NONLOCAL DIFFUSION PROBLEMS THAT APPROXIMATE A PARABOLIC EQUATION WITH SPATIAL DEPENDENCE

ALEXIS MOLINO AND JULIO D. ROSSI

Abstract. In this paper we show that smooth solutions to the Dirichlet problem for the parabolic equation

$$v_t(x,t) = \sum_{i,j=1}^{N} a_{ij}(x) \frac{\partial^2 v(x,t)}{\partial x_i \partial x_j} + \sum_{i=1}^{N} b_i(x) \frac{\partial v(x,t)}{\partial x_i} \qquad x \in \Omega,$$

with $v(x,t) = g(x,t)$, $x \in \partial\Omega$, can be approximated uniformly by solutions of nonlocal problems of the form

$$u_t^\varepsilon(x,t) = \int_{\mathbb{R}^n} K_\varepsilon(x,y)(u^\varepsilon(y,t) - u^\varepsilon(x,t))dy, \ x \in \Omega,$$

with $u^\varepsilon(x,t) = g(x,t)$, $x \notin \Omega$, as $\varepsilon \to 0$, for an appropriate rescaled kernel K_ε. In this way we show that the usual local evolution problems with spatial dependence can be approximated by non-local ones. In the case of an equation in divergence form we can obtain an approximation with symmetric kernels, that is, $K_\varepsilon(x,y) = K_\varepsilon(y,x)$.

1. Introduction

Nonlocal diffusion problems of the form

$$u_t(x,t) = \int_{\mathbb{R}^n} K(x,y)(u(y,t) - u(x,t))dy \tag{1}$$

and variations of it, have been extensively studied in recent years (see [1, 6, 5] and references therein). Here, the kernel $K : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ is a nonnegative, smooth function such that $\int_{\mathbb{R}^N} K(x,y)dx = 1$. A physical interpretation of (1) is the following: if $K(x,y)$ is the probability distribution that individuals jump from y to x and $u(x,t)$ is the density at position x at time t, then $\int_{\mathbb{R}^N} K(x,y)u(y,t)dy$ is the rate at which individuals are arriving to position x from all other locations y. Further, with the same reasoning, $\int_{\mathbb{R}^N} K(x,y)u(x,t)dy$ is interpreted as the rate at which they are leaving position x to all other places. Hence, in the absence of external or internal sources, the density $u(x,t)$ satisfies (1) (see [1, 11, 13, 17]). This kind of nonlocal diffusion equation is relevant in applications, for example, in the study of biological dispersal of species, image processing, particle systems, elasticity and coagulation models, [2, 3, 4, 11, 12, 13].

In this work we consider the following nonlocal diffusion problem: given a bounded domain $\Omega \subset \mathbb{R}^N$, $g \in L^1\left(\left(\mathbb{R}^N \setminus \Omega\right) \times (0,\infty)\right)$ and $u_0 \in L^1(\Omega)$, find $u(x,t)$

The first author is supported by MINECO grant MTM2015-68210-P (Spain) and Junta de Andalucía FQM-116 (Spain). The second author is supported by CONICET (Argentina).

such that

$$(P_K) \qquad \begin{cases} u_t(x,t) = \displaystyle\int_{\mathbb{R}^N} K(x,y)(u(y,t)-u(x,t))dy, & x \in \Omega,\ t>0, \\ u(x,t) = g(x,t), & x \notin \Omega,\ t>0, \\ u(x,0) = u_0(x), & x \in \Omega, \end{cases}$$

where the kernel $K(x,y)$ is a positive function with compact support contained in $\Omega \times B(0,d) \subset \mathbb{R}^N \times \mathbb{R}^N$ with

$$(2) \qquad 0 < \sup_{y \in B(0,d)} K(x,y) = R(x) \in L^\infty(\Omega).$$

As we mentioned before, the integral term in the problem takes into account the individuals arriving or leaving position $x \in \Omega$ from or to other places. In this model, imposing $u(x,t) = g(x,t)$ for $x \notin \Omega$, we are prescribing the values of u outside Ω. In the particular case $g = 0$, we mean that individuals that leave Ω, die (and therefore the density outside Ω is zero).

Existence and uniqueness of solutions of (P_K) is proved in Proposition 2.1 using a fixed point argument (see also Appendix A, for an alternative proof). In Proposition 2.2 we obtain an appropriate comparison principle.

As a local counterpart to our nonlocal evolution problem, we have the following second order local parabolic differential equation with Dirichlet boundary conditions

$$(Q) \qquad \begin{cases} v_t(x,t) = \displaystyle\sum_{i,j=1}^N a_{ij}(x) \frac{\partial^2 v(x,t)}{\partial x_i \partial x_j} + \sum_i^N b_i(x) \frac{\partial v(x,t)}{\partial x_i}, & x \in \Omega, t>0, \\ v(x,t) = g(x,t), & x \in \partial\Omega, t>0, \\ v(x,0) = u_0(x), & x \in \Omega, \end{cases}$$

where the coefficients $a_{ij}(x)$, $b_i(x)$ are smooth in $\overline{\Omega}$ and $(a_{ij}(x))$ is a symmetric positive definite matrix, i.e., $a_{ij} = a_{ji}$ and $\sum_{ij} a_{ij}(x)\xi_i\xi_j \geq \alpha|\xi|^2$ for every real vector $\xi = (\xi_1, \ldots, \xi_N) \neq 0$ and for some $\alpha > 0$.

It is important to stress that here we will use that (Q) has smooth solutions. In fact, under regularity assumptions on the boundary data g, the domain Ω and the initial condition u_0, we have that the solutions of (Q) are $\mathcal{C}^{2+\alpha,1+\alpha/2}\left(\overline{\Omega} \times [0,T]\right)$. For such a regularity result we refer to [14].

