

BENEATH THE KINETIC INTERPRETATION OF NOISE

CARLOS ESCUDERO AND HELDER ROJAS

ABSTRACT. Diffusion theory establishes a fundamental connection between stochastic differential equations and partial differential equations. The solution of a partial differential equation known as the Fokker-Planck equation describes the probability density of the stochastic process that solves a corresponding stochastic differential equation. The kinetic interpretation of noise refers to a prospective notion of stochastic integration that would connect a stochastic differential equation with a Fokker-Planck equation consistent with the Fick law of diffusion, without introducing correction terms in the drift. This work is devoted to identifying the precise conditions under which such a correspondence can occur. One of these conditions is a structural constraint on the diffusion tensor, which severely restricts its possible form and thereby renders the kinetic interpretation of noise a non-generic situation. This point is illustrated through a series of examples. Furthermore, the analysis raises additional questions, including the possibility of defining a stochastic integral inspired by numerical algorithms, the behavior of stochastic transport equations in heterogeneous media, and the development of alternative models for anomalous diffusion. All these topics are addressed using stochastic analytical tools similar to those employed to study the main problem: the existence of the kinetic interpretation of noise.

1. INTRODUCTION

Diffusion is usually considered as a fundamental transport process in many sciences, including physics, chemistry, biology, and even finance. Since the seminal work of Adolf Fick [25], diffusive transport has frequently been understood macroscopically through the constitutive law

$$\mathbf{J} = -\mathbf{D}\nabla u,$$

which postulates that the flux $\mathbf{J}$ of a conserved scalar quantity u (such as mass density) is proportional to the gradient of that quantity, where $\mathbf{D}$ denotes the diffusion tensor. This law is normally combined with the continuity equation to yield the well-known family of *convection-diffusion equations*. Today, this family constitutes one of the most common classes of mathematical models for transport phenomena. Despite its mathematical simplicity, the Fick law embodies a deep physical idea: it gives a macroscopic representation of the irreversible relaxation towards equilibrium. Half a century later, Einstein showed how this law emerges from the underlying stochastic dynamics of Brownian motion [14]. That connection was further underpinned by subsequent work on the *fluctuation-dissipation relation* [6, 36, 51]. But quite remarkably, the pioneering work that connected microscopic fluctuations to macroscopic diffusion was that of Bachelier in the field of mathematical finance [2]. This just highlights the interdisciplinary importance of diffusive transport.

Microscopically, the trajectories of diffusive particles (such as point masses) are commonly represented by stochastic differential equations (SDEs):

$$(1) \quad dX_t = b(X_t) dt + \sigma(X_t) dW_t,$$

where the noise amplitude σ is related to the diffusion tensor via $\mathbf{D} = \sigma\sigma^\top$. For systems with *state-dependent diffusivity*, i.e. $\mathbf{D} = \mathbf{D}(x)$, that equation becomes highly nontrivial and needs to be understood within the theory of stochastic integration by Itô [32, 33]. Although this theory is enough to connect microscopic trajectories to the macroscopic appearance of diffusion [37, 40], this has not

Date: October 24, 2025.

Key words and phrases. Stochastic differential equations, stochastic integrals, diffusion processes, partial differential equations, interpretations of noise, Itô vs Stratonovich dilemma, Fick law.

2010 *MSC:* 35K10, 35Q84, 60H05, 60H10, 60J60, 82C31.

stopped a broad discussion in the physical literature about the convenience of substituting the Itô integral by some newer notion of stochastic integration, prominently that of Stratonovich [48]. For some recent discussions on the topic, we refer the reader to [4, 5, 9, 13, 16, 17, 18]. This *interpretation of noise dilemma*, consequently, becomes relevant when it comes to diffusive transport in nonuniform and anisotropic media, where the diffusion tensor varies spatially. This happens, for example, across diverse geophysical settings. These include sediment and bedload transport, solute migration in heterogeneous and variably saturated aquifers, vertical mixing and buoyancy effects in the ocean, and other geomorphological and hydrogeological systems [1, 29, 39, 43, 44, 47].

Not surprisingly, more interpretations of noise have been explored throughout the years. Indeed, the dependence of a stochastic integral, thought of in a Riemann-sum-like fashion, on the evaluation point of the integrand has been a subject of study for decades. Herein, we depart from reference [19], where we carried out a comparison among three different interpretations of noise —namely, the Itô, Stratonovich, and Hänggi–Klimontovich (HK) calculi— within a stochastic differential equation. These interpretations are the most natural from a symmetry viewpoint, as they correspond to evaluating an integrand, respectively, at the left endpoint, middle point, and right endpoint of each subinterval in the partition generated by the Riemann sum approximation. That work showed how each convention alters both the paths of the process and the evolution of their probability densities from a rigorous mathematical perspective, highlighting the mutual comparison between them. In the present work, we depart from there but focus on the connection of convection-diffusion equations with stochastic processes. Precisely, we consider equations of the form

$$(2) \quad \begin{cases} \partial_t u = -\nabla \cdot (b(x) u) + \frac{1}{2} \nabla \cdot (\mathbf{D}(x) \nabla u), & x \in \mathbb{R}^d, t > 0, \\ u(x, 0) = u_0(x), & x \in \mathbb{R}^d, \end{cases}$$

posed in an arbitrary dimension d , and where the diffusion tensor $\mathbf{D}(x)$ and the velocity field $b(x)$ are spatially dependent. As already mentioned, such equations appear frequently in physics, where $u = u(x, t)$ commonly represents the density of a diffusive material at point x and time t . They describe the spatio-temporal dispersal of particles subjected simultaneously to a convective field and a heterogeneous diffusion that is compatible with Fick law (modeled, respectively, by the first and second terms in the right-hand side of that equation). One of the main goals of this paper is to investigate the precise connection between this type of convection-diffusion equations and associated multidimensional SDEs of the form

$$dX_t = b(X_t) dt + \sigma(X_t) \bullet dW_t,$$

where σ satisfies $\mathbf{D} = \sigma \sigma^\top$ and $\bullet$ denotes the Hänggi–Klimontovich integral (as mentioned before, the variant of the Stratonovich integral in which the evaluation point of the integrand is taken to be the right endpoint of each subinterval in the partition). That relation —namely, the former is the Fokker-Planck equation of the latter— is classical in one dimension [19], but known to fail in arbitrary dimensions [31]. Herein, we will show that under the structural condition $\nabla \cdot \mathbf{D} = 2\sigma \nabla \cdot \sigma^\top$, the probability density of the process X_t does obey the partial differential equation for u , provided enough regularity is granted. However, this does not hold when such a condition is not met, which is the generic case in any dimension equal to or greater than two. Reference [31] raises the question of the existence of a discretization scheme for SDEs that directly yielded equation (2) as its Fokker-Planck equation. Of course, the solution to that equation corresponds to the probability density of a process that can be shown to obey a family of Itô SDEs, all of which present explicit correction terms in the drift. However, the existence of a notion of stochastic integration —the *kinetic interpretation of noise* from [31]— that is able to provide the desired connection with neither the correction terms in the drift nor the fulfillment of the structural condition seems to be not possible. Showing this, as well as complementary questions, is the goal of the present work.

The outline of this paper is as follows. In Section 2 we make precise the connection between convection-diffusion equations and Itô SDEs we will employ thereafter. In Section 3 we introduce the multidimensional version of the HK integral. In Section 4 we explore the relation between HK and Itô SDEs. In Section 5 we establish the connection between convection-diffusion equations

and HK SDEs in the absence of drift correction, which relies on a key structure of the divergence of the diffusion tensor. In that section, we also explore examples where the condition on $\nabla \cdot \mathbf{D}$ either holds or fails, thereby illustrating the impact of the properties of the diffusion tensor on the coupling between HK SDEs and convection-diffusion equations. In Section 6 we address a number of complementary questions: the ill-posedness of an alternative discretization scheme for the kinetic interpretation of noise, which was motivated by a numerical algorithm; a complete derivation of another discretization scheme motivated by a different numerical algorithm, to illustrate how such a program can be carried out; the application of that scheme on a physically motivated problem, which moreover illustrates the difficulties that arise when one of our main assumptions is removed, namely, uniform ellipticity; and, lastly, the assessment of another physical model that replaces degenerate ellipticity by a more tractable analytical structure. Finally, in Section 7, we draw our main conclusions.

Basic notations and setting. Let d and m be positive integers. We consider $\mathbb{R}^m$ endowed with the Euclidean norm $|\cdot|$, and let $\mathcal{M}_{d \times m}(\mathbb{R})$ be the set of all $d \times m$ matrices with real coefficients, endowed with the Frobenius (Hilbert-Schmidt) norm, defined as

$$\|B\|_F := \left(\sum_{i=1}^d \sum_{j=1}^m b_{ij}^2 \right)^{1/2}.$$

The spectral norm, defined as

$$\|B\|_S := \max_{\{x \in \mathbb{R}^m : |x|=1\}} |Bx|,$$

where $B = (b_{ij}) \in \mathcal{M}_{d \times m}(\mathbb{R})$, will also be considered for matrices. Let $\mathcal{O}(d)$ be the set of all $d \times d$ orthogonal matrices, i.e. $\mathcal{O}(d) := \{R \in \mathcal{M}_{d \times d}(\mathbb{R}) : R^\top R = \mathbb{I}_d\}$. Furthermore, for some positive real number T , let $\Psi : \mathbb{R}^m \times [0, T] \rightarrow \mathcal{M}_{d \times m}(\mathbb{R})$ be a mapping for which the m -dimensional vector Ψ_{i*} denotes the i 'th row of the $d \times m$ matrix $\Psi = (\psi_{ij})$. Analogously, for $1 \leq k \leq m$, the d -dimensional vector $(\partial_k \Psi)_{*l}$ denotes the l 'th column of the $d \times m$ matrix $\partial_k \Psi = (\partial_k \psi_{ij})$. If $\sigma = (\sigma_{ij}) \in \mathcal{M}_{d \times m}(\mathbb{R})$, we write $\nabla \sigma = (\partial_k \sigma_{ij})_{ijk}$ and

$$\nabla \cdot \sigma = \left(\sum_{j=1}^m \partial_j \sigma_{ij} \right)$$

to denote the gradient and the divergence of σ , respectively. If $A = (a_{ijk})$ is a third-order tensor as, for example, in the case of $\nabla \sigma$, we write

$$A : B = \left(\sum_{j=1}^d \sum_{k=1}^m a_{ijk} b_{jk} \right)$$

to denote the double vertical dot product. We use the abbreviation $\|\cdot\|_p := \|\cdot\|_{L^p(\mathbb{R}^d)}$, $1 \leq p \leq \infty$, and $C_c^\infty(\mathbb{R}^d)$ denotes the class of smooth functions with compact support. In addition, for $U \subset \mathbb{R}^d$ open and $k \in \mathbb{N}_0$, we write $C_b^k(U)$ for the class of scalar functions $f : U \rightarrow \mathbb{R}$ such that, for every multi-index $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ with $|\alpha| := \sum_{i=1}^d \alpha_i \leq k$, the partial derivative $D^\alpha f := \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} f$ exists, is continuous, and satisfies

$$\sup_{x \in U} |D^\alpha f(x)| < \infty,$$

where $D^0 f = f$, $D^{e_i} f = \partial_i f$ and e_i is a member of the canonical basis of $\mathbb{R}^d$. When $U = \mathbb{R}^d$, we write $C_b^k(\mathbb{R}^d)$. For matrix fields $A : U \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ we use the componentwise convention: $A \in C_b^k(U)$ iff each entry $A_{ij} \in C_b^k(U)$. From now on, we will work on a filtered probability space $(\Omega, \mathbb{F}, \mathcal{F}_t, \mathbb{P})$ completed with the $\mathbb{P}$ -null sets and where a m -dimensional Brownian motion W_t is defined. Moreover, we assume that the right-continuous filtration $\mathcal{F}_t \supseteq \mathcal{F}_t^W$, where $\mathcal{F}_t^W$ is the natural filtration of W_t .

2. CONVECTION-DIFFUSION EQUATIONS AND ITÔ CALCULUS

The connection between partial differential equations of convection-diffusion type and Itô stochastic differential equations has been established since long ago. In this section, we start specifying some known facts, and then move to introduce the precise statements we will need in the remainder of this work.

Equation (2) can be rewritten as

$$(3) \quad \begin{cases} \partial_t u(x, t) = \mathcal{L}u(x, t), & x \in \mathbb{R}^d, t > 0, \\ u(x, 0) = u_0(x), & x \in \mathbb{R}^d, \end{cases}$$

where the convection-diffusion operator $\mathcal{L}$ is

$$(4) \quad \mathcal{L}u(x) := - \sum_{i=1}^d \partial_i(b_i(x)u(x)) + \frac{1}{2} \sum_{i,j=1}^d \partial_i(\mathbf{D}_{ij}(x)\partial_j u(x)),$$

where $b = (b_1 \dots, b_d) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\mathbf{D} : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ are (for the time being) continuously differentiable functions. From now on, unless explicitly specified, we assume $\mathcal{L}$ to be a uniformly elliptic operator in divergence form, so the resulting equation is parabolic [20]. Moreover, under suitable assumptions on the parameters, this equation possesses a solution that is unique and classical [20]. Note that the uniform ellipticity condition is standard within mathematical analysis. It also holds in many physical cases, but it cannot be considered universal within that realm. Sometimes, a degenerate elliptic operator is needed, as illustrated in Subsection 6.3, where some of the difficulties that might arise in such a case are shown.

