# BACKWARD KOLMOGOROV TRANSPORT: SAMPLING INVARIANT LAWS OF REVERSIBLE DIFFUSIONS FROM TRAJECTORY DATA

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

We consider a reversible diffusion with unknown drift, observed through trajectories, and construct samplers of its invariant law. The trajectories contain samples of this law, but a sampler built from them by density or score estimation inherits the errors of these estimates, in particular on the barriers between metastable states. The trajectories also contain the dynamics, and for a reversible diffusion the ratio between the law of the process and its invariant law solves the backward Kolmogorov equation, so that the eigenpairs of the generator propagate it in closed form. Backward Kolmogorov transport (BKT) evaluates this ratio with eigenpairs estimated from the trajectories and coefficients computed once from the initial particles, and moves the particles along the Wasserstein gradient flow of the Kullback-Leibler divergence. The particles do not interact, BKT defines a transport map, and no density, score or drift is estimated. With exact eigenpairs BKT transports the spectral projection of the initial law exactly when this projection is positive. For estimated eigenpairs we prove a transport identity in which truncation enters only through the initial law and estimation only through the residual of the eigenpairs, and a local stability bound in the 2-Wasserstein distance. With coefficients computed from the transported particles, the sampling error of the retained modes cancels, and a central limit theorem identifies the asymptotic variance. Experiments on Ornstein-Uhlenbeck processes, multi-well potentials, separable products and alanine dipeptide, including comparisons with particle methods based on density estimates, agree with these results, and with an accurate spectrum the errors of BKT reach or fall below those of independent samples.

## 1 INTRODUCTION

Consider a diffusion process with unknown drift that is observed through trajectories, as in molecular dynamics, stochastic partial differential equations and climate models, where the object of interest is the invariant law $\pi$ of the process. The theory of Koopman and transfer operators provides estimators of the generator from trajectories, with error bounds (Williams et al., 2015; Klus et al., 2020; Noé & Nüske, 2013; Wu & Noé, 2020; Kostic et al., 2023; 2024; Colbrook et al., 2023; Kostic et al., 2022; Meanti et al., 2023). A stationary trajectory also contains samples of $\pi$, and one may ask what a sampler adds to them. The recorded samples are finite in number, correlated and fixed, whereas a sampler generates any number of new samples of $\pi$, from a prescribed initial law, along the relaxation path of the process and without further simulation, as generative models produce new samples from their training data (Song et al., 2021; Lipman et al., 2023). The question is how the trajectory data should be used for this purpose.

The Wasserstein gradient flow of the Kullback-Leibler (KL) divergence transports any initial law to $\pi$ along the Fokker-Planck path of the diffusion (Jordan et al., 1998; Otto, 2001; Ambrosio et al., 2008), and its velocity is the negative gradient of the logarithm of the ratio between the current law and $\pi$. Particle methods estimate this ratio statistically during the evolution, by kernel density estimation (Degond & Mustieles, 1990; Russo, 1990; Lions & Mas-Gallic, 2001; Carrillo et al., 2019; Chertock, 2017; Klebanov, 2024) or by a fitted score (Maoutsa et al., 2020; Boffi & Vanden-Eijnden, 2023; Ilin et al., 2026; Zhou et al., 2026), and they require the drift or the score of the target,



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.

as does the weak generative sampler of Cai et al. (2026). From data, both the target and the current law of the particles must then be estimated from samples. Such estimates are biased by smoothing and noisy where the density is small, in particular on the barriers between metastable states through which the mass must be transported, and their errors do not decrease along the flow. A trajectory contains more than samples of $\pi$, namely the dynamics of the process. For a reversible diffusion the ratio between the current law and $\pi$ solves the backward Kolmogorov equation, so that the dynamics propagate it in closed form, through the eigenfunction expansion of the Koopman semigroup, whose spectrum encodes the relaxation rates and the metastable structure (Risken, 1989; Bakry et al., 2014; Pavliotis, 2014). This spectrum is estimated in weak form, by averages over the data, through the Dirichlet form $\int |\nabla f|^2 d\pi = \lim_{\tau\to 0}(2\tau)^{-1}\mathbb{E}_\pi|f(X_\tau) - f(X_0)|^2$ of the generator, which measures how observables decorrelate along the trajectories, and no density needs to be estimated. For the dihedral angles of alanine dipeptide, a kernel particle method built on the frames of a molecular dynamics trajectory stops at about three times the error of direct sampling, whereas the transport built on the spectrum estimated from the same frames reaches it (Figure 1).

With eigenpairs estimated from trajectories, this expansion underlies density forecasting (Berry et al., 2015; Giannakis, 2019; Zhao & Jiang, 2023; Wang & Araki, 2026) and the logarithmic-gradient control fields of rare-event simulation (Zhang et al., 2022a). The transport of a measure along the negative logarithmic gradient of the evolved ratio is the heat-flow map of Kim & Milman (2012), whose Lipschitz constant and stability are studied in Neeman (2022); Mikulincer & Shenfeld (2023); Brigati & Pedrotti (2025); Chewi et al. (2026b;a), and density-equalizing maps compute it for a uniform target with the Fourier expansion of the heat kernel (Gastner & Newman, 2004) or with the Laplace-Beltrami operator of a surface (Choi & Rycroft, 2018). In diffusion models the expansion of a known operator gives the score of a noising process that is run backward toward the data (Lou et al., 2023; Shen et al., 2025; Scarvelis et al., 2025), with the moments of the data computed once in Khoo et al. (2025), and Koopman representations linearise trained generative flows (Turan et al., 2026). In these works the operator is known, or the expansion is not used to sample the invariant law. Estimating the expansion from trajectories and using it to transport particles raises a question that does not arise in density forecasting, namely how the accuracy of the eigenpairs transfers to the particles, since the truncated ratio must remain accurate and positive along the path and its logarithmic gradient amplifies the error.

The samplers built so far on an estimated spectrum, the diffusion map particle systems of Li & Marzouk (2025) and the Koopman spectral Wasserstein gradient descent of Xu et al. (2025a), are data-driven versions of the Laplacian-adjusted Wasserstein gradient descent (LAWGD) of Chewi et al. (2020), which kernelises the Wasserstein gradient flow of the $\chi^2$ divergence as Stein variational gradient descent does (Liu & Wang, 2016; Teolis & de Hoop, 2026). Their particles interact, since the velocity of each particle depends on all current positions through moments re-estimated at every step. In backward Kolmogorov transport (BKT) the velocity field is determined once by the source and the spectrum, the particles do not interact, and the method defines a transport map. Each particle solves its own equation, the sampling error of the source is damped by the semigroup, and the error analysis concerns a single flow map, without propagation of chaos, as for probability-flow samplers and flow matching (Song et al., 2021; Chen et al., 2023b;a; Huang et al., 2025; Deveney et al., 2025; Shen et al., 2022; Kwon et al., 2022; Lipman et al., 2023; Albergo & Vanden-Eijnden, 2023; Benton et al., 2024).

The paper makes three contributions. First, we introduce BKT (Algorithm 1), which is, to our knowledge, the first method that moves non-interacting particles along the Fokker-Planck path of a process with the spectrum of its generator estimated from trajectories, and we show that with exact eigenpairs it transports a positive spectral projection of the source exactly (Proposition 3.3). Second, for estimated eigenpairs we prove a transport identity (Proposition 4.1) in which truncation enters only through the initial law, estimation only through the residual of the eigenpairs and the safeguards through a flux, a bound in total variation in which these terms carry no stability constant (Corollary 4.2), and, for truncations that change sign, where global stability estimates for probability flows fail, a bound in the 2-Wasserstein distance by localisation on validity regions (Theorem 4.4). Third, for coefficients computed from the transported particles, we show that the sampling error of the source cancels on the retained modes (Proposition 4.5) and prove a central limit theorem (Theorem 4.6) in which the asymptotic variance is that of an independent sample for population coefficients, is larger for coefficients from an independent sample, and for same-sample coefficients is that of the component orthogonal to the retained modes, at most of order $\lambda_{r+1}^{-1}$ for exact eigenpairs

up to terms due to the finite horizon, the truncation and the safeguards. This is the mechanism of moment matching in Monte Carlo integration (Glasserman, 2004), which also underlies kernel herding, support points, kernel thinning and the fixed points of Stein variational gradient descent (Chen et al., 2010; Mak & Joseph, 2018; Dwivedi & Mackey, 2024; Liu & Wang, 2018), and it explains the errors below independent sampling in Section 5.

The subject of the paper is the transport and its analysis, and the eigenpairs are taken as input. In high dimension, and in particular for many metastable states, the estimation of the slow eigenpairs is the main difficulty. It is a long-standing problem of transfer operator methods (Schütte et al., 2023) that is separate from the transport, and since the bounds of Section 4 hold for arbitrary estimated eigenpairs, the kernel and neural spectral estimators developed for it can be combined with BKT without change. Section 5 discusses where the dimension enters.

## 2 Setting and notation

Let $\pi$ be a probability measure on $\mathbb{R}^d$ with density proportional to $e^{-V}$, and let $X$ solve the stochastic differential equation

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2}\,dW_t, \tag{1}$$

whose invariant law is $\pi$. The generator of $X$ is $-L$ with $L \coloneqq -\Delta + \nabla V \cdot \nabla$, a nonnegative self-adjoint operator on $L^2(\pi)$ whose Dirichlet form is $\langle f, Lg\rangle_\pi = \int \nabla f \cdot \nabla g\, d\pi$, and $P_t \coloneqq e^{-tL}$ is the associated symmetric Markov semigroup, also called the Koopman semigroup, $(P_t g)(x) = \mathbb{E}[g(X_t) \mid X_0 = x]$ (Bakry et al., 2014). We assume that $L$ has discrete spectrum $0 = \lambda_0 < \lambda_1 \le \lambda_2 \le \cdots$ with orthonormal eigenfunctions $\varphi_k$, $\varphi_0 \equiv 1$, and write $\Pi_r$ for the orthogonal projection onto the span of $\varphi_0, \dots, \varphi_r$. Since $P_t$ is self-adjoint, for $\mu_0 \ll \pi$ with $\rho_0 = d\mu_0/d\pi \in L^2(\pi)$ the law $\mu_t$ of $X_t$ satisfies $d\mu_t/d\pi = P_t\rho_0$, that is

$$\rho_t \coloneqq \frac{d\mu_t}{d\pi} = 1 + \sum_{k\ge 1} e^{-\lambda_k t} c_k \varphi_k, \qquad c_k \coloneqq \mathbb{E}_{\mu_0}[\varphi_k], \tag{2}$$

and $\chi^2(\mu_t\|\pi) = \sum_{k\ge1} e^{-2\lambda_k t} c_k^2$ (Proposition A.1).

The family $(\mu_t)$ solves the Fokker-Planck equation

$$\partial_t \mu_t = \operatorname{div}\big(\mu_t \nabla \log(\mu_t/\pi)\big), \tag{3}$$

which is the Wasserstein gradient flow of the KL divergence $\mathrm{KL}(\mu\|\pi) \coloneqq \int \log(d\mu/d\pi)\, d\mu$, $\mu \ll \pi$ (Jordan et al., 1998), and $\mu_t$ is the law at time $t$ of the solution of the ordinary differential equation

$$\dot{x} = v_t(x), \qquad v_t = -\nabla \log \rho_t, \qquad \rho_t = d\mu_t/d\pi, \tag{4}$$

started from $\mu_0$ (Ambrosio et al., 2008, Ch. 8), a construction also used in generative modelling (Rozen et al., 2021). We denote by $\Phi_t$ the flow map of (4), so that $\mu_t = \Phi_t \# \mu_0$. Given pairs $(X_0, X_\tau)$ of a stationary trajectory, extended dynamic mode decomposition and its variational, generator and semigroup versions (gEDMD, SDMD) (Williams et al., 2015; Noé & Nüske, 2013; Klus et al., 2018; Wu & Noé, 2020; Klus et al., 2020; Kostic et al., 2024; Xu et al., 2025b), as well as diffusion maps built from samples of $\pi$ (Coifman & Lafon, 2006; Berry & Harlim, 2016), return approximations $(\hat\lambda_k, \hat\varphi_k, \nabla\hat\varphi_k)_{k\le r}$ of the leading eigenpairs as linear combinations of a dictionary of functions. If the trajectory is not stationary, the same estimators apply after the reweighting of Wu et al. (2017); Nüske et al. (2017). The accuracy of an approximate eigenfunction is measured in $L^2(\pi)$ and in the energy norm $\|f\|_{\dot H^1}^2 \coloneqq \int |\nabla f|^2 d\pi$, and error bounds are given in Kostic et al. (2023); Nüske et al. (2023); Kostic et al. (2024).

**Notation.** We write $\|f\|$ for the norm of $L^2(\pi)$. For $g$ with $\langle g, 1\rangle_\pi = 0$ we set $\|g\|_{\dot H^{-1}}^2 \coloneqq \sum_{k\ge1} \langle g, \varphi_k\rangle_\pi^2/\lambda_k$, and $W_2(\mu,\pi) \le 2\|d\mu/d\pi - 1\|_{\dot H^{-1}}$ for $\mu \ll \pi$ (Peyre, 2018). We denote by $\mathcal{P}_2(\mathbb{R}^d)$ the probability measures with finite second moment, by $W_2$ the 2-Wasserstein distance, by $\Phi\#\nu$ the image of $\nu$ under $\Phi$, by $\nu_0 \coloneqq M^{-1}\sum_j \delta_{x_j}$ the empirical measure of $M$ samples of the source and by $B$ a compact convex domain to which the particles are confined, and we set $\|\mu - \nu\|_{\mathrm{TV}} \coloneqq \sup_{|f|\le1} \int f\, d(\mu-\nu)$ and $\mathrm{Clip}_v(u) \coloneqq u \min\{1, v/|u|\}$ for $u \ne 0$, with $\mathrm{Clip}_v(0) = 0$. The assumptions are stated in Appendix A.

**Algorithm 1** Backward Kolmogorov transport

**Require:** eigenpairs $(\hat{\lambda}_k, \hat{\varphi}_k, \nabla\hat{\varphi}_k)_{k\le r}$ estimated from trajectories, source samples $x_1, \dots, x_M$, horizon $T$, step $h$

1: $\hat{c}_k \leftarrow M^{-1}\sum_j \hat{\varphi}_k(x_j)$ for $1 \le k \le r$, $\hat{c}_0 \leftarrow 1$, and $X^j \leftarrow x_j$
2: **for** $t = 0, h, \dots, T-h$ **do**
3: $\quad \hat{v}_s \leftarrow -\nabla\hat{\rho}_s/\hat{\rho}_s$ with $\hat{\rho}_s = \sum_{k=0}^{r} e^{-\hat{\lambda}_k s}\hat{c}_k\hat{\varphi}_k$ for $s \in \{t, t+\frac{h}{2}, t+h\}$ (Remark 3.1)
4: $\quad$ advance each $X^j$ by one fourth-order Runge-Kutta step for $\dot{x} = \hat{v}_s(x)$ and reflect at $\partial B$
5: **end for**
6: **return** $\hat{\mu}_T = M^{-1}\sum_j \delta_{X^j}$ and the spectral diagnostic $\hat{\chi}^2(t) = \sum_{k=1}^{r} e^{-2\hat{\lambda}_k t}\hat{c}_k^2$

## 3 Backward Kolmogorov transport

By (2) the velocity of the KL flow is $v_t = -\nabla \log \rho_t$ with $\rho_t = P_t\rho_0$, and the density ratio is determined by the eigenpairs and by the averages of the eigenfunctions over the source. Let $(\hat{\lambda}_k, \hat{\varphi}_k)_{k\le r}$ be approximations of the first $r$ nonconstant eigenpairs, with $\hat{\lambda}_0 = 0$, $\hat{\varphi}_0 = 1$ and $0 < \hat{\lambda}_1 \le \cdots \le \hat{\lambda}_r$, let $x_1, \dots, x_M$ be independent samples of the source $\mu_0$, an initial law of the particles that may be chosen freely, and let $T > 0$ be the time horizon. The approximate density ratio and the velocity are defined by

$$\hat{\rho}_t \coloneqq \sum_{k=0}^{r} e^{-\hat{\lambda}_k t}\hat{c}_k\hat{\varphi}_k, \qquad \hat{c}_k \coloneqq \int \hat{\varphi}_k \, d\nu_0, \qquad \hat{v}_t \coloneqq -\nabla \log \hat{\rho}_t. \tag{5}$$

The coefficient $\hat{c}_k$ is the empirical counterpart of $c_k$ in (2), computed from the particles that are subsequently transported. Each particle $j$ follows the ordinary differential equation

$$\dot{X}_t^j = \hat{v}_t(X_t^j), \qquad X_0^j = x_j, \qquad t \in [0, T], \tag{6}$$

with reflection at the boundary of $B$, and the output is the empirical measure $\hat{\mu}_T$ of the transported particles $X_T^1, \dots, X_T^M$. The velocity $\hat{v}_t$ depends on the particles only through the coefficients $\hat{c}_k$, which are computed once at $t = 0$. For $t > 0$ the particles therefore do not interact, and (6) is a system of $M$ decoupled equations, which Algorithm 1 integrates with the classical fourth-order Runge-Kutta method, although any integrator may be used. The spectral summation costs $O(Mrd)$ operations per time step, in addition to the evaluation of the eigenfunctions and their gradients.

**Coefficient regimes.** The average of $\hat{\varphi}_k$ over the source can be computed in three ways. In regime (P) it is computed exactly, $\bar{c}_k = \mathbb{E}_{\mu_0}[\hat{\varphi}_k]$, by quadrature when $\mu_0$ is known in closed form. In regime (I) it is computed from an independent sample $y_1, \dots, y_{M'}$ of $\mu_0$, $\check{c}_k = M'^{-1}\sum_j \hat{\varphi}_k(y_j)$, and in regime (S) from the transported particles themselves, $\hat{c}_k = M^{-1}\sum_j \hat{\varphi}_k(x_j)$, as in (5) and Algorithm 1. In all regimes the trajectory data used for the eigenpairs are independent of the transported sample. Theorem 4.4 covers (P) and (I), in which the field is independent of the transported particles. Regime (S) is the regime of the experiments, and Proposition 4.5 and Theorem 4.6 describe the cancellation of the sampling error on the retained modes.

*Remark* 3.1. A truncated expansion need not be positive, and the velocity in (5) is defined only where $\hat{\rho}_t > 0$. The computations therefore use the velocity

$$\mathrm{Clip}_{v_{\max}}\Big(-\frac{\nabla\hat{\rho}_t}{\max\{\hat{\rho}_t, \varepsilon_0\}}\Big)$$

with a threshold $\varepsilon_0 > 0$ and a bound $v_{\max}$. This field coincides with (5) where $\hat{\rho}_t \ge \varepsilon_0$ and $|\nabla \log \hat{\rho}_t| \le v_{\max}$, and it is continuous, locally Lipschitz and bounded by $v_{\max}$ on $\mathbb{R}^d$. On the set where it differs from (5), Theorem 4.4 uses only this bound and controls the mass of this set. The threshold is inactive at large times (Appendix A), and the safeguards act on few paths (Appendix E).

**Coefficients, diagnostic and source.** The coefficients decay at the rates of the dynamics, and the slow modes describe the transitions between metastable states. This propagation rule, known from density forecasting (Zhao & Jiang, 2023; Shen et al., 2025), here determines the velocity of the particles and damps the sampling error of the high modes, whereas recomputing the coefficients from the particles at each step, as in the data-driven $\chi^2$ flow of Xu et al. (2025a) without its preconditioning,

feeds it back. Along the exact flow the two rules coincide, since $\mathbb{E}_{\mu_t}[\varphi_k] = e^{-\lambda_k t} c_k$. The diagnostic $\hat{\chi}^2(t)$ equals $\chi^2(\pi\hat{\rho}_t \| \pi)$ when $\hat{\rho}_t \geq 0$ and the eigenfunctions are orthonormal. We call the source admissible when $\rho_0$ is bounded, and for a double well a source of compact support is admissible while a Gaussian source with full support is not.

**Lemma 3.2** (Residual identity). *Let $\hat{q}_t \coloneqq \pi\hat{\rho}_t$, $R_t \coloneqq \sum_{k \leq r} e^{-\hat{\lambda}_k t} \hat{c}_k (L - \hat{\lambda}_k)\hat{\varphi}_k$, and $\Omega_t \coloneqq \{x \in \mathbb{R}^d \mid \hat{\rho}_t(x) > 0\}$ the positivity set of the truncated ratio. On $\Omega_t$,*

$$\partial_t \hat{q}_t + \operatorname{div}\big(\hat{q}_t \, (-\nabla \log \hat{\rho}_t)\big) = \pi R_t.$$

If the eigenpairs are exact, then $R \equiv 0$ for any coefficients, and the truncated spectral density solves the continuity equation on $\Omega_t$. When the truncated ratio is positive, the conclusion holds on the whole space and for the transport map.

**Proposition 3.3** (Exact eigenpairs). *Let the eigenpairs be exact, with $\varphi_k \in C^2$, and let $\hat{\rho}_0 = 1 + \sum_{k=1}^{r} a_k \varphi_k$ be positive for some $a \in \mathbb{R}^r$. Then $\hat{\rho}_t = P_t \hat{\rho}_0$ and $\inf \hat{\rho}_t \geq \inf \hat{\rho}_0$ for $t \geq 0$, and the flow of the field* (5) *satisfies $\hat{\Phi}_{0,t} \# (\pi\hat{\rho}_0) = \pi\hat{\rho}_t$, so that BKT is the KL gradient flow started from $\pi\hat{\rho}_0$. For population coefficients, $a_k = c_k$,*

$$\|\hat{\Phi}_{0,T} \# \mu_0 - \pi\|_{\mathrm{TV}} \leq \|(I - \Pi_r)\rho_0\|_{L^1(\pi)} + e^{-\lambda_1 T} \|\Pi_r \rho_0 - 1\|.$$

Truncation thus replaces the source by its spectral projection, a classical property of eigenfunction expansions (Risken, 1989; Pavliotis, 2014) which here holds for the transport map, and by the maximum principle the threshold of Remark 3.1 remains inactive for all times when $\hat{\rho}_0 \geq \varepsilon_0$. A truncated expansion can nevertheless change sign, as $1 + e^{-t}x$ does for one Ornstein-Uhlenbeck mode, and Section 4 treats the general case.

