + \big\|\widetilde{B}_x(t)\big\|\, \big\|B_y(0) - \widetilde{B}_y(0)\big\| \\
&\le 4b^2\, \|U(t) - V_{\delta t}(t)\| \\
&\le 4b^2 K_1\, t\, \delta t^{\,r}.
\end{aligned} \tag{9}$$

Thus *for splitting order* r the correlator difference scales as

$$\sup_{|t|\le\tau_{\rm cut}} \left|C_{xy}(t)-\widetilde{C}_{xy}(t)\right| \;=\; O(\tau_{\rm cut}\,\delta t^{\,r}).$$

C. Observable error: cumulant expansion and operator-norm bounds

For any system observable O (with $\|O\|\le\|O\|_\infty$) define the reduced Heisenberg map

$$\Phi_H(t)[O] := \mathrm{Tr}_{\rm bath}\big(U_H^\dagger(t)\,(O\otimes I_{\rm bath})\,U_H(t)\,\rho_{\rm bath}\big).$$

Let $\widetilde{H}$ denote a model Hamiltonian (e.g. truncated random-bath circuit) with corresponding map $\Phi_{\tilde H}(t)$. Denote

$$\Delta\Phi(t) := \Phi_H(t)-\Phi_{\tilde H}(t), \qquad \Delta(t;O) := \Phi_H(t)[O]-\Phi_{\tilde H}(t)[O].$$

Define the interaction-picture bath operators $B_a(s)=e^{\mathrm{i}H_{\rm bath}s}B_ae^{-\mathrm{i}H_{\rm bath}s}$ and similarly for $\widetilde{B}_a(s)$. Denote connected (cumulant) n-point functions for the true bath by $\kappa^{(n)}_{a_1\dots a_n}(s_1,\dots,s_n)$ and for the model bath by $\widetilde{\kappa}^{(n)}$. In particular the two-point function is $C_{ab}(\tau):=\langle B_a(\tau)B_b(0)\rangle$.

We assume the bath satisfies clustering / decay so that position sums and time integrals below converge; precise decay assumptions are stated where needed.

Working in the interaction picture and expanding the time-ordered exponential, the reduced map can be written as a formal cumulant (connected) expansion (influence-functional form)

$$\Phi_H(t) \;=\; \mathcal{T}\exp\Big(\sum_{n=1}^{\infty}\mathcal{I}^{(n)}(t)\Big),$$

where each $\mathcal{I}^{(n)}(t)$ is a (super)operator built from the n-point connected cumulant $\kappa^{(n)}$ and nested commutators of the S-operators. The exact difference is therefore

$$\Delta\Phi(t)=\mathcal{T}\exp\Big(\sum_{n\ge1}\mathcal{I}^{(n)}\Big)-\mathcal{T}\exp\Big(\sum_{n\ge1}\widetilde{\mathcal{I}}^{(n)}\Big).$$

Expanding the difference of exponentials yields an infinite series; grouping by cumulant order, one may write formally

$$\Delta\Phi(t)=\sum_{n=1}^{\infty}\Big(\mathcal{I}^{(n)}(t)-\widetilde{\mathcal{I}}^{(n)}(t)\Big)+(\text{nested commutator / cross terms from the exponential}).$$

The leading contribution is the first-order term (linear in $\kappa^{(1)}$), the next is the second-order term (involving the two-point function), etc. The displayed "double-integral" formula

$$\mathcal{I}^{(2)}(t)-\widetilde{\mathcal{I}}^{(2)}(t)=\sum_{a,b}\int_0^t ds_1\int_0^{s_1} ds_2\;\mathcal{K}_{ab}(s_1,s_2)\,\big(C_{ab}(s_1-s_2)-\widetilde{C}_{ab}(s_1-s_2)\big)$$

is precisely the second-order cumulant contribution (with $\mathcal{K}_{ab}$ made explicit below). This contribution is *exact* if all higher connected cumulants agree (or vanish); in general higher-order terms must be added.

Explicit first- and second-order contributions Define the interaction-picture system operators $S_a(s):=e^{\mathrm{i}H_{\rm sys}s}S_ae^{-\mathrm{i}H_{\rm sys}s}$. A convenient explicit symmetric representation of the second-order superoperator acting on O is

$$\mathcal{I}^{(2)}(t)[O]=-\sum_{a,b}\int_0^t ds_1\int_0^{s_1} ds_2\;\big[S_a(s_1),\;\big[S_b(s_2),\,O\big]\big]\;C_{ab}(s_1-s_2),$$

and similarly for $\widetilde{\mathcal{I}}^{(2)}(t)$ with $\widetilde{C}_{ab}$. Hence the second-order difference is

$$(\mathcal{I}^{(2)}-\widetilde{\mathcal{I}}^{(2)})(t)[O]=-\sum_{a,b}\int_0^t ds_1\int_0^{s_1} ds_2\;\big[S_a(s_1),\;\big[S_b(s_2),\,O\big]\big]\;\big(C_{ab}(s_1-s_2)-\widetilde{C}_{ab}(s_1-s_2)\big).$$

The first-order term (linear in single-point expectation $\langle B_a(s)\rangle$) reads

$$\mathcal{I}^{(1)}(t)[O] = -\mathrm{i}\sum_a \int_0^t ds\, \big[S_a(s), O\big]\, \langle B_a(s)\rangle,$$

and the difference $\mathcal{I}^{(1)} - \widetilde{\mathcal{I}}^{(1)}$ depends on $\langle B_a(s)\rangle - \langle \widetilde{B}_a(s)\rangle$. Often one assumes zero-mean bath operators $\langle B_a(s)\rangle = 0$, in which case the first-order term vanishes identically.

Second-order bound. Using $\|[X,[Y,O]]\| \le 4\|X\|\|Y\|\|O\|$ and $\|S_a\| \le s$,

$$\begin{aligned}\big\|(\mathcal{I}^{(2)} - \widetilde{\mathcal{I}}^{(2)})(t)[O]\big\| &\le 4s^2\|O\|_\infty \sum_{a,b}\int_0^t ds_1 \int_0^{s_1} ds_2\, \big|C_{ab}(s_1-s_2) - \widetilde{C}_{ab}(s_1-s_2)\big| \\ &= 4s^2\|O\|_\infty \sum_{a,b}\int_0^t d\tau\,(t-\tau)\,\big|C_{ab}(\tau) - \widetilde{C}_{ab}(\tau)\big|, \qquad (10)\end{aligned}$$

where we changed variables $\tau = s_1 - s_2$ and used the triangular integration domain.

Equation ((10)) is exact and non-perturbative: it bounds the operator norm of the second-order contribution for arbitrary coupling strength s. Note the explicit $(t-\tau)$ factor: if the integrand $|C_{ab}(\tau) - \widetilde{C}_{ab}(\tau)|$ is supported or rapidly decaying in τ (e.g. up to τ_{cut}), then the integral over τ yields a time-independent bound for $\sup_{t\ge 0}$ (because for $t > \tau_{\mathrm{cut}}$ the factor $(t-\tau)$ can be bounded by t but the decaying integrand makes the overall expression saturate — see clustering/LR discussion below).