On the other hand, consider the Itô stochastic differential equation in $\mathbb{R}^d$:

$$(5) \quad dX_t = \left[b(X_t) + \frac{1}{2} \nabla \cdot \mathbf{D}(X_t) \right] dt + \sigma(X_t) dW_t, \quad X_0 = Z,$$

with W_t a m -dimensional Brownian motion and Z a $\mathcal{F}_0$ -measurable random variable in $\mathbb{R}^d$, where $\{\mathcal{F}_t\}_{t \geq 0}$ is the initially enlarged filtration $\mathcal{F}_t = \mathcal{F}_t^W \vee \sigma(Z)$; moreover, assume that the law of Z is absolutely continuous with respect to Lebesgue measure. For the time being, we take $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times m}(\mathbb{R})$ to be any function such that $\mathbf{D} = \sigma \sigma^\top$. From the diffusion theory for Itô SDEs it is known that, under adequate assumptions on the parameters and initial condition, the solution X_t of (5) is a Markov process with infinitesimal generator $\mathcal{L}^*$ defined by

$$(6) \quad \mathcal{L}^* \phi(x) := \sum_{i=1}^d \left[b(x) + \frac{1}{2} \nabla \cdot \mathbf{D}(x) \right]_i \partial_i \phi(x) + \frac{1}{2} \sum_{i,j=1}^d \mathbf{D}_{ij}(x) \partial_i \partial_j \phi(x), \quad \text{for all } \phi \in C_c^\infty(\mathbb{R}^d),$$

where $\mathcal{L}^*$ is the adjoint operator of $\mathcal{L}$ [37, 40, 45].

If we depart from equation (3), which we consider to be our datum, and we look for paths of a stochastic process that underlies it, in the sense that it comes as a strong solution to a precise SDE, we have already encountered our first trouble. Namely, that the correspondence of the partial differential equation (PDE) to the SDE is not unique. This happens in the simplest possible case: one-dimensional Brownian motion. In such a case, the heat equation (subjected to a unit Dirac mass initial condition):

$$\partial_t u(x, t) = \Delta u(x, t)$$

corresponds to two SDEs:

$$dX_t = +dW_t \quad \text{and} \quad dX_t = -dW_t,$$

from where we find two solutions $X_t = \pm W_t$ (up to an additive constant that equals the initial condition). Of course, both solutions are equal in law, but they differ pathwise. As herein we are mostly interested in strong solutions to SDEs, we need to specify which solution we select. In order to do so, first of all we need to take $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$, that is $m = d$. Although $m > d$ is perfectly possible, this only increases the number of solutions, and at the price of introducing more noises than physical dimensions, which we consider unnatural in the absence of a modeling

justification. Even in such a case, we have to extract the square root of a given $\mathbf{D}$, i.e. we look for a matrix σ such that $\mathbf{D} = \sigma\sigma^\top$. It is well known that there exist infinitely many solutions to this problem (except in the scalar case, which admits just two), but one can fix uniqueness by selecting the unique symmetric positive definite square matrix σ (more solutions are generated by multiplication of this by an orthogonal matrix). From now on we will assume this, the so-called *principal square root* of σ , for the sake of symmetry and due to the fact that, although different choices lead to solutions that differ pathwise, all share the same law. But note that different choices are very well possible, like for example the matrix given by the Cholesky decomposition of $\mathbf{D}$, which might be even more relevant for some purposes [27].

Remark 2.1. The existence of such a σ is guaranteed for any symmetric and positive definite matrix $\mathbf{D}$, which in turn is guaranteed by the assumption of uniform ellipticity. The symmetry of the diffusion tensor reflects the physical Onsager reciprocity and the mathematical fact that any asymmetric contribution can be absorbed by the drift term. Note, on the other hand, that if $\mathbf{D}$ is not symmetric, then there is no σ such that $\mathbf{D} = \sigma\sigma^\top$.

Remark 2.2. It is interesting to note that the existence of such a square root of $\mathbf{D}$ is guaranteed provided it is just positive semidefinite. This could correspond, however, to a degenerate elliptic operator and therefore will not be treated here, since that presents other complications (see Subsection 6.3). Nevertheless, in such a case it is also possible to fix uniqueness of the square root by selecting the only σ that is a symmetric positive semidefinite square matrix.

Now that we have made precise our selection of σ , we move on to establish the precise assumptions we will employ to build the relationship we seek between (3) and (5). Namely, we assume the following hypotheses:

- (H1) (Regularity of the parameters) $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\mathbf{D} : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ are C^4 and C^5 functions respectively, and there exists a constant $K > 0$ such that for all $x \in \mathbb{R}^d$, and $i, j, k, l = 1, \dots, d$,

$$|\partial_k b_i(x)| + |\partial_k \mathbf{D}_{ij}(x)| + |\partial_k \partial_l \mathbf{D}_{ij}(x)| \leq K.$$

- (H2) (Uniform ellipticity of the diffusion operator) There exists $\alpha > 0$ such that for all $x, \xi \in \mathbb{R}^d$,

$$\xi^\top \mathbf{D}(x) \xi \geq \alpha |\xi|^2.$$

- (H3) (Compatibility of the initial conditions) Z and $\{W_t\}_{t \geq 0}$ are independent, and $u_0 \in L^1(\mathbb{R}^d)$ is the probability density function (PDF) of Z , therefore $u_0 \geq 0$ and $\|u_0\|_1 = 1$.

Remark 2.3. Note that (H1) implies that the functions b and $\mathbf{D}$ are globally Lipschitz continuous, but they could be unbounded.

The following results establish a connection between PDE (3) and SDE (5) on a finite time interval $[0, T]$ with $T > 0$ arbitrary.

Lemma 2.4 (Standard Fokker-Planck form). *Under (H1), the initial value problems (3) and*

$$(7) \quad \begin{cases} \partial_t u(x, t) = \mathcal{L}_0 u(x, t), & x \in \mathbb{R}^d, t > 0, \\ u(x, 0) = u_0(x), & x \in \mathbb{R}^d, \end{cases}$$

where

$$(8) \quad \mathcal{L}_0 u = - \sum_{i=1}^d \partial_i (\tilde{b}_i u) + \frac{1}{2} \sum_{i,j=1}^d \partial_i \partial_j (D_{ij} u)$$

and $\tilde{b}_i := b_i + \frac{1}{2}(\nabla \cdot \mathbf{D})_i = b_i + \frac{1}{2} \sum_{j=1}^d \partial_j \mathbf{D}_{ij}$, coincide on classical solutions.

Proof. From (4), we have the following string of equalities

$$\begin{aligned}
\partial_t u &= -\nabla \cdot (bu) + \frac{1}{2} \nabla \cdot (\mathbf{D} \nabla u) \\
&= -\nabla \cdot (bu) + \frac{1}{2} \nabla \cdot (\nabla \cdot (\mathbf{D} u) - u \nabla \cdot \mathbf{D}) \\
&= -\nabla \cdot \left(\left[b + \frac{1}{2} \nabla \cdot \mathbf{D} \right] u \right) + \frac{1}{2} \nabla \cdot (\nabla \cdot (\mathbf{D} u)) \\
&= -\sum_{i=1}^d \partial_i \left(\left[b + \frac{1}{2} \nabla \cdot \mathbf{D} \right]_i u \right) + \frac{1}{2} \sum_{i,j=1}^d \partial_i \partial_j (\mathbf{D}_{ij} u),
\end{aligned}$$

provided u is a classical solution of either (3) or (7), i.e. $u \in C^1((0, T]; C(\mathbb{R}^d)) \cap C((0, T]; C^2(\mathbb{R}^d))$. $\square$

Lemma 2.5. *Let $\mathbf{D} : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ be a symmetric matrix that satisfies the following two conditions:*

(1) *(Uniform ellipticity) There exists $\alpha > 0$ such that for all $x, \xi \in \mathbb{R}^d$,*

$$\xi^\top \mathbf{D}(x) \xi \geq \alpha |\xi|^2.$$

(2) *(First-order regularity) There exists a constant M_k such that for all $k \in \{1, \dots, d\}$ it holds that*

$$\sup_{x \in \mathbb{R}^d} \max_{1 \leq i, j \leq d} |\partial_k \mathbf{D}_{ij}(x)| \leq M_k.$$

Let $\sigma : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ be the principal root of $\mathbf{D}$, i.e., the unique symmetric positive definite matrix satisfying $\mathbf{D} = \sigma \sigma^\top$. Then, for all $k \in \{1, \dots, d\}$ and $i, j \in \{1, \dots, d\}$, it holds that

$$\sup_{x \in \mathbb{R}^d} |\partial_k \sigma_{ij}(x)| \leq \frac{d^2}{2\sqrt{\alpha}} M_k.$$

Proof. We depart from the identity $\mathbf{D} = \sigma \sigma^\top$. Differentiating with respect to the k -th coordinate of x , we obtain

$$(9) \quad \sigma \sigma' + \sigma' \sigma = \mathbf{D}',$$

where $\sigma' = \partial_k \sigma(x)$ and $\mathbf{D}' = \partial_k \mathbf{D}(x)$. Since $\mathbf{D}$ is positive definite, then σ is also positive definite. Therefore, since the σ and $-\sigma$ matrices do not have common eigenvalues, using the spectral criterion for Sylvester equations, the linear Lyapunov operator

$$\mathcal{L}_\sigma : \mathcal{M}_{d \times d}(\mathbb{R}) \rightarrow \mathcal{M}_{d \times d}(\mathbb{R}), \quad \mathcal{L}_\sigma(X) = \sigma X + X \sigma,$$

is invertible, which ensures that the Sylvester equation (9) admits a unique solution $\sigma' = \mathcal{L}_\sigma^{-1}(\mathbf{D}')$, see [3, 30]. Subsequently, to obtain an explicit form of σ' , we need to obtain the spectral decompositions of $\mathbf{D}$ and σ . It is well known that for every $x \in \mathbb{R}^d$ there exists an orthogonal $P \in \mathcal{M}_{d \times d}(\mathbb{R})$ such that

$$\mathbf{D} = P \Lambda P^\top \quad \text{and} \quad \sigma = P \Lambda^{\frac{1}{2}} P^\top,$$

where $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_d)$ and $\Lambda^{\frac{1}{2}} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_d})$. Defining $H = (h_{ij}) := P^\top \mathbf{D}' P$ and $R = (r_{ij}) := P^\top \sigma' P$, and multiplying equation (9) from the left by $P^\top$ and from the right by P , yields

$$(10) \quad \Lambda^{\frac{1}{2}} R + R \Lambda^{\frac{1}{2}} = H.$$

In coordinates, equation (10) becomes

$$\sqrt{\lambda_i} r_{ij} + r_{ij} \sqrt{\lambda_j} = h_{ij},$$

where $1 \leq i, j \leq d$. Given that $\sqrt{\lambda_i} + \sqrt{\lambda_j} > 0$, we obtain

$$r_{ij} = \frac{h_{ij}}{\sqrt{\lambda_i} + \sqrt{\lambda_j}}.$$

Since by hypothesis **D** enjoys uniform ellipticity, that is, its smallest eigenvalue $\lambda_{\min} > \alpha \geq 0$, the precedent equality implies

$$(11) \quad |r_{ij}| \leq \frac{|h_{ij}|}{2\sqrt{\alpha}}.$$

Subsequently, from the fact that $\sigma' = PRP^\top$, we obtain the equality

$$\partial_k \sigma_{ij}(x) = e_i^\top PRP^\top e_j = u^\top Rv,$$

where e_i, e_j are members of the canonical basis of $\mathbb{R}^d$ and $u, v \in \mathbb{R}^d$ are orthonormal vectors. Therefore, for all $x \in \mathbb{R}^d$, we have that

$$(12) \quad |\partial_k \sigma_{ij}(x)| = |u^\top Rv| \leq \|R\|_S \leq \|R\|_F \leq d \left(\max_{a,b} |r_{ab}| \right),$$

where $1 \leq a, b \leq d$. Analogously, since by definition $H = P^\top \mathbf{D}' P$ for all $x \in \mathbb{R}^d$, we obtain

$$(13) \quad |h_{ij}(x)| \leq d \left(\max_{m,n} |\partial_k \mathbf{D}_{mn}(x)| \right),$$

where $1 \leq m, n \leq d$. Using the first-order regularity hypothesis on **D**, together with equations (11), (12), and (13), we obtain the final estimate

$$(14) \quad |\partial_k \sigma_{ij}(x)| \leq \frac{d^2}{2\sqrt{\alpha}} \left(\max_{m,n} |\partial_k \mathbf{D}_{mn}(x)| \right) \leq \frac{d^2}{2\sqrt{\alpha}} M_k.$$

Since it holds for all $x \in \mathbb{R}^d$ and all $i, j, k \in \{1, \dots, d\}$, this completes the proof. $\square$

Remark 2.6. Clearly, hypothesis (H1) implies the first-order regularity assumption in the statement of Lemma 2.5.

Lemma 2.7 (Semigroup generation). *If the domain of $\mathcal{L}_0$ is defined as $D(\mathcal{L}_0) := C_c^\infty(\mathbb{R}^d)$ and assumptions (H1) and (H2) are satisfied, then the closure of the convection-diffusion operator $\mathcal{L}_0$ generates a strongly continuous contraction semigroup $\mathcal{T}(t)$ on $L^1(\mathbb{R}^d)$.*

Proof. The operator $\mathcal{L}_0$ has the standard Fokker-Planck form

$$(15) \quad \mathcal{L}_0 f = - \sum_{i=1}^d \partial_i (\tilde{b}_i f) + \frac{1}{2} \sum_{i,j=1}^d \partial_i \partial_j (\mathbf{D}_{ij} f), \quad \text{for all } f \in C_c^\infty(\mathbb{R}^d).$$

From assumptions (H1) and (H2) along with Lemma 2.5 and Remark 2.6, and for some constant $K' > 0$, we have that

$$|\partial_k \tilde{b}_i(x)| + |\partial_k \sigma_{ij}(x)| \leq K',$$

for all $x \in \mathbb{R}^d$ and all $i, j, k = 1, \dots, d$. Therefore, applying Theorem 3.1 and Lemma 3.2 of [7] in (7), we can conclude that $(\mathcal{L}_0, D(\mathcal{L}_0))$ generates a strongly continuous contraction semigroup on $(L^1(\mathbb{R}^d), \|\cdot\|_1)$, which is denoted by $\mathcal{T}(t)$. $\square$

Definition 2.8. Let us denote by $C([0, T]; L^1(\mathbb{R}^d))$ the space of all functions $u(x, t)$ such that $\sup_{t \in [0, T]} \|u(\cdot, t)\|_1 < \infty$ and, for all $t \in [0, T]$, $\|u(\cdot, s) - u(\cdot, t)\|_1 \rightarrow 0$ as $s \rightarrow t$.