# 4 Error analysis

## 4.1 A transport identity

For the field of Remark 3.1, Lemma 3.2 extends to the whole path with an additional flux supported where the safeguards act. We state this extension for particles that are not confined, on $\mathbb{R}^d$ or on a torus, and for a field $\hat{v}$ that is bounded and Lipschitz in $x$ uniformly in $t \leq T$, with flow $\hat{\Phi}_{s,t}$. We set $\gamma_t \coloneqq \hat{\rho}_t \hat{v}_t + \nabla\hat{\rho}_t$, which vanishes where $\hat{v}_t = -\nabla \log \hat{\rho}_t$, and $F_T(f) \coloneqq \int_0^T \int \nabla(f \circ \hat{\Phi}_{s,T}) \cdot \gamma_s \, d\pi \, ds$.

**Proposition 4.1** (Transport identity). *Let $\hat{\rho}_t = \sum_{k=0}^{r} e^{-\hat{\lambda}_k t} \hat{c}_k \hat{\varphi}_k$, with coefficients of any regime, $\hat{\varphi}_k \in C^2$ and $\hat{\varphi}_k, |\nabla\hat{\varphi}_k|, L\hat{\varphi}_k \in L^1(\pi)$, and $R_t$ as in Lemma 3.2. For every probability measure $\nu$ and $f \in C_b^1$,*

$$\int f \, d(\hat{\Phi}_{0,T} \# \nu) = \int f \hat{\rho}_T \, d\pi + \int f \circ \hat{\Phi}_{0,T} \, d(\nu - \pi\hat{\rho}_0) - \int_0^T \langle f \circ \hat{\Phi}_{s,T}, R_s \rangle_\pi \, ds + F_T(f). \quad (7)$$

The particles carry the spectral density $\pi\hat{\rho}_T$, corrected by the transport of the initial discrepancy $\nu - \pi\hat{\rho}_0$, of the residual and of the flux of the safeguards (Appendix C). The identity is a duality argument for the continuity equation, of the kind used in a posteriori error estimation (Becker & Rannacher, 2001), and it separates the three sources of error. With exact eigenpairs and population coefficients, $\mu_0 - \pi\hat{\rho}_0 = \pi(I - \Pi_r)\rho_0$, and truncation enters only through the initial law.

**Corollary 4.2** (Total variation). *In regimes (P) and (I), if the safeguards are inactive on $[0, T]$,*

$$\|\hat{\Phi}_{0,T} \# \mu_0 - \pi\|_{\mathrm{TV}} \leq \|\rho_0 - \hat{\rho}_0\|_{L^1(\pi)} + \|\hat{\rho}_T - 1\|_{L^1(\pi)} + \sum_{k=1}^{r} |\hat{c}_k| \, \frac{\|(L - \hat{\lambda}_k)\hat{\varphi}_k\|_{L^1(\pi)}}{\hat{\lambda}_k}.$$

*Otherwise the same bound holds in the bounded-Lipschitz distance with the additional term $\sup_{s \leq T} \mathrm{Lip}(\hat{\Phi}_{s,T}) \int_0^T \|\gamma_s\|_{L^1(\pi)} ds$.*

Unlike in Theorem 4.4 below, no stability constant multiplies the truncation and residual terms, and a Lipschitz constant enters only through the flux of the safeguards. Each residual is weighted by the lifetime $1/\hat{\lambda}_k$ of its mode, and since $\partial_t \hat{\rho}_t = -L\hat{\rho}_t + R_t$, Duhamel's formula shows that the threshold is inactive when $\inf \hat{\rho}_0 - \int_0^T \|R_s\|_\infty ds \geq \varepsilon_0$ (Appendix A).

### 4.2 Stability in the Wasserstein distance

The bound concerns an approximation $\hat{\rho}$ of the density ratio that is twice differentiable in $x$ and either deterministic or fixed after conditioning on data independent of the transported particles. It is expressed through the error $e_t \coloneqq \hat{\rho}_t - \rho_t$ and the relative error $\eta_t \coloneqq e_t/\rho_t$. The assumptions (A0) to (A4) of Appendix A include $\rho_0 \in L^2(\pi) \cap L^\infty(\pi)$, a compact convex domain $B$ on which $\|\nabla^j \varphi_k\|_{\infty,B} \le K_B(1+\lambda_k)^{\alpha+j/2}$ for $j \le 2$, and a one-sided Lipschitz bound $-\nabla^2 \log \rho_t \preceq L_t^+ I$ for the exact velocity on the region defined below, with $|\nabla \log \rho_t| \le V_G$ and $\|\nabla^2 \log \rho_t\| \le H_G$ there. The bound is local because, for a source of compact support, the global one-sided Lipschitz constant of the exact flow itself is not integrable near $t = 0$.

**Definition 4.3** (Validity region)**.** *Let $\delta > 0$. A* validity region *for $\hat{\rho}$ is a family of closed sets $G_t \subset B$, $t \in [0,T]$, at distance at least $2\delta$ from $B^c$, together with numbers $\theta_t \ge 2\varepsilon_0$, such that*

$$\rho_t \ge \theta_t, \qquad |e_t| \le \tfrac{1}{2}\rho_t, \qquad |\nabla \log \hat{\rho}_t| \le v_{\max} \qquad \text{on } G_t.$$

*For such a region, the* expansion of the exact flow *on $G$ is $\Lambda_{s,t} \coloneqq \int_s^t (L_u^+)_+ \, du$ with $L_u^+ \coloneqq \sup_{G_u} \lambda_{\max}(-\nabla^2 \log \rho_u)$, and the* expansion of the approximate flow *is $\hat{\Lambda}_{s,t} \coloneqq \Lambda_{s,t} + \int_s^t \mathrm{pert}_u \, du$ with $\mathrm{pert}_u \coloneqq \sup_{G_u} \big(2\|\nabla^2 \eta_u\| + 4|\nabla \eta_u|^2\big)$. The* mass leaving the region *is $p_{\mathrm{exit}} \coloneqq \mu_0\big(\{x \mid \Phi_t(x) \notin G_t^\delta \text{ for some } t \le T\}\big)$, where $G_t^\delta \coloneqq \{y \in G_t \mid \bar{B}_\delta(y) \subset G_t\}$, and the* velocity errors *are*

$$J_p \coloneqq \int_0^T e^{\hat{\Lambda}_{s,T}} \, \big\|(\hat{v}_s - v_s)\mathbf{1}_{G_s}\big\|_{L^p(\mu_s)} \, ds, \qquad p \in \{1,2\}.$$

On $G_t$ the safeguards are inactive, and the quantity $\hat{\Lambda}$ measures the separation of nearby trajectories, $p_{\mathrm{exit}}$ the mass for which the region cannot be used, and $J_p$ the velocity error seen by the transported mass.

**Theorem 4.4** (Error bound for BKT)**.** *Assume (A0) to (A4) and $\pi \in \mathcal{P}_2(\mathbb{R}^d)$, let $\hat{\rho}$ be deterministic and $(G_t, \theta_t)$ a validity region for it, and consider the confined particle system of Appendix A with i.i.d. initial particles $x_j \sim \mu_0$. Let $S_M(P) \coloneqq \mathbb{E}\, W_2(M^{-1}\sum_{j \le M} \delta_{Z_j}, P)$ for $Z_j \sim P$ i.i.d. denote the error of $M$ independent samples of a law $P$. Then*

$$\begin{aligned} \mathbb{E}\, W_2(\hat{\mu}_T, \pi) &\le \underbrace{W_2(\mu_T, \pi)}_{\text{(i) finite time}} + \underbrace{S_M(\pi)}_{\text{(ii) sampling}} + \underbrace{C_T}_{\text{(iii) approximation}} \\ &\le 2\Big(\sum_{k \ge 1} \frac{e^{-2\lambda_k T} c_k^2}{\lambda_k}\Big)^{1/2} + S_M(\pi) + \big[4\mathcal{J}_2^2 + 4R_B^2 q_* + 2m_B\big]^{1/2}, \end{aligned} \tag{8}$$

*where $C_T \coloneqq [J_2^2 + 4R_B^2 q_* + 2m_B]^{1/2}$, $q_* \coloneqq \min\{1,\, p_{\mathrm{exit}} + \min(J_1/\delta, J_2^2/\delta^2)\}$, $R_B \coloneqq \sup_B |x|$, $m_B \coloneqq \|\rho_0\|_\infty \int_{B^c} |x|^2 d\pi$, and $J_p \le 2\mathcal{J}_p$ with*

$$\mathcal{J}_p \coloneqq \int_0^T e^{\hat{\Lambda}_{s,T}} \, \theta_s^{-(p-1)/2}\big(\|e_s\|_{\dot{H}^1} + V_G \|e_s\|\big) \, ds, \qquad p \in \{1,2\}.$$

For a field constructed from data independent of the particles, the bound holds conditionally on the data.

Term (i) is the distance of the exact flow to the target at time $T$, whose spectral bound decays at least like $e^{-\lambda_1 T}$. Term (ii) is the error of $M$ independent samples of the target (Fournier & Guillin, 2015) and carries no stability factor. Term (iii) consists of the velocity error on the validity region, the mass leaving the region and the mass outside the domain, and the errors of the eigenpairs and of the coefficients enter through $e_t$ (Corollaries B.4 and B.5). The amplification factor $e^{\hat{\Lambda}_{s,T}}$ is that of the exact flow up to a term that vanishes as $r \to \infty$, and it does not multiply the truncation and residual terms of Corollary 4.2. The proof is a stopped Grönwall argument, as in the analysis of probability flows (Chen et al., 2023a; Kwon et al., 2022; Benton et al., 2024) (Appendix B.1). The bound does not depend explicitly on $d$, and its version for the sliced distance of Section 5 has a sampling term of order $M^{-1/2}$ in every dimension (Remark B.3).

Validity regions are obtained from envelope bounds $|\nabla^j e_t| \le w\, a_j S_t$ on $B$, $j \le 2$, where $w$ is the growth envelope of the eigenfunctions and $S_t$ majorises the tail $\sum_{k>r} e^{-\lambda_k t}|c_k|$ and the

sampling error of the coefficients. The set where $wS_t \le \kappa\rho_t$ is a validity region, on which $\mathrm{pert}_t = O(a_2\kappa + a_1^2\kappa^2)$, so that $\int_0^T \mathrm{pert}_t\, dt \to 0$ as $r \to \infty$ for $\kappa$ of order $\lambda_r^{-3/2}$ (Lemma B.6), and $p_{\mathrm{exit}}$ is bounded by tail masses of $\mu_t$ (Lemma B.7). With exact eigenpairs the error splits as $e_t = e_t^{\mathrm{tr}} + e_t^{\mathrm{fl}}$ into a truncation error, with $\|e_t^{\mathrm{tr}}\| \le e^{-\lambda_{r+1}t}\tau_r$ and $\tau_r := \|(I - \Pi_r)\rho_0\|$, and a fluctuation of the coefficients, of order $M'^{-1/2}$ in regime (I) and zero in regime (P) (Corollary B.4). With estimated eigenpairs two further terms appear, the change of the coefficients due to the approximate eigenfunctions and the error of the retained modes, whose energy norms are bounded by weighted sums of $|\hat\lambda_k - \lambda_k|$ and $\|\hat\varphi_k - \varphi_k\|_{\dot H^1}$, $k \le r$ (Corollary B.5). The velocity error is thus controlled by the estimation theory of the spectral methods, and for Ornstein-Uhlenbeck targets with Gaussian sources these bounds decay geometrically in $r$ (Theorem D.2).

### 4.3 Same-sample coefficients

**Proposition 4.5** (Same-sample cancellation). *Let $\hat\varphi_0 = 1, \hat\varphi_1, \dots, \hat\varphi_r$ be orthonormal in $L^2(\pi)$, $\hat\Pi_r$ the orthogonal projection onto their span and $\hat c_k = M^{-1}\sum_j \hat\varphi_k(x_j)$ (regime (S)). For $f \in C_b^1$ and $g = f \circ \hat\Phi_{0,T}$ the particles $X_T^j = \hat\Phi_{0,T}(x_j)$ satisfy*

$$\frac{1}{M}\sum_{j=1}^{M} f(X_T^j) - \int f\, d\pi = \int f(\hat\rho_T - 1)\, d\pi + \frac{1}{M}\sum_{j=1}^{M}\big[(I - \hat\Pi_r)g\big](x_j) + E_T(f),$$

*where $E_T(f) := F_T(f) - \int_0^T \langle f \circ \hat\Phi_{s,T}, R_s\rangle_\pi\, ds$ collects the residual and flux terms of* (7)*.*

In regime (S) the sampling error of the source thus enters only through the component of $g$ orthogonal to the retained modes, and the fluctuation of the coefficients explicitly only through the first term, damped by $e^{-\hat\lambda_1 T}$. When $E_T(f) = 0$, the particles integrate exactly, with respect to $\pi\hat\rho_T$, every $f$ for which $f \circ \hat\Phi_{0,T}$ lies in the span of the retained eigenfunctions. Since $g$ depends on the sample through $\hat c$, the size of the remaining term is given by a central limit theorem. For $c \in \mathbb{R}^r$ let $\Phi^c$ be the flow of the field of Remark 3.1 with coefficients $(1, c)$, for particles that are not confined, $E_T^c(f)$ the corresponding term of Proposition 4.5, $m(c) := \int f \circ \Phi_{0,T}^c\, d\mu_0$, $\bar c_k := \mathbb{E}_{\mu_0}\hat\varphi_k$, $\bar g := f \circ \Phi_{0,T}^{\bar c}$ and $\hat\varphi := (\hat\varphi_1, \dots, \hat\varphi_r)$.

**Theorem 4.6** (Asymptotic variances). *In the setting of Proposition 4.5, assume (A2), let $\hat\varphi_k \in L^2(\mu_0)$, and assume that on a bounded neighbourhood $U$ of $\bar c$ the fields satisfy the hypotheses of Proposition 4.1, that $|f \circ \Phi_{0,T}^c - f \circ \Phi_{0,T}^{c'}| \le \kappa\,|c - c'|$ for some $\kappa \in L^2(\mu_0) \cap L^2(\pi)$, and that $m$ is differentiable at $\bar c$. Then, with $M' = M$ in regime (I), $\sqrt{M}\big(\int f\, d\hat\mu_T - m(\bar c)\big)$ converges in law to $N(0, \sigma^2)$, where*

$$\sigma_P^2 = \mathrm{Var}_{\mu_0}(\bar g), \qquad \sigma_I^2 = \sigma_P^2 + \mathrm{Var}_{\mu_0}\big(\nabla m(\bar c)\cdot\hat\varphi\big), \qquad \sigma_S^2 = \mathrm{Var}_{\mu_0}\big(\bar g + \nabla m(\bar c)\cdot\hat\varphi\big).$$

*Moreover $\nabla m(\bar c) = -(\langle\bar g, \hat\varphi_k\rangle_\pi)_{1\le k\le r} + \beta$ with $|\beta| \le e^{-\hat\lambda_1 T}\|f\| + \|\kappa\|\hat\tau_r + \ell_T(f)$, where $\hat\tau_r := \|(I - \hat\Pi_r)\rho_0\|$ and $\ell_T(f) := \limsup_{c\to\bar c} |E_T^c(f) - E_T^{\bar c}(f)|/|c - \bar c|$. Hence*

$$\sigma_S \le \|(I - \hat\Pi_r)\bar g\|_{L^2(\mu_0)} + \|\rho_0\|_\infty^{1/2}|\beta|,$$

*and $\|(I - \Pi_r)\bar g\|_{L^2(\mu_0)}^2 \le \|\rho_0\|_\infty \lambda_{r+1}^{-1}\|\bar g\|_{\dot H^1}^2$ for exact eigenpairs.*

The variance in regime (P) is that of $f$ under the output law $\Phi_{0,T}^{\bar c}\#\mu_0$, close to $\mathrm{Var}_\pi(f)$ by Sections 4.1 and 4.2, and regime (I) adds the independent fluctuation of the coefficients. In regime (S) the component of $\bar g$ along the retained modes is removed, and $\beta$ collects the finite horizon, the truncation and the residual and flux terms, with $\ell_T(f) = 0$ for exact eigenpairs when the safeguards are inactive on the whole space for $c$ near $\bar c$. For exact eigenpairs $\sigma_S^2 \le 2\|\rho_0\|_\infty(\lambda_{r+1}^{-1}\|\bar g\|_{\dot H^1}^2 + |\beta|^2)$, against $\mathrm{Var}_\pi(f) \le \lambda_1^{-1}\|f\|_{\dot H^1}^2$ for an independent sample. This is the mechanism of moment matching in Monte Carlo integration (Glasserman, 2004), in which a sample is transformed so that prescribed moments take their exact values, here those of the retained eigenfunctions transported by the flow. In the matched Ornstein-Uhlenbeck setting the observed standard deviations follow Theorem 4.6, and regime (S) reduces the standard deviation of an independent sample by factors between 4 and 70 for smooth test functions (Table 3). The gain is smaller in the Wasserstein distance, which involves test functions at all scales, and Corollary C.1 bounds the cost of the same-sample dependence in $W_2$ by a term of order $M^{-1/2}$ and the mass leaving the validity region.

| system | FD | Koopman | reference |
|---|---|---|---|
| 2D, $\beta = 0$ | $0.027 \pm 0.002$ | $0.028 \pm 0.003$ | $0.040 \pm 0.008$ |
| 2D, $\beta = 0.25$ | $0.028 \pm 0.002$ | $0.029 \pm 0.003$ | $0.041 \pm 0.008$ |
| 2D, $\beta = 0.5$ | $0.028 \pm 0.002$ | $0.031 \pm 0.004$ | $0.042 \pm 0.010$ |
| 2D, $\beta = 1$ | $0.027 \pm 0.002$ | $0.027 \pm 0.003$ | $0.042 \pm 0.012$ |
| 2D four wells | $0.024 \pm 0.003$ | $0.028 \pm 0.002$ | $0.043 \pm 0.008$ |
| 10D | $0.053 \pm 0.002$ | $0.061 \pm 0.002$ | $0.059 \pm 0.004$ |
| 50D | $0.0158 \pm 0.0002$ | $0.0188 \pm 0.0012$ | $0.0162 \pm 0.0008$ |

Table 1: Sliced $W_2$ error of BKT with the FD and the Koopman (RBF) eigenpairs, mean $\pm$ sd over ten realisations, for the double well with coupling $\beta$ (2D), the four-well product, the double well with nine harmonic coordinates (10D) and the fifty-fold product (50D).

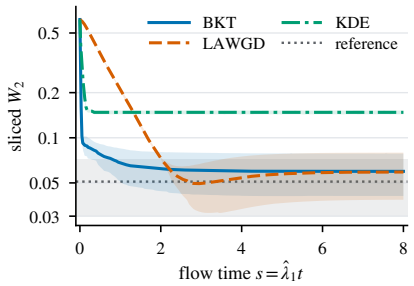



Figure 1: Alanine dipeptide, sliced $W_2$ against the flow time, mean $\pm$ sd over three realisations, and the reference level.

*Remark* 4.7 (Interacting spectral samplers). With the same eigenpairs, the $M$ particles of LAWGD follow $\dot{x}_i = -\sum_{k=1}^{r} \hat{\lambda}_k^{-1} m_k \nabla \hat{\varphi}_k(x_i)$ with $m_k = M^{-1} \sum_j \hat{\varphi}_k(x_j)$, the gradient flow of $\mathcal{E} = \frac{M}{2} \sum_{k=1}^{r} \hat{\lambda}_k^{-1} m_k^2$, a multiple of the squared maximum mean discrepancy between their empirical measure and $\pi$ for the kernel $\sum_{k=1}^{r} \hat{\lambda}_k^{-1} \hat{\varphi}_k \otimes \hat{\varphi}_k$ (Arbel et al., 2019). Its zeros integrate the retained eigenfunctions exactly, so that LAWGD attains the cancellation of Proposition 4.5 through the interaction when it reaches a zero of $\mathcal{E}$, whereas BKT imposes it at $t = 0$.

## 5 Numerical results

The experiments examine, for systems with a known invariant law, how BKT compares with particle methods that use the same data through density estimates, how the accuracy of the eigenpairs, computed with $r$ nonconstant modes, $J$ dictionary functions and $n$ trajectory pairs, transfers to the $M$ transported particles, and how the dimension and the number of metastable states enter. Unless stated otherwise the coefficients are those of regime (S). We write Koopman (RBF) and Koopman (Legendre) for eigenpairs estimated by reversible gEDMD (Klus et al., 2020) with Gaussian and Legendre dictionaries, and compare them with a finite-difference discretisation of the Dirichlet form on $N$ points per direction (FD-$N$) or with the analytical Hermite spectrum. The error is the sliced $W_2$ distance, with 64 fixed directions, between the transported particles and a target sample. The reference level is the mean $\pm$ standard deviation (sd) over 10 realisations of the same distance between two independent target samples, and an error below this mean plus one standard deviation is said to reach the reference level. The comparisons are with particle methods that, like BKT, involve no training of a neural network, and the horizon is $T = 8/\lambda_1$ (Appendix E).

**Dynamics against density estimates.** For the dihedral angles $(\phi, \psi)$ of alanine dipeptide, the eigenpairs of a reversible surrogate with unit mobility are estimated from 250,000 frames of a public molecular dynamics trajectory with a Fourier dictionary of degree 28, and 256 modes transport a source localised in one basin. The kernel particle method uses the same frames through a kernel estimate of the target density and a kernel estimate of the density of the particles (Degond & Mustieles, 1990; Chertock, 2017). It reaches a sliced error of 0.148 by $s = \hat{\lambda}_1 t = 1$ and remains there, about three times the error $0.051 \pm 0.021$ of direct sampling (Figure 1), since its velocity depends on two density estimates whose errors do not decrease along the flow. BKT reaches $0.060 \pm 0.019$ at $s = 8$, within the fluctuation of the reference, with a total variation of the basin masses of $0.033 \pm 0.018$ against $0.023 \pm 0.014$. LAWGD, which uses the same eigenpairs, reaches a comparable final error (Remark 4.7), and BKT attains its late-time level in 34 against 54 seconds (Appendix E.3). The same effect appears for the two-dimensional double wells, where the kernel method uses a drift regressed from the increments of the trajectory and gives 0.039 and 0.048 at $\beta = 0$ and 0.5, against 0.028 and 0.031 for BKT, with twice the spread at $\beta = 0.5$ and five times the cost (Figure 8).