General n-th order bound (higher orders) The n-th cumulant contribution acting on O has the symbolic structure

$$\mathcal{I}^{(n)}(t)[O] \sim \sum_{a_1,\dots,a_n} \int_{0\le s_n\le\cdots\le s_1\le t} ds_1\cdots ds_n\, \mathcal{C}^{(n)}_{a_1\dots a_n}(s_1,\dots,s_n)\, \kappa^{(n)}_{a_1\dots a_n}(s_1,\dots,s_n),$$

where $\mathcal{C}^{(n)}$ is a bounded nested-commutator superoperator with norm bounded by $(2s)^n\|O\|_\infty$. More precisely, using repeated commutator bounds one obtains the crude norm estimate

$$\big\|\big(\mathcal{I}^{(n)} - \widetilde{\mathcal{I}}^{(n)}\big)(t)[O]\big\| \le \frac{(2s)^n}{n!}\,\|O\|_\infty \sum_{a_1,\dots,a_n}\int_{[0,t]^n} ds_1\cdots ds_n\, \big|\kappa^{(n)}_{a_1\dots a_n}(s_1,\dots,s_n) - \widetilde{\kappa}^{(n)}(\cdot)\big|. \qquad (11)$$

The $1/n!$ compensates for the ordering volume when converting to an integral over the full hypercube; one may omit it and compensate in constants if desired.

Sufficient clustering hypothesis and summation of the series To give an explicit convergent bound for the total remainder one commonly employs a clustering / factorial bound on connected cumulants. Assume there exist constants $C_0 > 0$, $\xi_x > 0$, $\tau_c > 0$, and $a > 0$ such that for all $n \ge 2$ and all positions/time arguments

$$\big|\kappa^{(n)}_{a_1\dots a_n}(s_1,\dots,s_n)\big| \le C_0\,(a)^{n-2} \prod_{1\le i<j\le n} e^{-|x_{a_i}-x_{a_j}|/\xi_x} \prod_{i=1}^{n} e^{-|s_i|/\tau_c}. \qquad (12)$$

(Physically: spatial exponential clustering and exponential decay in each time argument; factorial/growth prefactor a^{n-2} absorbs combinatorics.) Suppose the same bound holds for the model cumulants and denote the difference by the same type of bound with a small prefactor when comparing.

Under ((12)) the spatial sums $\sum_{a_2,\dots,a_n}\prod_{i<j} e^{-|x_{a_i}-x_{a_j}|/\xi_x}$ produce only geometric factors (bounded by a constant G^{n-1} depending on lattice geometry and ξ_x). The time integrals $\int_{[0,t]^n}\prod_i e^{-|s_i|/\tau_c}ds_i$ produce a factor $(\tilde{\tau})^n$ with $\tilde{\tau} \sim \tau_c$ (uniform in t once $t \gtrsim \tau_c$). Combining these estimates into ((11)) yields, for some constants K, C',

$$\big\|\big(\mathcal{I}^{(n)} - \widetilde{\mathcal{I}}^{(n)}\big)(t)\big\| \le \|O\|_\infty\, C'\,\frac{\big(KaG\tilde{\tau}\big)^n}{n!}.$$

Summing over $n \ge 3$ and using the exponential / factorial decay of $1/n!$ yields convergence provided $KaG\tilde{\tau}$ is finite (it always is) and gives a finite remainder bound

$$\big\|\mathcal{R}_{\ge 3}(t)\big\| := \big\|\sum_{n\ge 3}\big(\mathcal{I}^{(n)} - \widetilde{\mathcal{I}}^{(n)}\big)(t)\big\| \le \|O\|_\infty\, C_{\mathrm{rem}},$$

where C_{rem} is an explicit finite constant depending on C_0, a, ξ_x, τ_c, s and geometry; importantly C_{rem} is *independent* of t (once $t \gtrsim \tau_c$).

D. Assembling the total observable error

a. Exact second-order contribution (operator form). For any system observable O the second-order cumulant contribution to the reduced map difference is

$$\Delta^{(2)}(t)[O] = \sum_{a,b} \int_0^t ds_1 \int_0^{s_1} ds_2 \; \mathcal{K}_{ab}(s_1, s_2)[O] \; \big(C_{ab}(s_1 - s_2) - \widetilde{C}_{ab}(s_1 - s_2)\big),$$

with $\mathcal{K}_{ab}(s_1, s_2)[\cdot] = -[S_a(s_1), [S_b(s_2), \cdot]]$. We bound norms using $\|[X, [Y, O]]\| \le 4\|X\|\|Y\|\|O\|$.

b. Decompose time and space domains. Change variables $\tau = s_1 - s_2$ and $u = s_2$. For fixed t the integration domain is $\{(u, \tau) : 0 \le u \le t - \tau,\ 0 \le \tau \le t\}$. Split the τ-integral at $\tau_{\rm cut}$ and the spatial sums at $\mathcal{R}/\mathcal{F}$. Then

$$\|\Delta^{(2)}(t)\| \le 4s^2 o \sum_{a,b} \int_0^t d\tau \, (t - \tau) \, \big|C_{ab}(\tau) - \widetilde{C}_{ab}(\tau)\big|$$

and we write

$$\sum_{a,b} = S_{RR} + S_{RF} + S_{FR} + S_{FF},$$

where e.g. S_{RF} denotes sum with $a \in \mathcal{R}, b \in \mathcal{F}$.

c. Error decomposition (spatial / temporal / mismatch / correlated). Using the above split and triangular inequality obtain the decomposition

$$\sup_{t\ge0} \|\Delta^{(2)}(t)\| \le E_{\rm spatial}(R, \tau_{\rm cut}) \; + \; E_{\rm temporal}(\tau_{\rm cut}, R) \; + \; E_{\rm match}(R, \tau_{\rm cut}) \; + \; E_{\rm corr}(R, \tau_{\rm cut}),$$

with the terms defined and bounded below.

(1) Spatial tail (purely spatial): errors from pairs where *both* indices are outside the kept region,

$$E_{\rm spatial} := 4s^2 o \sup_{t\ge0} \sum_{a,b\notin\mathcal{R}} \int_0^t d\tau \, (t - \tau) \, |C_{ab}(\tau)|.$$

Using clustering $|C_{ab}(\tau)| \le C_0 e^{-|x_a - x_b|/\xi_x} e^{-\tau/\tau_c}$ and shell counting $\sum_{b:d_b=r} 1 \le c_d r^{d-1}$, and performing the τ-integral, one obtains (for $t \gtrsim \tau_c$)

$$E_{\rm spatial}(R, \tau_{\rm cut}) \le C_{\rm s} \, P(R) \, e^{-(R - v_{\rm LR}\tau_{\rm cut})/\xi_x},$$

where $P(R)$ is a polynomial of degree $d - 1$, and $C_{\rm s}$ depends on C_0, s, o, τ_c and LR constants. (The $v_{\rm LR}\tau_{\rm cut}$ shift appears after including LR — see next paragraph.)

(2) Temporal tail (purely temporal): errors due to long-time lags (pairs inside kept region but $|\tau| > \tau_{\rm cut}$),

$$E_{\rm temporal} := 4s^2 o \sup_{t\ge0} \sum_{a,b\in\mathcal{R}} \int_{|\tau|>\tau_{\rm cut}} d\tau \, (t - \tau) \, |C_{ab}(\tau)|.$$

Using time decay $e^{-|\tau|/\tau_c}$ and $|\mathcal{R}|$ finite (volume $\sim R^d$),

$$E_{\rm temporal}(\tau_{\rm cut}, R) \le C_{\rm t} \, |\mathcal{R}| \, \tau_c e^{-\tau_{\rm cut}/\tau_c},$$

with $C_{\rm t}$ depending on C_0, s, o. This is uniform in t because the τ-integral converges.

