Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Poincaré and logarithmic Sobolev inequalities by decomposition of the energy landscape

by

Georg Menz and André Schlichting

Preprint no.: 40 2012

POINCARÉ AND LOGARITHMIC SOBOLEV INEQUALITIES BY DECOMPOSITION OF THE ENERGY LANDSCAPE

GEORG MENZ AND ANDRÉ SCHLICHTING

Abstract. We consider a diffusion on a potential landscape which is given by a smooth Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$ in the regime of low temperature ε. We proof the Eyring-Kramers formula for the optimal constant in the Poincaré (PI) and logarithmic Sobolev inequality (LSI) for the associated generator $L = \varepsilon\Delta - \nabla H \cdot \nabla$ of the diffusion. The proof is based on a refinement of the two-scale approach introduced by Grunewald, Otto, Westdickenberg and Villani [GOVW09]; and of the mean-difference estimate introduced by Chafaï and Malrieu [CM10]. The Eyring-Kramers formula follows as a simple corollary from two main ingredients: The first one shows that the PI and LSI constant of the diffusion restricted to a basin of attraction of a local minimum scales well in ε. This mimics the fast convergence of the diffusion to metastable states. The second ingredient is the estimation of a mean-difference by a weighted transport distance. It contains the main contribution to the PI and LSI constant, resulting from exponentially long waiting times of jumps between metastable states of the diffusion.

1. Introduction

Let us consider a diffusion on a potential landscape which is given by a sufficiently smooth *Hamiltonian function* $H : \mathbb{R}^n \to \mathbb{R}$. We are interested in the regime of low temperature $\varepsilon > 0$. The *generator* of the diffusion has the following form

$$L := \varepsilon\Delta - \nabla H \cdot \nabla. \tag{1.1}$$

The associated *Dirichlet form* is given for a test function $f : \mathbb{R}^n \to \mathbb{R}$ by

$$\mathcal{E}(f) := \int (-Lf) f \,\mathrm{d}\mu = \varepsilon \int |\nabla f|^2 \,\mathrm{d}\mu.$$

The corresponding diffusion ξ_t satisfies the stochastic differential equation, also called *over-damped Langevin equation* (cf. e.g. [LL10])

$$\mathrm{d}\xi_t = -\nabla H(\xi_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}B_t, \tag{1.2}$$

where B_t is the *Brownian motion* on $\mathbb{R}^n$. Under some growth assumption on H there exists an equilibrium measure of the according stochastic process, which is called *Gibbs measure* and is given by

$$\mu(\mathrm{d}x) = \frac{1}{Z_\mu} \exp\left(-\frac{H(x)}{\varepsilon}\right) \mathrm{d}x \qquad \text{with} \qquad Z_\mu = \int \exp\left(-\frac{H(x)}{\varepsilon}\right) \mathrm{d}x. \tag{1.3}$$

The evolution (1.2) of the stochastic process ξ_t can be translated into an evolution of the density of the process ξ_t. Namely, under the assumption that the law of the

---

Date: July 9, 2012.

2000 *Mathematics Subject Classification.* Primary 60J60; secondary 35P15; 49R05.

Key words and phrases. diffusion process, Eyring-Kramers formula, Kramers Law, metastability, Poincaré inequality, spectral gap, logarithmic Sobolev inequality, weighted transport distance.

initial state ξ_0 is absolutely continuous w.r.t. the Gibbs measure μ, the density $f_t\mu$ of the process ξ_t satisfies the *Fokker-Planck equation* (cf. e.g. [Øks98] or [Sch10])

$$\partial_t f_t = L f_t = \varepsilon \Delta f_t - \nabla H \cdot \nabla f_t.$$

We are particularly interested in the case where H has several local minima. Then for small ε, the process shows metastable behavior in the sense that there exists a separation of scales: On the fast scale, the process converges quickly to a neighborhood of a local minimum. On the slow scale, the process stays nearby a local minimum for an exponentially long waiting time after which it eventually jumps to another local minimum.

This behavior was first described in the context of chemical reactions. The exponential waiting time follows *Arrhenius' law* [Arr89] meaning that the mean exit time from one local minimum of H to another one is exponentially large in the energy barrier between them. By now, the Arrhenius law is well-understood even for non-reversible systems by the *Freidlin-Wentzell theory* [FW98], which is based on *large deviations*.

A refinement of the Arrhenius law is the *Eyring-Kramers formula* which additionally considers *pre-exponential* factors. The Eyring-Kramers formula for the Poincaré inequality (PI) goes back to Eyring [Eyr35] and Kramers [Kra40]. Both argue that also in high-dimensional problems of chemical reactions most reactions are nearby a single trajectory called *reaction pathway.* Evaluating the Hamiltonian along this *reaction coordinate* gives the classical picture of a double well potential (cf. Figure 1) in one dimension with an *energy barrier* separating the two local minima for which explicit calculations are feasible.

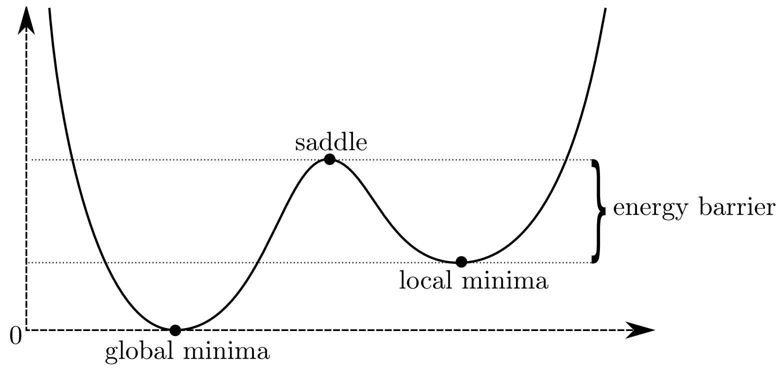

FIGURE 1. General double-well potential H on $\mathbb{R}$.

However, a rigorous proof of the Eyring-Kramers formula for the multidimensional case was open for a long time. For a special case, where all the minima of the potential as well as all the lowest saddle points in-between have the same energy, Sugiura [Sug95] defined an exponentially rescaled Markov chain on the set of minima in such a way that the pre-exponential factors become the transitions rates between the basins of the rescaled process. For the generic case, where the local minima and saddles have different energies, the group of Bovier, Eckhoff, Gayrard and Klein [BEGK04, BGK05] obtained first-order asymptotics that are sharp in the parameter ε. They also clarified the close connection between *mean exit times*, *capacities* and the exponentially small eigenvalues of the operator L given by (1.1). The main tool of [BEGK04, BGK05] is potential theory. The small eigenvalues are related to the mean exit times of appropriate subsets of the state space. Further, the mean exit times are given by *Newtonian capacities* which can explicitly be calculated in the regime of low temperature ε.

Shortly after, Helffer, Klein and Nier [HKN04, HN06, HN05] also deduced the Eyring-Kramers formula using the connection of the spectral gap estimate of the Fokker-Planck operator L given by (1.1) to the one of the *Witten Laplacian*. This approach makes it possible to get quantitative results with the help of semiclassical analysis. They deduced sharp asymptotics of the exponentially small eigenvalues of L and gave an explicit expansion in ε to theoretically any order. An overview on the Eyring-Kramers formula can be found in the review article of Berglund [Ber11].

In this work, we provide a new proof of the Eyring-Kramers formula for the first eigenvalue of the operator L, i.e. its *spectral gap*. The advantage of this new approach is that it extends to the logarithmic Sobolev inequality (LSI), which was not investigated before. The LSI was introduced by [Gro75] and is stronger than the PI. Compared to PI, the LSI has some advantages, which we outline in Remark 1.4. Usually, the LSI is harder to deduce than the PI due to its nonlinear structure.

By deducing the Eyring-Kramers formula for the LSI, we encounter a surprising effect: In the generic situation of having two local minima with different energies, the Eyring-Kramers formula for the LSI differs from the Eyring-Kramers formula for the PI by a term of inverse order in ϵ. However, in the symmetric situation of having local minima with the same energy, the Eyring-Kramers formula for the LSI coincides with the corresponding formula for the PI (cf. Corollary 2.15).

We conclude the introduction with an overview of the article:
In Section 1.1, we introduce PI and LSI.
In Section 1.2, we discuss the setting and the assumptions on the Hamiltonian H.
In Section 2, we outline the new approach and state the main results of this work.
In Section 3 and Section 4, we proof the main ingredients of our new approach. Namely, in Section 3 we deduce a local PI and a local LSI with optimal scaling in ε, whereas in Section 4 we estimate a mean-difference by using a weighted transport distance.
In the appendices, we provide for the convenience of the reader some basic but non-standard facts that are used in our arguments.

1.1. **Poincaré and logarithmic Sobolev inequality.**

Definition 1.1 (PI(ϱ) and LSI(ϱ))**.** Let X be an Euclidean space. A Borel probability measure μ on X satisfies the *Poincaré inequality* with constant $\varrho > 0$, if for all test functions $f : X \to \mathbb{R}^+$

$$\operatorname{var}_\mu(f) := \int \left(f - \int f \, \mathrm{d}\mu\right)^2 \mathrm{d}\mu \le \frac{1}{\varrho} \int |\nabla f|^2 \, \mathrm{d}\mu. \qquad \mathrm{PI}(\varrho)$$

In a similar way, the probability measure μ satisfies the *logarithmic Sobolev inequality* with constant $\alpha > 0$, if for all test function $f : X \to \mathbb{R}^+$ holds

$$\operatorname{Ent}_\mu(f) := \int f \, \log \frac{f}{\int f \, \mathrm{d}\mu \, \log\left(\int f \, \mathrm{d}\mu\right)} \, \mathrm{d}\mu \le \frac{1}{\alpha} \int \frac{|\nabla f|^2}{2f} \, \mathrm{d}\mu =: I(f\mu|\mu), \quad \mathrm{LSI}(\alpha)$$

where $I(f\mu|\mu)$ is called *Fisher information*. The gradient ∇ is determined by the Euclidean structure of X. Test functions are those functions for which the gradient exists and the right-hand-side in PI(ϱ) and LSI(α) is finite.

Remark 1.2 (Relation between PI(ϱ) and LSI(α)). Rothaus [Rot78] observed that LSI(α) implies PI(α). Therefore, we set $f = 1 + \eta g$ for η small and find

$$\operatorname{Ent}_\mu(f^2) = 2\eta^2 \operatorname{var}_\mu(g) + O(\eta^3) \quad \text{as well as} \quad \int |\nabla f|^2 \, \mathrm{d}\mu = \eta^2 \int |\nabla g|^2 \, \mathrm{d}\mu.$$

Hence, if μ satisfies LSI(α) then μ satisfies PI(α), which always implies $\alpha \le \varrho$.

The connection of the PI to the *spectral gap* of the operator L in (1.1) is the variational characterization of the latter one.

Lemma 1.3. *The spectral gap ϱ_{SG} of the operator L has the* variational characterization

$$\varrho_{\mathrm{SG}} := \inf_f \frac{\mathcal{E}(f)}{\operatorname{var}_\mu(f)} = \varepsilon \inf_f \frac{\int |\nabla f|^2 \, \mathrm{d}\mu}{\operatorname{var}_\mu(f)}, \qquad \mathrm{SG}(\varrho_{\mathrm{SG}})$$

where the infimum runs over all non-constant test functions $f : \mathbb{R}^n \to \mathbb{R}$.
From the defintion of $\mathrm{PI}(\varrho)$ and $\mathrm{SG}(\varrho_{\mathrm{SG}})$ follows that the operator L has a spectral gap of size $\varrho_{\mathrm{SG}} = \varrho\varepsilon$ if and only if the Gibbs measure μ satisfies $\mathrm{PI}(\varrho)$ with optimal constant ϱ.

Remark 1.4 (Difference between LSI and PI). Using a Gronwall type argument it is easy to see that the PI and the LSI imply exponential decay of the process ξ_t to equilibrium, i.e.

$$\begin{aligned} &\text{if } \mu \text{ satisfies } \mathrm{PI}(\varrho), \text{ then} && \operatorname{var}_\mu(f_t) \le \exp(-2\varrho\varepsilon t) \operatorname{var}_\mu(f_0), \quad \text{and} \\ &\text{if } \mu \text{ satisfies } \mathrm{LSI}(\varrho), \text{ then} && \operatorname{Ent}_\mu(f_t) \le \exp(-2\varrho\varepsilon t) \operatorname{Ent}_\mu(f_0). \end{aligned}$$

In order to give a meaning to the last two inequalites, one still has to estimate the intial variance $\operatorname{var}_\mu(f_0)$ or the initial entropy $\operatorname{Ent}_\mu(f_0)$. Using a simple toy example one can see an advantage of the LSI over the PI: We choose the Gibbs measure μ normally distributes as $\mathcal{N}(0, \mathrm{Id})$. The initially state $f_0\mu$ is also choosen to be normally distributed according to $\mathcal{N}(0, \sigma \,\mathrm{Id})$ for some $\sigma > 0$. Then direct calculation reveals

$$\operatorname{var}_\mu(f_0) = \begin{cases} \left(\frac{1}{\sigma(2-\sigma)}\right)^{\frac{n}{2}} - 1 & \text{if } \sigma < 2, \\ \infty & \text{if } \sigma \ge 2 \end{cases} \qquad \text{whereas} \qquad \operatorname{Ent}_\mu(f_0) = \frac{n}{2}(\sigma - 1 - \log\sigma).$$

1.2. **Setting and assumptions.** This article uses almost the same setting as found in [BEGK04, BGK05]. Before starting the precise assumptions on the Hamiltonian H, we introduce the notion of a *Morse function*.

Definition 1.5 (Morse function)**.** A smooth function $H : \mathbb{R}^n \to \mathbb{R}$ is a *Morse function*, if the Hessian $\nabla^2 H$ of H is non-degenerated on the set of critical points. More precisely, for some $1 \le C_H < \infty$ holds

$$\forall x \in \mathcal{S} := \{x \in \mathbb{R}^n : \nabla H = 0\} : \quad \frac{1}{C_H} \le \|\nabla^2 H(x)\| \le C_H. \tag{1.4}$$

We make the following assumption on the Hamiltonian H which despite the non-degeneracy matter if the domain of H is unbounded. Hereby, we have to assume stronger properties for H if we want to proof the LSI.

Assumption 1.6 (PI)**.** *We assume that $H \in C^3(\mathbb{R}^n, \mathbb{R})$ is a Morse function. Further, for some constants $C_H > 0$ and $K_H \ge 0$ holds*

$$\liminf_{|x|\to\infty} |\nabla H| \ge C_H. \tag{$\mathbf{A1}_{\mathtt{PI}}$}$$

$$\liminf_{|x|\to\infty} |\nabla H|^2 - \Delta H \ge -K_H. \tag{$\mathbf{A2}_{\mathtt{PI}}$}$$

Assumption 1.7 (LSI)**.** *We assume that $H \in C^3(\mathbb{R}^n, \mathbb{R})$ is a Morse function. Further, for some constants $C_H > 0$ and $K_H \ge 0$ holds*

$$\liminf_{|x|\to\infty} \frac{|\nabla H(x)|^2 - \Delta H(x)}{|x|^2} \ge C_H. \tag{$\mathbf{A1}_{\mathtt{LSI}}$}$$

$$\inf_x \nabla^2 H(x) \ge -K_H. \tag{$\mathbf{A2}_{\mathtt{LSI}}$}$$

Remark 1.8 (Discussion of assumptions). The Assumption 1.6 yields the following consequences for the Hamiltonian H:

- The condition ($\mathbf{A1}_{\mathtt{PI}}$) ensures that e^{-H} is integrable and can be normalized to a probability measure on $\mathbb{R}^n$. Hence, the Gibbs measure μ given by (1.3) is well defined.
- A combination of the condition ($\mathbf{A1}_{\mathtt{PI}}$) and ($\mathbf{A2}_{\mathtt{PI}}$) ensures that there exists a spectral gap for the operator L given by (1.1). Equivalently, this means by the variational characterization of the spectral gap of L (cf. Lemma 1.3) that the Gibbs measure μ given by (1.3) satisfies a PI for sufficiently small ε.
- The *Lyapunov-type* condition ($\mathbf{A2}_{\mathtt{PI}}$) allows to recover the Poincaré constant of the full Gibbs measure μ from the Poincaré constant of the Gibbs measure μ_R restricted to some ball B_R with radius $R > 0$ (cf. Section 3).
- The Morse Assumption (1.4) together with the growth condition ($\mathbf{A1}_{\mathtt{PI}}$) ensures that the set $\mathcal{S}$ of critical points is discrete and finite. In particular, it follows that the set of local minima $\mathcal{M} = \{m_1, \ldots, m_M\}$ is also finite, i.e. $M := \#\mathcal{M} < \infty$.

Similarly the Assumption 1.7 has the following consequences for the Hamiltonian H:

- The differences between the assumptions on H for the PI and the LSI is that ($\mathbf{A1}_{\mathtt{PI}}$) means at least linear growth at infinity for H, whereas a combination of condition ($\mathbf{A1}_{\mathtt{LSI}}$) and ($\mathbf{A2}_{\mathtt{LSI}}$) yields quadratic growth at infinity for H; that is
$$\liminf_{|x|\to\infty} \frac{|\nabla H(x)|}{|x|} \geq C_H. \qquad (\mathbf{A0}_{\mathtt{LSI}})$$
Moreover, quadratic growth at infinity is also a necessary condition to obtain LSI (cf. [Roy07, Theorem 3.1.21.]).
- In addition, ($\mathbf{A1}_{\mathtt{LSI}}$) and ($\mathbf{A2}_{\mathtt{LSI}}$) imply ($\mathbf{A1}_{\mathtt{PI}}$) and ($\mathbf{A2}_{\mathtt{PI}}$), which is only an indication that LSI(α) is stronger than PI(ϱ) in the sense of Remark 1.2. Whenever we refer to Assumption 1.6 the properties in question hold also under the Assumption 1.7.
- The condition ($\mathbf{A1}_{\mathtt{LSI}}$) also is a *Lyapunov type* condition. It only implies a defective WI-inequality (cf. Appendix G). To deduce LSI, one additionally has to assume the condition ($\mathbf{A2}_{\mathtt{LSI}}$) (cf. Section 3).

The final non-degeneracy assumption is not really needed for the proof of the Eyring-Kramers formula, but to keep the presentation feasible and clear. The saddle height $\widehat{H}(m_i, m_j)$ between two local minima m_i, m_j is defined by
$$\widehat{H}(m_i, m_j) := \inf\left\{ \max_{s\in[0,1]} H(\gamma(s)) : \gamma \in C([0,1], \mathbb{R}^n), \gamma(0) = m_i, \gamma(1) = m_j \right\}.$$

Assumption 1.9 (Non-degeneracy). *There exists $\delta > 0$ such that:*

(i) *The saddle height between two local minima m_i, m_j is attained at a unique critical point $s_{i,j} \in \mathcal{S}$, i.e. it holds $H(s_{i,j}) = \widehat{H}(m_i, m_j)$. The point $s_{i,j}$ is called* optimal *or* communicating saddle *between the minima m_i and m_j. It follows from Assumption 1.6 that $s_{i,j}$ is a saddle point of index one, i.e. $\left\{x \in \mathbb{R}^n : \left\langle \nabla^2 H(s_{i,j})x, x\right\rangle \leq 0\right\}$ is one-dimensional.*

(ii) *The set of local minima $\mathcal{M} = \{m_1, \ldots, m_M\}$ is ordered such that m_1 is the global minimum and for all $i \in \{3, \ldots, M\}$ yields*
$$H(s_{1,2}) - H(m_2) \geq H(s_{1,i}) - H(m_i) + \delta.$$

2. Outline of the new approach and main results

In this section we present the new approach to the Eyring-Kramers formula and formulate the main results of this article. Because the strategy is the same for the PI and LSI, we deal with both cases simultaneously. The approach uses ideas of the two-scale approach for LSI [GOVW09, OR07, Lel09] and the method by [CM10] to deduce PI and LSI estimates for mixtures of measures. However, the heuristics outlined in the introduction provide a good orientation for our proceeding. Remember that we have a splitting into two time-scales:

- the fast scale describes the *fast relaxation* to a local minima of H and
- the slow scale describes the *exponentially long transitions* between local equilibrium states.

Motivated by these two time scales, we specify in Section 2.1 a splitting of the measure μ into *local measures* living on the *basin of attraction* of the local minima of H. This splitting is lifted from the level of the measure to the level of the variance and entropy. In this way, we obtain *local variances and entropies*, which heuristically should correspond to the fast relaxation, and *coarse-grained variances and entropies*, which should correspond to the exponentially long transitions.

Now, we handle each contribution separately. The local variances and entropies are estimated by local PI (cf. Theorem 2.6) and local LSI respectively (cf. Theorem 2.7). The heuristics suggest that this contribution should be of higher order because this step only relies on the fast scale.

Before we can estimate the coarse-grained variances and entropies, we have to bring them in the form of mean-differences. This is automatically the case for the variances. However, for the coarse-grained entropies one has to apply a new weighted discrete LSI (cf. Section 2.2), which causes the difference between the PI and LSI in the Eyring-Kramers formula. The main contribution to the Eyring-Kramers formula (cf. Corollary 2.12 and Corollary 2.14) results from the estimation of the mean-difference, which is stated in Theorem 2.9.

At this point let us shortly summarize the the main results of this article:

- We provide good estimates for the local variances and entropies (cf. Section 2.3.1) and
- We provide sharp estimates for the mean-differences (cf. Section 2.3.2).
- From these main ingredients, the Eyring-Kramers formulas follow as simple corollaries (cf. Section 2.3.3).

We close this chapter with a discussion of the optimality of the Eyring-Kramers formula for the LSI in one dimension (cf. Section 2.4).

2.1. Partition of the state space.

The inspiration to use a partition of the state space comes from the work [JSTV04] for discrete Markov chains. Motivated by the fast convergence of the diffusion ξ_t given by (1.2) to metastable states, we decompose the Gibbs measure μ into a mixture of *local* Gibbs measures μ_i in the following way: To every local minimum $m_i \in \mathcal{M}$ for $i = 1, \dots, M$ we associate its *basin of attraction* Ω_i defined by

$$\Omega_i := \left\{ y \in \mathbb{R}^n : \lim_{t\to\infty} y_t = m_i, \dot{y}_t = -\nabla H(y_t), y_0 = y \right\}.$$

Up to sets of Lebesgue measure zero, the set $\mathcal{P}_M = \{\Omega_i\}_{i=1}^M$ is a partition of $\mathbb{R}^n$. We can associate the local Gibbs measure μ_i to each element of the partition Ω_i as the restriction of μ

$$\mu_i(\mathrm{d}x) := \frac{1}{Z_i Z_\mu}\, \mathbb{1}_{\Omega_i}(x)\, \exp\left(-\frac{H(x)}{\varepsilon}\right) \mathrm{d}x, \qquad \text{where} \quad Z_i = \mu(\Omega_i). \tag{2.1}$$

The marginal measure $\bar{\mu}$ is given by a sum of Dirac measures $\bar{\mu} = Z_1\delta_1 + \cdots + Z_M\delta_M$. We note that $\sum_i Z_i = 1$, since $\{\Omega_i\}_{i=1}^M$ is a partition of $\mathbb{R}^n$ and μ a probability measure. The starting point of the argument is a representation of the Gibbs measure μ as a *mixture* of the mutual singular measures μ_i, namely

$$\mu = Z_1\mu_1 + \cdots + Z_M\mu_M. \tag{2.2}$$

The decomposition of μ can be lifted to a decomposition of the variance $\operatorname{var}_\mu(f)$ and entropy $\operatorname{Ent}_\mu(f)$ by a straightforward substitution of the mixture representation (2.2) of μ. The equations below were also used in [CM10, Section 4.1].

Lemma 2.1 (Splitting of variance and entropy for partition)**.** *For all $f : \mathbb{R}^n \to \mathbb{R}$ holds*

$$\operatorname{var}_\mu(f) = \sum_{i=1}^M Z_i \operatorname{var}_{\mu_i}(f) + \sum_{i=1}^M \sum_{j>i} Z_i Z_j \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \tag{2.3}$$

$$\operatorname{Ent}_\mu(f) = \sum_{i=1}^M Z_i \operatorname{Ent}_{\mu_i}(f) + \operatorname{Ent}_{\bar{\mu}}\left(\bar{f}\right). \tag{2.4}$$

We call the terms $\operatorname{var}_{\mu_i}(f)$ *and* $\operatorname{Ent}_{\mu_i}(f)$ local variance *and* local entropy*. The term* $\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2$ *is called* mean-difference *and* $\operatorname{Ent}_{\bar{\mu}}(\bar{f})$*, denoted by* coarse-grained entropy*, is given by*

$$\operatorname{Ent}_{\bar{\mu}}\left(\bar{f}\right) = \sum_{i=1}^M Z_i \bar{f}_i \ \log \frac{\bar{f}_i}{\sum_{j=1}^M Z_j \bar{f}_j}, \tag{2.5}$$

where $\bar{f}(i) = \bar{f}_i = \mathbb{E}_{\mu_i}(f)$*.*

2.2. **Discrete logarithmic Sobolev type inequalities.** From (2.4) we have to estimate the *coarse-grained entropy* $\operatorname{Ent}_{\bar{\mu}}(\bar{f})$. From the heuristic explanation, we expect that the main contribution comes from this term, which we want to treat further. If H has only two minima, we can use the following discrete LSI for a Bernoulli random variable, which was found by Higuchi and Yoshida [HY95] and Diaconis and Saloff-Coste [DSC96, Theorem A.2.] at the same time.

Lemma 2.2 (Optimal logarithmic Sobolev inequality for Bernoulli measures)**.** *A Bernoulli measure μ_p on $X = \{0, 1\}$, i.e. a mixture of two Dirac measures $\mu_p = p\delta_0 + q\delta_1$ with $p + q = 1$ satisfies the discrete logarithmic Sobolev inequality*

$$\operatorname{Ent}_{\mu_p}(f^2) \le \frac{pq}{\Lambda(p,q)}(f(0) - f(1))^2 \tag{2.6}$$

with optimal constant given by the logarithmic mean *(cf. Appendix A)*

$$\Lambda(p,q) := \frac{p-q}{\log p - \log q}, \quad \textit{for } p \neq q \qquad \textit{and} \quad \Lambda(p,p) := \lim_{q\to p} \Lambda(p,q) = p.$$

We want to handle the general case with more than two minima. Therefore, we need to answer the question of how to generalize Lemma 2.2 to discrete measures with a state space with more than two elements. An application of the modified LSI for finite Markov chains of Diaconis and Saloff-Coste [DSC96, Theorem A.1.] would not lead to an optimal results (cf. [Sch12, Section 2.3.]). Even for a generic Markov chain on the 3-point space, the optimal logarithmic Sobolev constant is unknown. Therefore, we will use the following direct generalization of Lemma 2.2, which is a type of weighted LSI.