**Ornstein-Uhlenbeck process.** With the analytical spectrum, $M = 10^4$ and $T = 6$, the error equals 0.074, 0.045, 0.023, 0.011 and 0.011 at $r = 10, 15, 20, 30, 60$, against $0.019 \pm 0.005$ for $10^4$

independent samples, and lies below the reference level beyond $r = 20$ to 30 (Figure 3). Matched comparisons order the errors as (S) < (P) < (I), in agreement with Theorem 4.6 (Appendix D.1).

**Two-dimensional double well with coupling.** For $V = (x_1^2 - 1)^2 + x_2^2 + \frac{\beta}{2}(x_2 - x_1)^2$ and a source equal to a bump in the right well of $x_1$ times $N(0, 0.3)$ in $x_2$, BKT with the FD-201 and Koopman (RBF) eigenpairs lies below the reference level at $r = 64$ for every $\beta \in \{0, 0.25, 0.5, 1\}$, with differences between the values of $\beta$ within the fluctuation (Table 1). The reference level is reached between $r = 32$ and 64, since the ordering by eigenvalue interleaves the double-well modes with the harmonic ones (Figure 6). Relative to the reference mean, same-sample coefficients give $0.65$ to $0.73$ and coefficients from an independent sample $1.02$ to $1.23$, close to the values $2^{-1/2}$ and $(3/2)^{1/2}$ expected from Theorem 4.6 when the retained modes carry most of the fluctuation and the distance is evaluated against one target sample (Appendices C.3 and E.2).

**Multi-well potentials.** For the quartic four-well potential $V = (x^2 - 1)^2 + (y^2 - 1)^2$ with the source in one well, both spectra lie below the reference level at $r = 64$, and the four basin masses are recovered. With the FD spectrum BKT also reaches it for a nine-well potential ($0.033 \pm 0.003$ against $0.034 \pm 0.008$), and with the estimated spectrum it gives $0.055 \pm 0.006$, which isolates the error of the spectral estimation (Appendix E.2).

**Dimension and metastability.** The construction of BKT and the bounds of Section 4 do not depend on the dimension explicitly, and the dimension enters through the spectrum in two ways. First, the retained modes must represent the source. When the barriers are high, a potential with $n$ local minima has $n - 1$ nonzero eigenvalues that are exponentially small in the barrier heights and separated from the rest of the spectrum (Bovier et al., 2005), and a source concentrated in one well has in general nonzero coefficients on all of them, with $n = 2^d$ for the product $V = \sum_{i \le d}(x_i^2 - 1)^2$ of $d$ double wells. Second, these slow eigenpairs must be estimated, and with a fixed dictionary the number of functions grows exponentially with $d$. Both difficulties belong to the spectral description of metastable dynamics, the central problem of transfer operator methods in molecular dynamics (Schütte et al., 2023), and in the experiments the transport reaches the reference level whenever the spectrum is accurate, as for the coupled ten-dimensional Ornstein-Uhlenbeck process treated in its eigenbasis ($0.0195 \pm 0.0005$ against $0.0235 \pm 0.0018$, Appendix E.1) and for the fifty-fold product, where the product structure replaces the $17^{50}$ products of one-dimensional modes of a joint expansion by 16 modes per coordinate (Table 1). For $V = (x_1^2 - 1)^2 + \sum_{i=2}^{10} x_i^2$, eigenpairs estimated in $\mathbb{R}^{10}$ ($J = 657$, $r = 16$) also reach the reference level, since the source differs from $\pi$ only in the slow coordinate, and in the fifty-fold product the largest marginal errors ($0.039$ with the estimated factor, $0.020$ with the FD factor, $0.030$ for the reference) place the excess in the one-dimensional spectral estimation (Appendix E.2). The number of modes is thus set by the slow coordinates in which the source differs from $\pi$, and the source can be chosen accordingly. In applications the slow dynamics of high-dimensional systems is described on a few collective variables, as for alanine dipeptide, and coupled systems with many wells require eigenpairs from kernel or neural estimators (Meanti et al., 2023; Mardt et al., 2018; Kostic et al., 2024; Zhang et al., 2022b), which BKT accepts without change, since its bounds transfer the errors of any eigenpairs in energy norm and eigenvalue (Corollary B.5).

## 6 Concluding remarks

BKT turns the spectrum of the generator, estimated from trajectories, into a transport map that moves non-interacting particles along the KL gradient flow of a reversible diffusion, and it uses the trajectories through the dynamics rather than through density estimates. With exact eigenpairs it transports a positive spectral projection of the source exactly, its error for estimated eigenpairs is controlled by the decay of the exact flow, the sampling error of the source, the truncation and the residual of the eigenpairs, and same-sample coefficients remove the retained modes from the fluctuation of the source. For a nonreversible diffusion with the same invariant law and unit diffusion, BKT built on Dirichlet-form estimators still samples $\pi$ (Appendix A). In high dimension the limiting factor is the estimation of the slow eigenpairs of metastable dynamics, and the combination of BKT with neural spectral estimators, to which the bounds of Section 4 apply directly, is the next step, together with a non-asymptotic version of Theorem 4.6 in the Wasserstein distance.

### Reproducibility statement

The code, configurations and data needed to reproduce all experiments are available at `https://anonymous.4open.science/r/BKT-1E06`.

### AI use statement

An AI assistant was used to assist with the writing.

## References


Michael S. Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *International Conference on Learning Representations*, 2023.

Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics. ETH Zürich. Birkhäuser, second edition, 2008. doi: 10.1007/978-3-7643-8722-8.

Michael Arbel, Anna Korba, Adil Salim, and Arthur Gretton. Maximum mean discrepancy gradient flow. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and Geometry of Markov Diffusion Operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2014. doi: 10.1007/978-3-319-00227-9.

Roland Becker and Rolf Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001. doi: 10.1017/S0962492901000010.

Joe Benton, George Deligiannidis, and Arnaud Doucet. Error bounds for flow matching methods. *Transactions on Machine Learning Research*, 2024.

Tyrus Berry and John Harlim. Variable bandwidth diffusion kernels. *Applied and Computational Harmonic Analysis*, 40(1):68–96, 2016. doi: 10.1016/j.acha.2015.01.001.

Tyrus Berry, Dimitrios Giannakis, and John Harlim. Nonparametric forecasting of low-dimensional dynamical systems. *Physical Review E*, 91(3):032915, 2015. doi: 10.1103/PhysRevE.91.032915.

Sergey Bobkov and Michel Ledoux. One-dimensional empirical measures, order statistics, and Kantorovich transport distances. *Memoirs of the American Mathematical Society*, 261(1259), 2019. doi: 10.1090/memo/1259.

Nicholas M. Boffi and Eric Vanden-Eijnden. Probability flow solution of the Fokker–Planck equation. *Machine Learning: Science and Technology*, 4(3):035012, 2023. doi: 10.1088/2632-2153/ace2aa.

Anton Bovier, Véronique Gayrard, and Markus Klein. Metastability in reversible diffusion processes II: precise asymptotics for small eigenvalues. *Journal of the European Mathematical Society*, 7 (1):69–99, 2005. doi: 10.4171/JEMS/22.

Giovanni Brigati and Francesco Pedrotti. Heat flow, log-concavity, and Lipschitz transport maps. *Electronic Communications in Probability*, 30:1–12, 2025. doi: 10.1214/25-ECP717.

Zhiqiang Cai, Yu Cao, Yuanfei Huang, and Xiang Zhou. Weak generative sampler to efficiently sample invariant distribution of stochastic differential equation. *SIAM Journal on Scientific Computing*, 48(4):C708–C735, 2026. doi: 10.1137/24M1665271.

José Antonio Carrillo, Katy Craig, and Francesco S. Patacchini. A blob method for diffusion. *Calculus of Variations and Partial Differential Equations*, 58(2):53, 2019. doi: 10.1007/s00526-019-1486-3.

Sitan Chen, Sinho Chewi, Holden Lee, Yuanzhi Li, Jianfeng Lu, and Adil Salim. The probability flow ODE is provably fast. In *Advances in Neural Information Processing Systems*, volume 36, pp. 68552–68575. Curran Associates, Inc., 2023a.

Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, and Anru R. Zhang. Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions. In *International Conference on Learning Representations*, 2023b.

Yutian Chen, Max Welling, and Alexander J. Smola. Super-samples from kernel herding. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*, pp. 109–116, 2010.

Alina Chertock. A practical guide to deterministic particle methods. In *Handbook of Numerical Analysis*, volume 18, pp. 177–202. Elsevier, 2017. doi: 10.1016/bs.hna.2016.11.004.

Sinho Chewi, Thibaut Le Gouic, Chen Lu, Tyler Maunu, and Philippe Rigollet. SVGD as a kernelized Wasserstein gradient flow of the chi-squared divergence. In *Advances in Neural Information Processing Systems*, volume 33, pp. 2098–2109. Curran Associates, Inc., 2020.

Sinho Chewi, Katharina Eichinger, and Aram-Alexandre Pooladian. Near-Lipschitz stability of the Kim–Milman flow map. *arXiv preprint arXiv:2606.23383*, 2026a.

Sinho Chewi, Aram-Alexandre Pooladian, and Matthew S. Zhang. Stability of the Kim–Milman flow map. *Electronic Communications in Probability*, 31:1–13, 2026b. doi: 10.1214/26-ECP800.

Gary P. T. Choi and Chris H. Rycroft. Density-equalizing maps for simply connected open surfaces. *SIAM Journal on Imaging Sciences*, 11(2):1134–1178, 2018. doi: 10.1137/17M1124796.

Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006. doi: 10.1016/j.acha.2006.04.006.

Matthew J. Colbrook, Lorna J. Ayton, and Máté Szőke. Residual dynamic mode decomposition: robust and verified Koopmanism. *Journal of Fluid Mechanics*, 955:A21, 2023. doi: 10.1017/jfm.2022.1052.

Pierre Degond and Francisco-José Mustieles. A deterministic approximation of diffusion equations using particles. *SIAM Journal on Scientific and Statistical Computing*, 11(2):293–310, 1990. doi: 10.1137/0911018.

Teo Deveney, Jan Stanczuk, Lisa Kreusser, Chris Budd, and Carola-Bibiane Schönlieb. Closing the ODE–SDE gap in score-based diffusion models through the Fokker–Planck equation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 383 (2298):20240503, 2025. doi: 10.1098/rsta.2024.0503.

Timothée Devergne, Vladimir R. Kostic, Michele Parrinello, and Massimiliano Pontil. From biased to unbiased dynamics: an infinitesimal generator approach. In *Advances in Neural Information Processing Systems*, volume 37, pp. 75495–75521. Curran Associates, Inc., 2024.

Raaz Dwivedi and Lester Mackey. Kernel thinning. *Journal of Machine Learning Research*, 25 (152):1–77, 2024.

Nicolas Fournier and Arnaud Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3–4):707–738, 2015. doi: 10.1007/s00440-014-0583-7.

Michael T. Gastner and M. E. J. Newman. Diffusion-based method for producing density-equalizing maps. *Proceedings of the National Academy of Sciences*, 101(20):7499–7504, 2004. doi: 10.1073/pnas.0400280101.

Dimitrios Giannakis. Data-driven spectral decomposition and forecasting of ergodic dynamical systems. *Applied and Computational Harmonic Analysis*, 47(2):338–396, 2019. doi: 10.1016/j.acha.2017.09.001.

Paul Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53 of *Applications of Mathematics*. Springer, New York, 2004. doi: 10.1007/978-0-387-21617-1.

Daniel Zhengyu Huang, Jiaoyang Huang, and Zhengjiang Lin. Convergence analysis of probability flow ODE for score-based generative models. *IEEE Transactions on Information Theory*, 71(6): 4581–4601, 2025. doi: 10.1109/TIT.2025.3557050.

Vasily Ilin, Peter Sushko, and Jingwei Hu. Score-based deterministic density sampling. *Communications on Pure and Applied Analysis*, 32:153–170, 2026. doi: 10.3934/cpaa.2026028.

Jack Indritz. An inequality for Hermite polynomials. *Proceedings of the American Mathematical Society*, 12(6):981–983, 1961. doi: 10.1090/S0002-9939-1961-0132852-2.

Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998. doi: 10.1137/S0036141096303359.

Yuehaw Khoo, Mathias Oster, and Yifan Peng. Optimization-free diffusion model—a perturbation theory approach. *arXiv preprint arXiv:2505.23652*, 2025.

Young-Heon Kim and Emanuel Milman. A generalization of Caffarelli's contraction theorem via (reverse) heat flow. *Mathematische Annalen*, 354(3):827–862, 2012. doi: 10.1007/s00208-011-0749-x.

Ilja Klebanov. Deterministic Fokker–Planck transport—with applications to sampling, variational inference, kernel mean embeddings & sequential Monte Carlo. *arXiv preprint arXiv:2410.18993*, 2024.

Stefan Klus, Feliks Nüske, Péter Koltai, Hao Wu, Ioannis Kevrekidis, Christof Schütte, and Frank Noé. Data-driven model reduction and transfer operator approximation. *Journal of Nonlinear Science*, 28(3):985–1010, 2018. doi: 10.1007/s00332-017-9437-7.

Stefan Klus, Feliks Nüske, Sebastian Peitz, Jan-Hendrik Niemann, Cecilia Clementi, and Christof Schütte. Data-driven approximation of the Koopman generator: model reduction, system identification, and control. *Physica D: Nonlinear Phenomena*, 406:132416, 2020. doi: 10.1016/j.physd.2020.132416.

Vladimir Kostic, Pietro Novelli, Andreas Maurer, Carlo Ciliberto, Lorenzo Rosasco, and Massimiliano Pontil. Learning dynamical systems via Koopman operator regression in reproducing kernel Hilbert spaces. In *Advances in Neural Information Processing Systems*, volume 35, pp. 4017–4031. Curran Associates, Inc., 2022.

Vladimir Kostic, Karim Lounici, Pietro Novelli, and Massimiliano Pontil. Sharp spectral rates for Koopman operator learning. In *Advances in Neural Information Processing Systems*, volume 36, pp. 32328–32339. Curran Associates, Inc., 2023.

Vladimir R. Kostic, Karim Lounici, Hélène Halconruy, Timothée Devergne, and Massimiliano Pontil. Learning the infinitesimal generator of stochastic diffusion processes. In *Advances in Neural Information Processing Systems*, volume 37, pp. 137806–137846. Curran Associates, Inc., 2024.

Dohyun Kwon, Ying Fan, and Kangwook Lee. Score-based generative modeling secretly minimizes the Wasserstein distance. In *Advances in Neural Information Processing Systems*, volume 35, pp. 20205–20217. Curran Associates, Inc., 2022.

Fengyi Li and Youssef Marzouk. Diffusion map particle systems for generative modeling. *Foundations of Data Science*, 7(3):814–837, 2025. doi: 10.3934/fods.2024054.

Pierre-Louis Lions and Sylvie Mas-Gallic. Une méthode particulaire déterministe pour des équations diffusives non linéaires. *Comptes Rendus de l'Académie des Sciences, Série I*, 332 (4):369–376, 2001. doi: 10.1016/S0764-4442(00)01795-X.

Pierre-Louis Lions and Alain-Sol Sznitman. Stochastic differential equations with reflecting boundary conditions. *Communications on Pure and Applied Mathematics*, 37(4):511–537, 1984. doi: 10.1002/cpa.3160370408.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

Qiang Liu and Dilin Wang. Stein variational gradient descent: a general purpose Bayesian inference algorithm. In *Advances in Neural Information Processing Systems*, volume 29, pp. 2378–2386. Curran Associates, Inc., 2016.

Qiang Liu and Dilin Wang. Stein variational gradient descent as moment matching. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Aaron Lou, Minkai Xu, Adam Farris, and Stefano Ermon. Scaling Riemannian diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 80291–80305. Curran Associates, Inc., 2023.

Bethany Lusch, J. Nathan Kutz, and Steven L. Brunton. Deep learning for universal linear embeddings of nonlinear dynamics. *Nature Communications*, 9(1):4950, 2018. doi: 10.1038/s41467-018-07210-0.

Simon Mak and V. Roshan Joseph. Support points. *The Annals of Statistics*, 46(6A):2562–2592, 2018. doi: 10.1214/17-AOS1629.

Dimitra Maoutsa, Sebastian Reich, and Manfred Opper. Interacting particle solutions of Fokker–Planck equations through gradient-log-density estimation. *Entropy*, 22(8):802, 2020. doi: 10.3390/e22080802.

Andreas Mardt, Luca Pasquali, Hao Wu, and Frank Noé. VAMPnets for deep learning of molecular kinetics. *Nature Communications*, 9(1):5, 2018. doi: 10.1038/s41467-017-02388-1.

Giacomo Meanti, Antoine Chatalic, Vladimir Kostic, Pietro Novelli, Massimiliano Pontil, and Lorenzo Rosasco. Estimating Koopman operators with sketching to provably learn large scale dynamical systems. In *Advances in Neural Information Processing Systems*, volume 36, pp. 77242–77276. Curran Associates, Inc., 2023.

Dan Mikulincer and Yair Shenfeld. On the Lipschitz properties of transportation along heat flows. In *Geometric Aspects of Functional Analysis*, volume 2327 of *Lecture Notes in Mathematics*, pp. 269–290. Springer, 2023. doi: 10.1007/978-3-031-26300-2_9.

Joe Neeman. Lipschitz changes of variables via heat flow. *arXiv preprint arXiv:2201.03403*, 2022.

Frank Noé and Feliks Nüske. A variational approach to modeling slow processes in stochastic dynamical systems. *Multiscale Modeling & Simulation*, 11(2):635–655, 2013. doi: 10.1137/110858616.

Feliks Nüske, Hao Wu, Jan-Hendrik Prinz, Christoph Wehmeyer, Cecilia Clementi, and Frank Noé. Markov state models from short non-equilibrium simulations—analysis and correction of estimation bias. *The Journal of Chemical Physics*, 146(9):094104, 2017. doi: 10.1063/1.4976518.

Feliks Nüske, Sebastian Peitz, Friedrich Philipp, Manuel Schaller, and Karl Worthmann. Finite-data error bounds for Koopman-based prediction and control. *Journal of Nonlinear Science*, 33(1):14, 2023. doi: 10.1007/s00332-022-09862-1.

Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*, 26(1–2):101–174, 2001. doi: 10.1081/PDE-100002243.

Grigorios A. Pavliotis. *Stochastic Processes and Applications: Diffusion Processes, the Fokker–Planck and Langevin Equations*, volume 60 of *Texts in Applied Mathematics*. Springer, 2014. doi: 10.1007/978-1-4939-1323-7.

Rémi Peyre. Comparison between $W_2$ distance and $\dot{H}^{-1}$ norm, and localization of Wasserstein distance. *ESAIM: Control, Optimisation and Calculus of Variations*, 24(4):1489–1501, 2018. doi: 10.1051/cocv/2017050.

Hannes Risken. *The Fokker–Planck Equation: Methods of Solution and Applications*, volume 18 of *Springer Series in Synergetics*. Springer, Berlin, Heidelberg, second edition, 1989. doi: 10.1007/978-3-642-61544-3.

Noam Rozen, Aditya Grover, Maximilian Nickel, and Yaron Lipman. Moser flow: divergence-based generative modeling on manifolds. In *Advances in Neural Information Processing Systems*, volume 34, pp. 17669–17680. Curran Associates, Inc., 2021.

Giovanni Russo. Deterministic diffusion of particles. *Communications on Pure and Applied Mathematics*, 43(6):697–733, 1990. doi: 10.1002/cpa.3160430602.

Christopher Scarvelis, Haitz Sáez de Ocáriz Borde, and Justin Solomon. Closed-form diffusion models. *Transactions on Machine Learning Research*, 2025.

Christof Schütte, Stefan Klus, and Carsten Hartmann. Overcoming the timescale barrier in molecular dynamics: transfer operators, variational principles and machine learning. *Acta Numerica*, 32: 517–673, 2023. doi: 10.1017/S0962492923000016.

Zebang Shen, Zhenfu Wang, Satyen Kale, Alejandro Ribeiro, Amin Karbasi, and Hamed Hassani. Self-consistency of the Fokker Planck equation. In *Proceedings of Thirty Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pp. 817–841. PMLR, 2022.

Zheyang Shen, Huihui Wang, Marina Riabiz, and Chris J. Oates. Operator-informed score matching for Markov diffusion models. *arXiv preprint arXiv:2406.09084v2*, 2025.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

Trevor Teolis and Maarten V. de Hoop. Target-adapted Green–Bessel SVGD: uniform-in-time propagation of chaos and last-iterate consistency. *arXiv preprint arXiv:2609.08122*, 2026.

Erkan Turan, Ari Siozopoulos, Louis Martinez, Julien Gaubil, Emery Pierson, and Maks Ovsjanikov. Unfolding generative flows with Koopman operators: trajectory-preserving linearization. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, 2026.

A. W. van der Vaart. *Asymptotic Statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. doi: 10.1017/CBO9780511802256.

Cédric Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, 2003. doi: 10.1090/gsm/058.

Ziyue Wang and Yuko Araki. Functional time series forecasting of distributions: a Koopman–Wasserstein approach. *Behaviormetrika*, 53(2):571–597, 2026. doi: 10.1007/s41237-025-00278-1.

Matthew O. Williams, Ioannis G. Kevrekidis, and Clarence W. Rowley. A data-driven approximation of the Koopman operator: extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6):1307–1346, 2015. doi: 10.1007/s00332-015-9258-5.

Hao Wu and Frank Noé. Variational approach for learning Markov processes from time series data. *Journal of Nonlinear Science*, 30(1):23–66, 2020. doi: 10.1007/s00332-019-09567-y.

Hao Wu, Feliks Nüske, Fabian Paul, Stefan Klus, Péter Koltai, and Frank Noé. Variational Koopman models: slow collective variables and molecular kinetics from short off-equilibrium simulations. *The Journal of Chemical Physics*, 146(15):154104, 2017. doi: 10.1063/1.4979344.

Yuanchao Xu, Fengyi Li, Masahiro Fujisawa, Xiaoyuan Cheng, Youssef Marzouk, and Isao Ishikawa. Generative modeling through Koopman spectral analysis: an operator-theoretic perspective. *arXiv preprint arXiv:2512.18837*, 2025a.