(3) Correlator mismatch inside kept region (modeling error): model fails to reproduce correlator within kept region for $|\tau| \le \tau_{\rm cut}$,

$$E_{\rm match} := 4s^2 o \sup_{t\ge0} \sum_{a,b\in\mathcal{R}} \int_{|\tau|\le\tau_{\rm cut}} d\tau \, (t - \tau) \, |C_{ab}(\tau) - \widetilde{C}_{ab}(\tau)|.$$

If the model is fitted so that $\sup_{a,b\in\mathcal{R},\ |\tau|\le\tau_{\rm cut}} |C_{ab}(\tau) - \widetilde{C}_{ab}(\tau)| \le \delta_{\rm fit}$, then

$$E_{\rm match} \le 4s^2 o \, |\mathcal{R}|^2 \, \tau_{\rm cut} \, \delta_{\rm fit}.$$

This is linear in the fitting error $\delta_{\rm fit}$ and grows with the number of pairs in $\mathcal{R}$ (hence polynomial in R).

(4) Spatial–temporal correlated tail (mixed term): terms where one index is in $\mathcal{R}$ and the other is in $\mathcal{F}$, and/or where $|\tau| > \tau_{\rm cut}$ and one index far — these mix spatial and temporal truncation and are the cross/correlation terms:

$$E_{\rm corr} := 4s^2 o \sup_{t\ge 0}\Big(\sum_{a\in\mathcal{R}, b\notin\mathcal{R}} + \sum_{a\notin\mathcal{R}, b\in\mathcal{R}}\Big)\int_0^t d\tau\,(t-\tau)\,|C_{ab}(\tau) - \widetilde{C}_{ab}(\tau)|.$$

We bound the two symmetric double sums by twice the first one and focus on the sector with $a \in \mathcal{R}$, $b \notin \mathcal{R}$. Change integration variables to the time-difference $\tau = s_1 - s_2$ (with $0 \le \tau \le t$) and the base time $u = s_2$ (with $0 \le u \le t-\tau$). The factor $(t-\tau)$ is the remaining ordering volume; to obtain a uniform-in-time bound we will exploit the finite memory of the bath and restrict the effective u-integration to a bounded window of order $\tau_{\rm cut}$. Split the τ-integral at a memory cutoff $\tau_{\rm cut} > 0$ into short lags $\tau \in [0, \tau_{\rm cut}]$ and long lags $\tau > \tau_{\rm cut}$, and bound each contribution separately.

For the short-lag piece $\tau \le \tau_{\rm cut}$ we use the Lieb–Robinson locality of operator spreading together with the commutator-norm estimate implicit in the second-order kernel to obtain a spatio-temporal damping factor. Concretely, for a kept site a at distance $d_a \le R + \ell_0$ from the system and an omitted site b at distance $d_b \ge R$ one may bound the effective integrand (norm of the kernel times correlator) by a product of exponentials

$$\big|\mathcal{K}_{ab}(\cdot)\big|\,|C_{ab}(\tau)| \;\lesssim\; C_0\, K_{\rm kern}\, e^{-(d_a - v_{\rm LR}(u+\tau))/\lambda}\, e^{-(d_b - v_{\rm LR}u)/\lambda}\, e^{-\tau/\tau_c},$$

where $K_{\rm kern}$ is the nested-commutator prefactor (we may take the conservative $K_{\rm kern} = 4\,c_{\rm LR} s^2 o$). The two LR factors combine to produce an effective factor $e^{-(d_a+d_b-v_{\rm LR}(2u+\tau))/\lambda}$. The u-integral is effectively supported on $u \lesssim \tau_{\rm cut}$ (the bath memory), so for a uniform bound we replace the upper limit $t-\tau$ by $\tau_{\rm cut}$ and bound the u-integral by a finite factor proportional to $\tau_{\rm cut}\exp(2v_{\rm LR}\tau_{\rm cut}/\lambda)$. The remaining τ-integral over $[0, \tau_{\rm cut}]$ produces at most a finite factor $J(\tau_{\rm cut})$ (for physical baths $J(\tau_{\rm cut}) = O(\tau_c)$ when $1/\tau_c > v_{\rm LR}/\lambda$; otherwise J grows no faster than an exponential in $\tau_{\rm cut}$). The net pointwise bound for fixed a and b (short-lag) therefore reads

$$\mathcal{I}^{\rm short}_{ab} \le K_{\rm kern}\, C_0\, T_{\rm short}(\tau_{\rm cut})\, e^{-(d_a+d_b-v_{\rm LR}\tau_{\rm cut})/\lambda},$$

with $T_{\rm short}(\tau_{\rm cut})$ an explicit finite time factor. Since d_a is bounded by $R+\ell_0$ (finite), the dominant spatial dependence for omitted b is $e^{-d_b/\lambda}\, e^{-(R-v_{\rm LR}\tau_{\rm cut})/\lambda}$ up to constants. Summing over omitted b therefore yields the shell-sum

$$\sum_{b:\, d_b\ge R} e^{-d_b/\lambda} = \sum_{r\ge R} N(r)\, e^{-r/\lambda},$$

where $N(r)$ is the number of sites at graph distance r from the system region. On a d-dimensional lattice one may take the coarse bound $N(r) \le C_d(1+r)^{d-1}$; inserting this bound gives

$$\sum_{r\ge R} N(r)\, e^{-r/\lambda} \le C_d \sum_{r\ge R}(1+r)^{d-1} e^{-r/\lambda} \le P(R)\, e^{-R/\lambda},$$

with $P(R) = C'_d(1+R)^{d-1}$ a polynomial prefactor. Hence the short-lag mixed contribution summed over b is bounded by a constant times $P(R)\, e^{-(R-v_{\rm LR}\tau_{\rm cut})/\lambda}$.

For the long-lag piece $\tau > \tau_{\rm cut}$ the temporal clustering $e^{-\tau/\tau_c}$ dominates and the τ-integral is bounded by $\int_{\tau_{\rm cut}}^{\infty} e^{-\tau/\tau_c} d\tau = \tau_c e^{-\tau_{\rm cut}/\tau_c}$. The position dependence for each pair is bounded by the spatial clustering $e^{-|x_a - x_b|/\xi_x}$, and summing over omitted b again produces the same shell-sum $\sum_{r\ge R} N(r) e^{-r/\xi_x} \le P(R) e^{-R/\xi_x}$. Thus the long-lag mixed contribution is bounded by a constant times $P(R)\,\tau_c e^{-\tau_{\rm cut}/\tau_c}\, e^{-R/\xi_x}$.

