Electronic Journal of Probability

Electron. J. Probab. **20** (2015), no. 92, 1–39.
ISSN: 1083-6489 DOI: 10.1214/EJP.v20-4012

The Landau equation for Maxwellian molecules and the Brownian motion on $\mathrm{SO}_N(\mathbb{R})$

François Delarue* Stéphane Menozzi† Eulalia Nualart‡

Abstract

In this paper we prove that the spatially homogeneous Landau equation for Maxwellian molecules can be represented through the product of two elementary stochastic processes. The first one is the Brownian motion on the group of rotations. The second one is, conditionally on the first one, a Gaussian process. Using this representation, we establish sharp multi-scale upper and lower bounds for the transition density of the Landau equation, the multi-scale structure depending on the shape of the support of the initial condition.

Keywords: Landau equation for Maxwellian molecules; Stochastic analysis; Heat kernel estimates on groups; Large deviations.
AMS MSC 2010: 60H30; 60H40; 60H10.
Submitted to EJP on December 20, 2014, final version accepted on September 9, 2015.

1 Statement of the problem and existing results

The spatially homogeneous Landau equation for Maxwellian molecules is a common model in plasma physics. It can be obtained as a certain limit of the spatially homogeneous Boltzmann equation for N dimensional particles subject to pairwise interaction, when the collisions become grazing and when the interaction forces between particles at distance r are of order $1/r^{2N+1}$ (see Villani [24] and Guérin [15]).

The Landau equation reads as a nonlocal Fokker-Planck equation. Given an initial condition $(f(0,v), v \in \mathbb{R}^N)$, the solution is denoted by $(f(t,v), t \geq 0, v \in \mathbb{R}^N)$, $N \geq 2$, and satisfies

$$\partial_t f(t,v) = Lf(t,v), \tag{1.1}$$

where

$$Lf(t,v) = \frac{1}{2}\nabla \cdot \int_{\mathbb{R}^N} dv_*\, a(v-v_*)\left(f(t,v_*)\nabla f(t,v) - f(t,v)\nabla f(t,v_*)\right). \tag{1.2}$$

*Laboratoire J.A.Dieudonné, UMR 7351, Université Nice Sophia-Antipolis, France.
E-mail: delarue@unice.fr
†Laboratoire de Modélisation Mathématique d'Evry, UMR 8071, Université d'Evry Val d'Essonne, and Laboratory of Stochastic Analysis, Higher School of Economics, Moscow, Russian Federation.
E-mail: stephane.menozzi@univ-evry.fr,SMenozzi@hse.ru
‡Department of Economics and Business, Universitat Pompeu Fabra and Barcelona Graduate School of Economics, Spain. E-mail: eulalia@nualart.es, http://nualart.es

Here, a is an $N \times N$ nonnegative and symmetric matrix that depends on the collisions between binary particles. In the case of Maxwellian molecules, it is given by (up to a multiplicative constant)

$$a(v) = |v|^2 \mathrm{Id}_N - v \otimes v,$$

where Id_N denotes the identity matrix of size N, and $v \otimes v = vv^\top$, $v^\top$ denoting the transpose of v, v being seen as a column vector in $\mathbb{R}^N$. The unknown function $f(t,v)$ represents the density of particles of velocity $v \in \mathbb{R}^N$ at time $t \geq 0$ in a gas. It is assumed to be independent of the position of the particles (spatially homogeneous case).

The density $f(t,v)$ being given, the nonlocal operator L can be seen as a standard linear Fokker-Planck operator, with diffusion matrix $\overline{a}(t,v) = \int_{\mathbb{R}^N} a(v-v_*) f(t,v_*) dv_*$ and with drift $\overline{b}(t,v) = -(N-1) \int_{\mathbb{R}^N} (v-v_*) f(t,v_*) dv_*$. Such a reformulation permits to approach the Landau equation by means of the numerous tools that have been developed for linear diffusion operators. As a key fact in that direction, the diffusion matrix $\overline{a}$ can be shown to be uniformly elliptic for a wide class of initial conditions. This suggests that the solution $f(t,v)$ must share some of the generic properties of non-degenerate diffusion operators.

Such a remark is the starting point of the analysis initiated by Villani in [25, Proposition 4]. Therein, it is proved that, whenever the initial condition $f(0,v)$ is nonnegative and has finite mass and energy, the Landau PDE (1.1) admits a unique solution, which is bounded and $\mathcal{C}^\infty(\mathbb{R}^N)$ in positive time. Moreover, [25, Proposition 9] ensures that the solution satisfies the lower Gaussian bound

$$f(t,v) \geq C_t e^{-\delta_t \frac{|v|^2}{2}}, \ t > 0, \ v \in \mathbb{R}^N, \tag{1.3}$$

for some $C_t > 0$ and $\delta_t > 0$. The values of the constants C_t and δ_t are specified in Desvillettes and Villani [5, Theorem 9(ii)] in the case of hard potentials and when $N = 3$, under the additional condition that $f(0,v)$ has finite entropy, lies in some weighted L^2 space and is bounded from below by a strictly positive constant on a given ball. In this case, the lower bound (1.3) is established with $C_t = 1$ and $\delta_t = b_0 t + c_0/t$. This proves that, in finite time, the rate of propagation of the mass to the infinity is at least the same as for the heat equation. The key argument in [5] is to prove that the spectrum of $\bar{a}(t,v)$ is uniformly far away from zero, so that the mass can be indeed diffused to the whole space.

Anyhow, even if the lower bound (1.3) fits the off-diagonal decay of the heat kernel, it is worth mentioning that $\bar{a}(t,v)$ does not enter the required framework for applying two-sided Aronson's estimates for diffusion operators, see [1]. Indeed, the upper eigenvalue of $\bar{a}(t,v)$ can be shown to behave as $|v|^2$ when $|v|$ is large. The matrix $\bar{a}(t,v)$ thus exhibits several scales when $|v|$ tends to the infinity, which is the basic observation for motivating our analysis. Actually, a simple inspection will show that, for the same type of initial conditions as above, the quadratic form associated with $\bar{a}(t,v)$ has two regimes when $|v|$ is large. Along unitary vectors parallel to v, the quadratic form takes values of order 1. Along unitary vectors orthogonal to v, it takes values of order $|v|^2$. This suggests that the mass is spread out at a standard diffusive rate along *radial* directions, but at a much quicker rate along *tangential* directions. One of the main objective of the paper is to quantify this phenomenon precisely and to specify how it affects the lower bound (1.3), especially for highly anisotropic initial conditions. We also intend to discuss the sharpness of the bound by investigating the corresponding upper bound.

The strategy we have in mind is probabilistic. The starting point consists in deriving a probabilistic interpretation of the nonlinear operator L by means of a stochastic diffusion process $(X_t)_{t\geq 0}$ interacting with its own distribution, in the spirit of McKean to handle Vlasov type equations (see Sznitman [22]). Actually, McKean-Vlasov representations of the Landau equation were already investigated in earlier works by Funaki [9, 10, 11, 12] and more recently by Guérin [13, 14]. Part of the analysis developed in these series of papers is based on a very useful trick for representing the square root of the matrix $\overline{a}$, the square root of the diffusion matrix playing a key role in the dynamics of the stochastic process involved in the representation. In short, the key point therein is to enlarge the underlying probability space in order to identify the diffusive term with the stochastic integral of the root of a (and not the root of $\overline{a}$) with respect to a two-parameter white noise process. Basing the representation on the root of a makes it more tractable since $a(v)$ has a very simple geometric interpretation in terms of the orthogonal projection on the orthogonal $v^\perp$ of v. In this paper, we go one step forward into the explicitness of the representation. As a new feature, we show that the representation used by Funaki and Guérin can be linearized so that the stochastic process $(X_t)_{t\geq 0}$ solving the *enlarged* McKean-Vlasov equation reads as the product of two *auxiliary* basic processes:

$$X_t = Z_t\Gamma_t, \quad t \geq 0.$$

The first one is a (right) Brownian motion $(Z_t)_{t\geq 0}$ on the special group of rotations $\mathrm{SO}_N(\mathbb{R})$. The second one is, conditionally on $(Z_t)_{t\geq 0}$, a Gaussian process in $\mathbb{R}^N$ with a local covariance matrix given, at any time $t \geq 0$, by the second order moments of the density $f(t,v)$. Such a decomposition enlightens explicitly the coexistence of two scales in the dynamics of the Landau equation. It is indeed well seen that the Brownian motion on $\mathrm{SO}_N(\mathbb{R})$ cannot play any role in the diffusion of the mass along radial directions. Therefore, along such directions, only $(\Gamma_t)_{t\geq 0}$ can have an impact. Its covariance matrix can be proved to be uniformly non-degenerate for a wide class of initial conditions, explaining why, in such cases, the mass is transported along radial directions according to the standard heat propagation. The picture is different along tangential directions since, in addition to the fluctuations of $(\Gamma_t)_{t\geq 0}$, the process $(X_t)_{t\geq 0}$ also feels the fluctuations of the Brownian motion $(Z_t)_{t\geq 0}$ on $\mathrm{SO}_N(\mathbb{R})$. The effect of $(Z_t)_{t\geq 0}$ is all the more visible when the process $(X_t)_{t\geq 0}$ is far away from the origin: Because of the product form of the representation, the fluctuations in the dynamics of $(Z_t)_{t\geq 0}$ translate into multiplied fluctuations in the dynamics of $(X_t)_{t\geq 0}$ when $(X_t)_{t\geq 0}$ is of large size.

Our main result in that direction is Theorem 2.8, in which we provide two sided Gaussian bounds for the transition kernel of the process $(X_t)_{t\geq 0}$ when the initial condition X_0 is a centered random variable with a support not included in a line and finite variance. We then make appear the coexistence of two regimes in the transition density by splitting the off-diagonal decay of the density into a *radial cost* and a *tangential cost*. We explicitly show that the variance of the tangential cost increases at a quadratic rate when the starting point in the transition density tends to the infinity. The resulting bounds are sharp, which proves that our approach captures the behavior of the process in a correct way. The proof follows from our factorization of the process $(X_t)_{t\geq 0}$: Conditionally on the Brownian motion on $\mathrm{SO}_N(\mathbb{R})$, $(X_t)_{t\geq 0}$ is a Gaussian process with an explicit transition kernel. This gives a conditional representation of the transition density of $(X_t)_{t\geq 0}$ and this permits to reduce part of the work to the analysis of the heat kernel on the group $\mathrm{SO}_N(\mathbb{R})$. As a by-product, this offers an alternative to a more systematic probabilistic method based on the Malliavin calculus, as considered for instance in Guérin, Méléard and Nualart [16].

The conditional representation of the transition density of the process $(X_t)_{t\geq 0}$ also permits to consider the so-called *degenerate* case when the initial condition lies in a line. In that case, another inspection will show that the diffusion matrix $\bar{a}(t,v)$ degenerates as t tends to 0, the associated quadratic form converging to 0 with t along the direction of the initial condition. Obviously, this adds another difficulty to the picture given above: Because of the degeneracy of the matrix $\bar{a}$, the mass cannot be transported along radial directions as in standard heat propagation. In that framework, our representation provides a quite explicit description of the degeneracy rate of the system in small time. Indeed, conditionally on the realization of the Brownian motion $(Z_t)_{t\geq 0}$ on $\mathrm{SO}_N(\mathbb{R})$, the degeneracy is determined by the covariance matrix of the process $(\Gamma_t)_{t\geq 0}$, the form of which is, contrary to the non-degenerate case, highly sensitive to the realization of $(Z_t)_{t\geq 0}$. The crux is thus that, in the degenerate regime, the Brownian motion on the group of rotations also participates in the formation of the radial cost. Although quite exciting, this makes things rather intricate. In that direction, the thrust of our approach is to prove that large deviations of the process $(Z_t)_{t\geq 0}$ play an essential role in the shape of the off-diagonal decay of the transition density. Precisely, because of that large deviations, we can show that, when the initial condition of the transition is restricted to compact sets, the off-diagonal decay of the transition density is not Gaussian but is a mixture of an exponential and a Gaussian regimes, see Theorem 2.12.

Besides the density estimates, we feel that our representation of the solution raises several questions and could serve as a basis for further investigations. Obviously, the first one concerns possible extensions to more general cases, when the coefficients include a hard or soft potential (so that molecules are no more Maxwellian) or when the solution of the Landau equation also depends on the position of the particle (and not only on its velocity). In the same spirit, we could also wonder about a possible adaptation of this approach to the Boltzmann equation itself. Finally, the representation might be also useful to compute the solution numerically, providing a new angle to tackle with the particle approach developed by Fontbona *et al* [6] and Carrapatoso [3] or Fournier [7]. We leave all these questions to further prospects.

The paper is organized as follows. Main results are detailed in Section 2. In Section 3, we give some preliminary estimates concerning the Brownian motion on $\mathrm{SO}_N(\mathbb{R})$. Section 4 is devoted to the analysis of the non-degenerate case and Section 5 to the degenerate case.

In all the paper, we will use the notation $\langle \cdot, \cdot \rangle$ for the Euclidean scalar product in $\mathbb{R}^N$.

2 Strategy and Main Results

2.1 Representation of the Landau equation

The representation used in [9, 10, 11, 12, 13, 14] is based on a probabilistic set-up, which consists of

1. a complete probability space $(\Omega, \mathcal{F}, \mathrm{P})$, endowed with an N-dimensional space-time white noise $W = (W^1, ..., W^N)$ with independent entries, each of them with covariance measure $dsd\alpha$ on $\mathbb{R}_+ \times [0,1]$, where $d\alpha$ denotes the Lebesgue measure on $[0,1]$;
2. a random vector X_0 with values in $\mathbb{R}^N$ and finite second moment, independent of W, the augmented filtration generated by W and X_0 being denoted by $(\mathcal{F}_t)_{t\geq 0}$;
3. the auxiliary probability space $([0,1], \mathcal{B}([0,1]), d\alpha)$;
4. the symbols E, E_α for denoting the expectations and the symbols $\mathcal{L}, \mathcal{L}_\alpha$ for denoting the distributions of a random variable on $(\Omega, \mathcal{F}, \mathrm{P})$, $([0,1], \mathcal{B}([0,1]), d\alpha)$, respectively.

A couple of processes $(\mathcal{X}, Y)$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P}) \otimes ([0,1], \mathcal{B}([0,1]), d\alpha)$ is said to be a solution of the Landau SDE if $\mathcal{L}(\mathcal{X}) = \mathcal{L}_\alpha(Y)$, and for all $t \geq 0$, the following equation holds

$$\mathcal{X}_t = X_0 + \int_0^t \int_0^1 \sigma(\mathcal{X}_s - Y_s(\alpha)) W(ds, d\alpha) - (N-1) \int_0^t \int_0^1 (\mathcal{X}_s - Y_s(\alpha)) d\alpha ds, \qquad (2.1)$$

where σ is an $N\times N$ matrix such that $\sigma\sigma^\top = a$, the symbol $\top$ standing from now on for the transposition. The connection with (1.1) can be derived by computing $\mathrm{E}[\varphi(\mathcal{X}_t)]$, for a smooth function $\varphi:\mathbb{R}^N\to\mathbb{R}$ with a compact support. Indeed, assuming for a while that (2.1) has a solution, Itô's formula yields

$$\begin{aligned}\mathrm{E}\big[\varphi(\mathcal{X}_t)\big] = \mathrm{E}\big[\varphi(\mathcal{X}_0)\big] - (N-1)\mathrm{E}\Big\{\int_0^t \Big\langle \nabla\varphi(\mathcal{X}_s), \int_0^1 (\mathcal{X}_s - Y_s(\alpha))d\alpha\Big\rangle ds\Big\}\\ + \frac{1}{2}\mathrm{E}\Big\{\int_0^t \mathrm{Trace}\Big[\nabla^2\varphi(\mathcal{X}_s)\Big(\int_0^1 a(\mathcal{X}_s - Y_s(\alpha))d\alpha\Big)\Big]ds\Big\},\end{aligned}$$

where we used the relationship $a=\sigma\sigma^\top$. Since the law of $[0,1]\ni\alpha\mapsto Y_s(\alpha)\in\mathbb{R}^N$ fits $\mathcal{L}(\mathcal{X}_s)$, we deduce

$$\mathrm{E}\big[\varphi(\mathcal{X}_t)\big] = \mathrm{E}\big[\varphi(\mathcal{X}_0)\big] + \mathrm{E}\int_0^t L_{[\mathcal{L}(\mathcal{X}_s)]}\varphi(\mathcal{X}_s)ds, \tag{2.2}$$

where, for a probability measure μ on $\mathbb{R}^N$ such that $\int_{\mathbb{R}^N}|v|^2\mu(dv)<\infty$, we have denoted

$$\begin{aligned}\Big(L_{[\mu]}\varphi\Big)(v) = -(N-1)\Big\langle\nabla\varphi(v), \int_{\mathbb{R}^N}(v-v_*)\mu(dv_*)\Big\rangle\\ + \frac{1}{2}\mathrm{Trace}\Big[\nabla^2\varphi(v)\Big(\int_{\mathbb{R}^N} a(v-v_*)\mu(dv_*)\Big)\Big].\end{aligned}$$

Observe that (2.2) also rewrites:

$$\int_{\mathbb{R}^N}\varphi(v)\mathcal{L}(\mathcal{X}_t)(dv) = \int_{\mathbb{R}^N}\varphi(v)\mathcal{L}(\mathcal{X}_0)(dv) + \int_0^t\int_{\mathbb{R}^N} L_{[\mathcal{L}(\mathcal{X}_s)]}\varphi(v)\mathcal{L}(\mathcal{X}_s)(dv)ds, \tag{2.3}$$

thus establishing that for a given solution to (2.1), its law is a measure solution of the Landau equation $\partial_t\mu_t = L^*_{[\mu_t]}\mu_t$. Provided that X_0 has finite second moments, it has been established in [14] that (2.1) has a solution which is unique in law. If additionally, X_0 is not a Dirac mass and $(\mathcal{X}_s)_{s\geq 0}$ is the (unique in law) solution of (2.1), it is proved in [13] that, for any $s>0$, $\mathcal{L}(\mathcal{X}_s)$ is absolutely continuous with a smooth density $f(s,.)$ with respect to the Lebesgue measure, and f is as well smooth in time. Denoting with a slight abuse of notation by $L_{[f(s,.)]}$ the above generator for the measure $\mathcal{L}(\mathcal{X}_s)$ and observing that $\int_{\mathbb{R}^N} L_{[f(s,.)]}\varphi(v)f(s,v)dv = \int_{\mathbb{R}^N} Lf(s,v)\varphi(v)dv$ (with L as in (1.2)), one indeed deduces from (2.3) that f is a weak solution to (1.1). Since f is smooth, it is a classical solution as well.

Hence, if X_0 has finite moments of order two and is not a Dirac mass, which will be the considered framework from now on, the connection between (2.1) and the solution to (1.1) is fully justified.

Note that in the above computations, we have used the following relation for the (local) covariance of the martingale part driving (2.1),

$$\begin{aligned}\mathrm{E}\Big[\Big(\int_0^1\sigma(v-Y_s(\alpha))W(ds,d\alpha)\Big)\Big(\int_0^1\sigma(v-Y_s(\alpha))W(ds,d\alpha)\Big)^\top\Big]\Big|_{v=\chi_s}\\ = \int_{\mathbb{R}^N} a(v-v_*)f(s,v_*)dv_*\Big|_{v=\chi_s} ds = \bar{a}(s,\chi_s)ds,\end{aligned}$$

which can thus be identified with the local covariance in (2.1) with the diffusion matrix $\bar{a}$.

The starting point of our analysis is the geometric interpretation of the covariance matrix

$$a(v) = |v|^2\Pi(v),\quad \Pi(v) = \big(\mathrm{Id}_N - \frac{v\otimes v}{|v|^2}\big),\quad v\in\mathbb{R}^N\backslash\{0\}, \tag{2.4}$$

where, for $v \neq 0$, $\Pi(v)$ is the orthogonal projection onto $v^\perp$. Indeed, the key observation is that $a(v)$ also reads as the covariance matrix of the image of v by an antisymmetric standard Gaussian matrix of dimension $N \times N$:

(5) changing now the previous $W = ((W^i)_{1\leq i\leq N})$ into $W = ((W^{i,j})_{1\leq i,j\leq N})$ where the $(W^{i,j})_{1\leq i,j\leq N}$ are independent Gaussian white noises with covariance measure $dsd\alpha$ on $\mathbb{R}^+ \times [0,1]$,

it holds:

$$\frac{1}{2}\mathrm{E}\big[\big((W-W^\top)(ds,d\alpha)v\big)\otimes\big((W-W^\top)(ds,d\alpha)v\big)\big] = a(v)dsd\alpha, \quad v\in\mathbb{R}^N.$$

The proof is just a consequence of the fact

$$\begin{aligned}&\frac{1}{2}\sum_{k,\ell=1}^N \mathrm{E}\big[(W-W^\top)_{i,k}(ds,d\alpha)v_k\big((W-W^\top)_{j,\ell}(ds,d\alpha)v_\ell\big)\big]\\ &= \sum_{k,\ell=1}^N \big(\delta_{(i,k)}^{(j,\ell)} - \delta_{(i,k)}^{(\ell,j)}\big)v_kv_\ell dsd\alpha = \big(\delta_i^j|v|^2 - v_iv_j\big)dsd\alpha = \big(a(v)\big)_{i,j}dsd\alpha,\end{aligned} \tag{2.5}$$

where we have used the Kronecker symbol in the second line.

