Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 18, No. 4/2026

HARNACK INEQUALITY AND GAUSSIAN BOUNDS FOR PARABOLIC EQUATIONS WITH SINGULAR LOWER-ORDER COEFFICIENTS*

Mykola Ivanovich Yaremenko†

DOI 10.56082/annalsarscimath.2026.4.37

Abstract

We study the evolution family generated by second-order parabolic operators of the form

$$\partial_t u - \nabla \cdot (a(t,x)\nabla u) + b(t,x) \cdot \nabla u - \nabla \cdot (\tilde{b}(t,x)u) + V(t,x)u.$$

The diffusion matrix a is uniformly elliptic and Dini continuous in the spatial variable. The lower-order coefficients b, $\tilde{b}$, V belong to the parabolic Kato classes K_{l+1} and K_{l+2}, which are optimal for perturbation theory. Using Moser's iteration we prove the Harnack inequality and then the two-sided Gaussian bound for the fundamental solution:

$$c\,(t-s)^{-l/2} \exp\Big(-\frac{|x-y|^2}{C(t-s)}\Big) \le p(t,x;s,y)$$

$$\le C\,(t-s)^{-l/2} \exp\Big(-\frac{|x-y|^2}{c(t-s)}\Big),$$

with constants depending only on the ellipticity bounds and the Kato norms. As applications we obtain well-posedness of the Cauchy problem in $L^1 \cap L^\infty$ and a weighted maximum principle.

*Accepted for publication on July 01, 2026

†math.kiev@gmail.com, National Technical University of Ukraine *Igor Sikorsky Kyiv Polytechnic Institute* Kyiv, Ukraine

Keywords: Harnack inequality, Gaussian bounds, parabolic equations, fundamental solution, singular potentials, Kato class, Dini continuity, Moser iteration, perturbation theory, heat kernel.

MSC: 35K08, 35K10, 31B35, 35J10, 35B45, 35D30, 47D07.

1 Introduction

The study of fundamental solutions (heat kernels) for parabolic equations is a classical topic with deep connections to analysis, probability, and geometry. For operators in divergence form

$$\partial_t u - \nabla \cdot (a(t,x)\nabla u)$$

with uniformly elliptic, bounded measurable coefficients a, Nash [13] and Aronson [2,3] proved sharp two-sided Gaussian bounds. These bounds imply the Harnack inequality, independently established by Moser [11, 12], and form the cornerstone of the regularity theory for parabolic equations; see the monographs [6, 8, 9].

When lower-order terms (drift, adjoint drift, potential) are added, the problem becomes more delicate because these terms may be singular – they need not belong to any L^p space with a high exponent, yet they can still be treated as perturbations of the principal part. The natural framework for such singularities is provided by the **parabolic Kato classes**. For a measurable function f on $\mathbb{R}_+ \times \mathbb{R}^l$ and $p \geq 1$, we say $f \in K_p$ if

$$\lim_{h\to 0} \operatorname*{ess\,sup}_{(s,z)} \int_s^{s+h} \int_{\mathbb{R}^l} (4\pi(t-s))^{-l/2} \exp\Big(-\frac{|y-z|^2}{4(t-s)}\Big) |f(t,y)|^p \, dy \, dt = 0.$$

These classes are scaling-invariant and contain critical singularities: for example, $|x|^{-2} \in K_{l+2}$ in dimensions $l \geq 3$, and $|x|^{-1} \in K_{l+1}$ for $l \geq 2$. The parabolic Kato classes were introduced in the context of Schrödinger operators [1, 7, 14] and later extended to parabolic equations [15]. They are optimal for perturbation theory: if the lower-order coefficients belong to the appropriate Kato classes, the fundamental solution of the full operator inherits the Gaussian bounds of the principal part.

In this paper we consider the operator

$$Lu = \partial_t u - \nabla \cdot (a(t,x)\nabla u) + b(t,x) \cdot \nabla u - \nabla \cdot (\tilde{b}(t,x)u) + V(t,x)u, \quad (1)$$

where $x \in \mathbb{R}^l$, $t \geq 0$, and $l \geq 2$. The diffusion matrix a is assumed to be symmetric, uniformly elliptic, and Dini continuous in the spatial variable

(Assumptions 2 and 3). The lower-order coefficients satisfy

$$b, \tilde{b} \in K_{l+1}, \qquad V \in K_{l+2},$$

componentwise for the vector fields. These are exactly the critical exponents that make the perturbation series converge.

Our main result is the following.

Theorem 1 (Main Theorem)**.** *Under the above assumptions, the operator L has a fundamental solution $U_{t,s}(x,y)$ defined for $0 \le s < t$, $x, y \in \mathbb{R}^l$. For any $T > 0$ there exist constants $c, C > 0$ depending only on the ellipticity constants ν, μ, the Dini modulus ω, the Kato norms of $b, \tilde{b}, V$, and T such that for all $0 < t - s \le T$ and all $x, y \in \mathbb{R}^l$,*

$$c\,(t-s)^{-l/2} \exp\Big(-\frac{|x-y|^2}{C(t-s)}\Big) \le U_{t,s}(x,y) \le C\,(t-s)^{-l/2} \exp\Big(-\frac{|x-y|^2}{c(t-s)}\Big).$$

Moreover, the evolution family $U_{t,s}$ is strongly continuous on $L^1(\mathbb{R}^l) \cap L^\infty(\mathbb{R}^l)$ and satisfies $\|U_{t,s}\|_{1\to 1} \le C$, $\|U_{t,s}\|_{\infty\to\infty} \le C$.

The proof proceeds in three main steps.

1. Principal part. We first construct the fundamental solution p_0 of $L_0 = \partial_t - \nabla \cdot (a\nabla)$. This is done by the classical parametrix method, which exploits the Dini continuity of a to ensure convergence of the series. The result (Theorem 5) gives two-sided Gaussian bounds and derivative estimates for p_0.

2. Harnack inequality and lower bound. Using Moser's iteration [11,12] we prove the Harnack inequality for non-negative solutions of $L_0 u = 0$ (Theorem 6). From this inequality and the semigroup property we deduce the lower Gaussian bound for p_0. This part relies only on uniform ellipticity, not on Dini continuity.

3. Perturbation. The full operator L is treated as a perturbation of L_0. The Duhamel formula

$$U_{t,s} = P_{t,s} + \sum_{n=1}^{\infty} (-1)^n \int_{s<\tau_1<\cdots<\tau_n<t} P_{t,\tau_n} B(\tau_n) P_{\tau_n,\tau_{n-1}} \cdots B(\tau_1) P_{\tau_1,s}\, d\tau_1 \cdots d\tau_n$$

expresses the evolution family $U_{t,s}$ of L in terms of the unperturbed family $P_{t,s}$ of L_0 and the lower-order operator $B = b \cdot \nabla - \nabla \cdot (\tilde{b}\,\cdot) + V$. The

Kato class conditions guarantee that the integrated perturbation $\beta(h) = \operatorname{ess\,sup}_{z,s} \int_s^{s+h} \|B(\tau)P_{\tau,s}\delta_z\|_{L^1} d\tau$ tends to 0 as $h \to 0$. Hence for sufficiently small time steps the series converges in operator norm, and the Gaussian upper bound for $U_{t,s}$ follows by induction. The lower bound for small times is obtained by comparing $U_{t,s}$ with $P_{t,s}$ and using the smallness of β. For larger times a chaining argument (splitting the time interval into small pieces) extends the bounds to arbitrary T.

Applications and outline. The Gaussian bounds have immediate consequences, including well-posedness of the Cauchy problem in $L^1 \cap L^\infty$, a weighted maximum principle, and L^p–L^q ultracontractivity estimates. The paper is organised as follows. Section 2 collects the precise assumptions and recalls the definition of the parabolic Kato classes together with their basic properties. Section 3 is devoted to the construction of the fundamental solution of L_0 and the proof of Gaussian bounds and derivative estimates. In Section 4 we prove the Harnack inequality for L_0 via Moser iteration and derive the lower Gaussian bound for p_0. Section 5 contains the perturbation analysis: estimates for the Duhamel series, convergence, and the proof of the two-sided Gaussian bounds for $U_{t,s}$.

The results presented here improve upon earlier work by removing unnecessary regularity assumptions on the lower-order coefficients and by providing a unified treatment of drift, adjoint drift, and potential terms. The critical exponents $l+1$ and $l+2$ are optimal, as can be seen from counterexamples when the Kato norm is infinite. Moreover, the Dini continuity of a is essentially optimal for the parametrix method; if a is only VMO, the Gaussian bounds may fail in the sense that the heat kernel may not satisfy a two-sided estimate (see [6] for a discussion of the necessity of Dini conditions in the parametrix construction).

2 Preliminaries and assumptions

We work in $\mathbb{R}_+ \times \mathbb{R}^l$ with $l \geq 1$. For a function $f(t,x)$, the L^p norm in the spatial variable is $\|f(t,\cdot)\|_p = \left(\int |f(t,x)|^p dx\right)^{1/p}$, and when the dependence on t is not essential we simply write $\|f\|_p$. For a kernel $K(x,y)$, the operator norm from L^1 to L^1 is $\|K\|_{1\to 1} = \sup_y \int |K(x,y)|\, dx$.

2.1 The diffusion matrix

Let $a(t,x) = (a_{ij}(t,x))_{i,j=1}^l$ be a symmetric matrix satisfying:

Assumption 2 (Uniform ellipticity)**.** There exist constants $0 < \nu \le \mu < \infty$ such that for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^l$ and $\xi \in \mathbb{R}^l$,

$$\nu|\xi|^2 \le a_{ij}(t,x)\xi_i\xi_j \le \mu|\xi|^2.$$

Assumption 3 (Dini continuity)**.** There exists a non-decreasing modulus of continuity $\omega : [0,\infty) \to [0,\infty)$ with $\omega(0) = 0$ and $\int_0^1 \frac{\omega(\delta)}{\delta}\, d\delta < \infty$ such that for all $t \ge 0$ and $x, y \in \mathbb{R}^l$,

$$|a_{ij}(t,x) - a_{ij}(t,y)| \le \omega(|x-y|).$$

The Dini condition is a mild regularity assumption; it is satisfied, for example, by Hölder continuous coefficients with any exponent $\alpha > 0$ (since $\omega(\delta) = C\delta^\alpha$ gives $\int_0^1 \delta^{\alpha-1} d\delta < \infty$). It also allows logarithmic moduli such as $\omega(\delta) = |\log \delta|^{-1-\varepsilon}$.

2.2 Parabolic Kato classes

We recall the standard parabolic Kato classes, which are optimal for perturbation of the heat kernel. For a detailed treatment we refer to [1,14,15].

Definition 1 (Parabolic Kato class K_p)**.** Let $p \ge 1$. A measurable function $f : \mathbb{R}_+ \times \mathbb{R}^l \to \mathbb{R}$ is said to belong to the parabolic Kato class K_p if

$$\lim_{h \to 0} \operatorname*{ess\,sup}_{(s,z) \in \mathbb{R}_+ \times \mathbb{R}^l} \int_s^{s+h} \int_{\mathbb{R}^l} G(t-s, y-z)|f(t,y)|^p \, dy\, dt = 0,$$

where $G(t,x) := (4\pi t)^{-l/2} \exp\bigl(-|x|^2/(4t)\bigr)$ is the standard heat kernel.