Our main goal in this work is to show that the Dirichlet problem for the parabolic equation (Q) can be approximated by nonlocal problems of the form (P_K). More precisely, given $J : \mathbb{R}^N \to \mathbb{R}$ a nonnegative, radial and continuous function with compact support and finite second order momentum, we consider the rescaled kernel

$$(3) \qquad K_\varepsilon(x,y) = \frac{C(x)}{\varepsilon^{N+2}}\, a\Big(x - E(x)(x-y)\Big)\, J\left(L^{-1}(x)\frac{x-y}{\varepsilon}\right)$$

Here a is given by $a(s) = \sum_i (s_i + M)$, with M large enough to ensure $a(x) \geq \beta > 0$. The matrix $L(x)$ is the Cholesky's factor of $A(x)$, that is, $A(x) = L(x)L^t(x)$, the matrix $E(x)$ is related with the coefficients $(a_{ij}(x))$ and $b_i(x)$ and $C(x)$ is a normalizing function, see Section 3 for a precise definition. Then, we prove that u^ε, solutions of rescaled nonlocal problems (P_{K_ε}), approximate uniformly the solution of the corresponding Dirichlet problem for the parabolic equation. We can now formulate our main result.

Theorem 1.1. *Let $v \in \mathcal{C}^{2+\alpha,1+\alpha/2}\left(\overline{\Omega}\times[0,T]\right)$ be the solution to (Q). Let, for a given $\varepsilon>0$, u^ε be the solution to (P_{K_ε}), with initial condition $u_0(x)$ and external datum $g(x,t)$. Then, we have*
$$\|v-u^\varepsilon\|_{L^\infty(\Omega\times[0,T])}\to 0, \qquad \textit{as } \varepsilon\to 0.$$

To deal with an equation in divergence form
$$v_t(x,t)=\sum_{i,j=1}^N \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial v(x,t)}{\partial x_j}\right),$$
we can just take
$$b_i(x):=\sum_{j=1}^N \frac{\partial a_{ij}(x)}{\partial x_j}$$
and the previous approach works. However, in this case the resulting family of nonlocal approximating problems have non-symmetric kernels. Note that for symmetric kernels, i.e., $K(x,y)=K(y,x)$, one has the desirable property of an "integration by parts formula", that is,
$$\iint K(x,y)(u(y)-u(x))\varphi(x)dydx=-\frac{1}{2}\iint K(x,y)(u(y)-u(x))(\varphi(y)-\varphi(x))dydx.$$
This is similar to the usual integration by parts formula for divergence form operators,
$$\int \operatorname{div}(A(x)\nabla v(x))\varphi(x)dx=-\int A(x)\nabla v(x)\nabla\varphi(x)dx.$$
To obtain a family of symmetric kernels $K_\varepsilon(x,y)=K_\varepsilon(y,x)$ such that the corresponding solutions to the nonlocal problems converge as $\varepsilon\to 0$ to the solution to the Dirichlet problem in divergence form we consider

$$K_\varepsilon(x,y)=\frac{2}{C(J)\varepsilon^{N+2}}G\left(B^{-1}(x)\frac{x-y}{\varepsilon}\right)G\left(B^{-1}(y)\frac{x-y}{\varepsilon}\right), \tag{4}$$

where $G^2(s)=J(s)$ (J is a radially symmetric, compactly supported and smooth kernel), and $B(x)=(b_{ij}(x))$ is a $N\times N$ matrix such that
$$\det(B(x))B(x)B^t(x)=A(x).$$
Note that $B(x)$ is invertible since $A(x)$ is. In this way we obtain a family of nonlocal symmetric kernels such that the approximation result given in Theorem 1.1 holds.

For constant matrices A and $b_i(x)=0$ in problem (Q), the rescaled kernels (3) and (4) coincide.

We finish the introduction with a brief description of previous results. When one considers a convolution kernel J (as before, radially symmetric, compactly supported and smooth) and rescale it, that is, for

$$K_\varepsilon(x,y)=\frac{C}{\varepsilon^{N+2}}J\left(\frac{y-x}{\varepsilon}\right) \tag{5}$$

one finds in the limit as $\varepsilon \to 0$ solutions to the classical heat equation, $v_t = \Delta v$. This fact was proved in [8] for Dirichlet boundary conditions and in [7] for Newmann boundary conditions. For an evolution problem with the same kernel but with an inhomogeneous term $a(y)$ in front in the whole $\mathbb{R}^N$ we refer to [16] (see also [5]). In this case the limit equation is given by $v_t = \Delta(a(x)v)$. For approximations of models from elasticity (peridynamics) we refer to [2]. Concerning nonlinear nonlocal problems (approximating for example the $p-$Laplacian or the porous medium equation) we refer to the book [1] and the survey [18]. We remark that in the previously mentioned references the case of matrix dependent problems (like the ones included in this paper) was not treated (only scalar coefficients appear).

The rest of this paper is organized as follows: in Section 2, we prove existence and uniqueness for solutions to problem (P_K) using a fixed point theorem (Proposition 2.1). In addition, we show a comparison principle (Proposition 2.2). In Section 3, using Cholesky's decomposition of the matrix $A(x) = (a_{ij}(x))$, we prove the uniform convergence of u^ε to v, the solution of the local parabolic equation (Theorem 1.1). In Section 4 we deal with the divergence form equation proving the convergence result for a symmetric family of kernels. Finally, the Appendix is devoted to give an alternative proof of existence of solutions (Appendix A), additionally, a technical computation using in the proof of Theorem 1.1 is postponed to the second part of the Appendix (Appendix B).

2. Existence, Uniqueness and Comparison Principle

By a solution of problem (P_K) we mean a function $u \in \mathcal{C}([0,\infty); L^1(\Omega))$ which satisfies

$$u(x,t) = \int_0^t \int_{\mathbb{R}^N} K(x,y)(u(y,s) - u(x,s))dyds + u_0(x), \qquad x \in \Omega,\ t \geq 0,$$

here we understand that $u(y,s) = g(y,s)$ when $y \in \mathbb{R}^N \setminus \Omega$, $s > 0$. Consequently, due to the previous integral expression, we notice that $u \in \mathcal{C}^1([0,\infty); L^1(\Omega))$.

Proposition 2.1. *If $u_0 \in L^1(\Omega)$, there exists a unique solution of problem (P_K).*

Proof. Fixed $t_0 > 0$, we set the Banach space $X_{t_0} = \mathcal{C}\left([0,t_0]; L^1(\Omega)\right)$ with norm

$$|||v||| = \max_{0 \leq t \leq t_0} \|v(\cdot,t)\|_{L^1(\Omega)}.$$

Let $\mathcal{T} : X_{t_0} \longrightarrow X_{t_0}$ be the operator defined by

$$\mathcal{T}(v)(x,t) = \int_0^t \int_{\mathbb{R}^N} K(x,y)(v(y,s) - v(x,s))dyds + u_0(x),$$

with $v(x,t) = g(x,t)$ if $x \notin \Omega$.