Definition 2.9. A function $u \in C([0, T]; L^1(\mathbb{R}^d))$ is called a mild solution of equation (7) if

$$u(\cdot, t) = \int_0^t \overline{\mathcal{L}_0} u(\cdot, s) ds + u_0 \quad \text{in } L^1(\mathbb{R}^d) \text{ for all } 0 \leq t \leq T,$$

where $\overline{\mathcal{L}_0}$ is the closure of $\mathcal{L}_0$ in $L^1(\mathbb{R}^d)$ and the integral is a Bochner integral in $L^1(\mathbb{R}^d)$, possibly improper at the origin.

Theorem 2.10 (Relation between PDE and SDE). *Under assumptions (H1) – (H3), the SDE (5) admits a unique strong solution X_t on $[0, T]$ for any $T > 0$. Furthermore, the convection-diffusion equation (3) has a unique solution $u \in C^1((0, T]; C(\mathbb{R}^d)) \cap C((0, T]; C^2(\mathbb{R}^d)) \cap C([0, T]; L^1(\mathbb{R}^d))$. In addition, this solution $u(\cdot, t)$ of equation (3) corresponds to the PDF of X_t .*

Proof. Assumptions (H1) – (H3), by Remark 2.3 in the present work, include the assumptions for the existence and uniqueness of a strong solution X_t of Itô SDE (5) as can be found in [45] (see Theorem 7 in Chapter V). On the other hand, from our Lemma 2.7, we know that the closure of $\mathcal{L}_0$ is the generator of a strongly continuous semigroup $\mathcal{T}(t)$ on $(L^1(\mathbb{R}^d), \|\cdot\|_1)$. Therefore, we define the mapping

$$u : t \longmapsto u(\cdot, t) := \mathcal{T}(t)u_0,$$

from where it follows that $u \in C([0, T]; L^1(\mathbb{R}^d))$ due to the strong continuity of $\mathcal{T}(t)$. Then, from Proposition II.6.4 in [15], it follows that u is the unique mild solution of (7). Moreover, by Remark 2.5 in [7], this solution is classical, thus $u \in C^1((0, T]; C(\mathbb{R}^d)) \cap C((0, T]; C^2(\mathbb{R}^d)) \cap C([0, T]; L^1(\mathbb{R}^d))$. Besides, by our Lemma 2.4, it is the unique solution to initial value problem (3). As a final step to conclude the proof, we use Theorem 2.4 in [7] to find that $u(\cdot, t)$ is the PDF of X_t . $\square$

3. MULTIDIMENSIONAL DEFINITION OF THE HÄNGGI-KLIMONTOVICH INTEGRAL

The Hänggi-Klimontovich integral is the name employed in part of the physical literature to denote a variant of the Stratonovich integral. The difference from the latter lies in the evaluation point of the integrand in each subinterval of the partition in the Riemann sum approximation: instead of using the midpoint, the right endpoint is selected. Although the mathematical structure of this integral was already examined in [19], this section is devoted to analyze those of its properties that will be needed in the remainder of this work.

Definition 3.1. An $\mathbb{R}^d$ -valued Markov process $\{Y_t, 0 \leq t \leq T\}$ is called a *diffusion process* if it satisfies the following three conditions for every $t \in [0, T]$ and $x \in \mathbb{R}^d$, and any $\delta > 0$:

- (1) Continuity of sample paths:

$$\lim_{h \rightarrow 0^+} \frac{1}{h} \mathbb{P} \left(|Y_{t+h} - Y_t| > \delta \mid Y_t = x \right) = 0.$$

- (2) Existence of drift vector $b : \mathbb{R}^d \times [0, T] \longrightarrow \mathbb{R}^d$ such that

$$\lim_{h \rightarrow 0^+} \frac{1}{h} \mathbb{E} \left(Y_{t+h} - Y_t \mid Y_t = x \right) = b(x, t).$$

- (3) Existence of diffusion matrix $\mathbf{D} : \mathbb{R}^d \times [0, T] \longrightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ such that

$$\lim_{h \rightarrow 0^+} \frac{1}{h} \mathbb{E} \left((Y_{t+h} - Y_t)(Y_{t+h} - Y_t)^\top \mid Y_t = x \right) = \mathbf{D}(x, t).$$

From now on, we will take $\mathbf{D}(x, t)$ to be uniformly elliptic (see the previous section) and $\sigma(x, t) : \mathbb{R}^d \times [0, T] \longrightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ to be the principal square root of $\mathbf{D}(x, t)$, and we will refer to this function as the *amplitude of noise*.

Definition 3.2. Let $b : \mathbb{R}^d \times [0, T] \longrightarrow \mathbb{R}^d$ and $\sigma(x, t) : \mathbb{R}^d \times [0, T] \longrightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ be two measurable functions.

- (Lipschitz condition) It is said that b and σ satisfy the Lipschitz condition, in the first argument, if there exists a constant $C > 0$ such that for all $x, y \in \mathbb{R}^d$ and $0 \leq t \leq T$,

$$|b(x, t) - b(y, t)| \leq C|x - y|, \quad \|\sigma(x, t) - \sigma(y, t)\|_F \leq C|x - y|.$$

- (Linear growth condition). It is said that b and σ satisfy the linear growth condition, in the first argument, if there exists a constant $C' > 0$ such that for all $x \in \mathbb{R}^d$ and $0 \leq t \leq T$,

$$|b(x, t)| \leq C'(1 + |x|), \quad \|\sigma(x, t)\|_F \leq C'(1 + |x|).$$

Definition 3.3. Assume that $b(x, t)$ and $\sigma(x, t) = (\sigma_{ij}(x, t))$ are continuous on $\mathbb{R}^d \times [0, T]$ and both satisfy the Lipschitz and linear growth conditions in x . Furthermore, consider a d -dimensional diffusion process $(Y_t)_{t \geq 0} = (Y_t^1, \dots, Y_t^d)_{t \geq 0}$, with drift vector $b(x, t)$ and noise amplitude $\sigma(x, t)$. For any $\Psi : \mathbb{R}^d \times [0, T] \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$, for which all its partial derivatives $\partial_i \Psi(x, t)$ are continuous, the Hänggi-Klimontovich integral of $(\Psi(Y_t, t))_{t \geq 0} = [\Psi_{ij}(Y_t, t)]_{t \geq 0}$ in the interval $[0, T]$ with respect to the diffusion process $(Y_t)_{t \geq 0}$ is denoted by

$$(16) \quad \int_0^T \Psi(Y_t, t) \bullet dY_t$$

and defined to be the $\mathbb{R}^d$ -valued random variable whose i -th component is given by

$$(17) \quad \int_0^T \Psi_{i*}(Y_t, t) \bullet Y_t := \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n \Psi_{i*}(Y_{t_j}, t_j)(Y_{t_j} - Y_{t_{j-1}}) \quad \text{in probability,}$$

where $\Delta_n = \{t_0, t_1, \dots, t_{n-1}, t_n\}$ is a partition of $[0, T]$, and $\|\Delta_n\| = \max_{1 \leq k \leq n} (t_k - t_{k-1})$.

This integral is well-defined and its relation with the Itô integral is expressed in the following theorem.

Theorem 3.4. Under the assumptions established in Definition 3.3, the limit in equation (16) exists. Furthermore, for $\mathbf{D} = \sigma \sigma^\top$, it is connected to the Itô integral through the following identity

$$(18) \quad \int_0^T \Psi(Y_t, t) \bullet dY_t = \int_0^T \Psi(Y_t, t) dY_t + \sum_{l,k=1}^d \int_0^T (\partial_k \Psi(Y_t, t))_{*l} \mathbf{D}_{lk}(Y_t, t) dt,$$

which holds almost surely.

The proof of Theorem 3.4 can be found in [19].

Proposition 3.5. Let X_t be a d -dimensional diffusion process with drift $b(x, t)$, noise amplitude $\sigma(x, t)$, and such that $X_0 = x_0 \in \mathbb{R}^d$; let also W_t be a d -dimensional Brownian motion. Assume that $b(x, t)$ and $\sigma(x, t)$ obey the same assumptions as in Definition 3.3 and, moreover, that all partial derivatives $\partial_i \sigma(x, t)$ are continuous. Then the identity

$$(19) \quad \int_0^T \sigma(X_t, t) \bullet dW_t = \int_0^T \sigma(X_t, t) dW_t + \sum_{k,l=1}^d \int_0^T (\partial_k \sigma(X_t, t))_{*l} \sigma_{kl}(X_t, t) dt$$

holds almost surely.

Proof. Choose $Y_t = (X_t, W_t)$ and $\Psi(x, t) = (\mathbf{0}, \sigma(x, t))$, where $\mathbf{0}$ is a null matrix with dimension $d \times d$, such that

$$(20) \quad \int_0^T \Psi(Y_t, t) \bullet dY_t = \int_0^T (\mathbf{0}, \sigma(X_t, t)) \bullet d \begin{pmatrix} X_t \\ W_t \end{pmatrix} = \int_0^T \sigma(X_t, t) \bullet dW_t.$$

Under the stated assumptions, the process X_t can be written in the form [49]:

$$(21) \quad X_t = x_0 + \int_0^t u(X_t, t) dt + \int_0^t \sigma(X_t, t) dW_t.$$

In this case, by equations (10) and (11) in [48], the diffusion coefficient of the process Y_t is

$$(22) \quad A(x, t) = [a_{ij}(x, t)] = \begin{bmatrix} \mathbf{D} & \sigma \\ \sigma^\top & \mathbf{I} \end{bmatrix},$$

where $\mathbf{I}$ is the identity matrix with dimension $d \times d$. Therefore, the transformation formula (18) of Theorem 3.4 applied to equation (20) yields

$$(23) \quad \int_0^T \sigma(X_t, t) \bullet dW_t = \int_0^T \sigma(X_t, t) dW_t + \sum_{l=1}^{2d} \sum_{k=1}^{2d} \int_0^T (\mathbf{0}, \partial_k \sigma(X_t, t))_{*l} a_{lk} dt.$$

Note that $\partial_k \sigma(x, t) = [\partial_k \sigma_{ij}]$ is a null matrix for $d+1 \leq k \leq 2d$. Furthermore, the vector $(\partial_k \Psi)_{*l} = (\mathbf{0}, \partial_k \sigma(X_t, t))_{*l}$ is a null vector for $1 \leq l \leq d$. Finally, replace these two observations in equation (23) to find that the identity (19) is verified. $\square$

4. HÄNGGI-KLIMONTOVICH CALCULUS

As one of our present goals is to connect the integral introduced in the previous section with PDEs, herein we perform an intermediate step: we connect it with the Itô integral in the context of stochastic differential equations. We start introducing the concept of solution.

Definition 4.1. We say that a d -dimensional diffusion process X_t that satisfies a stochastic integral equation of the form

$$(24) \quad X_t = Z + \int_0^t b(X_s, s) ds + \int_0^t \sigma(X_s, s) \bullet dW_s,$$

for $0 \leq t \leq T$, some positive real number T , some $\mathcal{F}_0$ -measurable random variable Z whose distribution is absolutely continuous with respect to Lebesgue measure, and some functions b and σ such that all the expressions are well defined, is a solution to the Hänggi-Klimontovich stochastic differential equation (HK SDE)

$$(25) \quad dX_t = b(X_t, t) dt + \sigma(X_t, t) \bullet dW_t, \quad X_0 = Z.$$

Remark 4.2. As in Section 2, we take $\{\mathcal{F}_t\}_{t \geq 0}$ to be the initially enlarged filtration $\mathcal{F}_t = \mathcal{F}_t^W \vee \sigma(Z)$ and Z to be independent of $\{W_t\}_{t \geq 0}$.

Theorem 4.3 (Conversion rule for SDEs). *Let $b : \mathbb{R}^d \times [0, T] \rightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times [0, T] \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ be functions as those specified in Proposition 3.5. Furthermore, assume that the function $h := \nabla \cdot \mathbf{D} - \sigma \nabla \cdot \sigma^\top$ satisfies the Lipschitz and linear growth conditions in x . Then, the unique solution to the Itô SDE*

$$(26) \quad dX_t = [b(X_t, t) + h(X_t, t)] dt + \sigma(X_t, t) dW_t, \quad X_0 = Z,$$

solves equation (25) almost surely. Correspondingly, there almost surely exists a unique solution to (25) that is given by the solution to (26).