We derive the following result, which is at the core of the proof:

Lemma 2.1. *Given the process $(Y_t)_{t\geq 0}$, solution to Equation (2.1), consider the solution $(X_t)_{t\geq0}$ to the SDE*

$$\begin{aligned}X_t &= X_0\\ &+ \int_0^t\int_0^1 \frac{W-W^\top}{2^{1/2}}(ds,d\alpha)(X_s - Y_s(\alpha)) - (N-1)\int_0^t\int_0^1 (X_s - Y_s(\alpha))d\alpha ds.\end{aligned} \tag{2.6}$$

Then, $(X_t)_{t\geq0}$ has the same law as $(Y_t)_{t\geq0}$ and thus as the solution of the Landau SDE.

Proof. The process $(Y_t)_{t\geq0}$ being given, the above equation is linear and therefore has a unique (strong) solution. On the other hand, since (2.1) admits in our current setting a unique in law solution, it suffices to identify the bracket (in time) of the martingale part in (2.6) with $\bar{a}(t,X_t)dt$ (see (2.5)), to conclude that the law of $(X_s)_{s\geq0}$ in (2.6) coincides with the one of $(\mathcal{X}_s)_{s\geq0}$ solving (2.1). □

The representation (2.6) is linear and therefore factorizes through the resolvent. Namely,

Lemma 2.2. *The solution $(X_t)_{t\geq0}$ to (2.6) admits the following representation*

$$X_t = Z_t\left[X_0 - \int_0^t\int_0^1 Z_s^\top\frac{W-W^\top}{2^{1/2}}(ds,d\alpha)Y_s(\alpha)\right], \tag{2.7}$$

where letting

$$B_t = 2^{-1/2}\int_0^t\int_0^1 [W-W^\top](ds,d\alpha), \quad t\geq 0, \tag{2.8}$$

the process $(Z_t)_{t\geq0}$ solves the SDE:

$$Z_t = \mathrm{Id}_N + \int_0^t dB_sZ_s - (N-1)\int_0^t Z_sds = \mathrm{Id}_N + \int_0^t dB_s\circ Z_s, \tag{2.9}$$

where $dB_s\circ$ denotes the Stratonovitch integral.

The proof follows from a straightforward application of Itô's formula, noticing that the bracket $\int_0^1 [W-W^\top](dt,d\alpha)Z_t \cdot \int_0^1 Z_t^\top [W-W^\top](dt,d\alpha)Y_t(\alpha)$ is equal to

$$= \int_0^1 ([W-W^\top]\cdot[W-W^\top])(dt,d\alpha)Y_t(\alpha) = -2(N-1)\left(\int_0^1 \mathrm{Id}_N\, Y_t(\alpha)d\alpha\right)dt, \quad (2.10)$$

since $[W\cdot W]_{i,j}(dt,d\alpha) = \sum_{k=1}^N [W_{i,k}\cdot W_{k,j}](dt,d\alpha) = \delta_i^j dtd\alpha$ and $[W\cdot W^\top]_{i,j}(dt,d\alpha) = \sum_{k=1}^N [W_{i,k}\cdot W_{j,k}](dt,d\alpha) = N\delta_i^j dtd\alpha$.

The main feature is that $(B_t)_{t\geq 0}$ is a Gaussian process (with values in $\mathbb{R}^{N\times N}$) with $\mathrm{E}[B_t^{i,k}B_t^{j,\ell}] = t(\delta_{(i,k)}^{(j,\ell)} - \delta_{(i,k)}^{(\ell,j)})$ as covariance. In particular, $((B_t^{i,j})_{1\leq i<j\leq N})_{t\geq 0}$ is a standard Brownian motion with values in $\mathbb{R}^{N(N-1)/2}$. The matrix valued process B thus corresponds to the Brownian motion on the set $\mathcal{A}_N(\mathbb{R})$ of antisymmetric matrices. Recalling that $\mathcal{A}_N(\mathbb{R})$ is the Lie algebra of the special orthogonal group, this allows to identify $(Z_t)_{t\geq 0}$ with the right Brownian motion on $\mathrm{SO}_N(\mathbb{R})$ (see e.g. Chapter V in Rogers and Williams [20] and Chapter VII in Franchi and Le Jan [8]).

2.2 Conditional representation of the transition density

Throughout the paper, we shall assume that the centering condition

$$\mathrm{E}[X_0] = 0 \quad (2.11)$$

is in force. Actually, there is no loss of generality since, whenever $\mathrm{E}[X_0]\neq 0$, (2.6) ensures that, for all $t\geq 0, \mathrm{E}[X_t] = \mathrm{E}[X_0]$ and that $(X_t - \mathrm{E}[X_t])_{t\geq 0}$ also solves the equation.

The main representation of the conditional density is then the following:

Proposition 2.3. *Assume that X_0 is not a Dirac mass and is centered. Then for all $t>0$, the conditional law of X_t given $X_0 = x_0$ has a density, which can be expressed as*

$$f_{x_0}(t,v) = \mathrm{E}\left[(2\pi)^{-N/2}\det{}^{-1/2}(C_t)\exp\left(-\frac{1}{2}\big\langle v - Z_t x_0, C_t^{-1}\big(v - Z_t x_0\big)\big\rangle\right)\right], \quad (2.12)$$

for all $v\in\mathbb{R}^N$, where

$$C_t = \int_0^t Z_t Z_s^\top\big(\mathrm{E}[|X_s|^2]\mathrm{Id}_N - \mathrm{E}[X_s\otimes X_s]\big)\big(Z_t Z_s^\top\big)^\top ds. \quad (2.13)$$

The proof of Proposition 2.3 is postponed to Section 3.

From the above expression of the (stochastic) covariance matrix C_t, we introduce the (deterministic) matrix

$$\Lambda_s := \mathrm{E}[|X_s|^2]\mathrm{Id}_N - \mathrm{E}[X_s\otimes X_s]. \quad (2.14)$$

The matrix Λ_s then plays a key role for the control of the non-degeneracy of the diffusion matrix $\bar{a}(s,v)$, which, by (2.4), reads

$$\bar{a}(s,v) = \int_{\mathbb{R}^N} a(v-v_*)f(s,v_*)dv_* = \mathrm{E}\big[|X_s - v|^2\mathrm{Id}_N - (X_s - v)\otimes(X_s - v)\big], \quad v\in\mathbb{R}^N.$$

Since, for all $s\geq 0$, $\mathrm{E}[X_s] = \mathrm{E}[X_0] = 0$, we get that for all $v\in\mathbb{R}^N$:

$$\bar{a}(s,v) = \Lambda_s - \big(2\mathrm{E}[\langle X_s, v\rangle]\mathrm{Id}_N - \mathrm{E}[X_s\otimes v + v\otimes X_s]\big) + a(v) = \Lambda_s + a(v),$$

so that

$$\forall \xi\in\mathbb{R}^N,\ \langle \bar{a}(s,v)\xi,\xi\rangle = \langle\Lambda_s\xi,\xi\rangle + \langle a(v)\xi,\xi\rangle \geq \langle\Lambda_s\xi,\xi\rangle, \quad (2.15)$$

where we used that a is positive semidefinite for the last inequality.

The behavior of Λ_s can be summarized with the following result.

Proposition 2.4. *Assume that X_0 is not a Dirac mass and is centered. Then, for any $t > 0$, and for all $\xi \in \mathbb{R}^N$,*

$$\Psi_N(t, \overline{\lambda})|\xi|^2 \mathrm{E}\left[|X_0|^2\right] \leq \langle \xi, \Lambda_t \xi \rangle \leq \Psi_N(t, \underline{\lambda})|\xi|^2 \mathrm{E}\left[|X_0|^2\right].$$

where for all $(t, \beta) \in \mathbb{R}^+ \times [0, 1]$,

$$\Psi_N(t, \beta) := (1 - 1/N)(1 - \exp(-2Nt)) + (1 - \beta)\exp(-2Nt),$$

and

$$0 \leq \underline{\lambda} := \inf_{\xi \in \mathbb{R}^N, |\xi|=1} \frac{\mathrm{E}[|\langle \xi, X_0 \rangle|^2]}{\mathrm{E}[|X_0|^2]} \leq \sup_{\xi \in \mathbb{R}^N, |\xi|=1} \frac{\mathrm{E}[|\langle \xi, X_0 \rangle|^2]}{\mathrm{E}[|X_0|^2]} =: \overline{\lambda} \leq 1.$$

Proposition 2.4 will be proved in the next section. For any $t > 0$, it provides a lower bound for the spectrum of Λ_t. There are two cases. If $\overline{\lambda} < 1$, letting $\overline{\eta} := (1-\overline{\lambda}) \wedge (1-1/N) > 0$ (with the standard notations $a \wedge b := \min(a, b)$ and $a \vee b := \max(a, b)$), it holds that, for any $t \geq 0$ and $\xi \in \mathbb{R}^N$,

$$\langle \xi, \Lambda_t \xi \rangle \geq \overline{\eta}|\xi|^2 \mathrm{E}\left[|X_0|^2\right]. \tag{2.16}$$

so that Λ_t is *non-degenerate*, uniformly in time and space.

If $\overline{\lambda} = 1$, i.e. X_0 is embedded in a line, then for any $t > 0$ and $\xi \in \mathbb{R}^N$,

$$\langle \xi, \Lambda_t \xi \rangle \geq (1 - 1/N)(1 - \exp(-2Nt))|\xi|^2 \mathrm{E}\left[|X_0|^2\right],$$

so that Λ_t is non-degenerate in positive time, uniformly on any $[\varepsilon, +\infty) \times \mathbb{R}^N$, $\varepsilon > 0$. For t small, the lower bound for the spectrum behaves as $2(N-1)t$, so that Λ_t *degenerates* in small time.

In the following, we will call the case $\overline{\lambda} < 1$ (resp. $\overline{\lambda} = 1$) *non degenerate* (resp. *degenerate*).

Remark 2.5. Equations (2.15) and (2.16) entail and extend to arbitrary dimension in the case of Maxwellian molecules the previous non-degeneracy result of Desvillettes and Villani [5] (Proposition 4) on the diffusion matrix $\bar{a}$.

2.3 Estimates in the non-degenerate case

When $\overline{\lambda} < 1$, the spectrum of C_t in (2.13) can be easily controlled since $Z_t Z_s^{-1} = Z_t Z_s^\top \in \mathrm{SO}_N(\mathbb{R})$. In such a case, we then obtain from (2.12) the following first result for the conditional density of the Landau SDE:

Theorem 2.6. *Assume that X_0 is not a Dirac mass, is centered with variance 1, and its law is not supported on a line. Then, for all $t > 0$ and $v \in \mathbb{R}^N$,*

$$(2\pi\underline{\eta}t)^{-N/2}\mathrm{E}\left[\exp\left(-\frac{|v - Z_t x_0|^2}{2\overline{\eta}t}\right)\right] \leq f_{x_0}(t, v) \leq (2\pi\overline{\eta}t)^{-N/2}\mathrm{E}\left[\exp\left(-\frac{|v - Z_t x_0|^2}{2\underline{\eta}t}\right)\right],$$

where $\overline{\eta} := (1 - \overline{\lambda}) \wedge (1 - 1/N) \leq \underline{\eta} := (1 - \underline{\lambda}) \vee (1 - 1/N)$.

Remark 2.7. Observe that, since $(Z_s)_{s \geq 0}$ defines an isometry, the off-diagonal cost $|v - Z_t x_0|^2$ may be rewritten $|Z_t^\top v - x_0|^2$.

This formulation may be more adapted than the previous one when integrating the conditional density with respect to the initial law of X_0.

Now, exploiting the *Aronson like* heat kernel bounds for the marginal density of the rotation process $(Z_t)_{t \geq 0}$, see e.g. Varopoulos *et al.* [23] or Stroock [21], we actually derive in Section 4 the following control:

Theorem 2.8 (Explicit bounds for the conditional density)**.** *Under the assumptions of Theorem 2.6, there exists $C := C(N) \geq 1$ such that, for all $t > 0$, $x_0, v \in \mathbb{R}^N$,*

$$\frac{\delta_t^{N-1}}{Ct^{N/2}} \exp\big(-C\frac{I}{t}\big) \leq f_{x_0}(t,v) \leq \frac{C\delta_t^{N-1}}{t^{N/2}} \exp\big(-\frac{I}{Ct}\big), \tag{2.17}$$

where

$$\delta_t = \frac{1 \wedge \Big(\frac{t^{1/2}}{1 \vee (|x_0| \wedge |v|)}\Big)}{1 \wedge t^{1/2}}, \quad \textit{and} \quad I = \big||v| - |x_0|\big|^2 + (1 \wedge |v| \wedge |x_0|)^2 \big|\frac{v}{|v|} - \frac{x_0}{|x_0|}\big|^2.$$

If $|x_0| \wedge |v| \leq 1$, then δ_t is equal to 1 and I can be chosen as $I = |x_0 - v|^2$, which corresponds to a usual Gaussian estimate.

We stress the fact that the above bounds are sharp. The contribution in $||v| - |x_0||^2$ in I corresponds to a 'radial cost' and the contribution in $|v/|v| - x_0/|x_0||^2$ to a 'tangential cost'. The term $(1 \wedge |v| \wedge |x_0|)^2$ reads as the inverse of the variance along tangential directions. It must be compared with the variance along tangential directions in a standard Gaussian kernel, the inverse of which is of order $(|v| \wedge |x_0|)^2$ as shown in Remark 2.11 below. This says that, when $|x_0|$ and $|v|$ are greater than 1, $f_{x_0}(t,v)$ is super-diffusive in the tangential directions. This is in agreement with the observations made in Introduction: The non-Gaussian regime of the density for x_0 large occurs because of the super-diffusivity along iso-radial curves.

Anyhow, it is worth mentioning that the two-sided bounds become Gaussian when t tends to ∞. Indeed, noting that $\delta_t \to 1$ as $t \to \infty$ and that the tangential cost $(1 \wedge |v| \wedge |x_0|)^2 |v/|v| - x_0/|x_0||^2$ is bounded by 4, (2.17) yields

$$\frac{1}{Ct^{N/2}} \exp\big(-C\frac{\big||x_0| - |v|\big|^2}{t}\big) \leq f_{x_0}(t,v) \leq \frac{C}{t^{N/2}} \exp\big(-\frac{\big||x_0| - |v|\big|^2}{Ct}\big), \tag{2.18}$$

for t large enough (with respect to $|x_0|$, uniformly in $|v|$) and for a new constant C (independent of $|x_0|$ and $|v|$). This coincides with the asymptotic behavior of the N-dimensional Gaussian kernel: In the Gaussian regime, the variance along the tangential directions is $(|v| \wedge |x_0|)^2$, which is less than $|x_0|^2$ and which shows, in the same way as in (2.18), that the Gaussian tangential cost is also small in front of t, uniformly in v. However, some differences persist asymptotically when $|x_0|$ is large. Due to the super-diffusivity of the tangential directions in the Landau equation, the Landau tangential cost decays faster than the Gaussian one. Intuitively, the reason is that the 'angle' of the Landau process $(X_t)_{t \geq 0}$ reaches the uniform distribution on the sphere at a quicker rate than in the Gaussian regime. Clearly, the fact that the system forgets the initial angle of x_0 in long time could be recovered from Theorem 2.6 by replacing (at least formally) Z_t by a uniformly distributed random matrix on $\mathrm{SO}_N(\mathbb{R})$.

Of course, when the initial mass is already uniformly distributed along the spheres centered at 0, the marginal density of $(X_t)_{t \geq 0}$ already behaves in finite time as if the transition density was Gaussian. We illustrate this property in the following corollary (the proof of which is deferred to the next section):

Corollary 2.9. *Assume that X_0 admits an initial density of the radial form:*

$$f_0(x_0) = f(|x_0|),$$

for some Borel function $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$. Then, we can find a constant $C := C(N) \geq 1$ such that, for all $t > 0$,

$$\frac{1}{Ct^{N/2}} \int_{\mathbb{R}^N} f_0(x_0) g_N\big(C\frac{x_0 - v}{t^{1/2}}\big) dx_0 \leq f_t(v) \leq \frac{C}{t^{N/2}} \int_{\mathbb{R}^N} f_0(x_0) g_N\big(\frac{x_0 - v}{Ct^{1/2}}\big) dx_0, \tag{2.19}$$

where g_N denotes the standard Gaussian kernel of dimension N and where f_t is the solution of the Landau equation, which here reads

$$f_t(v) = \int_{\mathbb{R}^N} f_0(x_0) f_{x_0}(t,v) dx_0.$$

To conclude this subsection, notice that the Gaussian regime (that corresponds to $|x_0| \wedge |v| \le 1$ in the statement of Theorem 2.8) can be derived from (2.17) using the following Lemma and Remark.

Lemma 2.10. *Let $x_0 \in \mathbb{R}^N$ be given and $\Pi_{B_N(0,|x_0|)}$ denote the orthogonal projection from $\mathbb{R}^N$ onto the ball $B_N(0,|x_0|)$ of center 0 and radius $|x_0|$. Then, for all $v \in \mathbb{R}^N$ such that $|x_0| < |v|$,*

$$\begin{aligned} &|v - \Pi_{B_N(0,|x_0|)}(v)|^2 + |\Pi_{B_N(0,|x_0|)}(v) - x_0|^2 \\ &\quad \le |v - x_0|^2 \le 2|v - \Pi_{B_N(0,|x_0|)}(v)|^2 + 2|\Pi_{B_N(0,|x_0|)}(v) - x_0|^2. \end{aligned}$$

Proof. We write

$$\begin{aligned} |v - x_0|^2 &= |v - \Pi_{B_N(0,|x_0|)}(v) + \Pi_{B_N(0,|x_0|)}(v) - x_0|^2 \\ &= |v - \Pi_{B_N(0,|x_0|)}(v)|^2 + |\Pi_{B_N(0,|x_0|)}(v) - x_0|^2 \\ &\quad + 2\langle v - \Pi_{B_N(0,|x_0|)}(v), \Pi_{B_N(0,|x_0|)}(v) - x_0 \rangle. \end{aligned}$$

Now $\langle v - \Pi_{B_N(0,|x_0|)}(v), \Pi_{B_N(0,|x_0|)}(v) - x_0 \rangle \ge 0$, by orthogonal projection on a closed convex subset, and the lower bound follows. By convexity, we obtain the upper bound. □

Remark 2.11. Let us consider two given points $x_0, v \in \mathbb{R}^N$ such that $|x_0| \le |v|$. Noticing that $\Pi_{B_N(0,|x_0|)}(v) = (|x_0|/|v|)v$ and then that $|v - \Pi_{B_N(0,|x_0|)}(v)| = |v| - |x_0|$, we deduce from Lemma 2.10 that

$$\big||v| - |x_0|\big|^2 + |x_0|^2 \left| \frac{v}{|v|} - \frac{x_0}{|x_0|} \right|^2 \le |v - x_0|^2 \le 2\big||v| - |x_0|\big|^2 + 2|x_0|^2 \left| \frac{v}{|v|} - \frac{x_0}{|x_0|} \right|^2.$$

In particular, when $|x_0| \le 1$ we derive from (2.17) in Theorem 2.8 the usual two-sided Gaussian estimates. Now, if $|v| \le |x_0|$ and $|v| \le 1$, this still holds by symmetry.

2.4 Estimates in the degenerate case

We now discuss the case when the initial condition lies in a straight line, which by rotation invariance can be assumed to be the first vector e_1 of the canonical basis. By Proposition 2.4, we already know that the matrix Λ_t (see (2.14)) driving the ellipticity of the covariance matrix C_t (see (2.13)) becomes non-degenerate in positive time. This says that, after a positive time t_0, the system enters the same regime as the one discussed in Theorem 2.8, so that the transition density of the process satisfies, after t_0, the bounds (2.17). Anyhow, this leaves open the small time behavior of the transition kernel of the process.

Here, we thus go thoroughly into the analysis and specify both the on-diagonal rate of explosion and the off-diagonal decay of the conditional density in small time. Surprisingly, we show that the tail of the density looks much more like an exponential distribution rather than a Gaussian one. Precisely, we show that the off-diagonal decay of the density is of Gaussian type for 'untypical' values only, which is to say that, for values where the mass is effectively located, the decay is of exponential type. Put it differently, the two-sided bounds we provide for the conditional density read as a mixture of exponential and Gaussian distributions.