Lemma 2.3 (Weighted logarithmic Sobolev inequality)**.** *For $m \in \mathbb{N}$ let $\mu_m = \sum_{i=1}^m Z_i \delta_i$ be a discrete probability measure and assume that $\min_i Z_i > 0$. Then for a function $f : \{1, \dots, m\} \to \mathbb{R}_0^+$ holds the weighted logarithmic Sobolev inequality*

$$\operatorname{Ent}_{\mu_m}(f^2) \le \sum_{i=1}^m \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} (f_i - f_j)^2 . \tag{2.7}$$

Proof. We conclude by induction and find that for $m = 2$ the estimate (2.7) just becomes (2.6), which shows the base case. For the inductive step, let us assume that (2.7) holds for $m \ge 2$. Then, the entropy $\operatorname{Ent}_{\mu_{m+1}}(f^2)$ can be rewritten as follows

$$\operatorname{Ent}_{\mu_{m+1}}(f^2) = (1 - Z_{m+1}) \operatorname{Ent}_{\tilde{\mu}_m}(f^2) + \operatorname{Ent}_\nu(\tilde{f}),$$

where the probability measure $\tilde{\mu}_m$ lives on $\{1, \dots, m\}$ and is given by

$$\tilde{\mu}_m := \sum_{i=1}^m \frac{Z_i}{1 - Z_{m+1}} \delta_i .$$

Further, ν is the Bernoulli measure given by $\nu := (1 - Z_{m+1})\delta_0 + Z_{m+1}\delta_1$ and the function $\tilde{f} : \{0, 1\} \to \mathbb{R}$ is given with values

$$\tilde{f}_0 := \sum_{i=1}^m \frac{Z_i f_i^2}{1 - Z_{m+1}} \qquad \text{and} \qquad \tilde{f}_1 := f_{m+1}^2 .$$

Now, we apply the inductive hypothesis to $\operatorname{Ent}_{\tilde{\mu}_m}(f^2)$ and arrive at

$$\begin{aligned}(1 - Z_{m+1}) \operatorname{Ent}_{\tilde{\mu}_m}(f^2) &\le (1 - Z_{m+1}) \sum_{i=1}^m \sum_{j>i} \frac{Z_i Z_j}{(1 - Z_{m+1})^2} \frac{1 - Z_{m+1}}{\Lambda(Z_i, Z_j)} (f_i - f_j)^2 \\ &= \sum_{i=1}^m \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} (f_i - f_j)^2 ,\end{aligned}$$

where we used $\Lambda(\cdot, \cdot)$ being homogeneous of degree one in both arguments (cf. Appendix A), i.e. $\Lambda(\lambda a, \lambda b) = \lambda \Lambda(a, b)$ for $\lambda, a, b > 0$. We can apply the inductive base to the second entropy $\operatorname{Ent}_\nu(\tilde{f})$, which in this case is the discrete LSI for the two-point case (2.6)

$$\operatorname{Ent}_\nu(\tilde{f}) \le \frac{Z_{m+1}(1 - Z_{m+1})}{\Lambda(Z_{m+1}, 1 - Z_{m+1})} \left(\sqrt{\tilde{f}_0} - \sqrt{\tilde{f}_1} \right)^2 . \tag{2.8}$$

The last step is to apply the Jensen inequality to recover the square differences $(f_i - f_{m+1})^2$ from

$$\begin{aligned}\left(\sqrt{\tilde{f}_0} - \sqrt{\tilde{f}_1} \right)^2 &= \sum_{i=1}^m \frac{Z_i f_i^2}{1 - Z_{m+1}} - 2 \underbrace{\sqrt{\sum_{i=1}^m \frac{Z_i f_i^2}{1 - Z_{m+1}}}}_{\ge \sum_{i=1}^m \frac{Z_i f_i}{1 - Z_{m+1}}} f_{m+1} + f_{m+1}^2 . \\ &\le \sum_{i=1}^m \frac{Z_i}{1 - Z_{m+1}} (f_i - f_{m+1})^2 .\end{aligned}$$

We obtain in combination with (2.8) the following estimate

$$\operatorname{Ent}_\nu(\tilde{f}) \le \frac{Z_{m+1}}{\Lambda(Z_{m+1}, 1 - Z_{m+1})} \sum_{i=1}^m Z_i (f_i - f_{m+1})^2 .$$

To conclude the assertion, we first note that $1 - Z_{m+1} = \sum_{j=1}^m Z_j \ge Z_j$ for $j = 1, \dots, m$. Further, $\Lambda(a, \cdot)$ is monotone increasing for $a > 0$, i.e. $\partial_b \Lambda(a, b) > 0$ (cf.

Appendix A). Both properties imply that $\Lambda(Z_{m+1}, 1 - Z_{m+1}) \geq \Lambda(Z_{m+1}, Z_j)$ for $j = 1, \ldots, m$, which finally shows (2.7). □

We are now able to estimate the coarse-grained entropy $\operatorname{Ent}_{\bar\mu}(\overline{f^2})$ occuring in the splitting of the entropy (2.4) with the help of Lemma 2.3. This generalizes the approach of [CM10, Section 4.1.] to the case of finite mixtures with more than two components.

Lemma 2.4 (Estimate of the coarse-grained entropy)**.** *The coarse-grained entropy in* (2.5) *can be estimated by*

$$\operatorname{Ent}_{\bar\mu}\left(\overline{f^2}\right) \leq \sum_{i=1}^{M} \left(\sum_{j\neq i} \frac{Z_i Z_j \operatorname{var}_{\mu_i}(f)}{\Lambda(Z_i, Z_j)} + \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \right) \tag{2.9}$$

where $\overline{f^2} : \{1, \ldots, M\} \to \mathbb{R}$ *is given by* $\overline{f_i^2} := \mathbb{E}_{\mu_i}(f^2)$.

Proof. Since $\bar\mu = Z_1\delta_1 + \cdots + Z_M\delta_M$ is finite discrete probability measure, we can apply Lemma 2.3 to $\operatorname{Ent}_{\bar\mu}(\overline{f^2})$

$$\operatorname{Ent}_{\bar\mu}(\overline{f^2})) \leq \sum_{i=1}^{m} \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \left(\sqrt{\overline{f_i^2}} - \sqrt{\overline{f_j^2}} \right)^2 . \tag{2.10}$$

The square-root-mean-difference on the right-hand side of (2.10) can be estimated by using the Jensen inequality

$$\begin{aligned} \left(\sqrt{\mathbb{E}_{\mu_i}(f^2)} - \sqrt{\mathbb{E}_{\mu_j}(f^2)} \right)^2 &\leq \mathbb{E}_{\mu_i}(f^2) - 2 \underbrace{\sqrt{\mathbb{E}_{\mu_i}(f^2)\mathbb{E}_{\mu_j}(f^2)}}_{\geq \mathbb{E}_{\mu_i}(f)\mathbb{E}_{\mu_j}(f)} + \mathbb{E}_{\mu_j}(f^2) \\ &\leq \mathbb{E}_{\mu_i}(f^2) - 2\mathbb{E}_{\mu_i}(f)\mathbb{E}_{\mu_j}(f) + \mathbb{E}_{\mu_j}(f^2) \\ &= \operatorname{var}_{\mu_i}(f) + \operatorname{var}_{\mu_j}(f) + \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 . \end{aligned} \tag{2.11}$$

Now, we can combine (2.10) and (2.11) to arrive at the desired result (2.9). □

The combination of the splitting Lemma 2.1 with the above Lemma 2.4 results in an estimate of the entropy in terms of local variances, local entropies and mean-differences.

Corollary 2.5. *The entropy of* f *with respect to* μ *on a partition* $\{\Omega_i\}_{i=1}^M$ *with restricted probability measures* μ_i *on* Ω_i *can be estimated by*

$$\begin{aligned} \operatorname{Ent}_\mu(f^2) \leq& \sum_{i=1}^{M} Z_i \operatorname{Ent}_{\mu_i}(f^2) + \sum_{i=1}^{M} \sum_{j\neq i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \operatorname{var}_{\mu_i}(f) \\ &+ \sum_{i=1}^{M} \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 . \end{aligned} \tag{2.12}$$

2.3. **Main results.** The main results of this work are good estimates of the single terms on the right-hand side of (2.3) and (2.12). In detail, we need the *local PI* and the *local LSI* provided by Theorem 2.6 and Theorem 2.7. Furthermore, we need good control of the mean-differences, which will be the content of Theorem 2.9. Finally, the Eyring-Kramers formulas in Corollary 2.12 and Corollary 2.14 are simple consequences of these representations and estimates.

2.3.1. *Local Poincaré and logarithmic Sobolev inequalities.* Let us now turn to the estimation of the local variances and entropies. From the heuristic understanding of the process ξ_t given by (1.2), we expect a good behavior of the local Poincaré and logarithmic Sobolev constant for the local Gibbs measures μ_i as it resembles the fast convergence of ξ_t to a neighborhood of the next local minimum. Therefore, the local variances and entropies should not contribute to the leading order expansion of the total Poincaré and logarithmic Sobolev constant of μ. This idea is quantified in the next both theorems.

Theorem 2.6 (Local Poincaré inequality)**.** *Under Assumption 1.6, the local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf. (2.1)), satisfy* $\mathrm{PI}(\varrho_i)$ *with*

$$\varrho_i^{-1} = O(\varepsilon).$$

Theorem 2.7 (Local logarithmic Sobolev inequality)**.** *Under Assumption 1.7, the local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf. (2.1)), satisfy* $\mathrm{LSI}(\alpha_i)$ *with*

$$\alpha_i^{-1} = O(1).$$

Even if there are simple heuristics for the validity of Theorem 2.6 and Theorem 2.7, the proof is not straightforward. The reason is that our situation goes beyond the scope of the standard tools for PI and LSI:

- The Bakry-Émery criterion (cf. Theorem 3.1) cannot be applied because we do not have a convex Hamiltonian.
- A naive application of the Holley-Stroock perturbation principle (cf. Theorem 3.2) would yield an exponentially bad dependence on the parameter ε.
- One cannot apply a simple *Lyapunov argument*, because one cannot impose a drift condition due to the saddles on a basin of attraction.

Therefore, we are forced to apply a more subtle Lyapunov argument, which is based on a perturbation argument and on an explicit construction of a *Lyapunov function*. We outline the argument for Theorem 2.6 and Theorem 2.7 in Section 3.

Remark 2.8 (Optimality of Theorem 2.6 and Theorem 2.7). The one-dimensional case indicates that the results of Theorem 2.6 and Theorem 2.7 for a Gibbs measure obtained from restricting its Hamiltonian H to the basin of attraction Ω of a local minimum is the best behavior in ε, which one can expect in general. The optimality in the one-dimensional case was investigated in [Sch12, Section 3.3.] by using the *Muckenhoupt functional* [Muc72] and *Bobkov-Götze functional* [BG99].

2.3.2. *Mean-difference estimate.* Let us now turn to the estimation of the mean-difference $\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2$. From the heuristics and the splitting of the variance (2.3) and entropy (2.12), we expect to see in the estimation of the mean-difference the exponential long waiting times of the jumps of the diffusion ξ_t given by (1.2) between the basins of attraction Ω_i. We have to find a good upper bound for the constant C in the inequality

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \leq C \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Therefore, we introduce in Section 4.1 a *weighted transport distance* between probability measures which yields a variational bound on the constant C. By an approximation argument (cf. Section 4.2), we give an explicit construction of a transport interpolation (cf. Section 4.3), which allows for asymptotically sharp estimates of the constant C.

Theorem 2.9 (Mean-difference estimate)**.** *Under Assumption 1.6, the mean-differences between the measures μ_i and μ_j for $i = 1, \ldots M-1$ and $j = i+1, \ldots, M$ satisfy*

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \, e^{\frac{H(s_{i,j})}{\varepsilon}} \int |\nabla f|^2 \,\mathrm{d}\mu, \quad (2.13)$$

where $\lambda^-(s_{i,j})$ denotes the negative eigenvalue of the Hessian $\nabla^2 H(s_{i,j})$ at the 1-saddle $s_{i,j}$ defined in Assumption ii. The symbol $\lesssim$ means $\leq$ up to a multiplicative error term of the form

$$1 + O(\sqrt{\varepsilon}\,|\log \varepsilon|^{\frac{3}{2}}).$$

The proof of Theorem 2.9 is carried out in full detail in Section 4.

Remark 2.10 (Multiple minimal saddles). In Assumption 1.9, we demand that there is exactly one minimal saddle between the local minima m_i and m_j. The technique we will develop in Section 4 is flexible enough to handle also cases, in which there exists more than one minimal saddle between local minima. The according adaptions and the resulting theorem can be found in [Sch12, Section 4.5.].

Remark 2.11 (Relation to capacity). The quantity on the right-hand side of (2.13) is the inverse of the capacity of a small neighborhood around m_i w.r.t. to a small neighborhood around m_j. The capacity is the crucial ingredient of the works [BEGK04] and [BGK05].

2.3.3. *Eyring-Kramers formulas.* Now, let us turn to the Eyring-Kramers formulas. Starting from the splitting obtained in Lemma 2.1 and Corollary 2.5, we will see how a combination of Theorem 2.6, Theorem 2.7 and Theorem 2.9 immediately leads to the multidimensional Eyring-Kramers formulas for the PI (cf. [BGK05, Theorem 1.2]) and LSI.

Corollary 2.12 (Eyring-Kramers formula for Poincaré inequality)**.** *Under Assumption 1.6, the measure μ satisfies* PI(ϱ) *with*

$$\frac{1}{\varrho} \lesssim Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 (H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|} \, e^{\frac{H(s_{1,2})}{\varepsilon}}, \quad (2.14)$$

where $\lambda^-(s_{1,2})$ denotes the negative eigenvalue of the Hessian $\nabla^2 H(s_{1,2})$ at the 1-saddle $s_{1,2}$. Further, the order is given such that $H(m_1) \leq H(m_i)$ and $H(s_{1,2}) - H(m_2)$ is the energy barrier of the system in the sense of Assumption 1.9.

Proof. We start from the decomposition of the variance into local variances and mean-differences given by Lemma 2.1. Then, an application of Theorem 2.6 and Theorem 2.9 yields the estimate

$$\begin{aligned} \operatorname{var}_\mu(f) &\leq \sum_i Z_i \operatorname{var}_{\mu_i}(f) + \sum_i \sum_{j<i} Z_i Z_j \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \\ &\lesssim O(\varepsilon) \left(+ \sum_i \sum_{j>i} \frac{Z_i Z_j Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \, e^{\frac{H(s_{i,j})}{\varepsilon}} \right) \int |\nabla f|^2 \,\mathrm{d}\mu. \end{aligned} \quad (2.15)$$

The final step is to observe, that by Assumption 1.9, the exponential dominating term in (2.15) is given for $i = 1$ and $j = 2$. □

In [BGK05, Theorem 1.2]) it is also shown that the upper bound of (2.14) is optimal by an approximation of the harmonic function. Therefore, in the following we can assume that (2.14) holds with $\approx$ instead of $\lesssim$.

Remark 2.13 (Higher exponential small eigenvalues). The main result of [BGK05, Theorem 1.2] does not only characterize the second eigenvalue of L (i.e. the spectral gap) but also the higher exponentially small eigenvalues. In principle, these characterizations can be also obtained in the present approach: The dominating exponential modes in (2.15), i.e. those obtained by setting $i = 1$, correspond to the inverse eigenvalues of L for $j = 2, \ldots, M$. By using the variational characterization of the eigenvalues of the operator L (cf. Lemma 1.3), the other exponentially small eigenvalues may be obtained by restricting the class of test functions f to the orthogonal complement of the eigenspaces of smaller eigenvalues.

Corollary 2.14 (Eyring-Kramers formula for logarithmic Sobolev inequalities)**.** *Under Assumption 1.7, the measure* μ *satisfies* $\mathrm{LSI}(\alpha)$ *with*

$$\frac{2}{\alpha} \lesssim \frac{Z_1 Z_2}{\Lambda(Z_1, Z_2)} \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2(H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|}\, e^{\frac{H(s_{1,2})}{\varepsilon}} \approx \frac{1}{\Lambda(Z_1, Z_2)}\, \frac{1}{\varrho}, \tag{2.16}$$

where $\lambda^-(s_{1,2})$ *denotes the negative eigenvalue of the Hessian* $\nabla^2 H(s_{1,2})$ *at the* 1*-saddle* $s_{1,2}$*. Further, where we assume that the ordering is given such that* $H(m_1) \leq H(m_i)$ *and* $H(s_{1,2}) - H(m_2)$ *is the energy barrier of the system in the sense of Assumption 1.9.*

Proof. The starting point is the estimate in Corollary 2.5 from which we are left to bound the local entropies and variances as well as the mean-differences. The according bounds are the statements of Theorem 2.6, Theorem 2.7 and Theorem 2.9 and lead to

$$\begin{aligned} \mathrm{Ent}_\mu(f^2) \leq O(1) \sum_{i=1}^{M} Z_i \int |\nabla f|^2 \,\mathrm{d}\mu_i + O(\varepsilon) \sum_{i=1}^{M} \sum_{j \neq i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \int |\nabla f|^2 \,\mathrm{d}\mu_i \\ + \sum_{i=1}^{M} \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)}\ \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|}\, e^{\frac{H(s_{i,j})}{\varepsilon}} \int |\nabla f|^2 \,\mathrm{d}\mu. \end{aligned} \tag{2.17}$$

The first term on the right-hand side of (2.17) is just $O(1) \int |\nabla f|^2 \,\mathrm{d}\mu$. For estimating the second term, we have to take care of the expression and use the one-homogeneity of $\Lambda(\cdot,\cdot)$ (cf. Appendix A)

$$\frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} = Z_i \frac{\log \frac{Z_i}{Z_j}}{\frac{Z_i}{Z_j} - 1} = Z_i P\left(\tfrac{Z_i}{Z_j}\right), \qquad \text{where} \quad P(x) := \frac{\log x}{x-1}. \tag{2.18}$$

The function $P(x)$ is decreasing and has a logarithmic singularity in 0. Therefore, we have to take care when $\frac{Z_i}{Z_j}$ becomes small. Let us therefore do an asymptotic evaluation of $\frac{Z_i}{Z_j}$, which can be deduced from

$$Z_i Z_\mu = \int_{\Omega_i} e^{-\frac{H}{\varepsilon}} \,\mathrm{d}x = \left(\frac{(2\pi\varepsilon)^{\frac{n}{2}}}{\sqrt{\det \nabla^2 H(m_i)}} + O(\varepsilon^{\frac{n+1}{2}}) \right) e^{-\frac{H(m_i)}{\varepsilon}}. \tag{2.19}$$

This formula immediately leads to the identity

$$\frac{Z_i}{Z_j} \approx \frac{\sqrt{\nabla^2 H(m_j)}}{\sqrt{\nabla^2 H(m_i)}}\, e^{-\frac{H(m_i)-H(m_j)}{\varepsilon}}, \tag{2.20}$$

which becomes exponentially small provided that $H(m_i) > H(m_j)$. In particular, using the expression (2.19) in (2.18) results in

$$\frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} = Z_i O(\varepsilon^{-1}). \tag{2.21}$$

Hence, also the second term in (2.17) can be estimated by $O(1)\int |\nabla f|^2 \,\mathrm{d}\mu$. This shows that the third term dominates the first two on an exponential scale. This leads to the estimate

$$\operatorname{Ent}_\mu(f^2) \lesssim \sum_{i=1}^M \sum_{j>i} \frac{Z_i Z_j}{\Lambda(Z_i, Z_j)} \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} e^{\frac{H(s_{i,j})}{\varepsilon}} \int |\nabla f|^2 \,\mathrm{d}\mu.$$

From Assumption 1.9 together with (2.19) and (2.21) follows that the exponential leading order term is given for $i = 1$ and $j = 2$. □

Corollary 2.15 (Comparison of ϱ and α in special cases)**.** *Let us state two specific cases of* (2.14) *and* (2.16)*. Therefore, let* $\{\kappa_i^2\}_{i=1}^M$ *denote the Hessian of* H *evaluated at the minima* m_i*, i.e.*

$$\kappa_i^2 := \det \nabla^2 H(m_i).$$

Firstly, if $H(m_1) < H(m_2)$*, it holds*

$$\frac{1}{\varrho} \approx \frac{1}{\kappa_2} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2(H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|} e^{\frac{H(s_{1,2})-H(m_2)}{\varepsilon}}, \tag{2.22}$$

$$\frac{2}{\alpha} \lesssim \left(\frac{H(m_2)-H(m_1)}{\varepsilon} + \log\left(\frac{\kappa_1}{\kappa_2}\right)\right)\frac{1}{\varrho}. \tag{2.23}$$

A special case occurs when $H(m_1) = H(m_2)$ *and the constants take the form*

$$\frac{1}{\varrho} \approx \frac{1}{\kappa_1+\kappa_2} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2(H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|} e^{\frac{H(s_{1,2})-H(m_2)}{\varepsilon}}, \tag{2.24}$$

$$\frac{2}{\alpha} \lesssim \frac{1}{\Lambda(\kappa_1,\kappa_2)} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2(H(s_{1,2}))|}}{|\lambda^-(s_{1,2})|} e^{\frac{H(s_{1,2})-H(m_2)}{\varepsilon}}. \tag{2.25}$$

Proof. If $H(m_1) < H(m_2)$, then holds by (2.19) $Z_1 = 1 + O\left(e^{-\frac{H(m_2)-H(m_1)}{\varepsilon}}\right)$. Therefore, the factor $Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}$ evaluates with (2.19) to

$$Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \approx \frac{1}{\sqrt{\det \nabla^2 H(m_2)}} e^{-\frac{H(m_2)}{\varepsilon}},$$

which leads to (2.22). For the LSI, we additionally have to evaluate the factor $\frac{1}{\Lambda(Z_i,Z_j)}$ which can be done with the help of (2.20)

$$\begin{aligned}\frac{1}{\Lambda(Z_i,Z_j)} &= \log\frac{Z_i}{Z_j}\left(1 + O\left(e^{-\frac{H(m_2)-H(m_1)}{\varepsilon}}\right)\right)\\ &\overset{(2.20)}{\approx} \log\left(\frac{\sqrt{\nabla^2 H(m_j)}}{\sqrt{\nabla^2 H(m_i)}}\, e^{-\frac{H(m_i)-H(m_j)}{\varepsilon}}\right).\end{aligned}$$

That is already the estimate (2.23).
Let us turn now to the case $H(m_1) = H(m_2)$. Then, we can evaluate Z_μ like in (2.19) and obtain by assuming $H(m_1) = H(m_2) = 0$

$$Z_\mu = \left(\frac{(2\pi\varepsilon)^{\frac{n}{2}}}{\kappa_1} + \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{\kappa_2} + O(\varepsilon^{\frac{n+1}{2}})\right).$$

Therewith, the factor $Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}$ results with (2.19) in

$$Z_1 Z_2 \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} = \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_\mu} \frac{Z_1 Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{Z_2 Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} = \frac{1}{\frac{1}{\kappa_1}+\frac{1}{\kappa_2}} \frac{1}{\kappa_1} \frac{1}{\kappa_2} = \frac{1}{\kappa_1+\kappa_2},$$

which precisely leads to the expression (2.24). By using the homogeneity of $\Lambda(\cdot,\cdot)$ (cf. Appendix A) and (2.19) follows for the LSI

$$\frac{Z_1 Z_2}{\Lambda(Z_1,Z_2)} \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} = \frac{1}{\Lambda\left(\frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_2 Z_\mu}, \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_1 Z_\mu}\right)} \approx \frac{1}{\Lambda(\kappa_2,\kappa_1)}.$$

Finally, the result (2.25) is a consequence of the symmetry of $\Lambda(\cdot,\cdot)$. □

Remark 2.16 (Identification of α and ϱ). Remark 1.2 shows that always $\alpha \leq \varrho$. Let us introduce the shorthand notation $\kappa_i = \sqrt{\det \nabla^2 H(m_i)}$. We want to compare the case when $H(m_1) = H(m_2)$ where we observe by comparing (2.24) and (2.25)

$$1 \leq \frac{\varrho}{\alpha} \lesssim \frac{\frac{\kappa_1+\kappa_2}{2}}{\Lambda(\kappa_1,\kappa_2)}. \tag{2.26}$$

The quotient in (2.26) consists of the arithmetic and logarithmic mean. The lower bound of 1 can also be observed by applying the logarithmic-arithmetic mean inequality from Lemma A.1. Moreover equality only holds for $\kappa_1 = \kappa_2$. Hence, only in the symmetric case with $\kappa_1 = \kappa_2$ holds $\varrho \approx \alpha$.

Remark 2.17 (Relation to mixtures). If $H(m_1) < H(m_2)$, then (2.23) gives

$$\frac{\varrho}{\alpha} \lesssim \frac{1}{2}\log\left(\frac{\kappa_2}{\kappa_1} e^{\frac{H(m_2)-H(m_1)}{\varepsilon}}\right) \approx \frac{1}{2}\left|\log Z_2\right|, \quad \text{where} \quad Z_2 = \mu(\Omega_2). \tag{2.27}$$

which shows an inverse scaling in ε. Different scaling behavior between Poincaré and logarithmic Sobolev constants was also observed by Chafaï and Malrieu [CM10] in a different context. They consider mixtures of probability measures ν_0 and ν_1 satisfying PI(ϱ_i) and LSI(α_i), i.e. for $p \in [0,1]$ the measure ν_p given by

$$\nu_p = p\nu_0 + (1-p)\nu_1.$$

They deduce conditions under which also ν_p satisfies PI(ϱ_p) and LSI(α_p) and give bounds on the constants. They show in the one-dimensional case examples where the Poincaré constant stays bounded, whereas the logarithmic Sobolev constant blows up logarithmically, when the mixture parameter p goes to 0 or 1. The common feature of the examples they deal with is $\nu_1 \ll \nu_2$ or $\nu_2 \ll \nu_1$. This case can be generalized to the multidimensional case, where also a different scaling of the Poincaré and logarithmic Sobolev constants is observed. The details can be found in [Sch12, Chapter 6].
In the present case the Gibbs measure μ has also a mixture representation given in (2.2). In the two-component case it looks like

$$\mu = Z_1\mu_1 + Z_2\mu_2.$$

Let us emphasize, that $\mu_1 \perp \mu_2$. (2.27) shows also a logarithmic blow-up in the mixture parameter Z_2 for the ratio of the Poincaré and the logarithmic Sobolev constant.

2.4. **Optimality of the logarithmic Sobolev constant in one dimension.** In this section, we want to give a strong indication, that the result of Corollary 2.14 is optimal. Therefore, we will explicitly construct a function attaining equality in (2.16) for the one dimensional case. Therefore, let μ be a probability measure on $\mathbb{R}$ having as Hamiltonian H a generic double-well (cp. Figure 2). Namely, H has two minima m_1 and m_2 with $H(m_1) \leq H(m_2)$ and a saddle s in-between. Then, Theorem 2.14 shows

$$\inf_{g:\int g^2 \mathrm{d}\mu = 1} \frac{\int (g')^2 \mathrm{d}\mu}{\int g^2 \log g^2 \mathrm{d}\mu} \gtrsim \frac{\Lambda(Z_1,Z_2)}{Z_1 Z_2} \frac{\sqrt{2\pi\varepsilon}}{Z_\mu} \frac{\sqrt{|H''(s)|}}{2\pi\varepsilon}\, e^{\frac{H(s)-H(m_2)}{\varepsilon}}. \tag{2.28}$$

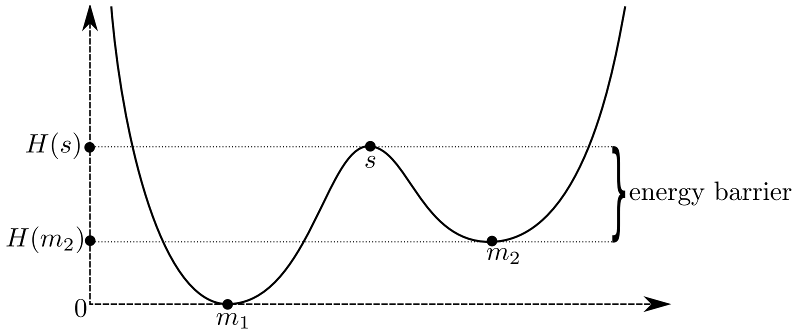

FIGURE 2. Double-well potential H on $\mathbb{R}$ (labeled).