Note that in the definition of the operator $\mathcal{T}$ we include the fact that we are taking $v(y,s) = g(y,s)$ when $y \notin \Omega$.

In this way, using Fubini's theorem we obtain

$$\|\mathcal{T}(v)(\cdot,t)\|_{L^1(\Omega)} \leq \|u_0\|_{L^1(\Omega)} + \int_0^t \left(\int_\Omega \int_{\mathbb{R}^N} K(x,y)|v(y,s)|dydx + \int_\Omega \int_{\mathbb{R}^N} K(x,y)|v(x,s)|dydx \right) ds.$$

Recalling hypothesis (2), let us denote by $R = \|R(x)\|_\infty$. We get

$$\int_\Omega K(x,y)|v(y,s)|dy \leq R(x)\|v(\cdot,s)\|_{L^1(\Omega)} \leq R\|v(\cdot,s)\|_{L^1(\Omega)},$$

$$\begin{aligned}\int_{\mathbb{R}^N\setminus\Omega} K(x,y)|v(y,s)|dy &= \int_{\mathbb{R}^N\setminus\Omega} K(x,y)|g(y,s)|dy \\ &\leq R\|g(\cdot,s)\|_{L^1((\mathbb{R}^N\setminus\Omega)\cap B(0,d))}\end{aligned}$$

and

$$\int_{\mathbb{R}^N} K(x,y)|v(x,s)|dy \leq R|B(0,d)||v(x,s)|.$$

Hence

$$\begin{aligned}\|\mathcal{T}(v)(\cdot,t)\|_{L^1(\Omega)} &\leq \|u_0\|_{L^1(\Omega)} + \tilde{C}\int_0^t \|v(\cdot,s)\|_{L^1(\Omega)}ds \\ &\quad + \tilde{C}\int_0^t \|g(\cdot,s)\|_{L^1((\mathbb{R}^N\setminus\Omega)\cap B(0,d))},\end{aligned} \tag{6}$$

where $\tilde{C} = C(|\Omega|, |B(0,d)|)$. Since $\|v(\cdot,s)\|_{L^1(\Omega)} \leq |||v|||$ it follows that

$$\|\mathcal{T}(v)(\cdot,t)\|_{L^1(\Omega)} \leq \|u_0\|_{L^1(\Omega)} + t\tilde{C}|||v||| + \tilde{C}\int_0^t \|g(\cdot,s)\|_{L^1((\mathbb{R}^N\setminus\Omega)\cap B(0,d))},$$

thus operator $\mathcal{T}$ is well defined and

$$|||\mathcal{T}(v)||| \leq \|u_0\|_{L^1(\Omega)} + t_0\tilde{C}|||v||| + \tilde{C}\int_0^{t_0} \|g(\cdot,s)\|_{L^1((\mathbb{R}^N\setminus\Omega)\cap B(0,d))}.$$

Now, choosing $t_0 < \tilde{C}^{-1}$, and noticing that the term involving g cancels when computing $\mathcal{T}(w-z)$ for every $w, z \in X_{t_0}$ we get

$$|||\mathcal{T}(w-z)||| < |||w-z|||.$$

Hence, $\mathcal{T}$ is a contraction on X_{t_0} which maps X_{t_0} into itself, then by the Banach contraction principle there exists a unique $u \in X_{t_0}$ such that $\mathcal{T}(u) = u$, i.e., we get local existence and uniqueness of problem (P_K) for $0 \leq t \leq t_0$. Moreover, taking the Banach space $X_{2t_0} = \mathcal{C}\left([t_0, 2t_0]; L^1(\Omega)\right)$ with norm $|||v||| = \max_{t_0 \leq t \leq 2t_0} \|v(\cdot,t)\|_{L^1(\Omega)}$, $\mathcal{T} : X_{2t_0} \longrightarrow X_{2t_0}$ defined by

$$\mathcal{T}(v)(x,t) = \int_{t_0}^t \int_{\mathbb{R}^N} K(x,y)(v(y,s) - v(x,s))dyds + u(x,t_0),$$

and arguing as above, there exists a unique solution in $[t_0, 2t_0]$ and consequently in $[0, 2t_0]$. By an iteration argument, we obtain a unique solution $u \in \mathcal{C}([0,\infty); L^1(\Omega))$ of problem (P_K). □

For an alternative proof we refer the reader to Appendix A.

By a subsolution (respectively supersolution) of problem (P_K) we mean a function $u \in \mathcal{C}^1([0,T]; L^1(\Omega))$ which satisfies the following inequalities

$$\begin{cases} u_t(x,t) \overset{(\geq)}{\leq} \displaystyle\int_{\mathbb{R}^N} K(x,y)(u(y,s) - u(x,s))dy, & x \in \Omega,\ t > 0, \\ u(x,t) \overset{(\geq)}{\leq} g(x,t), & x \notin \Omega,\ t > 0. \\ u(x,0) \overset{(\geq)}{\leq} u_0(x), & x \in \Omega. \end{cases}$$

Clearly, a solution is both a subsolution and a supersolution.

Proposition 2.2. *Let $u, v \in \mathcal{C}^1(\overline{\Omega} \times [0,T])$ be a subsolution and supersolution respectively of problem (P_K). Then $u \leq v$.*

Proof. We will denote by $w = v - u$. Obviously $w \in \mathcal{C}^1(\overline{\Omega} \times [0,T])$ and it satisfies

$$\begin{cases} w_t(x,t) \geq \displaystyle\int_{\mathbb{R}^N} K(x,y)(w(y,t) - w(x,t))dy, & x \in \Omega,\ t > 0, \\ w(x,t) \geq 0, & x \notin \Omega,\ t > 0. \\ w(x,0) \geq 0, & x \in \Omega. \end{cases}$$