Theorem 2.12. *Assume that the initial distribution of X_0 is compactly supported by e_1, i.e. there exists $C_0 > 0$ such that $X_0 \in [-C_0 e_1, C_0 e_1]$ a.s. Then, there exists $C := C(C_0) > 1$ such that, for $t \in (0, 1/C]$:*

$$\frac{1}{Ct^{(N+1)/2}} \exp\big(-CI(t, x_0, v)\big) \leq f_{x_0}(t, v) \leq \frac{C}{t^{(N+1)/2}} \exp\big(-\frac{I(t, x_0, v)}{C}\big),$$

where Ősormais au cas compact, pour lequel la derniŔre estimŐe rend compte de tout ce qui se passe:

$$I(t, x_0, v) = \frac{|v^1 - x_0^1|}{t} + \frac{|v^1 - x_0^1|^2}{t} + \sum_{i=2}^{N} \frac{|v^i|^2}{t}.$$

The reason why the conditional density follows a mixture of exponential and Gaussian rates may be explained as follows in the simplest case when $x_0 = 0$. The starting point is formula (2.12) in Proposition 2.3. When the initial condition is degenerate, the conditional covariance matrix C_t in (2.13) has two scales. As shown right below, the eigenvalues of C_t along the directions $e_2, \dots, e_N$ are of order t whereas the eigenvalue λ_t^1 of C_t along the direction e_1 is of order t^2 with large probability. Anyhow, with exponentially small probability, λ_t^1 is of order t: Precisely, the probability that it is of order ξt has logarithm of order $-\xi/t$ when $\xi \in (0, 1)$. Such large deviations of λ_t^1 follow from large deviations of $(Z_s)_{0 \leq s \leq t}$ far away from the identity. This rough description permits to compare the contributions of typical and rare events in the formula (2.13) for the density $f_{x_0}(t, v)$, when computed at a vector v parallel to the direction e_1. On typical scenarios, the off-diagonal cost $\langle C_t^{-1} v, v\rangle$ in the exponential appearing in (2.13) is of order $|v|^2/t^2$. In comparison with, by choosing ξ of order $|v|$, the events associated with large deviations of C_t generate an off-diagonal cost $\langle C_t^{-1} v, v\rangle$ of order $|v|/t$ with an exponentially small probability of logarithmic order -$|v|/t$: The resulting contribution in the off-diagonal decay is of order $|v|/t$, which is clearly smaller than $|v|^2/t^2$. This explains the exponential regime of $f_{x_0}(t, v)$. The Gaussian one follows from a threshold phenomenon: as $(Z_s)_{0 \leq s \leq t}$ takes values in $\mathrm{SO}_N(\mathbb{R})$, there is no chance for its elements to exceed 1 in norm. Basically, it means that, when $|v|$ is large, the best choice for ξ is not $|v|$ but 1: The corresponding off-diagonal cost is $|v|^2/t$, which occurs with probability of logarithmic order $-1/t$. This explains the Gaussian part of $f_{x_0}(t, v)$.

In the case when the conditioned initial position x_0 is not zero, specifically when it is far away from 0, things become much more intricate as the transport of the initial position x_0 by Z_t affects the density. This is the reason why we consider a compactly supported initial condition. To compare with, notice that, in the non-degenerate case, (2.17) gives Gaussian estimates when x_0 is restricted to a compact set. This is exactly what the statement of Theorem 2.8 says when $|x_0| \leq 1$, the argument working in the same way when $|x_0| \leq C_0$, for some $C_0 > 1$.

3 Conditional density of the Landau SDE: Derivation and Properties

3.1 Proof of Proposition 2.3

We claim

Lemma 3.1. *Recall X_0 is centered. Letting*

$$\bar{B}_t = 2^{-1/2} \int_0^t \int_0^1 [W - W^\top](ds, d\alpha) Y_s(\alpha), \tag{3.1}$$

the processes $(B_t)_{t \geq 0}$ and $(\bar{B}_t)_{t \geq 0}$ are independent. Also, the processes $(Z_t)_{t \geq 0}$ and $(\bar{B}_t)_{t \geq 0}$ are independent.

Proof. We know that setting $\tilde{Z}_t := \exp((N-1)t)Z_t$ then

$$\tilde{Z}_t = \mathrm{Id}_N + B_t + \int_0^t dB_s B_s + \cdots = \mathrm{Id}_N + \sum_{n\geq 1} \int_{0\leq t_n \leq \cdots \leq t_1 \leq t} dB_{t_1} dB_{t_2} \dots dB_{t_n}.$$

Hence it suffices to show that B and $\bar{B}$ are independent. As both are Gaussian processes, this can be easily proved by computing their covariance which turns out to be zero if X_0 is centered, see (2.10). □

Recalling that we can rewrite X_t as

$$X_t = Z_t \left[X_0 - \int_0^t Z_s^\top d\bar{B}_s \right], \quad t \geq 0, \tag{3.2}$$

X_0 being independent of $(B_t, \bar{B})_{t\geq 0}$, and using (2.5) to compute the covariance matrix of the Gaussian process $(\bar{B}_t)_{t\geq 0}$:

$$\frac{d}{dt}\mathrm{E}\big[\bar{B}_t \bar{B}_t^\top\big] = \int_0^1 \big\{\mathrm{Id}_N |Y_t(\alpha)|^2 - Y_t(\alpha)\otimes Y_t(\alpha)\big\} d\alpha = \mathrm{E}[|X_t|^2]\mathrm{Id}_N - \mathrm{E}[X_t \otimes X_t] = \Lambda_t, \tag{3.3}$$

the existence of the transition density and the representation (2.12) are direct consequences of (3.2) and Lemma 3.1. This proves Lemma 2.2.

3.2 Additional properties on the resolvent process

We give in this paragraph some additional properties on the process Z that are needed for the derivation of the density estimates. We will make use of the following lemma whose proof can be found in Franchi and Le Jan [8], see Theorem VII.2.1 and Remark VII.2.6.

Lemma 3.2. *Given $t > 0$, the process $(Z_t Z_{t-s}^\top)_{0\leq s\leq t}$ has the same law as the process $(Z_s)_{0\leq s\leq t}$.*

3.3 Proof of Proposition 2.4

Recall from (2.6) that the expectation is preserved, i.e.

$$\mathrm{E}[X_t] = \mathrm{E}[X_0], \qquad \text{for all } t \geq 0. \tag{3.4}$$

Since we also assumed that $\mathrm{E}[X_0] = 0$, the process $(X_t)_{t\geq 0}$ is centered. The point is then to compute

$$\langle \Lambda_t \xi, \xi \rangle = \mathrm{E}\left[|\xi|^2 |X_t|^2 - \langle \xi, X_t \rangle^2 \right], \quad \xi \in \mathbb{R}^N,\ t \geq 0.$$

Noting that $\mathrm{Trace}[a(v)] = (N-1)|v|^2$, for $v \in \mathbb{R}^N$, we get that

$$\begin{aligned}
\frac{d}{dt}\mathrm{E}\left[|X_t|^2\right] &= \int_0^1 \mathrm{E}\left[\mathrm{Trace}\left[a(X_t - Y_t(\alpha))\right]\right] d\alpha - 2(N-1)\int_0^1 \mathrm{E}\left[\langle X_t, X_t - Y_t(\alpha)\rangle\right] d\alpha \\
&= (N-1)\int_0^1 \mathrm{E}\left[|X_t - Y_t(\alpha)|^2\right] d\alpha - 2(N-1)\mathrm{E}\left[|X_t|^2\right] = 0.
\end{aligned}$$

To pass from the first to the second line, we used the fact that X_t and Y_t are centered and independent (as they are defined on each of the two components of the product space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P}) \otimes ([0,1], \mathcal{B}([0,1]), d\alpha)$, see the introduction of Subsection 2.1).

Therefore, the energy is preserved:

$$\mathrm{E}\left[|X_t|^2\right] = \mathrm{E}\left[|X_0|^2\right], \qquad \text{for all } t \geq 0. \tag{3.5}$$

Moreover, using the expression (2.1) of the Landau SDE (which implies that, just in the equation right below, W becomes again an N-dimensional space-time white noise), we see that

$$d\langle \xi, X_t\rangle = \int_0^1 \langle \sigma^\top (X_t - Y_t(\alpha))\xi, W(dt,d\alpha)\rangle - (N-1)\int_0^1 \langle \xi, (X_t - Y_t(\alpha))\rangle d\alpha dt.$$

Since $|\sigma^\top(y)\xi|^2 = |\xi|^2|y|^2 - \langle \xi, y\rangle^2$, we have

$$\begin{aligned}
&\frac{d}{dt}\mathrm{E}\left[\langle \xi, X_t\rangle^2\right] \\
&= \int_0^1 \mathrm{E}\left[|\sigma^\top\left(X_t - Y_t(\alpha)\right)\xi|^2\right] d\alpha - 2(N-1)\mathrm{E}\int_0^1 \langle \xi, X_t\rangle\langle \xi, (X_t - Y_t(\alpha))\rangle d\alpha \\
&= |\xi|^2 \int_0^1 \mathrm{E}\left[|X_t - Y_t(\alpha)|^2\right] d\alpha - \int_0^1 \mathrm{E}\left[\langle \xi, X_t - Y_t(\alpha)\rangle^2\right] d\alpha - 2(N-1)\mathrm{E}\left[\langle \xi, X_t\rangle^2\right] \\
&= 2|\xi|^2\mathrm{E}\left[|X_t|^2\right] - 2N\mathrm{E}\left[\langle \xi, X_t\rangle^2\right].
\end{aligned}$$

From (3.5), we deduce that

$$\frac{d}{dt}\mathrm{E}\left[\langle \xi, X_t\rangle^2\right] = 2|\xi|^2\mathrm{E}\left[|X_0|^2\right] - 2N\mathrm{E}\left[\langle \xi, X_t\rangle^2\right],$$

so that, for any $t \geq 0$,

$$\mathrm{E}\left[\langle \xi, X_t\rangle^2\right] = \exp(-2Nt)\left\{\mathrm{E}\left[\langle \xi, X_0\rangle^2\right] + 2\int_0^t \exp(2Ns)|\xi|^2\mathrm{E}\left[|X_0|^2\right] ds\right\}.$$

Finally,

$$\mathrm{E}\left[\langle \xi, X_t\rangle^2\right] = \exp(-2Nt)\mathrm{E}\left[\langle \xi, X_0\rangle^2\right] + \frac{1}{N}\left[1 - \exp(-2Nt)\right]|\xi|^2\mathrm{E}\left[|X_0|^2\right].$$

Therefore,

$$\begin{aligned}
\langle \Lambda_t\xi, \xi\rangle &= |\xi|^2\mathrm{E}\left[|X_t|^2\right] - \mathrm{E}\left[\langle \xi, X_t\rangle^2\right] \\
&= \frac{1}{N}\left[N - 1 + \exp(-2Nt)\right]|\xi|^2\mathrm{E}\left[|X_0|^2\right] - \exp(-2Nt)\mathrm{E}\left[\langle \xi, X_0\rangle^2\right].
\end{aligned} \tag{3.6}$$

Plugging the values of $\underline{\lambda}$ and $\overline{\lambda}$ in (3.6), we get the announced result.

3.4 Proof of Corollary 2.9

By Theorem 2.8, the result is straightforward when $|v| \leq 1$ (as the transition density has a Gaussian shape). When $|v| \geq 1$, the problem can be reformulated as follows. Given a constant $C > 0$, the point is to estimate

$$\begin{aligned}
q_t(v) := \int_{\mathbb{R}^N} &\frac{[\delta_t(|x_0|)]^{N-1}}{t^{N/2}} \\
&\times f_0(x_0)\exp\left(-\frac{C}{t}\left\{\big||v| - |x_0|\big|^2 + (1\wedge|x_0|)^2\Big|\frac{v}{|v|} - \frac{x_0}{|x_0|}\Big|^2\right\}\right) dx_0,
\end{aligned} \tag{3.7}$$

where we have let $\delta_t(|x_0|) := \dfrac{1 \wedge \left(\dfrac{t^{1/2}}{1 \vee (|x_0| \wedge |v|)}\right)}{1 \wedge t^{1/2}}$.

By a polar change of variable, we get

$$\begin{aligned}
q_t(v) = \frac{1}{t^{N/2}}\int_0^{+\infty} &d\rho\, \rho^{N-1}\left[\delta_t(\rho)\right]^{N-1} f_0(\rho)\exp\left(-\frac{C}{t}\big|\rho - |v|\big|^2\right) \\
&\times \int_{\mathbb{S}^{N-1}} \exp\left(-\frac{C}{t}(1\wedge\rho)^2\Big|s - \frac{v}{|v|}\Big|^2\right) d\nu_{\mathbb{S}^{N-1}}(s),
\end{aligned} \tag{3.8}$$

where $\nu_{\mathbb{S}^{N-1}}$ denotes the Lebesgue measure on the sphere $\mathbb{S}^{N-1}$ of dimension $N-1$.

As we shall make use of its renormalized version below, we normalize $\nu_{\mathbb{S}^{N-1}}$, so that $\nu_{\mathbb{S}^{N-1}}$ reads as a probability measure. Up to a multiplicative constant, the above expression remains unchanged. In particular, as we are just interested in lower and upper bounds of $q_t(v)$, we can keep the above as a definition for $q_t(v)$, with $\nu_{\mathbb{S}^{N-1}}$ being normalized.

Let us now recall the following two-sided heat kernel estimate on $\mathbb{S}^{N-1}$, see e.g. [21]. There exists $C' := C'(N) \geq 1$ such that, for all $t > 0$,

$$(C')^{-1} \leq \frac{1}{\Big(1 \wedge \dfrac{t^{1/2}}{1 \wedge \rho}\Big)^{N-1}} \int_{\mathbb{S}^{N-1}} \exp\left(-\frac{C(1\wedge\rho)^2}{t}\Big|s - \frac{v}{|v|}\Big|^2\right) d\nu_{\mathbb{S}^{N-1}}(s) \leq C'. \tag{3.9}$$

Therefore, what really counts in the expression of $q_t(v)$ is the product

$$\Big(1 \wedge \frac{t^{1/2}}{1\wedge\rho}\Big)\delta_t(\rho) = \begin{cases} 1 \wedge \dfrac{t^{1/2}}{\rho} & \text{if } \rho \leq 1, \\ 1 \wedge \dfrac{t^{1/2}}{\rho} & \text{if } 1 \leq \rho \leq |v|, \\ 1 \wedge \dfrac{t^{1/2}}{|v|} & \text{if } \rho > |v|, \end{cases} \tag{3.10}$$

Up to a redefinition of the function q_t, it is thus sufficient to consider

$$\begin{aligned} q_t(v) := {} & \frac{1}{t^{N/2}} \int_0^{|v|} f_0(\rho) \exp\left(-\frac{C}{t}\big|\rho - |v|\big|^2\right) \left\{1 \wedge \frac{t^{1/2}}{\rho}\right\}^{N-1} \rho^{N-1} d\rho \\ & + \frac{1}{t^{N/2}} \int_{|v|}^{+\infty} f_0(\rho) \exp\left(-\frac{C}{t}\big|\rho - |v|\big|^2\right) \left\{1 \wedge \frac{t^{1/2}}{|v|}\right\}^{N-1} \rho^{N-1} d\rho. \end{aligned} \tag{3.11}$$

Compare now with what happens when the convolution in (3.7) is made with respect to the Gaussian kernel. Basically $\delta_t(|x_0|)$ is replaced by 1 and $1 \wedge |x_0|$ is replaced by $|v| \wedge |x_0|$ (see Remark 2.11). Equivalently, $\delta_t(\rho)$ is replaced by 1 and $1 \wedge \rho$ by $|v| \wedge \rho$ in (3.8). This says that, in (3.9), $1 \wedge \rho$ is replaced by $|v| \wedge \rho$. Then, in (3.10), $\delta_t(\rho)$ is replaced by 1 and $1 \wedge \rho$ by $|v| \wedge \rho$, which leads exactly to the same three equalities. This shows that, in the Gaussian regime, the right quantity to consider is also (3.11).

4 Proof of the Density Estimates in the Non-Degenerate case

4.1 Preliminary results for the Haar measure and the heat kernel on $\mathrm{SO}_N(\mathbb{R})$

Starting from the representation Theorem 2.6, we want to exploit the *Aronson like* heat kernel estimates for the special orthogonal group. Precisely, from VIII.2.9 in Varopoulos *et al* [23], we derive that, for $t > 0$, the law of Z_t has a density, denoted by $p_{\mathrm{SO}_N}(t, \mathrm{Id}_N, \cdot)$, with respect to the probability Haar measure μ_{SO_N} of $\mathrm{SO}_N(\mathbb{R})$. Moreover, there exists a constant $\beta > 1$ such that, for any $g \in \mathrm{SO}_N(\mathbb{R})$ and for all $t > 0$:

$$\begin{aligned} & \frac{1}{\beta(1\wedge t)^{N(N-1)/4}} \exp\left(-\beta \frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N, g)}{t}\right) \\ & \leq p_{\mathrm{SO}_N}(t, \mathrm{Id}_N, g) \leq \frac{\beta}{(1\wedge t)^{N(N-1)/4}} \exp\left(-\frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N, g)}{\beta t}\right), \end{aligned} \tag{4.1}$$

where $d_{\mathrm{SO}_N}(\mathrm{Id}_N, g)$ denotes the Carnot distance between Id_N and g:

$$d_{\mathrm{SO}_N}(\mathrm{Id}_N, g) = \inf_{H \in \mathcal{A}_N(\mathbb{R}): e^H = g} \|H\|,$$

$\|\cdot\|$ standing for the usual matricial norm on $\mathcal{M}_N(\mathbb{R})$. Proof of the diagonal rate in (4.1) relies on the following volume estimate from Theorem V.4.1 in [23]: By compactness of $\mathrm{SO}_N(\mathbb{R})$, there exists $C_N \geq 1$ such that, for all $t > 0$,

$$C_N^{-1}(1 \wedge t)^{N(N-1)/4} \leq \mu_{\mathrm{SO}_N}\big(B_{\mathrm{SO}_N}(t^{1/2})\big) \leq C_N(1 \wedge t)^{N(N-1)/4}, \tag{4.2}$$

where $B_{\mathrm{SO}_N}(\rho) := \{g \in \mathrm{SO}_N(\mathbb{R}) : d_{\mathrm{SO}_N}(\mathrm{Id}_N, g) \leq \rho\}$, for $\rho > 0$, denotes the ball of radius ρ and center Id_N.

By local inversion of the exponential, it is well-checked that the Carnot distance is continuous with respect to the standard matricial norm on $\mathcal{M}_N(\mathbb{R})$. In particular, by compactness of $\mathrm{SO}_N(\mathbb{R})$, it is bounded on the whole group. Actually, we claim:

Lemma 4.1 (Equivalence between Carnot distance and matrix norm on the group)**.** *There exists a constant $C := C(N) > 1$ such that, for any $g \in \mathrm{SO}_N(\mathbb{R})$,*

$$C^{-1}\|\mathrm{Id}_N - g\| \leq d_{\mathrm{SO}_N}(\mathrm{Id}_N, g) \leq C\|\mathrm{Id}_N - g\|.$$

Proof. We first prove the upper bound. Without any loss of generality, we can restrict the analysis to the elements $g \in \mathrm{SO}_N(\mathbb{R})$ such that $\|\mathrm{Id}_N - g\| \leq \varepsilon$, for some $\varepsilon > 0$ small enough, the value of which is fixed right below. Indeed, for those g such that $\|\mathrm{Id}_N - g\| > \varepsilon$, the upper bound directly follows from the boundedness of the Carnot distance on the group.

Now, we can choose ε small enough so that the logarithm mapping on $\mathcal{M}_N(\mathbb{R})$ realizes a diffeomorphism from the ball of center Id_N and radius $\varepsilon > 0$ into some open subset around the null matrix. Then, letting $H := \ln(g)$, we deduce from the variational definition of the distance that $d_{\mathrm{SO}_N}(\mathrm{Id}_N, g) \leq \|H\|$. Writing $H = \ln(\mathrm{Id}_N + g - \mathrm{Id}_N)$, we obtain that $\|H\| \leq C\|g - \mathrm{Id}_N\|$ for some $C := C(N)$, which proves that $d_{\mathrm{SO}_N}(\mathrm{Id}_N, g) \leq C\|g - \mathrm{Id}_N\|$.

The converse is proved in a similar way. Given $g \in \mathrm{SO}_N(\mathbb{R})$, we deduce from the variational definition of the distance that there exists a matrix $H \in \mathcal{A}_N(\mathbb{R})$ such that $\exp(H) = g$ and $d_{\mathrm{SO}_N}(\mathrm{Id}_N, g) \geq \|H\|/2$. Since $d_{\mathrm{SO}_N}(\mathrm{Id}_N, g)$ is bounded by the diameter of the group, we get $\|H\| \leq C$ for some C independent of g. By the local Lipschitz property of the exponential, $\|g - \mathrm{Id}_N\| \leq C\|H\|$ (for a possibly new value of the constant C), which yields $\|g - \mathrm{Id}_N\| \leq 2C d_{\mathrm{SO}_N}(\mathrm{Id}_N, g)$. ☐

Part of our analysis relies on a specific parametrization of $\mathrm{SO}_N(\mathbb{R})$ by elements of $\mathbb{S}^{N-1} \times \mathrm{SO}_{N-1}(\mathbb{R})$, where $\mathbb{S}^{N-1}$ is the sphere of dimension $N-1$. Namely, for an element $h \in \mathrm{SO}_{N-1}(\mathbb{R})$, we denote by L_h the element of $\mathrm{SO}_N(\mathbb{R})$:

$$L_h := \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & h & \\ 0 & & & \end{array}\right).$$

Moreover, for an element $s \in \mathbb{S}^{N-1}$, we denote by V_s an element of $\mathrm{SO}_N(\mathbb{R})$ such that $V_s e_1 = s$. It is constructed in the following way. When $\langle s, e_1\rangle \neq 0$, the family $(s, e_2, \ldots, e_N)$ is free. We can orthonormalize it by means of the Gramm-Schmidt procedure. By induction, we let

$$u_1 := s, \quad u_i := e_i - \sum_{k=1}^{i-1}\langle e_i, u_k\rangle \frac{u_k}{|u_k|^2}, \quad i \in \{2, \cdots, N\}, \tag{4.3}$$

and then $s_i := u_i/|u_i|$, for all $i \in \{1, \ldots, N\}$, so that $s_1 = s$. Then, the family $(s_1, s_2, \ldots, s_N)$ is an orthonormal basis and V_s is given by the passage matrix expressing the $(s_i)_{1\leq i\leq N}$'

in the basis $(e_i)_{1\le i\le N}$. When $\langle s,e_1\rangle=0$, we consider $\langle s,e_2\rangle$. If $\langle s,e_2\rangle\neq 0$, then the family $(s,e_3,\dots,e_N,e_1)$ is free and we can apply the Gramm-Schmidt procedure. If $\langle s,e_2\rangle=0$, we then go on until we find some index $k\in\{3,\dots,N\}$ such that $\langle s,e_k\rangle\neq 0$. Such a construction ensures that the mapping $\mathbb{S}^{N-1}\ni s\mapsto V_s\in\mathrm{SO}_N(\mathbb{R})$ is measurable.