Proof. Under the stated assumptions, the Itô SDE (26) has a unique solution [45]. Now, to prove the equivalence between SDEs (25) and (26), we write the identity (19) using the stochastic differential notation

$$(27) \quad \sigma(X_t, t) \bullet dW_t = \sigma(X_t, t) dW_t + \left[\sum_{k,l=1}^d (\partial_k \sigma(X_t, t))_{*l} \sigma_{kl}(X_t, t) \right] dt.$$

If we denote by h the factor inside the square brackets on the right hand side of (27), then its i -th component is given by

$$\begin{aligned} [h]_i &= \sum_{l=1}^d \sum_{k=1}^d (\partial_k \sigma_{il}) \sigma_{kl} \\ &= [\nabla \sigma : \sigma^\top]_i, \end{aligned}$$

from where we obtain the following equality

$$(28) \quad h(x, t) = \nabla \sigma : \sigma^\top.$$

On the other hand, we have

$$\begin{aligned} [\nabla \cdot \mathbf{D}]_i &= [\nabla \cdot (\sigma \sigma^\top)]_i \\ &= \sum_{k=1}^d \left(\partial_k \left[\sum_{l=1}^d \sigma_{il} \sigma_{kl} \right]_{ik} \right) \\ &= \sum_{l=1}^d \sum_{k=1}^d (\partial_k \sigma_{il}) \sigma_{kl} + \sum_{l=1}^d \sigma_{il} \left(\sum_{k=1}^d \partial_k \sigma_{kl} \right) \\ &= [\nabla \sigma : \sigma^\top]_i + [\sigma \nabla \cdot \sigma^\top]_i, \end{aligned}$$

so we have another equality

$$(29) \quad \nabla \cdot \mathbf{D} = \nabla \sigma : \sigma^\top + \sigma \nabla \cdot \sigma^\top.$$

Equation (26) is easily deduced from identities (27), (28), and (29). Therefore, we conclude that the unique solution of (26) solves (25) almost surely. To check uniqueness assume there exist two solutions; by the same formula this enters in contradiction with the uniqueness of solution to equation (26). $\square$

Corollary 4.4. *Under the hypotheses established in Theorem 4.3 and the additional assumption $\sigma \nabla \cdot \sigma^\top = \frac{1}{2} \nabla \cdot \mathbf{D}$, there exists a unique solution to the Itô SDE*

$$(30) \quad dX_t = \left[b(X_t, t) + \frac{1}{2} \nabla \cdot \mathbf{D}(X_t, t) \right] dt + \sigma(X_t, t) dW_t, \quad X_0 = Z$$

that moreover solves equation (25) almost surely. Correspondingly, there almost surely exists a unique solution to (25) that is given by the solution to (30).

5. CONVECTION-DIFFUSION EQUATIONS AND HK CALCULUS

Now we state one of our main results, which connects HK SDEs with convection-diffusion PDEs. However, that connection is not general and relies on a very particular structure of the diffusion coefficient, as the following statement shows.

Theorem 5.1. *Let X_t be a d -dimensional diffusion process described by the autonomous HK SDE*

$$(31) \quad dX_t = b(X_t) dt + \sigma(X_t) \bullet dW_t, \quad X_0 = Z,$$

where the drift b , the diffusion coefficient $\mathbf{D} = \sigma \sigma^\top$, and the initial condition Z obey the hypotheses (H1) – (H3). Furthermore, assume that σ is the principal square root of $\mathbf{D}$ and the identity $\nabla \cdot \mathbf{D} = 2 \sigma \nabla \cdot \sigma^\top$ holds. Then, the equation

$$(32) \quad \partial_t u = -\nabla \cdot (b(x) u) + \frac{1}{2} \nabla \cdot (\mathbf{D}(x) \nabla u),$$

with initial condition $u(x, 0) = u_0(x) \in L^1(\mathbb{R}^d)$ such that $Z \sim u_0(\cdot)$, possesses a unique mild solution $u(x, t)$, which is in turn classical and corresponds to the probability density function of X_t .

Proof. Hypotheses (H1) – (H3) imply, by Lemma 2.5, the regularity assumptions in Theorem 4.3. This, together with the structural condition $\nabla \cdot \mathbf{D} = 2 \sigma \nabla \cdot \sigma^\top$, implies the result of Corollary 4.4. The statement then follows as an immediate consequence of this Corollary along with Theorem 2.10. $\square$

Of course, the equality $\nabla \cdot \mathbf{D} = 2 \sigma \nabla \cdot \sigma^\top$ holds trivially in one dimension but does not hold generically in any higher dimension due to the non-commutative character of the matrix algebra. Therefore, in the following, we illustrate this structural condition with a series of examples in which

it either does or does not hold. We anticipate that such a condition is rather restrictive and then fails in general. We can only expect it to hold for highly structured matrices.

5.1. Positive examples. In this subsection we list a number of cases in which the matrix σ does satisfy the identity

$$(33) \quad \nabla \cdot \mathbf{D} = 2 \sigma \nabla \cdot \sigma^\top,$$

which is the structural condition assumed in Theorem 5.1.

Case 1: Constant diffusion tensor. Assume $\mathbf{D} : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ is constant and positive definite, $\mathbf{D}(x) = D_0$ for all $x \in \mathbb{R}^d$. Let $\sigma(x) = D_0^{1/2}$ denote its principal square root, so that $\mathbf{D} = \sigma \sigma^\top$. Then, for all indices i, j, k and all $x \in \mathbb{R}^d$, we obtain

$$\partial_k \mathbf{D}_{ij}(x) = 0, \quad \partial_k \sigma_{ij}(x) = 0,$$

which meets the boundedness requirements in (H1). Furthermore, hypothesis (H2) is satisfied since $\mathbf{D}$ is positive definite. Hence,

$$(\nabla \cdot \mathbf{D})_i(x) = \sum_{j=1}^d \partial_j \mathbf{D}_{ij}(x) = 0, \quad (\nabla \cdot \sigma^\top)_i(x) = \sum_{j=1}^d \partial_j \sigma_{ji}(x) = 0,$$

and therefore, for all $x \in \mathbb{R}^d$, the identity (33) is verified.

Case 2: Isotropic diffusion tensor. Let $\mathbf{D} : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ be an isotropic tensor such that $\mathbf{D}(x) = g(x) \mathbb{I}_d$, where $g \in C_b^5(\mathbb{R}^d)$ and $\inf_{x \in \mathbb{R}^d} g(x) \geq \alpha > 0$. Furthermore, let $\sigma(x) = \sqrt{g(x)} \mathbb{I}_d$ be its principal square root. Since g belongs to C_b^5 , the smoothness and derivative bounds required in (H1) hold. Since g is bounded away from zero, the uniform ellipticity (H2) follows from $\xi^\top \mathbf{D}(x) \xi = g(x) |\xi|^2 \geq \alpha |\xi|^2$ for every $\xi \in \mathbb{R}^d$. Given the diagonal structure of $\mathbf{D}(x)$, we obtain

$$(\nabla \cdot \mathbf{D})_i(x) = \sum_{j=1}^d \partial_j \mathbf{D}_{ij}(x) = \partial_i g(x), \quad (\nabla \cdot \sigma^\top)_i(x) = \sum_{j=1}^d \partial_j \sigma_{ji}(x) = \partial_i \sqrt{g(x)} = \frac{\partial_i g(x)}{2\sqrt{g(x)}},$$

which ensures that the structural condition (33) is fulfilled uniformly.

Case 3: Diagonal anisotropic diffusion tensor. Let $g_k \in C_b^5(\mathbb{R}^d)$ be such that $\inf_{x \in \mathbb{R}^d} g_k(x) \geq \alpha_k > 0$ for each $k = 1, \dots, d$. Define

$$\mathbf{D}(x) = \text{diag}(g_1(x), \dots, g_d(x)).$$

Let σ denote its principal square root; then,

$$\sigma(x) = \text{diag}(\sqrt{g_1(x)}, \dots, \sqrt{g_d(x)}).$$

Analogously to Case 2, the assumptions on g_k imply that all required derivatives of $\mathbf{D}$ up to order 5 are bounded, and uniform ellipticity holds with $\alpha := \min_{1 \leq k \leq d} \alpha_k > 0$. Similarly to the previous case, we obtain

$$(\nabla \cdot \mathbf{D})_i(x) = \sum_{j=1}^d \partial_j \mathbf{D}_{ij}(x) = \partial_i g_i(x), \quad (\nabla \cdot \sigma^\top)_i(x) = \sum_{j=1}^d \partial_j \sigma_{ji}(x) = \partial_i \sqrt{g_i(x)} = \frac{\partial_i g_i(x)}{2\sqrt{g_i(x)}}.$$

Consequently,

$$(2 \sigma (\nabla \cdot \sigma^\top))_i(x) = 2 \sqrt{g_i(x)} \frac{\partial_i g_i(x)}{2\sqrt{g_i(x)}} = \partial_i g_i(x) = (\nabla \cdot \mathbf{D})_i(x),$$

which ensures that the structural condition (33) is met.

Case 4: Orthogonally diagonalizable diffusion tensor. Let $\mathbf{D} : \mathbb{R}^d \rightarrow \mathcal{M}_{d \times d}(\mathbb{R})$ be positive definite such that $\mathbf{D} \in C_b^5(\mathbb{R}^d)$. Suppose that there exist $R \in \mathbb{O}(d)$ constant and functions $g_k \in C_b^5(\mathbb{R}^d)$ where $\inf_{x \in \mathbb{R}^d} g_k(x) \geq \alpha_k > 0$ such that

$$\mathbf{D}(x) = R \operatorname{diag}(g_1(x), \dots, g_d(x)) R^\top.$$

Let

$$\sigma(x) = R \operatorname{diag}(\sqrt{g_1(x)}, \dots, \sqrt{g_d(x)}) R^\top$$

be its principal square root, so that $\mathbf{D} = \sigma \sigma^\top$. Since R is constant and $g_k \in C_b^5(\mathbb{R}^d)$ with $g_k \geq \alpha_k > 0$, the smoothness requirements in (H1) hold and the uniform ellipticity in (H2) follows with $\alpha := \min_{1 \leq k \leq d} \alpha_k$. Let r_k denote the k -th column of R ; then we obtain

$$\nabla \cdot \mathbf{D} = \sum_{k=1}^d r_k ((\nabla g_k)^\top r_k), \quad \nabla \cdot \sigma^\top = \sum_{k=1}^d r_k ((\nabla \sqrt{g_k})^\top r_k),$$

and

$$2\sigma \nabla \cdot \sigma^\top = \sum_{k=1}^d 2\sqrt{g_k} r_k ((\nabla \sqrt{g_k})^\top r_k) = \sum_{k=1}^d r_k ((\nabla g_k)^\top r_k) = \nabla \cdot \mathbf{D},$$

which verifies identity (33). Note that this case reduces to Case 3 when $R = \mathbb{I}_d$ and that any matrix is orthogonally diagonalizable if and only if it is symmetric (by the spectral theorem).

Remark 5.2. The structural condition might hold even if we considered a non-symmetric noise amplitude. This can be shown by applying a constant orthogonal rotation in Case 4: consider a constant matrix $Q \in \mathbb{O}(d)$ and define

$$\sigma(x) = \mathbf{D}^{1/2}(x) Q,$$

which is generally non-symmetric and satisfies $\mathbf{D} = \sigma \sigma^\top$. In this case, using the symmetry of $\mathbf{D}^{1/2}$ and the constancy of Q , we obtain

$$\sigma (\nabla \cdot \sigma^\top) = \mathbf{D}^{1/2} Q \nabla \cdot (Q^\top \mathbf{D}^{1/2}) = \mathbf{D}^{1/2} \nabla \cdot \mathbf{D}^{1/2},$$

which, together with the result obtained in Case 4, verifies identity (33).

Remark 5.3. It is even possible to go beyond constant rotations upon the inclusion of suitable assumptions on the rotation matrix. So, in contrast to what was presented in Remark 5.2, if we set

$$\sigma(x) = \mathbf{D}^{1/2}(x) Q(x), \quad Q : \mathbb{R}^d \rightarrow \mathbb{O}(d), \quad Q \in C_b^5(\mathbb{R}^d);$$

then, the identity (33) holds if and only if

$$\mathbf{D}^{1/2} Q (\nabla \cdot Q^\top) \mathbf{D}^{1/2} = 0 \quad \text{for all } x \in \mathbb{R}^d.$$

Two immediate sufficient conditions are: Q constant, which reproduces the case presented in Remark 5.2, or $\nabla \cdot Q^\top = 0$ on $\mathbb{R}^d$, which means that the rows of Q are divergence-free, that is, the rows of Q are solenoidal vector fields. Outside these classes, the equality above typically fails, and identity (33) does not hold.

Case 5: Diffusion in the presence of a privileged orientation. Consider now the case in which diffusion takes place in an anisotropic medium in which a single privileged direction exists, such as a fiber-aligned diffusion. Let $v \in \mathbb{R}^d$ with $\|v\| = 1$ be the privileged orientation and $f, g \in C_b^5(\mathbb{R}^d)$ with $\inf_x f(x) \geq \alpha_1 > 0$ and $\inf_x g(x) \geq \alpha_2 > 0$. Let us consider

$$\mathbf{D}(x) = f(x) v v^\top + g(x) (\mathbb{I}_d - v v^\top) = g(x) \mathbb{I}_d + (f(x) - g(x)) v v^\top.$$

Since v is constant and $f, g \in C_b^5$, all spatial derivatives of $\mathbf{D}$ up to order five are continuous and bounded, hence (H1) holds. For uniform ellipticity, we use the existence and uniqueness of the orthogonal decomposition relative to $\operatorname{span}\{v\}$: for every $\xi \in \mathbb{R}^d$ there exist unique vectors ξ_1, ξ_2 with

$$\xi = \xi_1 + \xi_2, \quad \xi_1 = (v^\top \xi) v, \quad \xi_2 \perp v, \quad |\xi|^2 = |\xi_1|^2 + |\xi_2|^2.$$

Since $vv^\top$ and $I_d - vv^\top$ are complementary orthogonal projectors (as $\|v\| = 1$), we obtain

$$\xi^\top \mathbf{D}(x) \xi = f(x) |\xi_1|^2 + g(x) |\xi_2|^2 \geq \alpha_1 |\xi_1|^2 + \alpha_2 |\xi_2|^2 \geq \alpha |\xi|^2, \quad \alpha := \min\{\alpha_1, \alpha_2\} > 0,$$

so $\mathbf{D}(x)$ is positive definite for every x and (H2) holds with ellipticity constant α . Subsequently, we diagonalize $\mathbf{D}$. Since

$$\mathbf{D}(x) = g(x) \mathbb{I}_d + (f(x) - g(x)) vv^\top,$$

we have $\mathbf{D}(x)v = f(x)v$, and for any $w \perp v$, $\mathbf{D}(x)w = g(x)w$. Thus, the spectrum is

$$\{f(x) \text{ (with multiplicity 1)}, g(x) \text{ (with multiplicity } d-1)\}.$$

Choose an orthonormal basis $\{v, e_2, \dots, e_d\}$ with $e_j \perp v$, $j = 2, \dots, d$, and let $R \in \mathbb{O}(d)$ be the constant orthogonal matrix whose columns are $v, e_2, \dots, e_d$. Then

$$\mathbf{D}(x) = R \operatorname{diag}(f(x), g(x), \dots, g(x)) R^\top.$$

The principal square root is

$$\sigma(x) = R \operatorname{diag}(\sqrt{f(x)}, \sqrt{g(x)}, \dots, \sqrt{g(x)}) R^\top, \quad \mathbf{D} = \sigma \sigma^\top.$$

Proceeding analogously to Case 4, since R is constant and $\Lambda(x) := \operatorname{diag}(\sqrt{f(x)}, \sqrt{g(x)}, \dots, \sqrt{g(x)})$ is diagonal, we have componentwise $\nabla \cdot \Lambda^2 = 2 \Lambda (\nabla \cdot \Lambda^\top)$. Therefore,

$$\nabla \cdot \mathbf{D} = 2 \sigma \nabla \cdot \sigma^\top,$$

and the structural identity (33) holds pointwise.