Now, we assume that $w(x,t)$ is not a nonnegative function, that is, there exists some point $(\tilde{x}, \tilde{t}) \in \Omega \times (0,T]$ such that $w(\tilde{x}, \tilde{t}) < 0$. Then, by the continuity of w, there exists $\varepsilon > 0$ such that $w(\tilde{x}, \tilde{t}) + \varepsilon \tilde{t}$ is also negative. Consider the function $w(x,t) + \varepsilon t \in \mathcal{C}(\overline{\Omega} \times [0,T])$, and let (x_0, t_0) be its minimum, thus

$$w_t(x_0, t_0) + \varepsilon \leq 0.$$

Conversely,

$$w_t(x_0, t_0) + \varepsilon > \int_{\mathbb{R}^N} K(x_0, y)(w(y, t_0) - w(x_0, t_0))dy \geq 0,$$

this leads to a contradiction and we conclude that $w(x,t)$ is a nonnegative function. □

3. Proof of Teorem 1.1

It is well known that given $A(x) = (a_{ij}(x))$ a symmetric and positive definite matrix there exists a unique lower triangular matrix $L(x) = (l_{ij}(x))$ with real and positive diagonal entries such that

$$A(x) = L(x)L^t(x), \tag{7}$$

where $L^t(x)$ denotes the transpose of $L(x)$ which is known as the Cholesky factor and (7) is known as the Cholesky factorization, see for instance [10].

Let $J : \mathbb{R}^n \to \mathbb{R}$ be a nonnegative, radially symmetric, continuous function with $\int_{\mathbb{R}^n} J(z)dz = 1$ and finite second order momentum. Assume also that J is strictly positive in $B(0,r)$ for some $r > 0$ and vanishes in $\mathbb{R}^n \setminus B(0,r)$.

Now we introduce some notations. Given a matrix $A(x) = (a_{ij}(x))$ with $\mathcal{C}^1(\bar{\Omega})$ coefficients we consider:

$$A_i(x) := \sum_{j=1}^{N} a_{ij}(x),$$

$$W(x) := \left(\begin{array}{cccc} b_1(x) & 0 & \dots & 0 \\ 0 & b_2(x) & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \dots & 0 & b_N(x) \end{array} \right)$$

We consider the rescaled kernel

$$K_\varepsilon(x,y) = \frac{C(x)}{\varepsilon^{N+2}}\, a\left(x - E(x)(x-y)\right)\, J\left(L^{-1}(x)\frac{x-y}{\varepsilon}\right). \tag{8}$$

Here a is defined as

$$a(s) = \sum_{i=1}^{N} (s_i + M),$$

for some constant $M > 0$ large enough to ensure $a(x) \geq \beta > 0$. The matrix $L(x)$ is given by (7) (note that we can take any $N \times N$ matrix $(l_{ij}(x))$, such that $A(x) = L(x)L^t(x))$, the function $C(x)$ is given by

$$C(x) = \frac{2}{C(J)a(x)(\det A(x))^{1/2}}$$

being $C(J) = \int J(z) z_1^2 dz$ and the matrix $E(x)$ by

$$E(x) = \frac{a(x)}{2} W(x) A^{-1}(x).$$

We remark that for this kernel, Proposition 2.1 and Proposition 2.2 can be used, since J is smooth, $a(x)$ is strictly positive and the coefficients of the envolved matrices are bounded. Therefore, for every $\varepsilon > 0$ we have existence, uniqueness and the comparison principle for the nonlocal problem.

Lemma 3.1. *Let u be a $\mathcal{C}^{2+\alpha,1+\alpha/2}\left(\mathbb{R}^N \times [0,T]\right)$ function and*

$$\mathcal{L}_\varepsilon(u) := \int_{\mathbb{R}^N} K_\varepsilon(x,y)(u(y,t) - u(x,t))dy.$$

Then

$$\left\| \mathcal{L}_\varepsilon(u) - \left(\sum_{i,j=1}^{N} a_{ij}(x) \frac{\partial^2 u(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{N} b_i(x) \frac{\partial u(x)}{\partial x_i} \right) \right\|_{L^\infty(\Omega \times [0,T])} \leq \theta(\varepsilon),$$

for some function $\theta(\varepsilon)$ that goes to zero as $\varepsilon \to 0$.

Proof. Under the change variables $y = x - \varepsilon L(x) z$, $\mathcal{L}_\varepsilon(u)$ becomes

$$\frac{C(x)\left(\det A(x)\right)^{1/2}}{\varepsilon^2} \int_{\mathbb{R}^N} a\left(x - \varepsilon D(x) z\right) J(z)(u(x - \varepsilon L(x)z, t) - u(x,t))dz$$

where $D(x) = \frac{a(x)}{2} W(x) \left(L^t(x)\right)^{-1}$. By a simple Taylor expansion we obtain

$$\begin{aligned}
\mathcal{L}_\varepsilon(u) = \quad & \frac{-C(x)\left(\det A(x)\right)^{1/2}}{\varepsilon} \sum_{i=1}^{N} \frac{\partial u}{\partial x_i} \sum_{j=1}^{N} l_{ij}(x) \int_{\mathbb{R}^N} a\left(x - \varepsilon D(x) z\right) J(z) z_j dz \\
& + \frac{1}{2} C(x)\left(\det A(x)\right)^{1/2} \sum_{i,j=1}^{N} \frac{\partial^2 u}{\partial x_i \partial x_j} \sum_{k,m=1}^{N} l_{ik}(x) l_{jm}(x) \\
& \qquad \times \int_{\mathbb{R}^n} a\left(x - \varepsilon D(x) z\right) J(z) z_k z_m dz + O(\varepsilon^\alpha) \\
= \quad & \mathcal{L}_\varepsilon^1(u) + \mathcal{L}_\varepsilon^2(u) + O(\varepsilon^\alpha).
\end{aligned}$$