With $s\mapsto V_s$ and $h\mapsto L_h$ at hand, we claim that the mapping $\phi:(s,h)\mapsto V_sL_h$ is bijective from $\mathbb{S}^{N-1}\times\mathrm{SO}_{N-1}(\mathbb{R})$ onto $\mathrm{SO}_N(\mathbb{R})$. Given some $g\in\mathrm{SO}_N(\mathbb{R})$, $g=V_sL_h$ if and only if $s=ge_1$ and $L_h=V_{ge_1}^\top g$. By construction of V_s and orthogonality of $V_{ge_1}^\top g$, we indeed check that $(V_{ge_1}^\top g)_{1,1}=1$ and $(V_{ge_1}^\top g)_{i,1}=(V_{ge_1}^\top g)_{1,i}=0$ for $i=2,\dots,N$. In other words, $V_{ge_1}^\top g$ always fits some L_h, the value of h being uniquely determined by the lower block $(V_{ge_1}^\top g)_{2\le i,j\le N}$, which proves the bijective property of ϕ. Denoting by Π_{N-1} the projection mapping:

$$\pi_{N-1}:\mathcal{M}_N(\mathbb{R})\ni(a_{i,j})_{1\le i,j\le N}\mapsto(a_{i,j})_{2\le i,j\le N},$$

we deduce that the converse of ϕ writes $\phi^{-1}:\mathrm{SO}_N(\mathbb{R})\ni g\mapsto(ge_1,\pi_{N-1}(V_{ge_1}^\top g))\in\mathbb{S}^{N-1}\times\mathrm{SO}_{N-1}(\mathbb{R})$.

The mapping ϕ allows us to disintegrate the Haar measure on $\mathrm{SO}_N(\mathbb{R})$ in terms of the product of the Lebesgue probability measure $\nu_{\mathbb{S}^{N-1}}$ on the sphere $\mathbb{S}^{N-1}$ and the Haar probability measure on $\mathrm{SO}_{N-1}(\mathbb{R})$. We have the following result, see e.g. Proposition III.3.2 in [8] for a proof:

Lemma 4.2 (Representation of the Haar measure on $\mathrm{SO}_N(\mathbb{R})$)**.** *Let f be a bounded Borel function from $\mathrm{SO}_N(\mathbb{R})$ to $\mathbb{R}$. Then (with $\nu_{\mathbb{S}^{N-1}}$ the normalized Lebesgue measure on $\mathbb{S}^{N-1}$),*

$$\int_{\mathrm{SO}_N(\mathbb{R})}f(g)d\mu_{\mathrm{SO}_N}(g):=\int_{\mathbb{S}^{N-1}\times\mathrm{SO}_{N-1}(\mathbb{R})}f(V_sL_h)d\nu_{\mathbb{S}^{N-1}}(s)d\mu_{\mathrm{SO}_{N-1}}(h).\tag{4.4}$$

4.2 Proof of Theorem 2.8

From (4.1) and Theorem 2.6, we derive the following two-sided bound for the conditional density. There exists $\tilde C:=\tilde C(N)\ge 1$ such that, for all $t>0$,

$$\begin{aligned}&\frac{1}{\tilde C\beta[t^{N/2}(1\wedge t)^{N(N-1)/4}]}\int_{\mathrm{SO}_N(\mathbb{R})}\exp\Big(-\Big\{\beta\frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N,g)}{t}+\tilde C\frac{|v-gx_0|^2}{t}\Big\}\Big)d\mu_{\mathrm{SO}_N}(g)\\&\quad\le f_{x_0}(t,v)\\&\qquad\le\frac{\tilde C\beta}{t^{N/2}(1\wedge t)^{N(N-1)/4}}\int_{\mathrm{SO}_N(\mathbb{R})}\exp\Big(-\Big\{\frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N,g)}{\beta t}+\frac{|v-gx_0|^2}{\tilde Ct}\Big\}\Big)d\mu_{\mathrm{SO}_N}(g),\end{aligned}\tag{4.5}$$

which will be the starting point to derive the bounds of Theorem 2.8.

4.2.1 Gaussian Regime

Let us first concentrate on the bounds when $|x_0|\wedge|v|\le 1$. Without loss of generality, we can assume by symmetry that $|x_0|\le 1$. Indeed, for all $g\in\mathrm{SO}_N(\mathbb{R})$, $|v-gx_0|=|g^\top v-x_0|$ and $d_{\mathrm{SO}_N}(\mathrm{Id}_N,g)=d_{\mathrm{SO}_N}(\mathrm{Id}_N,g^\top)$. Moreover, the Haar measure is invariant by transposition. This can be checked as follows. If Z is distributed according to the Haar measure, then, for any rotation ρ, $\rho Z^\top=(Z\rho^\top)^\top$. Since $Z\rho^\top$ has the same law as Z (as the group is compact, it is known the Haar measure is invariant both by left and right multiplications), we deduce that the law of $Z^\top$ is invariant by rotation.

Now, write:

$$\frac12\frac{|x_0-v|^2}{t}-\frac{\|\mathrm{Id}_N-g\|^2|x_0|^2}{t}\le\frac{|v-gx_0|^2}{t}\le 2\Big(\frac{\|\mathrm{Id}_N-g\|^2|x_0|^2}{t}+\frac{|x_0-v|^2}{t}\Big).$$

From (4.5) and the assumption $|x_0| \leq 1$, we get that:

$$
\begin{aligned}
f_{x_0}(t,v) &\geq \frac{(\tilde{C}\beta)^{-1}}{t^{N/2}} \exp\big(-2\tilde{C}\frac{|x_0-v|^2}{t}\big) \\
&\quad \times \left\{ \frac{1}{(1\wedge t)^{N(N-1)/4}} \int_{\mathrm{SO}_N(\mathbb{R})} \exp\Big(-\beta \frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N,g)}{t} - 2\tilde{C}\frac{\|\mathrm{Id}_N - g\|^2}{t}\Big) d\mu_{\mathrm{SO}_N}(g) \right\} \\
&\overset{\text{Lemma 4.1}}{\geq} \frac{\tilde{C}^{-1}}{t^{N/2}} \exp\big(-2\tilde{C}\frac{|x_0-v|^2}{t}\big) \\
&\quad \times \left\{ \frac{1}{(1\wedge t)^{N(N-1)/4}} \int_{\mathrm{SO}_N(\mathbb{R})} \exp\Big(-\tilde{C}\frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N,g)}{t}\Big) d\mu_{\mathrm{SO}_N}(g) \right\} \\
&\overset{(4.1)}{\geq} \frac{\tilde{C}^{-1}}{t^{N/2}} \exp\big(-2\tilde{C}\frac{|x_0-v|^2}{t}\big),
\end{aligned}
$$

the constant $\tilde{C}$ being allowed to increase from line to line. Observe that in the last inequality, we have used (4.1) at time $\tilde{t} = t/(\beta\tilde{C})$, where $\beta\tilde{C} > 1$, together with the fact that

$$
\frac{1}{(1\wedge t)^{N(N-1)/4}} = \frac{1}{(1\wedge(\beta\tilde{C}\tilde{t}))^{N(N-1)/4}} \geq \frac{\beta^{-1}}{(\beta\tilde{C})^{N(N-1)/4}} \frac{\beta}{(1\wedge\tilde{t})^{N(N-1)/4}}.
$$

On the other hand, using once again Lemma 4.1 and (4.1) and choosing $\tilde{C}$ large enough:

$$
\begin{aligned}
f_{x_0}(t,v) &\leq \frac{\tilde{C}\beta}{t^{N/2}} \exp\Big(-\frac{\tilde{C}^{-1}}{2}\frac{|x_0-v|^2}{t}\Big) \\
&\quad \times \left\{ \frac{1}{(1\wedge t)^{N(N-1)/4}} \int_{\mathrm{SO}_N(\mathbb{R})} \exp\Big(-\frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N,g)}{\beta t} + \frac{\|\mathrm{Id}_N - g\|^2}{\tilde{C}t}\Big) d\mu_{\mathrm{SO}_N}(g) \right\} \\
&\leq \frac{\tilde{C}}{t^{N/2}} \exp\Big(-\frac{|x_0-v|^2}{\tilde{C}t}\Big),
\end{aligned}
$$

where we have chosen $\tilde{C}$ such that, for all $g \in \mathrm{SO}_N(\mathbb{R})$,

$$
\exp\Big(\Big\{-\frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N,g)}{\beta t} + \frac{\|\mathrm{Id}_N - g\|^2}{\tilde{C}t}\Big\}\Big) \leq \exp\Big(-\frac{d^2_{\mathrm{SO}_N}(\mathrm{Id}_N,g)}{2\beta t}\Big),
$$

and we have applied (4.1) at time $\tilde{t} = 2\beta^2 t$.

4.2.2 Non Gaussian Regime

We now look at the case $|x_0| \wedge |v| > 1$. Starting from (4.5) and Lemma 4.1, we aim at giving, for given $c > 0$, upper and lower bounds, homogeneous to those of (2.17), for the quantity $t^{-N/2} p_{x_0}(t,v)$, where :

$$
p_{x_0}(t,v) := (1\wedge t)^{-N(N-1)/4} \int_{\mathrm{SO}_N(\mathbb{R})} \exp\Big(-c\frac{\|\mathrm{Id}_N - g\|^2 + |v - gx_0|^2}{t}\Big) d\mu_{\mathrm{SO}_N}(g). \quad (4.6)
$$

Above, we notice that $|v - gx_0|^2 = |g^\top v - x_0|^2$ and $\|\mathrm{Id}_N - g\|^2 = \|\mathrm{Id}_N - g^\top\|^2$. Since the Haar measure is invariant by transposition, the roles of v and x_0 can be exchanged in formula (4.6) and we can assume that $|v| \geq |x_0|$.

By Remark 2.11 (with x_0 replaced by gx_0), we know that

$$
\big||v| - |x_0|\big|^2 + \big|\frac{|x_0|}{|v|}v - gx_0\big|^2 \leq |v - gx_0|^2 \leq 2\Big(\big||v| - |x_0|\big|^2 + \big|\frac{|x_0|}{|v|}v - gx_0\big|^2\Big).
$$

Radial cost. The term $\big||v| - |x_0|\big|^2$ is referred to as the radial cost. Since it is independent of g, we can focus on the other one, called the tangential cost. Then, changing v into $(|x_0|/|v|)v$, we can assume that $|v| = |x_0|$.

Tangential cost. We now assume that $|v| = |x_0|$. By rotation, we can assume that $x_0 = |x_0|e_1$. Then, we can write $v = |x_0|he_1$ for some $h \in \mathrm{SO}_N(\mathbb{R})$. We then expand in (4.6)

$$\|\mathrm{Id}_N - g\|^2 + |x_0|^2|he_1 - ge_1|^2 = |e_1 - ge_1|^2 + |x_0|^2|he_1 - ge_1|^2 + \sum_{i=2}^N |e_i - ge_i|^2.$$

The strategy is then quite standard and consists in reducing the quadratic form $|e_1 - ge_1|^2 + |x_0|^2|he_1 - ge_1|^2$. We write

$$\begin{aligned}
&|e_1 - ge_1|^2 + |x_0|^2|he_1 - ge_1|^2 \\
&= \big(1 + |x_0|^2\big)|ge_1|^2 - 2\langle ge_1, e_1 + |x_0|^2 he_1\rangle + 1 + |x_0|^2 \\
&= \big(1 + |x_0|^2\big)\Big|ge_1 - \frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|^2 - \frac{1}{1 + |x_0|^2}\big|e_1 + |x_0|^2 he_1\big|^2 + 1 + |x_0|^2.
\end{aligned}$$

Since,

$$\begin{aligned}
\frac{1}{1 + |x_0|^2}\big|e_1 + |x_0|^2 he_1\big|^2 - \big(1 + |x_0|^2\big) &= \frac{1}{1 + |x_0|^2}\Big(\big|e_1 + |x_0|^2 he_1\big|^2 - \big(1 + |x_0|^2\big)^2\Big) \\
&= -\frac{2|x_0|^2}{1 + |x_0|^2}\big(1 - \langle e_1, he_1\rangle\big),
\end{aligned}$$

we finally get that

$$\begin{aligned}
&|e_1 - ge_1|^2 + |x_0|^2|he_1 - ge_1|^2 \\
&= \big(1 + |x_0|^2\big)\Big|ge_1 - \frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|^2 + \frac{2|x_0|^2}{1 + |x_0|^2}\big(1 - \langle e_1, he_1\rangle\big).
\end{aligned}$$

As the second term is independent of g, we write

$$\begin{aligned}
p_{x_0}(t, v) &= (1 \wedge t)^{-N(N-1)/4} \exp\Big(-\frac{c}{t}\frac{2|x_0|^2}{1 + |x_0|^2}\big(1 - \langle e_1, he_1\rangle\big)\Big) \\
&\times \int_{\mathrm{SO}_N(\mathbb{R})} \exp\Big(-\frac{c}{t}\big(1 + |x_0|^2\big)\Big|ge_1 - \frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|^2 - \frac{c}{t}\sum_{i=2}^N |e_i - ge_i|^2\Big) d\mu_{\mathrm{SO}_N}(g).
\end{aligned} \tag{4.7}$$

Now, we notice that

$$\Big|\frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big| \leq 1.$$

Since $|x_0|^2 > 1$, we have $|e_1 + |x_0|^2 he_1| > 0$. Therefore, we can proceed as in the previous paragraph: in the first term inside the second exponential in (4.7), we use Remark 2.11 to split the radial and tangential costs. The radial cost is here given by

$$\begin{aligned}
\Big(1 - \Big|\frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|\Big)^2 &= \Big(1 - \Big|\frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|^2\Big)^2\Big(1 + \Big|\frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|\Big)^{-2} \\
&= \Big[1 - \Big(1 - \frac{2|x_0|^2}{(1 + |x_0|^2)^2}\big(1 - \langle e_1, he_1\rangle\big)\Big)\Big]^2\Big(1 + \Big|\frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|\Big)^{-2} \\
&= \big[\frac{2|x_0|^2}{(1 + |x_0|^2)^2}\big]^2\big(1 - \langle e_1, he_1\rangle\big)^2\Big(1 + \Big|\frac{e_1 + |x_0|^2 he_1}{1 + |x_0|^2}\Big|\Big)^{-2}.
\end{aligned}$$

Using the fact that $|x_0|^2 \le 1+|x_0|^2$, the first factor in the last line above is less than $4(1+|x_0|^2)^{-2}$. Similarly, the second factor is less than $1-\langle e_1, he_1\rangle$ since $|\langle e_1, he_1\rangle| \le 1$. Finally, the last factor is obviously less than 1. We deduce that, up to a multiplicative constant, the last term above can be upper bounded by $(1+|x_0|^2)^{-2}(1-\langle e_1, he_1\rangle)$. In particular, up to a modification of the constant c in $p_{x_0}(t,v)$, we can see the radial cost as a part of the exponential pre-factor in (4.7). Therefore, without any ambiguity, we can slightly modify the definition of $p_{x_0}(t,v)$ and assume that it writes

$$
\begin{aligned}
p_{x_0}(t,v) &= (1\wedge t)^{-N(N-1)/4} \exp\biggl(-\frac{c}{t}\frac{2|x_0|^2}{1+|x_0|^2}\bigl(1-\langle e_1, he_1\rangle\bigr)\biggr) \\
&\times \int_{\mathrm{SO}_N(\mathbb{R})} \exp\biggl(-\frac{c}{t}\bigl(1+|x_0|^2\bigr)\Bigl|ge_1 - \frac{e_1+|x_0|^2 he_1}{|e_1+|x_0|^2 he_1|}\Bigr|^2 - \sum_{i=2}^N \frac{|e_i - ge_i|^2}{t}\biggr) d\mu_{\mathrm{SO}_N}(g).
\end{aligned}
\tag{4.8}
$$

Equation (4.4) now yields:

$$
\begin{aligned}
p_{x_0}(t,v) &= \frac{1}{(1\wedge t)^{N(N-1)/4}} \exp\biggl(-\frac{c}{t}\frac{2|x_0|^2}{1+|x_0|^2}\bigl(1-\langle e_1, he_1\rangle\bigr)\biggr) \\
&\times \int \exp\biggl(-\frac{c}{t}\biggl\{\bigl(1+|x_0|^2\bigr)\bigl|s-\bar{s}\bigr|^2 + \sum_{i=2}^N |e_i - V_s L_k e_i|^2\biggr\}\biggr) d\nu_{\mathbb{S}^{N-1}}(s) d\mu_{\mathrm{SO}_{N-1}}(k),
\end{aligned}
\tag{4.9}
$$

the integral being defined on $\mathbb{S}^{N-1}\times \mathrm{SO}_{N-1}(\mathbb{R})$, with

$$
\bar{s} = (e_1+|x_0|^2 he_1)/(|e_1+|x_0|^2 he_1|). \tag{4.10}
$$

Lower bound. Observe first that, for all $i\in\{2,\cdots,N\}$, $|e_i - V_s L_k e_i|^2 \le 2(|(\mathrm{Id}_N - V_s)e_i|^2 + |e_i - L_k e_i|^2)$, using that V_s defines an isometry for the last control. From Lemma 4.1, we now derive that $\sum_{i=2}^N |e_i - L_k e_i|^2 \le c_1 \|\mathrm{Id}_{N-1} - k\|^2 \le c_2 d^2_{\mathrm{SO}_{N-1}}(\mathrm{Id}_{N-1}, k)$, where $(c_1,c_2) := (c_1,c_2)(N)$. By (4.1), applied for $N-1$, we get that there exists $C := C(N)\ge 1$ (the value of which is allowed to increase below) such that

$$
\frac{1}{(1\wedge t)^{(N-1)(N-2)/4}} \int_{\mathrm{SO}_{N-1}(\mathbb{R})} \exp\left(-2cc_2 \frac{d^2_{\mathrm{SO}_{N-1}}(\mathrm{Id}_{N-1},k)}{t}\right) d\mu_{\mathrm{SO}_{N-1}}(k) \ge C^{-1}.
$$

Thus,

$$
\begin{aligned}
p_{x_0}(t,v) &\ge \frac{1}{C(1\wedge t)^{(N-1)/2}} \exp\biggl(-\frac{c}{t}\frac{2|x_0|^2}{1+|x_0|^2}\bigl(1-\langle e_1, he_1\rangle\bigr)\biggr) \\
&\times \int_{\mathbb{S}^{N-1}} \exp\biggl(-\frac{c}{t}\biggl\{\bigl(1+|x_0|^2\bigr)\bigl|s-\bar{s}\bigr|^2 + 2\sum_{i=2}^N |e_i - V_s e_i|^2\biggr\}\biggr) d\nu_{\mathbb{S}^{N-1}}(s).
\end{aligned}
$$

Let us restrict the integral to a neighborhood of $\bar{s}$ in $\mathbb{S}^{N-1}$ of the form

$$
\mathcal{V}_{\bar{s}} := \{s : \exists\, R\in \mathrm{SO}_N(\mathbb{R}),\ s = R\bar{s},\ \|R-\mathrm{Id}_N\| \le t^{1/2}/|x_0|\}. \tag{4.11}
$$

Then,

$$
\begin{aligned}
p_{x_0}(t,v) &\ge \frac{1}{C(1\wedge t)^{(N-1)/2}} \exp\biggl(-\frac{c}{t}\frac{2|x_0|^2}{1+|x_0|^2}\bigl(1-\langle e_1, he_1\rangle\bigr)\biggr) \\
&\times \int_{\mathcal{V}_{\bar{s}}} \exp\biggl(-\frac{2c}{t}\sum_{i=2}^N |e_i - V_s e_i|^2\biggr\}\biggr) d\nu_{\mathbb{S}^{N-1}}(s).
\end{aligned}
\tag{4.12}
$$

As the set of the s's such that $\langle s, e_1\rangle = 0$ is of zero measure, we can restrict the integral to the set of $s \in \mathbb{S}^{N-1}$ such that $\langle s, e_1\rangle \neq 0$. By construction (see (4.3)), $V_s e_i = s_i$, for $i \in \{1, \dots, N\}$, with $s_1 = s$ and

$$u_i = e_i - \sum_{k=1}^{i-1} \langle e_i, s_k\rangle s_k,\ s_i := \frac{u_i}{|u_i|},\ i \in \{2, \cdots, N\},\ u_1 = s. \tag{4.13}$$

We can write, for $i \in \{2, \dots, N\}$,

$$\begin{aligned} |e_i - V_s e_i|^2 = |e_i - s_i|^2 &\le 2\big(|e_i - u_i|^2 + |u_i - s_i|^2\big) \\ &= 2\big(|e_i - u_i|^2 + |1 - |u_i||^2\big) \\ &\le 2\big(|e_i - u_i|^2 + ||e_i| - |u_i||^2\big) \le 4\big(|e_i - u_i|^2\big). \end{aligned} \tag{4.14}$$

Now, by (4.13),

$$|e_i - u_i|^2 = \sum_{k=1}^{i-1} \langle e_i, s_k\rangle^2 = \langle e_i, s_1 \pm e_1\rangle^2 + \sum_{k=2}^{i-1} \langle e_i, s_k - e_k\rangle^2 \le |s_1 \pm e_1|^2 + \sum_{k=2}^{i-1} |s_k - e_k|^2. \tag{4.15}$$

Therefore, by (4.14) and (4.15) and by a standard induction, for all $i \in \{2, \dots, N\}$,

$$|e_i - s_i|^2 \le \bar{C}|e_1 \pm s_1|^2 = 2\bar{C}\big(1 \pm \langle e_1, s\rangle\big) = 2\bar{C}\frac{1 - \langle e_1, s\rangle^2}{1 \mp \langle e_1, s\rangle}. \tag{4.16}$$

In the above, we can always choose the sign in $\mp$ so that $1 \mp \langle e_1, s\rangle \ge 1$. Therefore, for all $i \in \{2, \dots, N\}$,

$$|e_i - s_i|^2 \le 2\bar{C}\big(1 - \langle e_1, s\rangle^2\big) = 2\bar{C}\sum_{k=2}^{N} \langle e_k, s\rangle^2. \tag{4.17}$$

Since, for $s \in \mathcal{V}_{\bar{s}}$, $|\langle s, e_k\rangle| \le |\langle \bar{s}, e_k\rangle| + t^{1/2}/|x_0|$, we deduce from (4.17):

$$\begin{aligned} \sum_{i=2}^{N} |e_i - s_i|^2 &\le \bar{C}\left(\sum_{i=2}^{N} |\langle \bar{s}, e_i\rangle|^2 + (N-1)t/|x_0|^2\right) \\ &\le \bar{C}\big(1 - \langle \bar{s}, e_1\rangle^2 + t/|x_0|^2\big) \le \bar{C}\big(2(1 - \langle \bar{s}, e_1\rangle) + t/|x_0|^2\big). \end{aligned}$$

We derive from (4.12) that

$$\begin{aligned} p_{x_0}(t, v) &\ge \frac{\nu_{\mathbb{S}^{N-1}}(\mathcal{V}_{\bar{s}})}{\bar{C}(1 \wedge t)^{(N-1)/2}} \exp\left(-\frac{\bar{C}}{t}\Big[\frac{2|x_0|^2}{1 + |x_0|^2}\big(1 - \langle e_1, h e_1\rangle\big) + \big(1 - \langle e_1, \bar{s}\rangle\big)\Big]\right) \\ &\ge \bar{C}^{-1}\delta_t^{N-1} \exp\left(-\frac{\bar{C}}{t}\Big[\frac{2|x_0|^2}{1 + |x_0|^2}\big(1 - \langle e_1, h e_1\rangle\big) + \big(1 - \langle e_1, \bar{s}\rangle\big)\Big]\right), \end{aligned}$$

denoting, as in Theorem 2.8, $\delta_t := \dfrac{1 \wedge \left(\frac{t^{1/2}}{|x_0|}\right)}{1 \wedge t^{1/2}}$ and using (4.11) for the last inequality.