Case 6: Scalar-modulated constant anisotropic diffusion. Let $B \in \mathcal{M}_{d \times d}(\mathbb{R})$ be constant, symmetric, and positive definite. Let $g \in C_b^5(\mathbb{R}^d)$ with $g(x) \geq \alpha > 0$ for every $x \in \mathbb{R}^d$. Consider

$$\mathbf{D}(x) = g(x) B.$$

Since B is constant and g belongs to C_b^5 , all spatial derivatives of $\mathbf{D}$ up to order five are continuous and bounded, which meets the requirement in (H1). Since B is positive definite, there exists a constant $c_B > 0$ such that

$$\xi^\top B \xi \geq c_B |\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d.$$

Additionally, since $g(x) \geq \alpha$, we obtain

$$\xi^\top \mathbf{D}(x) \xi \geq \alpha c_B |\xi|^2 \quad \text{for every } x \in \mathbb{R}^d.$$

Then $\mathbf{D}(x)$ is positive definite for all x and the uniform ellipticity assumption (H2) holds with constant αc_B . Take the principal square root $\sigma(x) = \sqrt{g(x)} B^{1/2}$, so that $\mathbf{D} = \sigma \sigma^\top$. Since B is constant,

$$\nabla \cdot \mathbf{D}(x) = B \nabla g(x), \quad (\nabla \cdot \sigma^\top)(x) = B^{1/2} \nabla \sqrt{g(x)}.$$

Hence,

$$2 \sigma(x) (\nabla \cdot \sigma^\top)(x) = 2 \sqrt{g(x)} B^{1/2} B^{1/2} \nabla \sqrt{g(x)} = B \nabla g(x) = \nabla \cdot \mathbf{D}(x),$$

and the structural identity (33) holds pointwise. One might also use a non-symmetric matrix $\tilde{\sigma}(x) := \sigma(x) Q$ with $Q \in \mathbb{O}(d)$ constant, and the identity remains valid.

Remark 5.4. As a specific example, we consider radial-modulated diffusion. Radial modulation is ubiquitous in rotationally symmetric models, for instance around point sources, and its assessment follows immediately from Case 6. To see this, let B be constant, symmetric, and positive definite as in that case. Let $g(x) = h(r)$ and $r = |x|$, where $h \in C_b^5([0, \infty))$, $h(r) \geq \alpha > 0$, and

$$h'(0) = h^{(3)}(0) = h^{(5)}(0) = 0.$$

On $\mathbb{R}^d \setminus \{0\}$,

$$\nabla g(x) = \frac{h'(r)}{r} x, \quad \nabla \cdot \mathbf{D}(x) = B \nabla g(x) = \frac{h'(r)}{r} B x.$$

The null-odd-derivative conditions ensure that g is C^5 at the origin, in particular, $\nabla g(0) = 0$, and that all derivatives of g up to order five extend continuously to $\mathbb{R}^d$. For $\sigma(x) = \sqrt{g(x)} B^{1/2}$, the

chain rule implies that the structural identity (33) holds for $x \neq 0$, and by continuity it also holds at $x = 0$. This case would arise whenever diffusivity depended only on the distance to the origin, for example, in a material whose properties depended on temperature and in the presence of a localized heat source. Note also that this remark covers anisotropic situations, given the generality of B .

5.2. Negative examples. In the following, we summarize a couple of cases in which the identity $\nabla \cdot \mathbf{D} = 2\sigma \nabla \cdot \sigma^\top$ does not hold. As already mentioned, this is the generic situation, even in dimension two.

Case 1: Bidimensional cross-coupled diffusion tensor. Let $\alpha, \beta > 0$ be constants and let $\tau \in C_b^5(\mathbb{R}^2)$ satisfy $\sup_x |\tau(x)| \leq \tau_{\max} < \sqrt{\alpha\beta}$. Consider

$$\mathbf{D}(x) = \begin{pmatrix} \alpha^2 + \tau(x)^2 & (\alpha + \beta)\tau(x) \\ (\alpha + \beta)\tau(x) & \beta^2 + \tau(x)^2 \end{pmatrix},$$

for $x \in \mathbb{R}^2$. Since $\tau \in C_b^5$, all spatial derivatives of $\mathbf{D}$ up to order five are continuous and bounded, hence (H1) holds. Also, let

$$\sigma(x) = \begin{pmatrix} \alpha & \tau(x) \\ \tau(x) & \beta \end{pmatrix}$$

denote its principal square root, $\mathbf{D} = \sigma\sigma^\top$. By the Rayleigh-Ritz theorem, for every $\xi \in \mathbb{R}^2$ we have

$$\xi^\top \mathbf{D}(x) \xi = |\sigma\xi|^2 \geq \lambda_{\min}(\sigma)^2 |\xi|^2,$$

where $\lambda_{\min}(\sigma)$ denotes the smallest eigenvalue of σ . Since σ is a symmetric 2×2 matrix, it follows that $\lambda_{\min}(\sigma) \geq \det \sigma / \text{tr } \sigma$. Here $\det \sigma = \alpha\beta - \tau(x)^2 \geq \alpha\beta - \tau_{\max}^2 =: \delta_1 > 0$ and $\text{tr } \sigma = \alpha + \beta$; hence

$$\xi^\top \mathbf{D}(x) \xi \geq \left(\frac{\delta_1}{\alpha + \beta} \right)^2 |\xi|^2.$$

Therefore, $\mathbf{D}$ is positive definite for every x and the uniform ellipticity hypothesis (H2) holds with constant $[\delta_1/(\alpha + \beta)]^2$. Subsequently, since in $\mathbf{D}$ only τ depends on x , we obtain

$$\nabla \cdot \mathbf{D} = \begin{pmatrix} 2\tau \partial_1 \tau + (\alpha + \beta) \partial_2 \tau \\ (\alpha + \beta) \partial_1 \tau + 2\tau \partial_2 \tau \end{pmatrix}, \quad \nabla \cdot \sigma^\top = \begin{pmatrix} \partial_2 \tau \\ \partial_1 \tau \end{pmatrix}.$$

Hence,

$$\Lambda(x) := \nabla \cdot \mathbf{D} - 2\sigma(\nabla \cdot \sigma^\top) = \begin{pmatrix} (\beta - \alpha) \partial_2 \tau \\ (\alpha - \beta) \partial_1 \tau \end{pmatrix}.$$

Thus, if $\alpha \neq \beta$ and $\nabla \tau$ is not identically zero, there exists x^* with $\Lambda(x^*) \neq 0$; therefore, the structural identity (33) is not fulfilled, despite $\mathbf{D}$ satisfies (H1)–(H2) and $\sigma = \mathbf{D}^{1/2}$ has been chosen as its principal square root.

Remark 5.5. The function $\tau(x)$ controls the off-diagonal part of $\mathbf{D}$ and σ . Therefore, intuitively, it measures how much the diffusive fluxes correlate between the x_1 and x_2 directions. If τ is coordinate dependent, that means the correlation also varies in space. Some admissible forms of τ that can be expressed in terms of elementary functions are:

- Periodic: $\tau(x) = \varepsilon \sin(k^\top x)$ with $k \in \mathbb{R}^2 \setminus \{0\}$, $\varepsilon \in \mathbb{R}$;
- Gaussian: $\tau(x) = \varepsilon \exp(-|x - x_0|^2/\ell^2)$ with $x_0 \in \mathbb{R}^2$, $\ell > 0$, $\varepsilon \in \mathbb{R}$;
- Front-like: $\tau(x) = \varepsilon \tanh((x_1 - c)/\ell)$ with $c \in \mathbb{R}$, $\ell > 0$, $\varepsilon \in \mathbb{R}$, and $x = (x_1, x_2)$.

All these examples give rise to regions where $\nabla \tau(x)$ does not cancel out. There, the function $\Lambda(x)$ does not vanish and the identity (33) fails as shown in Case 1.

Case 2: Separable cross-coupled diffusion tensor. Let $\varepsilon \neq 0$ be a constant, and let $a, b \in C_b^5(\mathbb{R})$. Assume

$$a(x_1) \geq a_0 > 0, \quad b(x_2) \geq b_0 > 0, \quad a(x_1)b(x_2) - \varepsilon^2 \geq \delta_2 > 0 \quad \text{for all } x = (x_1, x_2) \in \mathbb{R}^2.$$

Let us consider

$$\mathbf{D}(x) = \begin{pmatrix} a(x_1)^2 + \varepsilon^2 & \varepsilon(a(x_1) + b(x_2)) \\ \varepsilon(a(x_1) + b(x_2)) & b(x_2)^2 + \varepsilon^2 \end{pmatrix}.$$

Since $a, b \in C_b^5$, all spatial derivatives of $\mathbf{D}$ up to order five are continuous and bounded, hence (H1) holds. Moreover,

$$\det \mathbf{D}(x) = (a(x_1)b(x_2) - \varepsilon^2)^2 \geq \delta_2^2 > 0, \quad \mathbf{D}_{11}(x) = a(x_1)^2 + \varepsilon^2 > 0.$$

Therefore, by the Sylvester criterion, $\mathbf{D}(x)$ is positive definite for every x , and (H2) holds. Also, let us denote by

$$\sigma(x) = \begin{pmatrix} a(x_1) & \varepsilon \\ \varepsilon & b(x_2) \end{pmatrix}$$

its principal square root; then $\mathbf{D} = \sigma\sigma^\top$ and $\det \sigma(x) = a(x_1)b(x_2) - \varepsilon^2 \geq \delta_2 > 0$. Subsequently, by separability (that is, in $\mathbf{D}$ only a depends on x_1 and only b on x_2) we obtain

$$\nabla \cdot \mathbf{D} = \begin{pmatrix} 2a \partial_1 a + \varepsilon \partial_2 b \\ \varepsilon \partial_1 a + 2b \partial_2 b \end{pmatrix}, \quad \nabla \cdot \sigma^\top = \begin{pmatrix} \partial_1 a \\ \partial_2 b \end{pmatrix}.$$

Hence,

$$\Lambda(x) := \nabla \cdot \mathbf{D} - 2\sigma(\nabla \cdot \sigma^\top) = -\varepsilon \begin{pmatrix} \partial_2 b(x_2) \\ \partial_1 a(x_1) \end{pmatrix}.$$

Thus, if $\varepsilon \neq 0$ and either $\partial_1 a$ or $\partial_2 b$ is not identically zero, there exists x^* with $\Lambda(x^*) \neq 0$; therefore, the structural identity (33) is not fulfilled, despite $\mathbf{D}$ again satisfies (H1)–(H2) and $\sigma = \mathbf{D}^{1/2}$ is its principal square root.

Remark 5.6. The functions $a(x_1)$ and $b(x_2)$ regulate this cross-coupled diffusion in the directions x_1 and x_2 respectively. Assume $\varepsilon \neq 0$; then the structural identity (33) fails precisely in those regions in which either $\partial_1 a(x_1)$ or $\partial_2 b(x_2)$ does not vanish. Some admissible choices for a and b that can be written in terms of elementary functions are:

- Periodic: $a(x_1) = a_0 + \eta_a \sin(k_a x_1 + \theta_a)$, $b(x_2) = b_0 + \eta_b \cos(k_b x_2 + \theta_b)$, with $k_a, k_b > 0$, $\theta_a, \theta_b \in \mathbb{R}$, and amplitudes chosen so that $|\eta_a| < a_0$, $|\eta_b| < b_0$, and $(a_0 - |\eta_a|)(b_0 - |\eta_b|) - \varepsilon^2 \geq \delta_2$.
- Gaussian: $a(x_1) = a_0 + \eta_a e^{-(x_1 - x_1^0)^2/\ell_1^2}$, $b(x_2) = b_0 + \eta_b e^{-(x_2 - x_2^0)^2/\ell_2^2}$, with $x_1^0, x_2^0 \in \mathbb{R}$, $\ell_1, \ell_2 > 0$, and amplitudes satisfying $|\eta_a| < a_0$, $|\eta_b| < b_0$, and $(a_0 - |\eta_a|)(b_0 - |\eta_b|) - \varepsilon^2 \geq \delta_2$.
- Front-like: $a(x_1) = a_0 + \eta_a \tanh((x_1 - c_1)/\ell_1)$, $b(x_2) = b_0 + \eta_b \tanh((x_2 - c_2)/\ell_2)$, with $c_1, c_2 \in \mathbb{R}$, $\ell_1, \ell_2 > 0$, and amplitudes chosen so that $|\eta_a| < a_0$, $|\eta_b| < b_0$, and $(a_0 - |\eta_a|)(b_0 - |\eta_b|) - \varepsilon^2 \geq \delta_2$.