We have to construct a function g attaining the lower bound given in (2.28) satisfying $H(m_1) \leq H(m_2)$. We make the following ansatz for the function g and firstly define it on some small δ-neighborhoods around the minima m_1, m_2 and the saddle s:

$$g(x) := \begin{cases} g(m_1) & , x \in B_\delta(m_1) \\ g(m_1) + \frac{g(m_2)-g(m_1)}{\sqrt{2\pi\varepsilon\sigma}} \int_{m_1}^{x} e^{-\frac{(y-s)^2}{2\sigma\varepsilon}} \mathrm{d}y & , x \in B_\delta(s) \\ g(m_2) & , x \in B_\delta(m_2). \end{cases}$$

The ansatz contains the parameters $g(m_1)$, $g(m_2)$ and σ. Furthermore, we assume that in-between the δ-neighborhoods g is extended to a smooth function in a monotone fashion.

The measure μ is given by

$$\mu(\mathrm{d}x) = \frac{1}{Z_\mu} e^{-\frac{H(x)}{\varepsilon}}, \qquad \text{where} \qquad Z_\mu = \int e^{-\frac{H(x)}{\varepsilon}} \mathrm{d}x.$$

We fix Z_μ by assuming that $H(m_1) = 0$. We can represent μ as the mixture

$$\mu = Z_1\mu_1 + Z_2\mu_2, \qquad \text{where} \qquad \mu_1 = \mu \llcorner \Omega_1 \quad \text{and} \quad \mu_2 = \mu \llcorner \Omega_2,$$

hereby, $\Omega_1 = (-\infty, s)$ and $\Omega_2 = (s, \infty)$ and $Z_i = \mu(\Omega_i)$ for $i = 1, 2$, which implies $Z_1 + Z_2 = 1$. Then, for the ansatz g, we find via an asymptotic evaluation of $\int g^2 \mathrm{d}\mu$

$$\int g^2 \mathrm{d}\mu \approx Z_1 g^2(m_1) + Z_2 g^2(m_2) \stackrel{!}{=} 1.$$

This motivates the choice

$$g^2(m_1) = \frac{\tau}{Z_1} \qquad \text{and} \qquad g^2(m_2) = \frac{1-\tau}{Z_2} = \frac{1-\tau}{1-Z_1}, \quad \text{for some } \tau \in [0, 1]$$

Let us now calculate the denominator of (2.28)

$$\int g^2 \log g^2 \mathrm{d}\mu = \tau \log \frac{\tau}{Z_1} + (1-\tau) \log \frac{1-\tau}{Z_2}. \tag{2.29}$$

The final step is to evaluate the Dirichlet energy $\int (g')^2 \mathrm{d}\mu$. Therefore, we do a Taylor expansion of H around s. Furthermore, since s is a saddle, it holds $H''(s) < 0$

$$\begin{aligned}
\int (g')^2 \mathrm{d}\mu &\approx \frac{(g(m_2) - g(m_1))^2}{Z_\mu \; 2\pi\varepsilon\sigma} \int_{B_\delta(s)} e^{-\frac{(x-s)^2}{\sigma\varepsilon} - \frac{H(x)}{\varepsilon}} \mathrm{d}x \\
&\approx \frac{(g(m_2) - g(m_1))^2}{Z_\mu \; 2\pi\varepsilon\sigma} \int_{B_\delta(s)} e^{-\frac{1}{\varepsilon}\left(\frac{(x-s)^2}{\sigma} + H(s) + H''(s)\frac{(x-s)^2}{2}\right)} \mathrm{d}x \\
&\approx \frac{(g(m_2) - g(m_1))^2}{Z_\mu \; 2\pi\varepsilon\sigma} \; e^{-\frac{H(s)}{\varepsilon}} \int_{B_\delta(s)} e^{-\frac{(x-s)^2}{2\varepsilon}\left(\frac{2}{\sigma} + H''(s)\right)} \mathrm{d}x \\
&\approx \left(\sqrt{\frac{\tau}{Z_1}} - \sqrt{\frac{1-\tau}{Z_2}}\right)^2 \frac{\sqrt{2\pi\varepsilon}}{Z_\mu} \; e^{-\frac{H(s)}{\varepsilon}} \; \frac{1}{2\pi\varepsilon} \frac{1}{\sigma\sqrt{\frac{2}{\sigma} + H''(s)}},
\end{aligned} \tag{2.30}$$

where we assume that σ is small enough such that $\frac{2}{\sigma} + H''(s) > 0$. The last step is to minimize the right-hand side of (2.30) in σ, which means to maximize the expression $2\sigma + \sigma^2 H''(s)$ in σ. Elementary calculus results in $\sigma = -\frac{1}{H''(s)} = \frac{1}{|H''(s)|} > 0$ and therefore

$$\int (g')^2 \mathrm{d}\mu \approx \left(\sqrt{\frac{\tau}{Z_1}} - \sqrt{\frac{1-\tau}{Z_2}}\right)^2 \frac{\sqrt{2\pi\varepsilon}}{Z_\mu} \frac{\sqrt{|H''(s)|}}{2\pi\varepsilon} \; e^{-\frac{H(s)}{\varepsilon}}. \tag{2.31}$$

Hence, we have constructed by combining (2.29) and (2.31) an upper bound for the optimization problem (2.28) given by a one-dimensional optimization in the still free parameter $\tau \in (0,1)$

$$\min_{\tau \in (0,1)} \left(\frac{\left(\sqrt{\frac{\tau}{Z_1}} - \sqrt{\frac{1-\tau}{Z_2}}\right)^2}{\tau \log \frac{\tau}{Z_1} + (1-\tau)\log\frac{1-\tau}{Z_2}} \right) \frac{\sqrt{2\pi\varepsilon}}{Z_\mu} \frac{\sqrt{|H''(s)|}}{2\pi\varepsilon} \; e^{-\frac{H(s)}{\varepsilon}}.$$

The minimum in τ is attained according to Lemma A.3 for $\tau = Z_2$

$$\min_{\tau \in (0,1)} \frac{\left(\sqrt{\frac{Z_2}{Z_1}} - \sqrt{\frac{Z_1}{Z_2}}\right)^2}{Z_2 \log \frac{Z_2}{Z_1} + Z_1 \log \frac{Z_1}{Z_2}} = \frac{\Lambda(Z_1, Z_2)}{Z_1 Z_2}.$$

3. Local Poincaré and logarithmic Sobolev inequalities

In this section, we want to proof the local PI of Theorem 2.6 and the local LSI of Theorem 2.7. Therefore, we consider only one of the basins of attraction Ω_i for fixed i and we can omit the index i. We will write Ω and μ instead of Ω_i and μ_i respectively. Further, we assume w.l.o.g. that $0 \in \Omega$ is the unique minimum of H in Ω.

Let us begin with stating the two classical conditions for PI and LSI. The first one is the *Bakry-Émery criterion* which states the convexity of the Hamiltonian exhibits good mixing for the associate Gibbs measure.

Theorem 3.1 (Bakry-Émery criterion [BÉ85, Proposition 3, Corollaire 2]). *Let H be a Hamiltonian with Gibbs measure $\mu(\mathrm{d}x) = Z_\mu^{-1} e^{-\varepsilon^{-1} H(x)} \mathrm{d}x$ and assume that $\nabla^2 H(x) \geq \lambda > 0$ for all $x \in \mathbb{R}^n$. Then μ satisfies $\mathrm{PI}(\varrho)$ and $\mathrm{LSI}(\alpha)$ with*

$$\varrho \geq \frac{\lambda}{\varepsilon} \qquad \text{and} \qquad \alpha \geq \frac{\lambda}{\varepsilon}.$$

The second condition is the *Holley-Stroock perturbation principle*, which allows to show PI and LSI for a very large class of measures. However, in general the constant obtained from this principle will be not optimal in terms of scaling with the temperature ε.

Theorem 3.2 (Holley-Stroock perturbation principle [HS87, p. 1184])**.** *Let H be a Hamiltonian with Gibbs measure $\mu(\mathrm{d}x) = Z_\mu^{-1} e^{-\varepsilon^{-1} H(x)} \mathrm{d}x$. Further, let ψ be a bounded function and denote by $\tilde{\mu}$ the Gibbs measure with Hamiltonian $H + \psi$, i.e.*

$$\mu(\mathrm{d}x) = \frac{1}{Z_\mu} e^{-\frac{H(x)}{\varepsilon}} \mathrm{d}x \qquad \textit{and} \qquad \tilde{\mu}(\mathrm{d}x) = \frac{1}{Z_{\tilde{\mu}}} e^{-\frac{H(x)+\psi(x)}{\varepsilon}} \mathrm{d}x.$$

Then, if μ satisfies $\mathrm{PI}(\varrho)$ *or* $\mathrm{LSI}(\alpha)$ *then also $\tilde{\mu}$ satisfy* $\mathrm{PI}(\tilde{\varrho})$ *or respectively* $\mathrm{LSI}(\tilde{\alpha})$*. Hereby the constants satisfy the relations*

$$\tilde{\varrho} \geq e^{-\frac{\operatorname{osc}\psi}{\varepsilon}} \varrho \qquad \textit{and} \qquad \tilde{\alpha} \geq e^{-\frac{\operatorname{osc}\psi}{\varepsilon}} \alpha, \tag{3.1}$$

where $\operatorname{osc}\psi := \sup\psi - \inf\psi$*.*

For the proofs relying on semigroup theory of Theorem 3.1 and Theorem 3.2 we refer to the exposition by Ledoux [Led01, Corollary 1.4, Corollary 1.6 and Lemma 1.2]. The only difference is, that we always explicitly express the temperature ε and consider H being ε-independent.

Let us summarize the reasons, why we cannot directly apply the above standard criteria for the PI and LSI to a Hamiltonian restricted to the basin of attraction of a local minimum.

- The criterion of Bakry-Émery [BÉ85] does not cover the present situation, because in general H is not convex on the basin of attraction Ω.
- The perturbation principle of Holley-Stroock [HS87] cannot be applied naively because it would yield an exponentially bad dependence of the Poincaré constant ϱ on ε.

Nevertheless, we will use both of them in the proof. The perturbation principle of Holley-Stroock will be used very carefully. In particular, we will compare the measure μ with a measure $\tilde{\mu}$, which is obtained from the construction of a perturbed Hamiltonian $\tilde{H}_\varepsilon$ such that $\|H - \tilde{H}_\varepsilon\|_\infty = O(\varepsilon)$ in Ω. The condition of slight perturbation allows to compare the Poincaré and logarithmic Sobolev constants of μ and $\tilde{\mu}$ upto an ε-independent factor. The second step consists of a Lyapunov argument developed by Bakry, Barthe, Cattiaux, Guillin, Wang and Wu (cf. [BBCG08], [BCG08], [CGW10] and [CGWW09]). The Lyapunov conditions shows similarities to the characterization of the spectral gap by Donsker and Varadhan [DV76]. We will state a *Lyapunov function* for $\tilde{\mu}$, which will allow to compare the scaling behavior for the Poincaré and logarithmic Sobolev constants with the truncated Gibbs measure $\hat{\mu}_a$ (cf. Definition 3.5 and Lemma 3.6).

The following definition is motivated by the Holley-Stroock perturbation principle and becomes eminent from the subsequent Lemma 3.4.

Definition 3.3 (ε-modification $\tilde{H}_\varepsilon$ of H)**.** We say that $\tilde{H}_\varepsilon$ is a ε-modification of H, if for all ε small enough $\tilde{H}_\varepsilon$ is of class $C^2(\Omega) \cap C^0(\overline{\Omega})$ and satisfies the condition: $\tilde{H}_\varepsilon$ is ε-close to H, i.e. there exists an ε-independent constant $C_{\tilde{H}} > 0$ such that

$$\forall x \in \Omega: \ |\tilde{H}_\varepsilon(x) - H(x)| \leq C_{\tilde{H}} \varepsilon. \tag{$\tilde{\mathbf{H}}_\varepsilon$}$$

The associated modified Gibbs measure $\tilde{\mu}$ obtained from the ε-modified Hamiltonian $\tilde{H}_\varepsilon$ of H is given by

$$\tilde{\mu}(\mathrm{d}x) = \frac{1}{Z_{\tilde{\mu}}} e^{-\frac{\tilde{H}_\varepsilon}{\varepsilon}} \mathrm{d}x.$$

Lemma 3.4 (Perturbation by an ε-modification)**.** *If the ε-modified Gibbs measure $\tilde{\mu}$ satisfy* $\mathrm{PI}(\tilde{\varrho})$ *and* $\mathrm{LSI}(\tilde{\alpha})$ *then the associated measure μ also satisfies* $\mathrm{PI}(\varrho)$ *and* $\mathrm{LSI}(\alpha)$*, where the constants fulfill the estimate*

$$\varrho \geq e^{-2C_{\tilde{H}}} \tilde{\varrho} \qquad \textit{and} \qquad \alpha \geq e^{-2C_{\tilde{H}}} \tilde{\alpha}. \tag{3.2}$$

Proof. We just can apply Theorem 3.2 with H replaced by $\tilde{H}$ and $\psi = H - \tilde{H}$. Finally, observe that by ($\tilde{\mathbf{H}}_\varepsilon$) holds

$$\operatorname{osc} \psi = \sup(H - \tilde{H}) - \inf(H - \tilde{H}) \leq 2|H - \tilde{H}| \leq 2C_{\tilde{H}}\varepsilon.$$

Therewith, the bound (3.1) becomes (3.2). □

Definition 3.5 (Truncated Gibbs measure)**.** To the Gibbs measure μ we associate by $\hat{\mu}_a$ the truncated measure obtained from μ by restricting it to a ball of radius $a\sqrt{\varepsilon}$ around 0 for some $a > 0$

$$\hat{\mu}_a(\mathrm{d}x) = \frac{1}{Z_{\hat{\mu}_a}}\, \mathbb{1}_{B_{a\sqrt{\varepsilon}}}(x)\, e^{-\frac{H(x)}{\varepsilon}} \mathrm{d}x.$$

Lemma 3.6 (PI and LSI for truncated Gibbs measure)**.** *The measure $\hat{\mu}_a$ satisfies* PI($\hat{\varrho}$) *and* LSI($\hat{\alpha}$) *for ε small enough, where*

$$\frac{1}{\hat{\varrho}} = O(\varepsilon) \qquad \text{and} \qquad \frac{1}{\hat{\alpha}} = O(\varepsilon). \tag{3.3}$$

Proof. In the local minimum 0 of Ω the Hessian of H is non-degenerated by Assumption 1.6 or 1.7. Therefore, for ε small enough, H is strictly convex in $B_{a\sqrt{\varepsilon}}$ and satisfies by the Bakry-Émery criterion (cf. Theorem 3.1) PI($\hat{\varrho}$) and LSI($\hat{\alpha}$) with $\hat{\varrho}$ and $\hat{\alpha}$ obeying the relation (3.3). □

3.1. **Lyapunov conditions ...**

3.1.1. *...for Poincaré inequality.* In this subsection, we will show that there exists an ε-modified Hamiltonian $\tilde{H}_\varepsilon$ which ensures that the Poincaré constant of $\tilde{\mu}$ is of the same order as the Poincaré constant of the truncated measure $\hat{\mu}_a$ from Definition 3.5. Therefore, we will state a Lyapunov function for the measure $\tilde{\mu}$. Firstly, let us introduce the notion of a Lyapunov condition.

Definition 3.7 (Lyapunov condition for Poincaré inequality)**.** Let $H : \Omega \to \mathbb{R}$ be a Hamiltonian and let

$$\mu(\mathrm{d}x) = \frac{\mathbb{1}_\Omega(x)}{Z_\mu}\, e^{-\frac{H(x)}{\varepsilon}} \mathrm{d}x$$

denote the associated Gibbs measure μ at temperature ε. Then, $W : \Omega \to [1, \infty)$ is a Lyapunov function for H provided there exist constants $R > 0$, $b > 0$ and $\lambda > 0$ such that

$$\frac{1}{\varepsilon}\, LW \leq -\lambda W + b\, \mathbb{1}_{\Omega_R}. \tag{3.4}$$

[BBCG08] is recommended for further information on the use of Lyapunov conditions for deducing PI. The main ingredient of this technique is the following statement:

Theorem 3.8 (Lyapunov condition for PI [BBCG08, Theorem 1.4.])**.** *Suppose that H fulfills the Lyapunov condition* (3.4) *and that the restricted measure μ_R given by*

$$\mu_R(\mathrm{d}x) = \mu(\mathrm{d}x)\llcorner_{\Omega_R} = \frac{\mathbb{1}_{\Omega_R}(x)}{\mu(\Omega_R)}\, \mu(\mathrm{d}x), \qquad \text{where} \qquad \Omega_R = \Omega \cap B_R$$

satisfies PI(ϱ_R)*. Then, μ also satisfies* PI(ϱ) *with constant*

$$\varrho \geq \frac{\lambda}{b + \varrho_R}\, \varrho_R.$$

We want to apply Theorem 3.8 to our situation. Hence, we do not only have to verify the Lyapunov condition (3.4) but also have to investigate the dependence of the constants R, b and λ on the parameter ε. Therefore, we will explicitly construct an ε−modification $\tilde{H}_\varepsilon$ of H in the sense of Definition 3.3. More precisely, we deduce the following statement:

Lemma 3.9 (Lyapunov function for PI)**.** *Without loss of generality we may assume that $0 \in \Omega$ is the unique minimum of H in Ω. Then, there exits an $\varepsilon-$modification $\tilde{H}_\varepsilon$ of H in the sense of Definition 3.3 such that for some constant $a > 0$ large enough holds with $\lambda_0 > 0$*

$$\frac{1}{2\varepsilon}\Delta\tilde{H}_\varepsilon(x) - \frac{1}{4\varepsilon^2}|\nabla\tilde{H}_\varepsilon(x)|^2 \leq -\frac{\lambda_0}{\varepsilon} \qquad \text{for all} \quad |x| \geq a\sqrt{\varepsilon}. \tag{3.5}$$

In particular, $\tilde{H}_\varepsilon$ satisfies the Lyapunov condition (3.4) with Lyapunov function

$$W(x) = \exp\left(\frac{1}{2\varepsilon}\tilde{H}_\varepsilon(x)\right) \quad \text{and constants} \quad R = a\sqrt{\varepsilon},\ b \leq \frac{b_0}{\varepsilon},\ \text{and } \lambda \geq \frac{\lambda_0}{\varepsilon}. \tag{3.6}$$

If the above lemma holds true the content of the local PI of Theorem 2.6 is just a simple consequence of a combination of Theorem 3.8 and Lemma 3.4. We will outline the proof in Section 3.2. Likewise, the statement of Lemma 3.9 directly follows from the following two observations.

Lemma 3.10. *Assume that the Hamiltonian H satisfies the Assumption (**A2**$_{\mathtt{PI}}$). Then, there is a constant $0 \leq C_H < \infty$ and $0 \leq \tilde{R} < \infty$ such that*

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \leq -\frac{C_H}{\varepsilon} \qquad \text{for all } |x| \geq \tilde{R}. \tag{3.7}$$

Moreover, let us assume that H is a Morse function in the sense of Definition 1.5. Additionally, let $\mathcal{S}$ denote the set of all critical points of H in Ω; that is

$$\mathcal{S} = \{y \in \Omega \mid \nabla H(y) = 0\}.$$

Then, there exists a constant $0 < c_H$ depending only on H such for $a > 0$ and ε small enough holds

$$|\nabla H(x)| \geq c_H a\sqrt{\varepsilon} \qquad \text{for all } x \notin \bigcup_{y\in\mathcal{S}} B_{a\sqrt{\varepsilon}}(y). \tag{3.8}$$

In particular, this implies that there is a constant $C_H > 0$ such that

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \leq -\frac{C_H}{\varepsilon} \qquad \text{for all } x \in B_{\tilde{R}}(0)\backslash \bigcup_{y\in\mathcal{S}} B_{a\sqrt{\varepsilon}}(y). \tag{3.9}$$

Proof. The proof basically consists only of elementary calculations based on the non-degeneracy assumption on H. For showing (3.7) we use the assumptions (**A1**$_{\mathtt{PI}}$) and (**A2**$_{\mathtt{PI}}$). Therefore, we define $\tilde{R}$ such that

$$\forall\, |x| \geq \tilde{R}: \qquad |\nabla H| \geq \frac{C_H}{2} \quad \text{and} \quad |\nabla H| - \Delta H(x) \geq -2K_H.$$

Therewith, it is easy to show, that for $|x| \geq \tilde{R}$ holds

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \leq \frac{1}{\varepsilon}\left(K_H - \frac{1}{2}\left(\frac{1}{2\varepsilon} - 1\right)\frac{C_H^2}{4}\right) \leq -\frac{C_H^2}{32\varepsilon},$$

for $\varepsilon \leq \frac{1}{4}\frac{C_H^2}{C_H^2+8K_H}$, which proves the statement (3.7). The condition (3.8) is first checked for a δ-neighborhood around the critical points $y \in \mathcal{S}$. There, by the Morse assumption on H (cp. Assumption 1.6 and Definition 1.5), we can do a Taylor expansion of H around the critical point y and find for $x \in B_\delta(y)\backslash B_{a\sqrt{\varepsilon}}(y)$

$$|\nabla H(x)| \geq \left|\lambda_{\min}(\nabla^2 H(y))\right| a\sqrt{\varepsilon} + O(\delta^2). \tag{3.10}$$

This shows, that (3.8) holds for $x \in B_\delta(y)\backslash B_{a\sqrt{\varepsilon}}(y)$. To conclude, we assume that (3.8) does not hold for some critical point y, i.e. for every $\varepsilon > 0$ and $c_H > 0$ and $a > 0$ we find $x \notin B_{a\sqrt{\varepsilon}}(y)$ such that $|\nabla H(x)| \leq c_H a\sqrt{\varepsilon}$, which by (3.10) contradicts the Morse assumption (1.4) for ε small enough. Finally, (3.9) is a conclusion of a combination of (3.7) and (3.8). □

The second observation needed for the verification of Lemma 3.9 is given by the following statement, which is the main ingredient for the proof of the local PI.

Lemma 3.11. *On a basin of attraction Ω there exists an $\varepsilon-$modification $\tilde{H}_\varepsilon$ of H in the sense of Definition 3.3 satisfying*

(i) The modification $\tilde{H}_\varepsilon$ equals H except for small neighborhoods around the critical points except the local minimum of H, i.e.

$$\tilde{H}_\varepsilon(x) = H(x), \qquad \text{for all } x \notin \bigcup_{y\in\mathcal{S}\setminus\{0\}} B_{a\sqrt{\varepsilon}}(y).$$

(ii) There are constants $0 < C_H$ and $a > 1$ such that for all small ε it holds

$$\frac{\Delta \tilde{H}_\varepsilon(x)}{2\varepsilon} - \frac{|\nabla \tilde{H}_\varepsilon(x)|^2}{4\varepsilon^2} \leq -\frac{C_H}{\varepsilon} \qquad \text{for all} \quad x \in \bigcup_{y\in\mathcal{S}\setminus\{0\}} B_{a\sqrt{\varepsilon}}(y). \tag{3.11}$$

Proof of Lemma 3.11. By the property (i) of Lemma 3.11, it is sufficient to construct the ε-modification $\tilde{H}_\varepsilon$ on a small neighborhood of any critical point y, which is not the global minimum of H in Ω. By translation, we may assume w.l.o.g. that $y = 0$.
Because the Hamiltonian H is a Morse function in the sense of Definition 1.5, we may assume that u_i, $i \in \{1, \ldots, n\}$ are orthonormal eigenvectors w.r.t. the Hessian $\nabla^2 H(0)$. The corresponding eigenvalues are denoted by λ_i, $i \in \{1, \ldots, n\}$. Additionally, we may assume w.l.o.g. that $\lambda_1, \ldots, \lambda_\ell < 0$ and $\lambda_{\ell+1}, \ldots, \lambda_n > 0$ for some $\ell \in \{1, \ldots, n\}$. If all $\lambda_i < 0$, we set $\tilde{H}_\varepsilon(x) = H(x)$ on $B_{a\sqrt{\varepsilon}}(0)$ and directly observe the desired statement (ii).
For the construction of $\tilde{H}_\varepsilon$, we need a smooth auxiliary function $\xi : [0,\infty) \to \mathbb{R}$ that satisfies

$$\xi'(z) = -1 \quad \text{for } |z| \leq \frac{a}{2}\sqrt{\varepsilon} \qquad \text{and} \qquad -1 \leq \xi'(z) \leq 0 \quad \text{for } z \in [0,\infty) \tag{3.12}$$

as well as for some $C_\xi > 0$ and any $|z| \leq a\sqrt{\varepsilon}$

$$|\xi''(z)| \leq \frac{C_\xi}{\sqrt{\varepsilon}} \qquad \text{and} \qquad \xi(z) = \xi'(z) = \xi''(z) = 0 \quad \text{for } |z| \geq a\sqrt{\varepsilon}. \tag{3.13}$$

Let us choose a constant $\delta > 0$ small enough such that

$$-\tilde{\delta} := (n-2\ell)\delta + \sum_{i=1}^{\ell} \lambda_i < 0 \qquad \text{and} \qquad \delta \leq \frac{1}{2}\min\{\lambda_i : i = \ell+1, \ldots, n\}. \tag{3.14}$$

Because $u_1, \ldots, u_n$ is an orthonormal basis of $\mathbb{R}^n$, we introduce a norm $|\cdot|_\delta$ on $\mathbb{R}^n$ by

$$|x|_\delta^2 := \sum_{i=1}^{\ell} \frac{1}{2}\delta\,|\langle u_i, x\rangle|^2 + \sum_{i=\ell+1}^{n} \frac{1}{2}(\lambda_i - \delta)\,|\langle u_i, x\rangle|^2$$

The norm $|\cdot|_\delta$ is equivalent to the standard euclidean norm $|\cdot|$ and satisfies the estimate

$$\frac{\delta}{2}\,|x|^2 \;\leq\; |x|_\delta^2 \;\leq\; \frac{\lambda_{\max}^+ - \delta}{2}\,|x|^2 \leq, \tag{3.15}$$

where $\lambda_{\max}^+ = \max\{\lambda_i : i = \ell+1, \ldots, n\}$. With the help of the function

$$H_b(x) := \xi\left(|x|_\delta^2\right), \tag{3.16}$$

we define the ε-modification $\tilde{H}_\varepsilon$ of H on a small neighborhood of the critical point 0 as

$$\tilde{H}_\varepsilon(x) = H(x) + H_b(x).$$

Note that by definition of H_b holds $\tilde{H}_\varepsilon(x) = H(x)$ for all $|x|_\delta \geq a\sqrt{\varepsilon}$. Therefore, the property (i) of Lemma 3.10 is satisfied by the equivalence of norms on finite dimensional vectorspaces (3.15).