The exponent p is chosen according to the scaling properties of the operator. For a drift term $b \cdot \nabla u$, the natural condition is $b \in K_{l+1}$ (or equivalently $|b|^2 \in K_{l+2}$). For a potential Vu, the condition is $V \in K_{l+2}$. These classes are scaling-invariant under the parabolic scaling $(t,x) \to (\lambda^2 t, \lambda x)$ in the sense that if $f \in K_p$ then the scaled function $f_\lambda(t,x) = f(\lambda^2 t, \lambda x)$ also belongs to K_p. They contain functions with critical singularities, e.g., $V(x) = |x|^{-2}$ in dimension $l \ge 3$ belongs to K_{l+2}, and $b(x) = |x|^{-1}$ in dimension $l \ge 2$ belongs to K_{l+1}.

A useful property is that for $f \in K_p$, the map

$$(t,s,z) \mapsto \int_s^t \int_{\mathbb{R}^l} G(t-\tau, y-z)|f(\tau,y)|^p \, dy\, d\tau$$

is continuous in t, s and tends to 0 uniformly in z as $t - s \to 0$. This allows us to treat the lower-order terms as a small perturbation on short time intervals.

Lemma 1 (Basic estimate for Kato functions)**.** *If $f \in K_p$, then for any $\varepsilon > 0$ there exists $h_0 > 0$ such that for all $0 < h \le h_0$,*

$$\operatorname*{ess\,sup}_{(s,z)} \int_s^{s+h} \int_{\mathbb{R}^l} G(t-s, y-z)|f(t,y)|^p \, dy \, dt \le \varepsilon.$$

Proof. This follows directly from the definition of K_p. □

Assumption 4. The lower-order coefficients satisfy:

1. $b, \tilde{b} \in K_{l+1}$ componentwise,
2. $V \in K_{l+2}$.

For the adjoint drift term $-\nabla \cdot (\tilde{b}u)$ we work in the weak sense; estimates for this term are obtained by duality using only $\tilde{b} \in K_{l+1}$. No additional regularity on $\nabla \cdot \tilde{b}$ is required.

Remark. If $\tilde{b}$ is not regular enough to define $\nabla \cdot \tilde{b}$ classically, the operator $-\nabla \cdot (\tilde{b}u)$ is interpreted as the adjoint of $\tilde{b} \cdot \nabla$ in the usual distributional sense. The L^1 estimates needed for the perturbation series can be derived by duality from the estimates for the drift term; see [15] for details.

2.3 Notation for Gaussian kernels

For any $\alpha > 0$ and $t > 0$, $x \in \mathbb{R}^l$, define

$$K_\alpha(t,x) = (4\pi\alpha t)^{-l/2} \exp\Big(-\frac{|x|^2}{4\alpha t}\Big).$$

The standard heat kernel (corresponding to $\alpha = 1$) is denoted by $G(t,x) = K_1(t,x)$. These kernels satisfy the semigroup identity: for $0 < s < \tau < t$,

$$\int_{\mathbb{R}^l} K_\alpha(t-\tau, x-z)\, K_\alpha(\tau - s, z-y)\, dz = K_\alpha(t-s, x-y).$$

In the estimates that follow, we will often use K_α with different parameters α that arise from the ellipticity constants or from the Gaussian bounds of the fundamental solution. Specifically, the frozen kernel $p_0^y(t,x;s,y)$ is bounded between two such kernels with $\alpha = \nu$ and $\alpha = \mu$, and the derivative estimates involve K_α with $\alpha = \nu/2$ (see Theorem 5). The parameter α will always be a positive constant depending only on ν, μ, and we will systematically exploit the scaling and convolution properties of K_α.

We note that the notation K_α is not to be confused with the Kato classes K_p introduced in Definition 1; the distinction will be clear from the context (the former has a subscript $\alpha > 0$ and appears with arguments (t,x), while the latter is a function class denoted by K_p with $p \ge 1$).

2.4 The fundamental solution and evolution families

For a parabolic operator of the form $\partial_t - \nabla \cdot (a\nabla)$ with a satisfying Assumptions 2 and 3, there exists a unique fundamental solution $p_0(t, x; s, y)$ defined for $0 \le s < t$, $x, y \in \mathbb{R}^l$, with the properties:

1. $p_0(t, x; s, y)$ is twice continuously differentiable in x for $t > s$ (and the required spatial derivatives up to order two satisfy Gaussian bounds),
2. $\lim_{t \downarrow s} \int p_0(t, x; s, y)\phi(y)\, dy = \phi(x)$ for all $\phi \in C_c^\infty$,
3. $(\partial_t - \nabla_x \cdot (a(t, x)\nabla_x))p_0(t, x; s, y) = 0$ for $t > s$,
4. $(\partial_s + \nabla_y \cdot (a(s, y)\nabla_y))p_0(t, x; s, y) = 0$ for $s < t$.

The construction via the parametrix method is carried out in Section 3. The evolution family $\{P_{t,s}\}_{t \ge s}$ is defined by

$$(P_{t,s}f)(x) = \int_{\mathbb{R}^l} p_0(t, x; s, y) f(y)\, dy.$$

It satisfies $P_{t,s} = P_{t,\tau}P_{\tau,s}$ for $s < \tau < t$ and $\|P_{t,s}\|_{1\to 1} = 1$, $\|P_{t,s}\|_{\infty\to\infty} = 1$.

For the full operator $L = L_0 + B$, the evolution family $U_{t,s}$ is defined similarly via the Duhamel formula; its existence and basic properties are established in Section 5.

2.5 Function spaces and conventions

We will often use the following shorthand:

1. For a kernel $K(t, x; s, y)$, we write $K(t, x; s, y)$ as a function of (x, y) for fixed t, s.
2. The convolution in space is denoted
$$(K_1 * K_2)(x, y) = \int K_1(x, z)K_2(z, y)\, dz.$$
3. All constants c, C may change from line to line but depend only on the parameters indicated (e.g., ν, μ, ω, Kato norms, dimension l).
4. The notation $A \lesssim B$ means $A \le CB$ with a constant that is universal or depends only on the fixed parameters.

These preliminaries provide the necessary background for the subsequent sections. In the next section we construct the fundamental solution of L_0 and prove its Gaussian bounds.

3 Fundamental solution of the principal part

Consider the operator

$$L_0 u = \partial_t u - \nabla \cdot (a(t,x)\nabla u), \tag{2}$$

with a satisfying Assumptions 2 (uniform ellipticity) and 3 (Dini continuity in x). We construct its fundamental solution $p_0(t,x;s,y)$ for $0 \le s < t$, $x, y \in \mathbb{R}^l$ using the parametrix method [6, 9].

3.1 The frozen kernel

For a fixed $y \in \mathbb{R}^l$, freeze the coefficients at y and consider the constant-coefficient operator

$$L_0^y u = \partial_t u - a_{ij}(t,y)\partial_i \partial_j u.$$

Its fundamental solution is given explicitly by

$$p_0^y(t,x;s,y) = (4\pi)^{-l/2} (\det Q(s,t,y))^{-1/2} \exp\Big(-\frac{1}{4} x \cdot Q(s,t,y)^{-1} x\Big), \tag{3}$$

where $Q(s,t,y) = \int_s^t a(\tau,y)\, d\tau$. By ellipticity, $\nu(t-s)I \le Q(s,t,y) \le \mu(t-s)I$, hence

$$c_0 (t-s)^{-l/2} \exp\Big(-\frac{C_0|x|^2}{t-s}\Big) \le p_0^y(t,x;s,y) \le C_0 (t-s)^{-l/2} \exp\Big(-\frac{c_0|x|^2}{t-s}\Big), \tag{4}$$

with constants $c_0, C_0 > 0$ depending only on ν, μ. Moreover, for any multi-index α with $|\alpha| \le 2$,

$$|\partial_x^\alpha p_0^y(t,x;s,y)| \le C_\alpha (t-s)^{-(l+|\alpha|)/2} \exp\Big(-\frac{c_\alpha |x|^2}{t-s}\Big). \tag{5}$$

3.2 The error kernel and the parametrix equation

Define the error kernel

$$F(t,x;s,y) = (L_0 - L_0^y) p_0^y(t,x;s,y).$$

Since L_0 and L_0^y differ only in the diffusion matrix,

$$F(t,x;s,y) = (a_{ij}(t,x) - a_{ij}(t,y))\partial_i \partial_j p_0^y(t,x;s,y).$$

Using (5) with $|\alpha| = 2$ and the Dini continuity $|a_{ij}(t,x) - a_{ij}(t,y)| \le \omega(|x-y|)$, we obtain

$$|F(t,x;s,y)| \le C_F\,\omega(|x-y|)\,(t-s)^{-(l+2)/2}\exp\Big(-\frac{c_F|x-y|^2}{t-s}\Big), \tag{6}$$

with constants $C_F, c_F > 0$ depending only on ν, μ.

We seek the fundamental solution of L_0 in the form

$$p_0(t,x;s,y) = p_0^y(t,x;s,y) + \int_s^t \int_{\mathbb{R}^l} p_0^z(t,x;\tau,z)\,\Phi(\tau,z;s,y)\,dz\,d\tau,$$

where Φ is an unknown kernel. Substituting into $L_0 p_0 = 0$ leads to the Volterra equation

$$\Phi(t,x;s,y) = F(t,x;s,y) + \int_s^t \int_{\mathbb{R}^l} F(t,x;\tau,z)\,\Phi(\tau,z;s,y)\,dz\,d\tau. \tag{7}$$

3.3 Solution of the Volterra equation

Equation (7) is solved by successive approximations. It is a classical result (see [6, Chapter 1]) that under the Dini condition $\int_0^1 \frac{\omega(\delta)}{\delta} d\delta < \infty$, the series converges uniformly for $0 < t-s \le T$, and the solution Φ satisfies a Gaussian bound of the same type as the heat kernel:

$$|\Phi(t,x;s,y)| \le C_\Phi\,(t-s)^{-l/2}\exp\Big(-\frac{c_\Phi|x-y|^2}{t-s}\Big), \tag{8}$$

with constants $C_\Phi, c_\Phi > 0$ depending only on ν, μ, ω, T.