6. BEYOND THE KINETIC INTERPRETATION OF NOISE

Our developments so far have left a number of side questions open. The goal of this section is to address them. We start with an attempt of definition of stochastic integral which was aimed to constitute a direct discretization of the kinetic interpretation of noise.

6.1. The Hütter-Öttinger discretization. In reference [31], the authors propose a new discretization of the stochastic integral that would presumably give rise to the desired calculus. Instead of focusing on the general definition, we will examine a particular case that can be considered paradigmatic:

$$\int_0^T W_t \dot{W}_t := L^2(\Omega) - \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n \frac{1}{2} \left(\frac{W_{t_j}^2}{W_{t_{j-1}}} + W_{t_{j-1}} \right) (W_{t_j} - W_{t_{j-1}}),$$

where $\Delta_n = \{t_0, t_1, \dots, t_{n-1}, t_n\}$ is a partition of $[0, T]$, $\|\Delta_n\| = \max_{1 \leq k \leq n} (t_k - t_{k-1})$, and “ $\dot{W}_t$ ” is our notation for the new attempt of stochastic integration as described in [31]. This discretization presents the obvious drawback of being nonlinear for a linear integrand; but, moreover, it is almost surely ill-posed due to the fact that $W_{t_j} - W_{t_{j-1}} \sim \mathcal{N}(0, t_j - t_{j-1})$ along with $W_0 = 0$ with probability one (thus causing a divergence). In addition, in the interval $[0, T]$, for any $T > 0$, the Hausdorff dimension of the set of all zeros of Brownian motion is $1/2$ almost surely, as follows from Theorem 4.24 in [38] and the scale invariance of Brownian motion. In particular, the Brownian motion does not only vanish at the initial time $t = 0$, but also in an uncountable set of times within each finite interval starting from it. Just for comparison, consider the Riemann-Stieltjes-type discretization:

$$\int_0^T W_t \overset{\lambda}{\odot} dW_t := L^2(\Omega) - \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n W_{t_{j-1}^*} (W_{t_j} - W_{t_{j-1}}),$$

where $t_{j-1}^* \in [t_{j-1}, t_j]$, and therefore it can be expressed as $t_{j-1}^* = \lambda t_j + (1 - \lambda)t_{j-1}$ for any $\lambda \in [0, 1]$. It is well-known that this is a well-defined stochastic integral for any λ within this interval and yields the result:

$$\int_0^T W_t \overset{\lambda}{\odot} dW_t = \frac{W_T^2}{2} + \left(\lambda - \frac{1}{2} \right) T,$$

see Chapter 4 in [21] for the derivation. So, in contrast with the first discretization, the second choice (which is, in fact, a one-parameter family of discretizations) preserves the linearity of the integrand and gives rise to a well-posed integral.

This attempt of stochastic integral was inspired by a numerical algorithm [26]. Although our result in this section is negative, that does not mean that an algorithm cannot be reinterpreted as a stochastic integration scheme. Showing this is the goal of the next subsection.

6.2. The Fehlbeg 255/512-interpretation. In a sense, as mentioned in the previous subsection, one can think of noise interpretations as suitable limits of some numerical algorithms. In fact, the classical interpretations may be seen as certain limits of Runge-Kutta methods [28]. Precisely, the Itô interpretation would correspond to the forward Euler method, the HK interpretation to the backward Euler method, and the Stratonovich interpretation to the midpoint method. The last method differs in general to the Heun or trapezoidal method, but they coincide in the limit provided enough regularity is granted [45]; that can also be deduced, in a particular case, from the Riemann-Stieltjes-type discretization in the previous subsection. Moreover, most Runge-Kutta methods would reduce to one of these three integration schemes (with a clear preference towards the Stratonovich interpretation) in a suitable limit and if enough regularity is assumed. Interestingly, this is not the case for the first-order Runge-Kutta method introduced by Fehlbeg in Section IV of [23]. Assuming enough regularity, this scheme would correspond to the choice $\lambda = 255/512$ in the Riemann-Stieltjes scheme introduced in Subsection 6.1; obviously, the interpretations of Itô, Stratonovich, and HK correspond, respectively, to $\lambda = 0, 1/2, 1$. Since $255/512 = 1/2 - 1/512$, this interpretation could be considered as a perturbation of that of Stratonovich. We can introduce it in more generality as follows (from now on we assume $0 \leq a < b < \infty$).

Definition 6.1. For any $\Phi \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ the Fehlbeg 255/512–integral of $(\Phi(W_t))_{t \geq 0}$ in the interval $[a, b]$ with respect to the Brownian motion $\{W_t\}_{t \geq 0}$ is defined as

$$(34) \quad \int_a^b \Phi(W_t) \star dW_t := \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n \Phi(W_{t_j}^*) (W_{t_j} - W_{t_{j-1}}) \quad \text{in probability,}$$

where $\Delta_n = \{t_0, t_1, \dots, t_{n-1}, t_n\}$ is a partition of $[a, b]$, $\|\Delta_n\| := \max_{1 \leq j \leq n} (t_j - t_{j-1})$ is its diameter, and $t_{j-1}^* = \lambda t_j + (1 - \lambda)t_{j-1}$ with $\lambda = 255/512$.

As happens with the Stratonovich and HK integrals, Definition 6.1 poses a limit that exists and is connected to the Itô integral.

Theorem 6.2. For any $\Phi \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ the limit (34) exists and we have that the equality

$$(35) \quad \int_a^b \Phi(W_t) \star dW_t = \int_a^b \Phi(W_t) dW_t + \frac{255}{512} \int_a^b \Phi'(W_t) dt \quad \text{holds almost surely.}$$

Proof. Following [19], the right-hand side of equation (35) is almost surely well defined. Because $\Phi \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$, the Itô integral is defined as the limit in probability

$$(36) \quad \int_a^b \Phi(W_t) dW_t = \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n \Phi(W_{t_{j-1}}) (W_{t_j} - W_{t_{j-1}})$$

for any family of partitions Δ_n such that $\|\Delta_n\| \rightarrow 0$ as $n \rightarrow \infty$. For one such family Δ_n take the difference between the limits on the right-hand sides of (34) and (36) to get

$$(37) \quad \begin{aligned} D_{\Delta_n} &= \sum_{j=1}^n \left[\Phi(W_{t_j}^*) - \Phi(W_{t_{j-1}}) \right] (W_{t_j} - W_{t_{j-1}}) \\ &= \sum_{j=1}^n \Phi'((1 - \theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*) (W_{t_j}^* - W_{t_{j-1}})^2 \\ &\quad + \sum_{j=1}^n \Phi'((1 - \theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*) (W_{t_j} - W_{t_j}^*) (W_{t_j}^* - W_{t_{j-1}}), \end{aligned}$$

where we have used the continuous differentiability of Φ together with the mean value theorem; $0 < \theta_j < 1$, $j = 1, \dots, n$ are fixed parameters. Now, by Lemmata 7.2.1 and 7.2.3 in [37], we find

$$(38) \quad \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n \Phi'((1 - \theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*) (W_{t_j}^* - W_{t_{j-1}})^2 = \lambda \int_a^b \Phi'(W_t) dt$$

in probability.

To analyze the last line of equation (37), note that if Φ' were a bounded function (additionally to continuous), then

$$\begin{aligned}
(39) \quad & \mathbb{E} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*) (W_{t_j} - W_{t_j}^*) (W_{t_j}^* - W_{t_{j-1}}) \right| \right] \\
& \leq \mathbb{E} \left[\sum_{j=1}^n \left| \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*) \right| |W_{t_j}^* - W_{t_{j-1}}| \mathbb{E} \left[|W_{t_j} - W_{t_j}^*| \middle| \mathcal{F}_{t_j}^* \right] \right] \\
& \leq \sqrt{\frac{2(1-\lambda)\|\Delta_n\|}{\pi}} \mathbb{E} \left[\sum_{j=1}^n \left| \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*) \right| |W_{t_j}^* - W_{t_{j-1}}| \right] \\
& \leq \sqrt{\frac{2(1-\lambda)\|\Delta_n\|}{\pi}} \|\Phi'\|_\infty \mathbb{E} \left[\sum_{j=1}^n |W_{t_j}^* - W_{t_{j-1}}|^2 \right]^{1/2} \\
& = \sqrt{\frac{2\lambda(1-\lambda)\|\Delta_n\|}{\pi}} \|\Phi'\|_\infty \left[\sum_{j=1}^n (t_j - t_{j-1}) \right]^{1/2} \\
& = \sqrt{\frac{2\lambda(1-\lambda)\|\Delta_n\|}{\pi}} \|\Phi'\|_\infty \sqrt{b-a} \longrightarrow 0, \quad \text{as } n \rightarrow \infty,
\end{aligned}$$

by the triangle inequality, the tower property, the independence of the increments of Brownian motion, the value of the first absolute moment of Brownian motion, the Jensen inequality, and the variance of the Brownian increments. Thus, the last line of equation (37) converges to zero in $L^1(\Omega)$, and hence in probability (provided $\|\Phi'\|_\infty$ is well defined).

However, Φ' is continuous but not globally bounded in general; nevertheless, its continuity implies it is bounded on any closed interval, so the maximum

$$\max_{W_t \in [-l, l]} \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*)$$

is well defined, where the full trajectory of the Brownian motion, for $a \leq t \leq b$, is assumed to be constrained in $[-l, l]$, with $l > 0$ large enough. In consequence, by the almost sure continuity of Brownian motion and (39),

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j}^*) (W_{t_j} - W_{t_j}^*) (W_{t_j}^* - W_{t_{j-1}}) \right| > \epsilon \middle| \max_{a \leq t \leq b} |W_t| \leq l \right] = 0,$$

for all $\epsilon > 0$. Now, by the law of total probability,

$$\begin{aligned}
& \mathbb{P} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j^*})(W_{t_j} - W_{t_j^*})(W_{t_j^*} - W_{t_{j-1}}) \right| > \epsilon \right] \\
&= \mathbb{P} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j^*})(W_{t_j} - W_{t_j^*})(W_{t_j^*} - W_{t_{j-1}}) \right| > \epsilon \left| \max_{a \leq t \leq b} |W_t| \leq l \right] \right. \\
&\quad \times \mathbb{P} \left[\max_{a \leq t \leq b} |W_t| \leq l \right] \\
&\quad + \mathbb{P} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j^*})(W_{t_j} - W_{t_j^*})(W_{t_j^*} - W_{t_{j-1}}) \right| > \epsilon \left| \max_{a \leq t \leq b} |W_t| \not\leq l \right] \right. \\
&\quad \times \mathbb{P} \left[\max_{a \leq t \leq b} |W_t| \not\leq l \right] \\
(40) \quad &\leq \mathbb{P} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j^*})(W_{t_j} - W_{t_j^*})(W_{t_j^*} - W_{t_{j-1}}) \right| > \epsilon \left| \max_{a \leq t \leq b} |W_t| \leq l \right] \right. \\
&\quad \left. + \mathbb{P} \left[\max_{a \leq t \leq b} |W_t| \not\leq l \right] \right] \\
&\leq \mathbb{P} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j^*})(W_{t_j} - W_{t_j^*})(W_{t_j^*} - W_{t_{j-1}}) \right| > \epsilon \left| \max_{a \leq t \leq b} |W_t| \leq l \right] \right. \\
&\quad \left. + \frac{1}{l} \mathbb{E}(|W_b|) \right] \\
&= \mathbb{P} \left[\left| \sum_{j=1}^n \Phi'((1-\theta_j)W_{t_{j-1}} + \theta_j W_{t_j^*})(W_{t_j} - W_{t_j^*})(W_{t_j^*} - W_{t_{j-1}}) \right| > \epsilon \left| \max_{a \leq t \leq b} |W_t| \leq l \right] \right. \\
&\quad \left. + \frac{1}{l} \sqrt{\frac{2b}{\pi}} \rightarrow 0, \quad \text{as } l, n \rightarrow \infty, \right]
\end{aligned}$$

where we have used the Doob submartingale inequality in the previous-to-last step, see Theorem 4.5.1 in [37], and the value of the first absolute moment of Brownian motion in the final one.