Yuanchao Xu, Kaidi Shao, Isao Ishikawa, Yuka Hashimoto, Nikos Logothetis, and Zhongwei Shen. A data-driven framework for Koopman semigroup estimation in stochastic dynamical systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 35(10):103123, 2025b. doi: 10.1063/5.0283640.

Benjamin J. Zhang, Tuhin Sahai, and Youssef M. Marzouk. A Koopman framework for rare event simulation in stochastic differential equations. *Journal of Computational Physics*, 456:111025, 2022a. doi: 10.1016/j.jcp.2022.111025.

Wei Zhang, Tiejun Li, and Christof Schütte. Solving eigenvalue PDEs of metastable diffusion processes using artificial neural networks. *Journal of Computational Physics*, 465:111377, 2022b. doi: 10.1016/j.jcp.2022.111377.

Meng Zhao and Lijian Jiang. Data-driven probability density forecast for stochastic dynamical systems. *Journal of Computational Physics*, 492:112422, 2023. doi: 10.1016/j.jcp.2023.112422.

Mo Zhou, Stanley Osher, and Wuchen Li. Simulating Fokker–Planck equations via mean field control of score-based normalizing flows. *Journal of Computational Physics*, 566:115274, 2026. doi: 10.1016/j.jcp.2026.115274.

# A ASSUMPTIONS AND PRELIMINARY RESULTS

Throughout, $\pi \in \mathcal{P}_2(\mathbb{R}^d)$ has a positive $C^2$ density proportional to $e^{-V}$, $L = -\Delta + \nabla V \cdot \nabla$, so that $\langle f, Lg\rangle_\pi = \int \nabla f \cdot \nabla g \, d\pi$, and $P_t = e^{-tL}$. For $g$ with $\langle g, 1\rangle_\pi = 0$ we write $\|g\|_{\dot{H}^{-1}} = \sup\{\langle g, f\rangle_\pi \mid \|f\|_{\dot{H}^1} \le 1\}$, and $\|\mu - \pi\|_{\dot{H}^{-1}} = \|d\mu/d\pi - 1\|_{\dot{H}^{-1}}$ for $\mu \ll \pi$. Section 4 uses the following assumptions.

- (A0) *Regularity.* The semigroup $P_t$ has a positive symmetric kernel $p_t(x, y)$ of class $C^2$ in $x$ for $t > 0$. The ratio $\rho_t = P_t\rho_0$ is $C^2$ and positive for $t > 0$, $v_t = -\nabla \log \rho_t$ is locally Lipschitz, and $\mu_t = P_t^*\mu_0$ is the unique solution of $\partial_t\mu + \operatorname{div}(\mu v_t) = 0$ with $\mu|_{t=0} = \mu_0$ among curves with $d\mu_t/d\pi \in L^\infty$ (Ambrosio et al., 2008, Prop. 8.1.8, Thm. 8.3.1).
- (A1) *Spectrum.* The operator $L$ has discrete spectrum $0 = \lambda_0 < \lambda_1 \le \lambda_2 \le \cdots$ with orthonormal eigenfunctions $\varphi_k$, $\varphi_0 \equiv 1$, and $\sum_k (1 + \lambda_k)^a e^{-\lambda_k t} < \infty$ for all $t > 0$ and $a \ge 0$.
- (A2) *Source.* The density ratio $\rho_0 = d\mu_0/d\pi$ lies in $L^2(\pi) \cap L^\infty(\pi)$, and $c_k = \langle \rho_0, \varphi_k\rangle_\pi$.
- (A3) *Domain.* The set $B \subset \mathbb{R}^d$ is compact and convex, $R_B = \sup_B |x|$, $m_B = \|\rho_0\|_\infty \int_{B^c} |x|^2 d\pi$, and $\sup_B \|\nabla^j \varphi_k\| \le K_B(1 + \lambda_k)^{\alpha + j/2}$ for $j = 0, 1, 2$.
- (A4) *Exact flow.* On a neighbourhood of the sets $G_t$ of Definition 4.3, $-\nabla^2 \log \rho_t \preceq L_t^+ I$ with $L^+ \in L^1(0, T)$, $|\nabla \log \rho_t| \le V_G$ and $\|\nabla^2 \log \rho_t\| \le H_G$, and we set $\Lambda_{s,t} \coloneqq \int_s^t (L_u^+)_+ du$. The one-sided bound holds on the whole space for an Ornstein-Uhlenbeck target with a Gaussian source, but not for a source of compact support, since outside the support $-\nabla^2 \log \rho_t \approx I/(2t)$ as $t \to 0$. The locality of Theorem 4.4 is therefore imposed by the exact flow.

**The confined particle system.** The approximate ratio $\hat{\rho}$ is a function on $[0, T] \times \mathbb{R}^d$ of class $C^2$ in $x$, deterministic or fixed by conditioning on data independent of the transported sample, namely the trajectory data of the eigenpairs in regime (P) and, in regime (I), also the coefficient sample. We set $e_t \coloneqq \hat{\rho}_t - \rho_t$ and $\eta_t \coloneqq e_t/\rho_t$. The field $\hat{v}_t$ equals $\mathrm{Clip}_{v_{\max}}(-\nabla \log \hat{\rho}_t)$ where $\hat{\rho}_t \ge \varepsilon_0$, and elsewhere it is bounded and locally Lipschitz, as the field of Remark 3.1. The particle $\hat{X}_t(x)$ solves $\dot{\hat{X}} = \hat{v}_t(\hat{X})$ with reflection at $\partial B$ (Lions & Sznitman, 1984), starting from $x$ if $x \in B$ and otherwise from a point of $B$ chosen independently of the sample, and $\hat{\mu}_T = M^{-1} \sum_j \delta_{\hat{X}_T(x_j)}$. A computation without confinement coincides with this system for any domain that its particles do not reach, and the componentwise projection of Appendix E.2 is a time discretisation of the reflection. Since $\hat{\lambda}_k > 0$ for $k \ge 1$, $\hat{\rho}_t \to 1$ uniformly on $B$, and the threshold $\varepsilon_0$ is inactive after a finite time. We refer to the conditions of Definition 4.3 as (G1) $\rho_t \ge \theta_t$, (G2) $|e_t| \le \frac{1}{2}\rho_t$ and (G3) $|\nabla \log \hat{\rho}_t| \le v_{\max}$. By (G2), $\hat{\rho}_t \ge \frac{1}{2}\theta_t \ge \varepsilon_0$ on $G_t$, so that $\hat{v}_t = -\nabla \log \hat{\rho}_t$ there.

**Preliminary results.**

**Proposition A.1.** *Let $\rho_0 = d\mu_0/d\pi \in L^2(\pi)$ and let $\mu_t$ be the law of $X_t$ when $X_0 \sim \mu_0$. Then $d\mu_t/d\pi = P_t\rho_0 = 1 + \sum_{k \ge 1} e^{-\lambda_k t} c_k \varphi_k$ with $c_k = \mathbb{E}_{\mu_0}[\varphi_k]$, $\chi^2(\mu_t \| \pi) = \sum_{k \ge 1} e^{-2\lambda_k t} c_k^2$ and $\rho_t(x) = \mathbb{E}_{y \sim \mu_0}[p_t(x, y)]$ with $p_t(x, y) = \sum_k e^{-\lambda_k t} \varphi_k(x)\varphi_k(y)$. Under (A0), $\mu_t = \Phi_t \# \mu_0$, and $\rho_t$ solves the backward Kolmogorov equation $\partial_t \rho_t = -L\rho_t$.*

The first identity follows from $\int g \, d\mu_t = \int P_t g \, \rho_0 \, d\pi = \int g \, P_t\rho_0 \, d\pi$, and the others from Parseval's identity and (A0). The method takes its name from the last equation. The Lebesgue density $q_t = \pi\rho_t$ of $\mu_t$ solves the forward equation $\partial_t q_t = \Delta q_t + \operatorname{div}(q_t \nabla V)$, and $\nabla q_t + q_t \nabla V = \pi \nabla \rho_t$ gives $\pi \, \partial_t \rho_t = \operatorname{div}(\pi \nabla \rho_t) = -\pi L \rho_t$. The density ratio therefore evolves under the semigroup of the conditional expectations $(P_t g)(x) = \mathbb{E}[g(X_t) \mid X_0 = x]$, which the estimators of Section 2 approximate from trajectories. The estimators of Section 5 are built from the Gram matrix of the dictionary and from the matrix of its Dirichlet form, which are averages over the sampled states. The dynamics enter through the Dirichlet form, that is, through the invariant law and the diffusion matrix, which together determine the generator of a reversible diffusion. For a nonreversible diffusion with the same invariant law and unit diffusion matrix, these estimators approximate the eigenpairs of the symmetric part of its generator, which is $-L$, and BKT still has $\pi$ as its target.

**Lemma A.2.** *For $t \ge 0$, $\|\mu_t - \pi\|_{\dot{H}^{-1}}^2 = \sum_{k \ge 1} e^{-2\lambda_k t} c_k^2/\lambda_k$ and $W_2(\mu_t, \pi) \le 2\|\mu_t - \pi\|_{\dot{H}^{-1}}$.*

*Proof.* For $f = \sum f_k \varphi_k$ one has $\langle \rho_t - 1, f \rangle_\pi = \sum_{k \geq 1} e^{-\lambda_k t} c_k f_k$, and the Cauchy-Schwarz inequality with the weights $\lambda_k$ gives the identity, with equality for $f_k \propto e^{-\lambda_k t} c_k / \lambda_k$. The inequality is that of Peyre (2018), with $\pi$ as reference measure. □

**Lemma A.3.** *Let* $e_t^{\mathrm{tr}} = -\sum_{k>r} e^{-\lambda_k t} c_k \varphi_k$, $\tau_r = \|(I - \Pi_r)\rho_0\|$ *and* $m_r(t) = \sup_{\lambda \geq \lambda_{r+1}} \lambda e^{-2\lambda t}$. *Then* $\|e_t^{\mathrm{tr}}\| \leq e^{-\lambda_{r+1} t} \tau_r$, $\|e_t^{\mathrm{tr}}\|_{\dot{H}^1} \leq m_r(t)^{1/2} \tau_r$ *and* $\sup_B \|\nabla^j e_t^{\mathrm{tr}}\| \leq K_B \sum_{k>r} (1 + \lambda_k)^{\alpha + j/2} e^{-\lambda_k t} |c_k|$ *for* $j \leq 2$.

The first two bounds follow from Parseval's identity and the third from (A3), and the last sum is finite by (A1).

**Lemma A.4.** *Let* $P, Q \in \mathcal{P}_2(\mathbb{R}^d)$ *and let* $P_M$ *be the empirical measure of* $M$ *i.i.d. samples of* $P$. *Then* $\mathbb{E} W_2(P_M, Q) \leq W_2(P, Q) + \min\{S_M(P), S_M(Q)\}$.

*Proof.* Let $(U_j, Z_j)$ be i.i.d. pairs whose law is an optimal coupling of $P$ and $Q$, and let $Q_M$ be the empirical measure of the $Z_j$. Pairing equal indices and Jensen's inequality give $\mathbb{E} W_2(P_M, Q_M) \leq W_2(P, Q)$, and the triangle inequality through $Q_M$ or through $P$ gives the claim. □

*Proof of Lemma 3.2.* We have $\mathrm{div}(\hat{q}_t \nabla \log \hat{\rho}_t) = \mathrm{div}(\pi \nabla \hat{\rho}_t) = -\pi L \hat{\rho}_t$ and $\partial_t \hat{q}_t = -\pi \sum_k \hat{\lambda}_k e^{-\hat{\lambda}_k t} \hat{c}_k \hat{\varphi}_k$, and the identity follows by subtraction. For exact eigenpairs $L\hat{\varphi}_k = \lambda_k \hat{\varphi}_k$, so that $R \equiv 0$. □

*Proof of Proposition 3.3.* For exact eigenpairs $\hat{\rho}_t = 1 + \sum_{k=1}^r e^{-\lambda_k t} a_k \varphi_k = P_t \hat{\rho}_0$. Since $P_t$ is a Markov operator with a positive kernel, $P_t(\hat{\rho}_0 - \inf \hat{\rho}_0) \geq 0$, which gives $\inf \hat{\rho}_t \geq \inf \hat{\rho}_0$, and $\hat{\rho}_t > 0$ on $[0, T] \times \mathbb{R}^d$. The measure $\pi \hat{\rho}_0$ is a probability measure, since $\int \varphi_k \, d\pi = 0$ for $k \geq 1$, and by Proposition A.1 $\pi \hat{\rho}_t$ is the law at time $t$ of the diffusion started from it. The field $\hat{v}_t = -\nabla \log \hat{\rho}_t$ is locally Lipschitz on $[0, T] \times \mathbb{R}^d$, since $\hat{\rho}$ is continuous and positive there and $\varphi_k \in C^2$, and $\int_0^T \int |\hat{v}_t| \hat{\rho}_t \, d\pi \, dt = \int_0^T \|\nabla \hat{\rho}_t\|_{L^1(\pi)} dt \leq T \sum_k |a_k| \lambda_k^{1/2}$, since $\|\nabla \varphi_k\|^2 = \lambda_k$. By Lemma 3.2 with $R \equiv 0$, $\pi \hat{\rho}_t$ solves the continuity equation with velocity $\hat{v}_t$ on the whole space, and Ambrosio et al. (2008, Prop. 8.1.8) gives $\pi \hat{\rho}_t = \hat{\Phi}_{0,t} \# (\pi \hat{\rho}_0)$, with a flow defined $\pi \hat{\rho}_0$-almost everywhere, hence $\mu_0$-almost everywhere since $\hat{\rho}_0 > 0$. For population coefficients $\hat{\rho}_0 = \Pi_r \rho_0$, and

$$\hat{\Phi}_{0,T} \# \mu_0 - \pi = \hat{\Phi}_{0,T} \# (\mu_0 - \pi \hat{\rho}_0) + \pi(\hat{\rho}_T - 1).$$

The image of a signed measure under a measurable map has at most its total variation, $\|\mu_0 - \pi \hat{\rho}_0\|_{\mathrm{TV}} = \|(I - \Pi_r)\rho_0\|_{L^1(\pi)}$, and $\|\hat{\rho}_T - 1\|_{L^1(\pi)} \leq \|\hat{\rho}_T - 1\| = (\sum_{k=1}^r e^{-2\lambda_k T} c_k^2)^{1/2} \leq e^{-\lambda_1 T} \|\Pi_r \rho_0 - 1\|$. □

For estimated eigenpairs, $\partial_t \hat{\rho}_t = -\sum_k \hat{\lambda}_k e^{-\hat{\lambda}_k t} \hat{c}_k \hat{\varphi}_k = -L\hat{\rho}_t + R_t$. When the $\hat{\varphi}_k$ lie in the domain of $L$ and the residuals $(L - \hat{\lambda}_k)\hat{\varphi}_k$ are bounded, the variation of constants formula gives $\hat{\rho}_t = P_t \hat{\rho}_0 + \int_0^t P_{t-s} R_s \, ds$, and since $P_{t-s}$ is a Markov operator, $\inf \hat{\rho}_t \geq \inf \hat{\rho}_0 - \int_0^t \|R_s\|_\infty \, ds$.

# B Proof of the error bound and of its corollaries

## B.1 Proof of Theorem 4.4

Fix $x$ and write $X_t = \Phi_t(x)$, $\hat{X}_t = \hat{X}_t(x)$, $D_t = |\hat{X}_t - X_t|$, $\sigma_1 = \inf\{t \leq T \mid X_t \notin G_t^\delta\}$, $\sigma_2 = \inf\{t \leq T \mid D_t > \delta\}$ and $\sigma = \sigma_1 \wedge \sigma_2$. For $t < \sigma$ one has $X_t \in G_t^\delta$ and $\hat{X}_t \in \bar{B}_\delta(X_t) \subset G_t$, so that the reflection is inactive and the segment $[X_t, \hat{X}_t]$ lies in $G_t$, where $\hat{v}_t = -\nabla \log \hat{\rho}_t$. In particular, the sets $G_t$ need not be convex.

**Lemma B.1.** *On* $G_t$ *one has* $\hat{v}_t = v_t - \nabla \eta_t / (1 + \eta_t)$,

$$|\hat{v}_t - v_t| \leq 2|\nabla \eta_t| \leq \frac{2\big(|\nabla e_t| + V_G |e_t|\big)}{\rho_t}, \qquad -\nabla^2 \log \hat{\rho}_t \preceq \big(L_t^+ + \mathrm{pert}_t\big) I,$$

*and* $J_p \leq 2\mathcal{J}_p$ *for* $p \in \{1, 2\}$.

*Proof.* By (G2), $\hat{\rho} = \rho(1+\eta)$ with $1+\eta \geq \frac{1}{2}$, $\nabla\eta = (\nabla e - e\nabla\log\rho)/\rho$ and $-\nabla^2\log(1+\eta) = -\nabla^2\eta/(1+\eta) + \nabla\eta\otimes\nabla\eta/(1+\eta)^2$. By (G1), $\int_{G_s}|\hat{v}_s - v_s|^2 d\mu_s \leq 4\theta_s^{-1}\int(|\nabla e_s| + V_G|e_s|)^2 d\pi$, while in $L^1(\mu_s)$ the factor $\rho_s$ cancels the denominator, and Minkowski's inequality gives $J_p \leq 2\mathcal{J}_p$. □

**Lemma B.2.** *For $t < \sigma$ one has $D_t \leq I(x) := \int_0^T e^{\hat{\Lambda}_{s,T}}|\hat{v}_s - v_s|(X_s)\mathbf{1}_{G_s}(X_s)\,ds$. Moreover, $\mathbb{E}_{\mu_0} I = J_1$, $(\mathbb{E}_{\mu_0} I^2)^{1/2} \leq J_2$, and the event $E_{\mathrm{bad}} = \{\sigma \leq T\}$ satisfies $\mu_0(E_{\mathrm{bad}}) \leq q_*$.*

*Proof.* Along the segment $[X_t, \hat{X}_t] \subset G_t$ the Jacobian of $\hat{v}_t$ is symmetric and bounded above by $(L_t^+ + \mathrm{pert}_t)I$ (Lemma B.1), so that $\frac{d}{dt}D_t \leq (L_t^+ + \mathrm{pert}_t)D_t + |\hat{v}_t - v_t|(X_t)$ where $D_t > 0$, and Grönwall's inequality with $D_0 = 0$ gives the first bound. Tonelli's theorem and Minkowski's integral inequality give the moments of $I$, since $X_s \sim \mu_s$. Outside the event $\{\sigma_1 \leq T\}$, of probability $p_{\mathrm{exit}}$, a separation larger than $\delta$ requires $I \geq \delta$, and Markov's inequality applied to $I$ and $I^2$ bounds its probability by $\min\{J_1/\delta, J_2^2/\delta^2\}$. □

*Proof of Theorem 4.4.* Let $\widetilde{\mu}_T = \Phi_T\#\nu_0$. The points $\Phi_T(x_j)$ are i.i.d. with law $\mu_T$, so that the triangle inequality and Lemma A.4 give

$$\mathbb{E}W_2(\hat{\mu}_T, \pi) \leq \mathbb{E}W_2(\hat{\mu}_T, \widetilde{\mu}_T) + W_2(\mu_T, \pi) + \min\{S_M(\pi), S_M(\mu_T)\},$$

and the sampling error of the source is not propagated by the flow. Since the field is independent of the sample, pairing equal indices gives $\mathbb{E}W_2^2(\hat{\mu}_T, \widetilde{\mu}_T) \leq \mathbb{E}_{x\sim\mu_0} D_T(x)^2$, and it is this step that fails for same-sample coefficients. On $E_{\mathrm{bad}}^c$, $D_T \leq I$ by Lemma B.2. On $E_{\mathrm{bad}}$, confinement gives $D_T^2 \leq 2R_B^2 + 2|X_T|^2$, where $|X_T| \leq R_B$ unless $X_T \notin B$ and $\mathbb{E}[|X_T|^2; X_T \notin B] \leq m_B$ since $\|\rho_T\|_\infty \leq \|\rho_0\|_\infty$. Hence $\mathbb{E}D_T^2 \leq J_2^2 + 4R_B^2 q_* + 2m_B = C_T^2$, and Lemma A.2 gives the spectral bound on $W_2(\mu_T, \pi)$. □

For a field constructed from data $\mathcal{D}$ independent of the particles, the argument holds conditionally on $\mathcal{D}$, and all quantities in the bound depend on $\mathcal{D}$.

*Remark* B.3 (Dimension). The bound does not depend explicitly on $d$. The dimension enters through the number of modes needed to represent the source, which is small when the source differs from $\pi$ only in slow coordinates, through $S_M(\pi)$ and through the estimation of the eigenpairs, for which kernel, neural and generator-learning estimators remain applicable in high dimension (Meanti et al., 2023; Mardt et al., 2018; Lusch et al., 2018; Kostic et al., 2024; Devergne et al., 2024; Zhang et al., 2022b). For the sliced distance $SW_2(\mu,\nu)^2 := |\Theta|^{-1}\sum_{\theta\in\Theta} W_2^2(\theta_\#\mu, \theta_\#\nu) \leq W_2(\mu,\nu)^2$ of Section 5, with $\theta_\#$ the projection on a unit vector $\theta$, the proof gives (8) with $SW_2$ on the left and $S_M(\pi)$ replaced by the sliced error of an independent sample, of order $M^{-1/2}$ up to logarithmic factors in every dimension (Bobkov & Ledoux, 2019).