3.4 Fundamental solution of the principal part

Theorem 5. *Under Assumptions 2 and 3, the operator $L_0 = \partial_t - \nabla\cdot(a\nabla)$ has a fundamental solution $p_0(t,x;s,y)$ defined for $0 \le s < t$, $x, y \in \mathbb{R}^l$. There exist constants $c_1, C_1 > 0$ depending only on ν, μ, ω and the dimension l such that for all $t > s$:*

1. ***(Two-sided Gaussian bounds)***

$$c_1\,(t-s)^{-l/2}\exp\Big(-\frac{|x-y|^2}{C_1(t-s)}\Big)$$

$$\le p_0(t,x;s,y) \le C_1\,(t-s)^{-l/2}\exp\Big(-\frac{|x-y|^2}{c_1(t-s)}\Big).$$

2. ***(First spatial derivatives)***

$$|\nabla_x p_0(t,x;s,y)| \le C_1\,(t-s)^{-(l+1)/2}\exp\Big(-\frac{|x-y|^2}{C_1(t-s)}\Big).$$

(Estimates for higher derivatives are not required in this paper; they would demand additional regularity and are omitted.)

Proof. Step 1: Frozen kernel and parametrix. For fixed y, freeze the coefficients at y and consider the constant-coefficient operator $L_0^y u = \partial_t u - a_{ij}(t,y)\partial_i\partial_j u$. Its fundamental solution is

$$p_0^y(t,x;s,y) = (4\pi)^{-l/2}(\det Q(s,t,y))^{-1/2}\exp\Big(-\frac{1}{4}x\cdot Q(s,t,y)^{-1}x\Big),$$

where $Q(s,t,y)=\int_s^t a(\tau,y)\,d\tau$. Ellipticity gives $\nu(t-s)I \le Q \le \mu(t-s)I$, hence

$$c_0(t-s)^{-l/2}e^{-C_0|x|^2/(t-s)} \le p_0^y \le C_0(t-s)^{-l/2}e^{-c_0|x|^2/(t-s)}, \tag{9}$$

and

$$|\nabla_x p_0^y(t,x;s,y)| \le C_0(t-s)^{-(l+1)/2}e^{-c_0|x|^2/(t-s)}. \tag{10}$$

The constants c_0, C_0 depend only on ν,μ.

Write L_0 in non-divergence form: $L_0u = \partial_t u - a_{ij}\partial_i\partial_j u - (\nabla\cdot a)\cdot\nabla u$. The error kernel is

$$F = (L_0 - L_0^y)p_0^y = (a_{ij}(t,x)-a_{ij}(t,y))\partial_i\partial_j p_0^y - (\nabla\cdot a)\cdot\nabla p_0^y.$$

Using Dini continuity $|a_{ij}(t,x)-a_{ij}(t,y)| \le \omega(|x-y|)$ and the boundedness of $\nabla\cdot a$, we obtain

$$|F(t,x;s,y)| \le C_F\big(\omega(|x-y|)+(t-s)^{1/2}\big)(t-s)^{-(l+2)/2}\exp\Big(-\frac{c_F|x-y|^2}{t-s}\Big), \tag{11}$$

with $C_F, c_F > 0$ depending only on ν,μ.

We seek p_0 in the form

$$p_0(t,x;s,y) = p_0^y(t,x;s,y) + \int_s^t\int_{\mathbb{R}^l} p_0^z(t,x;\tau,z)\Phi(\tau,z;s,y)\,dz\,d\tau,$$

where Φ solves the Volterra equation

$$\Phi(t,x;s,y) = F(t,x;s,y) + \int_s^t\int_{\mathbb{R}^l} F(t,x;\tau,z)\Phi(\tau,z;s,y)\,dz\,d\tau. \tag{12}$$

Under the Dini condition $\int_0^1 \frac{\omega(\delta)}{\delta} d\delta < \infty$, the series converges uniformly on every finite time interval, and the solution Φ satisfies a Gaussian bound

$$|\Phi(t,x;s,y)| \le C_\Phi\,(t-s)^{-l/2} \exp\Big(-\frac{c_\Phi |x-y|^2}{t-s}\Big), \tag{13}$$

with C_Φ, c_Φ depending only on ν, μ, ω (and *not* on $t-s$). The proof is standard; see e.g. [6, Chapter 1].

Step 2: Removing the time horizon by scaling. The parametrix construction on an interval $0 < t-s \le T$ yields constants that may depend on T. However, the special form of the estimates (9)–(13) is scale-invariant: if we replace (t,s,x,y) by $(\lambda^2 t, \lambda^2 s, \lambda x, \lambda y)$ and define

$$\tilde{p}_0(t,x;s,y) = \lambda^l p_0(\lambda^2 t, \lambda x; \lambda^2 s, \lambda y),$$

then $\tilde{p}_0$ is the fundamental solution of the same equation with coefficients $\tilde{a}(t,x) = a(\lambda^2 t, \lambda x)$. The ellipticity constants ν, μ and the Dini modulus ω remain unchanged (the modulus for $\tilde{a}$ is $\tilde{\omega}(r) = \omega(\lambda r)$, which still satisfies the Dini condition). Consequently, it suffices to prove the Gaussian bounds for a single reference time, say $t-s=1$. The bounds for arbitrary $t-s$ then follow by scaling back: set $\lambda = (t-s)^{-1/2}$ and apply the scaling transformation to the solution for unit time. This argument shows that the constants c_1, C_1 in the final bounds can be chosen independent of $t-s$ (and hence of any T). We therefore only need to establish the bounds for the case $t-s=1$, which we do next.

Step 3: Upper Gaussian bound for p_0. From the representation,

$$p_0 = p_0^y + \int_s^t \int_{\mathbb{R}^l} p_0^z(\,\cdot\,,\tau,z)\,\Phi(\tau,z;s,y)\,dz\,d\tau =: I_0 + I_1.$$

By (9), $|I_0| \le C_0 (t-s)^{-l/2} e^{-c_0|x-y|^2/(t-s)}$. For I_1, use (9) for p_0^z and (13) for Φ:

$$|p_0^z(t,x;\tau,z)| \le C_0 (t-\tau)^{-l/2} e^{-c_0|x-z|^2/(t-\tau)},$$
$$|\Phi(\tau,z;s,y)| \le C_\Phi (\tau-s)^{-l/2} e^{-c_\Phi|z-y|^2/(\tau-s)}.$$

The spatial convolution gives

$$\int_{\mathbb{R}^l} e^{-c_0|x-z|^2/(t-\tau)} e^{-c_\Phi|z-y|^2/(\tau-s)} dz$$
$$= \Big(\frac{\pi}{\tilde{c}}\Big)^{l/2} \Big(\frac{(t-\tau)(\tau-s)}{t-s}\Big)^{l/2} e^{-\tilde{c}|x-y|^2/(t-s)},$$

where $\tilde{c} = \frac{c_0 c_\Phi}{c_0 + c_\Phi}$. Hence

$$\int_{\mathbb{R}^l} |p_0^z||\Phi|\, dz \le C_0 C_\Phi \Big(\frac{\pi}{\tilde{c}}\Big)^{l/2} (t-s)^{-l/2} e^{-\tilde{c}|x-y|^2/(t-s)}.$$

Integrating in τ yields a factor $(t-s)$. We absorb this factor into the exponential by noting that for any $r \ge 0$,

$$(t-s) e^{-\tilde{c}|x-y|^2/(t-s)} \le \frac{2}{\tilde{c}}\, e^{-(\tilde{c}/2)|x-y|^2/(t-s)}.$$

Thus

$$|I_1| \le C_0 C_\Phi \Big(\frac{\pi}{\tilde{c}}\Big)^{l/2} \frac{2}{\tilde{c}}\, (t-s)^{-l/2} e^{-(\tilde{c}/2)|x-y|^2/(t-s)}.$$

Choosing $c_1 = \min(c_0, \tilde{c}/2)$ and

$$C_1 = \max\Big(C_0,\ C_0 C_\Phi \Big(\frac{\pi}{\tilde{c}}\Big)^{l/2} \frac{2}{\tilde{c}}\Big),$$

we obtain the Gaussian upper bound

$$p_0(t,x;s,y) \le C_1 (t-s)^{-l/2} e^{-c_1|x-y|^2/(t-s)}.$$

Step 4: First derivative bound. Differentiate the representation:

$$\nabla_x p_0 = \nabla_x p_0^y + \int_s^t \int_{\mathbb{R}^l} \nabla_x p_0^z(t,x;\tau,z)\, \Phi(\tau,z;s,y)\, dz\, d\tau.$$

By (10), $|\nabla_x p_0^y| \le C_0 (t-s)^{-(l+1)/2} e^{-c_0|x-y|^2/(t-s)}$. The same convolution estimate as above, but now with $(t-\tau)^{-(l+1)/2}(\tau-s)^{-l/2}$, gives a factor $(t-s)^{-(l+1)/2}$ after integration. The extra $(t-s)$ from the time integral is again absorbed into the exponential. Hence there exists a constant C_1 (possibly enlarged) such that

$$|\nabla_x p_0(t,x;s,y)| \le C_1 (t-s)^{-(l+1)/2} e^{-c_1|x-y|^2/(t-s)}.$$

Step 5: Lower Gaussian bound via Harnack. For fixed y, s, $u(t,x) = p_0(t,x;s,y)$ is a positive weak solution of $L_0 u = 0$ for $t > s$. By Theorem 6, it satisfies the Harnack inequality with a constant C_H depending only on ν, μ, l. From this we derive the lower bound.

Diagonal lower bound. Using the already proved upper bound, for any $A > 0$,

$$\int_{|x-y| \ge A\sqrt{t-s}} u(t,x)\, dx \le C_1 e^{-c_1 A^2}.$$

Choose A so large that $C_1 e^{-c_1 A^2} \le 1/2$. Then

$$\int_{|x-y|\le A\sqrt{t-s}} u(t,x)\,dx \ge \frac{1}{2}.$$

Apply the Harnack inequality to the cylinder $(t-(A\sqrt{t-s})^2, t)\times B_{A\sqrt{t-s}}(y)$. A standard argument (see [3] or [8, Theorem 8.11]) yields a constant $c_* > 0$ such that

$$u(t,y) \ge c_*(t-s)^{-l/2}. \tag{14}$$

Near-diagonal lower bound. If $|x-y| \le \sqrt{t-s}$, then two applications of Harnack give

$$u(t,x) \ge \frac{1}{C_H^2} u(t,y) \ge \frac{c_*}{C_H^2}(t-s)^{-l/2}.$$

Thus there exists $c_{nd} > 0$ such that

$$|x-y| \le \sqrt{t-s} \quad\Longrightarrow\quad p_0(t,x;s,y) \ge c_{nd}(t-s)^{-l/2}. \tag{15}$$

Chaining for arbitrary x, y. Set $R = |x-y|$, $h = t-s$. Choose an integer $n = \lceil 4R^2/h \rceil$. Let $\delta = h/n$ and $x_0 = y, x_1, \dots, x_n = x$ equally spaced on the segment, so $|x_i - x_{i-1}| = R/n \le \frac{1}{2}\sqrt{\delta}$. By the semigroup property,

$$p_0(t,x;s,y) = \int_{\mathbb{R}^{l(n-1)}} \prod_{i=1}^{n} p_0(\tau_i, z_i; \tau_{i-1}, z_{i-1})\,dz_1 \cdots dz_{n-1},$$

with $\tau_i = s + i\delta$, $z_0 = y$, $z_n = x$. Restrict the integration over z_i to balls $B_i = B(x_i, \varepsilon\sqrt{\delta})$ with $\varepsilon = \frac{1}{4}$. Then for any $z_i \in B_i$, $z_{i-1} \in B_{i-1}$,

$$|z_i - z_{i-1}| \le 2\varepsilon\sqrt{\delta} + \frac{R}{n} \le \frac{1}{2}\sqrt{\delta} + \frac{1}{2}\sqrt{\delta} = \sqrt{\delta},$$

so (15) applies to each factor: $p_0(\tau_i, z_i; \tau_{i-1}, z_{i-1}) \ge c_{nd}\,\delta^{-l/2}$. The volume of each B_i is $\omega_l(\varepsilon\sqrt{\delta})^l = \omega_l \varepsilon^l \delta^{l/2}$. Hence

$$p_0(t,x;s,y) \ge (c_{nd})^n (\omega_l\varepsilon^l)^{n-1}\delta^{-l/2} = (c_{nd})^n(\omega_l\varepsilon^l)^{n-1} n^{l/2} h^{-l/2}.$$

Now $n \asymp R^2/h$, so $(c_{nd})^n \ge \exp(-CR^2/h)$ and $n^{l/2} \le \exp(CR^2/h)$ after adjusting constants. Therefore there exist $c, C > 0$ (depending only on ν, μ, ω, l) such that

$$p_0(t,x;s,y) \ge c\,h^{-l/2}\exp\Big(-\frac{|x-y|^2}{Ch}\Big).$$

This completes the proof of the lower Gaussian bound.