Expressions (36), (37), (38), and (40) prove the existence of the left-hand side of equation (34). Lastly, by taking a subsequence if necessary, we deduce

$$\int_a^b \Phi(W_t) \star dW_t = \int_a^b \Phi(W_t) dW_t + \frac{255}{512} \int_a^b \Phi'(W_t) dt \quad \text{almost surely,}$$

since $\lambda = 255/512$ by hypothesis. $\square$

Now, we will illustrate the use of these developments in an example. Let us assume that a given physical object moves with a velocity that can be assimilated to a constant multiple of Brownian motion; the constant will henceforth be denoted as k . If its mass were unitary, what can be assumed without loss of generality, its kinetic energy $Q_t = k^2 W_t^2 / 2$ would fulfill:

$$dQ_t = k^2 W_t dW_t + \frac{k^2}{2} dt,$$

as a consequence of the Itô lemma; so it obeys the stochastic differential equation

$$dQ_t = \frac{k^2}{2} dt + k \sqrt{2Q_t} dW_t.$$

We will suppose that the particle is initially at rest, so the initial condition is $Q_0 = 0$. This equation possesses a unique solution, as guaranteed by the Watanabe-Yamada theorem, see Theorem 1 in [50]. However, if we employed Stratonovich calculus, we would find

$$dQ_t = k^2 W_t \circ dW_t.$$

This equation is solved by $Q_t = k^2 W_t^2/2$, by $Q_t = 0$, and by the uncountable family of solutions

$$Q_t = \frac{k^2}{2} \left(\overleftrightarrow{W}_{t-\tau} \right)^2,$$

where $\overleftrightarrow{W}_t := W_t$ if $t > 0$ and $\overleftrightarrow{W}_t := 0$ if $t \leq 0$, with τ being any almost surely positive random variable independent of W_t . Note that, unlike $Q_t = k^2 W_t^2/2$ and $Q_t = 0$ that are strong solutions, all solutions in this family are weak, see Section 5.3 in [40]. We emphasize that this is not a problem since, first, weak solutions are those that matter in physics (where it is common to focus on the probability density function of the solution rather than on its paths), and also because infinitely many strong solutions can be built for this model following the procedures in [16] and [17]. The reason for this infinite multiplicity of solutions is the transformation of the origin from being an instantaneously reflecting boundary to being an absorbing one as we shift the interpretation from Itô to Stratonovich, see [9]. On the other hand, we could find the stochastic differential equation in Fehlbberg form

$$dQ_t = \frac{k^2}{1024} dt + k \sqrt{2Q_t} \star dW_t,$$

which is a direct consequence of Theorem 6.2. This equation approximates in a sense the Stratonovich interpretation, but preserves the reflecting character of the boundary at zero, so it is free from the infinite multiplicity of solutions. Of course, the same result follows from any other $\lambda = 1/2 - \epsilon$ with $0 < \epsilon < 1/2$. Even one can take the limit $\epsilon \searrow 0$ to recover the unique Stratonovich solution that owns physical meaning. The advantage of the choice $\lambda = 255/512$ is simply the preexistence of the Runge-Kutta rule introduced by Fehlbberg in [23]. So, in a sense, the choice $\epsilon = 1/512$ could be considered as a rule of thumb to select a noise interpretation close to that of Stratonovich but free from the multiplicity of solutions generated by a square-root diffusion term (a type of diffusion term that is not rare in physical modeling [9]).

As a final note to this subsection, let us mention that modeling the velocity of a physical object with a Brownian motion is not as common as describing it like an Ornstein-Uhlenbeck process; the latter choice is known as the Langevin model for the random dispersal of a particle. However, two things should be said in this regard. First, the Brownian motion approximates the Langevin dispersal in the limit of vanishing friction, so our Brownian model can be considered as an effective description of the object velocity in such a limit. Moreover, all the developments presented here for a particle moving with a Brownian speed can be immediately translated to the case of a Langevin particle. To this end, one just needs to repeat the arguments set forth in [16, 17, 19] *mutatis mutandis*.

The next subsection shows some possible consequences of removing one of our key assumptions: uniform ellipticity. This will be illustrated with the three classical interpretations of noise together with the one just introduced in this subsection.

6.3. Application to heterogeneous diffusions in the absence of uniform ellipticity. In this subsection we follow reference [8], where the following model (except for the different notation) for diffusion in heterogeneous media is studied:

$$\frac{dX_t}{dt} = k |X_t|^\alpha \xi(t),$$

where $k > 0$, $\alpha \in \mathbb{R}$, $\xi(t)$ is white noise, and the interpretation is that of Stratonovich. Clearly, its diffusion term is degenerated. While the physical consequences of such a model were investigated in [8], herein we will focus on its stochastic analytical properties. Our arguments will be akin to those employed in [9, 16, 17, 19].

As a first step, we translate the model into the precise stochastic differential equation

$$dX_t = \frac{\alpha k^2}{2} X_t^{2\alpha-1} dt + k X_t^\alpha dW_t,$$

which should be a reasonable proposal while the position of the particle that diffuses in the heterogeneous medium, X_t , stays nonnegative (since the absolute value in the original model has been removed). Of course, this means that we need to assume an initial condition $X_0 = x_0 \geq 0$. It turns out that this stochastic differential equation presents a rich phenomenology in terms of its solvability as the parameter α is varied. This is shown with particular cases in what follows. They illustrate how the lack of uniform ellipticity can be more or less problematic, depending on the specific value of α .

We start with the value $\alpha = 1$, for which the equation possesses a unique solution that is strong and global, and given by the explicit formula

$$X_t = x_0 \exp(k W_t).$$

This is a consequence of the classical existence theory (see, for instance, Theorem 10.3.5 in [37] and Theorem 5.2.1 in [40]). Moreover, the solution remains positive (null) for all times if the initial condition is positive (null), so removing the absolute value is appropriate in this case. Also, the Stratonovic formulation $dX_t = k X_t \circ dW_t$ is perfectly valid in this case too.

If $\alpha = 2$, we can use Itô calculus to find

$$X_t = \frac{1}{x_0^{-1} - k W_t},$$

whenever $x_0 > 0$ and $X_t = 0$ for $x_0 = 0$. So in the latter case we have global-in-time existence of the solution, while in the former it blows up at the stopping time

$$\tau_2 = \inf\{t > 0 \mid W_t = 1/(x_0 k)\},$$

which is almost surely positive and almost surely finite. In other words, the solution ceases to exist at a random time that is finite with probability one; but in the mean time it is positive, so the removal of the absolute value from the original model is legitimate. Additionally, during the lapse of existence of the solution, the equation might be converted to Stratonovich form without altering any of these results.

For $\alpha = 1/2$, the model reads

$$dX_t = \frac{k^2}{4} dt + k \sqrt{X_t} dW_t.$$

The change $X_t = Q_t/2$ reduces this model to exactly that of the previous subsection, so all said there about Itô, Stratonovich, and Fehlbeg applies here identically whenever $x_0 = 0$. For $x_0 > 0$, we have the strong solution

$$X_t = \left(\frac{k}{2} W_t + \sqrt{x_0} \right)^2,$$

which is nonnegative for all times (thus justifying the removal of the absolute value in the equation) and positive for $0 \leq t < \tau_{1/2}$, where

$$\tau_{1/2} = \inf\{t > 0 \mid W_t = -2\sqrt{x_0}/k\},$$

which is an almost surely finite and positive stopping time. Therefore, the solution gets to zero in finite time almost surely, and the same situation described in the previous subsection is reproduced again after the time lapse $\tau_{1/2}$.

For $\alpha = 3/4$ we have

$$dX_t = \frac{3k^2}{8} \sqrt{X_t} dt + k \sqrt[4]{X_t^3} dW_t \quad \text{and} \quad dX_t = k \sqrt[4]{X_t^3} \circ dW_t.$$

For $x_0 = 0$ we have as solution $X_t = 0$ and for $x_0 \geq 0$ we obtain

$$X_t = \left(\frac{k}{4} W_t + \sqrt[4]{x_0} \right)^4$$

both for Itô and Stratonovich. So the same arguments of the previous subsection and the previous paragraph can be reproduced in this case to prove that the infinite multiplicity of solutions happens in this case in finite time almost surely for *both* interpretations of noise.

Finally, consider $\alpha = 1/4$ to arrive at the model

$$dX_t = \frac{k^2}{8} \frac{1}{\sqrt{X_t}} dt + k \sqrt[4]{X_t} dW_t \quad \text{and} \quad dX_t = k \sqrt[4]{X_t} \circ dW_t.$$

While the Stratonovich version can be posed for $x_0 \geq 0$, the Itô one can only be posed for $x_0 > 0$; in this latter case the solution reads

$$(41) \quad X_t = \sqrt[3]{\left(\frac{3k}{4} W_t + \sqrt[4]{x_0^3} \right)^4}$$

for *both* interpretations. Moreover, the Stratonovich interpretation admits the solution $X_t = 0$ for $x_0 = 0$. Then, arguing as before, the Stratonovich interpretation produces an infinite multiplicity of solutions in finite time almost surely. This can be rectified by imposing an instantaneously reflecting boundary at the origin, which selects one solution out of the infinitely many; precisely, the one given by (41). The situation gets worse in the case of Itô, in which the solution ceases to exist in finite time almost surely due to the divergence in the drift.

All in all, we have shown the enormous consequences that changing the noise interpretation might have regarding existence and uniqueness of solution with a single model provided with a tunable parameter (and in the absence of uniform ellipticity). For some values of this parameter, the outcome is independent of the noise interpretation, while for others Itô is better, and even Stratonovich can be better for some. One can easily check that, in all the cases analyzed, the Fehlbeg interpretation behaves as the Itô one, while the HK interpretation can only worsen things or leave them the same; this latter claim follows easily from the arguments set forth in [19]. Also, in all these cases, the removal of the absolute value from the original stochastic differential equation is justified by the nonnegativity of the solutions. As a final note, let us mention that these conclusions affect only the cases characterized by the explicitly examined values of α . Due to the complexity of the model, each value of α leads to an equation that deserves its own study.

One of the advantages of these models is that they are able to produce anomalously diffusing random motions [8]. The next subsection will show that the same is accessible for other models that are free from the mathematical pathologies exposed here.

6.4. The case of scaled Brownian motion. The previous subsection focused on heterogeneous diffusion processes, which were introduced in the physical literature as capable of producing anomalously diffusing paths [8]. Yet another process that enjoys the same property is the scaled Brownian motion, which is the solution of

$$\frac{dX_t}{dt} = F(t) \xi(t),$$

where $\xi(t)$ is white noise and $F : [0, T] \rightarrow \mathbb{R}$ is deterministic, and has also been introduced in the physical literature [34]. The use of power laws, which are not necessarily continuously differentiable at the origin, as nonlinearities in the previous subsection created some technical difficulties. This raises the question of possible mathematical difficulties when power laws are used for $F(t)$, as is done in [34]. In this subsection we show that, contrary to what happened in the two previous subsections, the change in interpretation does not alter the results at all. Moreover, these are free from pathologies and reduce to the classical results one obtains for deterministic integrals under a mild regularity assumption. Since the equation for the scaled Brownian motion is just a stochastic integral, we start introducing what we mean by this sort of integration.

Definition 6.3. For any $0 \leq a < b < \infty$ and any $F : [a, b] \rightarrow \mathbb{R}$ continuous and of bounded variation, its λ -integral with respect to the Brownian motion $\{W_t\}_{t \geq 0}$ is defined as

$$(42) \quad \int_a^b F(t) \overset{\lambda}{\odot} dW_t := \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n F(t_j^*) (W_{t_j} - W_{t_{j-1}}) \quad \text{in probability,}$$

where $\Delta_n = \{t_0, t_1, \dots, t_{n-1}, t_n\}$ is a partition of $[a, b]$, $\|\Delta_n\| := \max_{1 \leq j \leq n} (t_j - t_{j-1})$ is its diameter, and $t_{j-1}^* = \lambda t_j + (1 - \lambda)t_{j-1}$ with $0 \leq \lambda \leq 1$.

The main result of this section proves that the limit exists and is uniform in λ . In other words, the equation is independent of the interpretation. Note that the result is obvious from the established literature for continuously differentiable functions, but here we assume much less regularity. In particular, $F(\cdot)$ is not necessarily absolutely continuous, so its Radon-Nikodym derivative does not need to be well defined.

Lemma 6.4. *The limit (42) exists and moreover*

$$\int_a^b F(t) \overset{\lambda}{\odot} dW_t = \int_a^b F(t) dW_t \quad \text{almost surely.}$$

Proof. The integral on the right-hand side is a well-defined Itô integral that admits the Riemann sum approximation in Definition 6.3 with $\lambda = 0$, see Theorem 21 of Section 5 in Chapter II of [45]; hence the case $\lambda = 0$ is a tautology. For $\lambda \neq 0$ compute

$$\int_a^b F(t) \overset{\lambda}{\odot} dW_t - \int_a^b F(t) dW_t = \mathbb{P} - \lim_{\|\Delta_n\| \rightarrow 0} D_{\Delta_n},$$

where

$$D_{\Delta_n} = \sum_{j=1}^n \left[F(t_j^*) - F(t_{j-1}) \right] (W_{t_j} - W_{t_{j-1}}).$$

Therefore

$$\begin{aligned} \mathbb{E} [|D_{\Delta_n}|] &= \mathbb{E} \left[\left| \sum_{j=1}^n \left[F(t_j^*) - F(t_{j-1}) \right] (W_{t_j} - W_{t_{j-1}}) \right| \right] \\ &\leq \sum_{j=1}^n |F(t_j^*) - F(t_{j-1})| \mathbb{E} [|W_{t_j} - W_{t_{j-1}}|] \\ &\leq \sqrt{\frac{2\|\Delta_n\|}{\pi}} \sum_{j=1}^n |F(t_j^*) - F(t_{j-1})| \\ &\leq \sqrt{\frac{2\|\Delta_n\|}{\pi}} \sum_{j=1}^n [|F(t_j) - F(t_j^*)| + |F(t_j^*) - F(t_{j-1})|] \rightarrow 0, \quad \text{as } n \rightarrow \infty, \end{aligned}$$

by the triangle inequality, the value of the first absolute moment of the Brownian increments, and the bounded variation of the integrand. Then,

$$L^1(\Omega) - \lim_{\|\Delta_n\| \rightarrow 0} D_{\Delta_n} = 0 \implies \mathbb{P} - \lim_{\|\Delta_n\| \rightarrow 0} D_{\Delta_n} = 0,$$

and, hence, the limit (42) exists. Now, by passing to a subsequence if necessary, the equality in the statement is established almost surely. $\square$

Remark 6.5. Note that the hypothesis on the bounded variation of the integrands covers the case of simple power laws studied in [34]. In fact, it covers the case of any function that is monotonic on the integration interval, assumed compact.

Corollary 6.6. *For any $0 \leq a < b < \infty$ and any $F : [a, b] \rightarrow \mathbb{R}$ continuous and of bounded variation, its λ -integral*

$$\int_a^b F(t) \overset{\lambda}{\odot} dW_t = F(b) W_b - F(a) W_a - \int_a^b W_t dF(t)$$

almost surely for every $0 \leq \lambda \leq 1$.