### B.2 Errors of the approximate density ratio

**Corollary B.4.** *Let the eigenpairs be exact and let the coefficients be computed from an i.i.d. sample $\nu'$ of $\mu_0$ of size $M'$, independent of the transported sample as in regime (I). Expectations below are over $\nu'$. Then $e_t = e_t^{\mathrm{tr}} + e_t^{\mathrm{fl}}$, with*

$$\|e_t^{\mathrm{tr}}\| \leq e^{-\lambda_{r+1}t}\tau_r, \qquad \|e_t^{\mathrm{tr}}\|_{\dot{H}^1} \leq m_r(t)^{1/2}\tau_r,$$

$$\mathbb{E}\|e_t^{\mathrm{fl}}\|_{\dot{H}^1}^2 = \frac{1}{M'}\sum_{k\leq r}\lambda_k e^{-2\lambda_k t}\mathrm{Var}_{\mu_0}(\varphi_k), \qquad \mathbb{E}\|e_t^{\mathrm{fl}}\|^2 \leq \frac{\mathbb{E}_{\mu_0}[p_{2t}(X,X)]}{M'},$$

*where $\tau_r := \|(I - \Pi_r)\rho_0\|$ and $m_r(t) := \sup_{\lambda\geq\lambda_{r+1}} \lambda e^{-2\lambda t}$. In regime (P), $e_t^{\mathrm{fl}} = 0$.*

*Proof.* The bounds on $e_t^{\mathrm{tr}}$ are Lemma A.3. For $e_t^{\mathrm{fl}} = \sum_{k\leq r} e^{-\lambda_k t}(\hat{c}_k - c_k)\varphi_k$, Parseval's identity and $\mathbb{E}(\hat{c}_k - c_k)^2 = \mathrm{Var}_{\mu_0}(\varphi_k)/M'$ give the identity, and $\sum_k e^{-2\lambda_k t}\mathrm{Var}_{\mu_0}(\varphi_k) \leq \mathbb{E}_{\mu_0}[p_{2t}(X,X)]$ gives the second bound. The same argument with (A3) bounds $\sup_B \|\nabla^j e_t\|$ for $j \leq 2$, as needed for the validity region, and it requires only an i.i.d. coefficient sample, so that it also applies to the source sample of regime (S). □

**Corollary B.5** (Estimated eigenpairs)**.** *Let the eigenpairs and the rank be fixed by data independent of the transported sample, with $\hat\lambda_k > 0$ and $\int \hat\varphi_k \, d\pi = 0$, and let $\hat c_k = \int \hat\varphi_k \, d\nu'$ as in regime (I). Write $\psi_k = \hat\varphi_k - \varphi_k$, $\tilde c_k = \int \varphi_k \, d\nu'$ and*

$$a_k(t) \coloneqq |e^{-\hat\lambda_k t} - e^{-\lambda_k t}| \le t|\hat\lambda_k - \lambda_k| e^{-(\lambda_k \wedge \hat\lambda_k)t}, \qquad A_{M'} \coloneqq \Big(\frac{\|\rho_0 - 1\|^2 + \|\rho_0\|_\infty / M'}{\lambda_1}\Big)^{1/2}.$$

*Then $e_t = e_t^{\mathrm{tr}} + e_t^{\mathrm{fl}} + d_t + e_t^{\mathrm{est}}$, where the first two terms are those of Corollary B.4 for the coefficients $\tilde c_k$, $d_t = \sum_{k=1}^r e^{-\lambda_k t}(\hat c_k - \tilde c_k)\varphi_k$ and $e_t^{\mathrm{est}} = \sum_{k=1}^r \hat c_k (e^{-\hat\lambda_k t}\hat\varphi_k - e^{-\lambda_k t}\varphi_k)$, and*

$$\|e_t^{\mathrm{est}}\|_{\dot H^1} \le \sum_{k=1}^r |\hat c_k| e^{-\hat\lambda_k t} \|\psi_k\|_{\dot H^1} + \Big(\sum_{k=1}^r \lambda_k |\hat c_k|^2 a_k(t)^2\Big)^{1/2},$$
$$\big(\mathbb{E}_{\nu'} \|d_t\|_{\dot H^1}^2\big)^{1/2} \le A_{M'} \Big(\sum_{k=1}^r \lambda_k e^{-2\lambda_k t} \|\psi_k\|_{\dot H^1}^2\Big)^{1/2}.$$

*The bounds in $L^2(\pi)$ are analogous, and for population coefficients $M' = \infty$ and $e_t^{\mathrm{fl}} = 0$.*

*Proof.* The decomposition follows by adding and subtracting the two expansions in the exact basis. With $\beta_k = \mathbb{E}_{\mu_0}\psi_k = \langle \rho_0 - 1, \psi_k\rangle_\pi$ and $V_k = \mathrm{Var}_{\mu_0}(\psi_k) \le \|\rho_0\|_\infty \|\psi_k\|^2$, one has $\mathbb{E}_{\nu'}|\hat c_k - \tilde c_k|^2 = \beta_k^2 + V_k/M' \le A_{M'}^2 \|\psi_k\|_{\dot H^1}^2$ by the Poincaré inequality, and the orthogonality of the $\varphi_k$ in $L^2(\pi)$ and in $\dot H^1$ gives the bound on $d_t$ without independence between modes. In $e_t^{\mathrm{est}} = \sum_k \hat c_k e^{-\hat\lambda_k t}\psi_k + \sum_k \hat c_k (e^{-\hat\lambda_k t} - e^{-\lambda_k t})\varphi_k$ the first sum is bounded by the triangle inequality and the second by orthogonality, and the bound on $a_k(t)$ is the mean value theorem. □

### B.3 Validity regions and the exit probability

**Lemma B.6** (Construction of a validity region)**.** *Suppose that for $t \le T$ there are a continuous function $w \ge 1$ on $B$ and numbers $S_t > 0$, $a_1, a_2$ such that $|\nabla^j e_t| \le w\, a_j S_t$ on $B$ for $j = 0, 1, 2$, with $a_0 = 1$. Let $\kappa \in (0, \frac12]$, $\kappa_t = \min\{\kappa, S_t/(2\varepsilon_0)\}$ and $G_t = \{x \in B_{2\delta} \mid w(x) S_t \le \kappa_t \rho_t(x)\}$, where $B_{2\delta}$ is the set of points of $B$ at distance at least $2\delta$ from its complement. Then $(G_t)$ is a validity region with $\theta_t \ge \max\{2\varepsilon_0, S_t/\kappa_t\}$ and $\mathrm{per}_t \le 2(a_2 + 2V_G a_1 + V_G^2 + H_G)\kappa_t + 4(a_1 + V_G)^2\kappa_t^2$, and the bound $v_{\max} \ge V_G + 2(a_1 + V_G)\kappa$ suffices.*

*Proof.* The set $G_t$ is closed and lies in $B_{2\delta}$. On $G_t$, $|e_t| \le wS_t \le \kappa_t\rho_t \le \rho_t/2$ and $\rho_t \ge S_t/\kappa_t \ge 2\varepsilon_0$, which gives (G1) and (G2). Differentiating $\eta\rho = e$ twice and using $|\nabla^j e| \le w a_j S_t \le a_j \kappa_t \rho_t$ gives $|\nabla\eta| \le (a_1 + V_G)\kappa_t$ and $\|\nabla^2\eta\| \le (a_2 + 2V_G a_1 + V_G^2 + H_G)\kappa_t$, hence the bound on $\mathrm{per}_t$, and $|\nabla \log\hat\rho| \le V_G + 2(a_1 + V_G)\kappa_t$ gives (G3). □

**Lemma B.7.** *(i) In dimension one, if $v_t$ is locally Lipschitz then $\Phi_t$ is increasing and $F_t \circ \Phi_t = F_0$, where $F_t$ is the distribution function of $\mu_t$. If $G_t^\delta = [a_t, b_t]$ is nonempty for all $t$, then $p_{\mathrm{exit}} \le \min\{1, \sup_t F_t(a_t^-) + \sup_t(1 - F_t(b_t))\}$, with equality when the two families of tails are disjoint. (ii) In any dimension, let $\bar v = \sup_{t\le T}\sup_B |v_t|$, $\Delta \le \delta/(2\bar v)$, $t_i = i\Delta$, and suppose that $G_{t_i}^{3\delta/2} \subset G_t^\delta$ for $t \in [t_i, t_{i+1}]$. Then $p_{\mathrm{exit}} \le \sum_{i \le T/\Delta} \mu_{t_i}\big((G_{t_i}^{2\delta})^c\big) + \mu_0(\{x \mid \Phi_t(x) \notin B \text{ for some } t \le T\})$.*

*Proof.* In dimension one the flow is increasing and is the quantile coupling of optimal transport (Villani, 2003, Ch. 2), so that a particle exits $G_t^\delta$ at time $t$ if and only if $F_0(x) < F_t(a_t^-)$ or $F_0(x) > F_t(b_t)$, which gives (i). For (ii), a particle in $G_{t_i}^{2\delta}$ moves by at most $\delta/2$ on $[t_i, t_{i+1}]$ and therefore remains in $G_{t_i}^{3\delta/2} \subset G_t^\delta$, and a union bound over the grid gives the claim. □

In dimension one, part (i) gives $p_{\mathrm{exit}}$ exactly. In higher dimension the hypothesis of (ii) asks only that the region move by less than $\delta/2$ over one grid interval, which is a condition on the step $\Delta$.

# C THE TRANSPORT IDENTITY AND SAME-SAMPLE COEFFICIENTS

*Proof of Proposition 4.1.* Let $\hat{q}_t = \pi\hat{\rho}_t$ and $\hat{\nu}_t = \hat{\Phi}_{0,t}\#\nu$. As in the proof of Lemma 3.2, and since $\hat{q}_t\hat{v}_t = \pi\gamma_t - \pi\nabla\hat{\rho}_t$,

$$\partial_t\hat{q}_t + \operatorname{div}(\hat{q}_t\hat{v}_t) = \pi R_t + \operatorname{div}(\pi\gamma_t), \qquad \partial_t\hat{\nu}_t + \operatorname{div}(\hat{\nu}_t\hat{v}_t) = 0$$

in the sense of distributions on the whole space, without any sign condition on $\hat{\rho}_t$. The function $g_s = f \circ \hat{\Phi}_{s,T}$ is bounded and Lipschitz and solves $\partial_s g_s + \hat{v}_s \cdot \nabla g_s = 0$ with $g_T = f$. Testing both equations against $g_s$, after mollification in $x$, gives $\frac{d}{ds}\int g_s\, d\hat{\nu}_s = 0$ and $\frac{d}{ds}\int g_s\hat{\rho}_s\, d\pi = \langle g_s, R_s\rangle_\pi - \int \nabla g_s\cdot\gamma_s\, d\pi$, and integration over $[0,T]$ gives (7). Where the safeguards are inactive, $\hat{v}_t = -\nabla\hat{\rho}_t/\hat{\rho}_t$ and $\gamma_t = 0$. □

*Proof of Corollary 4.2.* For $\nu = \mu_0$ and $|f| \le 1$, the second term of (7) is $\int (f\circ\hat{\Phi}_{0,T})(\rho_0 - \hat{\rho}_0)\, d\pi$, bounded by $\|\rho_0 - \hat{\rho}_0\|_{L^1(\pi)}$, and $|\int f\hat{\rho}_T\, d\pi - \int f\, d\pi| \le \|\hat{\rho}_T - 1\|_{L^1(\pi)}$. The residual term is at most $\int_0^T \|R_s\|_{L^1(\pi)} ds \le \sum_{k\le r} |\hat{c}_k|\, \|(L-\hat{\lambda}_k)\hat{\varphi}_k\|_{L^1(\pi)}/\hat{\lambda}_k$. If the safeguards are active, $f$ is also taken with $\mathrm{Lip}(f) \le 1$, and the flux term is at most $\int_0^T \mathrm{Lip}(\hat{\Phi}_{s,T})\|\gamma_s\|_{L^1(\pi)} ds$. □

## C.1 SAME-SAMPLE COEFFICIENTS AND ASYMPTOTIC VARIANCES

*Proof of Proposition 4.5.* Apply Proposition 4.1 with $\nu = \nu_0$, let $\tilde{c}_k$ be the coefficients of the field, with $\tilde{c}_0 = 1$, and let $\hat{\Pi}_r g = \sum_{k\le r}\langle g, \hat{\varphi}_k\rangle_\pi \hat{\varphi}_k$. By orthonormality $\int \hat{\varphi}_k\hat{\rho}_0\, d\pi = \tilde{c}_k$, so that $\int \hat{\Pi}_r g\, d(\nu_0 - \pi\hat{\rho}_0) = \sum_{k=1}^r \langle g, \hat{\varphi}_k\rangle_\pi \big(M^{-1}\sum_j \hat{\varphi}_k(x_j) - \tilde{c}_k\big)$, which vanishes in regime (S), while $\int (I - \hat{\Pi}_r)g\,\hat{\rho}_0\, d\pi = 0$ because $\hat{\rho}_0$ lies in the span of the $\hat{\varphi}_k$. If the $\hat{\varphi}_k$ have a Gram matrix $\Gamma \ne I$ in $L^2(\pi)$, for instance when they are normalised with respect to the trajectory data, and $\hat{\Pi}_r$ denotes the orthogonal projection onto their span, the same computation gives the additional term $b^\top(\Gamma^{-1} - I)\hat{m}$, where $b_k = \langle g, \hat{\varphi}_k\rangle_\pi$ and $\hat{m}_k = M^{-1}\sum_j \hat{\varphi}_k(x_j)$ for $0 \le k \le r$, whose fluctuation is of order $\|\Gamma - I\|M^{-1/2}$. □

*Proof of Theorem 4.6.* Write $G(c,x) = f(\Phi^c_{0,T}(x))$, $P$ for $\mu_0$ and $P_M$ for $\nu_0$, so that $\int f\, d\hat{\mu}_T = P_M G(c,\cdot)$ with $c = \bar{c}$, $\check{c}$ or $\hat{c}$ in regimes (P), (I) and (S). By the law of large numbers $\hat{c}, \check{c} \to \bar{c}$ almost surely, and $\sqrt{M}(\hat{c} - \bar{c}) = \sqrt{M}(P_M - P)\hat{\varphi}$ is asymptotically normal. Since $|G(c,\cdot) - G(c',\cdot)| \le \kappa|c - c'|$ on the bounded set $U$ with $\kappa \in L^2(\mu_0)$, the class $\{G(c,\cdot) \mid c \in U\}$ is $\mu_0$-Donsker (van der Vaart, 1998, Example 19.7 and Theorem 19.5). On the event $\hat{c} \in U$, whose probability tends to one, $P(G(\hat{c},\cdot) - \bar{g})^2 \le \|\kappa\|^2_{L^2(\mu_0)}|\hat{c} - \bar{c}|^2$, which tends to zero in probability, and asymptotic equicontinuity (van der Vaart, 1998, Lemma 19.24) gives $(P_M - P)(G(\hat{c},\cdot) - \bar{g}) = o_P(M^{-1/2})$. The differentiability of $m$ at $\bar{c}$ gives $m(\hat{c}) - m(\bar{c}) = \nabla m(\bar{c})\cdot(\hat{c} - \bar{c}) + o_P(M^{-1/2})$. Hence

$$P_M G(\hat{c},\cdot) - m(\bar{c}) = (P_M - P)\big(\bar{g} + \nabla m(\bar{c})\cdot\hat{\varphi}\big) + o_P(M^{-1/2}),$$

and the central limit theorem gives $\sigma_S^2$. In regime (I) the same expansion holds with $(P_M - P)\hat{\varphi}$ replaced by the independent fluctuation $\check{c} - \bar{c}$, and in regime (P) the second term is absent.

For the derivative of $m$, apply Proposition 4.1 with $\nu = \mu_0$ to the field with coefficients $c \in U$. Since $\hat{\rho}_0^c - \hat{\rho}_0^{\bar{c}} = \sum_{k=1}^r (c_k - \bar{c}_k)\hat{\varphi}_k$ and, by orthonormality, $\rho_0 - \hat{\rho}_0^{\bar{c}} = (I - \hat{\Pi}_r)\rho_0$,

$$m(c) = \langle f, 1\rangle_\pi + \sum_{k=1}^r e^{-\hat{\lambda}_k T} c_k \langle f, \hat{\varphi}_k\rangle_\pi + \langle G(c,\cdot), (I - \hat{\Pi}_r)\rho_0\rangle_\pi$$
$$- \sum_{k=1}^r (c_k - \bar{c}_k)\langle G(c,\cdot), \hat{\varphi}_k\rangle_\pi + E_T^c(f).$$

Subtracting the same expression at $\bar{c}$ and using $|\langle G(c,\cdot) - \bar{g}, h\rangle_\pi| \le \|\kappa\|\,\|h\|\,|c - \bar{c}|$ gives

$$m(c) - m(\bar{c}) = \sum_{k=1}^r (c_k - \bar{c}_k)\big(e^{-\hat{\lambda}_k T}\langle f, \hat{\varphi}_k\rangle_\pi - \langle \bar{g}, \hat{\varphi}_k\rangle_\pi\big) + \epsilon(c)$$

with $|\epsilon(c)| \le \big(\|\kappa\|\hat{\tau}_r + \ell_T(f) + o(1)\big)|c-\bar{c}|$, so that $\beta = \nabla m(\bar{c}) + (\langle \bar{g}, \hat{\varphi}_k\rangle_\pi)_k$ satisfies $|\beta| \le |(e^{-\hat{\lambda}_k T}\langle f, \hat{\varphi}_k\rangle_\pi)_k| + \|\kappa\|\hat{\tau}_r + \ell_T(f)$, and Bessel's inequality bounds the first term by $e^{-\hat{\lambda}_1 T}\|f\|$. Finally $\bar{g} + \nabla m(\bar{c})\cdot\hat{\varphi} = (I - \hat{\Pi}_r)\bar{g} + \langle \bar{g}, 1\rangle_\pi + \beta\cdot\hat{\varphi}$, where $\|\beta\cdot\hat{\varphi}\|_{L^2(\mu_0)} \le \|\rho_0\|_\infty^{1/2}|\beta|$ by orthonormality, and for exact eigenpairs $\|(I-\Pi_r)\bar{g}\|^2_{L^2(\mu_0)} \le \|\rho_0\|_\infty \|(I-\Pi_r)\bar{g}\|^2 \le \|\rho_0\|_\infty \lambda_{r+1}^{-1}\|\bar{g}\|^2_{\dot{H}^1}$. □

The hypotheses on $U$ hold on a torus, or for dictionaries with bounded derivatives of order at most two, since the safeguarded field is then Lipschitz in $x$ and in $c$, and the differentiability of $m$ holds when the safeguards are inactive along the paths or are replaced by smooth functions. In regime (P), $\sigma_P^2$ is the variance of $f$ under $\Phi^{\bar{c}}_{0,T}\#\mu_0$, and $\mathrm{Var}_\pi(f) \le \lambda_1^{-1}\|f\|^2_{\dot{H}^1}$ by the Poincaré inequality. The centring $m(\bar{c})$ differs from $\int f\, d\pi$ by $\int f(\hat{\rho}^{\bar{c}}_T - 1)\, d\pi + \langle \bar{g}, (I-\hat{\Pi}_r)\rho_0\rangle_\pi + E^{\bar{c}}_T(f)$, by Proposition 4.1 with $\nu = \mu_0$, and this bias is common to the three regimes.

**The energy of LAWGD.** With $m_k = M^{-1}\sum_j \hat{\varphi}_k(x_j)$ and $\mathcal{E} = \frac{M}{2}\sum_{k=1}^r \hat{\lambda}_k^{-1} m_k^2$, one has $\nabla_{x_i}\mathcal{E} = M\sum_k \hat{\lambda}_k^{-1} m_k \nabla_{x_i} m_k = \sum_k \hat{\lambda}_k^{-1} m_k \nabla\hat{\varphi}_k(x_i)$, which is minus the velocity of Remark 4.7, and $\frac{d}{dt}\mathcal{E} = -\sum_i |\dot{x}_i|^2$. Let $\nu_M$ be the empirical measure of the particles. Since $\int \hat{\varphi}_k\, d\pi = 0$, $\frac{2}{M}\mathcal{E} = \sum_{k=1}^r \hat{\lambda}_k^{-1}\big(\int \hat{\varphi}_k\, d(\nu_M - \pi)\big)^2$ is the squared maximum mean discrepancy between $\nu_M$ and $\pi$ for the kernel $\sum_{k=1}^r \hat{\lambda}_k^{-1}\hat{\varphi}_k \otimes \hat{\varphi}_k$, and for exact eigenpairs it is the truncation to the first $r$ modes of $\|d\nu_M/d\pi - 1\|^2_{\dot{H}^{-1}}$, defined by the same series.

### C.2 Same-sample coefficients in the 2-Wasserstein distance

**Corollary C.1** (Comparison with same-sample coefficients)**.** *Fix the eigenpairs and the rank independently of the transported sample, with $\hat{\lambda}_k > 0$ and $\mathrm{Var}_{\mu_0}(\hat{\varphi}_k) < \infty$ for $k \le r$, let $\bar{c}_k = \mathbb{E}_{\mu_0}\hat{\varphi}_k$ and $\hat{c}_k = M^{-1}\sum_j \hat{\varphi}_k(x_j)$, and let $\hat{\mu}_T^S$ be the measure transported with the coefficients $\hat{c}$ from the same confined initialisation. If Theorem 4.4 applies to the population field, both fields admit a common validity region with bounds $\theta_s$, $\hat{\Lambda}_{s,t}$ and $V_s$ (the latter for $|\nabla \log \hat{\rho}^{\bar{c}}_s|$), and $\|\hat{\varphi}_k\|_{\infty,B} + \|\nabla\hat{\varphi}_k\|_{\infty,B} \le K_r$, then*

$$\mathbb{E}\, W_2(\hat{\mu}_T^S, \pi) \le \mathcal{B}_T + \mathcal{A}_T\Delta_M + 2R_B\,\hat{p}_{\mathrm{exit}}^{1/2}, \qquad \mathcal{A}_T \coloneqq 2K_r\int_0^T e^{\hat{\Lambda}_{s,T}}\frac{1+V_s}{\theta_s}\,ds, \tag{9}$$

*with $\Delta_M \coloneqq \sum_{k\le r}(\mathrm{Var}_{\mu_0}(\hat{\varphi}_k)/M)^{1/2}$, where $\mathcal{B}_T$ is the bound of Theorem 4.4 for the population field and $\hat{p}_{\mathrm{exit}}$ the expected fraction of particles whose population-field trajectory leaves $G^\delta$, or separates from its same-sample counterpart by more than $\delta$, before $T$.*

The term $\mathcal{A}_T\Delta_M$ is of order $M^{-1/2}$ for fixed $r$, and the last term is controlled by the mass that leaves the validity region. The corollary bounds the cost of same-sample coefficients, and their gain is the cancellation of Proposition 4.5.