Thus all claims of Theorem 5 are established. □

4 Moser iteration and the Harnack inequality

In this section we prove the Harnack inequality for non-negative weak solutions of

$$L_0 u = \partial_t u - \nabla \cdot (a(t,x)\nabla u) = 0$$

under the sole assumption of uniform ellipticity (Assumption 2). The coefficients may depend on t in a measurable way; no spatial continuity is required. The constants depend only on the ellipticity constants ν, μ and the dimension l. The proof follows Moser's iteration [11, 12] as presented in [8].

4.1 Parabolic Sobolev inequality

Let $Q_R(t_0, x_0) = (t_0 - R^2, t_0) \times B_R(x_0)$. When the center is clear we write simply Q_R. The parabolic cylinders are scaled in the natural way: the time length is R^2, the spatial radius is R.

Lemma 2 (Parabolic Sobolev inequality)**.** *There exists a constant $C_S = C_S(l) > 0$ such that for any $u \in L^2(t_0 - R^2, t_0; H^1(B_R(x_0)))$,*

$$\Big(\frac{1}{|Q_R|}\iint_{Q_R} |u|^{2(1+2/l)}\,dx\,dt\Big)^{\frac{l}{l+2}} \le C_S \frac{1}{|Q_R|}\iint_{Q_R} \big(|u|^2 + R^2|\nabla u|^2\big)\,dx\,dt.$$

Proof. The inequality is a parabolic analogue of the classical Sobolev embedding. It is obtained by combining the Gagliardo–Nirenberg inequality in space with Hölder's inequality in time.

Scaling. If we rescale the cylinder to unit size by setting $\tilde{x} = (x - x_0)/R$, $\tilde{t} = (t - t_0)/R^2$, and define $\tilde{u}(\tilde{t}, \tilde{x}) = u(t_0 + R^2\tilde{t},\, x_0 + R\tilde{x})$, then $|\tilde{\nabla}\tilde{u}| = R|\nabla u|$ and $d\tilde{x}\,d\tilde{t} = R^{-(l+2)}dx\,dt$. The inequality for u on Q_R is equivalent to the same inequality for $\tilde{u}$ on the unit cylinder $Q_1(0,0)$ but with the term $|\nabla u|^2$ replaced by $|\tilde{\nabla}\tilde{u}|^2$ (the factor R^2 disappears from the right-hand side after the change of variables). The constant C_S therefore does not depend on R.

Proof for the unit cylinder. For the unit cylinder $Q_1 = (-1, 0) \times B_1(0)$, one uses the Sobolev inequality in the spatial variables: for each fixed $t \in (-1, 0)$,

$$\Big(\int_{B_1} |u(t,x)|^{2l/(l-2)}\,dx\Big)^{(l-2)/l} \le C_1\Big(\int_{B_1} |\nabla u(t,x)|^2\,dx + \int_{B_1} |u(t,x)|^2\,dx\Big),$$

where C_1 depends only on l. Then integrate in time and apply Hölder's inequality to raise the exponent to $2(1 + 2/l)$. The details are standard; see e.g. [8, Chapter 2] or [9]. □

Remark. The factor R^2 on the right-hand side is essential for the inequality to be dimensionally correct and invariant under parabolic scaling. Without it, the constant would depend on R in a non-trivial way. The inequality is valid for functions that vanish on the lateral boundary; for arbitrary functions in $L^2(H^1)$ it still holds after adding a cutoff, as we will use in the energy estimates with a smooth partition of unity.

4.2 Energy estimates with cutoffs

Let $u \geq 0$ be a weak solution of $L_0 u = 0$ in Q_{2R}. For any $p \geq 1$ we introduce a smooth cutoff function $\zeta(x)$ such that $0 \leq \zeta \leq 1$, $\zeta \equiv 1$ on B_R, $\zeta \equiv 0$ on $\mathbb{R}^l \setminus B_{2R}$, and $|\nabla \zeta| \leq C/R$. Multiply the equation by $\zeta^2 u^{p-1}$ and integrate over $\mathbb{R}^l$. Using integration by parts and the ellipticity,

$$\frac{1}{p}\frac{d}{dt}\int \zeta^2 u^p \, dx + \nu(p-1)\int \zeta^2 u^{p-2}|\nabla u|^2 \, dx \leq 2\int \zeta u^{p-1}|a\nabla u \cdot \nabla \zeta| \, dx.$$

By Young's inequality,

$$2\zeta u^{p-1}|a\nabla u \cdot \nabla \zeta| \leq \frac{\nu(p-1)}{2}\zeta^2 u^{p-2}|\nabla u|^2 + \frac{C(\nu,\mu)}{\nu(p-1)}u^p|\nabla \zeta|^2.$$

Thus

$$\frac{1}{p}\frac{d}{dt}\int \zeta^2 u^p \, dx + \frac{\nu(p-1)}{2}\int \zeta^2 u^{p-2}|\nabla u|^2 \, dx \leq C(\nu,\mu)\int u^p|\nabla \zeta|^2 \, dx.$$

Observe that

$$|\nabla(\zeta u^{p/2})|^2 = \zeta^2|\nabla(u^{p/2})|^2 + 2\zeta u^{p/2}\nabla\zeta \cdot \nabla(u^{p/2}) + u^p|\nabla\zeta|^2.$$

Using $|\nabla(u^{p/2})|^2 = \frac{p^2}{4}u^{p-2}|\nabla u|^2$, we can absorb the cross term again by Young's inequality. A standard computation (see [8, Lemma 8.1]) yields the following energy estimate: for any $t_1 < t_2$,

$$\sup_{t\in(t_1,t_2)}\int_{B_R} u(t)^p \, dx + \frac{4\nu(p-1)}{p}\int_{t_1}^{t_2}\int_{B_R}|\nabla(u^{p/2})|^2 \, dx\, dt \leq \frac{C}{R^2}\int_{t_1}^{t_2}\int_{B_{2R}} u^p \, dx\, dt + \int_{B_{2R}} u(t_1)^p \, dx. \tag{16}$$

The constant C depends only on ν, μ.

4.3 Local boundedness: from L^2 to L^∞

We now prove that for any $0 < \theta < 1$,

$$\sup_{Q_{\theta R}} u \le \frac{C}{(1-\theta)^{(l+2)/2}} \Big(\frac{1}{|Q_R|} \iint_{Q_R} u^2 \, dx\, dt \Big)^{1/2}, \tag{17}$$

where C depends only on ν, μ, l. The proof follows a standard Moser iteration; see [8, Theorem 8.2] for the complete details. Here we outline the main steps.

Define a decreasing sequence of radii

$$R_k = \theta_k R, \qquad \theta_k = \theta + (1-\theta)(1-2^{-k}), \quad k = 0, 1, 2, \dots$$

Thus $R_0 = \theta R$, $R_k \nearrow R$, and $R_{k+1} - R_k = (1-\theta) R\, 2^{-k-1}$. Corresponding time levels are

$$t_k = t_0 - R_k^2, \qquad Q_k = Q_{R_k}.$$

Set $p_k = 2(1 + 2/l)^k$, $k = 0, 1, 2, \dots$. Note that $p_0 = 2$ and $p_k \to \infty$.

We claim that for every $k \ge 0$,

$$\Big(\frac{1}{|Q_{k+1}|} \iint_{Q_{k+1}} u^{p_{k+1}} \, dx\, dt \Big)^{1/p_{k+1}} \le C_0^{1/p_k} \Big(\frac{1}{|Q_k|} \iint_{Q_k} u^{p_k} \, dx\, dt \Big)^{1/p_k}, \tag{18}$$

with a constant C_0 depending only on ν, μ, l and the ratio $R_k/(R_k - R_{k+1})$. Because $R_k - R_{k+1} \ge (1-\theta) R\, 2^{-k-2}$, the factor $(1-\theta)^{-1}$ will eventually appear.

To prove (18), apply the parabolic Sobolev inequality (Lemma 2) to the function $v = u^{p_k/2}$ on the cylinder Q_k. We obtain

$$\iint_{Q_k} u^{p_{k+1}} \, dx\, dt = \iint_{Q_k} v^{2(1+2/l)} \, dx\, dt$$

$$\le C_S \Big(\iint_{Q_k} (R_k^2 |\nabla v|^2 + v^2) \, dx\, dt \Big)^{1+2/l}.$$

Now use the energy estimate (16) with $p = p_k$ and with the time interval chosen so that the spatial domain is $B_{R_{k+1}}$ and the time interval is (t_{k+1}, t_k). A standard cutoff argument (see e.g. [8, Lemma 8.1]) shows that

$$\iint_{Q_k} (R_k^2 |\nabla v|^2 + v^2) \, dx\, dt \le \frac{C_1}{(1-\theta)^2} \iint_{Q_{k+1}} u^{p_k} \, dx\, dt,$$

where C_1 depends on ν, μ and the dimension. Substituting this into the previous inequality and taking the p_{k+1}-th root yields (18) with

$$C_0 = \bigl(C_S C_1^{1+2/l}\bigr)^{1/(p_k(1+2/l))} \cdot \bigl((1-\theta)^{-2(1+2/l)}\bigr)^{1/p_{k+1}}.$$

After iterating (18) from $k=0$ to $k=K-1$ we get

$$\Bigl(\frac{1}{|Q_K|}\iint_{Q_K} u^{p_K}\,dx\,dt\Bigr)^{1/p_K} \le \prod_{k=0}^{K-1} C_0^{1/p_k}\Bigl(\frac{1}{|Q_0|}\iint_{Q_0} u^2\,dx\,dt\Bigr)^{1/2}.$$

As $K\to\infty$, the left-hand side converges to $\sup_{Q_{\theta R}} u$ (by the Lebesgue differentiation theorem and the fact that u is continuous – solutions are Hölder continuous, see [13]). The product $\prod_{k=0}^{\infty} C_0^{1/p_k}$ is bounded by a constant depending only on ν, μ, l and $(1-\theta)^{-(l+2)/2}$ (the geometric series $\sum 1/p_k$ converges). This establishes (17).