Combining the short- and long-lag bounds, absorbing the finite a-sum and small shifts of order ℓ_0 into constants, produces the explicit uniform-in-time estimate for the mixed error

$$E_{\rm corr}(R, \tau_{\rm cut}) \le C_{\rm mix}\, P(R)\, \Big(e^{-(R-v_{\rm LR}\tau_{\rm cut})/\xi_x} + \tau_c e^{-\tau_{\rm cut}/\tau_c}\Big),$$

where the prefactor $C_{\rm mix}$ can be chosen explicitly as

$$C_{\rm mix} = 4\, c_{\rm LR}\, s^2\, o\, C_0\, T_{\rm short}(\tau_{\rm cut}),$$

and $P(R) = C'_d(1+R)^{d-1}$ is the polynomial shell-count coming from $\sum_{r\ge R}(1+r)^{d-1}e^{-r/\xi_x}$. Thus the mixed term is controlled by the product of an exponential spatial suppression (shifted by the LR light cone) and a polynomial shell-count prefactor; for fixed $\tau_{\rm cut}$ and $R > v_{\rm LR}\tau_{\rm cut}$ the exponential dominates and the mixed error decays exponentially in R.

d. Total second-order bound. Collecting all the pieces yields

$$\sup_{t\geq 0} \|\Delta^{(2)}(t)\| \leq C_1\Big(P(R)\, e^{-(R-v_{\rm LR}\tau_{\rm cut})/\xi_x} + |\mathcal{R}|\,\tau_c e^{-\tau_{\rm cut}/\tau_c} + |\mathcal{R}|^2\,\tau_{\rm cut}\,\delta_{\rm fit}\Big),$$

where C_1 depends on s, o, C_0 and LR constants and $P(R)$ is polynomial in R (degree $d-1$). The first and second terms are the spatial/temporal tails (with LR shift), the third is the in-region correlator fitting error; the mixed/correlation contributions are included in the first term (through the $v_{\rm LR}\tau_{\rm cut}$ shift) and in a constant factor multiplying $P(R)$.

e. Higher-order contributions. Under a cumulant clustering hypothesis of the form (for $n \geq 3$)

$$|\kappa^{(n)}_{a_1\ldots a_n}(s_1,\ldots,s_n)| \leq C_0'\, a^{n-2}\prod_{i<j} e^{-|x_{a_i}-x_{a_j}|/\xi_x}\prod_i e^{-|s_i|/\tau_c},$$

the n-th order difference $\Delta^{(n)}$ can be bounded analogously by a term

$$\|\Delta^{(n)}\| \leq C_n\, P_n(R)\,\big(e^{-(R-v_{\rm LR}\tau_{\rm cut})/\xi_x} + |\mathcal{R}|\,\tau_c e^{-\tau_{\rm cut}/\tau_c} + |\mathcal{R}|^n \tau_{\rm cut}^{n-1}\delta^{(n)}\big),$$

with $\delta^{(n)}$ the model mismatch of n-point cumulants inside the kept spacetime block, and P_n polynomial. Summation over $n \geq 3$ converges to $C_{\rm rem}$ provided the cumulants satisfy factorial/geometric suppression (typical physically: large bath, mixing).

f. Final uniform bound. Therefore, combining second-order and higher orders,

$$\sup_{t\geq 0} \|\Delta\Phi(t)\| \leq C\Big(P(R)\, e^{-(R-v_{\rm LR}\tau_{\rm cut})/\xi_x} + |\mathcal{R}|\,\tau_c e^{-\tau_{\rm cut}/\tau_c} + |\mathcal{R}|^2\,\tau_{\rm cut}\,\delta_{\rm fit}\Big) + C_{\rm rem},$$

where $C_{\rm rem}$ is the finite remainder coming from $n \geq 3$ terms (small if higher cumulants are suppressed) and C collects constants. The mixed / correlated spatial–temporal effects have been explicitly accounted for in the $e^{-(R-v_{\rm LR}\tau_{\rm cut})/\xi_x}$ factor and in polynomial prefactors $P(R)$.

g. Guidance for parameter choice. To achieve tolerance ε:

- choose $\tau_{\rm cut}$ so $|\mathcal{R}|\,\tau_c e^{-\tau_{\rm cut}/\tau_c} \lesssim \varepsilon/4$,
- choose R so $P(R)\, e^{-(R-v_{\rm LR}\tau_{\rm cut})/\xi_x} \lesssim \varepsilon/4$,
- fit correlators so $|\mathcal{R}|^2\,\tau_{\rm cut}\,\delta_{\rm fit} \lesssim \varepsilon/4$,
- ensure cumulant remainder $C_{\rm rem} \lesssim \varepsilon/4$ (increase bath degrees or match higher cumulants if needed).

II. BATH CONSTRUCTION

For an exponentially decaying bath, we may choose a random bath generated by a Haar random unitary. The features of such a bath should be much fewer. They may be easily inferred by measuring the local physical quantities , without knowing the microscopic details

A. Randomly compiled circuit baths and matching two-point correlators (general)

We replace the deterministic short-time product formula on the kept region $\mathcal{R}$ by a stroboscopic random circuit ensemble. Time is discretized in steps of size δt and we denote the discrete times $t_n = n\delta t$. At each step a random unitary layer U_n (brickwork of local gates) acts on $\mathcal{R}$; different layers have correlated random parameters chosen from a specified stochastic model. The Heisenberg evolution of a bath operator under one circuit realisation is $\widetilde{B}_x(t_n) = U_0^\dagger U_1^\dagger \cdots U_{n-1}^\dagger B_x U_{n-1}\cdots U_1 U_0$. Averaging over the circuit ensemble produces an averaged Heisenberg superoperator Λ_n acting on operators: for any bath operator X,

$$\mathbb{E}_{\rm ens}\big[\widetilde{X}(t_n)\big] \;=\; \Lambda_n[X], \qquad \Lambda_n \;=\; \mathbb{E}_{\rm ens}\big[\mathrm{Ad}_{U_0}\circ \mathrm{Ad}_{U_1}\circ\cdots\circ \mathrm{Ad}_{U_{n-1}}\big],$$

where $\mathrm{Ad}_U(X) = U^\dagger X U$. The two-point correlator produced by the circuit ensemble (for a fixed bath initial state $\rho_{\rm bath}$) is

$$\widetilde{C}_{ab}(t_n) \;=\; \mathbb{E}_{\rm ens}\big[\langle \widetilde{B}_a(t_n)\, B_b(0)\rangle_{\rho_{\rm bath}}\big] \;=\; \mathrm{Tr}\big(B_a\,\Lambda_n[B_b]\,\rho_{\rm bath}\big).$$

Hence matching $\widetilde{C}_{ab}(t_n)$ to the target correlator $C_{ab}(t_n) = \mathrm{Tr}(B_a\, B_b(t_n)\rho_{\mathrm{bath}})$ is equivalent to choosing the ensemble so the averaged superoperators Λ_n act on the local operator subspace in a prescribed way.

Introduce an orthonormal operator basis $\{O_\mu\}_{\mu=1}^{D}$ for operators supported on the kept region $\mathcal{R}$ (for example tensor products of Pauli strings truncated by radius). Flatten superoperators into $D \times D$ matrices via the Hilbert–Schmidt inner product $\langle X, Y\rangle = \mathrm{Tr}(X^\dagger Y \rho_{\mathrm{bath}})$ (or the trace inner product if ρ_{bath} is full rank). Denote the matrix representation of Λ_n by $M^{(n)}$ with matrix elements $M^{(n)}_{\mu\nu} = \langle O_\mu, \Lambda_n[O_\nu]\rangle$. The target action of the true bath on basis elements is given by matrices $T^{(n)}$ with elements $T^{(n)}_{\mu\nu} = \langle O_\mu, B_\nu(t_n)\rangle$ (these are directly related to the target correlators: $C_{ab}(t_n) = \sum_{\mu,\nu} c_{a\mu} T^{(n)}_{\mu\nu} c_{b\nu}$ where $B_x = \sum_\mu c_{x\mu} O_\mu$). Matching correlators on the kept region and on the sample times $\{t_n\}$ is equivalent to the matrix fitting problem

$$\text{find ensemble parameters} \quad \text{s.t.} \quad M^{(n)} \approx T^{(n)} \quad \text{for all } |t_n| \le \tau_{\mathrm{cut}}.$$

There are two complementary constructive strategies.