For the first expression, $\mathcal{L}_\varepsilon^1(u)$, having in mind the definition of the function $a(s)$ and that J is a radial function, more specifically, we use that $\int J(z) z_j dz = 0$ and $\int J(z) z_m z_j dz = 0$ if $m \neq j$, we get

$$\begin{aligned}
\lim_{\varepsilon \to 0} \mathcal{L}_\varepsilon^1(u) &= C(x)\left(\det A(x)\right)^{1/2} \sum_{i=1}^{N} \frac{\partial u}{\partial x_i} \sum_{j=1}^{N} l_{ij}(x) \sum_{k,m=1}^{N} d_{km}(x) \int_{\mathbb{R}^N} J(z) z_m z_j dz \\
&= C(x)\left(\det A(x)\right)^{1/2} C(J) \sum_{i=1}^{N} \frac{\partial u}{\partial x_i} \sum_{j=1}^{N} l_{ij}(x) \sum_{k=1}^{N} d_{jk}^t(x),
\end{aligned}$$

here $d_{jk}^t(x)$ denotes the (j,k)-term of the matrix $D^t(x)$. Finally, since

$$\sum_{j=1}^{N} l_{ij}(x) \sum_{k=1}^{N} d_{jk}^t(x) = \sum_{k=1}^{N} (L(x) D^t(x))_{ik} = \frac{a(x)}{2} b_i(x),$$

it follows that

$$\lim_{\varepsilon \to 0} \mathcal{L}_\varepsilon^1(u) = \sum_{i=1}^{N} \frac{\partial u(x,t)}{\partial x_i} b_i(x).$$

On the other hand, letting $\varepsilon \to 0$ in $\mathcal{L}_\varepsilon^2(u)$ taking into account the choice of the matrix $L(x)$ we have

$$\lim_{\varepsilon \to 0} \mathcal{L}_\varepsilon^2(u) = \sum_{i,j=1}^{N} \frac{\partial^2 u(x,t)}{\partial x_i \partial x_j} \sum_{k=1}^{N} l_{ik}(x) l_{kj}^t(x) = \sum_{i,j=1}^{N} a_{ij}(x) \frac{\partial^2 u(x,t)}{\partial x_i \partial x_j},$$

which concludes the proof. □

Remark 3.2. We want to point out that the use of Cholesky's decomposition is not necessary for the proof. In fact, any matrix $L(x)$ satisfying (7) is also allowed. The reason to choose Cholesky's factor is to ensure the uniqueness of the rescaled kernel K_ε defined in (8).

In order to prove our main result, let $\tilde{v}$ be a $\mathcal{C}^{2+\alpha, 1+\alpha/2}\left(\mathbb{R}^N \times [0,T]\right)$ extension of v, the solution of the parabolic problem (Q). Therefore, $\tilde{v}$ verifies

$$\begin{cases}
\tilde{v}_t(x,t) = \Lambda(\tilde{v}(x,t)), & x \in \Omega,\ t \in (0,T], \\
\tilde{v}(x,t) = G(x,t), & x \notin \Omega,\ t \in (0,T], \\
\tilde{v}(x,0) = u_0(x), & x \in \Omega,
\end{cases}$$

where $G(x,t) = g(x,t)$ if $x \in \partial\Omega$ and

$$\Lambda(\tilde v(x,t)) = \sum_{i,j=1}^{N} a_{ij}(x) \frac{\partial^2 \tilde v(x,t)}{\partial x_i \partial x_j} + \sum_{i=1}^{N} b_i(x) \frac{\partial \tilde v(x,t)}{\partial x_i}.$$

Moreover, as G is smooth we get

$$G(x,t) = g(x,t) + O(\varepsilon), \ \text{ if } \operatorname{dist}(x, \partial\Omega) \le a\varepsilon, \tag{9}$$

where $a = r\sqrt{\lambda_{min}}$. Here λ_{min} denotes the $\max_{x\in\overline{\Omega}} \lambda_{min}(A(x)) > 0$. For more details we refer the reader to Appendix B.

Proof of Theorem 1.1. Set $w^\varepsilon := \tilde v - u^\varepsilon$ which satisfies

$$\begin{cases} w_t(x,t) = \Lambda(\tilde v) - \mathcal{L}_\varepsilon(\tilde v) + \mathcal{L}_\varepsilon(w^\varepsilon), & x \in \Omega, \ t \in (0,T], \\ w^\varepsilon(x,t) = G(x,t) - g(x,t), & x \notin \Omega, \ t \in (0,T], \\ w^\varepsilon(x,0) = 0, & x \in \Omega. \end{cases} \tag{10}$$

First, we claim that $\bar w(x,t) = K_1\theta(\varepsilon)t + K_2\varepsilon$ is a supersolution with $K_1, K_2 > 0$ sufficiently large but independent of ε. Indeed, taking into account Lemma 3.1 and that $\mathcal{L}_\varepsilon(\bar w) = 0$ we have

$$\bar w_t(x,t) = K_1\theta(\varepsilon) \ge \Lambda(\tilde v) - \mathcal{L}_\varepsilon(\tilde v) + \mathcal{L}_\varepsilon(\bar w).$$

Moreover, $\bar w(x,0) > 0$ and by (9) we obtain that $\bar w(x,t) \ge K_2\varepsilon \ge O(\varepsilon)$, for $t \in (0,T]$ and $x \notin \Omega$ such that $\operatorname{dist}(x,\partial\Omega) \le a\varepsilon$, which is our claim. From the comparison result we get

$$\tilde v - u^\varepsilon \le \bar w(x,t) = K_1\theta(\varepsilon)t + K_2\varepsilon.$$

Similar arguments applied to the case $\underline{w}(x,t) = -\bar w(x,t)$ leads us to assert that $\underline{w}(x,t)$ is a subsolution of problem (10). We conclude, using again the comparison principle stated in Proposition 2.2, that

$$-K_1\theta(\varepsilon)t - K_2\varepsilon \le \tilde v - u^\varepsilon \le K_1\theta(\varepsilon)t + K_2\varepsilon,$$

and hence

$$\|v - u^\varepsilon\|_{L^\infty(\Omega\times[0,T])} \le K_1 T\theta(\varepsilon) + K_2\varepsilon \to 0.$$

$\square$

Remark 3.3. It is worth pointing out that the particular case $A(x) = I$ and $b_i(x) = 0$, which corresponds to the heat equation, the rescaled kernel (5) considered by Cortázar et al. in [8] is the same K_ε considered here.