4.4 Reverse Hölder inequality and the Harnack estimate

To obtain a lower bound for $\inf u$, we consider $v = u^{-1}$. Because $u>0$ in Q_{2R} (by the strong maximum principle for parabolic equations, which holds for divergence-form operators with bounded coefficients), a direct computation shows that v satisfies

$$\partial_t v - \nabla\cdot(a\nabla v) + 2a\nabla \ln u\cdot\nabla v = 0.$$

This is a linear parabolic equation with a lower-order term that is locally integrable. The Moser iteration can be adapted to such equations; the crucial point is that the drift $a\nabla\ln u$ is in L^2_{loc} because of the energy estimate (16) with $p=2$. The same iterative argument applied to v yields

$$\inf_{Q_{\theta R}} u \ge \frac{c}{(1-\theta)^{(l+2)/2}}\Bigl(\frac{1}{|Q_R|}\iint_{Q_R} u^{-2}\,dx\,dt\Bigr)^{-1/2}. \tag{19}$$

The constant $c>0$ depends only on ν, μ, l. (See [11] or [8, Theorem 8.7] for the detailed proof.)

Now combine (17) and (19) with $\theta = 1/2$. By Hölder's inequality,

$$\frac{1}{|Q_R|}\iint_{Q_R} u^2\,dx\,dt \cdot \frac{1}{|Q_R|}\iint_{Q_R} u^{-2}\,dx\,dt \ge 1,$$

hence

$$\Bigl(\frac{1}{|Q_R|}\iint_{Q_R} u^2\,dx\,dt\Bigr)^{1/2} \le \Bigl(\frac{1}{|Q_R|}\iint_{Q_R} u^{-2}\,dx\,dt\Bigr)^{-1/2}.$$

Therefore,

$$\sup_{Q_{R/2}} u \le \frac{C}{(1/2)^{(l+2)/2}} \Big(\frac{1}{|Q_R|}\iint_{Q_R} u^2\,dx\,dt\Big)^{1/2}$$

$$\le \frac{C}{2^{(l+2)/2}} \Big(\frac{1}{|Q_R|}\iint_{Q_R} u^{-2}\,dx\,dt\Big)^{-1/2} \le C_H \inf_{Q_{R/2}} u,$$

where $C_H = \frac{C}{2^{(l+2)/2}} \cdot \frac{1}{c}$ (with c from (19)) depends only on ν, μ, l. This is the Harnack inequality.

Theorem 6 (Harnack inequality for L_0)**.** *There exists a constant $C_H > 0$ depending only on ν, μ and the dimension l such that for any non-negative weak solution u of $L_0 u = 0$ in Q_{2R},*

$$\sup_{Q_{R/2}} u \le C_H \inf_{Q_{R/2}} u.$$

4.5 From Harnack to the lower Gaussian bound for p_0

We now apply the Harnack inequality (Theorem 6) to the fundamental solution $p_0(t, x; s, y)$ of L_0. For fixed y and s, the function $u(t, x) = p_0(t, x; s, y)$ is a positive weak solution for $t > s$. The Harnack inequality gives a constant $C_H > 0$ depending only on ν, μ, l such that for any parabolic cylinder Q_{2R},

$$\sup_{Q_{R/2}} u \le C_H \inf_{Q_{R/2}} u.$$

Diagonal lower bound. Using the Gaussian upper bound (already proved for p_0 with scaling-invariant constants) and the fact that $\int u(t, x)\,dx = 1$ for all $t > s$, one can show that there exists $c_d > 0$ such that

$$p_0(t, y; s, y) \ge c_d\,(t - s)^{-l/2}, \qquad 0 < t - s \le 1. \tag{20}$$

The proof is standard: choose A large so that the Gaussian upper bound gives $\int_{|x-y|\ge A\sqrt{t-s}} u(t, x)\,dx \le 1/2$. Then the complement ball contains at least half the mass. Apply Harnack to relate the value at the center to the average over that ball, and use the fact that the volume of the ball is $\omega_l (A\sqrt{t-s})^l$. The constant c_d depends only on ν, μ, l and A, but A is fixed once and for all by the upper bound constants, so c_d is universal. For $t - s > 1$, the bound follows by scaling (see below).

Near-diagonal lower bound. Let $|x-y| \le \sqrt{t-s}$. Choose an intermediate time $\tau = t - |x-y|^2$. Then both (t,x) and (t,y) lie in the top half of the cylinders with radius $|x-y|$ centered at y and x respectively. Two applications of the Harnack inequality (one forward, one backward) give

$$p_0(t,x;s,y) \ge \frac{1}{C_H^2}\, p_0(t,y;s,y) \ge \frac{c_d}{C_H^2}\,(t-s)^{-l/2}.$$

Thus there exists $c_{nd} > 0$ such that

$$|x-y| \le \sqrt{t-s} \quad \Longrightarrow \quad p_0(t,x;s,y) \ge c_{nd}\,(t-s)^{-l/2}. \tag{21}$$

Chaining to arbitrary x, y (for $t-s \le 1$). Let $R = |x-y|$, $h = t-s \le 1$. Choose an integer $n = \lceil 4R^2/h \rceil$. Set $\delta = h/n$ and partition the time interval into n equal subintervals: $\tau_i = s + i\delta$, $i = 0, \dots, n$, with $\tau_0 = s$, $\tau_n = t$. On the spatial side, take points $x_0 = y, x_1, \dots, x_n = x$ equally spaced along the line segment, so $|x_i - x_{i-1}| = R/n \le \frac{1}{2}\sqrt{\delta}$. By the semigroup property,

$$p_0(t,x;s,y) = \int_{\mathbb{R}^{l(n-1)}} \prod_{i=1}^{n} p_0(\tau_i, z_i; \tau_{i-1}, z_{i-1})\, dz_1 \cdots dz_{n-1},$$

with $z_0 = y$, $z_n = x$. Restrict each integration z_i to the ball $B_i = B(x_i, \varepsilon\sqrt{\delta})$ with $\varepsilon = 1/4$. Then for any $z_i \in B_i$, $z_{i-1} \in B_{i-1}$,

$$\begin{aligned} |z_i - z_{i-1}| &\le |z_i - x_i| + |x_i - x_{i-1}| + |x_{i-1} - z_{i-1}| \\ &\le \varepsilon\sqrt{\delta} + \frac{R}{n} + \varepsilon\sqrt{\delta} \le \frac{1}{2}\sqrt{\delta} + \frac{1}{2}\sqrt{\delta} = \sqrt{\delta}. \end{aligned}$$

Hence each factor satisfies the near-diagonal bound (21): $p_0(\tau_i, z_i; \tau_{i-1}, z_{i-1}) \ge c_{nd}\,\delta^{-l/2}$. The volume of each ball B_i is $\omega_l(\varepsilon\sqrt{\delta})^l = \omega_l(1/4)^l\delta^{l/2}$. Therefore,

$$p_0(t,x;s,y) \ge (c_{nd})^n(\omega_l(1/4)^l)^{n-1}\delta^{-l/2} = (c_{nd})^n(\omega_l(1/4)^l)^{n-1}n^{l/2}h^{-l/2}.$$

Since $n \asymp R^2/h$, we have $(c_{nd})^n \ge \exp(-CR^2/h)$ and $n^{l/2} \le \exp(CR^2/h)$ after adjusting constants. Consequently, there exist $c, C > 0$ depending only on ν, μ, l such that for all $0 < h \le 1$,

$$p_0(t,x;s,y) \ge c\,h^{-l/2}\exp\Big(-\frac{C|x-y|^2}{h}\Big).$$

Extension to arbitrary $t-s$ by scaling. For $t-s>1$, set $\lambda=(t-s)^{-1/2}\le 1$ and use the scaling argument described in Section 3. The scaled fundamental solution $\tilde{p}_0(\tau,\xi;\sigma,\eta)=\lambda^l p_0(\lambda^{-2}\tau,\lambda^{-1}\xi;\lambda^{-2}\sigma,\lambda^{-1}\eta)$ satisfies the same bounds with $t-s=1$. Rescaling back yields the lower bound for all $t-s>0$ with constants c,C that are independent of the time horizon.

Thus, together with the upper bound from Theorem 5 (which is also scale-invariant), we obtain the two-sided estimate

$$c\,(t-s)^{-l/2}\exp\Big(-\frac{C|x-y|^2}{t-s}\Big)$$

$$\le p_0(t,x;s,y)\le C\,(t-s)^{-l/2}\exp\Big(-\frac{c|x-y|^2}{t-s}\Big)$$

for all $0<t-s$ and all $x,y\in\mathbb{R}^l$, where $c,C>0$ depend only on ν,μ,ω,l (and not on T).

5 Perturbation by lower-order terms

Now consider the full operator

$$L=L_0+B,\qquad Bu=b\cdot\nabla u-\nabla\cdot(\tilde{b}u)+Vu,$$

with coefficients satisfying Assumption 4. Let $U_{t,s}$ be the evolution family generated by L on $L^1(\mathbb{R}^l)\cap L^\infty(\mathbb{R}^l)$ (its existence will be proved via the Duhamel series). Let $P_{t,s}$ be the evolution family generated by $L_0=\partial_t-\nabla\cdot(a\nabla)$; its kernel $p_0(t,x;s,y)$ satisfies the two-sided Gaussian bounds and derivative estimates from Theorem 5. The families are related by the Duhamel formula

$$U_{t,s}=P_{t,s}-\int_s^t U_{t,\tau}B(\tau)P_{\tau,s}\,d\tau, \tag{22}$$

which is understood as an identity on L^1. Iterating gives the series expansion

$$U_{t,s}=\sum_{n=0}^{\infty}(-1)^n\int_{s<\tau_1<\cdots<\tau_n<t}P_{t,\tau_n}B(\tau_n)P_{\tau_n,\tau_{n-1}}\cdots B(\tau_1)P_{\tau_1,s}\,d\tau_1\cdots d\tau_n, \tag{23}$$

with the $n=0$ term being $P_{t,s}$. Our goal is to show that this series converges in operator norm on L^1 and L^∞ and that its kernel satisfies two-sided Gaussian bounds.