One method is direct statistical parametrisation and least-squares fitting. Choose a parametric stochastic model for the random layer $U(\theta)$ (e.g. each two-site gate is $\exp(-\mathrm{i}\theta_{ij} H_{ij})$ with θ_{ij} sampled from a Gaussian process with mean $\mu_{ij}(n)$ and covariance $K_{(ij,n),(i'j',m)}$). The averaged single-step superoperator is $\bar{S} := \mathbb{E}_\theta[\mathrm{Ad}_{U(\theta)}]$. If successive steps are independent and identically distributed (i.i.d.), then $\Lambda_n = \bar{S}^n$. In the i.i.d. case the fitting problem reduces to choosing $\bar{S}$ so that $\bar{S}^n$ approximates $T^{(n)}$ for n up to $N_{\mathrm{cut}} = \tau_{\mathrm{cut}}/\delta t$. This is a matrix approximation (nonlinear in ensemble parameters but linear in the unknown $\bar{S}$). A practical algorithm is: compute empirical target matrices $T^{(n)}$ (from the true bath or its samples), form the loss $\mathcal{L}(\bar{S}) = \sum_{n=1}^{N_{\mathrm{cut}}} \|\bar{S}^n - T^{(n)}\|_F^2$ and optimise $\bar{S}$ over the convex set of unital CP maps that are reachable by the ensemble (projecting onto physically allowed maps if needed). In many cases one parameterises $\bar{S}$ in a low-dimensional basis (Kraus matrices built from local generators) so optimisation is tractable. The gradient of $\mathcal{L}$ with respect to ensemble parameters can be computed by chain rule using matrix powers; stochastic gradient methods are effective when the matrix dimension D is large.

A second (often more efficient) method leverages the transfer-matrix / spectral decomposition. Model the ensemble as a stationary Markov process on layer parameters, so that $\Lambda_n = \Theta^n$ where Θ is a one-step averaged superoperator (the transfer operator). Diagonalize or Schur-decompose Θ on the relevant operator subspace: $\Theta = V\Lambda V^{-1}$ with eigenvalues $\{\lambda_j\}$. Then $\Theta^n = V\Lambda^n V^{-1}$ exhibits multi-exponential time dependence. Fit the target time series of matrix entries $T^{(n)}_{\mu\nu}$ by matching the modal decomposition, i.e. solve a Prony/linear prediction problem to extract dominant decay/oscillation exponents λ_j and mode projectors. This reduces the fitting to matching a small number r of dominant eigenmodes so that for $n \le N_{\mathrm{cut}}$ one has $T^{(n)} \approx \sum_{j=1}^{r} \lambda_j^n\, P_j$ with rank-P_j. Once $\{\lambda_j, P_j\}$ are chosen to fit the target, one synthesises an ensemble that realizes Θ with that spectrum (explicit construction available for local transfer operators, see e.g. parametrised Kraus maps or correlated random gate ensembles). This spectral approach is numerically stable when the bath correlator has a few dominant relaxation/oscillation modes (typical for baths with exponential memory).

Both approaches reduce to a finite-dimensional optimization problem. For the direct least-squares approach the normal equations are linear in the unknown entries of $\bar{S}$ when powers are held fixed, but nonlinearity arises because $\bar{S}$ appears inside powers; one may use alternating minimisation: first fit $\bar{S}$ to match $\{T^{(n)}\}$ by linearising around an initial guess, then reproject to physically admissible maps. For the spectral approach one fits the eigenvalues λ_j and eigenprojections P_j using linear prediction / Hankel matrix SVD.

Error propagation to observable dynamics is controlled by the influence-functional bound derived earlier: a correlator mismatch $\sup_{a,b,|\tau|\le\tau_{\mathrm{cut}}} |C_{ab}(\tau) - \widetilde{C}_{ab}(\tau)| \le \delta_{\mathrm{fit}}$ produces a reduced-map operator norm error bounded by a constant times $|\mathcal{R}|^2 \tau_{\mathrm{cut}}\, \delta_{\mathrm{fit}}$. Combined with the spatial and temporal tail bounds, the total uniform observable error satisfies

$$\sup_{t\ge 0} \left|\langle O(t)\rangle_H - \langle O(t)\rangle_{\mathrm{rand}}\right| \le C_{\mathrm{tails}}\left(e^{-(R - v_{\mathrm{LR}}\tau_{\mathrm{cut}})/\xi_x} + \tau_c e^{-\tau_{\mathrm{cut}}/\tau_c}\right) + C_{\mathrm{fit}}\, |\mathcal{R}|^2 \tau_{\mathrm{cut}}\, \delta_{\mathrm{fit}} + C_{\mathrm{rem}},$$

with $C_{\mathrm{fit}} = 4s^2\|O\|$ as in the second-order bound.

In practice one proceeds as follows. Choose a basis $\{O_\mu\}$ on $\mathcal{R}$ and sample the true correlator time series $T^{(n)}_{\mu\nu}$ for $|t_n| \le \tau_{\mathrm{cut}}$. Use SVD/Prony to extract a small number of dominant modes or directly minimise the Frobenius loss $\sum_n \|M^n - T^{(n)}\|_F^2$ over admissible $\bar{S}$ (or over parameters of the stochastic gate model) subject to CPTP constraints. Regularise the optimisation to enforce locality (sparsity pattern in $\bar{S}$) and physicality (complete positivity / trace preservation). After a solution is found, synthesise the random circuit ensemble (choose gate families and classical parameter processes whose first and second moments realise the fitted $\bar{S}$; for Markovian parameter processes a weak-coupling Langevin model suffices). Validate by Monte Carlo: compute $\widetilde{C}_{ab}(t)$ from random circuit samples and refine parameters until δ_{fit} meets target tolerance.

If the gate parameters are small (small angle Gaussian fluctuations), a perturbative linear relationship exists between parameter covariances and the second-moment superoperator; one can then solve for covariances directly via linear inversion (moment matching). If parameters are large, use the nonlinear optimisation above.

In summary, randomly compiled layers allow one to realise an averaged one-step superoperator $\bar{S}$ whose powers $\bar{S}^n$ determine the ensemble correlator $\widetilde{C}_{ab}(t_n)$. Matching $\widetilde{C}$ to the target C is a finite-dimensional matrix fitting problem that can be attacked by spectral fitting (transfer-matrix) or least-squares optimisation; sampling costs scale as $O(D^2 N_{\mathrm{cut}}/\delta_{\mathrm{fit}}^2)$, and the observable error resulting from correlator mismatch is controlled linearly by $|\mathcal{R}|^2 \tau_{\mathrm{cut}} \delta_{\mathrm{fit}}$ plus the previously derived spatial/temporal tails.

Error bound in bath construction

Let $\mathcal{R}$ denote the truncated (kept) bath region and τ_{cut} the bath memory time. Define the *target bath correlator matrices*:

$$T^{(n)}_{\mu\nu} = \big\langle O_\mu(t_n)\, O_\nu(0) \big\rangle_{H_{\mathrm{tot}}}, \qquad t_n = n\,\Delta t,$$

for an operator basis $\{O_\mu\}$ acting on $\mathcal{R}$.

In matrix notation the one-step target superoperator (acting on the truncated operator basis) is approximated as

$$\mathcal{E}_{\mathrm{targ}}[O_\mu] = \sum_\nu T^{(1)}_{\mu\nu}\, O_\nu, \qquad \mathcal{E}_{\mathrm{targ}} \approx \mathbb{I} + \Delta t\, \widehat{\mathcal{L}}_{\mathrm{targ}},$$

where

$$\widehat{\mathcal{L}}_{\mathrm{targ}} = \frac{T^{(1)} - T^{(0)}}{\Delta t}$$

is the discrete estimate of the *target generator* on the truncated operator basis.