Assume first that $\langle e_1, h e_1\rangle \le 0$. The above equation yields

$$p_{x_0}(t, v) \ge \bar{C}^{-1}\delta_t^{N-1} \exp\big(-\frac{\bar{C}}{t}\big) \ge \bar{C}^{-1}\delta_t^{N-1} \exp\Big(-\frac{\bar{C}}{t}\big(1 - \langle e_1, h e_1\rangle\big)\Big).$$

Recalling that, for the tangential cost analysis, we have assumed $|x_0| = |v|$, we derive

$$1 - \langle e_1, h e_1\rangle = 1 - \langle \frac{x_0}{|x_0|}, \frac{v}{|x_0|}\rangle = \frac{1}{2|x_0|^2}|x_0 - v|^2,$$

which gives the claim.

Assume now that $\langle e_1, he_1\rangle \geq 0$. It can be checked from the definition of $\bar{s}$ in (4.10) that $\langle e_1, \bar{s}\rangle \geq \langle e_1, he_1\rangle$ so that we eventually get:

$$p_{x_0}(t,v) \geq \bar{C}^{-1}\delta_t^{N-1}\exp\Big(-\frac{\bar{C}}{t}\big(1-\langle e_1, he_1\rangle\big)\Big).$$

We conclude by the same argument as above.

Upper bound. Going back to (4.9) and using the fact that $V_s \in \mathrm{SO}_N(\mathbb{R})$ for any $s \in \mathbb{S}^{N-1}$, we get

$$\begin{aligned} p_{x_0}(t,v) &= \frac{1}{(1\wedge t)^{N(N-1)/4}}\exp\Big(-\frac{c}{t}\frac{2|x_0|^2}{1+|x_0|^2}\big(1-\langle e_1, he_1\rangle\big)\Big) \\ &\times \int \exp\Big(-\frac{c}{t}\Big\{\big(1+|x_0|^2\big)\big|s-\bar{s}\big|^2 + \sum_{i=2}^{N}|\hat{s}_i - L_k e_i|^2\Big\}\Big) d\nu_{\mathbb{S}^{N-1}}(s) d\mu_{\mathrm{SO}_{N-1}}(k), \end{aligned} \tag{4.18}$$

where $\hat{s}_i = V_s^\top e_i,\ i \in \{2,\cdots,N\}$. We then focus on the integral with respect to k, namely

$$q_t(s) := (1\wedge t)^{-(N-1)(N-2)/4}\int_{\mathrm{SO}_{N-1}(\mathbb{R})}\exp\Big(-\frac{c}{t}\sum_{i=2}^{N}|\hat{s}_i - L_k e_i|^2\Big) d\mu_{\mathrm{SO}_{N-1}}(k),$$

for a given $s \in \mathbb{S}^{N-1}$, the normalization $(1\wedge t)^{(N-1)(N-2)/4}$ standing for the order of the volume of the ball of radius $t^{1/2}$ in $\mathrm{SO}_{N-1}(\mathbb{R})$. Denoting by $\hat{s}^{2,N}$ the $N-1$ square matrix made of the column vectors $((\hat{s}_2)_j)_{2\leq j\leq N}, \ldots, ((\hat{s}_N)_j)_{2\leq j\leq N}$, where $(\hat{s}_i)_j$ stands for the jth coordinate of $\hat{s}_i$, we get

$$\sum_{i=2}^{N}|\hat{s}_i - L_k e_i|^2 \geq \|\hat{s}^{2,N} - k\|^2.$$

Now, we distinguish two cases. For a given $\varepsilon > 0$ to be specified next, we first consider the case when $\|\hat{s}^{2,N} - k\| \geq \varepsilon$ for any $k \in \mathrm{SO}_{N-1}(\mathbb{R})$. Then, there exists a constant $c' := c'(\varepsilon) > 0$ such that $\|\hat{s}^{2,N} - k\| \geq c' d_{\mathrm{SO}_{N-1}}(\mathrm{Id}_{N-1}, k)$, so that (up to a modification of c)

$$q_t(s) \leq (1\wedge t)^{-(N-1)(N-2)/4}\int_{\mathrm{SO}_{N-1}(\mathbb{R})}\exp\Big(-\frac{c}{t}d^2_{\mathrm{SO}_{N-1}}(\mathrm{Id}_{N-1},k)\Big) d\mu_{\mathrm{SO}_{N-1}}(k) \leq \bar{C},$$

for a constant $\bar{C} := \bar{C}(N)$.

Let us now assume that there exists $k_0 \in \mathrm{SO}_{N-1}(\mathbb{R})$ such that $\|\hat{s}^{2,N} - k_0\| \leq \varepsilon$. By invariance by rotation of the Haar measure, we notice that $q_t(s)$ can be bounded by

$$\begin{aligned} q_t(s) &\leq (1\wedge t)^{-(N-1)(N-2)/4}\int_{\mathrm{SO}_{N-1}(\mathbb{R})}\exp\Big(-\frac{c}{t}\|\hat{s}^{2,N} - k_0 k\|^2\Big) d\mu_{\mathrm{SO}_{N-1}}(k) \\ &= (1\wedge t)^{-(N-1)(N-2)/4}\int_{\mathrm{SO}_{N-1}(\mathbb{R})}\exp\Big(-\frac{c}{t}\|k_0^\top \hat{s}^{2,N} - k\|^2\Big) d\mu_{\mathrm{SO}_{N-1}}(k). \end{aligned}$$

Letting $\tilde{s}^{2,N} := k_0^\top \hat{s}^{2,N}$, we notice that $\|\tilde{s}^{2,N} - \mathrm{Id}_{N-1}\| \leq \varepsilon$. This permits to define $\tilde{S}^{2,N} := \ln(\tilde{s}^{2,N})$ (provided ε is chosen small enough).

Again, we distinguish two cases, according to the value of the variable k in the integral. When $\|\tilde{s}^{2,N} - k\| \geq \varepsilon$, we can use the same trick as before and say that $\|\tilde{s}^{2,N} - k\| \geq c d_{\mathrm{SO}_N}(\mathrm{Id}_{N-1}, k)$. Repeating the computations, we get

$$(1\wedge t)^{-(N-1)(N-2)/4}\int_{\|\tilde{s}^{2,N}-k\|\geq\varepsilon}\exp\Big(-\frac{c}{t}\|\tilde{s}^{2,N} - k\|^2\Big) d\mu_{\mathrm{SO}_{N-1}}(k) \leq \bar{C}.$$

When $\|\tilde{s}^{2,N} - k\| \leq \varepsilon$, we have $\|\mathrm{Id}_{N-1} - k\| \leq 2\varepsilon$, so that k may be inverted by the logarithm and written as $k = \exp(K)$ for some antisymmetric matrix K of size $N-1$. By local Lipschitz property of the logarithm, we deduce that, for such a k (and for a new value of c'),

$$\|\tilde{s}^{2,N} - k\| \geq c' \|\tilde{S}^{2,N} - K\|.$$

We then denote $\tilde{H}^{2,N}$ the orthogonal projection of $\tilde{S}^{2,N}$ on $\mathcal{A}_{N-1}(\mathbb{R})$. We get

$$\|\tilde{s}^{2,N} - k\| \geq c' \|\tilde{H}^{2,N} - K\|.$$

Clearly, $\tilde{H}^{2,N}$ is in the neighborhood of 0. By local Lipschitz property of the exponential, we finally obtain (again, for a new value of c')

$$\|\tilde{s}^{2,N} - k\| \geq c' \|\exp(\tilde{H}^{2,N}) - k\|.$$

Letting $\tilde{h}^{2,N} := \exp(\tilde{H}^{2,N})$, we end up with

$$\begin{aligned}
&(1 \wedge t)^{-(N-1)(N-2)/4} \int_{\|\tilde{s}^{2,N} - k\| \leq \varepsilon} \exp\Big(-\frac{c}{t} \|\tilde{s}^{2,N} - k\|^2\Big) d\mu_{\mathrm{SO}_{N-1}}(k) \\
&\leq (1 \wedge t)^{-(N-1)(N-2)/4} \int_{\|\tilde{s}^{2,N} - k\| \leq \varepsilon} \exp\Big(-\frac{c}{t} d^2_{\mathrm{SO}_{N-1}}(\tilde{h}^{2,N}, k)\Big) d\mu_{\mathrm{SO}_{N-1}}(k),
\end{aligned}$$

where we have used Lemma 4.1 on $\mathrm{SO}_{N-1}(\mathbb{R})$ to get the second line. By a new rotation argument,

$$(1 \wedge t)^{-(N-1)(N-2)/4} \int_{\|\tilde{s}^{2,N} - k\| \leq \varepsilon} \exp\Big(-\frac{c}{t} d^2_{\mathrm{SO}_{N-1}}(\tilde{h}^{2,N}, k)\Big) d\mu_{\mathrm{SO}_{N-1}}(k) \leq \bar{C},$$

which shows that $q_t(s) \leq \bar{C}$.

Equation (4.18) thus yields:

$$\begin{aligned}
p_{x_0}(t, v) \leq \frac{\bar{C}}{(1 \wedge t)^{(N-1)/2}} \exp\bigg(-\frac{c}{t} \frac{2|x_0|^2}{1 + |x_0|^2} \big(1 - \langle e_1, h e_1 \rangle\big)\bigg) \\
\times \int_{\mathbb{S}^{N-1}} \exp\bigg(-\frac{c(1 + |x_0|^2)}{t} |s - \bar{s}|^2\bigg) d\nu_{\mathbb{S}^{N-1}}(s).
\end{aligned}$$

Observing now that there exists $\bar{c} > 1$ such that $\bar{c}^{-1}|s - \bar{s}| \leq d(s, \bar{s}) \leq \bar{c}|s - \bar{s}|$, where d stands for the Riemannian metric on the sphere $\mathbb{S}^{N-1}$, we then deduce from the heat kernel estimates in Stroock [21] that

$$\frac{1}{1 \wedge \Big(\frac{t^{1/2}}{1 + |x_0|}\Big)^{N-1}} \exp\bigg(-\frac{c(1 + |x_0|^2)}{t} |s - \bar{s}|^2\bigg) \leq \bar{C} p_{\mathbb{S}^{N-1}}\left(\frac{t}{1 + |x_0|^2}, s, \bar{s}\right),$$

where $p_{\mathbb{S}^{N-1}}$ stands for the heat kernel on $\mathbb{S}^{N-1}$. Since we have assumed $|x_0| \geq 1$ we finally derive up to a modification of $\bar{C}$:

$$p_{x_0}(t, v) \leq \bar{C} \delta_t^{N-1} \exp\bigg(-\frac{c}{t} \frac{2|x_0|^2}{1 + |x_0|^2} \big(1 - \langle e_1, h e_1 \rangle\big)\bigg) \leq \bar{C} \delta_t^{N-1} \exp\Big(-\frac{1 - \langle e_1, h e_1 \rangle}{\bar{C} t}\Big),$$

which gives an upper bound homogeneous to the lower bound and completes the proof.

5 The degenerate case

The strategy to complete the proof of Theorem 2.12 relies on an expansion of Z_t in terms of iterated integrals of the Brownian motion on the Lie algebra $\mathcal{A}_N(\mathbb{R})$ of

$\mathrm{SO}_N(\mathbb{R})$. In that framework, it is worth mentioning that we do not exploit anymore the underlying group structure. Instead, we explicitly make use of the *Euclidean* structure of $\mathcal{A}_N(\mathbb{R})$. Indeed the analysis relies on precise controls of events described by the whole trajectory of Z. We manage to handle the probability of those events by controlling the corresponding trajectories of the $\mathcal{A}_N(\mathbb{R})$-valued Brownian motion B. In that perspective, the heat kernel estimates (4.1) for the marginals of Z in $\mathrm{SO}_N(\mathbb{R})$ are not sufficient, as once again, the distribution of the whole path is needed to carry on the analysis.

5.1 Set-up

In the whole section, we will assume that degeneracy occurs along the first direction of the space, that is X_0 has the form:

$$X_0 = X_0^1 e_1,$$

where e_1 is the first vector for the canonical basis and X_0^1 is a square integrable real-valued random variable. Because of the isotropy of the original equation, this choice is not restrictive. To make things simpler, additionally to the centering assumption, recall $\mathrm{E}\big[X_0^1\big] = 0$, we will also suppose (without any loss of generality) that X_0^1 is reduced, that is

$$\mathrm{E}\big[\big(X_0^1\big)^2\big] = 1.$$

Given a real x_0^1, we will work under the conditional measure given $\{X_0^1 = x_0^1\}$, which we will still denote by P. Therefore, recalling (2.7) and (3.2), we will write in the whole section $(X_t)_{t\geq 0}$ as

$$X_t = Z_t(x_0^1 e_1) - Z_t \int_0^t Z_s^\top d\bar{B}_s, \quad t \geq 0, \tag{5.1}$$

which is understood as the conditional version of the original process $(X_t)_{t\geq 0}$ given the initial condition $X_0 = x_0^1 e_1$. In this framework, the typical scales of X_t in small time t are given by:

$$\mathrm{E}\left[|X_t^1 - (Z_t x_0^1 e_1)^1|^2\right] \sim_{t\to 0} t^2, \quad \mathrm{E}\left[|X_t^i - (Z_t x_0^1 e_1)^i|^2\right] \sim_{t\to 0} t, \quad 2 \leq i \leq N, \tag{5.2}$$

showing that the fluctuations of the density is t in the first component and $t^{1/2}$ and the other ones. Eq. (5.2) will be proved below.

5.2 Small time expansions

The key point in the whole analysis lies in small time expansions of the process $(Z_t)_{t\geq 0}$ and of the 'conditional covariance' matrix C_t in (2.13). The precise strategy is to expand both of them in small times, taking care of the tails of the remainders in the expansion (recalling that the covariance matrix is random). We thus remind the reader of the so-called Bernstein equality, that will play a major role in the whole proof, see e.g. Revuz and Yor [19]:

Proposition 5.1. *Let $(M_t)_{t\geq 0}$ be a continuous scalar martingale satisfying $M_0 = 0$. Then, for any $A > 0$ and $\sigma > 0$,*

$$\mathrm{P}\Big(M_t^* \geq A,\ \langle M\rangle_t \leq \sigma^2\Big) \leq 2\exp\big(-\frac{A^2}{2\sigma^2}\big),$$

where we have used the standard notation $M_t^* := \sup_{0\leq s\leq t} |M_s|$.

Remark 5.2 (Notation for supremums)**.** With a slight abuse of notation, for a process $(Y_t)_{t\geq 0}$ with values in $\mathbb{R}^\ell$, $\ell \geq 1$, we will denote $Y_t^* := \max_{i\in\{1,\dots,\ell\}}(Y_t^\ell)^*$. Identifying $\mathbb{R}^\ell \otimes \mathbb{R}^k$ with $\mathbb{R}^{\ell\times k}$, we will also freely use those notations for matrix valued processes.

5.2.1 Landau notations revisited

In order to express the remainders in the expansion of the covariance matrix in a quite simple way, we will make a quite intensive use of Landau notations, but in various forms:

Definition 5.3 (Landau notations)**.** *Given some $T > 0$, we let:*

(i) *Given a deterministic function $(\psi_t)_{0\le t\le T}$ (scalar, vector or matrix valued), we write $\psi_t = \mathcal{O}(t^\alpha)$, for some $\alpha \ge 0$ and for any $t \in [0,T]$ if there exists a constant $C := C(N,T)$ such that $|\psi_t| \le Ct^\alpha$.*

(ii) *Given a process $(\Psi_t)_{0\le t\le T}$ (scalar, vector or matrix valued), we write $\Psi_t = O(t^\alpha)$, for some $\alpha \ge 0$ and for any $t \in [0,T]$ if there exists a constant $C := C(N,T)$ such that $|\Psi_t| \le Ct^\alpha$ a.s. Moreover, we write $\Psi_t = O_{\mathrm{P}}(t^\alpha)$, for some $\alpha \ge 0$ and for any $t \in [0,T]$ if, for all $p \in \mathbb{N}^*$, there exists a constant $C := C(N,T,p)$ such that $\mathrm{E}[|\Psi_t|^p]^{1/p} \le Ct^\alpha$.*

5.2.2 Small time expansion of the Brownian motion on $\mathrm{SO}_N(\mathbb{R})$

Following the proof of Lemma 3.1, we then expand Z_t according to

$$Z_t = \exp\big(-(N-1)t\big)\left(\mathrm{Id}_N + B_t + S_t\right) = \exp\big(-(N-1)t\big)\left(\mathrm{Id}_N + B_t + \int_0^t dB_s B_s + R_t\right),$$

for $t \ge 0$, with

$$S_t = \int_0^t dB_s \int_0^s dB_r \tilde{Z}_r, \quad R_t = \int_0^t dB_s \int_0^s dB_r \int_0^r dB_u \tilde{Z}_u, \quad \tilde{Z}_t = \exp\big((N-1)t\big) Z_t.$$

Given some time horizon $T > 0$, the remainders $(S_t)_{0\le t\le T}$ and $(R_t)_{0\le t\le T}$ can be controlled as follows on $[0,T]$:

Lemma 5.4. *There exists $C := C(N,T) > 0$ such that, for all $t \in [0,T]$ and $y > 0$,*

$$\mathrm{P}\big(S_t^* \ge y\big) \le 4\exp\big(-\frac{y}{Ct}\big), \quad \mathrm{P}\big(R_t^* \ge y\big) \le 6\exp\big(-\frac{y^{2/3}}{Ct}\big).$$

Proof. Applying Bernstein's inequality componentwise and using the fact that $\|\tilde{Z}_r\| \le \exp((N-1)T)$ for $r \in [0,T]$, there exists a constant $C > 0$ such that, for all $t \in [0,T]$,

$$\mathrm{P}\left(\sup_{0\le s\le t}\left|\int_0^s dB_r \tilde{Z}_r\right| \ge y_1\right) \le 2\exp\left(-\frac{y_1^2}{Ct}\right), \tag{5.3}$$

for any $y_1 > 0$. By Bernstein's inequality again,

$$\mathrm{P}\left(S_t^* \ge y_2, \int_0^t \left|\int_0^s dB_r \tilde{Z}_r\right|^2 ds \le t y_1^2\right) \le 2\exp\left(-\frac{y_2^2}{Cty_1^2}\right), \tag{5.4}$$

for any $y_2 > 0$. Similarly,

$$\mathrm{P}\left(R_t^* \ge y_3, \int_0^t |S_s|^2 ds \le t y_2^2\right) \le 2\exp\left(-\frac{y_3^2}{Cty_2^2}\right), \tag{5.5}$$

for $y_3 > 0$. Choosing $y_2 = y$ and $y_1 = y^{1/2}$, we complete the proof of the first inequality by adding (5.3) and (5.4). Choosing $y_3 = y$, $y_2 = y^{2/3}$ and $y_1 = y^{1/3}$, we complete the proof of the second inequality by adding (5.3), (5.4) and (5.5). $\square$