5.1 Estimates for the perturbation on short time intervals

For a function f on $\mathbb{R}^l$, denote the Dirac delta at z by δ_z. For the operator $B(\tau)$ we have

$$B(\tau)P_{\tau,s}\delta_z = b(\tau,\cdot)\cdot\nabla P_{\tau,s}(\cdot,z) - \nabla\cdot\big(\tilde{b}(\tau,\cdot)P_{\tau,s}(\cdot,z)\big) + V(\tau,\cdot)P_{\tau,s}(\cdot,z).$$

We need to estimate the L^1 norm of this quantity. Using the derivative estimate from Theorem 5:

$$|\nabla_x p_0(\tau,x;s,z)| \le C_0(\tau-s)^{-(l+1)/2}\exp\Big(-\frac{|x-z|^2}{C_0(\tau-s)}\Big)$$

$$\le C_1(\tau-s)^{-1/2}K_\chi(\tau-s,x-z),$$

where $K_\chi(t,x) = (4\pi\chi t)^{-l/2}e^{-|x|^2/(4\chi t)}$ with $\chi = \nu/2$ (the constant C_1 depends only on ν,μ). Similarly, $|p_0(\tau,x;s,z)| \le CK_\chi(\tau-s,x-z)$.

Estimate for the adjoint drift term. We consider the operator $B_{\mathrm{adj}}(\tau)u = -\nabla\cdot(\tilde{b}(\tau,\cdot)u)$. For fixed z and s, set $u_\tau(x) = P_{\tau,s}(x,z)$. We need an L^1 bound for $B_{\mathrm{adj}}(\tau)u_\tau$.

Lemma 3. *For every $\tau > s$,*

$$\|B_{adj}(\tau)P_{\tau,s}\delta_z\|_{L^1} \le C\int_{\mathbb{R}^l}|\tilde{b}(\tau,y)|\,\big|\nabla_y P_{\tau,s}(y,z)\big|\,dy,$$

where the constant C depends only on the dimension l and the ellipticity constants.

Proof. The L^1 norm of a function (or measure) w can be expressed via duality:

$$\|w\|_{L^1} = \sup_{\phi\in C_c^\infty,\ \|\phi\|_\infty\le 1}\int w(x)\,\phi(x)\,dx.$$

For $w = B_{\mathrm{adj}}(\tau)P_{\tau,s}\delta_z$, integration by parts gives

$$\int\big(-\nabla\cdot(\tilde{b}u_\tau)\big)(x)\,\phi(x)\,dx = \int\tilde{b}(\tau,x)u_\tau(x)\cdot\nabla\phi(x)\,dx.$$

Now we use the fact that the dual space of $W_0^{1,\infty}(\mathbb{R}^l)$ (Lipschitz functions with compact support and zero boundary) is the space of divergence-measure vector fields, and the norm is

$$\|\nabla\cdot v\|_{L^1} = \sup_{\phi\in C_c^\infty,\ \|\nabla\phi\|_\infty\le 1}\int v\cdot\nabla\phi.$$

Thus we may restrict the supremum to test functions with $\|\nabla\phi\|_\infty \le 1$. For such ϕ,

$$\Big|\int \tilde{b} u_\tau \cdot \nabla\phi\Big| \le \int |\tilde{b}|\, u_\tau\, |\nabla\phi| \le \int |\tilde{b}|\, u_\tau.$$

This gives $\|B_{\mathrm{adj}}(\tau)u_\tau\|_{L^1} \le \int |\tilde{b}| u_\tau$, which involves u_τ rather than its gradient. To obtain the desired bound with $|\nabla u_\tau|$, we use a duality argument with the adjoint evolution.

Consider the operator $B_{\mathrm{adj}}(\tau)P_{\tau,s}$ as a map from L^1 to L^1. Its adjoint maps L^∞ to L^∞ and is given by

$$\big(B_{\mathrm{adj}}(\tau)P_{\tau,s}\big)^* = P^*_{\tau,s}\,\tilde{b}(\tau,\cdot)\cdot\nabla,$$

where $P^*_{\tau,s}$ is the adjoint of $P_{\tau,s}$ (the evolution of the backward equation). The kernel of $P^*_{\tau,s}$ is the same as that of $P_{\tau,s}$ with variables swapped: $p_0(\tau, y; s, z)$. In particular, it satisfies the same Gaussian bounds and derivative estimates as p_0.

For any $g \in L^\infty$,

$$\big(P^*_{\tau,s}(\tilde{b}\cdot\nabla g)\big)(z) = \int_{\mathbb{R}^l} \tilde{b}(\tau,y)\cdot\nabla_y p_0(\tau,y;s,z)\, g(y)\, dy.$$

Hence

$$\big|\big(P^*_{\tau,s}(\tilde{b}\cdot\nabla g)\big)(z)\big| \le \int |\tilde{b}(\tau,y)|\, |\nabla_y p_0(\tau,y;s,z)|\, dy\ \|g\|_\infty.$$

Therefore,

$$\|P^*_{\tau,s}(\tilde{b}\cdot\nabla g)\|_\infty \le \Big(\int |\tilde{b}(\tau,y)|\, |\nabla_y p_0(\tau,y;s,z)|\, dy\Big)\|g\|_\infty.$$

Now, for any g with $\|g\|_\infty \le 1$,

$$\big\langle B_{\mathrm{adj}}(\tau)P_{\tau,s}\delta_z,\, g\big\rangle = \big\langle \delta_z,\, P^*_{\tau,s}\tilde{b}\cdot\nabla g\big\rangle$$

$$= \big(P^*_{\tau,s}\tilde{b}\cdot\nabla g\big)(z) \le \int |\tilde{b}(\tau,y)|\, |\nabla_y p_0(\tau,y;s,z)|\, dy.$$

Taking the supremum over all such g yields

$$\|B_{\mathrm{adj}}(\tau)P_{\tau,s}\delta_z\|_{L^1} \le \int |\tilde{b}(\tau,y)|\, |\nabla_y p_0(\tau,y;s,z)|\, dy,$$

which is exactly the desired estimate (with $C = 1$). This completes the proof. □

Lemma 4. *Under Assumption 4, define*

$$\beta(h) = \operatorname*{ess\,sup}_{z,s} \int_s^{s+h} \|B(\tau)P_{\tau,s}\delta_z\|_{L^1} d\tau.$$

Then $\beta(h) \to 0$ as $h \to 0$. More precisely, there exists a constant C depending only on ν, μ, l such that

$$\beta(h) \le C\Big(\beta_b(h)^{\frac{1}{l+1}} + \beta_{\tilde{b}}(h)^{\frac{1}{l+1}} + \beta_V(h)^{\frac{1}{l+2}}\Big),$$

where for $f \in K_p$ we set

$$\beta_f(h) = \operatorname*{ess\,sup}_{s,z} \int_s^{s+h} \int_{\mathbb{R}^l} G(t-s, y-z)|f(t,y)|^p dy\, dt.$$

Proof. We estimate each term separately. For the drift term,

$$\int_s^{s+h} (\tau - s)^{-1/2} \int |b(\tau, y)| K_\chi(\tau - s, y - z)\, dy\, d\tau.$$

Apply Hölder's inequality in space with exponent $l+1$:

$$\int |b| K_\chi \le \Big(\int |b|^{l+1} K_\chi\Big)^{\frac{1}{l+1}} \Big(\int K_\chi\Big)^{\frac{l}{l+1}} = \Big(\int |b|^{l+1} K_\chi\Big)^{\frac{1}{l+1}}.$$

Thus the drift contribution is bounded by

$$C \int_s^{s+h} (\tau - s)^{-1/2} \Big(\int |b|^{l+1} K_\chi\Big)^{\frac{1}{l+1}} d\tau.$$

Now use Hölder's inequality in time with exponents $l+1$ and $(l+1)/l$:

$$\int_s^{s+h} (\tau - s)^{-1/2} (\cdots)^{\frac{1}{l+1}} d\tau$$

$$\le \Big(\int_s^{s+h} (\tau - s)^{-\frac{l+1}{2l}} d\tau\Big)^{\frac{l}{l+1}} \Big(\int_s^{s+h} \int |b|^{l+1} K_\chi dy\, d\tau\Big)^{\frac{1}{l+1}}.$$

Since $-\frac{l+1}{2l} > -1$ for all $l \ge 1$, the first factor equals $Ch^{\frac{l-1}{2(l+1)}}$. Hence the drift term contributes at most

$$Ch^{\frac{l-1}{2(l+1)}} \beta_b(h)^{\frac{1}{l+1}}.$$

Because $h^{\frac{l-1}{2(l+1)}} \le 1$ for $h \le 1$, we can absorb it into the constant. Therefore

$$\text{drift contribution} \le C\beta_b(h)^{\frac{1}{l+1}}.$$

The same estimate holds for the adjoint drift term with $\tilde{b}$.

For the potential term,

$$\int_s^{s+h} \int |V(\tau,y)| K_\chi(\tau - s, y - z)\, dy\, d\tau$$

$$\le \Big(\int_s^{s+h} \int |V|^{l+2} K_\chi dy\, d\tau\Big)^{\frac{1}{l+2}} \Big(\int_s^{s+h} \int K_\chi dy\, d\tau\Big)^{\frac{l+1}{l+2}}.$$

The second factor is $Ch^{\frac{l+1}{l+2}}$. Thus the potential contribution is

$$\le Ch^{\frac{l+1}{l+2}} \beta_V(h)^{\frac{1}{l+2}} \le C\beta_V(h)^{\frac{1}{l+2}}.$$

Combining the three estimates gives the stated bound. By definition of the Kato classes, $\beta_f(h) \to 0$ as $h \to 0$ (Lemma 1), hence $\beta(h) \to 0$. □

5.2 Convergence of the Duhamel series in L^1 and L^∞

Choose $h_0 > 0$ such that $\beta(h_0) \le 1/2$. For $0 < t - s \le h_0$, we estimate the operator norm of the n-th term in (23). Because $\|P_{t,\tau}\|_{1\to 1} = 1$ and $\|P_{\tau,s}\|_{1\to 1} = 1$, we have

$$\|P_{t,\tau_n} B(\tau_n) P_{\tau_n,\tau_{n-1}} \cdots B(\tau_1) P_{\tau_1,s}\|_{1\to 1} \le \prod_{i=1}^{n} \|B(\tau_i) P_{\tau_i,\tau_{i-1}}\|_{1\to 1},$$

where we set $\tau_0 = s$. Integrating over the simplex $s < \tau_1 < \cdots < \tau_n < t$ yields

$$\int_{s<\tau_1<\cdots<\tau_n<t} \prod_{i=1}^{n} \|B(\tau_i) P_{\tau_i,\tau_{i-1}}\|_{1\to 1}\, d\tau_1 \cdots d\tau_n$$

$$\le \frac{1}{n!} \Big(\int_s^t \|B(\tau) P_{\tau,s}\|_{1\to 1} d\tau\Big)^n \le \frac{\beta(t-s)^n}{n!}.$$

Hence the L^1 operator norm of the n-th term is at most $\beta(t-s)^n/n!$. Summing over n gives

$$\|U_{t,s}\|_{1\to 1} \le \sum_{n=0}^{\infty} \frac{\beta(t-s)^n}{n!} = e^{\beta(t-s)} \le e^{1/2} < \infty.$$

For L^∞, note that the adjoint of $U_{t,s}$ satisfies the same Duhamel series with B replaced by its formal adjoint $B^*u = -b\cdot\nabla u+\tilde{b}\cdot\nabla u+Vu-(\nabla\cdot b)u+(\nabla\cdot\tilde{b})u$. Under the same Kato class assumptions (and assuming $b, \tilde{b}$ are such that $\nabla\cdot b, \nabla\cdot\tilde{b}$ are also in the appropriate Kato classes, or by a density argument), the same estimate holds. For simplicity, we work with the kernel estimates.