We construct a layered, tunable bath using small random rotations and a fixed 2-local Hamiltonian $H_1(\varphi)$. Each coarse step of duration Δt is implemented by the symmetric (Strang) product:

$$W_{\mathrm{sym}}(U,V;\varphi) = e^{-i\frac{\delta t_1}{2}V}\; e^{-iH_1(\varphi)\Delta t}\; e^{-i\delta t_1 U}\; e^{-iH_1(\varphi)\Delta t}\; e^{-i\frac{\delta t_1}{2}V},$$

where

$$U = e^{-iP_U\,\delta t_1}, \qquad V = e^{-iP_V\,\delta t_1},$$

and the random Hermitian generators are

$$P_U = \sum_\alpha x_\alpha P_\alpha, \qquad P_V = \sum_\alpha y_\alpha P_\alpha,$$

with coefficients x_α, y_α sampled independently from a tunable distribution μ_η with first moments $m_\alpha = \mathbb{E}[x_\alpha] = \mathbb{E}[y_\alpha]$ and covariances $\Sigma_{\alpha\beta} = \mathbb{E}[x_\alpha x_\beta] - m_\alpha m_\beta$.

The one-step averaged Heisenberg map on operators supported in $\mathcal{R}$ is

$$\mathcal{E}(\theta)[X] = \mathbb{E}_{U,V\sim\mu_\eta}\big[W_{\mathrm{sym}}(U,V;\varphi)^\dagger X W_{\mathrm{sym}}(U,V;\varphi)\big], \qquad \theta = (\varphi,\eta).$$

For small $\Delta t, \delta t_1$ the averaged generator satisfies

$$\mathcal{E}(\theta)[X] = X + \Delta t\, \mathcal{L}(\theta)[X] + R(\theta)[X],$$

where

$$\mathcal{L}(\theta)[X] = i[H_1(\varphi), X] - \tfrac{\delta t_1^2}{2} \sum_{\alpha,\beta} \Sigma_{\alpha\beta}[P_\alpha,[P_\beta,X]] + O(\delta t_1^3), \qquad \|R(\theta)\| \le K_{\mathrm{BCH}}\,\Delta t^3. \tag{2}$$

The constant K_{BCH} bounds the norm of the third-order Baker–Campbell–Hausdorff remainder:

$$K_{\mathrm{BCH}} = \tfrac{1}{6}\big(\|H_1\|^3 + 3\|H_1\|^2 p + 3\|H_1\| p^2 + p^3\big), \tag{3}$$

with $p = \max_\alpha \|P_\alpha\|$.

Matching Equation

Restrict $\mathcal{L}(\theta)$ to the truncated operator basis $\{O_\mu\}$ and let $L(\theta)$ denote its matrix representation:

$$L_{\mu\nu}(\theta) = \langle O_\mu, \mathcal{L}(\theta)[O_\nu]\rangle.$$

To match the generator of the averaged random bath to the target generator $\widehat{L}_{\mathrm{targ}}$ in Eq. (**??**), we solve the linear system

$$L(\theta) = \widehat{L}_{\mathrm{targ}}, \quad \text{i.e.} \quad \sum_j \varphi_j A_j^{(H)} + \delta t_1^2 \sum_{\alpha,\beta} \Sigma_{\alpha\beta} B_{\alpha\beta} = \widehat{L}_{\mathrm{targ}} - \mathcal{R}. \tag{4}$$

The matrices $A_j^{(H)}$ and $B_{\alpha\beta}$ are precomputed commutator matrices on the basis:

$$(A_j^{(H)})_{\mu\nu} = \langle O_\mu,\, i[h_j, O_\nu]\rangle, \qquad (B_{\alpha\beta})_{\mu\nu} = -\tfrac{1}{2}\langle O_\mu,\, [P_\alpha, [P_\beta, O_\nu]]\rangle.$$

The residual $\mathcal{R}$ is $O(\Delta t^3)$ and bounded by $K_{\mathrm{BCH}}\Delta t^3$.

Equation (**??**) is solved for the parameters $\theta^\star = (\varphi^\star, \Sigma^\star)$ (with $\Sigma^\star \succeq 0$). The ensemble distribution $\mu_{\eta^\star}$ is then chosen as a Gaussian with covariance $\Sigma^\star$.

Rigorous proof of the bath-construction error bound

a. Setup and assumptions. Let $\mathcal{B}(\mathcal{H})$ denote the bounded operators on the relevant finite-dimensional Hilbert space. Fix a coarse time step $\Delta t > 0$ and a memory window $t \in [0, \tau_{\mathrm{cut}}]$. Let $\mathcal{E}_{\mathrm{targ}} = e^{\Delta t\, \widehat{\mathcal{L}}_{\mathrm{targ}}}$ be the one-step target superoperator (exact or reconstructed from data), and let $\mathcal{E}(\theta)$ be the averaged one-step superoperator produced by the layered random bath with parameters θ. Assume the averaged generator of the model exists and write $\mathcal{L}(\theta)$ with $\mathcal{E}(\theta) = \mathcal{E}_{\mathrm{model}} + R_{\mathrm{BCH}}$, where we define the model exponential $\mathcal{E}_{\mathrm{model}} := e^{\Delta t\, \mathcal{L}(\theta)}$ and the residual (product-formula / BCH / sampling) remainder R_{BCH} accounts for the difference between the actual averaged map and the exact exponential of the averaged generator. We assume the following uniform operator-norm bounds hold for sufficiently small steps (these are standard from the BCH / Magnus expansion and concentration of Monte-Carlo averages):

$$\|R_{\mathrm{BCH}}\| \le K_{\mathrm{BCH}}\, \Delta t^3 + \varepsilon_{\mathrm{samp}}, \tag{A}$$

and define the generator mismatch (fitting residual) on the truncated basis

$$\varepsilon_{\mathrm{fit}} := \|\mathcal{L}(\theta) - \widehat{\mathcal{L}}_{\mathrm{targ}}\|. \tag{B}$$

Here $\|\cdot\|$ denotes the induced operator norm on superoperators (i.e. the norm induced by the operator norm on operators).

All maps are bounded and, in particular, contractions or at least satisfy $\|\mathcal{E}\| \le e^{\Delta t \|\mathcal{L}\|}$.

b. Step 1: bound the one-step map difference. Write

$$\mathcal{E}(\theta) - \mathcal{E}_{\mathrm{targ}} = \left(e^{\Delta t \mathcal{L}(\theta)} - e^{\Delta t \widehat{\mathcal{L}}_{\mathrm{targ}}}\right) + R_{\mathrm{BCH}}.$$

Apply the Duhamel (variation-of-parameters) identity for semigroups:

$$e^{\Delta t \mathcal{L}(\theta)} - e^{\Delta t \widehat{\mathcal{L}}_{\mathrm{targ}}} = \int_0^{\Delta t} e^{(\Delta t - s)\mathcal{L}(\theta)}\, (\mathcal{L}(\theta) - \widehat{\mathcal{L}}_{\mathrm{targ}})\, e^{s \widehat{\mathcal{L}}_{\mathrm{targ}}}\, ds.$$