What really counts in the sequel is the first column $(Z_t^{\cdot,1})$ of the matrix Z_t. By antisymmetry of the matrix-valued process $(B_t)_{t\geq 0}$, the entries of the column $(Z_t^{\cdot,1})$ write

$$\begin{aligned}
Z_t^{1,1} &= \exp[-(N-1)t]\left(1+\sum_{j=2}^N \int_0^t dB_s^{1,j}B_s^{j,1} + R_t^{1,1}\right) \\
&= 1 - \frac{N-1}{2}t - \frac{1}{2}\sum_{j=2}^N (B_t^{j,1})^2 + O\big(t^2 + t|B_t^{\cdot,1}|^2 + |R_t|\big), \qquad (5.6)\\
Z_t^{i,1} &= \exp[-(N-1)t]\left(B_t^{i,1}+S_t^{i,1}\right) = \big(1+\mathcal{O}(t)\big)\left(B_t^{i,1}+S_t^{i,1}\right), \quad i\in\{2,\dots,N\}.
\end{aligned}$$

5.2.3 Expression of the covariance matrix

By (2.13) and (3.3), we know

$$C_t = \int_0^t Z_t Z_s^\top \frac{d}{ds}\langle \bar{B}\rangle_s \big(Z_t Z_s^\top\big)^\top ds. \qquad (5.7)$$

By (3.3) and (3.6), we have

$$\frac{d}{ds}\langle \bar{B}\rangle_s = \Lambda_s = \frac{1}{N}\left[N-1+\exp(-2Ns)\right]\mathrm{Id}_N - \exp(-2Ns)e_1\otimes e_1. \qquad (5.8)$$

We then notice that C_t reads

$$C_t = \int_0^t \bar{Z}_s \Lambda_{t-s}\bar{Z}_s^\top ds,$$

where we have denoted $\bar{Z}_s := Z_t Z_{t-s}^\top,\ s\in[0,t]$. By the invariance in law of Lemma 3.2, we know that $(Z_s)_{0\leq s\leq t}$ and $(\bar{Z}_s)_{0\leq s\leq t}$ have the same law. In particular, noting that $Z_t = \bar{Z}_t$, the following identity in law holds:

$$(Z_t, C_t) \overset{(\text{law})}{=} (Z_t, \bar{C}_t), \qquad (5.9)$$

where

$$\bar{C}_t := \int_0^t Z_s \Lambda_{t-s} Z_s^\top ds. \qquad (5.10)$$

Proposition 2.3 thus yields:

$$f_{x_0}(t,v) = \mathrm{E}\left[(2\pi)^{-N/2}\det{}^{-1/2}(\bar{C}_t)\exp\left(-\frac{1}{2}\big\langle v - Z_t x_0, \bar{C}_t^{-1}\big(v-Z_t x_0\big)\big\rangle\right)\right]. \qquad (5.11)$$

Now, by (5.8) and (5.10), we can expand $\bar{C}_t$ into

$$\bar{C}_t := \int_0^t Z_s\big((1-2(t-s))\mathrm{Id}_N - (1-2N(t-s))e_1\otimes e_1\big)Z_s^\top ds + O(t^3). \qquad (5.12)$$

5.2.4 Expansion of the covariance matrix

We now expand the integrand that appears in (5.12).

$$\begin{aligned}
&Z_s\big((1-2(t-s))\mathrm{Id}_N - (1-2N(t-s))e_1\otimes e_1\big)Z_s^\top \\
&= (1-2(t-s))\mathrm{Id}_N - (1-2N(t-s))(Z_s e_1)\otimes(Z_s e_1) \qquad (5.13)\\
&= (1-2(t-s))\mathrm{Id}_N - (1-2N(t-s))Z_s^{\cdot,1}\otimes Z_s^{\cdot,1}.
\end{aligned}$$

ŐjĹ dit plus tot. By (5.6), the entries of $(Z_s^{\cdot,1}) \otimes (Z_s^{\cdot,1}) =: \mathcal{T}_s$ write for all $s \in [0,T]$ and all $i,j \in \{2,\cdots,N\}$,

$$\begin{aligned}
\mathcal{T}_s^{1,1} &= 1-(N-1)s - \sum_{j=2}^{N}\big(B_s^{j,1}\big)^2 + O\big(s^2 + |B_s^{\cdot,1}|^4 + |R_s| + |R_s|^2\big),\\
\mathcal{T}_s^{1,i} &= \mathcal{T}_s^{i,1} = \big(1+\mathcal{O}(s)\big)\big(B_s^{i,1}+S_s^{i,1}\big)\big(1+S_s^{1,1}\big)\\
&= B_s^{i,1} + O\big(|S_s^{-1,1}| + |S_s|^2 + |B_s^{i,1}||S_s| + s|B_s^{i,1}|\big),\\
\mathcal{T}_s^{i,j} &= \mathcal{T}_s^{j,i} = \big(1+\mathcal{O}(s)\big)\big(B_s^{i,1}+S_s^{i,1}\big)\big(B_s^{j,1}+S_s^{j,1}\big)\\
&= B_s^{i,1}B_s^{j,1} + O\big(|B_s^{\cdot,1}||S_s| + \big|S_s\big|^2 + s|B_s^{\cdot,1}|^2\big),
\end{aligned} \tag{5.14}$$

where we have used the identity $Z_s^{1,1} = (1+\mathcal{O}(s))(1+S_s^{1,1})$ in the second line and the notation $S_s^{-1,1} := (0, S_s^{2,1}, \ldots, S_s^{N,1})$ in the third one.

Denoting by $\bar{\mathcal{T}}_s := (1-2(t-s))\mathrm{Id}_N - (1-2N(t-s))\mathcal{T}_s$ the last term in (5.13), it can be expanded as

$$\begin{aligned}
\bar{\mathcal{T}}_s^{1,1} &= 2(N-1)(t-s) + (N-1)s + \sum_{j=2}^{N}\big(B_s^{j,1}\big)^2 + O\big(t^2 + ((B_t^{\cdot,1})^*)^4 + R_t^* + (R_t^*)^2\big),\\
\bar{\mathcal{T}}_s^{i,i} &= 1-(B_s^{i,1})^2 + O\big(t + (B_t^{\cdot,1})^* S_t^* + (S_t^*)^2 + t((B_t^{\cdot,1})^*)^2\big),\\
\bar{\mathcal{T}}_s^{1,i} &= \bar{\mathcal{T}}_s^{i,1} = -B_s^{i,1} + O\big(t + (S_t^{-1,1})^* + (S_t^*)^2 + (B_t^{\cdot,1})^* S_t^* + t(B_t^{\cdot,1})^*\big),\\
\bar{\mathcal{T}}_s^{i,j} &= \bar{\mathcal{T}}_s^{j,i} = -B_s^{i,1}B_s^{j,1} + O\big(t + (B_t^{\cdot,1})^* S_t^* + (S_t^*)^2 + t((B_t^{\cdot,1})^*)^2\big), \quad i \neq j,
\end{aligned}$$

for $(i,j) \in \{2,\ldots,N\}^2$. By (5.12) and by integration, we thus derive the following expansions for the entries of $\bar{C}_t$: for all $(i,j) \in \{2,\ldots,N\}^2$,

$$\begin{aligned}
(\bar{C}_t)^{1,1} &= \int_0^t \sum_{j=2}^{N}(B_s^{j,1})^2 ds + \frac{3}{2}(N-1)t^2 + tO\big(t^2 + ((B_t^{\cdot,1})^*)^4 + R_t^* + (R_t^*)^2\big),\\
(\bar{C}_t)^{i,i} &= t - \int_0^t (B_s^{i,1})^2 ds + tO\big(t + (B_t^{\cdot,1})^* S_t^* + (S_t^*)^2 + t((B_t^{\cdot,1})^*)^2\big),\\
(\bar{C}_t)^{1,i} &= (\bar{C}_t)^{i,1} = -\int_0^t B_s^{i,1} ds + tO\big(t + (S_t^{-1,1})^* + (S_t^*)^2 + (B_t^{\cdot,1})^* S_t^* + t(B_t^{\cdot,1})^*\big),\\
(\bar{C}_t)^{i,j} &= (\bar{C}_t)^{j,i} = -\int_0^t B_s^{i,1}B_s^{j,1} ds + tO\big(t + (B_t^{\cdot,1})^* S_t^* + (S_t^*)^2 + t((B_t^{\cdot,1})^*)^2\big), \quad i \neq j.
\end{aligned} \tag{5.15}$$

By (5.1), Eq. (5.2) follows from the bounds for $(\bar{C})_{1,1}$ and $(\bar{C}_{i,i})_{2\le i\le N}$.

5.3 Proof of the Lower Bound in Theorem 2.12

We start from the representation formula (5.11) derived from the identity in law (5.9). We insist here that we choose some 'untypical' events for the Brownian path on $\mathcal{A}_N(\mathbb{R})$ to derive the bounds of Theorem 2.12.

5.3.1 First Step

The point is to find some relevant scenarios to explain the typical behavior of $f_{x_0}(t,v)$ in (5.11). Given $\xi \in (0,1]$ such that $t/\xi^2 \le 1$ and $\gamma \in (0,1]$, we thus introduce the events

$$\mathcal{B}^1 = \bigcap_{j=2}^{N}\left\{\sup_{0\le s\le t}|B_s^{j,1} - \frac{s}{t}\xi| \le \gamma\frac{t}{\xi}\right\}, \quad \mathcal{B}^{i,j} = \left\{\sup_{0\le s\le t}\left|\int_0^s dB_r^{i,j}B_r^{j,1}\right| \le t\right\}, \tag{5.16}$$

for $i,j\in\{2,\dots,N\}$. We then let

$$\mathcal{B}=\mathcal{B}^1\cap\bigcap_{i,j=2}^N\mathcal{B}^{i,j}. \tag{5.17}$$

Lemma 5.5. *There exists a constant $c>0$ such that*

$$\mathrm{P}\left(\mathcal{B}\right)\geq c\gamma^{(N-1)/2}\Big(1\wedge(\frac{t^{1/2}}{\xi})\Big)^{N(N-1)/2}\exp\Big(-(N-1)\frac{\xi^2}{t}\Big).$$

Proof. On the event $\mathcal{B}^1$, it holds, for all $j\in\{2,\dots,N\}$,

$$(B_t^{j,1})^*=\sup_{0\leq s\leq t}|B_s^{j,1}|\leq\Big(\xi+\gamma\frac{t}{\xi}\Big)\leq 2\xi,$$

since we have $\gamma t/\xi^2\leq 1$.

By independence of $B^{j,i}$ and $B^{k,1}$ for $i,j,k\in\{2,\dots,N\}$, we also know that, conditionally on $\mathcal{B}^1$, the process $(\int_0^s dB_r^{i,j}B_r^{j,1})_{0\leq s\leq t}$ behaves as a Wiener integral, with a variance process less than $(4\xi^2 s)_{0\leq s\leq t}$. Therefore, using a Brownian change of time, we obtain

$$\mathrm{P}\left(\mathcal{B}^{i,j}|\mathcal{B}^1\right)\geq\mathrm{P}\big(\sup_{0\leq s\leq 4\xi^2 t}|\beta_s|\leq t\big),$$

where $(\beta_s)_{s\geq 0}$ is a 1D Brownian motion. We deduce that there exists a constant $c>0$ (which value is allowed to increase from line to line) such that

$$\mathrm{P}\left(\mathcal{B}^{i,j}|\mathcal{B}^1\right)\geq c\Big(1\wedge(\frac{t^{1/2}}{\xi})\Big).$$

In fact, we must bound from below the conditional probability $\mathrm{P}(\cap_{i,j=2}^N\mathcal{B}^{i,j}|\mathcal{B}^1)$.

By antisymmetry of the matrix B and conditional independence of the processes $(B^{i,j})_{2\leq i<j\leq N}$, we deduce that

$$\mathrm{P}\Big(\bigcap_{i,j=2}^N\mathcal{B}^{i,j}\,|\mathcal{B}^1\Big)=\prod_{2\leq i<j\leq N}\mathrm{P}\left(\mathcal{B}^{i,j}|\mathcal{B}^1\right)\geq c^{(N-1)(N-2)/2}\left(\min(1,\xi^{-1}t^{1/2})\right)^{(N-1)(N-2)/2}.$$

It thus remains to bound $\mathrm{P}(\mathcal{B}^1)$ from below. For some $j\in\{2,\dots,N\}$, we deduce from Girsanov's theorem that

$$\begin{aligned}\mathrm{P}\Big(\sup_{0\leq s\leq t}|B_s^{j,1}-\frac{s}{t}\xi|\leq\frac{\gamma t}{\xi}\Big)&=\mathrm{E}\Big[\exp\Big(-\frac{\xi}{t}\beta_t-\frac{\xi^2}{2t}\Big)\mathbf{1}_{\{\sup_{0\leq s\leq t}|\beta_s|\leq\gamma t/\xi\}}\Big]\\&\geq\exp\Big(-1-\frac{\xi^2}{2t}\Big)\mathrm{P}\Big(\sup_{0\leq s\leq t}|\beta_s|\leq\frac{\gamma t}{\xi}\Big)\geq c\gamma\Big(1\wedge(\frac{t^{1/2}}{\xi})\Big)\exp\Big(-\frac{\xi^2}{2t}\Big),\end{aligned}$$

where $(\beta_s)_{s\geq 0}$ is a 1D Brownian motion. By independence of the processes $(B^{1,j})_{2\leq j\leq N}$, we deduce that

$$\mathrm{P}\left(\mathcal{B}^1\right)\geq c^{N-1}\gamma^{N-1}\Big(1\wedge(\frac{t^{1/2}}{\xi})\Big)^{N-1}\exp\Big(-(N-1)\frac{\xi^2}{2t}\Big).$$

We finally deduce that

$$\mathrm{P}\left(\mathcal{B}\right)\geq c^{N(N-1)/2}\gamma^{N-1}\exp\Big(-(N-1)\frac{\xi^2}{2t}\Big)\Big(1\wedge(\frac{t^{1/2}}{\xi})\Big)^{N(N-1)/2}.$$

$\square$

5.3.2 Second Step

We now plug the analysis of the covariance matrix performed in §5.2.4 into the previous step: We compute the typical values of the conditional covariance matrix on the event $\mathcal{B} \cap \mathcal{R}$, where

$$\mathcal{R} = \{S_t^* \le \xi^{3/2}\} \cap \{R_t^* \le \xi^{9/4}\}, \tag{5.18}$$

so that, by Lemma 5.4, $\mathrm{P}(\mathcal{R}^\complement) \le c^{-1}\exp(-c\xi^{3/2}/t)$. Therefore,

$$\mathrm{P}(\mathcal{B} \cap \mathcal{R}) \ge c^{N(N-1)/2}\gamma^{N-1}\exp\bigl(-(N-1)\frac{\xi^2}{2t}\bigr)\Bigl(1 \wedge \bigl(\frac{t^{1/2}}{\xi}\bigr)\Bigr)^{N(N-1)/2} - c^{-1}\exp\bigl(-c\frac{\xi^{3/2}}{t}\bigr),$$

which proves that there exists a constant $C := C(N) \ge 1$ (which value is allowed to increase from line to line) such that

$$\mathrm{P}(\mathcal{B} \cap \mathcal{R}) \ge C^{-1}\gamma^{N-1}\exp\bigl(-2(N-1)\frac{\xi^2}{2t}\bigr) - c^{-1}\exp\bigl(-c\frac{\xi^{3/2}}{t}\bigr),$$

using the fact that $1 \vee (\xi/t^{1/2}) \le C\exp[\xi^2/(Nt)]$. Therefore, for ξ small enough,

$$\mathrm{P}(\mathcal{B} \cap \mathcal{R}) \ge C^{-1}\exp\bigl(-2(N-1)\frac{\xi^2}{2t}\bigr). \tag{5.19}$$

On $\mathcal{B} \cap \mathcal{R}$ (see (5.16) and (5.17) for the definitions of $\mathcal{B}$ and (5.18) for the definition of $\mathcal{R}$), we have

$$S_t^* \le \xi^{3/2},\ R_t^* \le \xi^{9/4},\ (S_t^{-1,1})^* \le t,\ (B_t^{\cdot,1})^* \le 2\xi,\ B_s^{1,i}B_s^{1,j} = (s/t)^2\xi^2 + O(\gamma t), \tag{5.20}$$

the last expansion holding true for all $i, j \in \{2, \dots, N\}$ and following from the fact that $\gamma^2 t^2/\xi^2 \le \gamma t$. Recall indeed from the first lines in §5.3.1 that $\xi, \gamma \in (0,1]$ and $t/\xi^2 \le 1$.

We deduce from (5.15) that, for all $i, j \in \{2, \dots, N\}$, on $\mathcal{B} \cap \mathcal{R}$,

$$\begin{aligned}
(\bar{C}_t)^{1,1} &= t(N-1)\frac{\xi^2}{3} + O\bigl(t^2 + t\xi^{9/4}\bigr),\\
(\bar{C}_t)^{i,i} &= t\bigl(1 - \frac{\xi^2}{3}\bigr) + O\bigl(t^2 + t\xi^{9/4}\bigr),\\
(\bar{C}_t)^{1,i} &= (\bar{C}_t)^{i,1} = -t\frac{\xi}{2} + O\bigl(\gamma t\xi + t^2 + t\xi^{9/4}\bigr),\\
(\bar{C}_t)^{i,j} &= (\bar{C}_t)^{j,i} = -t\frac{\xi^2}{3} + O\bigl(t^2 + t\xi^{9/4}\bigr), \quad i \ne j,
\end{aligned}$$

the $O(\gamma t)$ in the third expansion following from the fact that $\gamma t^2/\xi = \gamma t\xi(t/\xi^2) \le \gamma t\xi$. Therefore, we can write, on $\mathcal{B} \cap \mathcal{R}$,

$$\begin{aligned}
\bar{C}_t &= \bar{C}_t^0 + O\left(t^2 + t\xi^{9/4}\right),\\
\bar{C}_t^0 &= t\ \mathrm{diag}(\xi, 1, \dots, 1)\bar{C}^{00}\mathrm{diag}(\xi, 1, \dots, 1),
\end{aligned} \tag{5.21}$$

where $\mathrm{diag}(\xi, 1, \dots, 1)$ denotes the diagonal matrix with $(\xi, 1, \dots, 1)$ as diagonal and where, for every $i \in \{2, \dots, N\}$,

$$\begin{aligned}
&(\bar{C}^{00})^{1,1} = \frac{N-1}{3}, \quad (\bar{C}_t^{00})^{i,i} = \bigl(1 - \frac{\xi^2}{3}\bigr),\\
&(\bar{C}^{00})^{1,i} = (\bar{C}^{00})^{i,1} = -\frac{1}{2} + \mathcal{O}(\gamma), \quad (\bar{C}^{00})^{i,j} = (\bar{C}^{00})^{j,i} = -\frac{\xi^2}{3}, \quad i \ne j.
\end{aligned}$$

5.3.3 Third Step

We go thoroughly into the analysis of $\bar{C}^0_t$. When $\xi = \gamma = 0$, the determinant of $\bar{C}^{00}$ can be computed explicitly by adding $1/2$ times the column i to the first column, for any $i = 2, \dots, N$. We obtain as a result

$$\left[\det(\bar{C}^{00})\right]_{|\xi=\gamma=0} = \frac{N-1}{12}.$$

We deduce that

$$\det(\bar{C}^0_t) = t^N \xi^2 \left[\frac{N-1}{12} + \mathcal{O}\left(\gamma + \xi^2\right)\right]. \tag{5.22}$$

In a similar way,

$$\left(\bar{C}^0_t\right)^{-1} = t^{-1}\mathrm{diag}(1/\xi, 1, \dots, 1) \left(\bar{C}^{00}\right)^{-1} \mathrm{diag}(1/\xi, 1, \dots, 1), \tag{5.23}$$

where, for γ and ξ^2 small enough,

$$\left(\bar{C}^{00}\right)^{-1} = \left[\left(\bar{C}^{00}\right)^{-1}\right]_{|\xi=\gamma=0} \left(\mathrm{Id}_N + \mathcal{O}(\gamma + \xi^2)\right),$$

with

$$\left[\left(\bar{C}^{00}\right)^{-1}\right]_{|\xi=\gamma=0} = \mathcal{O}(1), \tag{5.24}$$

so that, by (5.23), $(\bar{C}^0_t)^{-1} = \mathcal{O}(t^{-1}\xi^{-2})$. Therefore, referring to (5.21), we write

$$\bar{C}_t = \bar{C}^0_t + M_t, \tag{5.25}$$

with $M_t = O(t^2 + t\xi^{9/4})$ on $\mathcal{B} \cap \mathcal{R}$, and we let

$$\mathrm{Id}_N + M'_t := \left(\bar{C}^0_t\right)^{1/2} \left(\bar{C}^0_t + M_t\right)^{-1} \left(\bar{C}^0_t\right)^{1/2},$$

where the exponent $1/2$ indicates the symmetric square root. Indeed, when $\gamma = \xi = 0$, $\bar{C}^{00}$ is the covariance matrix of the vector

$$\left(\left(\frac{N-1}{12}\right)^{1/2} \zeta_1 - \sum_{i=2}^{N} \frac{\zeta_i}{2}, \zeta_2, \dots, \zeta_N\right),$$

with $(\zeta_1, \dots, \zeta_N) \overset{(\mathrm{law})}{=} \mathcal{N}^{\otimes N}(0,1)$, so that it is a non-negative symmetric matrix; since its determinant is positive, it is a positive symmetric matrix. By continuity, this remains true for ξ and γ small enough. For the same values of ξ and γ, (5.21) says that $\bar{C}^0_t$ is also symmetric and positive. Then,

$$\left(\bar{C}^0_t + M_t\right)^{-1} = \left(\bar{C}^0_t\right)^{-1/2} \left(\mathrm{Id}_N + M'_t\right) \left(\bar{C}^0_t\right)^{-1/2}. \tag{5.26}$$

As $M_t(\bar{C}^0_t)^{-1} = O(t/\xi^2 + \xi^{1/4})$ is small when t/ξ^2 and ξ are small, we can write

$$\begin{aligned}
\mathrm{Id}_N + M'_t &= \left(\bar{C}^0_t\right)^{1/2} \left(\bar{C}^0_t\right)^{-1} \left[\mathrm{Id}_N + M_t \left(\bar{C}^0_t\right)^{-1}\right]^{-1} \left(\bar{C}^0_t\right)^{1/2} \\
&= \left(\bar{C}^0_t\right)^{-1/2} \sum_{n \geq 0} \left[-M_t \left(\bar{C}^0_t\right)^{-1}\right]^n \left(\bar{C}^0_t\right)^{1/2} \\
&= \mathrm{Id}_N + \sum_{n \geq 0} (-1)^{n+1} \left(\bar{C}^0_t\right)^{-1/2} \left[M_t \left(\bar{C}^0_t\right)^{-1}\right]^n M_t \left(\bar{C}^0_t\right)^{-1/2}.
\end{aligned}$$

By (5.23), $(\bar{C}_t^0)^{-1/2} = O(t^{-1/2}\xi^{-1})$, so that $M_t(\bar{C}_t^0)^{-1/2} = O(t^{3/2}\xi^{-1}+t^{1/2}\xi^{5/4})$. Therefore,

$$\sum_{n\geq 0}(-1)^{n+1}\left(\bar{C}_t^0\right)^{-1/2}\left[M_t\left(\bar{C}_t^0\right)^{-1}\right]^n M_t\left(\bar{C}_t^0\right)^{-1/2} = \sum_{n\geq 0}\left[O(t/\xi^2+\xi^{1/4})\right]^{n+1}$$
$$= O(t/\xi^2+\xi^{1/4}),$$

provided t/ξ^2 and ξ are small enough.