For the verification of the statement (ii) of Lemma 3.10, it is sufficient to deduce the following two facts: The first one is the estimate

$$\Delta \tilde{H}(x) \leq -\frac{\tilde{\delta}}{2} \qquad \text{for all } |x|_\delta \leq \frac{a}{2}\sqrt{\varepsilon}. \tag{3.17}$$

The second one is that there is a constant $0 < C_H$ such that for a large $a > 1$ and small enough ε holds

$$\frac{\Delta \tilde{H}_\varepsilon(x)}{2} - \frac{|\nabla \tilde{H}_\varepsilon(x)|^2}{4\varepsilon} \leq -C_H \qquad \text{for all } \frac{a}{2}\sqrt{\varepsilon} \leq |x|_\delta \leq a\sqrt{\varepsilon}. \tag{3.18}$$

Let us have a look at (3.17). Because the function ξ has derivative -1 for $|x|_\delta \leq \frac{a}{2}\sqrt{\varepsilon}$, straightforward calculation yields

$$\nabla^2 \tilde{H}(x) = \nabla^2 H(x) - \sum_{i=1}^{\ell} \delta\; u_i \otimes u_i - \sum_{i=\ell+1}^{n} (\lambda_i - \delta)\; u_i \otimes u_i.$$

Taking the trace in the above identity results in

$$\Delta \tilde{H}(x) = \Delta H(x) - \sum_{i=\ell+1}^{n} \lambda_i + (n - 2\ell)\delta.$$

By the Taylor formula there is a constant $0 \leq C < \infty$ such that

$$|\Delta H(x) - \Delta H(0)| \leq C|x|.$$

Therefore, we get for $|x|_\delta \leq \frac{a}{2}\sqrt{\varepsilon}$

$$\begin{aligned}\Delta \tilde{H}(x) &= \Delta H(0) - \sum_{i=\ell+1}^{n} \lambda_i + (n-2\ell)\delta + \Delta H(x) - \Delta H(0) \\ &\leq \sum_{i=1}^{\ell} \lambda_i + (n-2\ell)\delta + C\frac{a}{2}\sqrt{\varepsilon} \overset{(3.14)}{\leq} -\tilde{\delta} + Ca\sqrt{\varepsilon} \leq -\frac{\tilde{\delta}}{2},\end{aligned}$$

for $\sqrt{\varepsilon} \leq \frac{2\tilde{\delta}}{Ca}$, which yields the desired statement (3.17).

Let us turn to the verification of (3.18). On the one hand, straightforward calculation reveals that there exists a constant $0 < C_\Delta < \infty$ such that

$$\Delta \tilde{H}(x) \leq C_\Delta \qquad \text{for all } \frac{a}{2}\sqrt{\varepsilon} < |x|_\delta < a\sqrt{\varepsilon}. \tag{3.19}$$

Indeed, we observe

$$\begin{aligned}\Delta \tilde{H}_\varepsilon(x) &= \Delta H(x) + \xi''(|x|_\delta^2)\left|\nabla |x|_\delta^2\right|^2 + \underbrace{\xi'(|x|_\delta^2)}_{\leq 0}\;\underbrace{\Delta |x|_\delta^2}_{\geq 0} \\ &\overset{(3.13)}{\leq} \Delta H(x) + \frac{C_\xi}{\sqrt{\varepsilon}}\;\left|\sum_{i=1}^{\ell} \delta\,\langle u_i, x\rangle\, u_i + \sum_{i=\ell+1}^{n} (\lambda_i - \delta)\,\langle u_i, x\rangle\, u_i\;\right|^2 \\ &\leq \Delta H(x) + \frac{C_\xi}{\sqrt{\varepsilon}}\;\lambda_{\max}^{+}\,|x|^2 \leq C_H + \frac{C_\xi}{\sqrt{\varepsilon}}\;a^2\varepsilon \leq C_H + C_\xi a^2\sqrt{\varepsilon} \leq C_\Delta,\end{aligned}$$

for some C_Δ and ε small enough, which yields (3.19).

On the other hand, we will deduce that there is a constant $0 < c_\nabla < \infty$ such that

$$|\nabla \tilde{H}_\varepsilon(x)|^2 \geq c_\nabla\; a^2\; \varepsilon \qquad \text{for all } \frac{a}{2}\sqrt{\varepsilon} < |x|_\delta < a\sqrt{\varepsilon}. \tag{3.20}$$

We want to note that the observations (3.19) and (3.20) already yield the desired statement (3.18). Indeed, we get for $a^2 \geq 4\frac{C_\Delta}{c_\nabla}$

$$\frac{\Delta \tilde{H}_\varepsilon(x)}{2\varepsilon} - \frac{|\nabla \tilde{H}_\varepsilon(x)|^2}{4\varepsilon^2} \leq \frac{C_\Delta}{2\varepsilon} - \frac{c_\nabla a^2}{4\varepsilon} \leq -\frac{C_\Delta}{2\varepsilon} \qquad \text{for all } \frac{a}{2}\sqrt{\varepsilon} < |x|_\delta < a\sqrt{\varepsilon},$$

which is the desired statement (3.18). Therefore, it is only left to deduce the estimate (3.20). By the definition of $\tilde{H}_\varepsilon$ from above, we can write

$$|\nabla \tilde{H}_\varepsilon(x)|^2 = |\nabla H(x)|^2 + |\nabla H_b(x)|^2 + 2\left\langle \nabla H(x), \nabla H_b(x)\right\rangle. \tag{3.21}$$

Let us have a closer look at each term on the right-hand side of the last identity and let us start with the first term. By Taylor's formula we obtain

$$|\nabla H(x) - \nabla^2 H(0)x| \leq C_\nabla |x|_\delta^2 \tag{3.22}$$

where $0 < C_\nabla < \infty$ denotes a generic constant. Therefore, we can estimate

$$|\nabla H(x)|^2 \geq |\nabla^2 H(0)x|^2 - C_\nabla a^4 \varepsilon^2 \qquad \text{for } |x|_\delta \leq a\sqrt{\varepsilon}. \tag{3.23}$$

By the definition of $\lambda_1, \ldots \lambda_n$, we also know

$$|\nabla^2 H(0)x|^2 = \sum_{i=1}^n \lambda_i^2 \left|\langle u_i, x\rangle\right|^2. \tag{3.24}$$

Let us have a closer look at the second term in (3.21), namely $|\nabla H_b(x)|^2$. From the definition (3.16) of $|\nabla H_b(x)|^2$ follows

$$\begin{aligned} |\nabla H_b(x)|^2 &= |\xi'(|x|_\delta^2)|^2 \left(\sum_{i=1}^\ell \delta^2 \left|\langle u_i, x\rangle\right|^2 + \sum_{i=\ell+1}^n (\lambda_i - \delta)^2 \left|\langle u_i, x\rangle\right|^2\right) \\ &\leq 2\lambda_{\max}^+ |x|_\delta^2. \end{aligned} \tag{3.25}$$

Now, we turn the the analysis of the last term, namely $2\left\langle \nabla H(x), \nabla H_b(x)\right\rangle$. By using the estimates (3.22) and (3.25), we get for $|x|_\delta \leq a\sqrt{\varepsilon}$.

$$\begin{aligned} \langle \nabla H(x), \nabla H_b(x)\rangle &= \left\langle \nabla^2 H(0)x, \nabla H_b(x)\right\rangle + \left\langle \nabla H(x) - \nabla^2 H(0)x, \nabla H_b(x)\right\rangle \\ &\overset{(3.22)}{\geq} \left\langle \nabla^2 H(0)x, \nabla H_b(x)\right\rangle - 2C_\nabla \lambda_{\max} |x|_\delta^3 \\ &\geq -\sum_{i=1}^\ell \lambda_i \delta \left|\xi'\left(|x|_\delta^2\right)\right| \left|\langle u_i, x\rangle\right|^2 \\ &\quad - \sum_{i=\ell+1}^n \lambda_i(\lambda_i - \delta) \left|\xi'\left(|x|_\delta^2\right)\right| \left|\langle u_i, x\rangle\right|^2 - 2C_\nabla \lambda_{\max} a^3 \varepsilon^{\frac{3}{2}}. \end{aligned} \tag{3.26}$$

Combining now the estimates and identities (3.21), (3.23), (3.24), (3.25) and (3.26), we arrive for $|x|_\delta \leq a\sqrt{\varepsilon}$ at

$$\begin{aligned} |\nabla \tilde{H}_\varepsilon(x)|^2 \geq & \sum_{i=1}^\ell \Big(\lambda_i - \delta \left|\xi'\left(|x|_\delta^2\right)\right|\Big)^2 \left|\langle u_i, x\rangle\right|^2 \\ &+ \sum_{i=\ell+1}^n \Big(\lambda_i - (\lambda_i - \delta) \left|\xi'\left(|x|_\delta^2\right)\right|\Big)^2 \left|\langle u_i, x\rangle\right|^2 - 4C_\nabla \lambda_{\max}^+ a^3 \varepsilon^{\frac{3}{2}}. \end{aligned}$$

By (3.12) holds $|\xi'(|x|_\delta^2)| \leq 1$, which applied to the last inequality yields

$$|\nabla \tilde{H}_\varepsilon(x)|^2 \geq \delta^2 \sum_{i=1}^n \left|\langle u_i, x\rangle\right|^2 - 4C_\nabla \lambda_{\max}^+ a^3 \varepsilon^{\frac{3}{2}}.$$

Because $u_1, \ldots u_n$ is an orthonormal basis of $\mathbb{R}^n$, the desired statement (3.20) follows for $\frac{a\sqrt{\varepsilon}}{2} \leq |x|_\delta \leq a\sqrt{\varepsilon}$ from

$$|\nabla \tilde{H}_\varepsilon(x)|^2 \geq = \delta^2 |x|^2 - 4C_\nabla \lambda^+_{\max} a^3 \varepsilon^{\frac{3}{2}} \overset{(3.15)}{\geq} \frac{2\delta^2}{\lambda^+_{\max}} |x|^2_\delta - 4C_\nabla \lambda^+_{\max} a^3 \varepsilon^{\frac{3}{2}}$$
$$\geq \frac{\delta^2}{2\lambda^+_{\max}}\, a^2 \varepsilon - 4C_\nabla \lambda^+_{\max} a^3 \varepsilon^{\frac{3}{2}} \geq c_\nabla a^2 \varepsilon$$

for some $c_\nabla < \frac{\delta^2}{2\lambda^+_{\max}}$ and ε small enough. □

We have collected all auxiliary results needed in the proof of Lemma 3.9.

Proof of Lemma 3.9. The condition (3.5) is a consequence of (3.7) from Lemma 3.10 and of (3.11) from Lemma 3.11. Now, we verify the Lyapunov condition (3.4) and calculate with $W = \exp\left(\frac{1}{2\varepsilon}\tilde{H}_\varepsilon\right)$

$$\frac{1}{\varepsilon}\frac{LW}{W} = \frac{1}{2\varepsilon}\Delta \tilde{H}_\varepsilon + \frac{1}{4\varepsilon^2}|\nabla \tilde{H}_\varepsilon|^2 - \frac{1}{2\varepsilon^2}|\nabla \tilde{H}_\varepsilon|^2 = \frac{1}{2\varepsilon}\Delta \tilde{H}_\varepsilon - \frac{1}{4\varepsilon^2}|\nabla \tilde{H}_\varepsilon|^2.$$

Choosing $|x| \geq a\sqrt{\varepsilon} := R$ one obtains from (3.5) the estimate $\lambda \geq \frac{\lambda_0}{\varepsilon}$. If $|x| \leq a\sqrt{\varepsilon}$, we note that $\tilde{H}_\varepsilon = H$ in $B_{a\sqrt{\varepsilon}}(0)$. Furthermore, H is quadratic around 0 and therefore is bounded by $H(x) \leq C_H a^2 \varepsilon$ for $|x| \leq a\sqrt{\varepsilon}$. Using, this in the definition of W, we arrive at the bound for $|x| \leq a\sqrt{\varepsilon}$

$$W(x) = e^{\frac{1}{2\varepsilon}H(x)} \leq e^{\frac{C_H a^2}{2}}.$$

This yields the desired estimate on the constant b, namely for $|x| \leq a\sqrt{\varepsilon}$

$$\frac{1}{\varepsilon} LW(x) \leq \frac{1}{2\varepsilon}\Delta H(x) W(x) \leq \frac{C_H e^{\frac{C_H a^2}{2}}}{\varepsilon} =: \frac{b_0}{\varepsilon},$$

which finishes the proof. □

3.1.2. *...for logarithmic Sobolev inequality.* The Lyapunov condition for proving LSI is stronger than the one for PI. Nevertheless, the construction of the ε-modified Hamiltonian $\tilde{H}_\varepsilon$ from the previous section carries over and we can mainly use the same Lyapunov function as for the PI. The Lyapunov condition for LSI goes back to the work of Cattiaux, Guillin, Wand and Wu [CGWW09]. We will apply the results in the form of the work [CGW10]. We will restate the proofs of the main results in [CGW10] for two reasons: Firstly, to adopt the notation to the low temperature regime and more importantly, to work out the explicit dependence between the constants of the Lyapunov condition, the logarithmic Sobolev constant and especially their ε-dependence.

Theorem 3.12 (Lyapunov condition for LSI [CGW10, Theorem 1.2])**.** *Suppose that there exists a C^2-function $W : \Omega \to [1, \infty)$ and constants $\lambda, b > 0$ such that for $L = \varepsilon\Delta - \nabla H \cdot \nabla$ holds*

$$\forall x \in \Omega : \frac{1}{\varepsilon}\frac{LW}{W} \leq -\lambda\, |x|^2 + b. \tag{3.27}$$

Further assume, that $\nabla^2 H \geq -K_H$ for some $K_H > 0$ and μ satisfies PI(ϱ), *then μ satisfies* LSI(α) *with*

$$\frac{1}{\alpha} \leq 2\sqrt{\frac{1}{\lambda}\left(\frac{1}{2} + \frac{b + \lambda\mu(|x|^2)}{\varrho}\right) + \frac{K_H}{2\varepsilon\lambda} + \frac{K_H(b + \lambda\mu(|x|^2)) + 2\varepsilon\lambda}{\varrho\varepsilon\lambda}}. \tag{3.28}$$

where $\mu(|x|^2)$ denotes the second moment of μ.

Lemma 3.13 ([CGW10, Lemma 3.4])**.** *Assume that U is a non-negative locally Lipschitz function such that for some lower bounded function ϕ*

$$\frac{Le^U}{e^U} = LU + \varepsilon \left|\nabla U\right|^2 \leq -\varepsilon\phi \tag{3.29}$$

in the distributional sense. Then for any g holds

$$\int \phi g^2 \mathrm{d}\mu \leq \int \left|\nabla g\right|^2 \mathrm{d}\mu.$$

Proof. We can assume w.l.o.g. that g is smooth with bounded support and ϕ is bounded. For the verification of the desired statement, we need the symmetry of L in $L^2(\mu)$:

$$\int (-Lf) g \mathrm{d}\mu = \int f(-L) g \mathrm{d}\mu = \varepsilon \int \nabla f \cdot \nabla g \mathrm{d}\mu, \tag{3.30}$$

and the simple estimate

$$2g\nabla U \cdot \nabla g \leq \left|\nabla U\right|^2 g^2 + \left|\nabla g\right|^2. \tag{3.31}$$

An application of the assumption (3.29) yields

$$\begin{aligned} \varepsilon \int \phi g^2 \mathrm{d}\mu &\overset{(3.29)}{\leq} \int \left(-LU - \varepsilon \left|\nabla U\right|^2\right) g^2 \mathrm{d}\mu \\ &\overset{(3.30)}{=} \varepsilon \int \left(2g\nabla U \cdot \nabla g - \left|\nabla U\right|^2 g^2\right) \mathrm{d}\mu \overset{(3.31)}{\leq} \varepsilon \int \left|\nabla g\right|^2 \mathrm{d}\mu, \end{aligned}$$

which is the desired estimate. □

The proof of Theorem 3.12 relies on an interplay of some other functional inequalities, which will not occur anywhere else. Therefore, in Appendix G a condensed summary may be found.

Proof of Theorem 3.12. The argument of [CGW10] is a combination of the Lyapunov condition (3.27) leading to a defective WI inequality and the use of the HWI inequality of Otto and Villani [OV00]. In the following, we will use the measure ν given by $\nu(\mathrm{d}x) = h(x)\mu(\mathrm{d}x)$, where we can assume w.l.o.g. that ν is a probability measure, i.e. $\int h \,\mathrm{d}\mu = 1$. The first step is to estimate the Wasserstein distance in terms of the total variation (cf. Theorem G.2 and [Vil09, Theorem 6.15])

$$W_2^2(\nu, \mu) \leq 2 \| \left|\cdot\right|^2 (\nu - \mu)\|_{TV}. \tag{3.32}$$

For every function g with $|g| \leq \phi(x) := \lambda \left|x\right|^2$, where λ is from the Lyapunov condition (3.27) we get

$$\begin{aligned} \int g \,\mathrm{d}(\nu - \mu) &\leq \int \phi \,\mathrm{d}\nu + \int \phi \,\mathrm{d}\mu \\ &= \int (\lambda \left|x\right|^2 - b) h(x) \,\mu(\mathrm{d}x) + \int b \,\mathrm{d}\nu + \mu(\phi), \end{aligned} \tag{3.33}$$

We can apply to $\int (\lambda \left|x\right|^2 - b) h \,\mathrm{d}\mu$ Lemma 3.13, where the assumption is exactly the Lyapunov condition (3.27) by choosing $U = \log W$ and arrive at

$$\int (\lambda \left|x\right|^2 - b) h \,\mathrm{d}\mu \leq \int \left|\nabla \sqrt{h}\right|^2 \mathrm{d}\mu = \int \frac{\left|\nabla h\right|^2}{4h} \mathrm{d}\mu = \frac{1}{2} I(\nu|\mu), \tag{3.34}$$

by the definition of the Fisher information. Taking the supremum over g in (3.33) and combining the estimate with (3.32) and (3.34) we arrive at the defective WI inequality

$$\frac{\lambda}{2} W_2^2(\nu, \mu) \leq \lambda \| \left|\cdot\right|^2 (\nu - \mu)\|_{TV} \leq \frac{1}{2} I(\nu|\mu) + b + \mu(\phi). \tag{3.35}$$

The next step is to use the HWI inequality (cf. Theorem G.6 and [OV00, Theorem 3]), which holds by the assumption $\nabla^2 H \geq -K_H$

$$\mathrm{Ent}_\mu(h) \leq W_2(\nu,\mu)\sqrt{2I(\nu|\mu)} + \frac{K_H}{2\varepsilon} W_2^2(\nu,\mu).$$

Substituting the defective WI inequality into the HWI inequality and using the Young inequality $ab \leq \frac{\tau}{2}a^2 + \frac{1}{2\tau}b^2$ for $\tau > 0$ results in

$$\begin{aligned}\mathrm{Ent}_\mu(h) &\leq \tau I(\nu|\mu) + \left(\frac{1}{2\tau} + \frac{K_H}{2\varepsilon}\right) W_2^2(\nu,\mu) \\ &\overset{(3.35)}{\leq} \left(\tau + \frac{1}{2\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right)\right) I(\nu|\mu) + \frac{1}{\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right)(b + \mu(\phi)).\end{aligned} \tag{3.36}$$

The last inequality is of the type $\mathrm{Ent}_\mu(h) \leq \frac{1}{\alpha_d} I(\nu|\mu) + B\int h \mathrm{d}\mu$ and is often called defective logarithmic Sobolev inequality dLSI(α_d, B). It is well-known, that a defective logarithmic Sobolev inequality can be tightened by PI(ϱ) to LSI(α) with constant (cf. Proposition G.9)

$$\frac{1}{\alpha} = \frac{1}{\alpha_d} + \frac{B+2}{\varrho}. \tag{3.37}$$

A combination of (3.36) and (3.37) reveals

$$\begin{aligned}\frac{1}{\alpha} &= \tau + \frac{1}{2\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right) + \frac{1}{\varrho}\left(\frac{1}{\lambda}\left(\frac{1}{\tau} + \frac{K_H}{\varepsilon}\right)(b + \mu(\phi)) + 2\right) \\ &= \tau + \frac{1}{\tau\lambda}\left(\frac{1}{2} + \frac{b+\mu(\phi)}{\varrho}\right) + \frac{K_H}{2\varepsilon\lambda} + \frac{K_H(b+\mu(\phi)) + 2\varepsilon\lambda}{\varrho\varepsilon\lambda} =: \tau + \frac{c_1}{\tau} + c_2.\end{aligned}$$

The last step is to optimize in τ, which leads to $\tau = \sqrt{c_1}$ and therefore $\frac{1}{\alpha} = 2\sqrt{c_1} + c_2$. The final result (3.28) follows by recalling the definition of $\phi(x) = \lambda\,|x|^2$. □

For proofing the Lyapunov condition (3.27) we can use the construction of an ε-modification done in Lemma 3.11.

Lemma 3.14 (Lyapunov function for LSI)**.** *There exists an ε-modification $\tilde H_\varepsilon$ of H satisfying the Lyapunov condition (3.27) with Lyapunov function*

$$W(x) = \exp\left(\frac{1}{2\varepsilon}\tilde H_\varepsilon(x)\right) \quad \text{and constants} \quad b = \frac{b_0}{\varepsilon}, \ \text{and } \lambda \geq \frac{\lambda_0}{\varepsilon}$$

for some $b_0, \lambda_0 > 0$ and Hessian $\nabla^2 \tilde H(x) \geq -K_{\tilde H}$ for some $K_{\tilde H} \geq 0$.

The proof consists of three steps, which correspond to three regions of interests. First, we will consider a neighborhood of ∞, i.e. we will fix some $\tilde R > 0$ and only consider $|x| \geq \tilde R$. This will be the analog estimate to formula (3.7) of Lemma 3.10. Then, we will look at an intermediate regime for $a\sqrt{\varepsilon} \leq |x| \leq \tilde R$, where we will have to take special care for the neighborhoods around critical points and use the construction of Lemma 3.11. The last regime is for $|x| \leq a\sqrt{\varepsilon}$, which will be the simplest case.

Therefore, besides the construction done in the proof of Lemma 3.11, we need an analogous formulation of Lemma 3.10 under the stronger assumption (**A1**$_{\mathrm{LSI}}$).

Lemma 3.15. *Assume that the Hamiltonian H satisfies Assumption (**A1**$_{\mathrm{LSI}}$). Then, there is a constant $0 \leq C_H < \infty$ and $0 \leq \tilde R < \infty$ such that for ε small enough*

$$\frac{\Delta H(x)}{2\varepsilon} - \frac{|\nabla H(x)|^2}{4\varepsilon^2} \leq -\frac{C_H}{\varepsilon}\,|x|^2 \qquad \text{for all } |x| \geq \tilde R. \tag{3.38}$$

We skip the proof of the Lemma 3.15, because it would work in the same way as for Lemma 3.10 and only consists of elementary calculations based on the non-degeneracy assumption on H. The only difference, is that we now demand the stronger statement (3.38), which is a consequence of the stronger assumption (**A1**$_{\mathrm{LSI}}$) in comparison to assumption (**A2**$_{\mathrm{PI}}$).

Now, we have collected the auxiliary statements and can proof Lemma 3.14.

Proof of Lemma 3.14. First, let us check the lower bound on the Hessian of $\tilde{H}$. We will use the same construction as of the PI in Lemma 3.11. Therefore, the support of $\tilde{H} - H$ is compact and $\tilde{H}$ is composed only of smooth functions, which already implies the lower bound on the Hessian for compact domains. Outside a sufficient large domain, the lower bound is just the Assumption (**A2**$_{\mathrm{LSI}}$). Now we can turn to verify the Lyapunov condition (3.27) and calculate with $W = \exp(\frac{1}{2\varepsilon}\tilde{H}_\varepsilon)$

$$\frac{1}{\varepsilon}\frac{LW}{W} = \frac{1}{2\varepsilon}\Delta\tilde{H}_\varepsilon + \frac{1}{4\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2 - \frac{1}{2\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2 = \frac{1}{2\varepsilon}\Delta\tilde{H}_\varepsilon - \frac{1}{4\varepsilon^2}|\nabla\tilde{H}_\varepsilon|^2.$$

If $|x| \geq \tilde{R}$ with $\tilde{R}$ given in Lemma 3.15, we apply (3.38) and have the Lyapunov condition fulfilled with constant $\lambda = \frac{C_H}{\varepsilon}$. This allows us to only consider $x \in B_{\tilde{R}} \cap \Omega$, which is of course bounded. In this case, Lemma 3.9 yields for $a\sqrt{\varepsilon} \leq |x| \leq \tilde{R}$ the estimate

$$\frac{1}{\varepsilon}\frac{LW}{W} \leq -\frac{\lambda_0}{\varepsilon} \leq -\frac{\lambda_0}{\tilde{R}^2\varepsilon}\,|x|^2\,. \tag{3.39}$$

For $|x| \leq a\sqrt{\varepsilon}$ holds by the representation (3.39) since H is smooth and strictly convex in $B_{a\sqrt{\varepsilon}}$ the bound

$$\frac{1}{\varepsilon}\frac{LW}{W} \leq \frac{1}{2\varepsilon}\Delta H(x) \leq \frac{b_0}{\varepsilon}. \tag{3.40}$$

A combination of (3.39) and (3.40) is the desired estimate (3.27). □

3.2. **Proof of the local inequalities.** In the previous Section 3.1, we were able to construct Lyapunov functions for the Hamiltonian restricted to the basin of attraction for each minimum. This is sufficient to finally prove the local PI and LSI of Theorem 2.6 and Theorem 2.7, which consist of mainly checking, whether the constants in the Lyapunov conditions show the right scaling behavior in ε. Let us start by restating the local PI.

Theorem 2.6 (Local Poincaré inequality)**.** *Under Assumption 1.6, the local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf. (2.1)), satisfy* PI(ϱ_i) *with*

$$\varrho_i^{-1} = O(\varepsilon).$$

Proof. We prove the theorem for each μ_i individually and omit the index i. The first step is the application of the Holley-Stroock perturbation principle in Lemma 3.4, which ensures that whenever $\tilde{H}_\varepsilon$ is an ε-modification of H, i.e. $\sup_{x\in\Omega}|\tilde{H}_\varepsilon(x) - H(x)| \leq C_{\tilde{H}}\varepsilon$, the Poincaré constants are of the same order in terms of scaling in ε, i.e.

$$\varrho \geq e^{-2C_{\tilde{H}}}\,\tilde{\varrho}. \tag{3.41}$$

In the next step, we construct an explicit ε-modification $\tilde{H}$ satisfying the Lyapunov condition Definition 3.7. Therefore, we can apply Theorem (3.8) with constant λ and b satisfying the bounds (3.6) from Lemma 3.9. This leads to a lower bound for $\tilde{\varrho}$ by

$$\tilde{\varrho} \geq \frac{\lambda\varrho_R}{b+\varrho_R} \geq \frac{\lambda_0\varrho_R}{b_0+\varepsilon\varrho_R}. \tag{3.42}$$

The final step is to observe that, since $R = a\sqrt{\varepsilon}$, we can assume that the measure $\tilde{\mu}_R = \tilde{\mu} \llcorner B_{a\sqrt{\varepsilon}}$ is just the measure $\hat{\mu}_a$. Therefore, it holds $\varrho_R^{-1} = O(\varepsilon)$ by Lemma 3.6, which leads by combining the estimates (3.41) and (3.42) to the conclusion $\varrho^{-1} = O(\varepsilon)$. □

Before continuing with the proof of the local LSI of Theorem 2.7, we want to remark, that the Lyapunov condition for the PI and in particular for the LSI imply an estimate of the second moment of μ.