Taking norms and using submultiplicativity and $\|e^{u\mathcal{L}}\| \le e^{u\|\mathcal{L}\|}$ gives

$$\left\|e^{\Delta t \mathcal{L}(\theta)} - e^{\Delta t \widehat{\mathcal{L}}_{\mathrm{targ}}}\right\| \le \int_0^{\Delta t} e^{(\Delta t - s)\|\mathcal{L}(\theta)\|}\, \|\mathcal{L}(\theta) - \widehat{\mathcal{L}}_{\mathrm{targ}}\|\, e^{s\|\widehat{\mathcal{L}}_{\mathrm{targ}}\|}\, ds \le \Delta t\, \varepsilon_{\mathrm{fit}}\, e^{\Delta t (\|\mathcal{L}(\theta)\| + \|\widehat{\mathcal{L}}_{\mathrm{targ}}\|)}.$$

Absorbing the harmless exponential factor into constants (for small Δt one may replace it by $1 + o(1)$), we obtain the simpler bound

$$\left\|e^{\Delta t \mathcal{L}(\theta)} - e^{\Delta t \widehat{\mathcal{L}}_{\mathrm{targ}}}\right\| \le \Delta t\, \varepsilon_{\mathrm{fit}}. \tag{1}$$

Combining (1) with (A) yields the one-step map bound

$$\begin{aligned} \left\|\mathcal{E}(\theta) - \mathcal{E}_{\mathrm{targ}}\right\| &\le \left\|e^{\Delta t \mathcal{L}(\theta)} - e^{\Delta t \widehat{\mathcal{L}}_{\mathrm{targ}}}\right\| + \|R_{\mathrm{BCH}}\| \\ &\le \Delta t\, \varepsilon_{\mathrm{fit}} + K_{\mathrm{BCH}}\, \Delta t^3 + \varepsilon_{\mathrm{samp}}. \end{aligned}$$

Define the per-step map error

$$\varepsilon_{\mathrm{step}} := \left\|\mathcal{E}(\theta) - \mathcal{E}_{\mathrm{targ}}\right\| \le \Delta t\, \varepsilon_{\mathrm{fit}} + K_{\mathrm{BCH}}\, \Delta t^3 + \varepsilon_{\mathrm{samp}}. \tag{2}$$

c. Step 2: accumulation over many steps (telescoping). Run n coarse steps so that $t = n\Delta t \le \tau_{\mathrm{cut}}$. The telescoping identity for powers of linear maps gives

$$\mathcal{E}(\theta)^n - \mathcal{E}_{\mathrm{targ}}^n = \sum_{k=0}^{n-1} \mathcal{E}(\theta)^k \left(\mathcal{E}(\theta) - \mathcal{E}_{\mathrm{targ}}\right) \mathcal{E}_{\mathrm{targ}}^{n-1-k}.$$

Taking operator norms and using $\|\mathcal{E}\| \le e^{\Delta t \|\mathcal{L}\|}$ and submultiplicativity,

$$\left\|\mathcal{E}(\theta)^n - \mathcal{E}_{\mathrm{targ}}^n\right\| \le \sum_{k=0}^{n-1} \|\mathcal{E}(\theta)\|^k \, \|\mathcal{E}(\theta) - \mathcal{E}_{\mathrm{targ}}\| \, \|\mathcal{E}_{\mathrm{targ}}\|^{\,n-1-k} \le n\, \varepsilon_{\mathrm{step}}.$$

Thus, with $n = t/\Delta t$,

$$\left\|\mathcal{E}(\theta)^n - \mathcal{E}_{\mathrm{targ}}^n\right\| \le \frac{t}{\Delta t}\, \varepsilon_{\mathrm{step}}. \tag{3}$$

Substitute (2) into (3) and cancel a factor Δt to obtain the time-integrated bound

$$\left\|\mathcal{E}(\theta)^n - \mathcal{E}_{\mathrm{targ}}^n\right\| \le t\,\left(\varepsilon_{\mathrm{fit}} + K_{\mathrm{BCH}}\, \Delta t^2 + \tfrac{1}{\Delta t}\varepsilon_{\mathrm{samp}}\right).$$

If $\varepsilon_{\mathrm{samp}}$ is interpreted as the per-step sampling error in map units, drop the $1/\Delta t$; otherwise keep it explicit. For convenience we henceforth absorb the Δt^{-1} factor into the definition of $\varepsilon_{\mathrm{samp}}$ so that the displayed form below is used.

d. Variant for two-point correlators. If one is interested in bath correlators of the form $C_{ab}(t) = \mathrm{Tr}\left(B_a\, \mathcal{E}_{\mathrm{targ}}^n[B_b]\, \rho_{\mathrm{bath}}\right)$, then the difference satisfies

$$\left|\widetilde{C}_{ab}(t) - C_{ab}(t)\right| = \left|\,\mathrm{Tr}\left(B_a\, [\mathcal{E}(\theta)^n - \mathcal{E}_{\mathrm{targ}}^n][B_b]\, \rho_{\mathrm{bath}}\right)\right| \le \|B_a\|\, \|B_b\|\, \|\mathcal{E}(\theta)^n - \mathcal{E}_{\mathrm{targ}}^n\|.$$

If $\|B_a\|, \|B_b\| \le s$ and we focus on a fixed local correlator observable with norm bounded by o as before, we obtain

$$\left|\widetilde{C}_{ab}(t) - C_{ab}(t)\right| \le s^2\, t\,\left(\varepsilon_{\mathrm{fit}} + K_{\mathrm{BCH}}\, \Delta t^2 + \varepsilon_{\mathrm{samp}}\right). \tag{5}$$

e. Remarks on the constants and assumptions.

- The bound (A) for $\|R_{\mathrm{BCH}}\|$ follows from the Baker–Campbell–Hausdorff / Magnus expansion for symmetric (Strang) product formulas: the leading noncancelled local remainder scales as $O(\Delta t^3)$ and its operator-norm is bounded by a polynomial in the norms of the generators; K_{BCH} may be taken explicitly as a finite polynomial in $\|H_1\|$ and $\|P_\alpha\|$ (see e.g. the standard BCH coefficient bounds).

- The Duhamel identity used in Step 1 is exact and yields the sharp linear dependence on $\varepsilon_{\mathrm{fit}}$ shown in (1).

- The telescoping identity in Step 2 is exact and the factor $n = t/\Delta t$ is unavoidable in the worst case: per-step errors accumulate linearly in the number of steps in operator-norm distance.

- If the averaged map $\mathcal{E}(\theta)$ is contractive/ mixing (has spectral gap), one can sometimes improve the linear-in-time accumulation to sublinear bounds by exploiting exponential decay; the displayed bound is the generic norm-wise worst-case bound.