4. Divergence form operators

In this section, we consider the following rescaled kernel

$$K_\varepsilon(x,y) = \frac{2}{C(J)\varepsilon^{N+2}} G\left(B^{-1}(x)\frac{x-y}{\varepsilon}\right) G\left(B^{-1}(y)\frac{x-y}{\varepsilon}\right), \tag{11}$$

where $G^2(s) = J(s)$ and $B(x) = (b_{ij}(x))$ is a $N \times N$ matrix such that

$$\det(B(x))B(x)B^t(x) = A(x).$$

Note that the kernels given in (11) are symmetric, that is, they verify

$$K_\varepsilon(x,y) = K_\varepsilon(y,x).$$

For this family of symmetric kernels Proposition 2.1 and Proposition 2.2 can be used. Therefore, we have that the approximation result stated in Theorem 1.1 holds for the divergence form equation

$$v_t(x,t) = \sum_{i,j=1}^N \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial v(x,t)}{\partial x_j}\right).$$

This can be proved exactly as before as soon as one has the following result.

Lemma 4.1. *Let u be a $\mathcal{C}^{2+\alpha}\left(\mathbb{R}^N\right)$ function and*

$$\mathcal{L}_\varepsilon(u) := \int_{\mathbb{R}^N} K_\varepsilon(x,y)(u(y)-u(x))dy.$$

Then

$$\left\| \mathcal{L}_\varepsilon(u) - \sum_{i=1}^N \frac{\partial}{\partial x_i}\left(a_{ij}(x)\frac{\partial u(x)}{\partial x_j}\right)\right\|_{L^\infty(\Omega\times[0,T])} \leq \theta(\varepsilon),$$

for some function $\theta(\varepsilon)$ that goes to zero as $\varepsilon \to 0$.

Proof. In this proof we will use the following notations for partial derivatives and for the coefficients of the inverse and the adjoint of a matrix,

$$(f(s))_i' = \frac{\partial f(s)}{\partial s_i}, \quad B^{-1}(x) = (b_{ij}^{-1}(x)), \quad B^*(x) = (b_{ij}^*(x)).$$

Using the change of variable $z = \frac{x-y}{\varepsilon}$ and Taylor's expansions we get

$$\mathcal{L}_\varepsilon(u)(x) = F_{1,\varepsilon}(x) + F_{2,\varepsilon}(x) + O(\varepsilon^{2+\alpha})$$

with

$$F_{1,\varepsilon}(x) = \frac{-2}{C(J)\varepsilon}\sum_{i=1}^N \frac{\partial u(x)}{\partial x_i}\int_{\mathbb{R}^N} G(B^{-1}(x-\varepsilon z)z)G(B^{-1}(x)z)z_i\,dz$$

and

$$F_{2,\varepsilon}(x) = \frac{1}{C(J)}\sum_{i,j=1}^N \frac{\partial^2 u(x)}{\partial x_i \partial x_j}\int_{\mathbb{R}^N} G(B^{-1}(x-\varepsilon z)z)G(B^{-1}(x)z)z_i z_j\,dz.$$

Let us first analyze the limit as $\varepsilon \to 0$ of $F_{1,\varepsilon}(x)$. As $\int J(B^{-1}(x)z)z_i dz = 0$ (this follows changing z by $-z$), we can use L'Hopital's rule to obtain

$$\begin{aligned}\lim_{\varepsilon\to 0} F_{1,\varepsilon}(x) &= \frac{2}{C(J)}\sum_{i=1}^N \frac{\partial u(x)}{\partial x_i} \\ &\quad \times \int_{\mathbb{R}^N}\sum_{j=1}^N G_j'(B^{-1}(x)z)\sum_{k,m=1}^N (b_{jk}^{-1})_m'(x)z_k z_m G(B^{-1}(x)z)z_i\,dz.\end{aligned}$$

Now we observe that

$$G_j'(s)G(s) = \frac{1}{2}J_j'(s)$$

and hence

$$\lim_{\varepsilon\to 0} F_{1,\varepsilon}(x) = \frac{1}{C(J)} \sum_{i,j,k,m=1}^{N} \frac{\partial u(x)}{\partial x_i} (b_{jk}^{-1})'_m(x) \int_{\mathbb{R}^N} J'_j(B^{-1}(x)z) z_k z_m z_i \, dz.$$

Changing variables as $w = B^{-1}(x)z$ we have

$$\begin{aligned} \lim_{\varepsilon\to 0} F_{1,\varepsilon}(x) = \frac{det(B(x))}{C(J)} \sum_{i,j,k,m,p,q,r=1}^{N} & \frac{\partial u(x)}{\partial x_i} (b_{jk}^{-1})'_m(x) b_{ip}(x) b_{kq}(x) b_{mr}(x) \\ & \times \int_{\mathbb{R}^N} J'_j(w) w_p w_q w_r \, dw. \end{aligned}$$

To continue we have to find the value of the last integral. We have that

$$\int_{\mathbb{R}^N} J'_j(w) w_p w_q w_r \, dw = 0,$$

except for the following cases:

Case 1. $p = q = r = j$. In this case we have

$$\int_{\mathbb{R}^N} J'_j(w)(w_j)^3 \, dw = -3 \int_{\mathbb{R}^N} J(w)(w_j)^2 \, dw = -3C(J).$$

Case 2. ($p = j$ and $q = r \neq j$) or ($q = j$ and $p = r \neq j$) or ($r = j$ and $p = q \neq j$). In any of these cases one index is equal to j and the other two indexes are the same but different from j. Hence, in this case we get

$$\int_{\mathbb{R}^N} J'_j(w) w_j (w_q)^2 \, dw = - \int_{\mathbb{R}^N} J(w)(w_q)^2 \, dw = -C(J).$$

Collecting these cases we obtain

$$\lim_{\varepsilon\to 0} F_{1,\varepsilon}(x) = \sum_{i=1}^{N} \frac{\partial u(x)}{\partial x_i} H_i(x).$$

with

$$
\begin{aligned}
H_i(x) = -det(B(x)) \Bigg\{ & \sum_{j,k,m=1}^{N} 3(b_{jk}^{-1})'_m(x) b_{ij}(x) b_{kj}(x) b_{mj}(x) \\
& + \sum_{j,k,m \ p\neq j}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ij}(x) b_{kp}(x) b_{mp}(x) \Big] \\
& + \sum_{j,k,m,p\neq j}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ip}(x) b_{kj}(x) b_{mp}(x) \Big] \\
& + \sum_{j,k,m \ p\neq j}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ip}(x) b_{kp}(x) b_{mj}(x) \Big] \Bigg\} \\
= -det(B(x)) \Bigg\{ & \sum_{j,k,m,p=1}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ij}(x) b_{kp}(x) b_{mp}(x) \Big] \\
& + \sum_{j,k,m,p=1}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ip}(x) b_{kj}(x) b_{mp}(x) \Big] \\
& + \sum_{j,k,m,p=1}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ip}(x) b_{kp}(x) b_{mj}(x) \Big] \Bigg\} = A_1 + A_2 + A_3.
\end{aligned}
$$