*Proof of Corollary C.1.* Condition on the eigenpair data and fix the source sample. With $\delta_s = \hat{\rho}^{\hat{c}}_s - \hat{\rho}^{\bar{c}}_s = \sum_{k\le r} e^{-\hat{\lambda}_k s}(\hat{c}_k - \bar{c}_k)\hat{\varphi}_k$ and $\hat{\rho}^{\hat{c}}_s \ge \theta_s/2$ on the common region,

$$\sup_{G_s}|\hat{v}^{\hat{c}}_s - \hat{v}^{\bar{c}}_s| = \sup_{G_s}\Big|\frac{\nabla\delta_s - \delta_s\nabla\log\hat{\rho}^{\bar{c}}_s}{\hat{\rho}^{\hat{c}}_s}\Big| \le \frac{2K_r(1+V_s)}{\theta_s}\sum_{k=1}^r|\hat{c}_k - \bar{c}_k|.$$

For a pair of trajectories that remains in the region, the argument of Lemma B.2 gives $|X^S_{T,j} - X^P_{T,j}| \le \mathcal{A}_T\sum_k|\hat{c}_k - \bar{c}_k|$, and the other pairs are at distance at most $2R_B$. Pairing the particles, the bound $\mathbb{E}|\hat{c}_k - \bar{c}_k| \le (\mathrm{Var}_{\mu_0}(\hat{\varphi}_k)/M)^{1/2}$ and the triangle inequality through the population measure give the claim. The comparison is pathwise, so that no independence between the field and the particles is needed. □

### C.3 Resolution of a fixed evaluation sample

In Section 5 the $M$ transported particles are compared with one fixed target sample $Y_1$ of size $M$, and the reference level is the mean distance between two independent target samples of size $M$,

over ten pairs. To first order, the sliced distance between point clouds is the $L^2(du\,d\theta)$ norm of the difference of their projected quantile fluctuations about those of $\pi$, and independent fluctuations add in squares, so that $\mathbb{E}\,SW_2^2(X,Y_1) \approx \mathbb{E}\,SW_2^2(X,\pi) + \mathbb{E}\,SW_2^2(Y_1,\pi)$ for particles $X$ independent of $Y_1$ and $\mathbb{E}\,SW_2^2(Y,Y') \approx 2\,\mathbb{E}\,SW_2^2(Y,\pi)$ for independent samples $Y, Y'$ of size $M$. Particles without fluctuation therefore have an error of about $2^{-1/2}$ times the reference level, particles with the fluctuation of an independent sample the reference level, and particles with twice this fluctuation on the modes that dominate the distance about $(3/2)^{1/2}$ times the reference level. By Theorem 4.6 these are the situations of regimes (S), (P) and (I) when the retained modes carry most of the fluctuation.

## D THE ORNSTEIN-UHLENBECK CASE

Let $\pi = N(0,1)$, $\lambda_k = k$, $\varphi_k = He_k/\sqrt{k!}$ and $\mu_0 = N(m_0, s_0^2)$ with $s_0 < 1$, $q := 1 - s_0^2$, $m_t := m_0 e^{-t}$ and $\sigma_t^2 := 1 - q e^{-2t}$. Then $\rho_t = N(\cdot; m_t, \sigma_t^2)/N(\cdot; 0, 1)$, the velocity $v_t(x) = (\sigma_t^{-2} - 1)x - m_t \sigma_t^{-2}$ is affine, the flow map $\Phi_{s,t}(x) = m_t + \sigma_t \sigma_s^{-1}(x - m_s)$ has Lipschitz constant at most $1/s_0$, $L_t^+ = \sigma_t^{-2} - 1$, and $W_2(\mu_t, \pi) = (m_t^2 + (\sigma_t - 1)^2)^{1/2}$.

**Proposition D.1.** *The coefficients of the source are $c_k = q^{k/2} He_k(m_0/\sqrt{q})/\sqrt{k!}$, and $c_k^2 \le e^{m_0^2/(2q)} q^k$ and $\tau_r^2 \le e^{m_0^2/(2q)} q^{r+1}/(1-q)$.*

The formula follows from the generating function $\sum_k t^k \mathbb{E}[He_k(X)]/k! = \exp\{t m_0 - q t^2/2\}$, and the bounds from Cramér's inequality $|He_k(y)| \le \sqrt{k!}\, e^{y^2/4}$ (Indritz, 1961). For $m_0 = 1.5$ the formula gives $\tau_r < 10^{-2}$ first at $r = 9$, 32 and 102 when $s_0 = 0.8$, 0.5 and 0.3.

**Theorem D.2** (Explicit rates)**.** *Assume $s_0 < 1$, $b \ge |m_0| + 3$, exact eigenpairs and population coefficients $c_k$ (regime (P)), threshold $\varepsilon_0$, bound $v_{\max} \ge V_B + C'(q, V_B)$ with $V_B := b(s_0^{-2} - 1) + |m_0| s_0^{-2}$, rank $r$ and a fixed horizon $T$. Then, with constants $C, c$ depending only on $(m_0, s_0, b, \delta, T)$,*

$$\mathbb{E}\,W_2(\hat{\mu}_T, \pi) \le W_2(\mu_T, \pi) + S_M(\mu_T) + C(r+1)^c\, \mathcal{T}_r(\varepsilon_0) + \mathcal{E}_{\mathrm{tail}}(b), \tag{10}$$

*where $\mathcal{T}_r(\varepsilon_0) := q^{\frac{r+1}{2}} \min\{\varepsilon_0^{-1/2}, q^{-\frac{r+1}{4}}\} + b\, q^{\frac{r+1}{2(2-s_0^2)}} + b\, \varepsilon_0^{\frac{1}{1+q}}$ up to factors $e^{O(\sqrt{\log(1/\varepsilon_0)})}$ and $e^{O(\sqrt{r})}$, subpolynomial in $1/\varepsilon_0$ and in $q^{-r}$, and $\mathcal{E}_{\mathrm{tail}}(b) = O(e^{-(b-|m_0|-1)^2/4})$ up to a polynomial factor in $b$.*

*Sketch of proof.* Assumptions (A0) to (A4) hold on $B = [-b, b]$ with $K_B = e^{b^2/4}$, $\alpha = 0$, $V_G \le V_B$ and $H_G \le s_0^{-2} - 1$, by Cramér's inequality and the explicit flow. With $z = \sqrt{q} e^{-t}$ and $K' = e^{m_0^2/(4q)}$, Proposition D.1 gives the envelope bounds of Lemma B.6 with $w(x) = e^{x^2/4}$, $S_t = K' z^{r+1}/(1-z)$, $a_1 = O(\sqrt{r+1})$ and $a_2 = O(r+1)$. For $\kappa = (r+1)^{-3/2}$ the lemma gives a validity region with $\mathrm{pert}_t = O((r+1)^{-1/2})$, so that $e^{\hat{\Lambda}_{0,T}} \le s_0^{-1} e^{CT/\sqrt{r+1}}$. Since $\log \rho_t(x) - x^2/4$ is a concave quadratic, the region is an interval whose half-width $w_t$ satisfies $w_t^2 \ge \big((r+1)\log(1/z) - \frac{3}{2}\log(r+1) - c\big)/\big(\frac{1}{2\sigma_t^2} - \frac{1}{4}\big)$. Lemma B.7(i) and Gaussian tails then give $p_{\mathrm{exit}} = O\big((r+1)^c q^{(r+1)/(2-s_0^2)} + \varepsilon_0^{2/(1+q)} + e^{-(b-|m_0|-1)^2/2}\big)$, where the term in $\varepsilon_0$ comes from the set $\{\rho_t < 2\varepsilon_0 e^{x^2/4}\}$. Lemma A.3 and $\theta_t \ge \max\{2\varepsilon_0, S_t(r+1)^{3/2}\}$ give $J_2 = O\big(\tau_r \min\{\varepsilon_0^{-1/2}(r+1)^{-1/2}, (r+1)^{-5/4} q^{-(r+1)/4}\}\big)$. Inserting these bounds into Theorem 4.4 with $q_* \le p_{\mathrm{exit}} + J_2^2/\delta^2$ and $S_M(\mu_T) = \sigma_T S_M(\pi)$ gives (10), and the factors $e^{O(\sqrt{\log(1/\varepsilon_0)})}$ and $e^{O(\sqrt{r})}$ come from the shift of the centre and the margin in the Gaussian tail. □

The per-mode ratio $\max\{q^{1/4}, q^{1/(2(2-s_0^2))}\}$ equals 0.93 for $s_0 = 0.5$. It comes from the conversion of $L^2(\mu_s)$ norms to $L^2(\pi)$ norms at the edge of the validity region. In the bound of Corollary 4.2 the truncation term decays with the sharper ratio $q^{1/2} = 0.87$ per mode, since $\|(I - \Pi_r)\rho_0\|_{L^1(\pi)} \le \tau_r$, and the flux of the safeguards is the remaining term. With coefficients from a finite sample, regime (S) is covered by Proposition 4.5, Theorem 4.6 and Corollary C.1, and the rank dependence observed in Section 5 is that of same-sample BKT. For $\pi = N(0, I_d)$ and a product Gaussian source all objects tensorise, and the number of modes needed for a given accuracy is the product of the numbers needed in each coordinate. When source and target factorise, the coordinate-wise application of BKT in the product experiments avoids this product.

Table 2: Matched Ornstein-Uhlenbeck computations at $h = 0.025$, error to the target, paired endpoint discrepancy $D_T$ and residual-norm bound $\bar{D}_T$ (numerical quadrature), mean $\pm$ sd over ten matched realisations, for population (P), independent-sample (I) and same-sample (S) coefficients.

| $r$ | coefficients | $W_2$ | $D_T$ | $\bar{D}_T$ |
|---|---|---|---|---|
| 10 | P | $0.1020 \pm 0.0180$ | $0.0953 \pm 0.0058$ | $0.2529 \pm 0.0203$ |
| 10 | I | $0.1104 \pm 0.0144$ | $0.1011 \pm 0.0108$ | $0.2664 \pm 0.0260$ |
| 10 | S | $0.0961 \pm 0.0094$ | $0.1020 \pm 0.0096$ | $0.2563 \pm 0.0168$ |
| 20 | P | $0.0533 \pm 0.0135$ | $0.0249 \pm 0.0029$ | $0.0433 \pm 0.0062$ |
| 20 | I | $0.0668 \pm 0.0151$ | $0.0415 \pm 0.0116$ | $0.0590 \pm 0.0129$ |
| 20 | S | $0.0382 \pm 0.0065$ | $0.0456 \pm 0.0164$ | $0.0629 \pm 0.0191$ |
| 40 | P | $0.0447 \pm 0.0151$ | $0.0015 \pm 0.0008$ | $0.0200 \pm 0.0006$ |
| 40 | I | $0.0609 \pm 0.0170$ | $0.0330 \pm 0.0119$ | $0.0504 \pm 0.0146$ |
| 40 | S | $0.0227 \pm 0.0045$ | $0.0411 \pm 0.0163$ | $0.0545 \pm 0.0152$ |

### D.1 A MATCHED COMPARISON OF THE COEFFICIENT REGIMES

For the Ornstein-Uhlenbeck flow the deviation of computed paths from the exact ones is explicit. For points $x_j$ and absolutely continuous paths $Y_j$ with $Y_j(0) = x_j$, the residual $\mathcal{R}_j(t) = \dot{Y}_j(t) - v_t(Y_j(t))$ satisfies $Y_j(T) - \Phi_T(x_j) = \int_0^T K(s,T)\mathcal{R}_j(s)\,ds$ with $K(s,T) = \sigma_T/\sigma_s \le \max\{1, s_0^{-1}\}$, since $v_t$ is affine. With $\nu_0$ and $\hat{\mu}_T$ the empirical measures of the $x_j$ and of the $Y_j(T)$, this gives

$$W_2(\hat{\mu}_T, \pi) \le W_2(\Phi_T \# \nu_0, \pi) + D_T \le W_2(\Phi_T \# \nu_0, \pi) + \bar{D}_T, \quad \bar{D}_T = \int_0^T K(s,T)\|\mathcal{R}(s)\|_M\,ds,$$

where $\|z\|_M = (M^{-1}\sum_j |z_j|^2)^{1/2}$ and $D_T = \|Y(T) - \Phi_T(x)\|_M$, for any paths and in particular for those of regime (S). For the Runge-Kutta paths the residual is evaluated on the piecewise linear interpolant, and it includes the effects of the step, of the threshold and of the velocity bound. The one-dimensional error against the Gaussian target is computed by the exact formula for the monotone coupling on the quantile intervals.

**Matched experiment.** The parameters are $m_0 = 1.5$, $s_0 = 0.5$, $T = 6$ and $M = 2000$, with the analytical Hermite eigenpairs for $r \in \{10, 20, 40\}$, RK4 steps $h \in \{0.05, 0.025\}$ and ten realisations of the source, matched across ranks, steps and regimes. In (P) the coefficients are the population values, in (I) they are estimated from an independent source sample of size $M' = 2000$, and in (S) from the $M$ transported points. Within each realisation the transported sample, the coefficient sample and a target reference cloud of the same size are independent. The threshold and the velocity bound are those of Appendix E.1, dispersions are sample standard deviations over the ten realisations, and comparisons between regimes or steps are paired within a realisation.

**Results.** Table 2 reports the finer step and Figure 2 both steps. The exact-flow value $W_2(\Phi_T \# \nu_0, \pi)$ of the source clouds is $0.0448 \pm 0.0152$ and the independent-target reference $0.0391 \pm 0.0096$, and the population distance is $W_2(\mu_6, \pi) = 0.0037$. At $h = 0.025$, increasing $r$ from 10 to 40 reduces the distance to the target from $0.102$ to $0.045$ in (P), from $0.110$ to $0.061$ in (I) and from $0.096$ to $0.023$ in (S). At $r = 40$ the population coefficients track the exact endpoints most closely ($D_T = 0.0015 \pm 0.0008$ against $0.033 \pm 0.012$ for (I) and $0.041 \pm 0.016$ for (S)). The ordering (S) $<$ (P) $<$ (I) at every rank agrees with Theorem 4.6, in which $\sigma_I \ge \sigma_P$ and $\sigma_S$ is small when the retained modes carry most of the fluctuation. The same-sample field corrects the sampling error of the retained modes of the source, so that its trajectories depart most from the exact paired trajectories and end closest to the target. The bound $W_2(\Phi_T \# \nu_0, \pi) + \bar{D}_T$ has the right scale ($0.065$, $0.095$, $0.099$ against measured $0.045$, $0.061$, $0.023$ at $r = 40$). Halving the step changes the error by at most $0.001$ at $r = 20$ and 40, and by about $0.015$ at $r = 10$, where the threshold and the velocity bound are active, while the population computations at $r = 40$ activate neither.

**Test functions.** For the smooth test functions $\sin x$, $\tanh x$ and $e^{-x^2}$, Table 3 gives the standard deviation of the particle mean $M^{-1}\sum_j f(X_T^j)$ over 200 realisations of the source, with $M = M' = 2000$, $h = 0.025$ and the same threshold and velocity bound, together with the asymptotic standard

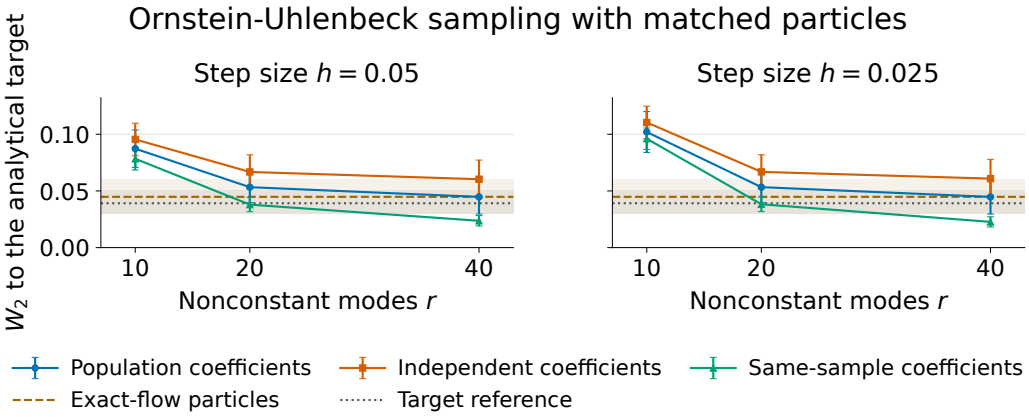



Figure 2: Matched Ornstein-Uhlenbeck sampling with analytical eigenpairs. Distance $W_2$ to the target for population (P), independent-sample (I) and same-sample (S) coefficients at both steps. Bars and bands show the mean $\pm$ sd over ten realisations. The exact-flow value is computed from the source clouds of BKT, and the reference from an independent target cloud of the same size.

| | | | (P) | | (I) | | (S) | |
|---|---|---|---|---|---|---|---|---|
| $f$ | $r$ | independent | observed | Theorem 4.6 | observed | Theorem 4.6 | observed | Theorem 4.6 |
| $\sin x$ | 20 | 14.7 | 15.8 | 14.7 | 22.3 | 20.9 | 0.75 | 0.69 |
| $\tanh x$ | 20 | 14.0 | 15.4 | 14.0 | 21.3 | 19.8 | 0.76 | 0.80 |
| $e^{-x^2}$ | 20 | 7.55 | 6.87 | 7.57 | 9.51 | 10.2 | 2.06 | 1.95 |
| $\sin x$ | 40 | 14.7 | 14.1 | 14.7 | 19.1 | 21.0 | 0.35 | 0.48 |
| $\tanh x$ | 40 | 14.0 | 13.5 | 14.0 | 18.4 | 19.8 | 0.20 | 0.19 |
| $e^{-x^2}$ | 40 | 7.55 | 7.45 | 7.54 | 10.8 | 10.7 | 0.52 | 0.53 |

Table 3: Matched Ornstein-Uhlenbeck setting with $M = 2000$. Standard deviation, in units of $10^{-3}$, of the particle mean of $f$ over 200 realisations in the three regimes, the asymptotic values $\sigma/\sqrt{M}$ of Theorem 4.6, and $\mathrm{sd}_\pi(f)/\sqrt{M}$ for an independent sample.

deviations $\sigma/\sqrt{M}$ of Theorem 4.6. These are computed from the population flow of $4 \cdot 10^5$ source points, with $\nabla m(\bar{c})$ obtained by central differences of step $10^{-3}$ on $2 \cdot 10^4$ common source points. The relative sampling error of the observed standard deviations is about $5\%$, and the predictions agree with the observations within ten percent, except for $\sin x$ at $r = 40$ in regime (S). There $|\beta| = 0.17$, against $|\beta| < 0.01$ for $\tanh x$ and $e^{-x^2}$, since $\sin x$ varies in the tails, where the safeguards act and the flux term of Proposition 4.1 enters $\beta$. Regime (P) has the standard deviation of an independent sample, regime (I) a larger one, and regime (S) reduces the standard deviation of an independent sample by factors between 4 and 20 at $r = 20$ and between 14 and 70 at $r = 40$. The factor is smallest for $e^{-x^2}$, whose composition with the flow is a narrow bump in the source variable.

## E Details of the numerical experiments

**Cost of the velocity.** With eigenfunctions given as combinations of $J$ dictionary functions, the coefficients $e^{-\hat{\lambda}_k t}\hat{c}_k$, which are common to all particles, are contracted with the eigenvectors first, in $O(rJ)$ operations, and an evaluation of the velocity then costs one evaluation of the dictionary and of its gradients at the $M$ particles, $O(MJd)$ operations in general.

### E.1 Ornstein-Uhlenbeck processes

**Settings.** The process is $dX = -X\,dt + \sqrt{2}\,dW$ with target $N(0, 1)$, the source is $N(1.5, 0.5^2)$, and the spectrum is the analytical one, $\lambda_k = k$ and $\varphi_k = He_k/\sqrt{k!}$. The $M = 10^4$ particles are transported up to $T = 6$ by the fourth-order Runge-Kutta method with step $0.05$, threshold

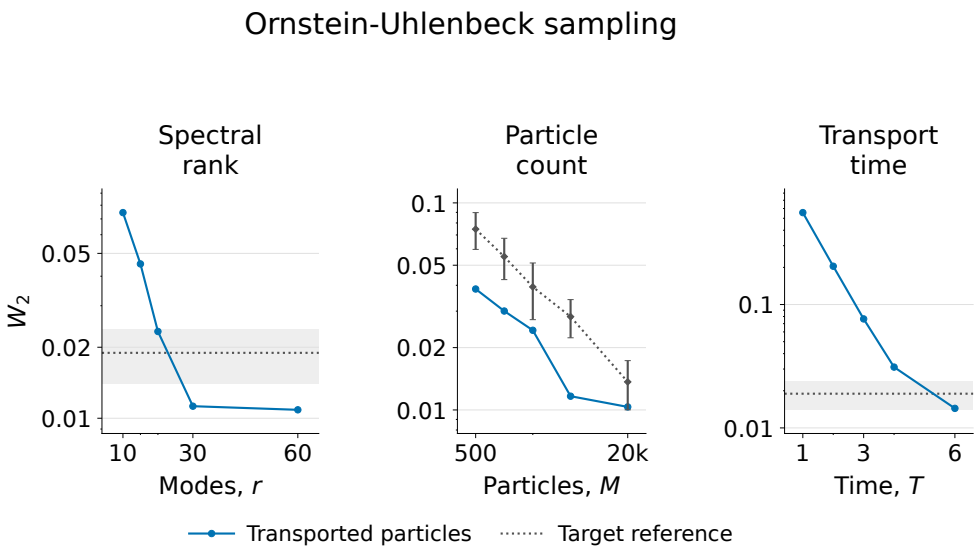



Figure 3: Ornstein-Uhlenbeck process with the analytical spectrum and particle coefficients. Error against the number of modes ($M = 10^4$, $T = 6$, left), against the number of particles ($r = 40$, $T = 6$, middle, with the reference for each $M$) and against the horizon ($r = 40$, $M = 10^4$, right). The dotted line, band and bars give the reference mean $\pm$ sd over 10 repetitions.

$\varepsilon_0 = 10^{-3}$ and bound $v_{\max} = 20$, without spatial clipping or reflection, and the stability diagnostic below uses a grid on $[-4, 4]$. The coefficients $\hat{c}_k = \frac{1}{M}\sum_i \varphi_k(X_0^i)$, $\hat{c}_0 = 1$, are computed once from the initial particles that are then transported (regime (S)) and propagated by $e^{-\lambda_k t}$. The error is the empirical $W_2$ distance between the sorted particles and the quantiles of $N(0,1)$, and the reference is the same distance for an independent sample of $N(0,1)$ of the same size, with mean $\pm$ sd over 10 independent samples ($0.0189 \pm 0.0049$ for $M = 10^4$).