Lemma 3.16 (Second moment estimate)**.** *If H fulfills the Lyapunov condition (3.4), then μ has finite second moment and it holds*

$$\int |x|^2 \, \mu(\mathrm{d}x) \leq \frac{1 + bR^2}{\lambda} \tag{3.43}$$

Proof. As it is outlined in [BBCG08], the Lyapunov condition (3.4) yields the following estimate: for any function f and $m \in \mathbb{R}$ it holds

$$\int (f - m)^2 \, \mathrm{d}\mu \leq \frac{1}{\lambda} \int |\nabla f|^2 \, \mathrm{d}\mu + \frac{b}{\lambda} \int_{\Omega_R} (f - m)^2 \, \mathrm{d}\mu.$$

We set $f(x) = |x|$ and $m = 0$ to observe the estimate (3.43). □

As a direct consequence, we get the desired estimate on the second moment.

Corollary 3.17. *If H fulfills the assumptions ($\mathbf{A1}_{\mathtt{PI}}$) and ($\mathbf{A2}_{\mathtt{PI}}$), then μ has finite second moment and it holds*

$$\int |x|^2 \, \mu(\mathrm{d}x) = O(\varepsilon).$$

Proof. We cannot apply the previous Lemma 3.16, but first have to do a change of measure to a measure $\tilde{\mu}$, where $\tilde{\mu}$ comes from an ε-modified Hamiltonian $\tilde{H}_\varepsilon$ of H

$$\int |x|^2 \, \mathrm{d}\mu \leq e^{2C_{\tilde{H}}} \int |x|^2 \, \mathrm{d}\tilde{\mu}.$$

Moreover, Lemma 3.9 ensures that $\tilde{H}_\varepsilon$ satisfies the Lyapunov condition (3.4) with constants $\lambda \geq \frac{\lambda_0}{\varepsilon}$, $b \leq \frac{b_0}{\varepsilon}$ and $R = a\sqrt{\varepsilon}$. Now, we can apply the previous Lemma 3.16 and immediately observe the result. □

Theorem 2.7 (Local logarithmic Sobolev inequality)**.** *Under Assumption 1.7, the local measures $\{\mu_i\}_{i=1}^M$, obtained by restricting μ to the basin of attraction Ω_i of the local minimum m_i (cf. (2.1)), satisfy* LSI(α_i) *with*

$$\alpha_i^{-1} = O(1).$$

Proof. For the same reason as in the proof of Theorem 2.6, we omit the index i. The first step is also the same as in the proof of Theorem 2.6. By Lemma 3.4 we obtain that, whenever $\tilde{H}_\varepsilon$ is an ε-modification of μ in the sense of Definition 3.3, the logarithmic Sobolev constants α and $\tilde{\alpha}$ of μ and $\tilde{\mu}$ satisfy $\alpha \geq \exp(-2C_{\tilde{H}})\tilde{\alpha}$.
The next step is to construct an explicit ε-modification $\tilde{H}$ satisfying the Lyapunov condition (3.27) of Theorem 3.12, which is provided by Lemma 3.14.
Additionally, the logarithmic Sobolev constant $\tilde{\alpha}$ depends on the second moment of $\tilde{\mu}$. Since $\tilde{H}_\varepsilon$ satisfies by Lemma 3.9 in particular the Lyapunov condition for PI (3.4) with constants $\lambda \geq \frac{\lambda_0}{\varepsilon}$, $b \leq \frac{b_0}{\varepsilon}$ and $R = a\sqrt{\varepsilon}$, we can apply Lemma 3.16 and arrive at

$$\int |x|^2 \, \mathrm{d}\tilde{\mu} \leq \frac{1 + R^2 b}{\lambda} \leq \frac{1 + b_0 a^2}{\lambda_0} \, \varepsilon = O(\varepsilon).$$

Now, we have control on all constants occurring in (3.28) and can determine the logarithmic Sobolev constant $\tilde{\alpha}$ of $\tilde{\mu}$. Let us estimate term by term of (3.28) and use the fact from Theorem (2.6), that $\tilde{\mu}$ satisfies PI($\tilde{\varrho}$) with $\tilde{\varrho}^{-1} = O(\varepsilon)$

$$2\sqrt{\frac{1}{\lambda}\left(\frac{1}{2} + \frac{b + \lambda\tilde{\mu}(|x|^2)}{\varrho}\right)} \leq 2\sqrt{\frac{\varepsilon}{\lambda_0}\left(\frac{1}{2} + O(1)\right)} = O(\sqrt{\varepsilon}).$$

The second term evaluates to $\frac{K_H}{2\varepsilon\lambda} = O(1)$ and finally the last one

$$\frac{K_H(b + \lambda\tilde{\mu}(|x|^2)) + 2\varepsilon\lambda}{\varrho\varepsilon\lambda} = O(\varepsilon)\left(K_H\left(\frac{b_0}{\varepsilon} + O(\varepsilon)\right) + O(1)\right) = O(1).$$

A combination of all the results leads to the conclusion $\tilde{\alpha}^{-1} = O(1)$ and since $\tilde{H}_\varepsilon$ is only an ε-modification of H also $\alpha^{-1} = O(1)$. □

4. Mean-difference estimates – weighted transport distance

This section is devoted to the proof of Theorem 2.9. We want to estimate the mean-difference $\left(\mathbb{E}_{\mu_i} f - \mathbb{E}_{\mu_j} f\right)^2$ for i and j fixed. The proof consists of four steps:

In the first step, we introduce the *weighted transport distance* in Section 4.1. This distance depends on the transport speed similarly to the Wasserstein distance, but in addition weights the speed of a transported particle w.r.t. the reference measure μ. The weighted transport distance allows in general for a variational characterization of the constant C in the inequality

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \leq C \int |\nabla f|^2 \,\mathrm{d}\mu.$$

The problem of finding good estimates of the constant C is then reduced to the problem of finding a good transport between the measures μ_i and μ_j w.r.t. to the weighted transport distance.

For measures as general as μ_i and μ_j, the construction of an explicit transport interpolation is not feasible. Therefore, the second step consists of an approximation, which is done in Section 4.2. There, the restricted measures μ_i and μ_j are replaced by *simpler* measures ν_i and ν_j, namely truncated Gaussians. We show in Lemma 4.6 that this approximation only leads to higher order error terms.

The most import step, the third one, consists of the estimation of the mean-difference w.r.t. the approximations ν_i and ν_j. Because the structure of ν_i and ν_j is very simple, we can explicitly construct a transport interpolation between ν_i and ν_j (see Lemma 4.11 in Section 4.3). The last step consists of collecting and controlling the error (cf. Section 4.4).

4.1. Mean-difference estimates by transport.

At the moment let us consider two arbitrary measures $\nu_0 \ll \mu$ and $\nu_1 \ll \mu$. The starting point of the estimation is a representation of the mean-difference as a transport interpolation. This idea goes back to Chafaï and Malrieu [CM10]. However, they used a similar but non-optimal estimate for our purpose. Hence, let us consider a transport interpolation $(\Phi_s : \mathbb{R}^n \to \mathbb{R}^n)_{s\in[0,1]}$ between ν_0 and ν_1, i.e. the family $(\Phi_s)_{s\in[0,1]}$ satisfies

$$\Phi_0 = \mathrm{Id}, \quad (\Phi_1)_\sharp \nu_0 = \nu_1, \quad \text{and} \quad (\Phi_s)_\sharp \nu_0 =: \nu_s.$$

The representation of the mean-difference as a transport interpolation is attained by using the fundamental theorem of calculus, i.e.

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 = \left(\int_0^1 \int \langle \nabla f(\Phi_s), \dot{\Phi}_s\rangle \,\mathrm{d}\nu_0 \,\mathrm{d}s\right)^2.$$

At this point it is tempting to apply the Cauchy-Schwarz inequality in $L^2(\mathrm{d}\nu_0 \times \mathrm{d}s)$ leading to the estimate of Chafaï and Malrieu [CM10]. However, this strategy would not yield the pre-exponential factors in the Eyring-Kramers formula (2.14) (cf. Remark 4.2). On Stephan Luckhaus' advice the authors realized the fact that it really matters on which integral you apply the Cauchy-Schwarz inequality. This insight lead to the following proceeding

$$\begin{aligned}
(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 &= \left(\int_0^1 \int \langle \nabla f, \dot\Phi_s \circ \Phi_s^{-1} \rangle \,\mathrm{d}\nu_s \,\mathrm{d}s \right)^2 \\
&= \left(\int \left\langle \nabla f, \int_0^1 \dot\Phi_s \circ \Phi_s^{-1} \,\frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \right\rangle \,\mathrm{d}\mu \right)^2 \\
&\leq \int \left| \int_0^1 \dot\Phi_s \circ \Phi_s^{-1} \,\frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \right|^2 \mathrm{d}\mu \int |\nabla f|^2 \,\mathrm{d}\mu. \qquad (4.1)
\end{aligned}$$

Note that in the last step we have applied the Cauchy-Schwarz inequality only in $L^2(\mathrm{d}\mu)$ and that the desired Dirichlet integral $\int |\nabla f|^2 \,\mathrm{d}\mu$ is already recovered.

The prefactor in front of the the Dirichlet energy on the right-hand side of (4.1) only depends on the transport interpolation $(\Phi_s)_{s\in[0,1]}$. Hence, we can infimize over all possible admissible transport interpolations and arrive at the following definition.

Definition 4.1 (Weighted transport distance $\mathcal{T}_\mu$)**.** Let μ be an absolutely continuous probability measure on $\mathbb{R}^n$ with connected support. Additionally, let ν_0 and ν_1 be two probability measures such that $\nu_0 \ll \mu$ and $\nu_1 \ll \mu$, then define the *weighted transport distance* by

$$\mathcal{T}_\mu^2(\nu_0, \nu_1) := \inf_{\Phi_s} \int \left| \int_0^1 \dot\Phi_s \circ \Phi_s^{-1} \,\frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \right|^2 \,\mathrm{d}\mu. \qquad (4.2)$$

The family $(\Phi_s)_{s\in[0,1]}$ is chosen absolutely continuous in the parameter s such that $\Phi_0 = \mathrm{Id}$ on $\operatorname{supp}\nu_0$ and $(\Phi_1)_\sharp \nu_0 = \nu_1$. For a fixed family and $(\Phi_s)_{s\in[0,1]}$ and a point $x \in \operatorname{supp}\mu$ the *cost density* is defined by

$$\mathcal{A}(x) := \left| \int_0^1 \dot\Phi_s \circ \Phi_s^{-1}(x) \,\nu_s(x) \,\mathrm{d}s \,\right|. \qquad (4.3)$$

Remark 4.2 (Relation of $\mathcal{T}_\mu$ to [CM10]). The transport distance $\mathcal{T}_\mu(\nu_0, \nu_1)$ is always smaller than the constant obtained by Chafaï and Malrieu [CM10, Section 4.6]. Indeed, applying the Cauchy-Schwarz inequality on $L^2(\mathrm{d}s)$ in (4.2) yields

$$\begin{aligned}
\mathcal{T}_\mu^2(\nu_0, \nu_1) &\leq \inf_{\Phi_s} \int \int_0^1 |\dot\Phi_s \circ \Phi_s^{-1}|^2 \,\frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \int_0^1 \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \,\mathrm{d}\mu \\
&\leq \inf_{\Phi_s} \left(\sup_x \left(\int_0^1 \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu}(x) \,\mathrm{d}s \right) \int \int_0^1 |\dot\Phi_s|^2 \,\mathrm{d}s \,\mathrm{d}\nu_0 \right),
\end{aligned}$$

where we used the assumption that $\nu_s \ll \mu$ for all $s \in [0,1]$ in the last L^1-L^∞-estimate.

Remark 4.3 (Relation of $\mathcal{T}_\mu$ to the L^2-Wasserstein distance W_2). If the support of μ is convex, we can set the transport interpolation $(\Phi_s)_{s\in[0,1]}$ to the linear interpolation map $\Phi_s(x) = (1-s)x + sU(x)$. Assuming that U is the optimal W_2-transport map between ν_0 and ν_1, the estimate in Remark 4.2 becomes

$$\mathcal{T}_\mu^2(\nu_0, \nu_1) \leq \left(\sup_x \int_0^1 \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu}(x) \,\mathrm{d}s \right) W_2^2(\nu_0, \nu_1).$$

Remark 4.4 (Invariance under time rescaling). The cost density $\mathcal{A}$ given by (4.3) is independent of rescaling the transport interpolation in the parameter s. Indeed, we observe that

$$\mathcal{A}(x) = \left| \int_0^1 \dot\Phi_s \circ \Phi_s^{-1}(x)\, \nu_s(x)\, \mathrm{d}s \right| = \left| \int_0^T \dot\Phi_t^T \circ (\Phi_t^T)^{-1}(x)\, \nu_t^T(x)\, \mathrm{d}t \right|,$$

where $\Phi_t^T = \Phi_{t/T}$ and $\nu_t^T = \nu_{t/T}$.

Remark 4.5 (Relation to negative Sobolev-norms). The weighted transport distance is a dynamic formulation for the negative Sobolev norm $H^{-1}(\mathrm{d}\mu)$ like Benamou and Brenier did for the Wasserstein distance [BB00]. Precisely, for $\nu_0 = \varrho_0\mu$ and $\nu_1 = \varrho_1\mu$ holds

$$\mathcal{T}_\mu(\nu_0, \nu_1) = \|\varrho_0 - \varrho_1\|_{H^{-1}(\mathrm{d}\mu)} = \inf_J \left\{ \int |J|^2 \,\mathrm{d}\mu : \varepsilon\nabla\cdot J - \nabla H \cdot J = \varrho_0 - \varrho_1 \right\}.$$

In fact, it is possible to define a whole class of weighted Wasserstein type distances interpolating between the negative Sobolev norm and the Wasserstein distance. Theses transports were found by Dolbeault, Nararet and Savaré [DNS09]. However, the authors were unaware of their work during the preparation of this article.

4.2. **Approximation of the local measures μ_i.** In this subsection we show that it is sufficient to consider only the mean-difference w.r.t. some auxiliary measures ν_i approximating μ_i for $i = 1, \ldots, M$. More precisely, the next lemma shows that there are nice measures ν_i which are close to the measures μ_i in the sense of the mean-difference.

Lemma 4.6 (Mean-difference of approximation)**.** *For $i = 1, \ldots, M$ let ν_i be a truncated Gaussian measure centered around the local minimum m_i with covariance matrix $\Sigma_i = (\nabla^2 H(m_i))^{-1}$, more precisely*

$$\nu_i(\mathrm{d}x) = \frac{1}{Z_{\nu_i}}\, e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}}\, \mathbb{1}_{E_i}(x)\, \mathrm{d}x, \quad \text{where} \quad Z_{\nu_i} = \int_{E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}}\, \mathrm{d}x, \qquad (4.4)$$

where we write $A[x] := \langle x, Ax\rangle$. The restriction E_i is given by an ellipsoid

$$E_i = \{x \in \mathbb{R}^n : |\Sigma_i^{-\frac{1}{2}}(x - m_i)| \le \sqrt{2\varepsilon}\,\omega(\varepsilon)\}. \qquad (4.5)$$

Additionally, assume that μ_i satisfies $\mathrm{PI}(\varrho_i)$ with $\varrho_i^{-1} = O(\varepsilon)$.
Then the following estimate holds

$$(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 \le O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)) \int |\nabla f|^2 \,\mathrm{d}\mu, \qquad (4.6)$$

where the function $\omega(\varepsilon) : \mathbb{R}^+ \to \mathbb{R}^+$ in (4.5) and (4.6) is smooth and monotone satisfying

$$\omega(\varepsilon) \ge |\log\varepsilon|^{\frac{1}{2}} \qquad \text{for } \varepsilon < 1.$$

The first step towards the proof of Lemma 4.6 is the following statement.

Lemma 4.7. *Let ν_i be a probability measure satisfying $\nu_i \ll \mu_i$. Moreover, if μ_i satisfies $\mathrm{PI}(\varrho_i)$ for some $\varrho_i > 0$, then the following estimate holds*

$$(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 \le \frac{1}{\varrho_i}\, \mathrm{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right) \int |\nabla f|^2 \,\mathrm{d}\mu_i. \qquad (4.7)$$

Proof of Lemma 4.7. The result is a consequence from the representation of the mean-difference as a covariance. Therefore, we note that $\mathrm{d}\nu_i = \frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i$ since $\nu_i \ll \mu_i$ and use the Cauchy-Schwarz inequality for the covariance

$$\begin{aligned}(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f))^2 &= \int f\,\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i - \int f \mathrm{d}\mu_i \int \frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i \\ &= \mathrm{cov}^2_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}, f\right) \leq \mathrm{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right)\mathrm{var}_{\mu_i}(f).\end{aligned}$$

Using the fact that μ_i satisfies a PI results in (4.7). □

The above lemma tells us that we only need to construct ν_i approximating μ_i in variance for $i = 1, \ldots M$. The following lemma provides exactly this.

Lemma 4.8 (Approximation in variance)**.** *Let the measures ν_i be given by (4.4). Then the partition sum Z_{ν_i} satisfies for ε small enough*

$$Z_{\nu_i} = (2\pi\varepsilon)^{\frac{n}{2}}\sqrt{\det \Sigma_i}\;(1 + O(\sqrt{\varepsilon})). \tag{4.8}$$

Additionally, ν_i is a good approximation in variance of μ_i, i.e.

$$\mathrm{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right) = O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)). \tag{4.9}$$

Proof of Lemma 4.8. The proof of (4.8) reduces to an estimate of a Gaussian integral on the complementary domain $\mathbb{R}^n \backslash E_i$

$$\begin{aligned}Z_{\nu_i} &= \int_{E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}}\mathrm{d}x \\ &= (2\pi\varepsilon)^{\frac{n}{2}}\sqrt{\det \Sigma_i}\left(1 - \frac{1}{(2\pi\varepsilon)^{\frac{n}{2}}\sqrt{\det\Sigma_i}}\int_{\mathbb{R}^n\backslash E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}}\mathrm{d}x\right).\end{aligned}$$

The integral on the complementary domain $\mathbb{R}^n\backslash E_i$ evaluates by the change of variables $x \mapsto y = (2\varepsilon\Sigma_i)^{-\frac{1}{2}}(x - m_i)$ to

$$\begin{aligned}\frac{1}{(2\pi\varepsilon)^{\frac{n}{2}}\sqrt{\det\Sigma_i}}\int_{\mathbb{R}^n\backslash E_i} e^{-\frac{\Sigma_i^{-1}[x-m_i]}{2\varepsilon}}\mathrm{d}x &= \frac{1}{\pi^{\frac{n}{2}}}\int_{\mathbb{R}^n\backslash B_{\omega(\varepsilon)}} e^{-y^2}\mathrm{d}y \\ = \frac{n}{\Gamma(\frac{n}{2}+1)}\int_{\omega(\varepsilon)}^{\infty} r^{n-1}e^{-r^2}\mathrm{d}r &= \frac{\Gamma(\frac{n}{2},\omega^2(\varepsilon))}{\Gamma(\frac{n}{2})},\end{aligned}$$

where $\Gamma(\frac{n}{2},\omega^2(\varepsilon))$ is the complementary incomplete Gamma function. It has the asymptotic expansion [Olv97, p. 109-112] given by

$$\Gamma\left(\frac{n}{2},\omega^2(\varepsilon)\right) = O(e^{-\omega^2(\varepsilon)}\omega^{n-2}(\varepsilon)), \quad \text{for } \omega(\varepsilon) \geq \sqrt{n}.$$

We obtain (4.8) by the choice of $\omega(\varepsilon) \geq |\log\varepsilon|^{\frac{1}{2}}$, since the error becomes

$$O(e^{-\omega^2(\varepsilon)}\omega^{n-2}(\varepsilon)) = O(\varepsilon\,|\log\varepsilon|^{\frac{n}{2}-1}) = O(\sqrt{\varepsilon}), \qquad \text{for } \varepsilon \leq e^{-n}.$$

For the proof of (4.9), we compare the asymptotic expression for $Z_{\mu_i} = Z_i Z_\mu e^{\varepsilon^{-1}m_i}$ from (2.19) and Z_{ν_i} and obtain

$$Z_{\mu_i} = Z_{\nu_i} + O(\sqrt{\varepsilon}). \tag{4.10}$$

The relative density of ν_i w.r.t. μ_i can be estimated by Taylor expanding H around m_i. By the definition of ν_i given in (4.4), we obtain that $\Sigma_i^{-1}[y - m_i] - H_i(y) = O(|y - m_i|^3)$. This observation together with (4.10) leads to

$$\begin{aligned}\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}(y) &= \frac{Z_{\mu_i}}{Z_{\nu_i}}e^{-\frac{1}{2\varepsilon}\Sigma_i^{-1}[y-m_i]+\frac{1}{2\varepsilon}H_i(y)}\mathbb{1}_{E_i}(y) = \frac{Z_{\mu_i}}{Z_{\nu_i}}e^{\frac{O(|y-m_i|^3)}{\varepsilon}}\mathbb{1}_{E_i}(y) \\ &= 1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)).\end{aligned}$$

Now, the conclusion directly follows from the definition of the variance

$$\begin{aligned}\operatorname{var}_{\mu_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right) &= \int_{E_i}\left(\frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\right)^2 \mathrm{d}\mu_i - \left(\int \frac{\mathrm{d}\nu_i}{\mathrm{d}\mu_i}\,\mathrm{d}\mu_i\right)^2 \\ &= \int_{E_i} 1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\,\mathrm{d}\mu_i - \left(\int_{E_i}\mathrm{d}\nu_i\right)^2 \\ &= 1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon)) - 1.\end{aligned}$$

□

Proof of Lemma 4.6. A combination of Lemma 4.7 and Lemma 4.8 together with the assumption $\varrho_i^{-1} = O(\varepsilon)$ immediately reveals

$$\left(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\mu_i}(f)\right)^2 \overset{(4.7),(4.9)}{\leq} O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)) \int |\nabla f|^2 \,\mathrm{d}\mu_i.$$

□

4.3. **Affine transport interpolation.** The aim of this section is to estimate $\left(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f)\right)^2$ with the help of the weighted transport distance $\mathcal{T}_\mu(\nu_i,\nu_j)$ introduced in Section 4.1. The main result of this section estimates the weighted transport distance $\mathcal{T}_\mu(\nu_i,\nu_j)$ and is formulated in Lemma 4.11. For the proof of Lemma 4.11, we construct an explicit transport interpolation between ν_i and ν_j w.r.t. the measure μ. We start with a class of possible transport interpolations and optimize the weighted transport cost in this class.

Let us state the main idea of this optimization procedure. Therefore, we recall that the measures ν_i and ν_j are truncated Gaussians by the approximation we have done in the previous Section 4.2. Hence, the measures ν_i and ν_j are characterized by their mean and covariance matrix. We will choose the transport interpolation (cf. Section 4.3.1) such that the push forward measures $\nu_s := (\Phi_s)_\sharp \nu_0$ are again truncated Gaussians. Hence, it is sufficient to optimize among all paths γ connecting the minima m_i and m_j and all covariance matrices interpolating between Σ_i and Σ_j.

4.3.1. *Definition of regular affine transport interpolations.* Let us state in this section the class of transport interpolation among we want to optimize the weighted transport cost.

Definition 4.9 (Affine transport interpolations)**.** Assume that the measures ν_i and ν_j are given by Lemma 4.6. In detail, $\nu_i = \mathcal{N}(m_i, \varepsilon^{-1}\Sigma_i)\llcorner E_i$ and $\nu_j = \mathcal{N}(m_j, \varepsilon^{-1}\Sigma_j)\llcorner E_j$ are truncated Gaussians centered in m_i and m_j with covariance matrices $\varepsilon^{-1}\Sigma_i$ and $\varepsilon^{-1}\Sigma_j$. The restriction E_i and E_j are given for $l = 1,\dots,M$ by the ellipsoids

$$E_l = \{x \in \mathbb{R}^n : |\Sigma_l^{-\frac{1}{2}}(x - m_l)| \leq \sqrt{2\varepsilon}\,\omega(\varepsilon)\}, \qquad \text{where } \omega(\varepsilon) \geq |\log\varepsilon|^{\frac{1}{2}}.$$

A transport interpolation Φ_s between ν_i and ν_j is called *affine transport interpolation* if there exists

- an interpolation path $(\gamma_s)_{s\in[0,T]}$ between $m_i = \gamma_0$ and $m_j = \gamma_T$ satisfying
$$\gamma = (\gamma_s)_{s\in[0,T]} \in C^2([0,T],\mathbb{R}^n) \quad \text{and} \quad \forall s \in [0,T] : \dot\gamma_s \in S^{n-1}, \tag{4.11}$$
- an interpolation path $(\Sigma_s)_{s\in[0,T]}$ of covariance matrices between Σ_i and Σ_j satisfying
$$\Sigma = (\Sigma_s)_{s\in[0,T]} \in C^2([0,T],\mathbb{R}^{n\times n}_{\mathrm{sym},+}), \quad \Sigma_0 = \Sigma_i \quad \text{and} \quad \Sigma_T = \Sigma_j,$$

such that the transport interpolation $(\Phi_s)_{s\in[0,T]}$ is given by

$$\Phi_s(x) = \Sigma_s^{\frac{1}{2}} \Sigma_0^{-\frac{1}{2}} (x - m_0) + \gamma_s. \tag{4.12}$$

Since the cost density $\mathcal{A}$ given by (4.3) is invariant under rescaling of time (cf. Remark 4.4), one can always assume that the interpolation path γ_s is parameterized by arc-length. Hence, the condition $\dot\gamma_s \in S^{n-1}$ (cf. (4.11)) is not restricting.

We want to emphasize that for an affine transport interpolation $(\Phi_s)_{s\in[0,T]}$ the push forward measure $(\Phi_s)_\sharp \nu_0 = \nu_s$ is again a truncated Gaussian $\mathcal{N}(\gamma_s, \varepsilon^{-1}\Sigma_s)\llcorner E_s$, where E_s is the support of ν_s being again an ellipsoid in $\mathbb{R}^n$ given by

$$E_s = \{x \in \mathbb{R}^n : |\Sigma_s^{-\frac{1}{2}}(x - \gamma_s)| \le \sqrt{2\varepsilon}\, \omega(\varepsilon)\}. \tag{4.13}$$

Therewith, the partition sum of ν_s is given by (cf. (4.8))

$$Z_{\nu_s} = (2\pi\varepsilon)^{\frac{n}{2}} \sqrt{\det \Sigma_s}(1 + O(\sqrt{\varepsilon})). \tag{4.14}$$

By denoting $\sigma_s = \Sigma_s^{\frac{1}{2}}$ and using the definition (4.12) of the affine transport interpolation $(\Phi_s)_{s\in[0,T]}$, we arrive at the relations

$$\begin{aligned} \dot\Phi_s(x) &= \dot\sigma_s \sigma_0^{-1}(x - m_0) + \dot\gamma_s, \\ \Phi_s^{-1}(y) &= \sigma_0 \sigma_s^{-1}(y - \gamma_s) + m_0, \\ \dot\Phi_s \circ \Phi_s^{-1}(y) &= \dot\sigma_s \sigma_s^{-1}(y - \gamma_s) + \dot\gamma_s. \end{aligned}$$

Among all possible affine transport interpolations we are considering only those satisfying the following regularity assumption.