Therefore, for t/ξ^2 and ξ small enough, the matrix $\mathrm{Id}_N + M_t'$, which is symmetric by construction, has all its eigenvalues between $1/2$ and 2, so that, for given a vector $v = (v_1,\dots,v_N)^\top$, (5.26) yields

$$\begin{aligned}\langle v, \left(\bar{C}_t^0 + M_t\right)^{-1} v\rangle &\leq C\big\langle \left(v_1,v_2,\dots,v_N\right)^\top, \left(\bar{C}_t^0\right)^{-1}\left(v_1,v_2,\dots,v_N\right)^\top\big\rangle \\ &\leq Ct^{-1}\Big\langle \Big(\frac{v_1}{\xi},v_2,\dots,v_N\Big)^\top, \left(\bar{C}_t^{00}\right)^{-1}\Big(\frac{v_1}{\xi},v_2,\dots,v_N\Big)^\top\Big\rangle \\ &\leq Ct^{-1}\bigg(\frac{v_1^2}{\xi^2}+\sum_{i=2}^N v_i^2\bigg).\end{aligned} \tag{5.27}$$

5.3.4 Final Step

We can summarize what we have proved in the following way: There exists a constant $K := K(N) \geq 1$ such that, for $\max(t/\xi^2,\xi^2,\gamma) \leq 1/K$, Eq. (5.27) holds for any $(v_1,\dots,v_N) \in \mathbb{R}^N$ on the event $\mathcal{B}\cap\mathcal{R}$.

The point is now to plug $(v^1 - Z_t^{1,1}x_0^1, v^2 - Z_t^{2,1}x_0^1,\dots,v^N - Z_t^{N,1}x_0^1)$ instead of $(v^1,\dots,v^N)$ in (5.27). Put it differently, we are to bound:

$$\inf_{Kt\leq\xi^2\leq 1/K} I(t,x_0,v,\xi), \quad I(t,x_0,v,\xi) := \Big[\frac{|v^1 - Z_t^{1,1}x_0^1|^2}{\xi^2} + \sum_{i=2}^N |v^i - Z_t^{i,1}x_0^1|^2\Big]. \tag{5.28}$$

By (5.6) and (5.20), on $\mathcal{B}\cap\mathcal{R}$,

$$\begin{aligned}Z_t^{1,1} &= 1 + O\big(t+\xi^2\big) = 1 + O(\xi^2),\\ Z_t^{i,1} &= \big(1+\mathcal{O}(t)\big)\big(\frac{\gamma t}{\xi}+\xi+t\big) = O(\xi),\end{aligned}$$

where we have used $t\leq\xi^2$ in both expansions. Pay attention that this step is crucial as, together with the previous paragraph, it gives the joint behavior of $(Z_t^{\cdot,1},\bar{C}_t)$ on $\mathcal{B}\cap\mathcal{R}$.

Therefore, we can find a constant $C := C(N) > 0$ such that

$$\begin{aligned}\big|v^1 - Z_t^{1,1}x_0^1\big| &\leq \big|v^1 - x_0^1\big| + C\xi^2|x_0^1|,\\ \big|v^i - Z_t^{i,1}x_0^1\big| &\leq |v^i| + C\xi|x_0^1|, \quad i\neq 1.\end{aligned} \tag{5.29}$$

The value of C being allowed to increase from line to line, we get:

$$I(t,x_0,v,\xi) \leq C\bigg[\frac{|v^1 - x_0^1|^2}{\xi^2} + |x_0^1|^2\xi^2 + \sum_{i=2}^N |v^i|^2\bigg].$$

We now handle the minimization problem in (5.28) according to the value of

$$\varsigma := \frac{|v^1 - x_0^1|}{1\vee|x_0^1|}$$

If $\varsigma \leq Kt$, we choose $\xi^2 = Kt$ in the infimum. We obtain

$$\inf_{Kt\leq\xi^2\leq 1/K} I(t,x_0,v,\xi) \leq C\bigg(2Kt\big(1\vee|x_0^1|^2\big) + \sum_{i=2}^N |v^i|^2\bigg).$$

If $\varsigma \geq 1/K$, we choose $\xi^2 = 1/K$ in the infimum. We obtain

$$\inf_{Kt \leq \xi^2 \leq 1/K} I(t, x_0, v, \xi) \leq C\biggl(2K|v^1 - x_0^1|^2 + \sum_{i=2}^{N} |v^i|^2\biggr).$$

If $\varsigma \in [Kt, 1/K]$, we choose $\xi^2 = \varsigma$ in the infimum. We obtain

$$\inf_{Kt \leq \xi^2 \leq 1/K} I(t, x_0, v, \xi) \leq C\biggl(2\bigl(1 \vee |x_0^1|\bigr)|v^1 - x_0^1| + \sum_{i=2}^{N} |v^i|^2\biggr).$$

This gives a lower bound for the exponential factor in (5.27) on the event $\mathcal{B} \cap \mathcal{R}$. When $x_0 \in [-C_0, C_0]$, we can modify C (allowing it to depend on C_0) in such a way that, in any of three cases,

$$\inf_{Kt \leq \xi^2 \leq 1/K} I(t, x_0, v, \xi) \leq C\biggl(|v^1 - x_0^1| + |v^1 - x_0^1|^2 + \sum_{i=2}^{N} |v^i|^2\biggr), \tag{5.30}$$

which fits the off-diagonal cost in the statement of Theorem 2.12. Notice that the dependence of C upon C_0 can be made explicit.

It remains to discuss the diagonal rate. By (5.22) and (5.26), on $\mathcal{B} \cap \mathcal{R}$,

$$\det(\bar{C}_t) = \det(\bar{C}_t^0 + M) \leq C' t^N \xi^2 = C' t^{N+1}(\xi^2/t),$$

for some constant C'. Now,

$$\frac{\xi^2}{t}\begin{cases} = K & \text{if } \varsigma \leq Kt, \\ = \dfrac{\varsigma}{t} \leq \exp\Bigl(\dfrac{|v^1 - x_0^1|}{t}\Bigr) & \text{if } \varsigma \in [Kt, 1/K], \\ = \dfrac{1}{Kt} \leq \dfrac{\varsigma}{t} \leq \exp\Bigl(\dfrac{|v^1 - x_0^1|}{t}\Bigr) & \text{if } \varsigma \geq 1/K. \end{cases}$$

Therefore, modifying C' if necessary,

$$\bigl[\det(\bar{C}_t)\bigr]^{-1/2} \geq (C')^{-1/2} t^{-(N+1)/2} \exp\Bigl(-\frac{|v^1 - x_0^1|}{t}\Bigr). \tag{5.31}$$

In the same way, (5.19) implies

$$\mathrm{P}\,(\mathcal{B} \cap \mathcal{R}) \geq (C')^{-1} \exp\Bigl(-C' \frac{|v^1 - x_0^1|}{t}\Bigr). \tag{5.32}$$

By (5.11), (5.27), (5.30), (5.31) and (5.32), we complete the proof of the lower bound. Indeed, for $x_0 \in [-C_0, C_0]$ and $C := C(N, C_0)$,

$$f_{x_0}(t, v) \geq \frac{1}{Ct^{(N+1)/2}} \exp\biggl(-C\biggl[\frac{|v^1 - x_0^1|}{t} + \frac{|v^1 - x_0^1|^2}{t} + \sum_{i=2}^{N} \frac{|v^i|^2}{t}\biggr]\biggr).$$

5.4 Proof of the Upper Bound in Theorem 2.12

Let us restart from the expression of the conditional density given by Proposition 2.3 that we recall here. For all $(t, x_0, v) \in \mathbb{R}^{+*} \times (\mathbb{R}^N)^2$ we have:

$$f_{x_0}(t, v) = \mathrm{E}\left[\frac{1}{(2\pi)^{N/2} \det(C_t)^{1/2}} \exp\left(-\tfrac{1}{2}\langle C_t^{-1}(v - Z_t x_0), v - Z_t x_0\rangle\right)\right].$$

In order to handle the degeneracy in the first coordinate, we introduce the *rescaled* covariance matrix (pay attention that the notation M below has nothing to do with the one used in (5.25))

$$M_t := t^{-1}\mathbb{T}_t^{-1}C_t\mathbb{T}_t^{-1}, \tag{5.33}$$

where $\mathbb{T}_t$ is the $N \times N$-diagonal matrix:

$$\mathbb{T}_t := \mathrm{diag}(t^{1/2}, 1, \dots, 1), \tag{5.34}$$

the matrix $t^{1/2}\mathbb{T}_t$ expressing the different scales in the fluctuations of the system, as emphasized in (5.2). Writing $C_t = t^{1/2}\mathbb{T}_t M_t(t^{1/2}\mathbb{T}_t)$ in the definition of $f_{x_0}(t,v)$, we deduce that

$$\begin{aligned} f_{x_0}(t,v) = \mathrm{E}\biggl[&\frac{1}{(2\pi)^{N/2}\mathrm{det}(M_t)^{1/2}t^{(N+1)/2}} \\ &\times \exp\Bigl(-\tfrac{1}{2}\bigl\langle M_t^{-1}\bigl[t^{-1/2}\mathbb{T}_t^{-1}(v - Z_t x_0)\bigr], t^{-1/2}\mathbb{T}_t^{-1}(v - Z_t x_0)\bigr\rangle\Bigr)\biggr]. \end{aligned} \tag{5.35}$$

This representation makes the explosion rate of the density along the diagonal appear, provided the determinant of the matrix M_t is well-controlled as t tends to 0. In order to get the off-diagonal decay of the density, we have in mind to perform a Gaussian integration by parts, in its most direct version, in order to bound the density by the tails of the marginal distributions of the process

$$Z_t\Gamma_t,\ \Gamma_t := \int_0^t Z_s^\top d\bar{B}_s.$$

Such a strategy is inspired from the approach based on Malliavin calculus for estimating densities, see e.g. Kusuoka and Stroock [18], but here we take benefit of the underlying Gaussian structure to make the integration by parts directly and thus avoid any further reference to Malliavin calculus.

5.4.1 Main step

We now establish the upper bound of Theorem 2.12 for $f_{x_0}(t,.)$ in (5.35). Rewrite first

$$f_{x_0}(t,v) = \frac{1}{t^{(N+1)/2}}\mathrm{E}\bigl[p_t\bigl(t^{-1/2}\mathbb{T}_t^{-1}(v - Z_t x_0)\bigr)\bigr], \tag{5.36}$$

where

$$p_t(y) := \frac{1}{(2\pi)^{N/2}\mathrm{det}(M_t)^{1/2}}\exp\bigl(-\tfrac{1}{2}\langle M_t^{-1}y, y\rangle\bigr), \quad y \in \mathbb{R}^N,$$

stands for the conditional density at time t of $t^{-1/2}\mathbb{T}_t^{-1}Z_t\Gamma_t$ given the σ-field $\mathcal{F}_t^Z := \sigma((Z_u)_{0\le u\le t})$. Note indeed that, conditional on $\mathcal{F}_t^Z$, $Z_t\Gamma_t$ is an N-dimensional Gaussian random variable with zero mean and covariance C_t given by (5.7), so that the conditional covariance matrix of $t^{-1/2}\mathbb{T}_t^{-1}Z_t\Gamma_t$ given $\mathcal{F}_t^Z$ is exactly M_t defined in (5.33) (pay attention that M_t is random).

Since p_t is smooth, we directly have

$$\begin{aligned} p_t(y) &= (-1)^N\int_{\prod_{i=1}^N\{\mathrm{sign}(y_i)z_i>|y_i|\}}\partial_{z_1,\cdots,z_N}p_t(z)dz \\ &= \frac{(-1)^N}{(2\pi)^{N/2}\mathrm{det}(M_t)^{1/2}}\int_{\prod_{i=1}^N\{\mathrm{sign}(y_i)z_i>|y_i|\}}\partial_{z_1,\cdots,z_N}\Bigl\{\exp\bigl(-\tfrac{1}{2}\langle M_t^{-1}z, z\rangle\bigr)\Bigr\}dz. \end{aligned} \tag{5.37}$$

Let now, for any $1 \le i \le N$,

$$P_t^i(z) := \bigl(\partial_{z_i,\cdots,z_1}\bigl\{\exp\bigl(-\tfrac{1}{2}\langle M_t^{-1}z, z\rangle\bigr)\bigr\}\bigr)\exp\bigl(\tfrac{1}{2}\langle M_t^{-1}z, z\rangle\bigr), \quad z \in \mathbb{R}^N,$$

which is a polynomial of the variable z with degree i. Similarly to the Hermite polynomials, it can be defined by induction

$$\begin{aligned}&\forall z \in \mathbb{R}^N,\ P_t^1(z) = -(M_t^{-1}z)_1,\\&\forall i \in \{2,\cdots,N\}, \forall z \in \mathbb{R}^N,\ P_t^i(z) = \partial_{z_i} P_t^{i-1}(z) - (M_t^{-1}z)_i P_t^{i-1}(z).\end{aligned} \tag{5.38}$$

The highest order term in $P_t^i(z)$ writes $(-1)^i \prod_{j=1}^i (M_t^{-1}z)_j$. Moreover, if N is odd (resp. even), there are only contributions of odd (resp. even) degrees of z in $P_t^N(z)$.

In particular, we can compute, for any $z \in \mathbb{R}^N$,

$$\begin{aligned}P_t^2(z) &= \prod_{i=1}^2 (M_t^{-1}z)_i - (M_t^{-1})_{1,2},\\P_t^3(z) &= -\prod_{i=1}^3 (M_t^{-1}z)_i + \sum_{\sigma\in\mathfrak{S}_3} (M_t^{-1})_{\sigma(1),\sigma(2)} (M_t^{-1}z)_{\sigma(3)},\end{aligned}$$

where $\mathfrak{S}_3$ is the symmetric group on $\{1,2,3\}$. More generally, for any $1 \le i \le N$, we can find a polynomial function $\mathscr{P}^i$ on $\mathbb{R}^i \times \mathbb{R}^{i(i-1)/2}$ such that

$$P_t^i(z) = \mathscr{P}^i\Big(\big((M_t^{-1}z)_j\big)_{1\le j\le i}, \big((M_t^{-1})_{j,k}\big)_{1\le j<k\le i}\Big).$$

The family $\mathscr{P}^1,\ldots,\mathscr{P}^N$ can be defined by induction by means of (5.38):

$$\begin{aligned}\mathscr{P}^i\big((\zeta_j)_{1\le j\le i}, (\vartheta_{j,k})_{1\le j<k\le i}\big) &= \sum_{\ell=1}^{i-1} \vartheta_{\ell,i}\partial_{\zeta_\ell}\mathscr{P}^{i-1}\big((\zeta_j)_{1\le j\le i}, (\vartheta_{j,k})_{1\le j<k\le i-1}\big)\\&\quad - \zeta_i \mathscr{P}^{i-1}\big((\zeta_j)_{1\le j\le i}, (\vartheta_{j,k})_{1\le j<k\le i-1}\big).\end{aligned}$$

Denoting by $M_t^{-1/2}$ the symmetric square root of M_t^{-1}, we can express both $M_t^{-1}z$ and M_t^{-1} in terms of quadratic combinations of $M_t^{-1/2}z$ and $M_t^{-1/2}$. Therefore, we can find a polynomial function $\mathscr{Q}^N$ on $\mathbb{R}^N \times \mathbb{R}^{N^2}$ such that

$$P_t^N(z) = \mathscr{Q}^N\Big(\big((M_t^{-1/2}z)_j\big)_{1\le j\le N}, \big((M_t^{-1/2})_{j,k}\big)_{1\le j,k\le N}\Big).$$

Then,

$$\begin{aligned}&\partial_{z_1,\cdots,z_N}\Big\{\exp\big(-\tfrac12\langle M_t^{-1}z, z\rangle\big)\Big\}\\&= P_t^N(z)\exp\Big(-\tfrac12|M_t^{-1/2}z|^2\Big)\\&= \mathscr{Q}^N\Big(\big((M_t^{-1/2}z)_j\big)_{1\le j\le N}, \big((M_t^{-1/2})_{j,k}\big)_{1\le j,k\le N}\Big)\exp\Big(-\tfrac12|M_t^{-1/2}z|^2\Big),\end{aligned}$$

which permits to 'absorb' the polynomial terms in $\big((M_t^{-1/2}z)_j\big)_{1\le j\le N}$ inside the exponential. There exists a constant $c := c(N) \in (0,1]$ such that

$$\big|\partial_{z_1,\cdots,z_N}\big\{\exp\big(-\tfrac12\langle M_t^{-1}z,z\rangle\big)\big\}\big| \le J_t(N)\exp\Big(-c|M_t^{-1/2}z|^2\Big), \tag{5.39}$$

where $J_t(N)$ reads

$$J_t(N) := \Big|\mathscr{R}^N\Big(\big((M_t^{-1/2})_{j,k}\big)_{1\le j,k\le N}\Big)\Big|, \tag{5.40}$$

for a polynomial function $\mathscr{R}^N$ on $\mathbb{R}^{N^2}$. Plugging (5.39) into (5.37) we obtain:

$$p_t(y) \le \frac{J_t(N)}{c^{N/2}} \int_{\prod_{i=1}^N\{\mathrm{sign}(y_i)z_i>|y_i|\}} \exp\Big(-\frac{c}{2}|M_t^{-1/2}z|^2\Big)\frac{c^{N/2}\det(M_t^{-1/2})dz}{(2\pi)^{N/2}}.$$

The covariance matrix M_t being given, the integral in the right-hand side can be interpreted as the probability that an N-dimensional centered Gaussian random vector with $c^{-1}M_t$ as covariance matrix be in the set $\{z \in \mathbb{R}^N : \text{sign}(y_i)z_i > |y_i|,\ i = 1,\dots,N\}$. Conditionally on $\mathcal{F}_t^Z$, we know that $c^{-1/2}t^{-1/2}\mathbb{T}_t^{-1}Z_t\Gamma_t$ is precisely a centered Gaussian vector with $c^{-1}M_t$ as covariance. Therefore, choosing $y = t^{-1/2}\mathbb{T}_t^{-1}(v - Z_t x_0)$, we deduce from (5.36):

$$\begin{aligned} f_{x_0}(t,v) &\le \frac{1}{t^{(N+1)/2}c^{N/2}} \\ &\quad \times \mathrm{E}\Big[J_t(N)\mathrm{P}\Big(\bigcap_{i=1}^N \Big\{\big|\big(t^{-1/2}\mathbb{T}_t^{-1}Z_t\Gamma_t\big)^i\big| > c^{1/2}\big|\big(t^{-1/2}\mathbb{T}_t^{-1}(v - Z_t x_0)\big)^i\big|\Big\}\big|\mathcal{F}_t^Z\Big)\Big]. \end{aligned}$$

Since the matrix $\mathbb{T}_t$ is diagonal, Cauchy-Schwarz inequality gives:

$$\begin{aligned} f_{x_0}(t,v) &\le \frac{\mathrm{E}[J_t(N)^2]^{1/2}}{t^{(N+1)/2}c^{N/2}}\mathrm{E}\Big[\mathrm{P}\Big(\bigcap_{i=1}^N \big\{|(Z_t\Gamma_t)^i| > c^{1/2}\big|(v - Z_t x_0)^i\big|\big\}|\mathcal{F}_t^Z\Big)^2\Big]^{1/2} \\ &\le \frac{\mathrm{E}[J_t(N)^2]^{1/2}}{t^{(N+1)/2}c^{N/2}}\mathrm{P}\Big(\bigcap_{i=1}^N \big\{|(Z_t\Gamma_t)^i| > c^{1/2}\big|(v - Z_t x_0)^i\big|\big\}\Big)^{1/2}. \end{aligned} \tag{5.41}$$

Formula (5.41) is the starting point for the upper bound that follows from the next two Lemmas.