**Dependence on the parameters.** Figure 3 gives the three variations quoted in Section 5. The error equals $0.0744$, $0.0451$, $0.0233$, $0.0112$ and $0.0109$ at $r = 10, 15, 20, 30, 60$. For $r = 40$ it equals $0.0384$, $0.0300$, $0.0242$, $0.0117$ and $0.0103$ at $M = 500, 10^3, 2{\cdot}10^3, 5{\cdot}10^3, 2{\cdot}10^4$, against references $0.0747 \pm 0.0151$, $0.0550 \pm 0.0124$, $0.0393 \pm 0.0120$, $0.0282 \pm 0.0059$ and $0.0137 \pm 0.0037$, and $0.556$, $0.204$, $0.076$, $0.031$ and $0.014$ at $T = 1, 2, 3, 4, 6$. The error of BKT lies below the reference at every $M$ in this variation. The coefficients are computed from the transported initial particles (regime (S)), and the reduction below the reference is the cancellation of Proposition 4.5.

### A TEN-DIMENSIONAL PROCESS WITH NEAREST-NEIGHBOUR COUPLING

**Model and analytical eigenpairs.** Let $V_\gamma(x) = \frac{1}{2}|x|^2 + \frac{\gamma}{2}\sum_{j=1}^{9}(x_{j+1} - x_j)^2$, $A_\gamma = I + \gamma E^\top E$ and $\pi_\gamma = N(0, A_\gamma^{-1})$, with $\gamma = 0.15$, for which the correlations of neighbouring coordinates lie between $0.1170$ and $0.1235$. For $\gamma = 0$ the target is the product Gaussian. Write $A_\gamma = Q\,\mathrm{diag}(a_1, \dots, a_{10})Q^\top$ with $Q$ orthogonal and $z_j = \sqrt{a_j}\, q_j^\top x$. The eigenfunctions and decay rates of $L$ are $\varphi_{\boldsymbol{n}}(x) = \prod_j h_{n_j}(z_j)$ and $\lambda_{\boldsymbol{n}} = \sum_j n_j a_j$, with $h_k = He_k/\sqrt{k!}$. The isotropic source $N(0.5\mathbf{1}, 0.25I)$ and the target remain products in the coordinates $z$, so the ratio factorises, $\hat{\rho}_t(x) = \prod_j \hat{\rho}_{j,t}(z_j)$ with $\hat{\rho}_{j,t}(z) = 1 + \sum_{k \le 30} e^{-k a_j t}\hat{c}_{jk}h_k(z)$ and $\hat{c}_{jk} = M^{-1}\sum_i h_k(z_{j,0}^i)$, the velocity is $\dot{z}_j = -a_j \partial_{z_j} \log \hat{\rho}_{j,t}$, and the particles are mapped back to the coordinates $x$ for the evaluation. The coupling is thus removed by the rotation, and the experiment concerns a correlated ten-dimensional target with analytical eigenpairs.

**Protocol and results.** The computations use $M = 5000$ particles, ten realisations of the source, with the same initial cloud across cases, step $0.05$, threshold $10^{-3}$, a componentwise velocity bound $20$ in the coordinates $z$ and no domain restriction. The error is the sliced $W_2$ (64 fixed directions)

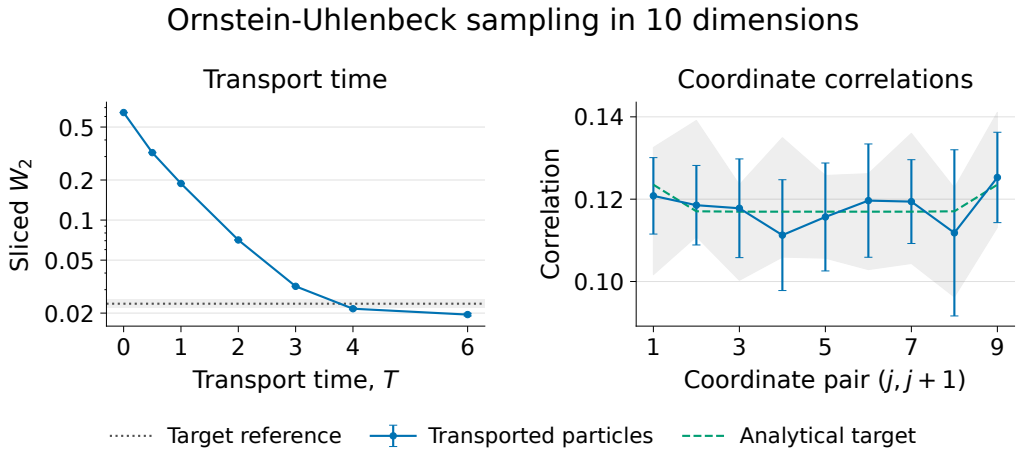



Figure 4: Ten-dimensional Ornstein-Uhlenbeck process with coupling $\gamma = 0.15$. Sliced $W_2$ against the transport time (left) and neighbouring-coordinate correlations of the transported particles against the target (right), mean $\pm$ sd over ten realisations. The reference is direct sampling from the analytical Gaussian.

of the transported particles to the analytical projected Gaussian laws, and the reference is the same distance for 10 independent target samples of size $M$ ($0.0235 \pm 0.0018$ at $\gamma = 0.15$ and $0.0262 \pm 0.0026$ at $\gamma = 0$). The error decreases at the rate $e^{-T}$ and then approaches the finite-particle level, from $0.643$ at $T = 0$ to $0.0195 \pm 0.0005$ at $T = 6$ (Figure 4, and $0.0219 \pm 0.0004$ against $0.0262 \pm 0.0026$ at $\gamma = 0$). The relative covariance error at $T = 6$ is $0.042$ against $0.047$ for the reference, and the neighbouring correlations and the $(x_1, x_2)$ marginal agree with the target (Figures 4 and 5). The paths of $0.26 \pm 0.10\%$ of the particles meet the threshold or the componentwise velocity bound at $\gamma = 0.15$ ($0.38 \pm 0.10\%$ at $\gamma = 0$), and the error of the remaining particles is $0.0198 \pm 0.0006$.

**Verification of the expansion.** The truncated expansion $1 + \sum_{k \le r} e^{-kt} c_k \varphi_k$ with population coefficients is compared with the closed-form ratio $\rho_t = N(1.5e^{-t}, 1 - 0.75e^{-2t})/N(0,1)$ on the grid, and the closed-form $\chi^2$ sum $\sum_k e^{-2kt} c_k^2$ with quadrature. The differences at finite $r$ are truncation and quadrature errors and decrease as $r$ grows.

**Positivity and stability diagnostics (particle coefficients).** On the domain, the minimum of the truncated ratio is $-4.89$, $-0.44$, $-0.40$ at $t = 0$ for $r = 10, 20, 40$, becomes positive at $t \approx 0.70$, $0.45$, $0.30$, and is $0.031$ at $t = 1$ for every $r$. Truncation produces negative densities at early times, which is why Section 4 confines the analysis to a validity region. The one-sided Lipschitz integral $\int_0^T (L_t^+)_+ dt$ of the approximate velocity is $213$, $97$, $48$ over the whole domain and $36.2$, $4.9$, $4.5$ over a screened region, against the exact value $\log(1/s_0) = 0.69$. The screening thus removes almost all of the growth, which is concentrated where the truncated ratio is small and where the analysis uses only the bound of the field. With $\epsilon_t := \max_{\text{grid}} |\hat{\rho}_t - \rho_t|$, the grid statistics use the points where $\rho_t \ge \sqrt{\epsilon_t}$, and the particle statistics the particles with $|x| \le b$ and $\rho_t(x) \ge 2\epsilon_t$. Along the particle paths, $99.6\%$ of the particle steps fall in the screened region for $r = 20$ and $40$, with no activation of the threshold or of the velocity bound inside it. Counted per particle, the paths of $3.2\%$, $0.9\%$ and $0.02\%$ of the particles meet the threshold or the velocity bound for $r = 10$, $15$ and $20$, none for $r = 30$ and $0.02\%$ for $r = 40$.

### E.2 Double wells, multi-wells and products

**Protocol.** The repeated experiments of this section have ten realisations, and the rank and dictionary scans use the first one. A realisation consists of a new source sample and, for the estimated spectra, of a new trajectory from which the eigenpairs are estimated. For the double-well, multi-well and product systems the transported particles are compared with one fixed target sample, and the reference compares two independent target samples. For the ten-dimensional Ornstein-Uhlenbeck process both are compared with the analytical projected Gaussian laws (Appendix E.1),

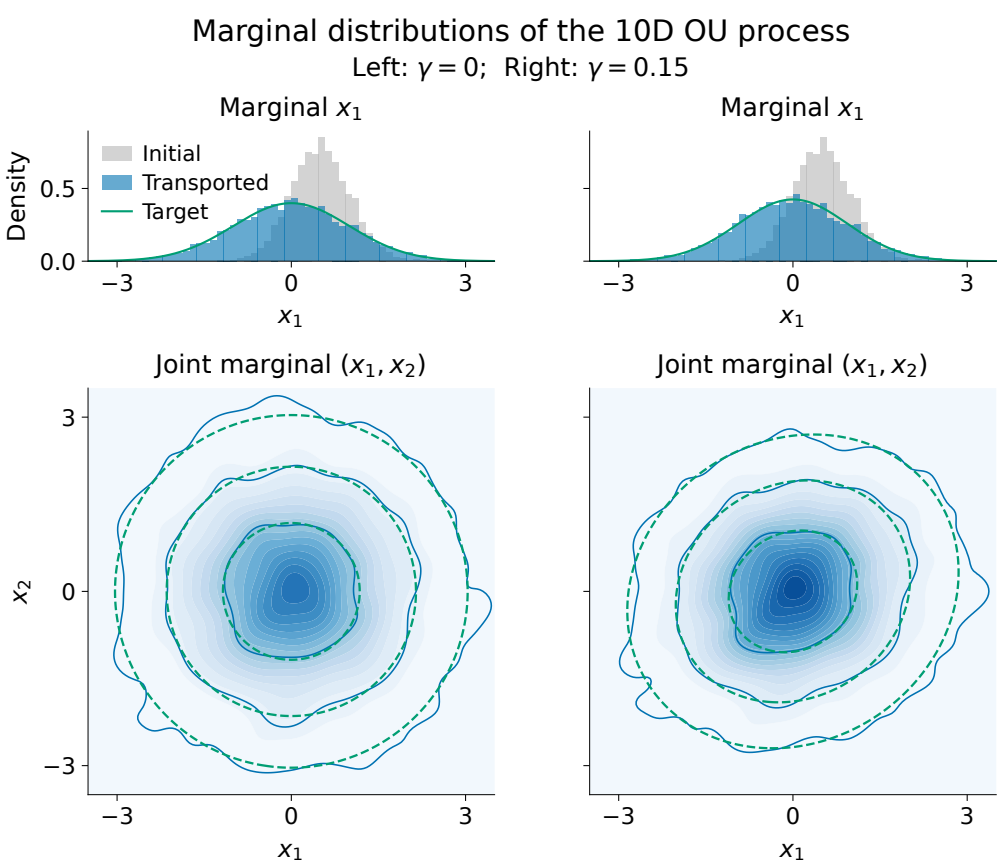



Figure 5: Marginals of the transported particles in ten dimensions at $T = 6$ (first realisation of the source), $\gamma = 0$ (left) and $\gamma = 0.15$ (right). The top row shows the $x_1$ marginal of the initial particles, of the transported particles and of the analytical target, the bottom row the $(x_1, x_2)$ joint marginal of the transported particles (kernel density estimate) against the analytical contours enclosing $50\%$, $90\%$ and $99\%$ of the probability.

and in one dimension both the transported and the reference particles are compared with the target quantiles. The main dictionary has $J = 145$ functions in two dimensions and $J = 657$ in ten, and a one-dimensional factor spectrum (FD-1401) is used for the separable systems. The diffusion $dX = -\nabla V\, dt + \sqrt{2}\, dW$ is simulated with step $2 \cdot 10^{-3}$ after 20 time units of equilibration, with the drift-implicit Euler scheme for the polynomial multi-wells, and pairs are taken at lag $0.1$. The sample sizes $10^4$ and $10^5$ are initial segments of the trajectory of $2 \cdot 10^5$ pairs of each realisation. The eigenpairs are computed by reversible gEDMD from the matrices $G$ and $D$, with whitening of $G$ at tolerance $10^{-8}$ and normalisation in $L^2(\pi)$. In two dimensions the dictionaries are $12 \times 12$ Gaussians plus the constant ($J = 145$, width $7.5\%$ of the domain) and tensor Legendre polynomials of total degree at most 16 ($J = 153$). In ten dimensions they are eight Gaussians per coordinate with products over neighbouring pairs ($J = 657$) and Legendre polynomials of total degree at most 3 ($J = 286$). For the products they are 80 Gaussians plus the constant on $[-2.8, 2.8]$ (width $0.14$, whitening $10^{-10}$) or Legendre polynomials of degree 32 per coordinate ($J = 33$). BKT uses fourth-order Runge-Kutta with step $0.05$ (double wells) or $0.02$ (multi-wells and products), $\varepsilon_0 = 10^{-3}$, $v_{\max} = 20$, horizon $T = 8/\lambda_1$ with $\lambda_1$ from the FD spectrum or from the fitted generator. In ten dimensions the candidate modes are those whose generator and lag eigenvalues agree within a relative tolerance of $0.5$, and among them the $r$ modes with the largest coefficients, computed from an independent source sample of size $10^5$, are used, so that the selected modes do not depend on the transported particles. The comparisons across sample sizes, dictionaries and coefficient regimes are paired within a realisation. For the two-dimensional double wells the source is a bump of centre 1 and width $0.6$ in $x_1$ times $N(0, 0.3)$ in $x_2$, for which $\rho_0$ is bounded when $\beta \le 1$ (the quadratic coefficient of $\log \rho_0$ in $x_2$ is $-(\frac{2}{3} - \frac{\beta}{2}) < 0$). In the ten-dimensional double well the source has

| system | $n$ | $M$ | $r$ | well masses (target) |
|---|---|---|---|---|
| 2D double well, $\beta = 0, 0.25, 0.5, 1$ | $2 \cdot 10^5$ | 2000 | 64 | – |
| 2D four-well product | $2 \cdot 10^5$ | 2000 | 64 | 0.23 to 0.27 (0.25 each) |
| 2D nine wells | $2 \cdot 10^5$ | 2000 | 64 | basins of Table 6 |
| 10D double-well product, $\beta = 0$ | $10^5$ | 1000 | 16 | left $0.55 \pm 0.01$ (0.50) |
| 10D and 50D double-well products | $2 \cdot 10^5$ | $2 \cdot 10^4$ | 16 per coordinate | – |

Table 4: Settings of the experiments of Table 1, with ten realisations each. The columns give the sample size $n$, the number $M$ of transported particles, the number $r$ of modes and the well masses of the transported particles (RBF, range or mean $\pm$ sd over the realisations) against the target.

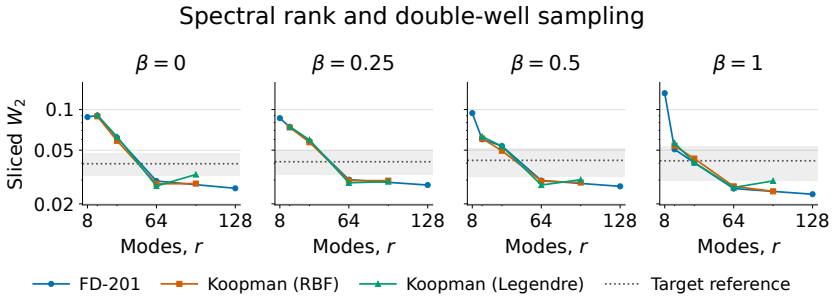



Figure 6: Two-dimensional double well. Error against the number of modes for one realisation of the source ($n = 2 \cdot 10^5$), with the FD-201, Koopman (RBF) and Koopman (Legendre) eigenpairs, for $\beta = 0, 0.25, 0.5, 1$. The dotted line and band show the reference mean $\pm$ sd over 10 repetitions.

the same bump in $x_1$ and nine independent $N(0, 0.5)$ harmonic coordinates, which are at equilibrium. The domain $B = [-2.8, 2.8] \times [-2.5, 2.5]$ of the two-dimensional double wells is enforced by componentwise projection of the trajectory data, of the reference samples and of the transported particles, the latter at the arguments of the Runge-Kutta stages and at the end of each step, so that the finite-step reference law is that of a projected Euler scheme. The paths of fewer than $0.3\%$ of the particles meet the threshold, the velocity bound or the projection ($0.06$ to $0.28\%$ over the spectra and couplings), and the error of the remaining particles differs from that of all particles by at most $0.003$. Halving the transport step changes the $\beta = 0.5$ error from $0.0305$ to $0.0309$ (RBF) and by at most $0.0010$ for the other spectra. The FD-201 eigenpairs of the multi-wells use Neumann tail truncation at the points where $U = 30$. The polynomial targets are defined on the whole plane and the particles are never clipped or reflected. Their normalisation, distribution function, basin masses (separated by the saddles) and barriers are obtained by direct integration of $e^{-U}$ on $40\,001$ points up to $U = 40$. The sample size is $n = 2 \cdot 10^5$ for the two-dimensional systems and $10^5$ for the ten-dimensional ones, and Table 7 gives the error at all three sizes.

Here *product* refers to the invariant density, since an additive potential $V(x) = \sum_i V_i(x_i)$ gives $\pi(x) = \prod_i \pi_i(x_i)$. The uncoupled double well has one double-well factor and Gaussian factors, whereas the four-well target and the ten- and fifty-dimensional double-well products have two, ten and fifty double-well factors, respectively.

**Dictionary sensitivity.** Table 5, Table 6 and Figure 6 compare the two dictionaries at the sample sizes of Table 4. Since BKT itself is unchanged, the comparison measures the sensitivity of the spectral estimation to the dictionary. On the two-dimensional double wells and the four-well the two dictionaries are equivalent within the fluctuation between realisations, and both are below the reference. On the nine-well the Legendre dictionary fails in five of the ten realisations (errors $0.90$ to $2.71$, against $0.036$ to $0.052$ in the other five), with a non-monotone dependence on $r$ ($0.9$ to $2.3$ for $r = 8$ to $96$ in the first realisation). The polynomials grow without bound outside the domain, where the estimated eigenfunctions are unreliable, and a few particles are driven far into the tails. The paths of $0.2\%$ of the particles meet a safeguard, and the error of the remaining particles is $0.042 \pm 0.007$. The same instability appears on the four-well at $r = 32$, in nine of ten realisations (median $2.51$), and not at $r = 64$ ($0.026 \pm 0.002$). The RBF dictionary, whose functions are bounded, is stable in all realisations. In ten dimensions it gives the smaller sliced error ($0.061$ against $0.072$). The instability

| system | FD | Koopman (RBF) | Koopman (Legendre) | reference |
|---|---|---|---|---|
| 2D double-well product, $\beta = 0$ | $0.027 \pm 0.002$ | $0.028 \pm 0.003$ | $0.028 \pm 0.003$ | $0.040 \pm 0.008$ |
| 2D double well, $\beta = 0.25$ | $0.028 \pm 0.002$ | $0.029 \pm 0.003$ | $0.030 \pm 0.003$ | $0.041 \pm 0.008$ |
| 2D double well, $\beta = 0.5$ | $0.028 \pm 0.002$ | $0.031 \pm 0.004$ | $0.029 \pm 0.003$ | $0.042 \pm 0.010$ |
| 2D double well, $\beta = 1$ | $0.027 \pm 0.002$ | $0.027 \pm 0.003$ | $0.028 \pm 0.003$ | $0.042 \pm 0.012$ |
| 2D four-well product | $0.024 \pm 0.003$ | $0.028 \pm 0.002$ | $0.026 \pm 0.002$ | $0.043 \pm 0.008$ |
| 2D nine wells | $0.033 \pm 0.003$ | $0.055 \pm 0.006$ | fails in 5 of 10 | $0.034 \pm 0.008$ |
| 10D double-well product, $\beta = 0$ | $0.053 \pm 0.002$ | $0.061 \pm 0.002$ | $0.072 \pm 0.008$ | $0.059 \pm 0.004$ |
| 10D double-well product | $0.0151 \pm 0.0003$ | $0.0195 \pm 0.0021$ | $0.0195 \pm 0.0021$ | $0.0156 \pm 0.0014$ |
| 50D double-well product | $0.0158 \pm 0.0002$ | $0.0188 \pm 0.0012$ | $0.0188 \pm 0.0012$ | $0.0162 \pm 0.0008$ |

Table 5: Dictionary comparison (settings of Table 4), sliced $W_2$ with the RBF and the Legendre dictionaries, mean $\pm$ sd over ten realisations, and the reference (mean $\pm$ sd over 10 repetitions). The nine-well Legendre entry is discussed in the text.

| system | spectrum | $r = 8$ | 16 | 32 | 64 | 96 | 128 |
|---|---|---|---|---|---|---|---|
| 4 wells | FD-201 | 0.129 | 0.124 | 0.051 | 0.024 | – | 0.022 |
| | Koopman (RBF) | – | 0.125 | 0.050 | 0.029 | 0.029 | – |
| | Koopman (Legendre) | – | 0.128 | 2.44 | 0.027 | 0.027 | – |
| 9 wells | FD-201 | 0.325 | 0.199 | 0.049 | 0.036 | – | 0.032 |
| | Koopman (RBF) | 0.309 | 0.188 | 0.063 | 0.063 | 0.062 | – |
| | Koopman (Legendre) | 0.91 | 1.56 | 2.34 | 1.64 | 1.88 | – |

Table 6: Dependence on the number of modes for the multi-wells, for one realisation of the source and the sample sizes of Table 4, in sliced $W_2$ against $Y_1$. At $r = 3$ the four-well gives 0.32 with every spectrum, and the reference levels are 0.043 and 0.034. The double wells are shown in Figure 6, and at $\beta = 0.5$ the RBF errors are 0.061, 0.050, 0.030 and 0.029 at $r = 16, 32, 64, 96$.