5.3 Pointwise Gaussian upper bound for $U_{t,s}$

We prove that for sufficiently small $h_0 > 0$ (to be chosen later) and for all $0 < t - s \le h_0$, the kernel of $U_{t,s}$ satisfies

$$|U_{t,s}(x,y)| \le C\,(t-s)^{-l/2}\exp\Big(-\frac{c|x-y|^2}{t-s}\Big). \tag{24}$$

The extension to arbitrary $t-s \le T$ is then obtained by a standard chaining argument (see the end of this subsection).

5.3.1 Kernel estimates for the elementary building blocks

Let $P_{t,s}$ be the evolution family of L_0. Its kernel $p_0(t,x;s,y)$ satisfies the two-sided Gaussian bounds of Theorem 5. In particular there exist constants $C_0, c_0 > 0$ such that

$$0 \le p_0(t,x;s,y) \le C_0\,(t-s)^{-l/2}\exp\Big(-\frac{c_0|x-y|^2}{t-s}\Big). \tag{25}$$

For the first-order part $B(\tau) = b\cdot\nabla-\nabla\cdot(\tilde{b}\,\cdot)+V$ we consider the operator $Q_{\tau,s} = B(\tau)P_{\tau,s}$. Its kernel $q_{\tau,s}(z,y)$ is a distribution, but under the Kato class assumptions it is a function that satisfies a Gaussian estimate with a small factor.

Lemma 5. *There exist constants $C_1, c_1 > 0$ (depending only on the ellipticity and the Kato norms) such that for all $0 < \tau - s \le h_0$ and all $z, y \in \mathbb{R}^l$,*

$$|q_{\tau,s}(z,y)| \le C_1\,\beta(\tau-s)\,(\tau-s)^{-l/2}\exp\Big(-\frac{c_1|z-y|^2}{\tau-s}\Big), \tag{26}$$

where $\beta(h)$ is defined in Lemma 4 and satisfies $\beta(h) \to 0$ as $h \to 0$.

Proof. Recall that $Q_{\tau,s} = B(\tau)P_{\tau,s}$ and $q_{\tau,s}(z,y)$ is the kernel of this operator, i.e.,

$$(Q_{\tau,s}f)(z) = \int_{\mathbb{R}^l} q_{\tau,s}(z,y)f(y)\,dy, \qquad q_{\tau,s}(z,y) = (B(\tau)P_{\tau,s}\delta_y)(z).$$

We have three contributions from $B(\tau) = b\cdot\nabla - \nabla\cdot(\tilde{b}\,\cdot) + V$.

1. Drift term $b \cdot \nabla P_{\tau,s}$. Using the derivative estimate from Theorem 5,

$$|\nabla_x p_0(\tau, x; s, y)| \le C_0(\tau - s)^{-(l+1)/2} \exp\Big(-\frac{|x-y|^2}{C_0(\tau - s)}\Big)$$

$$\le C_d(\tau - s)^{-1/2} K_\chi(\tau - s, x - y),$$

where $K_\chi(t, x) = (4\pi\chi t)^{-l/2} \exp(-|x|^2/(4\chi t))$ with $\chi = \nu/2$ (the constant C_d depends only on ν, μ). Hence the kernel of $b \cdot \nabla P_{\tau,s}$ is bounded by

$$|b(\tau, z)|\, |\nabla_z p_0(\tau, z; s, y)| \le C_d(\tau - s)^{-1/2} |b(\tau, z)|\, K_\chi(\tau - s, z - y).$$

2. Adjoint drift term $-\nabla \cdot (\tilde{b} P_{\tau,s})$. By Lemma 3, for each fixed y the operator $-\nabla \cdot (\tilde{b} P_{\tau,s})$ maps the Dirac delta δ_y to an L^1 function. Moreover, the duality argument in the proof of that lemma shows that for almost every z,

$$\big|\big(-\nabla \cdot (\tilde{b} P_{\tau,s})\delta_y\big)(z)\big| \le \int_{\mathbb{R}^l} |\tilde{b}(\tau, w)|\, |\nabla_w p_0(\tau, w; s, y)|\, dw.$$

Indeed, approximating δ_z by a sequence of smooth functions $\phi_\varepsilon(x - z)$ with $\|\nabla \phi_\varepsilon\|_\infty \le C/\varepsilon$ and letting $\varepsilon \to 0$ yields the pointwise bound. (The right-hand side is independent of z, so the inequality holds for every z outside a null set.) Using the Gaussian estimate for $\nabla_w p_0$, we obtain

$$|q^{\mathrm{adj}}_{\tau,s}(z, y)| \le C_d(\tau - s)^{-1/2} \int_{\mathbb{R}^l} |\tilde{b}(\tau, w)|\, K_\chi(\tau - s, w - y)\, dw.$$

3. Potential term $V P_{\tau,s}$. The kernel is simply $V(\tau, z) p_0(\tau, z; s, y)$, and we have

$$|V(\tau, z) p_0(\tau, z; s, y)| \le C_0 |V(\tau, z)|\, (\tau - s)^{-l/2} K_\chi(\tau - s, z - y).$$

4. Combining the three contributions. All three terms have the form

$$(\tau - s)^{-l/2} \int_{\mathbb{R}^l} |f(\tau, w)|\, K_\chi(\tau - s, w - y)\, dw$$

multiplied by a constant that may also include a factor $(\tau - s)^{-1/2}$ for the drift and adjoint drift. However, the time integral in the definition of $\beta(\tau - s)$ already contains the factor $(\tau - s)^{-1/2}$ from the derivative estimate. To obtain a unified estimate, we define

$$\beta_b(h) = \operatorname*{ess\,sup}_{s,z} \int_s^{s+h} \int_{\mathbb{R}^l} K_\chi(t - s, w - z) |b(t, w)|^{l+1} dw\, dt,$$

and similarly for $\tilde{b}$ and V with exponents $l+1$ and $l+2$, respectively. Using Hölder's inequality exactly as in Lemma 4, we obtain for each term an estimate of the form

$$|q_{\tau,s}(z,y)|$$

$$\leq C\big(\beta_b(\tau-s)^{\frac{1}{l+1}} + \beta_{\tilde{b}}(\tau-s)^{\frac{1}{l+1}} + \beta_V(\tau-s)^{\frac{1}{l+2}}\big)(\tau-s)^{-l/2}\exp\Big(-\frac{c|z-y|^2}{\tau-s}\Big),$$

where $c>0$ depends only on χ and the constants from the Gaussian bounds. By definition, $\beta(h)$ in Lemma 4 is precisely this sum (up to a constant factor). Therefore, there exist constants $C_1, c_1 > 0$ such that

$$|q_{\tau,s}(z,y)| \leq C_1\,\beta(\tau-s)\,(\tau-s)^{-l/2}\exp\Big(-\frac{c_1|z-y|^2}{\tau-s}\Big)$$

for all $0<\tau-s\leq h_0$. The factor $\beta(\tau-s)$ tends to 0 as $\tau-s\to 0$ by Lemma 4, which completes the proof. □

5.3.2 Induction for the iterated kernels

Recall that the n-th term in the Duhamel series (23) is

$$K_n(t,x;s,y)$$

$$= \int_{s<\tau_1<\cdots<\tau_n<t}\int_{\mathbb{R}^{ln}} p_0(t,x;\tau_n,z_n)\prod_{i=1}^{n} q_{\tau_i,\tau_{i-1}}(z_i,z_{i-1})\,dz_1\cdots dz_n\,d\tau_1\cdots d\tau_n,$$

where we set $\tau_0=s$, $z_0=y$ and z_n is the integration variable for the outermost p_0. For $n=0$ we define $K_0=p_0$.

Lemma 6. *For every $n\geq 0$ and all $0<t-s\leq h_0$ (with $h_0\leq 1$ and $\beta(h_0)\leq 1/2$) we have*

$$|K_n(t,x;s,y)| \leq C^{n+1}\,\frac{(\beta(t-s))^n}{n!}\,(t-s)^{-l/2}\exp\Big(-\frac{c|x-y|^2}{t-s}\Big), \tag{27}$$

where $c=\min(c_0,c_1)$ and $C=\max\big(C_0,\,C_1(\pi/c)^{l/2},\,1\big)$.

Proof. We proceed by induction on n.

Base case $n=0$: By (25) and $c\leq c_0$ we have $|K_0|\leq C_0(t-s)^{-l/2}e^{-c_0|x-y|^2/(t-s)}\leq C^1\frac{(\beta)^0}{0!}(t-s)^{-l/2}e^{-c|x-y|^2/(t-s)}$, so (27) holds.

Induction step: Assume (27) holds for some $n \geq 0$. Then, we have

$$K_{n+1}(t,x;s,y) = \int_s^t \int_{\mathbb{R}^l} p_0(t,x;\tau,z) \Big(\int_{s<\tau_1<\cdots<\tau_n<\tau} \int_{\mathbb{R}^{ln}} q_{\tau,\tau_n}(z,z_n) \prod_{i=1}^{n} q_{\tau_i,\tau_{i-1}}(z_i,z_{i-1})\, dz_1 \cdots dz_n\, d\tau_1 \cdots d\tau_n \Big)\, dz\, d\tau.$$

The inner double integral is exactly the kernel of the operator $Q_{\tau,\tau_n} \cdots Q_{\tau_1,s}$ evaluated at (z,y). By the same inductive estimate applied to the product of n kernels q (instead of one p_0 and n q's) we obtain

$$\Big| \int \cdots \int \Big| \leq C_1^n\, (\pi/c)^{nl/2}\, \frac{\big(\int_s^\tau \beta(\sigma - s)\, d\sigma\big)^n}{n!}\, (\tau - s)^{-l/2} \exp\Big(-\frac{c|z-y|^2}{\tau - s}\Big),$$

where we used that the convolution of n Gaussians with parameter c produces a factor $(\pi/c)^{nl/2}$ and the time integrals give $(\int_s^\tau \beta)^n/n!$. (A straightforward induction on the number of q-factors yields this bound; the proof is identical to the one we are giving, with C_0 replaced by C_1.)