Assumption 4.10 (Regular affine transport interpolations)**.** *An affine transport interpolation* $(\gamma_s, \Sigma_s)_{s\in[0,T]}$ *belongs to the class of* regular affine transport interpolations *if the length* $T < T^*$ *is bounded by some uniform* $T^* > 0$ *large enough. Further, for a uniform constant* $c_\gamma > 0$ *holds*

$$\inf \{r(x, y, z) : x, y, z \in \gamma, x \neq y \neq z \neq x\} \ge c_\gamma, \tag{4.15}$$

where $r(x, y, z)$ *denotes the radius of the unique circle through the three distinct points* x, y *and* z*. Furthermore, there exists a uniform constant* $C_\Sigma \ge 1$ *for which*

$$C_\Sigma^{-1}\,\mathrm{Id} \le \Sigma_s \le C_\Sigma\,\mathrm{Id} \qquad \textit{and} \qquad \|\dot\Sigma_s\| \le C_\Sigma. \tag{4.16}$$

The infimum in condition (4.15) is called *global radius of curvature* (cf. [GMSvdM02]). It ensures that a small neighborhood of size $\frac{c_\gamma}{2}$ around γ is not self-intersecting, since the infimum can only be attained for the following three cases:

(i) All three points in a minimizing sequence of (4.15) coalesce to a point at which the radius of curvature is minimal.
(ii) Two points coalesce to a single point and the third converges to another point, such that the both points are a pair of closest approach.
(iii) Two points coalesce to a single point and the third converges to the starting or ending point of γ.

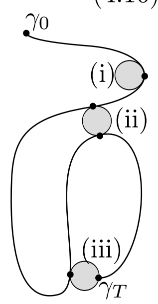

FIGURE 3. Global radius of curvature

In the following calculations, there often occurs a multiplicative error of the form $1+O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))$. Therefore, let us introduce for convenience the notation "$\approx$" meaning "$=$" up to the multiplicative error $1+O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))$. The symbols "$\lesssim$" and "$\gtrsim$" have the analogous meaning.

Now, we can formulate the key ingredient for the proof of Theorem 2.9, namely the estimation of the weighted transport distance $\mathcal{T}_\mu(\nu_i, \nu_j)$.

Lemma 4.11. *Assume that ν_i and ν_j are given by Lemma 4.6. Then the weighted transport distance $\mathcal{T}_\mu(\nu_i,\nu_j)$ can be estimated as*

$$\begin{aligned}
\mathcal{T}_\mu^2(\nu_i,\nu_j) &= \inf_{\Phi_s} \int \left(\int_0^1 |\dot{\Phi}_s \circ \Phi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right)^2 \, \mathrm{d}\mu \\
&\le \inf_{\Psi_s} \int \left(\int_0^1 |\dot{\Psi}_s \circ \Psi_s^{-1}| \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right)^2 \, \mathrm{d}\mu \\
&\lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, 2\pi\varepsilon \left(\frac{\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} + \frac{T(C_\Sigma)^{\frac{n-1}{2}}}{\sqrt{2\pi\varepsilon}} e^{-\omega^2(\varepsilon)} \right) \, e^{\frac{H(s_{i,j})}{\varepsilon}},
\end{aligned} \tag{4.17}$$

where the infimum over Ψ_s only considers regular affine transport interpolations Ψ_s in the sense of Assumption 4.10.
In particular, if we choose $\omega(\varepsilon) \ge |\log\varepsilon|^{\frac{1}{2}}$, which is enforced by Lemma 4.6, we get the estimate

$$\mathcal{T}_\mu^2(\nu_i,\nu_j) \le \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, \frac{2\pi\varepsilon\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \, e^{\frac{H(s_{i,j})}{\varepsilon}} \, \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right). \tag{4.18}$$

Before turning to the proof of Lemma 4.11, we want to anticipate the structure of the affine transport interpolation (γ,Σ) which realizes the desired estimate (4.18): Having a closer look at the structure of the weighted transport distance $\mathcal{T}_\mu^2(\nu_i,\nu_j)$, it becomes heuristically clear that the mass should be transported from E_i to E_j over the saddle point $s_{i,j}$ into the direction of the eigenvector to the negative eigenvalue $\lambda^-(s_{i,j})$ of $\nabla^2 H(s_{i,j})$. There, only the region around the saddle gives the main contribution to the estimate (4.18). Then, we only have one more free parameter to choose for our affine transport interpolation (γ,Σ): It is the covariance structure Σ_{τ^*} of the interpolating truncated Gaussian measure ν_{τ^*} at the passage time τ^* at the saddle point $s_{i,j}$. In the proof of Lemma 4.11 below, we will see by an optimization procedure that the best Σ_{τ^*} is given by $\Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{i,j})$, restricted to the stable manifold of the saddle point $s_{i,j}$.

The proof of Lemma 4.11 presents the core of the proof of the Eyring-Kramers formulas and consists of three steps carried out in the following sections:

- In Section 4.3.2, we carry out some preparatory work: We introduce tube coordinates on the support of the transport cost $\mathcal{A}$ given by (4.3) (cf. Lemma 4.12), we deduce a pointwise estimate on the transport cost $\mathcal{A}$ and we give a rough a priori estimate on the transport cost $\mathcal{A}$.
- In Section 4.3.3, we split the transport cost into a transport cost around the saddle and the complement. We also estimate the transport cost of the complement yielding the second summand in the desired estimate (4.17).
- In Section 4.3.4, we finally deduce a sharp estimate of the transport cost around the saddle yielding the first summand in the desired estimate (4.17).

4.3.2. *Preparations and auxiliary estimates.* The main reason for making the regularity Assumption 4.10 on affine transport interpolations is that we can introduce tube coordinates around the path γ. In these coordinates, the calculation of the cost density $\mathcal{A}$ given by (4.3) becomes a lot handier.

We start with defining the caps E_0^- and E_T^+ as

$$E_0^- := \{x \in E_0 : \langle x - \gamma_0, \dot{\gamma}_0\rangle < 0\} \quad \text{and} \quad E_T^+ := \{x \in E_T : \langle x - \gamma_T, \dot{\gamma}_T\rangle > 0\},$$

The caps E_0^- and E_T^+ have no contribution to the total cost but unfortunately need some special treatment. Further, we define the slices V_s with $s \in [0,T]$

$$V_s = \{x \in \operatorname{span}\{\dot{\gamma}_s\}^\perp : |\Sigma_s^{-\frac{1}{2}} x| \le \sqrt{2\varepsilon}\,\omega(\varepsilon)\}.$$

In span V_s we can choose a basis $e_s^2, \ldots, e_s^n$ smoothly depending on the parameter s. In particular, there exists a family $(Q_s)_{s\in[0,T]} \in C^2([0,T], SO(n))$ satisfying the same regularity assumption as the family $(\Sigma_\tau)_{\tau\in[0,T]}$ such that

$$Q_s e^1 = \dot\gamma_s, \qquad Q_s e^i = e_s^i, \qquad \text{for } i = 2, \ldots, n, \tag{4.19}$$

where $(e^1, \ldots, e^n)$ is the canonical basis of $\mathbb{R}^n$.

Let use now define the tube E as

$$E = \bigcup_{s\in[0,T]} (\gamma_s + V_s)\,.$$

The support of the cost density $\mathcal{A}$ given by (4.3) is now given by

$$\operatorname{supp}\mathcal{A} = E_0^- \cup E \cup E_T^+. \tag{4.20}$$

By the definition (4.13) of E_s and the uniform bound (4.16) on Σ_s holds

$$\operatorname{diam} V_s \le 2\sqrt{2\varepsilon C_\Sigma}\,\omega(\varepsilon). \tag{4.21}$$

Therewith, we find

$$\operatorname{supp}\mathcal{A} \subset B_{2\sqrt{2\varepsilon C_\Sigma}\,\omega(\varepsilon)}((\gamma_\tau)_{\tau\in[0,T]}) := \left\{x\in\mathbb{R}^n : |x-\gamma_\tau| \le 2\sqrt{2\varepsilon C_\Sigma}\,\omega(\varepsilon)\right\}.$$

The assumption (4.13) ensures that $B_{2\sqrt{2\varepsilon C_\Sigma}\,\omega(\varepsilon)}((\gamma_\tau)_{\tau\in[0,T]})$ is not self-intersecting for any ε small enough. The next lemma just states that by changing to tube coordinates in E one can asymptotically neglect the Jacobian determinant $\det J$.

Lemma 4.12 (Change of coordinates)**.** *The change of coordinates $(\tau, z) \mapsto x = \gamma_\tau + z_\tau$ with $z_\tau \in V_\tau$ satisfies for any function ξ on E*

$$\int_E \xi(x)\,\mathrm{d}x \approx \int_0^T \int_{V_\tau} \xi(\gamma_\tau + z_\tau)\,\mathrm{d}z_\tau\,\mathrm{d}\tau.$$

Proof of Lemma 4.12. We use the representation of the tube coordinates via (4.19). Therewith, it holds that $x = \gamma_\tau + Q_\tau z$, where $z \in \{0\}\times\mathbb{R}^{n-1}$. Then, the Jacobian J of the coordinate change $x \mapsto (\tau, Q_\tau z)$ is given by

$$J = (\dot\gamma_\tau + \dot Q_\tau z, (Q_\tau)_2, \ldots, (Q_\tau)_n) \in \mathbb{R}^{n\times n},$$

where $(Q_\tau)_i$ denotes the i-th column of Q_τ. By the definition (4.19) of Q_τ follows $\dot\gamma_\tau = (Q_\tau)_1$. Hence, we have the representation $J = Q_\tau + \dot Q_\tau z \otimes e_1$. The determinant of J is then given by

$$\det\left(Q_\tau + \dot Q_\tau z\otimes e_1\right) = \underbrace{\det(Q_\tau)}_{=1}\,\det\left(\operatorname{Id} + (Q_\tau^\top \dot Q_\tau z)\otimes e_1\right) = 1 + \left(Q_\tau^\top \dot Q_\tau z\right)_1.$$

By Assumption 4.10 holds $\|\dot Q_\tau\| \le C_\Sigma$ implying $(Q_\tau^\top \dot Q_\tau z)_{1,1} = O(z)$. Since $Q_\tau z \in V_\tau$, we get $O(z) = O(\sqrt{\varepsilon}\,\omega(\varepsilon))$ by (4.21). Hence we get

$$\det J = 1 + O(\sqrt{\varepsilon}\,\omega(\varepsilon)),$$

which concludes the proof. □

An important tool is the following auxiliary estimate.

Lemma 4.13 (Pointwise estimate of the cost-density $\mathcal{A}$)**.** *For $x \in \operatorname{supp}\mathcal{A}$ we define*

$$\tau = \operatorname*{arg\,min}_{s\in[0,T]} |x - \gamma_s| \qquad \textit{and} \qquad z_\tau := x - \gamma_\tau. \tag{4.22}$$

Then the following estimate holds

$$\mathcal{A}(x) \lesssim (2\pi\varepsilon)^{-\frac{n-1}{2}} \sqrt{\det_{1,1}(Q_\tau^\top \tilde\Sigma_\tau^{-1} Q_\tau)}\;\; e^{-\frac{\tilde\Sigma_\tau^{-1}[z_\tau]}{2\varepsilon}} =: P_\tau\; e^{-\frac{\tilde\Sigma_\tau^{-1}[z_\tau]}{2\varepsilon}}, \tag{4.23}$$

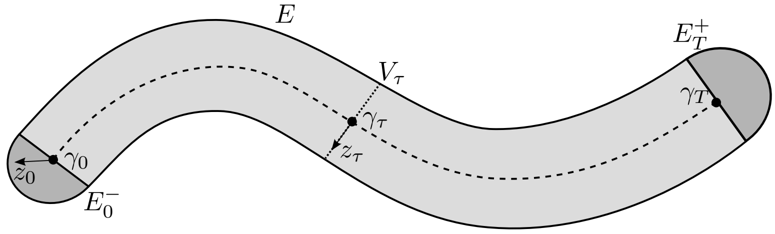

FIGURE 4. The support of $\mathcal{A}$ in tube coordinates.

where Q_τ is defined in (4.19) and $\tilde{\Sigma}_\tau^{-1}$ is given by

$$\tilde{\Sigma}_\tau^{-1} = \Sigma_\tau^{-1} - \frac{1}{\Sigma_\tau^{-1}[\dot{\gamma}_\tau]} \Sigma_\tau^{-1}\dot{\gamma}_\tau \otimes \Sigma_\tau^{-1}\dot{\gamma}_\tau. \tag{4.24}$$

Further, $\det_{1,1} A$ *is the determinant of the matrix obtained from A removing the first row and column.*

Remark 4.14. With a little bit of additionally work, one could show that (4.23) holds with "$\approx$" instead of "$\lesssim$". It follows from (4.24) that the matrix $\tilde{\Sigma}_\tau^{-1}$ is positive definite. Hence, $\mathcal{A}$ is an $\mathbb{R}^{n-1}$-dimensional Gaussian on the slice $\gamma_\tau + V_\tau$ up to approximation errors.

Proof of Lemma 4.13. We start the proof with some preliminary remarks and results. By the regularity Assumption 4.10 on the transport interpolation, we find that for all $x \in \operatorname{supp}\mathcal{A}$ holds uniformly

$$I_T(x) := \{s : E_s \ni x\} \quad \text{satisfies} \quad \mathcal{H}^1(I_T(x)) = O(\sqrt{\varepsilon}\,\omega(\varepsilon)).$$

This allows to linearize the transport interpolation around τ given in (4.22). It holds for s such that $x \in E_s$

$$\begin{aligned} \Sigma_s^{-1}[x-\gamma_s] &= \Sigma_\tau^{-1}[\gamma_\tau + z_\tau - \gamma_s] + O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)) \\ &= \Sigma_\tau^{-1}[(\tau - s)\dot{\gamma}_\tau + z_\tau] + O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)). \end{aligned} \tag{4.25}$$

For similar reasons, we can linearize the determinant $\det\Sigma_s$ and have $\det\Sigma_s = \det\Sigma_\tau + O(\sqrt{\varepsilon}\,\omega(\varepsilon))$. Finally, we have the following bound on the transport speed

$$\begin{aligned} |\dot{\Phi}_s \circ \Phi_s^{-1}(x)| \mathbb{1}_{E_s}(x) &= \left|\dot{\sigma}_s \sigma_s^{-1}(x-\gamma_s) + \dot{\gamma}_s\right| \mathbb{1}_{E_s}(x) \\ &\leq \left(\left|\dot{\sigma}_s \sigma_s^{-1}(x-\gamma_s)\right| + |\dot{\gamma}_s|\right) \mathbb{1}_{E_s}(x) \\ &\leq \left(C_\Sigma\,|x-\gamma_s| + 1\right) \mathbb{1}_{E_s}(x) = \left(1 + O(\sqrt{\varepsilon}\,\omega(\varepsilon))\right) \mathbb{1}_{E_s}(x). \end{aligned} \tag{4.26}$$

Let us first consider the case $x \in E$. We use (4.14), (4.25) and (4.26) to arrive with $x = \gamma_\tau + z_\tau$ where $z_\tau \in V_\tau$ at

$$\begin{aligned} \mathcal{A}(x) &= \int_{I_T(x)} |\dot{\Phi}_s \circ \Phi_s^{-1}(x)|\, \frac{1}{Z_{\nu_s}}\, \exp\left(-\frac{1}{2\varepsilon}\Sigma_s^{-1}[x-\gamma_s]\right) \mathbb{1}_{E_s}(x)\,\mathrm{d}s \\ &\leq \frac{1}{(2\pi\varepsilon)^{\frac{n}{2}}} \int_{I_T(x)} \frac{1 + O(\sqrt{\varepsilon}\,\omega(\varepsilon))}{\sqrt{\det\Sigma_s}}\, \exp\left(-\frac{1}{2\varepsilon}\Sigma_s^{-1}[x-\gamma_s]\right) \mathrm{d}s \\ &\lesssim \frac{1}{(2\pi\varepsilon)^{\frac{n}{2}}\sqrt{\det\Sigma_\tau}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2\varepsilon}\Sigma_\tau^{-1}[(\tau-s)\dot{\gamma}_\tau + z_\tau]\right) \mathrm{d}s \\ &= \frac{\sqrt{\det\Sigma_\tau^{-1}}}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{\sqrt{2\pi\varepsilon}}{\sqrt{\Sigma_\tau^{-1}[\dot{\gamma}_\tau]}} \exp\left(-\frac{1}{2\varepsilon}\tilde{\Sigma}_\tau^{-1}[z_\tau]\right) \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right), \end{aligned}$$

where the last step follows by an application of a partial Gaussian integration (cf. Lemma B.1). Finally, by using the relation (C.2), we get that

$$\frac{\Sigma_\tau^{-1}}{\Sigma_\tau^{-1}[\dot\gamma_\tau]} = \det{}_{1,1}(Q_\tau^\top \tilde\Sigma_\tau^{-1} Q_\tau),$$

and conclude the hypothesis for this case.

Let us now consider the case $x \in E_0^- \cup E_T^+$. For convenience, we only consider the case $x \in E_0^-$. By the definition of E_0^- holds $\tau = 0$. The integration domain $I_T(x)$ is now given by

$$I_T(x) = [0, s^*) \quad \text{with} \quad s^* = O(\sqrt{\varepsilon}\,\omega(\varepsilon)). \tag{4.27}$$

Therewith, we can estimate $\mathcal{A}(x)$ in the same way as for $x \in E$ and conclude the proof. □

We only need one more ingredient for the proof of Lemma 4.11. It is an a priori estimate on the cost density $\mathcal{A}$.

Lemma 4.15 (A priori estimates for the cost density $\mathcal{A}$)**.** *For $\mathcal{A}$ it holds:*

$$\int \mathcal{A}(x)\,\mathrm{d}x \lesssim T, \qquad \textit{and} \tag{4.28}$$

$$\mathcal{A}(x) \lesssim \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{\frac{n-1}{2}} \qquad \textit{for } x \in \operatorname{supp}\mathcal{A}. \tag{4.29}$$

Proof of Lemma 4.15. Let us first consider the estimate (4.28). It follows from the characterization (4.20) of the support of $\mathcal{A}$ that

$$\int \mathcal{A}(x)\,\mathrm{d}x = \int_E \mathcal{A}(x)\,\mathrm{d}x + \int_{E_0^- \cup E_T^+} \mathcal{A}(x)\,\mathrm{d}x. \tag{4.30}$$

Now, we estimate the first term on the right-hand side of the last identity. Using the change to tube coordinates of Lemma 4.12 and noting that the upper bound (4.23) is a $(n-1)$-dimensional Gaussian density on V_τ for $\tau \in [0,T]$, we can easily infer that

$$\int_E \mathcal{A}(x)\,\mathrm{d}x \lesssim |\gamma| = T$$

Let us turn to the second term on the right-hand side of (4.30). For convenience, we only consider the integral w.r.t. the cap E_0^-. It follows from (4.26) and (4.27) that

$$\begin{aligned}\int_{E_0^-} \mathcal{A}(x)\,\mathrm{d}x &\lesssim \int_{E_0^-}\int_0^1 \nu_s(x)\,\mathrm{d}s\,\mathrm{d}x = \int_0^{s^*}\int_{E_0^-} \nu_s(x)\,\mathrm{d}x\,\mathrm{d}s \\ &\lesssim \int_0^{s^*}\int \nu_s(x)\,\mathrm{d}x\,\mathrm{d}s = s^* = O(\sqrt{\varepsilon}\,\omega(\varepsilon)),\end{aligned}$$

which yields the desired statement (4.28).

Let us now consider the estimate 4.29. Note by Remark 4.14 the matrix $\tilde\Sigma_\tau^{-1}$ given by (4.24) is positive definite and the matrix we subtract is also positive definite. Therefore, it holds in the sense of quadratic forms

$$0 < \tilde\Sigma_\tau^{-1} = \Sigma_\tau^{-1} - \frac{1}{\Sigma_\tau^{-1}[\dot\gamma_\tau]}\Sigma_\tau^{-1} \otimes \Sigma_\tau^{-1} \le \Sigma_\tau^{-1}.$$

Now, the uniform bound (4.16) yields

$$\sqrt{\det{}_{1,1}(Q_\tau^\top \tilde\Sigma_\tau^{-1} Q_\tau)} \le C_\Sigma^{\frac{n-1}{2}}.$$

Then, the desired statement (4.29) follows directly from the estimate (4.23). □

4.3.3. *Reduction to neighborhood around the saddle.* Firstly, observe that from (4.29) follows the a priori estimate

$$\frac{\mathcal{A}^2(x)}{\mu(x)} \lesssim \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{n-1} Z_\mu \, e^{\frac{1}{\varepsilon}H(x)}. \tag{4.31}$$

Hence, on an exponential scale, the leading order contribution to the cost comes from neighborhoods of points where $H(x)$ is large. Therefore, we want to make the set, where H is comparable to its value at the optimal connecting saddle $s_{i,j}$, as small as possible. For this purpose, let us define the following set

$$\Xi_{\gamma,\Sigma} := \left\{x \in \operatorname{supp} \mathcal{A} : H(x) \geq H(s_{i,j}) - \varepsilon\omega^2(\varepsilon)\right\}. \tag{4.32}$$

Therewith, we obtain by denoting the complement $\Xi^c_{\gamma,\Sigma} := \operatorname{supp} \mathcal{A} \backslash \Xi_{\gamma,\Sigma}$ the splitting

$$\mathcal{T}^2_\mu(\nu_i, \nu_j) \leq \int_{\Xi_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)} \, \mathrm{d}x + \int_{\Xi^c_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)} \, \mathrm{d}x.$$

The integral on $\Xi^c_{\gamma,\Sigma}$ can be estimated with the a priori estimate (4.31) and Lemma 4.15 as follows

$$\begin{aligned}
\int_{\Xi^c_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)} \, \mathrm{d}x &\overset{(4.32)}{\leq} Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon} - \omega^2(\varepsilon)} \int_{\Xi^c_{\gamma,\Sigma}} \mathcal{A}^2(x) \, \mathrm{d}x \\
&\overset{(4.29)}{\lesssim} Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon} - \omega^2(\varepsilon)} \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{\frac{n-1}{2}} \int \mathcal{A}(x) \, \mathrm{d}x \qquad (4.33)\\
&\overset{(4.28)}{\lesssim} Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon} - \omega^2(\varepsilon)} \left(\frac{C_\Sigma}{2\pi\varepsilon}\right)^{\frac{n-1}{2}} T.
\end{aligned}$$

We observe that estimate (4.33) is the second summand in the desired bound (4.17).

4.3.4. *Cost estimate around the saddle.* The aim of this subsection is to deduce the estimate

$$\int_{\Xi_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)} \, \mathrm{d}x \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, e^{\frac{H(s_{i,j})}{\varepsilon}} \, \frac{2\pi\varepsilon\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|}. \tag{4.34}$$

Note that this estimate would yield the missing ingredient for the verification of the desired estimate (4.17).

By the non-degeneracy Assumption 1.9, we can assume that ε is small enough such that $E_0^- \cup E_T^+ \subset \Xi^c_{\gamma,\Sigma}$. It follows that $\Xi_{\gamma,\Sigma} \subset E$. We claim that the transport interpolation Φ_s can be chosen such that there exists a connected interval $I_T \subset [0,T]$ satisfying

$$\Xi_{\gamma,\Sigma} \subset \bigcup_{s \in I_T} (V_s + \gamma_s) \qquad \text{and} \qquad \mathcal{H}^1(I_T) = O(\sqrt{\varepsilon}\, \omega(\varepsilon)). \tag{4.35}$$

Indeed, the level set $\left\{x \in \mathbb{R}^n : H(x) \leq H(s_{i,j}) - \varepsilon\omega^2(\varepsilon)\right\}$ consists of at least two connected components M_i and M_j such that $m_i \in M_i$ and $m_j \in M_j$. Further, it holds

$$\operatorname{dist}(M_i, M_j) = \inf_{x \in M_i, y \in M_j} |x - y| = O(\sqrt{\varepsilon}\, \omega(\varepsilon)),$$

which follows from expanding H around $s_{i,j}$ in direction of the eigenvector corresponding to the negative eigenvalue of $\nabla^2 H(s_{i,j})$. We can choose the path γ in direction of this eigenvector in a neighborhood of size $O(\sqrt{\varepsilon}\, \omega(\varepsilon))$ around $s_{i,j}$, which shows (4.35).