Let us compute each one of the last three terms A_1, A_2 and A_3. First, using that

$$
\sum_{k=1}^{N} b_{ik}^{-1}(x) b_{kj}(x) = \begin{cases} 1 & i = j, \\ 0 & i \neq j, \end{cases}
$$

we obtain

$$
\sum_{k=1}^{N} (b_{ik}^{-1})'_m(x) b_{kj}(x) = - \sum_{k=1}^{N} b_{ik}^{-1}(x) (b_{kj})'_m(x). \tag{12}
$$

Using this property, we get

$$
\begin{aligned}
A_1 &= -det(B(x)) \sum_{j,k,m,p=1}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ij}(x) b_{kp}(x) b_{mp}(x) \Big] \\
&= det(B(x)) \sum_{j,k,m,p=1}^{N} \Big[b_{ij}(x) b_{jk}^{-1}(x) (b_{kp})'_m(x) b_{mp}(x) \Big] \\
&= det(B(x)) \sum_{m,p=1}^{N} \Big[(b_{kp})'_m(x) b_{mp}(x) \Big].
\end{aligned}
$$

Now, for A_2, using again (12) we have

$$
\begin{aligned}
A_2 &= -det(B(x)) \sum_{j,k,m,p=1}^{N} (b_{jk}^{-1})'_m(x) \Big[b_{ip}(x) b_{kj}(x) b_{mp}(x) \Big] \\
&= det(B(x)) \sum_{j,k,m,p=1}^{N} \Big[b_{jk}^{-1}(x) (b_{kj})'_m(x) b_{ip}(x) b_{mp}(x) \Big].
\end{aligned}
$$

As
$$b_{jk}^{-1}(x) = \frac{1}{det(B(x))}(b_{jk}^*(x))^t = \frac{1}{det(B(x))} b_{kj}^*(x)$$
we get
$$A_2 = \sum_{m,p=1}^N b_{ip}(x) b_{mp}(x) \sum_{k,j=1}^N b_{kj}^*(x)(b_{kj})'_m(x).$$
Now we use the formula for the derivative of the determinant (see [9] for a simple proof),
$$(det(B(x)))'_m = \sum_{k,j=1}^N b_{kj}^*(x)(b_{kj})'_m(x),$$
to obtain
$$A_2 = \sum_{m,p=1}^N b_{ip}(x) b_{mp}(x)(det(B(x)))'_m.$$

Finally, for A_3, using (12) one more time, we have
$$\begin{aligned} A_3 &= -det(B(x)) \sum_{j,k,m,p=1}^N (b_{jk}^{-1})'_m(x)\Big[b_{ip}(x) b_{kp}(x) b_{mj}(x)\Big] \\ &= det(B(x)) \sum_{j,k,m,p=1}^N \Big[b_{ip}(x)(b_{kp})'_m(x) b_{mj}(x) b_{jk}^{-1}(x)\Big] \\ &= det(B(x)) \sum_{m,p=1}^N \Big[(b_{mp})'_m(x) b_{ip}(x)\Big]. \end{aligned}$$

Hence, collecting these expressions for A_i we obtain
$$\begin{aligned} H_i(x) = \sum_{j=1}^N \Big[& det(B(x))(B'_j(x)B^t(x))_{ij} \\ & + (det(B(x)))'_j (B(x)B^t(x))_{ij} \\ & + det(B(x))(B(x)(B^t)'_j(x))_{ij}\Big] = \sum_{j=1}^N \frac{\partial a_{ij}(x)}{\partial x_j}. \end{aligned}$$

Therefore, we have obtained
$$\lim_{\varepsilon\to 0} F_{1,\varepsilon}(x) = \sum_{i=1}^N \frac{\partial u(x)}{\partial x_i} \sum_{j=1}^N \frac{\partial a_{ij}(x)}{\partial x_j}. \tag{13}$$

Next, we deal with the limit as $\varepsilon \to 0$ of $F_{2,\varepsilon}(x)$. It holds that
$$\lim_{\varepsilon\to 0} F_{2,\varepsilon}(x) = \frac{1}{C(J)} \sum_{i,j=1}^N \frac{\partial^2 u(x)}{\partial x_i \partial x_j} \int_{\mathbb{R}^N} G^2(B^{-1}(x)z) z_i z_j \, dz.$$
Changing variables as $w = B^{-1}(x)z$ we get
$$\lim_{\varepsilon\to 0} F_{2,\varepsilon}(x) = \frac{det(B(x))}{C(J)} \sum_{i,j=1}^N \frac{\partial^2 u(x)}{\partial x_i \partial x_j} \int_{\mathbb{R}^N} J(w) \sum_{k=1}^N b_{ik}(x) w_k \sum_{m=1}^N b_{jm}(x) w_m \, dw.$$

Now we only have to observe that

$$\int_{\mathbb{R}^N} J(w) w_k w_m \, dw = \begin{cases} C(J) & k = m, \\ 0 & k \neq m, \end{cases}$$

to obtain

$$\lim_{\varepsilon\to 0} F_{2,\varepsilon}(x) = \sum_{i,j=1}^{N} \frac{\partial^2 u(x)}{\partial x_i \partial x_j} det(B(x)) \sum_{k=1}^{N} b_{ik}(x) b_{jk}(x) = \sum_{i,j=1}^{N} \frac{\partial^2 u(x)}{\partial x_i \partial x_j} a_{ij}(x). \tag{14}$$