Lemma 5.6 (Diagonal Controls). *Given $T > 0$, there exists a constant $C := C(N,T)$ such that, for all $t \in [0,T]$,*

$$\mathrm{E}[J_t(N)^2]^{1/2} \le C.$$

Lemma 5.7 (Tail estimates). *Given $T > 0$, there exists a constant $C := C(N,T) \ge 1$ such that for all $t \in [0,T]$ and $v \in \mathbb{R}^N$:*

$$\begin{aligned} &\mathrm{P}\Big(\bigcap_{i=1}^N \big\{|(Z_t\Gamma_t)^i| > c^{1/2}\big|(v - Z_t x_0)^i\big|\big\}\Big) \\ &\quad \le C\mathrm{E}\Big[\exp\Big(-\frac{C^{-1}}{t}\Big\{|(v - Z_t x_0)^1| + |(v - Z_t x_0)^1|^2 + \sum_{i=2}^N |(v - Z_t x_0)^i|^2\Big\}\Big)\Big]^{1/2}. \end{aligned}$$

The proofs of Lemmas 5.6 and 5.7 are given in the subsections 5.4.2 and 5.4.3 respectively.

5.4.2 Derivation of the diagonal controls

This subsection is dedicated to the proof of Lemma 5.6. Usually, in the Malliavin calculus approach to density estimates, this step is the most involved and requires a precise control of the determinant of the Malliavin covariance matrix, see e.g. Kusuoka and Stroock [17] or Bally [2]. In the current framework the determinant of the 'covariance' matrix M_t still plays a key role but the specific structure of that matrix, especially the fact that $(Z_s)_{s\ge 0}$ defines an isometry, yields the required estimate almost for free.

Precisely, we have the following Proposition.

Proposition 5.8 (Control of the determinant of the covariance). *For a given $T > 0$, there exists $C := C(N,T)$ such that, for all $t \in [0,T]$, almost surely,*

$$\det(M_t)^{-1} \le C.$$

Proof. Since $M_t = t^{-1}\mathbb{T}_t^{-1} C_t \mathbb{T}_t^{-1}$ and $\det(t^{1/2}\mathbb{T}_t)^{-2} = t^{-(N+1)}$, we here concentrate on $\det(C_t)$. The claim of the proposition indeed follows from the bound

$$\det(C_t) \geq C t^{N+1}. \tag{5.42}$$

To derive (5.42) we recall the 'variational' formulation of the determinant for symmetric matrices (see for instance [4]).

Lemma 5.9 (Variational expression of the determinant). *Let A be a symmetric $N \times N$ matrix. Then*

$$\det{}^{1/N}(A) = \frac{1}{N} \inf\{\mathrm{Tr}(aA),\, a \in \mathcal{S}_N^+(\mathbb{R}),\, \det(a) = 1\},$$

where $\mathcal{S}_N^+(\mathbb{R})$ stands for the set of symmetric positive $N \times N$ matrices.

Recall the expression of C_t from (5.7). Since Z_t is an isometry, we have $\det(C_t) = \det(\hat{C}_t)$ where $\hat{C}_t = \int_0^t Z_s^\top d\langle \bar{B}\rangle_s Z_s$. From (5.8), we get that $\hat{C}_t = \int_0^t Z_s^\top D_s Z_s ds$ with

$$D_s := \mathrm{diag}\Big((N-1)\frac{1-\exp(-2Ns)}{N}, 1 - \frac{1-\exp(-2Ns)}{N}, \dots, 1 - \frac{1-\exp(-2Ns)}{N}\Big), \tag{5.43}$$

so that $\det(D_s) \geq Cs$ for $s \in [0,T]$ and for some constant $C := C(N,T) > 0$. Therefore, we derive from Lemma 5.9 that, for any $a \in \mathcal{S}_N^+$ with $\det(a) = 1$,

$$\begin{aligned}
\left\{\frac{1}{N}\mathrm{Tr}\left[\left(\int_0^t Z_s^\top D_s Z_s ds\right) a\right]\right\}^N &= \left\{\int_0^t \frac{1}{N}\mathrm{Tr}\left(Z_s^\top D_s Z_s a\right) ds\right\}^N \\
&\geq \left\{\int_0^t \det\left(Z_s^\top D_s Z_s\right)^{1/N} ds\right\}^N \\
&\geq \left\{\int_0^t \det(D_s)^{1/N} ds\right\}^N \geq C t^{N+1},
\end{aligned}$$

for a new value of C. Taking the infimum over a and reapplying Lemma 5.9, this proves (5.42) and thus the proposition. □

To achieve the proof of Lemma 5.6, it therefore remains to check that the entries of the matrix M_t are bounded in any $L^p(\mathrm{P})$, $p \geq 1$ (uniformly on $[0,T]$). With the notation of Definition 5.3, Lemma 5.6 will follow from the control

$$\forall (i,j) \in \{1, \cdots, N\}^2,\ (M_t)_{i,j} = O_{\mathrm{P}}(1). \tag{5.44}$$

Associated with Proposition 5.8, this will indeed yield that $M_t^{-1/2}$ also satisfies (5.44) (by controlling from above and below the eigenvalues of M_t in terms of its determinant and its norm). Equation (5.44) is easily derived from (5.9), (5.15) and the definition of the scale matrix $t^{1/2}\mathbb{T}_t$ in (5.34). □

5.4.3 Derivation of the tail estimates

This subsection is dedicated to the proof of Lemma 5.7. Conditioning with respect to $\mathcal{F}_t^B := \sigma((B_s)_{0 \leq s \leq t})$ (which is independent of $(\bar{B}_s)_{s \geq 0}$),

$$\begin{aligned}
\pi &:= \mathrm{P}\Big(\bigcap_{i=1}^N \big\{|(Z_t\Gamma_t)^i| > c^{1/2}\big|(v - Z_t x_0)^i\big|\big\}\Big) \\
&= \mathrm{E}\Big[\mathrm{P}\Big(\bigcap_{i=1}^N \Big\{\Big|\Big(Z_t \int_0^t Z_s^\top d\bar{B}_s\Big)^i\Big| > c^{1/2}\big|(v - Z_t x_0)^i\big|\Big\}\big|\mathcal{F}_t^B\Big)\Big].
\end{aligned}$$

Since $(Z_s)_{0\le s\le t}$ is an isometry, it is bounded and so is $(Z_t Z_s^\top)_{0\le s\le t}$. Moreover, by (5.8), $d\langle \bar{B}_t\rangle/dt$ is less than Id_N (in the sense of symmetric matrices). Therefore, By Proposition 5.1 (Bernstein inequality) applied to the conditionally Gaussian variables $\big((\int_0^t Z_t Z_s^\top d\bar{B}_s)^i\big)_{i\in\{1,\cdots,N\}}$, there exists a constant $\bar{c} := \bar{c}(N) \ge 1$ such that

$$\pi \le \bar{c}\mathrm{E}\bigg[\exp\bigg(-\frac{1}{\bar{c}}\frac{|v^1-(Z_t x_0)^1|^2}{t} - \frac{1}{\bar{c}}\sum_{i=2}^N \frac{|(v-Z_t x_0)^i|^2}{t}\bigg)\bigg]. \tag{5.45}$$

Equation (5.45) provides us with the Gaussian part of the estimate. To derive the exponential one, we apply Chebychev inequality: for any $\gamma > 0$,

$$\begin{aligned}
\pi &\le \mathrm{E}\bigg[\exp\bigg(-\gamma c^{1/2}\frac{|v^1-(Z_t x_0)^1|}{t}\bigg)\mathrm{E}\Big[\exp\big(\frac{\gamma}{t}(Z_t\Gamma_t)^1\big)\mathbb{I}_{\cap_{i=1}^N\{(Z_t\Gamma_t)^i > c^{1/2}|(v-Z_t x_0)^i|\}}\Big|\mathcal{F}_t^B\Big]\bigg] \\
&\le \bar{c}^{1/2}\mathrm{E}\bigg[\exp\bigg(-\gamma c^{1/2}\frac{|v^1-(Z_t x_0)^1|}{t} - \frac{1}{2\bar{c}}\frac{|(v^1-(Z_t x_0)^1|^2}{t} - \frac{1}{2\bar{c}}\sum_{i=2}^N\frac{|(v-Z_t x_0)^i|^2}{t}\bigg) \\
&\qquad \times \mathrm{E}\bigg[\exp\bigg(\frac{2\gamma}{t}\bigg(Z_t\int_0^t Z_s^\top d\bar{B}_s\bigg)^1\bigg)\bigg|\mathcal{F}_t^B\bigg]^{1/2}\bigg] \\
&\le \bar{c}^{1/2}\mathrm{E}\bigg[\exp\bigg(-2\gamma c^{1/2}\frac{|v^1-(Z_t x_0)^1|}{t} - \frac{1}{\bar{c}}\frac{|(v^1-(Z_t x_0)^1|^2}{t} - \frac{1}{\bar{c}}\sum_{i=2}^N\frac{|(v-Z_t x_0)^i|^2}{t}\bigg)\bigg]^{1/2} \\
&\qquad \times \mathrm{E}\bigg[\exp\bigg(\frac{2\gamma}{t}\bigg(Z_t\int_0^t Z_s^\top d\bar{B}_s\bigg)^1\bigg)\bigg]^{1/2},
\end{aligned} \tag{5.46}$$

recalling $(Z_t\Gamma_t)^1 = (Z_t\int_0^t Z_s^\top d\bar{B}_s)^1$ and using also the Cauchy-Schwarz and Bernstein inequalities (similarly to (5.45)) to pass from the first to the second line. Recalling (5.7) and using the Gaussian character of the conditional distribution of $\int_0^t Z_s^\top d\bar{B}_s$ given $\mathcal{F}_t^B$, we get

$$\mathrm{E}\bigg[\exp\bigg(\frac{2\gamma}{t}\bigg(\int_0^t Z_t Z_s^\top d\bar{B}_s\bigg)^1\bigg)\bigg|\mathcal{F}_t^B\bigg] = \mathrm{E}\Big[\exp\big(\frac{2\gamma^2}{t^2}C_t^{1,1}\big)\big|\mathcal{F}_t^B\Big] \tag{5.47}$$

When taking the expectation, we know from the identity in law (5.9) that we can replace C_t by $\bar{C}_t$. By (5.12) and (5.13),

$$\begin{aligned}
\bar{C}_t^{1,1} &= \int_0^t \big(1-2(t-s)-(1-2N(t-s))\big(Z_s^{1,1}\big)^2\big)ds + O(t^3) \\
&= \int_0^t \big(1-\big(Z_s^{1,1}\big)^2\big)ds + O(t^2) = \int_0^t \big(1-Z_s^{1,1}\big)\big(1+Z_s^{1,1}\big)ds + O(t^2).
\end{aligned}$$

We then write $Z_s^{1,1} = (1+\mathcal{O}(s))(1+S_s^{1,1})$, which leads to a simplified version of (5.15):

$$\bar{C}_t^{1,1} = tO\big(S_t^*\big) + O(t^2).$$

The point is then to plug the above expansion into the expectation of (5.47). We thus compute the moments of the right-hand side above. We make use of Lemma 5.4, which says that S_t^*/t has an exponential tail. Therefore, choosing γ small enough, we can bound the last factor in the right-hand side in (5.47) by $\bar{C} := \bar{C}(N,T)$. This completes the proof of Lemma 5.7 □

5.4.4 Conclusion

Combining Lemmas 5.6 and 5.7 we derive that, for $t \in [0,T]$,

$$\begin{aligned} &f_{x_0}(t,v) \\ &\leq \frac{C}{t^{(N+1)/2}} \mathrm{E}\biggl[\exp\biggl(-\frac{1}{Ct}\biggl[|v^1-(Z_t x_0)^1| + |v^1-(Z_t x_0)^1|^2 + \sum_{i=2}^N |v^i-(Z_t x_0)^i|^2\biggr]\biggr)\biggr]^{1/2}, \end{aligned} \tag{5.48}$$

with $C := C(N,T)$. Using Cauchy-Schwarz inequality, it suffices to bound

$$\begin{aligned} F_1 &:= \mathrm{E}\biggl[\exp\biggl(-\frac{2}{Ct}\Bigl[|v^1-(Z_t x_0)^1| + |v^1-(Z_t x_0)^1|^2\Bigr]\biggr)\biggr]^{1/4}, \\ F_2 &:= \mathrm{E}\biggl[\exp\biggl(-\frac{2}{Ct}\sum_{i=2}^N |v^i-(Z_t x_0)^i|^2\biggr)\biggr]^{1/4}. \end{aligned}$$

We start with F_2. By the inequality, $-2|v^i-(Z_t x_0)^i|^2 \leq -|v^i|^2 + 2|(Z_t x_0)^i|^2$, we obtain

$$\begin{aligned} F_2 &\leq \exp\biggl(-\frac{1}{4Ct}\sum_{i=2}^N |v^i|^2\biggr)\mathrm{E}\biggl[\exp\biggl(\frac{2}{Ct}\sum_{i=2}^N |(Z_t x_0)^i|^2\biggr)\biggr]^{1/4} \\ &= \exp\biggl(-\frac{1}{4Ct}\sum_{i=2}^N |v^i|^2\biggr)\mathrm{E}\biggl[\exp\biggl(\frac{2|x_0^1|^2}{Ct}\sum_{i=2}^N \bigl(Z_t^{i,1}\bigr)^2\biggr)\biggr]^{1/4}. \end{aligned} \tag{5.49}$$

Now, $\sum_{i=2}^N (Z_t^{i,1})^2 = 1-(Z_t^{1,1})^2 = O(1-Z_t^{1,1}) = O(S_t^{1,1}+t)$. Therefore, for $|x_0^1| \leq C_0$, we deduce from Lemma 5.4 that we can choose $C := C(N,T,C_0)$ large enough in (5.48) such that the second factor in the last line is bounded by C.

It remains to bound F_1. Given some $A > 0$, we split the expectation according to the events $\{|Z_t^{1,1}-1||x_0^1| \leq A\}$ and $\{|Z_t^{1,1}-1||x_0^1| > A\}$. Using the inequalities $-2|v^1-(Z_t x_0)^1|^2 \leq -|v^1-x_0^1|^2 + 2|(Z_t^{1,1}-1)x_0^1|^2$ and $-2|v^1-(Z_t x_0)^1| \leq -|v^1-x_0^1| + 2|(Z_t^{1,1}-1)x_0^1|$, we obtain

$$F_1 \leq \exp\biggl(-\frac{1}{4Ct}\Bigl[|v^1-x_0^1| + |v^1-x_0^1|^2\Bigr] + \frac{1}{2Ct}\bigl(A+A^2\bigr)\biggr) + \mathrm{P}\bigl(|Z_t^{1,1}-1||x_0^1| > A\bigr),$$

for $C := C(N,T)$. Now, by Proposition 5.1 (Bernstein inequality), we have

$$\mathrm{P}\bigl(|(Z_t^{1,1}-1)x_0^1| \geq A\bigr) \leq 2\exp\bigl(-\frac{A^2}{2(x_0^1)^2 t}\bigr).$$

On the other hand, since $Z_t^{1,1}-1 = O(S_t^{1,1}+t)$, Lemma 5.4 yields (for a possibly new value of C)

$$\mathrm{P}\bigl(|(Z_t^{1,1}-1)x_0^1| \geq A\bigr) \leq C\exp\bigl(-\frac{A}{C|x_0^1|t}\bigr).$$

Choosing $A = |v^1-x_0^1|/8$, we deduce that (once again, modifying C if necessary)

$$F_1 \leq C\exp\biggl(-\frac{1}{Ct}\biggl[\frac{|v^1-x_0^1|}{1\vee|x_0^1|} + \frac{|v^1-x_0^1|^2}{1\vee|x_0^1|^2}\biggr]\biggr). \tag{5.50}$$

By (5.48), (5.49) and (5.50), we get an upper bound for $f_{x_0}(t,v)$. When $x_0 \in [-C_0,C_0]$, we can choose C, depending upon N, T and C_0, in order to get the required estimate. (As in the lower bound, the dependence of C upon C_0 can be made explicit.)

References

[1] Aronson, D. G. (1967) Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, **73**, 890–896. MR-0217444

[2] Bally, V. (1990) On the connection between the Malliavin covariance matrix and Hörmander's condition. *Journal of Functional Analysis*, **96**, 219–255 MR-1101258

[3] Carrapatoso, K. (2012) Propagation of chaos for the spatially homogeneous Landau equation for Maxwellian molecules. arXiv:1212.3724

[4] Delarue, F. (2012), Stochastic Analysis for the Complex Monge-Ampère Equation. (An Introduction to Krylov's Approach). *Complex Monge-Ampère equations and Geodesics in the Space of Kähler Metrics. Lecture Notes in Mathematics*, **2028**, 55-198. MR-2932443

[5] Desvillettes, L. and Villani, C. (2000), On the spatially homogeneous Landau equation for hard potentials. Part I. Existence, uniqueness and smoothness. *Comm. Partial Differential Equations*, **25**, 179-259. MR-1737547

[6] Fontbona, J. and Guérin, H. and Méléard, S. (2009), Measurability of optimal transportation and convergence rate for Landau type interacting particle systems. *Probab. Theory Related Fields*, **134**, 3-4, 329-351. MR-2475665

[7] Fournier, N. (2009), Particle approximation of some Landau equations. *Kinet. Relat. Models*, **2, 3**, 451 - 464. MR-2525721

[8] Franchi, J. and Le Jan, Y. (2012), Hyperbolic dynamics and Brownian motion. An introduction. *Oxford University Press.* MR-2977883

[9] Funaki, T. (1983), The diffusion approximation of the Boltzmann equation of Maxwellian molecules, *Publ. Res. Inst. Math. Sci.*, **19**, 841-886. MR-0716978

[10] Funaki, T. (1984), A certain class of diffusion processes associated with nonlinear parabolic equations. *Z. Wahrsch. Verw. Gebiete*, **67**, 331-348. MR-0762085

[11] Funaki, T. (1985), The diffusion approximation of the spatially homogeneous Boltzmann equation. *Duke Mathematical Journal*, **52**, 1-23. MR-0791288

[12] Funaki, T. (1986), Construction of stochastic processes associated with the Boltzmann equation and its applications. *Stochastic Processes and Their Applications, Lecture Notes in Mathematics*, **1203**, 51-65. MR-0872101

[13] Guérin, H. (2002), Existence and regularity of a weak function-solution for some Landau equation with a stochastic approach, *Stochastic Process. Appl.*, **101**, 303-325. MR-1931271.

[14] Guérin, H. (2003), Solving Landau equation for some soft potentials through a probabilistic approach. *Annals of Applied Probability*, **13**, 515-539. MR-1970275

[15] Guérin, H. (2004), Pointwise convergence of Boltzmann solutions for grazing collisions in a Maxwell gas via a probabilitistic interpretation. *ESAIM: Probability and Statistics*, **8**, 36-55. MR-2085604

[16] Guérin, H., Méléard, S. and Nualart, E. (2006), Estimates for the density of a nonlinear Landau process, *Journal of Functional Analysis*, **238**, 649-677. MR-2253737

[17] Kusuoka, S. and Stroock, D. (1985) Applications of the Malliavin calculus II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **32**, 1–76. MR-0783181

[18] Kusuoka, S., Stroock, D. (1987) Applications of the Malliavin calculus III, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, **34**, 391-442. MR-0914028

[19] Revuz, D. and Yor, M. (1999) *Continuous martingales and Brownian motion. 3rd ed.*. Springer. MR-1725357

[20] Rogers, L.C.G. and Williams D. (1985) *Markov processes and Martingales, Volume II: Itô Calculus*. Cambridge University Press MR-1780932.

[21] Stroock, D. (2006) Estimates for transition probabilities on a compact manifold. *Journal of Functional Analysis*, **242**, 295–303. MR-2274025

[22] Sznitman, A. S. (1991) Topics in propagation of chaos. *Lecture Notes in Mathematics, Springer*, **1464**, 165-251. MR-1108185

[23] Varopoulos, N. T. and Saloff-Coste, L. and Coulhon, T. (1992) Analysis and Geometry on Groups. *Cambridge University Press.* MR-1218884

[24] Villani, C. (1998), A new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations, *Arch. Rational Mech. Anal.*, **143**, 273-307. MR-1650006

[25] Villani, C. (1998), On the spatially homogeneous Landau equation for Maxwellian molecules, *Math. Models Methods Appl. Sci.*, **8**, 957-983. MR-1646502

Acknowledgments. Part of the work for this paper was done when the authors where visiting the Centre Interfacultaire Bernoulli of the École Polytechnique Fédérale de Lausanne, during the semester on Stochastic Analysis and Applications organized by Robert Dalang, Marco Dozzi, Franco Flandoli and Francesco Russo. We are very grateful for their invitation and hospitality. For the second author, the article was prepared within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program. Third author acknowledges support from the European Union programme FP7-PEOPLE-2012-CIG under grant agreement 333938.