Combining the covering (4.35) and Lemma 4.12 yields the estimate

$$\int_{\Xi_{\gamma,\Sigma}} \frac{\mathcal{A}^2(x)}{\mu(x)} \, \mathrm{d}x \leq \int_{I_T} \int_{V_s} \frac{\mathcal{A}^2(\gamma_s + z_s)}{\mu(\gamma_s + z_s)} \, \mathrm{d}z_s \, \mathrm{d}s. \tag{4.36}$$

Recalling the definition (4.19) of the family of rotations $(Q_\tau)_{\tau\in[0,T]}$, it holds that $z_\tau = Q_\tau z$ with $z \in \{0\} \times \mathbb{R}^{n-1}$. Hence, the following relation holds

$$\int_{I_T} \int_{V_\tau} \frac{\mathcal{A}^2(\gamma_\tau + z_\tau)}{\mu(\gamma_\tau + z_\tau)} \, \mathrm{d}z_\tau \, \mathrm{d}\tau = \int_{\{0\}\times\mathbb{R}^{n-1}} \int_{I_T} \mathbb{1}_{V_\tau}(Q_\tau z) \frac{\mathcal{A}^2(\gamma_\tau + Q_\tau z)}{\mu(\gamma_\tau + Q_\tau z)} \, \mathrm{d}\tau \, \mathrm{d}z. \tag{4.37}$$

The next step is to rewrite $H(\gamma_\tau + Q_\tau z)$. We can assume, that γ actually passes the saddle $s_{i,j}$ at time $\tau^* \in (0,T)$. Then, by the reason that $|z_\tau| = O(\sqrt{\varepsilon}\ \omega(\varepsilon))$ for $z_\tau \in V_\tau$ and the global non-degeneracy assumption (1.4), we can Taylor expand $H(\gamma_\tau + z_\tau)$ around $s_{i,j} = \gamma_{\tau^*}$ for $\tau \in I_T$ and $z_\tau = Q_\tau z \in V_\tau$. More precisely, we get

$$\begin{aligned}
&H(\gamma_\tau + Q_\tau z) - H(s_{i,j}) \\
&= \frac{1}{2}\nabla^2 H(s_{i,j})[\gamma_\tau + Q_\tau z - s_{i,j}] + O(|\gamma_\tau + Q_\tau z - s_{i,j}|^3) \\
&= \frac{1}{2}\nabla^2 H(s_{i,j})[\gamma_\tau - \gamma_{\tau^*}] + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_\tau z] \\
&\qquad + \left\langle Q_\tau z, \nabla^2 H(s_{i,j})(\gamma_\tau - \gamma_{\tau^*})\right\rangle + O(|\gamma_\tau + Q_\tau z - \gamma_{\tau^*}|^3)
\end{aligned}$$

Now, further expanding γ_τ and Q_τ in τ leads to

$$\begin{aligned}
\gamma_\tau &= \gamma_{\tau^*} + \dot\gamma_{\tau^*}(\tau - \tau^*) + O(|\tau - \tau^*|), \quad \text{and} \\
Q_\tau z &= Q_{\tau^*} z + O(|\tau - \tau^*|\ |z|).
\end{aligned}$$

For the expansion of H, we arrive at the identity

$$\begin{aligned}
&H(\gamma_\tau + Q_\tau z) - H(s_{i,j}) = \\
&\quad \frac{1}{2}\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}(\tau - \tau^*) + O(|\tau - \tau^*|^2)] + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_{\tau^*} z + O(|\tau - \tau^*|\,|z|)] \\
&\quad + \left\langle Q_{\tau^*} z + O(|\tau - \tau^*|\,|z|), \nabla^2 H(s_{i,j}) \left(\dot\gamma_{\tau^*}(\tau - \tau^*) + O(|\tau - \tau^*|^2)\right)\right\rangle \\
&\quad + O(|\gamma_\tau + Q_\tau z - \gamma_{\tau^*}|^3) \\
&= \frac{1}{2}\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}]\ (\tau - \tau^*)^2 + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_{\tau^*} z] \\
&\quad + \left\langle Q_{\tau^*} z, \nabla^2 H(s_{i,j})\dot\gamma_{\tau^*}\right\rangle (\tau - \tau^*) \\
&\quad + O(|\tau - \tau^*|^3, |z|\,|\tau - \tau^*|^2, |z|^2\,|\tau - \tau^*|, |z|^3).
\end{aligned}$$

Using $|\tau - \tau^*| = O(\sqrt{\varepsilon}\ \omega(\varepsilon))$ and $|z| = O(\sqrt{\varepsilon}\ \omega(\varepsilon))$ we obtain for the error the estimate

$$O(|\tau - \tau^*|^3, |z|\,|\tau - \tau^*|^2, |z|^2\,|\tau - \tau^*|, |z|^3) = O(\varepsilon^{\frac{3}{2}}\ \omega^3(\varepsilon)).$$

The term $\left\langle Q_{\tau^*} z, \nabla^2 H(s_{i,j})\dot\gamma_{\tau^*}\right\rangle (\tau - \tau^*)$ in the expansion of H has no sign and has to vanish. This is only the case, if we choose $\dot\gamma_{\tau^*}$ as an eigenvector of $\nabla^2 H(s_{i,j})$ to the negative eigenvalue $\lambda^-(s_{i,j})$, because then

$$\left\langle Q_{\tau^*} z, \nabla^2 H(s_{i,j})\dot\gamma_{\tau^*}\right\rangle (\tau - \tau^*) = \lambda^-(s_{i,j}) \left\langle Q_{\tau^*} z, \dot\gamma_{\tau^*}\right\rangle = 0.$$

Additionally, by this choice of $\dot\gamma_{\tau^*}$ the quadratic form $\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}]$ evaluates to

$$\nabla^2 H(s_{i,j})[\dot\gamma_{\tau^*}] = \lambda^-(s_{i,j})|\dot\gamma_{\tau^*}|^2 = \lambda^-(s_{i,j}).$$

Therefore, we deduced the desired rewriting of $H(\gamma_\tau + Q_\tau z)$ as

$$H(\gamma_\tau + Q_\tau z) = H(s_{i,j}) - |\lambda^-(s_{i,j})|(\tau - \tau^*)^2 + \frac{1}{2}\nabla^2 H(s_{i,j})[Q_{\tau^*} z] + O(\varepsilon^{\frac{3}{2}}\ \omega^3(\varepsilon)). \tag{4.38}$$

From the regularity assumptions on the transport interpolation we can deduce that

$$\begin{aligned}\tilde{\Sigma}_\tau^{-1}[Q_\tau z] &= \tilde{\Sigma}_{\tau^*}^{-1}[Q_\tau z] + O(|\tau-\tau^*|\,|z|^2)\\ &= \tilde{\Sigma}_{\tau^*}^{-1}[Q_{\tau^*} z + O(|\tau-\tau^*|\,|z|)] + O(|\tau-\tau^*|\,|z|^2)\\ &= \tilde{\Sigma}_{\tau^*}^{-1}[Q_{\tau^*} z] + O(\varepsilon^{\frac{3}{2}}\omega^3(\varepsilon)).\end{aligned}$$

Then, it follows easily from the definition (4.23) of P_τ that

$$P_\tau \approx P_{\tau^*} \qquad \text{for } \tau \in I_T. \tag{4.39}$$

Applying the cost estimate (4.23) of Lemma 4.13, the representation (4.38) and the identity (4.39) yields the estimate for $\gamma_\tau + Q_\tau z) in \Xi_{\gamma,\Sigma}$

$$\frac{\mathcal{A}^2(\gamma_\tau + Q_\tau z)}{\mu(\gamma_\tau + Q_\tau z)} \lesssim Z_\mu e^{\frac{H(s_{i,j})}{\varepsilon}} P_{\tau^*}^2 e^{-\frac{(2\tilde{\Sigma}_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))[Q_{\tau^*} z]}{2\varepsilon} - \frac{|\lambda^-(s_{i,j})|(\tau-\tau^*)^2}{2\varepsilon}}. \tag{4.40}$$

The exponentials are densities of two Gaussian, if we put an additional constraint on the transport interpolation. Namely, we postulate

$$2\tilde{\Sigma}_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}) > 0 \quad \text{on} \quad \operatorname{span} V_{\tau^*}$$

in the sense of quadratic forms. It holds that $\operatorname{span} V_{\tau^*} = Q_{\tau^*}(\{0\} \times \mathbb{R}^{n-1}) = \operatorname{span}\{\dot{\gamma}_{\tau^*}\}^\perp$ is the tangent space of the stable manifold in the 1-saddle $s_{i,j}$. With this preliminary considerations we finally are able to estimate the right-hand side of (4.37) as follows

$$\begin{aligned}&\int_{\{0\}\times\mathbb{R}^{n-1}} \int_{I_T} \mathbb{1}_{V_\tau}(Q_\tau z)\, \frac{\mathcal{A}^2(\gamma_\tau + Q_\tau z)}{\mu(\gamma_\tau + Q_\tau z)}\, \mathrm{d}\tau\, \mathrm{d}z\\ &\overset{(4.40)}{\lesssim} Z_\mu\, e^{\frac{H(s_{i,j})}{\varepsilon}} \int_{\{0\}\times\mathbb{R}^{n-1}} \int_{I_T} P_{\tau^*}^2\, e^{-\frac{(2\tilde{\Sigma}_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))[Q_{\tau^*} z]}{2\varepsilon} - \frac{\left|\lambda^-(s_{i,j})\right|(\tau-\tau^*)^2}{2\varepsilon}}\, \mathrm{d}\tau\, \mathrm{d}z\\ &\le Z_\mu\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, \frac{\sqrt{2\pi\varepsilon}}{\sqrt{|\lambda^-(s_{i,j})|}} \int_{\{0\}\times\mathbb{R}^{n-1}} P_{\tau^*}^2\, e^{-\frac{(2\tilde{\Sigma}_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))[Q_{\tau^*} z]}{2\varepsilon}}\, \mathrm{d}z\\ &= Z_\mu\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, \frac{\sqrt{2\pi\varepsilon}}{\sqrt{|\lambda^-(s_{i,j})|}}\, P_{\tau^*}^2\, \frac{(2\pi\varepsilon)^{\frac{n-1}{2}}}{\sqrt{\det_{1,1}\left(Q_{\tau^*}^\top (2\tilde{\Sigma}_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))Q_{\tau^*}\right)}}\\ &= \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, \frac{2\pi\varepsilon}{\sqrt{|\lambda^-(s_{i,j})|}}\, \underbrace{\frac{\det_{1,1}(Q_{\tau^*}^\top \tilde{\Sigma}_{\tau^*}^{-1} Q_{\tau^*})}{\sqrt{\det_{1,1}\left(Q_{\tau^*}^\top (2\tilde{\Sigma}_{\tau^*}^{-1} - \nabla^2 H(s_{i,j}))Q_{\tau^*}\right)}}}_{\text{to optimize!}}.\end{aligned} \tag{4.41}$$

The final step consists of optimizing the choice of $\tilde{\Sigma}_{\tau^*}$. Let us use the notation $A = Q_{\tau^*}^\top \tilde{\Sigma}_{\tau^*}^{-1} Q_{\tau^*}$ and $B = Q_{\tau^*}^\top H(s_{i,j}) Q_{\tau^*}$. Then the minimization problem has the structure

$$\inf_{A\in\mathbb{R}^{n\times n}_{\mathrm{sym},+}} \left\{ \frac{\det_{1,1} A}{\sqrt{\det_{1,1}(2A - B)}} : 2A - B > 0 \quad \text{on} \quad \{0\}\times\mathbb{R}^{n-1} \right\}. \tag{4.42}$$

In the appendix, we show in Lemma D.1 that the optimal value of (4.42) is attained at $\tilde{\Sigma}_{\tau^*}^{-1} = \nabla^2 H(s_{i,j})$ restricted V_{τ^*}. The optimal value is given by

$$\frac{\det_{1,1} A}{\sqrt{\det_{1,1}(2A - B)}} = \sqrt{\det_{1,1}(Q_{\tau^*}^\top \nabla^2 H(s_{i,j}) Q_{\tau^*})}.$$

Because V_{τ^*} is the tangent space of the stable manifold of the saddle $s_{i,j}$, it holds

$$\det_{1,1}(Q_{\tau^*}^\top \nabla^2 H(s_{i,j}) Q_{\tau^*}^\top) = \frac{\det(\nabla^2 H(s_{i,j}))}{\lambda^-(s_{i,j})} = \frac{|\det(\nabla^2 H(s_{i,j}))|}{|\lambda^-(s_{i,j})|}. \tag{4.43}$$

The final step is a combination of (4.36), (4.37), (4.41) and (4.43) to obtain the desired estimate (4.34).

For the verification of Lemma 4.11, it is only left to deduce the estimate (4.18). For that purpose we analyze the error terms in the estimate (4.17) i.e.

$$\mathcal{T}_\mu^2(\nu_i,\nu_j) \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, 2\pi\varepsilon \left(\underbrace{\frac{\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|}}_{=O(1)} + \underbrace{\frac{T(C_\Sigma)^{\frac{n-1}{2}}}{\sqrt{2\pi\varepsilon}} e^{-\omega^2(\varepsilon)}}_{=O(\varepsilon^{-\frac12} e^{-\omega^2(\varepsilon)})} \right).$$

By the choice of $\omega(\varepsilon) \geq |\log\varepsilon|^{\frac12}$, enforced by Lemma 4.6, we see that

$$O(\varepsilon^{-\frac12} e^{-\omega^2(\varepsilon)}) = O(\sqrt{\varepsilon}).$$

Recalling, that "$\lesssim$" means "$\leq$" up to a multiplicative error of order $1+O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))$ we get

$$\mathcal{T}_\mu^2(\nu_i,\nu_j) \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}\, e^{\frac{H(s_{i,j})}{\varepsilon}}\, 2\pi\varepsilon \frac{\sqrt{|\det(\nabla^2 H(s_{i,j})|}}{|\lambda^-(s_{i,j})|} \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right).$$

The last inequality already yields the desired estimate (4.18) by using the observation

$$\left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right)\, \left(1 + O(\sqrt{\varepsilon})\right) = \left(1 + O(\sqrt{\varepsilon}\,\omega^3(\varepsilon))\right).$$

4.4. **Conclusion of the mean-difference estimate.** With the help of Lemma 4.6 and Lemma 4.11 the proof of Theorem 2.9 is straightforward. We can estimate the mean-differences w.r.t. to the measure μ_i by introducing the means w.r.t. the approximations ν_i and ν_j.

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 = \left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\nu_i}(f) + \mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f) + \mathbb{E}_{\nu_j}(f) - \mathbb{E}_{\mu_j}(f)\right)^2$$

We apply the Young inequality with a weight that is motivated by the final total multiplicative error term $R(\varepsilon)$ in Theorem 2.9. More precisely,

$$\begin{aligned}\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 &\leq (1+\varepsilon^{\frac12}\omega^3(\varepsilon))\left(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f)\right)^2 + \\ &\quad + 2(1+\varepsilon^{-\frac12}\omega^{-3}(\varepsilon))\left(\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\nu_i}(f)\right)^2 + \left(\mathbb{E}_{\mu_j}(f) - \mathbb{E}_{\nu_j}(f)\right)^2\right).\end{aligned}$$

Then, the estimate (4.6) of Lemma 4.6 yields

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \leq (1+\sqrt{\varepsilon}\,\omega^3(\varepsilon))\left(\mathbb{E}_{\nu_i}(f) - \mathbb{E}_{\nu_j}(f)\right)^2 + O(\varepsilon)\int |\nabla f|^2\,\mathrm{d}\mu, \tag{4.44}$$

which justifies the statement, that the approximation only leads to higher-order error terms in ε. An application of (4.1) to the estimate (4.44) transfers the mean-difference to the Dirichlet form with the help of the weighted transport distance

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \leq \left(\left(1+\sqrt{\varepsilon}\,\omega^3(\varepsilon)\right)\, \mathcal{T}_\mu^2(\nu_i,\nu_j) + O(\varepsilon)\right)\int |\nabla f|^2\,\mathrm{d}\mu,$$

The weighted transport distance $\mathcal{T}_\mu(\nu_i,\nu_j)$ is dominating the above estimate. Finally, we arrive at the estimate

$$\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\mu_j}(f)\right)^2 \lesssim \mathcal{T}_\mu^2(\nu_i,\nu_j)\int |\nabla f|^2\,\mathrm{d}\mu.$$

Now, the Theorem 2.9 follows directly from an application of the estimate (4.18) of Lemma 4.11 and setting $\omega(\varepsilon) = |\log\varepsilon|^{\frac12}$.

Appendix A. Properties of the logarithmic mean Λ

In this part of the appendix, we collect some properties of the logarithmic mean $\Lambda(\cdot,\cdot)$. Let us start with a collection of some essential properties for this section. A more complete study can be found in [Car72] and the recent review [Bha08].

Let us first recall the definition of $\Lambda(\cdot,\cdot) : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$

$$\Lambda(a,b) = \frac{a-b}{\log a - \log b}, \quad a \neq b \qquad \text{and} \qquad \Lambda(a,a) = a.$$

The value of $\Lambda(a,b)$ is also given by the logarithmic average of a and b

$$\Lambda(a,b) = \int_0^1 a^s\, b^{1-s}\, \mathrm{d}s = \frac{1}{\log a - \log b} \;\left[a^s b^{1-s}\right]_{s=0}^1 . \tag{A.1}$$

The equation (A.1) justifies the statement, that $\Lambda(\cdot,\cdot)$ is a mean, since one immediately recovers the simple bounds $\min\{a,b\} \leq \Lambda(a,b) \leq \max\{a,b\}$. Furthermore, the following representations hold for $1/\Lambda(\cdot,\cdot)$

$$\frac{1}{\Lambda(a,b)} = \int_0^1 \frac{\mathrm{d}\tau}{\tau a + (1-\tau) b} = \int_0^\infty \frac{\mathrm{d}\tau}{(a+\tau)(b+\tau)} \tag{A.2}$$

Some immediate properties are:

- $\Lambda(\cdot,\cdot)$ is symmetric
- $\Lambda(\cdot,\cdot)$ is homogeneous of degree one, i.e. for $\lambda > 0$ holds $\Lambda(\lambda a, \lambda b) = \lambda \Lambda(a,b)$.

The derivatives of $\Lambda(\cdot,\cdot)$ are given by straight-forward calculus

$$\partial_a \Lambda(a,b) = \frac{1}{\log a - \log b}\left(1 - \frac{\Lambda(a,b)}{a}\right) > 0 \qquad \text{and}$$
$$\partial_b \Lambda(a,b) = \frac{1}{\log b - \log a}\left(1 - \frac{\Lambda(a,b)}{b}\right) > 0.$$

Hence $\Lambda(\cdot,\cdot)$ is strictly monotone increasing in both arguments.

The following result is almost classical.

Lemma A.1. *The logarithmic mean can be bounded below by the geometric mean and above by the arithmetic mean*

$$\sqrt{ab} \leq \Lambda(a,b) \leq \frac{a+b}{2}, \tag{A.3}$$

with equality if and only if $a = b$.

There exists at least four proofs of the inequality A.3

- [Car72, Theorem 1] uses the representation (A.2)
- [Mie11, Appendix A] starts with (A.1) and uses the convexity of $s \mapsto a^s b^{1-s}$
- [Bha08] gives an argument by simple calculus.
- Again [Bha08] relates the terms in question to hyperbolic trigonometric functions, which allow for a quantification of the error, in the case with no equality. We will present his proof here.

Proof. Since w.l.o.g. $a, b > 0$, we can switch to exponential variables and set $a = e^x$ and $b = e^x$. Therewith we arrive for the quotient of geometric and logarithmic mean at

$$\frac{\sqrt{ab}}{\Lambda(a,b)} = e^{\frac{x+y}{2}} \frac{x-y}{e^x - e^y} = \frac{x-y}{e^{\frac{x-y}{2}} - e^{\frac{y-x}{2}}} = \frac{\frac{x-y}{2}}{\sinh\left(\frac{x-y}{2}\right)}. \tag{A.4}$$

It is easy to verify, that the function $t \mapsto \frac{t}{\sinh t}$ is symmetric and strictly decreasing in $|t|$, hence it has a unique maximum for $t = 0$ with 1. This proves $\sqrt{ab} \leq \Lambda(a,b)$ with equality only if $a = b$.

By the same reasoning, we obtain for the quotient of arithmetic and logarithmic mean in exponential variables

$$\frac{\frac{a+b}{2}}{\Lambda(a,b)} = \frac{\frac{x-y}{2}}{\tanh\left(\frac{x-y}{2}\right)}.$$

Again, one can check that the function $t \mapsto \frac{t}{\tanh t}$ is symmetric and strictly increasing in $|t|$, hence it has a unique minimum for $t = 0$ with value 1. This proves $\frac{a+b}{2} \geq \Lambda(a,b)$ with equality only if $a = b$. □

The bounds in (A.3) are good, if a is of the same order as b, whereas the following bound is particular good if $\frac{a}{b}$ becomes very small or very large.

Lemma A.2. *It holds for $p \in (0,1)$ the following bound*

$$\frac{\Lambda(p,1-p)}{p(1-p)} < \min\left\{ \frac{1}{p\log\frac{1}{p}}, \frac{1}{(1-p)\log\frac{1}{1-p}} \right\}. \tag{A.5}$$

Proof. Let us first consider the case $0 < p < \frac{1}{2}$. Then, it is enough to show, that

$$\frac{\Lambda(p,1-p)}{p(1-p)}\, p\log\frac{1}{p} = \frac{(1-2p)\log\frac{1}{p}}{(1-p)\log\frac{1-p}{p}} \overset{!}{<} 1. \tag{A.6}$$

This follows easily from the following lower bound on the denominator

$$(1-p)\log\frac{1-p}{p} = (1-2p)\log\frac{1}{p} + p\log\frac{1}{p} - (1-p)\log\frac{1}{1-p} > (1-2p)\log\frac{1}{p},$$

since $p\log\frac{1}{p} - (1-p)\log\frac{1}{1-p} > 0$ for $0 < p < \frac{1}{2}$. The case $\frac{1}{2} < p < 1$ follows by symmetry under the variable change $p \mapsto 1-p$. It remains to check the case $p = \frac{1}{2}$. The left-hand side of (A.6) evaluates for $p = \frac{1}{2}$ to

$$\lim_{p\to\frac{1}{2}} \frac{\Lambda(p,1-p)}{p(1-p)}\, p\log\frac{1}{p} = \log 2 < 1.$$

□

The logarithmic mean also occurs in the following optimization problem, which appears in the proof of the optimality of the Eyring-Kramers formula for the logarithmic Sobolev constant in one dimension (cf. Section 2.4).

Lemma A.3. *For $p \in (0,1)$ and $t \in (0,1)$ we define the function $h_p(t)$ according to*

$$h_p(t) = \frac{\left(\sqrt{\frac{t}{p}} - \sqrt{\frac{1-t}{1-p}}\right)^2}{t\log\frac{t}{p} + (1-t)\log\frac{1-t}{1-p}}.$$

Then it holds

$$\min_{t\in(0,1)} h_p(t) = \frac{\Lambda(p,1-p)}{p(1-p)}. \tag{A.7}$$

The minimum in (A.7) is attained for $t = 1-p$.

Proof. Let us introduce the function $f_p : (0,1) \to \mathbb{R}$ and $g_p : (0,1) \to \mathbb{R}$ given by the nominator and denominator of (A.7), namely

$$f_p(t) := \left(\sqrt{\frac{t}{p}} - \sqrt{\frac{1-t}{1-p}}\right)^2 \qquad \text{and} \qquad g_p(t) := t\log\frac{t}{p} + (1-t)\log\frac{1-t}{1-p}.$$

It is easy to verify, that the following relations for the derivatives hold true

$$f_p'(t) = \left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right)\left(\tfrac{1}{\sqrt{tp}} + \tfrac{1}{\sqrt{(1-p)(1-t)}}\right), \qquad g_p'(t) = \log\tfrac{t}{p} - \log\tfrac{1-t}{1-p}, \tag{A.8}$$
$$f_p''(t) = \sqrt{\tfrac{(1-t)t}{(1-p)p}}\ \tfrac{1}{2(1-t)^2t^2} > 0, \qquad g_p''(t) = \tfrac{1}{(1-t)t} > 0.$$

Hence, both functions f_p an g_p are strictly convex and have a unique minimum for $t = p$, where they are both zero. The derivative of the quotient of f_p and g_p has the form

$$h_p'(t) := \left(\frac{f_p(t)}{g_p(t)}\right)' = \frac{1}{g_p^2(t)}\left(f_p'(t)g_p(t) - f_p(t)g_p'(t)\right) \tag{A.9}$$

The representation (A.8) for g_p' leads to

$$h_p'(t)g_p^2(t) = \left(tf_p'(t) - f_p(t)\right)\log\tfrac{t}{p} + \left((1-t)f_p'(t) + f_p(t)\right)\log\tfrac{1-t}{1-p}. \tag{A.10}$$

Now, we can use (A.8) for f_p' to find

$$\begin{aligned} tf_p'(t) - f_p(t) &= \left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right)\left(\sqrt{\tfrac{t}{p}} + \tfrac{t}{\sqrt{(1-p)(1-t)}} - \sqrt{\tfrac{t}{p}} + \sqrt{\tfrac{1-t}{1-p}}\right) \\ &= \tfrac{1}{\sqrt{(1-p)(1-t)}}\left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right) \end{aligned} \tag{A.11}$$

and likewise

$$(1-t)f_p'(t) + f_p(t) = \tfrac{1}{\sqrt{tp}}\left(\sqrt{\tfrac{t}{p}} - \sqrt{\tfrac{1-t}{1-p}}\right). \tag{A.12}$$

Using (A.11) and (A.12) in (A.10) leads by (A.9) to

$$h_p'(t) = \frac{1}{g_p^2(t)}\underbrace{\left(\sqrt{\frac{t}{p}} - \sqrt{\frac{1-t}{1-p}}\right)}_{=:v_p(t)}\underbrace{\left(\frac{\log\frac{t}{p}}{\sqrt{(1-p)(1-t)}} + \frac{\log\frac{1-t}{1-p}}{\sqrt{tp}}\right)}_{=:w_p(t)}.$$

Since $g_p(p) = g_p'(p) = 0$ and $g_p''(p) > 0$, the function $\frac{1}{g_p^2(t)}$ has a pole of order 4 in $t = p$. Moreover, the function $v_p(t)$ has a simple zero in $t = p$. We have to do some more investigations for the function $w_p(t)$. First, we observe that $w_p(t)$ can be rewritten as

$$w_p(t) = \underbrace{\frac{t-p}{\sqrt{(1-t)t(1-p)p}}}_{=:\hat{w}_p(t)}\underbrace{\left(\frac{\sqrt{tp}\ \log\frac{t}{p}}{(t-p)} - \frac{\sqrt{(1-t)(1-p)}\ \log\frac{1-t}{1-p}}{(p-t)}\right)}_{:=\tilde{w}_p(t)}.$$

The function $\tilde{w}_p(t)$ can be expressed in terms of the logarithmic mean

$$\tilde{w}_p(t) = \frac{\sqrt{tp}}{\Lambda(t,p)} - \frac{\sqrt{(1-t)(1-p)}}{\Lambda(1-t,1-p)} \tag{A.13}$$

and is measuring the defect in the geometric-logarithmic mean inequality (A.3). Let us switch to exponential variables and set

$$x(t) := \log\sqrt{\frac{t}{p}} \quad \text{and} \quad y(t) := \log\sqrt{\frac{1-t}{1-p}}.$$

Note, that either $x(t) \leq 0 \leq y(t)$ for $t \leq p$ or $y(t) \leq 0 \leq x(t)$ for $t \geq 0$ with equality only for $t = p$. Therewith, (A.13) becomes with the same argument as in (A.4)

$$\tilde{w}_p(t) = \frac{x(t)}{\sinh(x(t))} - \frac{y(t)}{\sinh(y(t))}.$$

By making use of the fact, that the function $x \mapsto \frac{x}{\sinh x}$ is symmetric, monotone decreasing in $|x|$ and has a unique maximum in 1, we can conclude that

$$\tilde{w}_p(t) = 0 \qquad \text{if and only if} \qquad x(t) = -y(t).$$

The solutions to the equation $x(t) = -y(t)$ are given for $t \in \{p, 1-p\}$. Let us first consider the case $t = p$, then $x(t) = y(t) = 0$ and $w_p(p)$ is a zero of order 2, since the function $x \mapsto \frac{x}{\sinh(x)}$ is strictly concave for $t = 0$. Now, we can go back to $h_p'(t)$ and argue with the representation

$$\lim_{t\to p} h_p'(t) = \lim_{t\to p} \frac{v_p(t)\hat{w}_p(t)\tilde{w}_p(t)}{g_p^2(t)} \overset{!}{\neq} 0.$$

This is a consequence of counting the zeros for $t = p$ in the nominator and denominator according to their order. For the denominator $g_p^2(p)$ is a zero of order 4. For the nominator we have $v_p(p)$ is a zero of order 1, $\hat{w}_p(p)$ is a zero of order 1 and $\tilde{w}_p(p)$ is a zero of order 2, which leads in total again to a zero of order 4 exactly compensating the zero of the denominator.
The other case is $t = 1-p$. Let us evaluate $h_p(1-p)$, which is given by

$$\begin{aligned} h_p(1-p) &= \frac{\frac{1}{p(1-p)}\left(p-(1-p)\right)^2}{(1-p)\log\frac{1-p}{p} + p\log\frac{p}{1-p}} \\ &= \frac{1}{p(1-p)} \frac{(p-(1-p))^2}{(p-(1-p))\log\frac{p}{1-p}} = \frac{\Lambda(p,1-p)}{p(1-p)}. \end{aligned}$$

Since, $t = 1-p$ is the only critical point of $h_p(t)$ inside $(0,1)$, it remains to check whether the boundary values are larger than $h_p(1-p)$. They are given by

$$\lim_{t\to 0} h_p(t) = \frac{1}{(1-p)\log\frac{1}{1-p}} \qquad \text{and} \qquad \lim_{t\to 1} h_p(t) = \frac{1}{p\log\frac{1}{p}}.$$

We observe, that the demanded inequality to be in a global minimum

$$h_p(1-p) = \frac{\Lambda(p,1-p)}{p(1-p)} \overset{!}{<} \min\left\{\frac{1}{p\log\frac{1}{p}}, \frac{1}{(1-p)\log\frac{1}{1-p}}\right\}$$

is just (A.5) of Lemma A.2. □

Appendix B. Partial Gaussian integrals

This section is devoted to proof the representation for partial or incomplete Gaussian integrals. Lemma (B.1) is an ingredient to evaluate the weighted transport cost in Section 4.3.