Finally, from (13) and (14) we conclude that

$$\begin{aligned} \lim_{\varepsilon\to 0} \mathcal{L}_\varepsilon(u)(x) &= \sum_{i,j=1}^{N} \frac{\partial^2 u(x)}{\partial x_i \partial x_j} a_{ij}(x) + \sum_{i=1}^{N} \frac{\partial u(x)}{\partial x_i} \sum_{j=1}^{N} \frac{\partial a_{ij}(x)}{\partial x_j} \\ &= \sum_{i,j=1}^{N} \frac{\partial}{\partial x_i} \left(a_{ij}(x) \frac{\partial u(x)}{\partial x_j} \right) \end{aligned}$$

as we wanted to show. □

5. Appendix

Appendix A. For any arbitrary $T > 0$ we claim that $\mathcal{T}$ is a contraction on $\mathcal{C}\left([0,T]; L^1(\Omega)\right)$ with norm

$$|||v||| = \max_{0\le t\le T} e^{-Mt} \|v(\cdot,t)\|_{L^1(\Omega)}$$

being M some constant greater than $\tilde{C} = C(|\Omega| + |B(0,d)|)$. Indeed, from (6)

$$\|\mathcal{T}(v)(\cdot,t)\|_{L^1(\Omega)} \le \|u_0\|_{L^1(\Omega)} + \frac{\tilde{C}}{M}\left(e^{Mt} - 1\right) |||v|||,$$

therefore

$$|||T(v)||| \le \max_{0\le t\le T} \left(e^{-Mt} \|u_0\|_{L^1(\Omega)} + \frac{\tilde{C}}{M}\left(1 - e^{-Mt}\right) |||v||| \right) \le \|u_0\|_{L^1(\Omega)} + \frac{\tilde{C}}{M} |||v|||,$$

and the claim is proved. The rest of the proof is similar in spirit to the proof of Proposition 2.1.

Appendix B. Given $B(x)$, matrix $n \times n$ defined for each $x \in \overline{\Omega}$, we wish to recall that the induced matrix norm to the euclidian matrix norm

$$\|B(x)\|_2 = \sup_{y\neq 0} \frac{\|B(x)y\|_2}{\|y\|_2}$$

is the spectral norm, i.e., $\|B(x)\|_2 = \sqrt{\lambda_{Max}(B^t(x)B(x))}$. Thus

$$\|L^{-1}(x)\|_2 = \sqrt{\lambda_{Max}(A^{-1}(x))} = (\lambda_{min}(A(x)))^{-1/2},$$

and hence $L^{-1}(x)\dfrac{x-y}{\varepsilon} \in B(0,r)$ if $y \in B(x, \frac{r\varepsilon}{\|L^{-1}(x)\|_2}) \subset B(x, a\varepsilon)$.

Ackonwledgement

This work was carried out during a stay by the first author at Universidad of Buenos Aires (Argentina) supported by Secretaría de Estado de Investigación, Desarrollo e Innovación EEBB2014 (Spain).

References

[1] F. Andreu-Vaillo, J. M. Mazón, J. D. Rossi and J. Toledo. Nonlocal Diffusion Problems. Mathematical Surveys and Monographs, vol. 165. AMS, 2010.

[2] F. Bobaru, M. Yang, L. Frota Alves, S. A. Silling, E. Askari and J. Xu, Convergence, adaptive refinement, and scaling in 1D peridynamics, *Int. J. Numer. Meth. Engng*, **77**, (2009) 852–877.

[3] M. Bodnar, J.J.L. Velazquez, An integro-differential equation arising as a limit of individual cell-bases models, *J. Differential Equations*, **222**, (2006) 341–380.

[4] C. Carrillo, P. Fife, Spatial effects in discrete generation population models, *J. Math. Biol.* **50**, (2005) 161–188.

[5] C. Cortázar, J. Coville, M. Elgueta, S. Martínez, A nonlocal inhomogeneous dispersal process, *J. Differential Equations* **241**, 332–358 (2007).

[6] C. Cortázar, M. Elgueta, J. D. Rossi and N. Wolanski, Boundary fluxes for nonlocal diffusion, *J. Differential Equations* **234**, 360–390 (2007).

[7] C. Cortazar, M. Elgueta, J. D. Rossi and N. Wolanski. How to approximate the heat equation with Neumann boundary conditions by nonlocal diffusion problems. *Arch. Rat. Mech. Anal.* **187(1)**, 137–156, (2008).

[8] C. Cortázar, M. Elgueta, J. D. Rossi, Nonlocal Diffusion problems that approximate the heat equation with Dirichlet boundary conditions, *Israel Journal of Mathematics* **170**, 53-60 (2009).

[9] M.A. Goldberg, The Derivative of a Determinant, *The American Mathematical Monthly* **79** No. 10, 1124-1126 (1972).

[10] A. Householder, The Theory of Matrices in Numerical Analysis, *Dover, New York, NY.* (1964).

[11] P. Fife, Some nonlocal trends in parabolic and parabolic-like evolutions, in Trends in Nonlinear Analysis, *Springer, Berlin* vol. **129**, 153–191 (2001).

[12] N. Fournier, P. Laurencot, Well-posedness of Smoluchowskis coagulation equation for a class of homogeneous kernels, *J. Funct. Anal.* **233**, (2006) 351–379.

[13] V. Hutson, S. Martínez, K. Mischaikow, G.T. Vickers, The evolutions of dispersal, *J. Math. Biol.* **47**, 483–517 (2003).

[14] G.M. Lieberman, Second Order Parabolic Differential Equations, *World Scientific Publishing Co. Pte. Ltd.* (1996).

[15] S. A. Silling and R. B. Lehoucq, Convergence of Peridynamics to Classical Elasticity Theory, *J. Elasticity* **93**, (2008), 13–37.

[16] J. W. Sun, W. T. Li, F. Y. Yang, Approximate the Fokker-Planck equation by a class of nonlocal dispersal problems, *Nonlinear Analysis*, **74**, (2011), 3501–3509.

[17] E. Valdinoci, From the long jump random walk to the fractional Laplacian, *Bol. Soc. Esp. Mat. Apl. SeMA*, **49**, (2009), 33–44.

[18] J. L. Vazquez, Recent progress in the theory of nonlinear diffusion with fractional Laplacian operators *Discr. Cont. Dyn. Sys. - Series S.* **7(4)**, (2014), 857–885.

J. D. Rossi: Departamento de Matemática, FCEyN, Universidad de Buenos Aires, Ciudad Universitaria, Pab 1 (1428), Buenos Aires, Argentina. jrossi@dm.uba.ar

A. Molino: Departamento de Análisis Matemático, Campus Fuentenueva S/N, Universidad de Granada 18071 - Granada, Spain. amolino@ugr.es