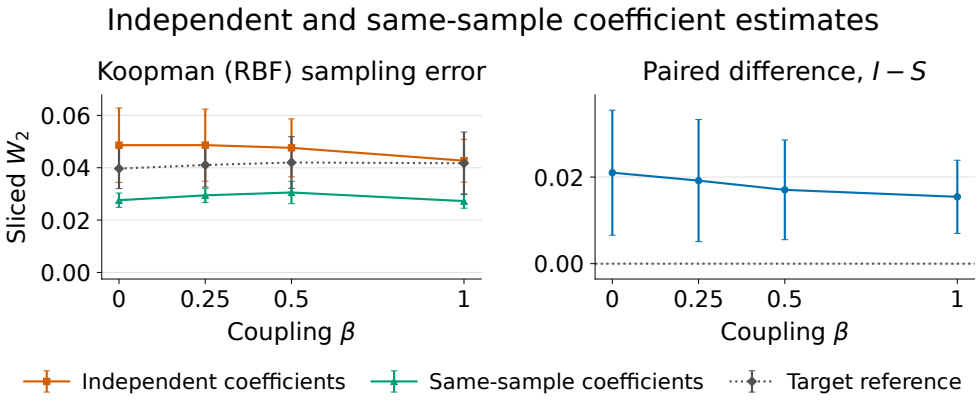



Figure 7: Two-dimensional double well. Error with coefficients from the transported particles (regime (S)) and from an independent source sample (regime (I)), mean $\pm$ sd over ten realisations, and reference mean $\pm$ sd over 10 repetitions.

is therefore a property of the unbounded polynomial dictionary, and BKT inherits the stability of the bounded one.

**Coefficients from an independent sample.** On the double wells with the RBF eigenpairs, coefficients from an independent source sample of size $M' = M = 2000$ (regime (I)) give $0.049 \pm 0.014$, $0.049 \pm 0.014$, $0.048 \pm 0.011$ and $0.043 \pm 0.008$ at $\beta = 0, 0.25, 0.5, 1$, against $0.028$, $0.029$, $0.031$ and $0.027$ with the coefficients of the transported particles (regime (S)). The paired difference is $0.015$ to $0.021$ and is positive in 39 of the 40 pairs (Figure 7). Relative to the reference mean, regime (S) gives $0.65$ to $0.73$ and regime (I) $1.02$ to $1.23$, against $2^{-1/2} \approx 0.71$ for an exact representation of $\pi$ and $(3/2)^{1/2} \approx 1.22$ for a doubled sampling error of the retained modes (Appendix C.3), in agreement with Proposition 4.5, by which the sampling error of the retained modes cancels in (S) and adds to the fluctuation of the coefficients in (I).

**Sample size.** Table 7 gives the RBF error at the three sample sizes. It decreases with $n$ in two and in ten dimensions, and the experiments of Table 1 use $n = 2 \cdot 10^5$ in two dimensions and $10^5$ in ten.

| system | $n = 10^4$ | $n = 10^5$ | $n = 2 \cdot 10^5$ |
|---|---|---|---|
| 2D double-well product, $\beta = 0$ | $0.035 \pm 0.004$ | $0.029 \pm 0.003$ | $0.028 \pm 0.003$ |
| 2D double well, $\beta = 0.25$ | $0.036 \pm 0.004$ | $0.031 \pm 0.004$ | $0.029 \pm 0.003$ |
| 2D double well, $\beta = 0.5$ | $0.037 \pm 0.008$ | $0.031 \pm 0.004$ | $0.031 \pm 0.004$ |
| 2D double well, $\beta = 1$ | $0.053 \pm 0.020$ | $0.028 \pm 0.002$ | $0.027 \pm 0.003$ |
| 2D four-well product | $0.048 \pm 0.013$ | $0.029 \pm 0.004$ | $0.028 \pm 0.002$ |
| 2D nine wells | $0.087 \pm 0.024$ | $0.056 \pm 0.006$ | $0.055 \pm 0.006$ |
| 10D double-well product, $\beta = 0$ | $0.068 \pm 0.004$ | $0.061 \pm 0.002$ | $0.060 \pm 0.004$ |

Table 7: Koopman (RBF) error (mean $\pm$ sd over ten realisations) against the sample size $n$ of the trajectory data.

| | sliced $W_2$ | | time (s) | |
|---|---|---|---|---|
| | $\beta = 0$ | $\beta = 0.5$ | $\beta = 0$ | $\beta = 0.5$ |
| BKT | $0.028 \pm 0.003$ | $0.031 \pm 0.004$ | 9.5 | 8.7 |
| kernel particle flow | $0.039 \pm 0.013$ | $0.048 \pm 0.027$ | 53.6 | 45.9 |
| reference | $0.040 \pm 0.008$ | $0.042 \pm 0.010$ | – | – |

Table 8: Two-dimensional double well, BKT and the kernel particle flow, mean $\pm$ sd over ten realisations, and the reference (mean $\pm$ sd over 10 repetitions). The times include the estimation of the eigenpairs (3.1 and 3.2 s) and of the drift (1.3 s).

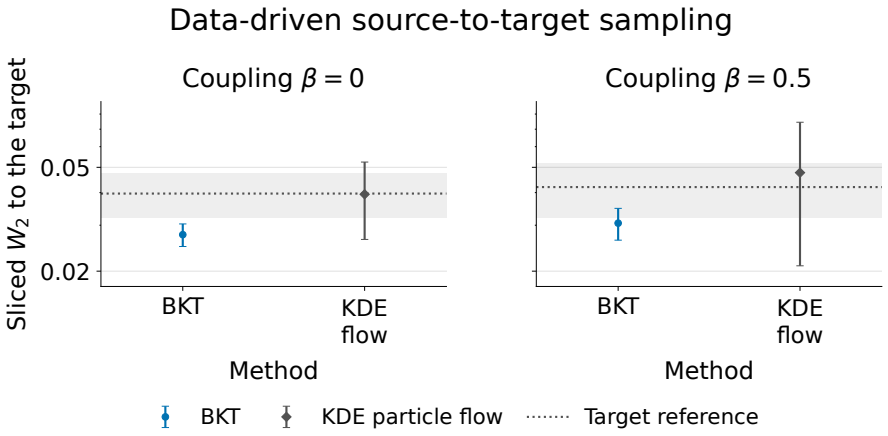



Figure 8: Two-dimensional double well at $\beta = 0$ and 0.5. Sliced $W_2$ of BKT and of the kernel particle flow, mean $\pm$ sd over ten realisations. The dotted line and the band show the reference mean $\pm$ sd over 10 repetitions.

**Comparison with a kernel particle method.** On the double wells at $\beta = 0$ and 0.5, with $n = 2 \cdot 10^5$ pairs and 2000 output particles (ten realisations, source as above), BKT is compared with the deterministic particle method for diffusion (Degond & Mustieles, 1990; Chertock, 2017), an interacting particle flow along $\hat{b} - \nabla \log \hat{\mu}_t$, where $\hat{b}$ is the drift estimated by regression of the increments on the gradients of the dictionary and $\hat{\mu}_t$ a Gaussian kernel density estimate of the current particles (Silverman bandwidth). Table 8 and Figure 8 give the errors and the computation times on one core, estimation included. With the coupling $\beta = 0.5$ the kernel flow has a larger mean error and twice the spread ($0.048 \pm 0.027$ against $0.039 \pm 0.013$ at $\beta = 0$), while the error of BKT is unchanged within the fluctuation between realisations ($0.031 \pm 0.004$ against $0.028 \pm 0.003$). At the prescribed observation time the error falls below the reference mean plus one standard deviation in ten realisations out of ten for BKT, and in nine and eight for the kernel flow at $\beta = 0$ and 0.5. Unlike a subsample of stationary data, BKT provides a map from any admissible source to the target, which also applies to non-stationary data.

**Nine wells.** On this system BKT reaches the reference with the FD spectrum, and the comparison with the estimated spectra isolates the spectral estimation. BKT with the FD-201 eigenpairs gives $0.033 \pm 0.003$ at $r = 64$ ($0.032$ at $r = 128$ in the first realisation) against a reference of $0.034 \pm 0.008$,

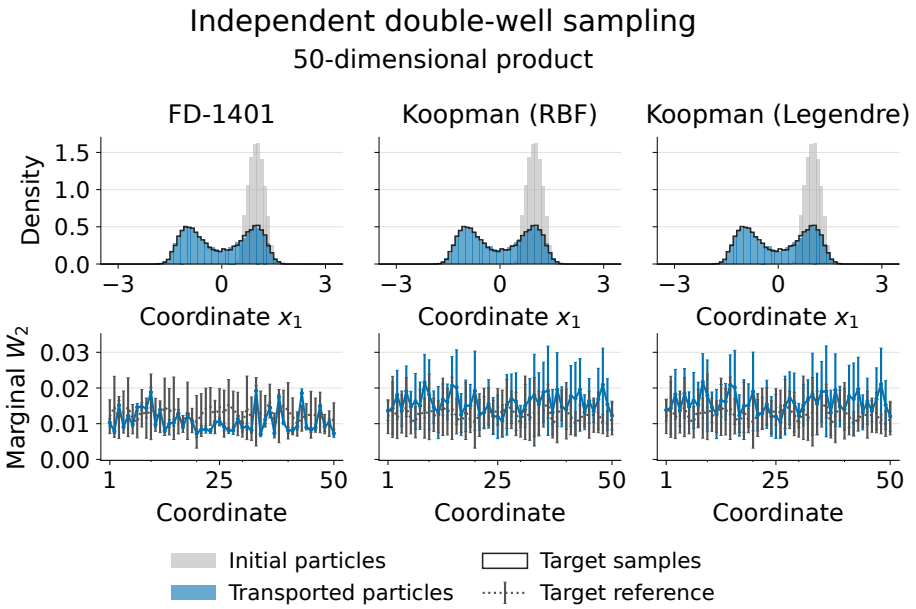



Figure 9: Fifty-dimensional double-well product, histograms of the first coordinate of the initial particles, the transported particles and a target sample (top, first realisation) and marginal errors per coordinate, mean ± sd over ten realisations, with the reference mean ± sd (bottom), for the FD-1401 factor spectrum and the Koopman (RBF) and Koopman (Legendre) factor spectra.

while Koopman (RBF) gives $0.055 \pm 0.006$, does not improve beyond $r = 32$ (Table 6), and the Legendre dictionary fails in half of the realisations (Table 5). A source concentrated in one narrow well needs many modes, and the estimated spectrum supplies about 30 accurate ones. With accurate eigenpairs BKT thus reaches the reference, and the remaining error is that of the spectral estimation.

**Separable products.** For $V = \sum_i (x_i^2 - 1)^2$ BKT acts coordinate by coordinate with the FD-1401 factor spectrum or with coordinate-wise gEDMD eigenpairs ($n = 2 \cdot 10^5$, $r = 16$ per coordinate, $M = 2 \cdot 10^4$, $T = 8/\lambda_1 = 10.69$, step 0.02), the source and the target being products (Figure 9 for $d = 50$). The sliced distances are $0.0151 \pm 0.0003$ (FD), $0.0195 \pm 0.0021$ (RBF) and $0.0195 \pm 0.0021$ (Legendre) against $0.0156 \pm 0.0014$ for $d = 10$, and $0.0158 \pm 0.0002$, $0.0188 \pm 0.0012$ and $0.0188 \pm 0.0012$ against $0.0162 \pm 0.0008$ for $d = 50$. The energy distances are $0.0050 \pm 0.0010$ for both dictionaries against $0.0052 \pm 0.0013$ in ten dimensions and $0.0116 \pm 0.0009$ against $0.0120 \pm 0.0010$ in fifty. The total variation between the distribution of the number of negative coordinates and its target is $0.0092 \pm 0.0021$ against $0.0095 \pm 0.0035$ and $0.0191 \pm 0.0028$ against $0.0165 \pm 0.0032$ for the RBF dictionary. The largest marginal errors are $0.0356 \pm 0.0056$ (RBF) and $0.0354 \pm 0.0057$ (Legendre) in ten dimensions and $0.0387 \pm 0.0047$ and $0.0387 \pm 0.0048$ in fifty, against references of $0.0260 \pm 0.0055$ and $0.0297 \pm 0.0045$, with the FD factor at $0.0188 \pm 0.0014$ and $0.0201 \pm 0.0009$, so that the excess of the estimated factors comes from the one-dimensional spectral estimation. In each coordinate the paths of 0.12 to 0.15% of the particles meet the velocity bound. The two dictionaries give the same results, and the number of modes is additive over the coordinates.

### E.3 The dihedral marginal of alanine dipeptide

**Data and target.** The three public 250 ns trajectories of alanine dipeptide in explicit solvent (backbone dihedrals $(\phi, \psi)$ at 1 ps, 250,000 frames each, available through `mdshare`, Nüske et al., 2017) are assigned one each to estimation, validation and evaluation. The target is the smoothed empirical marginal of $(\phi, \psi)$ on the torus. The eigenpairs are those of the unit-mobility reversible diffusion with this invariant law, estimated by the Gram and Dirichlet matrices of the estimation frames (smoothed by a wrapped Gaussian of bandwidth 0.10 rad) with a real tensor-product Fourier dictionary of degree 28 in each angle (3,249 functions), a relative Gram cutoff of $10^{-12}$ and $r = 256$

| | draws | sliced $W_2$ | basin-mass TV | reference $W_2$, TV |
|---|---|---|---|---|
| initial source | 3 | $0.615 \pm 0.003$ | $0.604 \pm 0.019$ | – |
| Fourier degree 28, $r = 256$ | 3 | $0.060 \pm 0.019$ | $0.033 \pm 0.018$ | $0.051 \pm 0.021$, $0.023 \pm 0.014$ |

Table 9: Alanine dipeptide, sliced $W_2$ and basin-mass total variation of 1000 transported particles against 1000 evaluation frames, at $\hat{\lambda}_1 T = 8$, mean $\pm$ sd over the three paired designs. The last column gives the reference distances between two further disjoint clouds of evaluation frames.

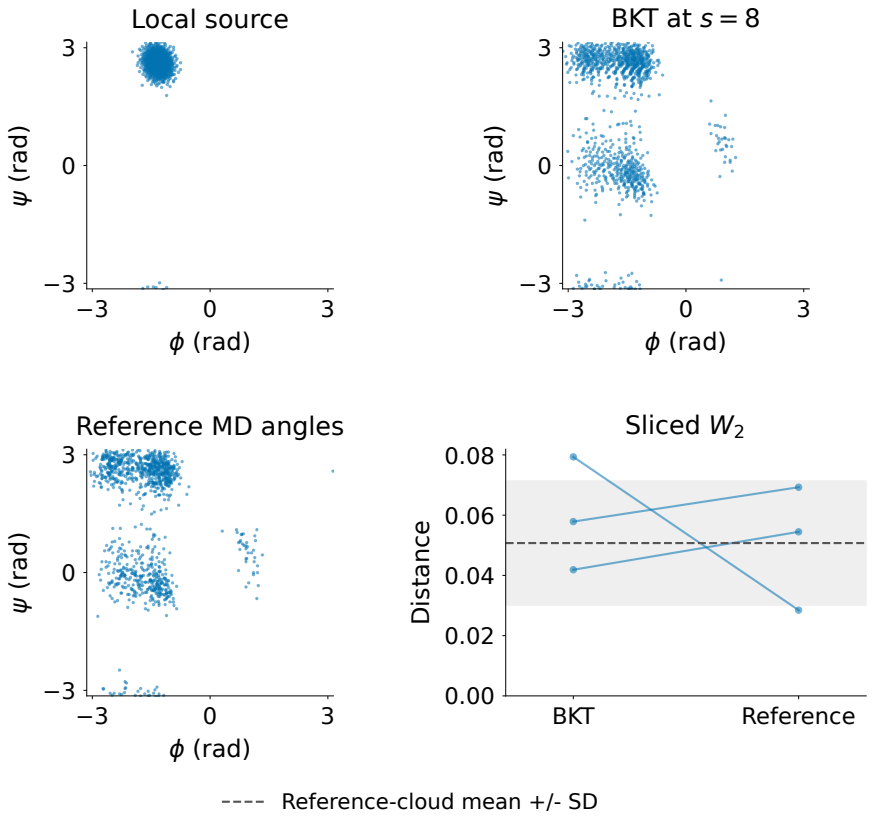



Figure 10: Alanine dipeptide with 256 Fourier modes, the local source, the transported particles at $s = 8$ and 1000 evaluation frames in the $(\phi, \psi)$ plane for the first design, followed by the sliced $W_2$ of the final particles and the paired reference for each of the three designs. The reference band denotes its mean $\pm$ sd.

nonconstant modes. The construction uses the Gram and Dirichlet forms of the frames, and the experiment transports one prescribed local source to the smoothed angular distribution.

**Transport and evaluation.** The source is a wrapped Gaussian localised in one basin, represented by 1000 initial points from a scrambled Sobol design. The coefficients are the empirical moments of these points (regime (S)). The equations are integrated to $s = \hat{\lambda}_1 T = 8$ with an adaptive eighth-order Runge-Kutta method (DOP853, relative and absolute tolerances $10^{-8}$ and $10^{-10}$, maximum step 0.05 in $s$). The density threshold is $10^{-3}$ and the velocity bound is 20 in unscaled time, or $20/\hat{\lambda}_1$ in $s$. The error is the sliced $W_2$ (32 fixed directions) in the embedding $(\cos\phi, \cos\psi, \sin\phi, \sin\psi)$ between the transported particles and 1000 frames of the evaluation trajectory, and the total variation between the masses of the four Ramachandran basins. The reference is the same distance between two further disjoint clouds of 1000 evaluation frames, that is, the error of a subsample of the trajectory. Three paired source and reference designs are used with the eigenpairs fixed. Figure 10 shows the final distributions of these runs, which are also those of the comparison below.

| $s = \hat{\lambda}_1 t$ | 0.1 | 1 | 4 | 8 |
|---|---|---|---|---|
| BKT | $0.091 \pm 0.011$ | $0.068 \pm 0.017$ | $0.060 \pm 0.019$ | $0.060 \pm 0.019$ |
| LAWGD | $0.578 \pm 0.003$ | $0.210 \pm 0.004$ | $0.054 \pm 0.021$ | $0.059 \pm 0.020$ |
| KDE | $0.204 \pm 0.005$ | $0.148 \pm 0.004$ | $0.148 \pm 0.004$ | $0.148 \pm 0.004$ |

Table 10: Alanine dipeptide, sliced $W_2$ to the evaluation frames along the normalised flow time, mean $\pm$ sd over three paired realisations, for BKT and the Laplacian-adjusted Wasserstein gradient descent with 256 eigenpairs, and for the kernel-density particle method with bandwidth $0.10$ rad.

**Results.** The final sliced error is $0.060 \pm 0.019$ against $0.051 \pm 0.021$ between reference clouds, and two of the three paired errors are at most their reference. The total variation of the four-basin masses is $0.033 \pm 0.018$ against $0.023 \pm 0.014$, and the mean mass of the rare positive-$\phi$ basin is $0.029$ against $0.042$, a difference smaller than the total variation between two reference clouds. The local-width ratio is $1.17 \pm 0.09$, computed from the median minor-axis standard deviation of twelve-neighbour local clouds, relative to the evaluation cloud (Figure 10, Table 9), consistent with the smoothing of the target at bandwidth $0.10$ rad. With i.i.d. initial points in place of the Sobol design the error is $0.058 \pm 0.021$. The paths of $0.4$ to $0.8\%$ of the particles meet the velocity bound and $0.1$ to $0.2\%$ the threshold.

**Comparison with interacting particle methods.** BKT and the Laplacian-adjusted Wasserstein gradient descent (LAWGD) of Chewi et al. (2020) are both run with the 256 nonconstant eigenpairs of the Fourier dictionary of degree 28, 1000 initial particles and their empirical coefficients (regime (S)). BKT keeps these coefficients fixed, whereas LAWGD recomputes the current moments in its velocity $-\sum_k \hat{\lambda}_k^{-1} \nabla \hat{\varphi}_k \, M^{-1} \sum_i \hat{\varphi}_k(X_t^i)$. The third method, KDE, adapts the classical diffusion-velocity particle method (Degond & Mustieles, 1990; Chertock, 2017) to periodic kernels. Its target density $\hat{\pi}$ is fitted once to the 250,000 estimation frames with bandwidth $0.10$ rad, and its current density $\hat{\mu}_s$ uses all 1000 particles, including self terms, with the same bandwidth. In normalised time $s = \hat{\lambda}_1 t$, with $\hat{\lambda}_1 = 0.0013776$, its velocity is $(\nabla \log \hat{\pi} - \nabla \log \hat{\mu}_s)/\hat{\lambda}_1$. KDE thus follows the same gradient flow with both densities estimated by kernels and without the generator, whereas BKT and LAWGD take the velocity from the eigenpairs of the generator, which carry the relaxation rates and the metastable structure of the dynamics. The discretisations and integration tolerances of the Fourier and kernel representations are specified in the accompanying code.

Figure 1 and Table 10 show three paired realisations of the source and of the reference with common estimation frames. BKT reaches the late-time error level by $s = 1$ ($0.068$, against $0.210$ for LAWGD and $0.148$ for KDE) and then remains near $0.060$. LAWGD reaches a comparable level at $s = 4$, while KDE remains near $0.148$. The two methods built on the spectrum of the generator thus end at $0.06$, and the method built on density estimates at more than twice this error. The kernel method converges to a stationary point of its own velocity $(\nabla \log \hat{\pi} - \nabla \log \hat{\mu}_s)/\hat{\lambda}_1$, at which the logarithmic gradients of the two kernel estimates agree at the particles. The final particles therefore retain the smoothing error of both estimates and the fluctuation of the estimate from 1000 particles, which are largest where the target density is small, and a longer flow does not reduce them. BKT estimates no density during the transport, since the law of the particles at time $s$ is represented by the propagated coefficients $e^{-\hat{\lambda}_k t}\hat{c}_k$. The velocity of BKT uses only the fixed initial coefficients, whereas LAWGD and KDE recompute empirical quantities from the evolving particles. The BKT error reaches its late-time level before that of LAWGD and varies little thereafter. On one thread the estimation of the eigenpairs takes 25 seconds and the integration to $s = 8$ takes $72 \pm 3$ seconds for BKT, $70 \pm 4$ for LAWGD and $1056 \pm 145$ for KDE, and each method first reaches its own late-time level after $34 \pm 6$, $54 \pm 5$ and $831 \pm 90$ seconds.