Lemma B.1 (Partial Gaussian integral). *Let $\Sigma^{-1} \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$ be a symmetric positive definite matrix and let $\eta \in S^{n-1}$ be a unit vector. Therewith, $\left\{r\eta + z^\perp\right\}_{r\in\mathbb{R}}$ is for $z^\perp \in \operatorname{span}\{\eta\}^\perp$ an affine subspace of $\mathbb{R}^n$. The integral of a centered Gaussian w.r.t. to this subspace is given by*

$$\int_{\mathbb{R}} \exp\left(-\frac{1}{2}\Sigma^{-1}[r\eta + z^\perp]\right) \mathrm{d}r = \frac{\sqrt{2\pi}}{\sqrt{\Sigma^{-1}[\eta]}} \exp\left(-\tilde{\Sigma}^{-1}[z^\perp]\right),$$
$$\textit{with} \quad \tilde{\Sigma}^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1}\eta \otimes \Sigma^{-1}\eta}{\Sigma^{-1}[\eta]}.$$

Proof. To evaluate this integral on an one-dimensional subspace of $\mathbb{R}^n$, we have to expand the quadratic form $\Sigma^{-1}[r\eta + z^\perp]$ and arrive at the relation

$$\begin{aligned}
&\int_{\mathbb{R}} \exp\left(-\frac{1}{2}\Sigma^{-1}[r\eta + z^\perp]\right) \mathrm{d}r \\
&\quad = \exp\left(-\frac{1}{2}\Sigma^{-1}[z^\perp]\right) \int_{\mathbb{R}} \exp\left(-\frac{r^2}{2}\Sigma^{-1}[\eta] + r\left\langle \eta, \Sigma^{-1} z^\perp\right\rangle\right) \mathrm{d}r \\
&\quad = \exp\left(-\frac{1}{2}\Sigma^{-1}[z^\perp]\right) \frac{\sqrt{2\pi}}{\sqrt{\Sigma^{-1}[\eta]}} \exp\left(\frac{\left\langle \eta, \Sigma^{-1} z^\perp\right\rangle^2}{2\Sigma^{-1}[\eta]}\right) \\
&\quad = \frac{\sqrt{2\pi}}{\sqrt{\Sigma^{-1}[\eta]}} \exp\left(-\frac{1}{2}\left(\Sigma^{-1} - \frac{\Sigma^{-1}\eta \otimes \Sigma^{-1}\eta}{\Sigma^{-1}[\eta]}\right)[z^\perp]\right),
\end{aligned}$$

which concludes the hypothesis. □

Appendix C. Subdeterminants, adjugates and inverses

Let $A \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$, then define for $\eta \in S^{n-1}$ the matrix

$$\tilde{A} := A - \frac{A\eta \otimes A\eta}{A[\eta]}. \tag{C.1}$$

The matrix $\tilde{A}$ has at least rank $n-1$, since we subtracted from the positive definite matrix A a rank-1 matrix. Further, from the representation it is immediate, that $\tilde{A}$ has rank $n-1$ if and only if η is an eigenvector of A. In this case $\ker A = \operatorname{span} \eta$. It is easy to show that

$$\tilde{A} > 0 \quad \text{on } \operatorname{span}\{\eta\}^\perp .$$

Let $V = \operatorname{span}\{\eta\}^\perp$ be the $(n-1)$-dimensional subspace perpendicular to η. Then for a matrix $A \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$ we want to calculate the determinant of A restricted to this subspace V. This determinant is obtained by first choosing $Q \in SO^n$ such that $Q(\{0\} \times \mathbb{R}^{n-1}) = V$ and then evaluating the determinant of the minor consisting of the $(n-1)\times(n-1)$ lower right submatrix of $Q^\top AQ$ denoted by $\det_{1,1}(Q^\top AQ)$. Hence, we have

$$\det{}_{1,1}(Q^\top AQ), \qquad \text{with } Q \in SO(n) : Q^\top \eta = e^1 = (1,0,\dots,0)^\top .$$

Since $V = \operatorname{span}\{\eta\}^\perp$, it follows that the first column of Q is given by η and we can decompose $Q^\top AQ$ into

$$Q^\top AQ = \begin{pmatrix} A[\eta] & \widehat{Q^\top A\eta} \\ \widehat{Q^\top A\eta}^\top & \widehat{Q^\top AQ} \end{pmatrix},$$

where for a matrix M, $\widehat{M}$ is the lower right $(n-1)\times(n-1)$ submatrix of M and for a vector v, $\widehat{v}$ the $(n-1)$ lower subvector of v. Therewith, we find a similarity transformation which applied to $Q^\top AQ$ results in

$$\begin{aligned}
\det A = \det Q^\top AQ &= \det\left(\begin{pmatrix} A[\eta] & \widehat{Q^\top A\eta} \\ \widehat{Q^\top A\eta}^\top & \widehat{Q^\top AQ} \end{pmatrix}\begin{pmatrix} 1 & -\frac{\widehat{Q^\top A\eta}}{A[\eta]} \\ 0 & \mathrm{Id}_{n-1} \end{pmatrix}\right) \\
&= \det\begin{pmatrix} A[\eta] & 0 \\ \widehat{Q^\top A\eta}^\top & \widehat{Q^\top AQ} - \frac{\widehat{A\eta}\otimes\widehat{A\eta}}{A[\eta]} \end{pmatrix} \\
&= A[\eta] \det{}_{1,1}\left(Q^\top AQ - \frac{Q^\top A\eta \otimes Q^\top A\eta}{A[\eta]}\right).
\end{aligned}$$

The determinant of the minor is given by

$$\det_{1,1}\left(Q^\top AQ - \frac{Q^\top A\eta \otimes Q^\top A\eta}{A[\eta]}\right) = \det_{1,1}\left(Q^\top\left(A - \frac{A\eta\otimes A\eta}{A[\eta]}\right)Q\right).$$

Hence, by the definition (C.1) of $\tilde{A}$ and the subdeterminant we found the identity

$$\det A = A[\eta]\det_{1,1}(Q^\top AQ). \tag{C.2}$$

Appendix D. A matrix optimization

Lemma D.1. *Let $B \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$, then it holds*

$$\inf_{A\in\mathbb{R}^{n\times n}_{\mathrm{sym},+}}\left\{\frac{\det A}{\sqrt{\det(2A-B)}} : 2A > B\right\} = \sqrt{\det B}$$

and for the optimal A holds $A = B$.

Proof. We note that

$$\frac{\det A}{\sqrt{\det(2A-B)}} = \frac{1}{\sqrt{\det(A^{-1})\det(2\,\mathrm{Id} - A^{-\frac12}BA^{-\frac12})}}.$$

Therewith, it is enough to maximize the radical of the root. Therefore, we substitute $A^{-\frac12} = CB^{-\frac12}$ with $C > 0$ not necessarily symmetric and observe that $A^{-\frac12} = B^{-\frac12}C^\top$. We obtain

$$\det(A^{-1})\det(2\,\mathrm{Id} - A^{-\frac12}BA^{-\frac12}) = \det(B^{-1})\det(CC^\top)\det(2\,\mathrm{Id} - CC^\top).$$

Note, that $CC^\top \in \mathbb{R}^{n\times n}_{\mathrm{sym},+}$ and it is enough to calculate

$$\sup_{\tilde{C}\in\mathbb{R}^{n\times n}_{\mathrm{sym},+}}\left\{\det(\tilde{C})\det(2\,\mathrm{Id} - \tilde{C}) : \tilde{C} < 2\,\mathrm{Id}\right\}.$$

From the constraint $0 < \tilde{C} < 2\,\mathrm{Id}$ we can write $\tilde{C} = \mathrm{Id} + D$, where D is symmetric and satisfies $-\mathrm{Id} < D < \mathrm{Id}$ in the sense of quadratic forms. From here, we finally observe

$$\det(\tilde{C})\det(2\,\mathrm{Id} - \tilde{C}) = \det(\mathrm{Id} + D)\det(\mathrm{Id} - D) = \det(\mathrm{Id} - D^2).$$

Since $D^2 \geq 0$, we find the optimal $\tilde{C}$ given by Id, which yields that $A = B$. □

Appendix E. Jacobi's formula

Lemma E.1 (Jacobi's formula)**.** *Let $\mathbb{R} \ni t \mapsto \Phi_t \in \{A \in \mathbb{R}^{n\times n} : \det A \neq 0\}$ be a differentiable function, then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\log\det\Phi_t = \mathrm{tr}\left(\Phi_t^{-1}\dot{\Phi}_t\right).$$

Proof. We first note that the determinant of $\Phi(t)$ is a multilinear function d of the columns $\phi_t^1, \ldots, \phi_t^n$, i.e. $\det\Phi_t = d(\phi_t^1, \ldots, \phi_t^n)$ Then, it follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\det\Phi_t = d(\dot{\phi}_t^1, \phi_t^2, \ldots, \phi_t^n) + \cdots + d(\phi_t^1, \ldots, \phi_t^{n-1}, \dot{\phi}_t^n).$$

Now, the proof consists of two steps. We first proof the identity (E.1) for $\Phi_t = \mathrm{Id}$ and then generalize this result. If we assume w.l.o.g. that $\Phi_0 = \mathrm{Id}$. By expanding the determinant $d(\dot{\phi}_t^1, \phi_t^2, \ldots, \phi_t^n)$ along its first column it immediately follows that

$$d(\dot{\phi}_t^1, \phi_t^2, \ldots, \phi_t^n) = \dot{\phi}_t^{1,1}.$$

From here we conclude that

$$\frac{\mathrm{d}}{\mathrm{d}t}\det\Phi_t = \mathrm{tr}\,\dot{\Phi}_t.$$

Now, let $\Phi_t = A$ be a general invertible matrix. Hence, we can apply the result from the first step to $A^{-1}\Phi_t$ and arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\det\left(A^{-1}\Phi_t\right) = \operatorname{tr}\left(A^{-1}\dot{\Phi}_t\right).$$

The results follows by substituting A back. □

Appendix F. Jacobi matrices

For a smooth function $f : \mathbb{R}^n \to \mathbb{R}^n$ denotes $Df(x)$ the *Jacobi matrix* of the partial derivatives of f in $x \in \mathbb{R}^n$ given by

$$Df(x) := \left(\frac{\mathrm{d}f_i}{\mathrm{d}x_j}(x)\right)_{i,j=1}^n.$$

Lemma F.1. *Let $A, B \in \mathbb{R}^{n\times n}$, then it holds*

$$\nabla\left|Ax + f(Bx)\right| = (A + Df(x)B)^\top \frac{Ax + f(Bx)}{|Ax + f(Bx)|}, \tag{F.1}$$

$$D\frac{f(x)}{|f(x)|} = \frac{1}{|f(x)|}\left(\mathrm{Id} - \frac{f(x)}{|f(x)|} \otimes \frac{f(x)}{|f(x)|}\right) Df(x). \tag{F.2}$$

Proof. Let us first check the relation (F.1) and calculate the partial derivative

$$\frac{\mathrm{d}\left|Ax + f(Bx)\right|}{\mathrm{d}x_i} = \frac{1}{2\left|Ax + f(Bx)\right|}\sum_j \frac{\mathrm{d}}{\mathrm{d}x_i}\left(\sum_k A_{jk}x_k + f_j(Bx)\right)^2 \tag{F.3}$$

The inner derivative of (F.3) evaluates to

$$\frac{\mathrm{d}}{\mathrm{d}x_i}\left(\sum_k A_{jk}x_k + f_j(Bx)\right)^2 = 2\left(\sum_k A_{jk}x_k + f_j(Bx)\right)\left(A_{ji} + \frac{\mathrm{d}f_j(Bx)}{\mathrm{d}x_i}\right). \tag{F.4}$$

The derivative of $f_j(Bx)$ becomes

$$\begin{aligned}\frac{\mathrm{d}f_j(Bx)}{\mathrm{d}x_i} &= \frac{\mathrm{d}f_j\left(\sum_k B_{1k}x_k, \dots, \sum_k B_{nk}x_k\right)}{\mathrm{d}x_i} \\ &= \sum_{k=1}^n \partial_k f_j(Bx) B_{ki} = (Df(Bx)B)_{ji}.\end{aligned} \tag{F.5}$$

Hence, a combination of (F.3), (F.4) and (F.5) leads to

$$\begin{aligned}\frac{\mathrm{d}\left|Ax + f(Bx)\right|}{\mathrm{d}x_i} &= \frac{1}{|Ax + f(Bx)|}\sum_j \left((Ax)_j + f_j(Bx)\right)\left(A_{ji}(Df(Bx)B)_{ji}\right) \\ &= \sum_j (A + Df(Bx)B)^\top_{ij}\frac{(Ax + f(Bx))_j}{|Ax + f(Bx)|},\end{aligned}$$

which shows (F.1). For the equation (F.2), let us first consider the Jacobian of the function $F(x) = \frac{x}{|x|}$, which is given by

$$DF(x) = \frac{1}{|x|}\left(\mathrm{Id} - \frac{x}{|x|} \otimes \frac{x}{|x|}\right).$$

Then, by the chain rule, we observe that

$$D\frac{f(x)}{|f(x)|} = D(F \circ f)(x) = DF(f(x))Df(x),$$

which is just (F.2). □

Appendix G. Some more functional inequalities

We already introduced in Section 1.1 the both functional inequalities PI(ϱ) and LSI(α). The inequalities PI(ϱ) and LSI(α) can be thought as the extremes of a whole family of inequalities, from which we want introduce at least two more in this short section.

G.1. **Horizontal, vertical and infinitesimal distances.**

Let us introduce a new functional inequality incorporating the Wasserstein transportation distance. The interplay of the different functional inequalities was discovered by Otto and Villani [OV00].

Definition G.1 (Wasserstein distance)**.** For any two probability measures μ, ν on an Euclidean space X, the *Wasserstein distance* of between μ and ν is defined by the formula

$$W_2^2(\nu,\mu) = \inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2\,\pi(\mathrm{d}x,\mathrm{d}y),$$

where $\Pi(\nu,\mu)$ is the set of all *couplings*, i.e. all measures π on $\mathbb{R}^n\times\mathbb{R}^n s$ with first marginal ν and second marginal μ, i.e. $\int_{\mathbb{R}^n}\pi(\cdot,\mathrm{d}y)=\nu(\cdot)$ and $\int_{\mathbb{R}^n}\pi(\mathrm{d}x,\cdot)=\mu(\cdot)$.

Since the Wasserstein distance measures the displacement between two measure, it can be thought as a *horizontal* distance[1] on the space of probability measures. On the contrary, classical distances like the total variation, variance or relative entropy are *vertical* distances, since they measure the pointwise difference of the densities between two measures. Often, one is interested in the interplay between a horizontal and vertical distances and how a distance of the one kind can be bounded by a distance of the other kind. The following theorem provides a simple and in general rough bound of the Wasserstein distance between two measures in terms of the second moment of the total variation of the difference of the two measures.

Theorem G.2 (Control by total variation [Vil09, Theorem 6.15])**.** *Let μ and ν be two probability measures on an Euclidean space X, then*

$$W_2^2(\nu,\mu) \leq 2\int |x|^2\;|\nu-\mu|(\mathrm{d}x) = 2\left\| |\cdot|^2\,(\nu-\mu)\right\|_{TV}.$$

More difficult is the question, whether a horizontal distance can be estimated by an *infinitesimal* distance, like the Dirichlet energy or Fisher information, which somehow measure the local relative fluctuations between two measures. The prototype and extensively studied inequality of this type is the *transportation-information* inequality.

Definition G.3 (Transportation-information inequality WI)**.** A probability measure μ on an Euclidean space X satisfies WI(ρ) with constant $\rho>0$, if for all test functions $f>0$ with $\int f\mathrm{d}\mu = 1$ holds

$$W_2^2(f\mu,\mu) \leq \frac{1}{\rho^2}\int \frac{|\nabla f|^2}{f}\,\mathrm{d}\mu. \qquad \mathrm{WI}(\rho)$$

In the abbreviation WI, W stands for the Wasserstein distance and I stands for the Fisher information.

It turns out, that the WI inequality is just in-between the PI and LSI.

Lemma G.4 (Relation between LSI(ρ), WI(ρ) and PI(ρ))**.** *Let μ be a probability measure on an Euclidean space X. Then the following implications hold*

$$\mu \textit{ satisfies } \mathrm{LSI}(\rho) \quad\Rightarrow\quad \mu \textit{ satisfies } \mathrm{WI}(\rho) \quad\Rightarrow\quad \mu \textit{ satisfies } \mathrm{PI}(\rho),$$

[1]The notion of horizontal and vertical distances is adopted from a talk of Nicola Gigli on the recent preprint [AGS12]

where all of the implications are strict.

Remark G.5. The first implication in WI(ρ) is on of the result in [OV00]. An example satisfying WI(ρ) but not LSI(ρ) was constructed in [CG06]. For the second implication, one uses a linearization argument, like we already presented in Remark 1.2. To proof that the implication is sharp, consider the measure $\mu(\mathrm{d}x) = Z^{-1}\exp(-|x|)\,\mathrm{d}x$ on the real line. Then, the condition [Goz07, Theorem 6] states that μ does not satisfy WI(ρ), but for instance by the Muckenhoupt functional [Muc72] one can check (cf. [Sch12, Section 5.3.]), that μ satisfies PI(ρ).

PI as well as LSI are also in this class and bound a vertical distance, i.e. the variance respectively the relative entropy, by an infinitesimal distance, i.e. the Dirichlet form respectively the Fisher information. An inequality showing the interplay between all three kinds of distances, i.e. vertical, horizontal and infinitesimal, was discovered by Otto and Villani [OV00]. The name HWI-inequality comes from the quantities in question, since the inequality bounds the relative entropy H in terms of the Wasserstein distance W and the Fisher information I.

Theorem G.6 (HWI inequality [OV00, Theorem 3])**.** *Let $\mu(\mathrm{d}x) = e^{-H(x)}\mathrm{d}x$ a probability measure on $\mathbb{R}^n$, with finite moments of order 2, such that $H \in C^2(\mathbb{R}^n)$, $\nabla^2 H \geq K_H$, $K_H \in \mathbb{R}$ (not necessarily positive). Then, for all test functions f with $\int f \mathrm{d}\mu = 1$ holds*

$$\mathrm{Ent}_\mu(f) = H(f\mu|\mu) \leq W_2(f\mu,\mu)\sqrt{2I(f\mu|\mu)} - \frac{K_H}{2} W_2^2(f\mu,\mu). \qquad \text{HWI}$$

Remark G.7 (A covariance estimate in terms of $\mathcal{T}_\mu$). A special case of the "mean-difference" estimate occurs, by setting $\nu_0 = g\mu$ and $\nu_1 = \mu$, where $g \geq 0$ and $\int g \,\mathrm{d}\mu = 1$, then we arrive at the following covariance estimate

$$\mathrm{cov}_\mu^2(f,g) = (\mathbb{E}_{g\mu}(f) - \mathbb{E}_\mu(f))^2 \leq \mathcal{T}_\mu^2(g\mu,\mu) \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Finally, setting $f = g$ results in

$$\mathrm{var}_\mu(f) \leq \mathcal{T}_\mu(f\mu,\mu)\sqrt{\int |\nabla f|^2 \,\mathrm{d}\mu}. \tag{G.1}$$

The estimate (G.1) has the same structure as the HWI inequality in the sense, that it connects a vertical with the product of a horizontal and the square root of an infinitesimal distance. However, the estimate (G.1) does not demand a lower bound on the Hessian of the exponential density of μ.

G.2. **Defective logarithmic Sobolev inequality.** Let us present how a *defective* LSI can be tightened to LSI with the help of PI.

Definition G.8 (Defective logarithmic Sobolev inequality dLSI(α_d, B))**.** A measure μ on $\mathbb{R}^n$ satisfies the *defective logarithmic Sobolev inequality* dLSI(α_d, B) with constants $\alpha_d, B > 0$, if for all test function $f : \mathbb{R}^n \to \mathbb{R}^+$ holds

$$\mathrm{Ent}_\mu(f) \leq \frac{1}{\alpha_d}\int \frac{|\nabla f|^2}{2f}\mathrm{d}\mu + B\int f \mathrm{d}\mu. \qquad \mathrm{dLSI}(\alpha_d, B)$$

Proposition G.9 (dLSI(α_d, B) and PI(ϱ) imply LSI(α))**.** *Assume that μ satisfies* dLSI(α_d, B) *and* PI(ϱ)*, then μ satisfies* LSI(α) *with*

$$\frac{1}{\alpha} = \frac{1}{\alpha_d} + \frac{B+2}{\varrho}.$$

Proof. The argument is from [Led99] and is a simple consequence of the estimate

$$\operatorname{Ent}_\mu(f^2) \le \operatorname{Ent}_\mu\left((f - \mathbb{E}_\mu(f))^2\right) + 2\operatorname{var}_\mu(f),$$

which is due to [Rot86] and [DS89]. An application of dLSI(α_d, B) leads to

$$\operatorname{Ent}_\mu(f^2) \le \frac{1}{\alpha_d}\int 2\left|\nabla f\right|^2 \mathrm{d}\mu + (B+2)\operatorname{var}_\mu(f).$$

The result follow from applying PI(ϱ) to the variance in the last term. □

Acknowledgment

This work is part of the Ph.D. thesis of the second author, written under the supervision of Stephan Luckhaus at the University of Leipzig.
The authors are greatly indebted to Felix Otto for drawing the authors' attention to the low temperature regime and several helpful discussions.
They wish to thank the Max-Planck-Institute for Mathematics in the Sciences in Leipzig where the paper was written under great working conditions. The second author thanks the DFG project FOR718, the IMPRS of the Max-Planck-Institute and the University of Leipzig for financial support.

References

[Arr89] S. Arrhenius. On the reaction velocity of the inversion of cane sugar by acids. *Zeitschrift für physikalische Chemie*, 4:226ff, 1889.

[BB00] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[BBCG08] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13:60–66, 2008.

[BCG08] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008.

[BÉ85] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[BEGK04] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.

[Ber11] N. Berglund. Kramers' law: Validity, derivations and generalisations. *arXiv:1106.5799*, 2011(x):24, June 2011.

[BG99] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.

[BGK05] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc. (JEMS)*, 7(1):69–99, 2005.

[Bha08] R. Bhatia. The logarithmic mean. *Resonance*, 13(6):583–594, September 2008.

[Car72] B. C. Carlson. The logarithmic mean. *Amer. Math. Monthly*, 79:615–618, 1972.

[CG06] P. Cattiaux and A. Guillin. On quadratic transportation cost inequalities. *J. Math. Pures Appl. (9)*, 86(4):341–361, 2006.

[CGW10] P. Cattiaux, A. Guillin, and L.-M. Wu. A note on Talagrand's transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148(1-2):285–304, 2010.

[CGWW09] P. Cattiaux, A. Guillin, F.-Y. Wang, and L. Wu. Lyapunov conditions for super Poincaré inequalities. *J. Funct. Anal.*, 256(6):1821–1841, 2009.

[CM10] D. Chafaï and F. Malrieu. On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):72–96, 2010.

[DNS09] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.

[DS89] J.-D. Deuschel and D. W. Stroock. *Large deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1989.

[DSC96] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 1996.

[DV76] M. D. Donsker and S. R. S. Varadhan. On the principal eigenvalue of second-order elliptic differential operators. *Comm. Pure Appl. Math.*, 29(6):595–621, 1976.

[Eyr35] H. Eyring. The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107, 1935.

[FW98] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.

[GMSvdM02] O. Gonzalez, J. H. Maddocks, F. Schuricht, and H. von der Mosel. Global curvature and self-contact of nonlinearly elastic curves and rods. *Calc. Var. Partial Differential Equations*, 14(1):29–68, 2002.

[AGS12] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *arXiv:1109.0222*, pages 1–60, 2012.

[GOVW09] N. Grunewald, F. Otto, C. Villani, and M. G. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(2):302–351, 2009.

[Goz07] N. Gozlan. Characterization of Talagrand's like transportation-cost inequalities on the real line. *J. Funct. Anal.*, 250(2):400–425, 2007.

[Gro75] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[HKN04] B. Helffer, M. Klein, and F. Nier. Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach. *Mat. Contemp.*, 26:41–85, 2004.

[HN05] B. Helffer and F. Nier. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.

[HN06] B. Helffer and F. Nier. Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach: the case with boundary. *Mém. Soc. Math. Fr. (N.S.)*, (105):vi+89, 2006.

[HS87] R. Holley and D. Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Statist. Phys.*, 46(5-6):1159–1194, 1987.

[HY95] Y. Higuchi and N. Yoshida. Analytic conditions and phase transition for ising models., 1995.

[JSTV04] M. Jerrum, J.-B. Son, P. Tetali, and E. Vigoda. Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *Ann. Appl. Probab.*, 14(4):1741–1765, 2004.

[Kra40] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.

[Led99] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 120–216. Springer, Berlin, 1999.

[Led01] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 167–194. Springer, Berlin, 2001.

[Lel09] T. Lelièvre. A general two-scale criteria for logarithmic Sobolev inequalities. *J. Funct. Anal.*, 256(7):2211–2221, 2009.

[LL10] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *Nonlinearity*, 23(9):2131–2163, 2010.

[Mie11] A. Mielke. Geodesic convexity of the relative entropy in reversible markov chains. *WIAS Preprint*, (1650):1–31, 2011.

[Muc72] B. Muckenhoupt. Hardy's inequality with weights. *Studia Math.*, 44:31–38, 1972. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, I.

[Øks98] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, fifth edition, 1998. An introduction with applications.

[Olv97] F. W. J. Olver. *Asymptotics and special functions*. AKP Classics. A K Peters Ltd., Wellesley, MA, 1997. Reprint of the 1974 original [Academic Press, New York; MR0435697 (55 #8655)].

[OR07] F. Otto and M. G. Reznikoff. A new criterion for the logarithmic Sobolev inequality and two applications. *J. Funct. Anal.*, 243(1):121–157, 2007.

[OV00] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[Rot78] O. S. Rothaus. Lower bounds for eigenvalues of regular Sturm-Liouville operators and the logarithmic Sobolev inequality. *Duke Math. J.*, 45(2):351–362, 1978.

[Rot86] O. S. Rothaus. Hypercontractivity and the Bakry-Emery criterion for compact Lie groups. *J. Funct. Anal.*, 65(3):358–367, 1986.

[Roy07] G. Royer. *An initiation to logarithmic Sobolev inequalities*, volume 14 of *SMF/AMS Texts and Monographs*. American Mathematical Society, Providence, RI, 2007. Translated from the 1999 French original by Donald Babbitt.

[Sch10] Z. Schuss. *Theory and applications of stochastic processes*, volume 170 of *Applied Mathematical Sciences*. Springer, New York, 2010. An analytical approach.

[Sch12] A. Schlichting. *The Eyring-Kramers formula for Poincaré and logarithmic Sobolev inequalities*. PhD thesis, University of Leipzig, 2012.

[Sug95] M. Sugiura. Metastable behaviors of diffusion processes with small parameter. *J. Math. Soc. Japan*, 47(4):755–788, 1995.

[Vil09] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

Georg Menz, Max Planck Institute for Mathematics in the Sciences, Leipzig

André Schlichting, University of Bonn
E-mail address: Georg.Menz@mis.mpg.de, Schlichting@iam.uni-bonn.de