Now insert this estimate into the expression for K_{n+1} and use the bound (25) for p_0:

$$|K_{n+1}| \leq \int_s^t \int_{\mathbb{R}^l} C_0 (t-\tau)^{-l/2} e^{-c_0|x-z|^2/(t-\tau)} \times C_1^n (\pi/c)^{nl/2} \frac{\big(\int_s^\tau \beta\big)^n}{n!} (\tau - s)^{-l/2} e^{-c|z-y|^2/(\tau-s)}\, dz\, d\tau.$$

Since $c \leq c_0$, we may replace $e^{-c_0|x-z|^2/(t-\tau)}$ by the larger quantity $e^{-c|x-z|^2/(t-\tau)}$. Then the spatial convolution becomes

$$\int_{\mathbb{R}^l} e^{-c|x-z|^2/(t-\tau)} e^{-c|z-y|^2/(\tau-s)} dz = \Big(\frac{\pi(t-\tau)(\tau-s)}{c(t-s)}\Big)^{l/2} \exp\Big(-\frac{c|x-y|^2}{t-s}\Big).$$

Multiplying by $(t-\tau)^{-l/2}(\tau-s)^{-l/2}$ yields $(\pi/c)^{l/2}(t-s)^{-l/2} e^{-c|x-y|^2/(t-s)}$. Therefore

$$|K_{n+1}| \leq C_0 C_1^n (\pi/c)^{(n+1)l/2}\, \frac{1}{n!} (t-s)^{-l/2} e^{-c|x-y|^2/(t-s)} \int_s^t \Big(\int_s^\tau \beta(\sigma - s)\, d\sigma \Big)^n d\tau.$$

The time integral is elementary:

$$\int_s^t \Big(\int_s^\tau \beta(\sigma - s)\, d\sigma \Big)^n d\tau = \frac{1}{n+1} \Big(\int_s^t \beta(\sigma - s)\, d\sigma \Big)^{n+1}.$$

Because $t-s \le h_0 \le 1$ and β is non-decreasing, we have $\int_s^t \beta(\sigma - s)\,d\sigma \le \beta(t-s)(t-s) \le \beta(t-s)$. Hence

$$\frac{1}{n+1}\Big(\int_s^t \beta\Big)^{n+1} \le \frac{(\beta(t-s))^{n+1}}{(n+1)!}.$$

Now set $C = \max\big(C_0,\, C_1(\pi/c)^{l/2},\, 1\big)$. Then $C_0 \le C$ and $C_1(\pi/c)^{l/2} \le C$. Consequently

$$C_0 C_1^n (\pi/c)^{(n+1)l/2} = C_0\big(C_1(\pi/c)^{l/2}\big)^n (\pi/c)^{l/2} \le C \cdot C^n \cdot C = C^{n+2}.$$

Thus

$$|K_{n+1}| \le C^{n+2}\,\frac{(\beta(t-s))^{n+1}}{(n+1)!}\,(t-s)^{-l/2} e^{-c|x-y|^2/(t-s)},$$

which is exactly (27) with n replaced by $n+1$. This completes the induction. □

5.3.3 Summation and small-time upper bound

For $0 < t-s \le h_0$ we have $\beta(t-s) \le 1/2$. Summing the series (23) and using Lemma 6 gives

$$|U_{t,s}(x,y)| \le \sum_{n=0}^{\infty} C^{n+1}\frac{(\beta(t-s))^n}{n!}(t-s)^{-l/2} e^{-c|x-y|^2/(t-s)}$$

$$= Ce^{C\beta(t-s)}(t-s)^{-l/2} e^{-c|x-y|^2/(t-s)} \le Ce^{C/2}(t-s)^{-l/2} e^{-c|x-y|^2/(t-s)}.$$

Replacing $Ce^{C/2}$ by a new constant (still denoted C) we obtain (24).

5.3.4 Extension to arbitrary time intervals

For any $T > 0$ and any $0 < t-s \le T$, choose an integer $m \ge 1$ such that $(t-s)/m \le h_0$. Then split the interval $[s,t]$ into m equal subintervals:

$$U_{t,s} = U_{t,t_{m-1}} U_{t_{m-1},t_{m-2}} \cdots U_{t_1,s}, \qquad t_i = s + i\frac{t-s}{m}.$$

Each factor satisfies the Gaussian upper bound (24) with the same constants C, c. Convolving m such kernels yields (see e.g. [3, Lemma 2]) a bound of the form

$$|U_{t,s}(x,y)| \le C^m\,(t-s)^{-l/2}\exp\Big(-\frac{c|x-y|^2}{t-s}\Big)\cdot(\text{polynomial factor}),$$

where the polynomial factor comes from the finite number of convolutions and is bounded by a constant depending only on m (hence on T). Since $C^m = \exp(m \ln C) \le \exp\big((T/h_0 + 1) \ln C\big) =: C(T)$, we finally obtain

$$|U_{t,s}(x,y)| \le C(T)\,(t-s)^{-l/2} \exp\Big(-\frac{c|x-y|^2}{t-s}\Big)$$

for all $0 < t - s \le T$. This is the required Gaussian upper bound for the full evolution family.

5.4 Lower Gaussian bound via perturbation

For small times, start from the Duhamel formula (22):

$$U_{t,s} = P_{t,s} - \int_s^t U_{t,\tau} B(\tau) P_{\tau,s} d\tau.$$

Taking pointwise values and using the upper bound for $U_{t,\tau}$ (proved above) and the estimate for $\|B(\tau)P_{\tau,s}\delta_y\|_{L^1}$, we obtain

$$|U_{t,s}(x,y) - P_{t,s}(x,y)| \le \int_s^t \int |U_{t,\tau}(x,z)|\, \|B(\tau)P_{\tau,s}\delta_y\|_{L^1} dz\, d\tau$$

$$\le C\beta(t-s)(t-s)^{-l/2} \exp\Big(-\frac{|x-y|^2}{c_0(t-s)}\Big),$$

where we used that $\int |U_{t,\tau}(x,z)| dz \le C$ uniformly (by the L^1 bound on the kernel). Now recall that $P_{t,s}$ satisfies the lower bound

$$P_{t,s}(x,y) \ge c_0 (t-s)^{-l/2} \exp\Big(-\frac{|x-y|^2}{C_0(t-s)}\Big),$$

with constants $c_0, C_0 > 0$ from Theorem 5. Choose $h_1 \le h_0$ such that

$$C\beta(h_1) \le \frac{c_0}{2}.$$

Then for $0 < t - s \le h_1$,

$$U_{t,s}(x,y) \ge P_{t,s}(x,y) - \frac{c_0}{2}(t-s)^{-l/2} \exp\Big(-\frac{|x-y|^2}{c_0(t-s)}\Big).$$

If the Gaussian exponent in the error term is larger (i.e., c_0 is larger than the exponent in the lower bound of P), we need to adjust. In fact, the lower bound of P has constant C_0 in the exponent, while the error term has

constant c_0. Without loss of generality we may assume $C_0 \geq c_0$ (otherwise swap the names). Then

$$\exp\Big(-\frac{|x-y|^2}{c_0(t-s)}\Big) \leq \exp\Big(-\frac{|x-y|^2}{C_0(t-s)}\Big)$$

because $c_0 \leq C_0$ implies the exponent is larger (more negative). Thus the error term is bounded by the same Gaussian as P. Hence

$$U_{t,s}(x,y) \geq \frac{c_0}{2}(t-s)^{-l/2}\exp\Big(-\frac{|x-y|^2}{C_0(t-s)}\Big).$$

This gives the lower bound for small times.

For larger times, we again use the semigroup property. Write $t-s = mh_1 + r$ with $0 \leq r < h_1$. Then

$$U_{t,s} = U_{s+mh_1+r,s+mh_1}U_{s+mh_1,s+(m-1)h_1}\cdots U_{s+h_1,s}.$$

Each factor satisfies the lower bound with constant $c_0/2$ and exponent constant C_0. The convolution of $m+1$ such kernels yields

$$U_{t,s}(x,y) \geq \big(\tfrac{c_0}{2}\big)^{m+1}(t-s)^{-l/2}\exp\Big(-\frac{|x-y|^2}{C_0(t-s)}\Big).$$

Since $m+1 \leq (t-s)/h_1 + 1$, we have

$$\big(\tfrac{c_0}{2}\big)^{m+1} = \exp\big((m+1)\ln(c_0/2)\big) \geq \exp\big(-\frac{t-s}{h_1}|\ln(c_0/2)| - |\ln(c_0/2)|\big).$$

This factor can be absorbed into the exponential by replacing C_0 with a slightly larger constant C (because $\exp(-a(t-s))$ times $\exp(-|x-y|^2/(C_0(t-s)))$ is still a Gaussian with a modified constant). Thus we obtain

$$U_{t,s}(x,y) \geq c\,(t-s)^{-l/2}\exp\Big(-\frac{|x-y|^2}{C(t-s)}\Big)$$

for all $0 < t-s \leq T$, with constants depending only on ν, μ, ω, T and the Kato norms.

5.5 Completion of the proof of the main theorem

Combining the upper and lower bounds for $U_{t,s}$ and the existence of the evolution family via the convergent Duhamel series, we have proved Theorem 1. The constants c, C depend on the parameters as stated.

Acknowledgements. The author would like to express his sincere gratitude to the anonymous reviewer for the careful reading of the manuscript and for numerous insightful comments and suggestions that significantly improved the quality of this work.

References

[1] M. Aizenman and B. Simon, Brownian motion and Harnack inequality for Schrödinger operators, *Commun. Pure Appl. Math.* 35 (1982), 209-273.

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

[3] D.G. Aronson, Non-negative solutions of linear parabolic equations, *Ann. Scuola Norm. Sup. Pisa* 22 (1968), 607-694.

[4] E.B. Davies, *Heat Kernels and Spectral Theory*, Cambridge Tracts in Mathematics, vol. 92, Cambridge University Press, 1989.

[5] E.B. Fabes and D.W. Stroock, A new proof of Moser's parabolic Harnack inequality via the old ideas of Nash, *Arch. Rational Mech. Anal.* 96 (1986), 327-338.

[6] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, 1964.

[7] T. Kato, Schrödinger operators with singular potentials, *Israel J. Math.* 13 (1973), 135-148.

[8] N.V. Krylov, *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, Graduate Studies in Mathematics, vol. 12, Amer. Math. Soc., 1996.

[9] O.A. Ladyzhenskaya, V.A. Solonnikov and N.N. Uraltseva, *Linear and Quasi-Linear Equations of Parabolic Type*, Amer. Math. Soc., 1968.

[10] V. Liskevich and Y. Semenov, Estimates for fundamental solutions of second-order parabolic equations, *J. London Math. Soc.* 53 (1996), 567-580.

[11] J. Moser, A Harnack inequality for parabolic differential equations, *Commun. Pure Appl. Math.* 17 (1964), 101-134.

[12] J. Moser, On a pointwise estimate for parabolic differential equations, *Commun. Pure Appl. Math.* 24 (1971), 727-740.

[13] J. Nash, Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.* 80 (1958), 931-954.

[14] B. Simon, Schrödinger semigroups, *Bull. Amer. Math. Soc.* 7 (1982), 447-526.

[15] K.-T. Sturm, Heat kernel bounds on manifolds, *Math. Ann.* 292 (1992), 573-586.

[16] Q.S. Zhang, Gaussian bounds for the fundamental solutions of a class of parabolic equations, *J. Differ. Equations* 140 (1997), 317-337.