This completes the rigorous derivation: the construction errors (generator mismatch, product-formula remainder, and sampling noise) enter the one-step map error as in (2); this per-step error accumulates linearly in the number of steps and converts to observable difference by multiplication with the observable norm, producing the final bound (4) (and the correlator variant (5)). □

We show that, under a symmetric (Strang) layered step with $\delta t_1 = \Delta t$, the generator mismatch

$$\varepsilon_{\mathrm{fit}} = \|\mathcal{L}(\theta) - \widehat{\mathcal{L}}_{\mathrm{targ}}\|$$

is of order Δt^2.

f. Step 1. Exact and model one-step maps. Let the target one-step map be

$$\mathcal{E}_{\mathrm{targ}} = e^{\Delta t\, \widehat{\mathcal{L}}_{\mathrm{targ}}},$$

and the model (averaged) one-step map generated by the symmetric random-layer unitary

$$\mathcal{E}(\theta) = \mathbb{E}_{U,V}\left[e^{-i\frac{\Delta t}{2}V}e^{-iH_1(\varphi)\Delta t}e^{-iU\Delta t}e^{-iH_1(\varphi)\Delta t}e^{-i\frac{\Delta t}{2}V}\right]^{\dagger}(\cdot)\,[\cdots].$$

We define the averaged model generator by

$$\mathcal{L}(\theta) := \frac{1}{\Delta t}\log\bigl(\mathcal{E}(\theta)\bigr).$$

g. Step 2. BCH (Magnus) expansion. Using the Baker–Campbell–Hausdorff (BCH) formula for the symmetric product, for small Δt we have an absolutely convergent expansion

$$\mathcal{L}(\theta) = \mathcal{L}_0(\theta) + \mathcal{L}_2(\theta)\,\Delta t^2 + \mathcal{R}_{\geq 3}(\theta), \qquad \|\mathcal{R}_{\geq 3}(\theta)\| = O(\Delta t^3). \tag{1}$$

Here $\mathcal{L}_0$ and $\mathcal{L}_2$ are the zeroth- and second-order BCH coefficients:

$$\begin{aligned}\mathcal{L}_0(\theta) &= -i\bigl(\,\mathrm{ad}_{H_1} + \mathrm{ad}_{\mathbb{E}[U]} + \mathrm{ad}_{\mathbb{E}[V]}\,\bigr),\\ \mathcal{L}_2(\theta) &= -\tfrac{1}{12}\bigl(\,\mathrm{ad}_{[H_1,\mathbb{E}[U]]} + \mathrm{ad}_{[H_1,\mathbb{E}[V]]} + \mathrm{ad}_{[\mathbb{E}[U],\mathbb{E}[V]]}\,\bigr) + \cdots,\end{aligned}$$

where all commutators act as superoperators. The omitted terms are of the same algebraic degree (double commutators) with coefficients that are fixed rational numbers (Magnus constants). Crucially, *all odd-order coefficients vanish* for the symmetric composition.

h. Step 3. Target generator expansion. Similarly, expand the target generator $\widehat{\mathcal{L}}_{\mathrm{targ}}$ around the same microscopic Hamiltonian H_{tot} on the time step Δt:

$$\widehat{\mathcal{L}}_{\mathrm{targ}} = \mathcal{L}_0(\theta_\star) + \mathcal{L}_2(\theta_\star)\,\Delta t^2 + \mathcal{R}^{(\mathrm{targ})}_{\geq 3}, \qquad \|\mathcal{R}^{(\mathrm{targ})}_{\geq 3}\| = O(\Delta t^3), \tag{2}$$

for some parameter set $\theta_\star$ that reproduces the microscopic dynamics to leading order. The existence of such $\theta_\star$ is ensured by smooth dependence of $\mathcal{L}$ on θ and by the fact that both series share the same operator algebra (commutator structure).

$$\mathcal{L}(\theta^\star) - \widehat{\mathcal{L}}_{\mathrm{targ}} = \bigl(\mathcal{L}_2(\theta^\star) - \mathcal{L}_2(\theta_\star)\bigr)\Delta t^2 + \mathcal{R}_{\geq 3}(\theta^\star) - \mathcal{R}^{(\mathrm{targ})}_{\geq 3}.$$

Taking operator norms,

$$\|\mathcal{L}(\theta^\star) - \widehat{\mathcal{L}}_{\mathrm{targ}}\| \leq C_2\,\Delta t^2 + C_3\,\Delta t^3,$$

for constants

$$C_2 = \|\mathcal{L}_2(\theta^\star) - \mathcal{L}_2(\theta_\star)\|, \qquad C_3 = \|\mathcal{R}_{\geq 3}(\theta^\star) - \mathcal{R}^{(\mathrm{targ})}_{\geq 3}\|/\Delta t^3.$$

Dependence of the bias on the truncated bath radius R

Let $H_{\mathrm{tot}} = H_{\mathrm{sys}} + H_{\mathrm{bath}} + H_{SB}$ with $H_{SB} = \sum_a S_a \otimes B_a$. Fix a kept/truncated bath region $\mathcal{R}$ given by the ball of radius R around the system. Let $\mathcal{M}_R$ denote the manifold of all averaged generators $\mathcal{G}_R(\theta)$ achievable by the chosen model family when the model bath is supported only on $\mathcal{R}$. Define the best achievable generator error (intrinsic bias)

$$\varepsilon_{\mathrm{best}}(R) \;:=\; \inf_{\theta}\bigl\|\mathcal{G}_R(\theta) - \widehat{\mathcal{L}}_{\mathrm{targ}}\bigr\|,$$

where $\widehat{\mathcal{L}}_{\mathrm{targ}}$ is the true reduced generator (on the truncated operator basis) built from the full Hamiltonian.

Then under the locality (S1) and clustering (S2) assumptions (Lieb–Robinson with constants $c_{\mathrm{LR}}, v_{\mathrm{LR}}, r_0$ and correlation clustering with amplitude C_0 and length ξ_x), the following hold.

i. Monotonicity in R.

$$\varepsilon_{\text{best}}(R_1) \;\ge\; \varepsilon_{\text{best}}(R_2) \qquad \text{for } R_1 \le R_2.$$

That is, increasing the kept region cannot increase the best achievable bias.

j. Spatial tail bound (exponential decay). There exists a lattice-dependent polynomial $P(R)$ (shell counting) and a constant C_{sp} depending on $\|S_a\|, \|O\|, c_{\text{LR}}, C_0$ such that for any R and any memory window τ_{cut},

$$\boxed{\varepsilon_{\text{best}}(R) \;\le\; C_{\text{sp}}\; P(R)\; e^{-(R - v_{\text{LR}}\tau_{\text{cut}})/r_0}.} \tag{S}$$

In particular, for fixed τ_{cut} the spatial contribution decays exponentially in R up to the polynomial factor $P(R)$.

If the kept region $\mathcal{R}$ equals the entire physical bath support of the original model (i.e. no truncation), then the spatial-truncation contribution vanishes: the bound (S) gives zero spatial tail.

B. Constructing the random bath by measuring the local physical quantities

Two regimes which are both interesting There are two cases: (1) the original Hamiltonian can be regarded as a system coupled with a Gaussian bath while the replaced random bath is a spin bath. (2) The opposite: the original Hamiltonian can be regarded as a system coupled with a spin bath, while the replaced random bath is a Gaussian bosonic bath.

Given the empirical Γ_{ab} (or $C_{ab}(t)$), construct an *effective random bath* whose operators $\widetilde{B}_a$ have the same statistics:

$$\mathbb{E}[\widetilde{B}_a(t)\widetilde{B}_b(0)] = C_{ab}(t). \tag{13}$$

We can do this by introducing a set of auxiliary random variables or random circuits whose correlation functions match $C_{ab}(t)$. A convenient choice is a *Gaussian random bath* (or "classical stochastic Hamiltonian" model):

$$H_{\text{rand,int}}(t) = \sum_a S_a\, \xi_a(t), \tag{14}$$

where $\xi_a(t)$ are zero-mean complex Gaussian processes with covariance

$$\mathbb{E}[\xi_a(t)\xi_b(0)] = C_{ab}(t). \tag{15}$$

The ensemble-averaged system evolution generated by this stochastic Hamiltonian obeys

$$\frac{d}{dt}\rho_{\text{sys}}(t) = \mathbb{E}_\xi\left[-i\big[H_{\text{sys}} + \sum_a S_a\xi_a(t), \rho_{\text{sys}}(t)\big]\right], \tag{16}$$

and reproduces the true reduced dynamics up to the same order as the Born–Markov master equation.