Proof. First note that the integral on the right-hand side is almost surely well defined as a Riemann–Stieltjes integral because the integrator is of bounded variation and the integrand is continuous with probability one. Moreover, the integral on the left-hand side is well defined by Lemma 6.4. Again by this lemma, we have that

$$\begin{aligned} \int_a^b F(t) \overset{\lambda}{\odot} dW_t &= \int_a^b F(t) dW_t \\ &= \mathbb{P} - \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n F(t_{j-1}) (W_{t_j} - W_{t_{j-1}}). \end{aligned}$$

On the other hand, the integral

$$\begin{aligned} \int_a^b W_t dF(t) &= \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n W_{t_{j-1}} [F(t_j) - F(t_{j-1})] \\ &= \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n W_{t_j} [F(t_j) - F(t_{j-1})] \end{aligned}$$

almost surely by the properties of Riemann–Stieltjes integration. Now, since almost sure convergence implies convergence in probability, we find

$$\begin{aligned} &\int_a^b F(t) \overset{\lambda}{\odot} dW_t + \int_a^b W_t dF(t) \\ &= \mathbb{P} - \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n F(t_{j-1}) (W_{t_j} - W_{t_{j-1}}) + \mathbb{P} - \lim_{\|\Delta_n\| \rightarrow 0} \sum_{j=1}^n W_{t_j} [F(t_j) - F(t_{j-1})] \\ &= F(b) W_b - F(a) W_a, \end{aligned}$$

because the cancellation of the mixed terms leads to a telescopic sum. By passing to a subsequence if necessary, the equality is the statement becomes established. $\square$

Remark 6.7. This corollary shows that the classical integration by parts formula of the Riemann–Stieltjes integral is still valid in this case.

Note that the two results introduced in this subsection extend, for the particular case of a deterministic integrand, Theorem 8.3.4 in [37] and Theorem 30 in Section 5 of Chapter V in [45] (with, obviously, many more simplifying assumptions in the case of this second statement). Note also that they directly imply existence and uniqueness of solution to the scaled Brownian motion equation. Some types of singular functions that can be integrated within the present framework, but not in the classical one, are devil staircases [22], the Minkowski question mark function [42], the Lebesgue function [35], and their extensions [24].

Overall, models for anomalously diffusing particles have been introduced in [8] and [34]. While both types of models rely on the presence of power laws, this type of functional dependence is introduced in different manners. This implies that the first sort of models presents technical difficulties that are absent in the second, as we have proven in this and the previous subsection.

7. CONCLUSIONS

The notion of kinetic interpretation of noise was introduced in [31], although it might have been introduced under another format elsewhere, due to the natural character of this idea. Indeed, it could seem appealing to consider the diffusion tensor as more fundamental than the noise amplitude (which is its square root), given its appearance in the fluctuation-dissipation relation. As such, it can be considered a reasonable datum to provide as input for a mathematical model. If this model were a Fokker-Planck equation, we would have only two possibilities: the Itô form (7) and the Fick form (3). Any other form of the Fokker-Planck equation relies on a fractional power of the diffusion tensor rather than on the tensor itself (this can be seen, for instance, in equation (9) of [41]). Since there is a clear connection between the Itô form of the Fokker-Planck equation and SDEs (see Section 2), one could be tempted to conjecture some similar relation starred by the Fick form. However, this relation only exists under a very restrictive structural condition on the diffusion tensor (at least, if one wants to remove the extra drift term), precisely $\nabla \cdot \mathbf{D} = 2\sigma \nabla \cdot \sigma^\top$, what makes it highly non-generic (except in one dimension) due to the non-commutativity of the matrix algebra. Section 5 was devoted to state this condition as well as to give examples of diffusion tensors that either fulfill it or not. Still, some of the very structured tensors that fulfill it are of physical relevance.

Section 6 addresses a number of problems that arise from the analysis of the kinetic interpretation of noise. The first is the notion of stochastic integration introduced in [31], which we have seen to be ill-posed. This notion was motivated by the classical *Fixman algorithm* [26]. To show that it is possible to introduce notions of stochastic integration from numerical algorithms, we illustrated so with the Fehlberg 255/512–interpretation in that section. Subsequently, we used it, along with the other three more classical interpretations, to model stochastic transport in heterogeneous media. Moreover, we employed the resulting models to illustrate what happens when one of our main hypotheses, the uniform ellipticity assumption, is removed. Finally, we analyzed a family of models that are capable of producing anomalously diffusing paths, as the previous ones, but, in contrast to those, enjoy much more mathematical tractability.

Although our work is primarily methodological and aims to establish a mathematical framework for the rigorous analysis of physical models, it may also have practical implications. As already mentioned, the attempt of stochastic integration in [31] was motivated by the Fixman algorithm [26]. More recent numerical schemes depart from this [10, 11, 12]. Our work suggests reconsidering numerical analytical methods that arise from the kinetic interpretation of noise. In particular, it points to a direct numerical approach: determining the Itô SDE corresponding to a given convection-diffusion equation as shown in Section 2, and subsequently employing the numerical method in [7]. However, given the complexity of fluctuating hydrodynamics, it is possible that a different approach is required in that field.

We also believe that our developments may help lay the groundwork for a theoretical framework for designing and analyzing media capable of manipulating diffusive trajectories. Specifically, spatially engineered diffusion tensors, coupled with an appropriate choice of stochastic interpretation, can give rise to the phenomenon of *stochastic cloaking*, in which certain regions of space are effectively hidden from particles undergoing random diffusive motion. Stochastic cloaking thus represents a potential application of the principles of stochastic calculus introduced here to problems involving the control of random transport processes. In this sense, our framework may provide part of the mathematical foundation for the design of metamaterials capable of concealing regions of space from diffusing particles. In [46], it is shown how the properties of the diffusion tensor, together with the choice of stochastic interpretation, can be combined to controllably modify diffusion in space. In particular, that work suggests that the choice of the Itô interpretation is key to achieving this effect, despite the seemingly more natural character of the Fick law in that context.

Reference [46] provides yet another example, along with [4, 5, 9, 13, 16, 17, 18, 19], of the greater flexibility of the Itô interpretation in a wide range of physical applications. In part, our present work points in the same direction, but, simultaneously, highlights a specific structure of the diffusion

tensor that permits giving a precise meaning to the kinetic interpretation of noise. Perhaps that interpretation is not as flexible as that of Itô, but the highlighted structure could be of relevance in some physical problems.

ACKNOWLEDGEMENTS

This work has been partially supported by the Government of Spain (Ministerio de Ciencia, Innovación y Universidades) and the European Union through Projects PID2024-158823NB-I00 and CPP2024-011557.

REFERENCES

- [1] C. Ancey, P. Bohorquez, and J. Heyman. *Stochastic interpretation of the advection-diffusion equation and its relevance to bed load transport*. Journal of Geophysical Research: Earth Surface, 120:2529–2551, 2015.
- [2] L. Bachelier. *Théorie de la spéculation*. Annales Scientifiques de l'École Normale Supérieure, 3:21–86, 1900.
- [3] Rajendra Bhatia. *Positive Definite Matrices*. Princeton University Press, Princeton, 2009.
- [4] A. Bhattacharyay. *Generalization of Stokes-Einstein relation to coordinate dependent damping and diffusivity: an apparent conflict*. Journal of Physics A: Mathematical and Theoretical, 53:075002, 2020.
- [5] A. Bhattacharyay. Brownian motion near a wall: the dilemma of Itô or Stratonovich. Journal of Physics A: Mathematical and Theoretical, 58:213001, 2025.
- [6] H. B. Callen and T. A. Welton. *Irreversibility and generalized noise*. Physical Review, 83:34–40, 1951.
- [7] L. Chen, E. R. Jakobsen, and A. Naess. *On numerical density approximations of solutions of SDEs with unbounded coefficients*. Advances in Computational Mathematics, 44:693–721, 2018.
- [8] A. G. Cherstvy, A. V. Chechkin, and R. Metzler. *Anomalous diffusion and ergodicity breaking in heterogeneous diffusion processes*. New Journal of Physics, 15:083039, 2013.
- [9] Á. Corrales and C. Escudero. *Itô vs Stratonovich in the presence of absorbing states*. Journal of Mathematical Physics, 60:123301, 2019.
- [10] S. Delong, F. Balboa Usabiaga, R. Delgado-Buscalioni, B. E. Griffith, and A. Donev. *Brownian dynamics without Green's functions*. The Journal of Chemical Physics, 140:134110, 2014.
- [11] S. Delong, B. E. Griffith, E. Vanden-Eijnden, and A. Donev. *Temporal integrators for fluctuating hydrodynamics*. Physical Review E, 87:033302, 2013.
- [12] S. Delong, Y. Sun, B. E. Griffith, E. Vanden-Eijnden, and A. Donev. *Multiscale temporal integrators for fluctuating hydrodynamics*. Physical Review E, 90:063312, 2014.
- [13] A. Dhawan and A. Bhattacharyay. *Itô-distribution from Gibbs measure and a comparison with experiment*. Physica A: Statistical Mechanics and its Applications, 637:129599, 2024.
- [14] A. Einstein. *Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen*. Annalen der Physik, 17:549–560, 1905.
- [15] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer-Verlag, New York, 2000.
- [16] C. Escudero. *Kinetic energy of the Langevin particle*. Studies in Applied Mathematics, 145:719–738, 2020.
- [17] C. Escudero. *Fluctuation-dissipation relation, Maxwell-Boltzmann statistics, equipartition theorem, and stochastic calculus*. Physica Scripta, 98:055214, 2023.
- [18] C. Escudero and C. Manada. *Itô versus Stratonovich in a stochastic cosmological model*. Letters in Mathematical Physics, 112:12, 2022.
- [19] C. Escudero and H. Rojas. *Itô versus Hänggi-Klimontovich*. arXiv preprint arXiv:2309.03654, 2023.
- [20] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, 1998.

- [21] L. C. Evans. *An Introduction to Stochastic Differential Equations*. American Mathematical Society, Providence, 2013.
- [22] K. J. Falconer. *One-sided multifractal analysis and points of non-differentiability of devil's staircases*. Mathematical Proceedings of the Cambridge Philosophical Society, 136:167–174, 2004.
- [23] E. Fehlberg. *Low-order classical Runge-Kutta formulas with stepsize control and their application to some heat transfer problems*. Technical Report R-315, NASA, 1969.
- [24] J. Fernández Sánchez, P. Viader, J. Paradís, and M. Díaz Carrillo. *A singular function with a non-zero finite derivative*. Nonlinear Analysis: Theory, Methods & Applications, 75:5010–5014, 2012.
- [25] A. Fick. *Ueber diffusion*. Annalen der Physik, 94:59–86, 1855.
- [26] M. Fixman. *Simulation of polymer dynamics. I. General theory*. The Journal of Chemical Physics, 69:1527–1537, 1978.
- [27] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, 2013.
- [28] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer-Verlag, Berlin, 1993.
- [29] C. V. Henri and E. Diamantopoulos. *Unsaturated transport modeling: Random-walk particle-tracking as a numerical-dispersion free and efficient alternative to Eulerian methods*. J. Adv. Model. Earth Syst., 14:e2021MS002812, 2022.
- [30] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1994.
- [31] M. Hütter and H. C. Öttinger. *Fluctuation-dissipation theorem, kinetic stochastic integral and efficient simulations*. Journal of the Chemical Society, Faraday Transactions, 94:1403–1406, 1998.
- [32] K. Itô. *Stochastic integral*. Proceedings of the Imperial Academy, Tokyo, 20:519–524, 1944.
- [33] K. Itô. *On a stochastic integral equation*. Proceedings of the Imperial Academy, Tokyo, 22:32–35, 1946.
- [34] J.-H. Jeon, A. V. Chechkin, and R. Metzler. *Scaled Brownian motion: a paradoxical process with a time dependent diffusivity for the description of anomalous diffusion*. Phys. Chem. Chem. Phys., 16:15811–15817, 2014.
- [35] K. Kawamura. *On the set of points where Lebesgue's singular function has the derivative zero*. Proc. Japan Acad., Ser. A, Math. Sci., 87:162–166, 2011.
- [36] R. Kubo. *The fluctuation-dissipation theorem*. Reports on Progress in Physics, 29:255–284, 1966.
- [37] H. H. Kuo. *Introduction to Stochastic Integration*. Springer, New York, 2006.
- [38] P. Mörters and Y. Peres. *Brownian Motion*. Cambridge University Press, Cambridge, 2010.
- [39] T. Nordam, R. Kristiansen, R. Nepstad, E. van Sebille, and A. M. Booth. *A comparison of Eulerian and Lagrangian methods for vertical particle transport in the water column*. Geosci. Model Dev., 16:5339–5363, 2023.
- [40] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, Berlin, 2003.
- [41] A. Pacheco-Pozo, I. M. Sokolov, R. Metzler, and D. Krapf. *Heterogeneous diffusion in a harmonic potential: the role of the interpretation*. New J. Phys., 27:064602, 2025.
- [42] J. Paradís, P. Viader, and L. Bibiloni. *The derivative of Minkowski's $\varphi(x)$ function*. Journal of Mathematical Analysis and Applications, 253:107–125, 2001.
- [43] L. Perez, J. J. Hidalgo, and M. Dentz. *Reactive random walk particle tracking and its equivalence with the advection-diffusion-reaction equation*. Water Resources Research, 55:847–855, 2019.
- [44] G. Pérez-Illanes and D. Fernández-García. *MODPATH-RW: A random walk particle tracking code for solute transport in heterogeneous aquifers*. Groundwater, 62:617–634, 2024.
- [45] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin, 2005.

- [46] C. Roberts, Z. Zhang, H. Rojas, S. Bo, C. Escudero, S. Guenneau, and G. Pruessner. *Stochastic cloaking: concealing a region from diffusive particles*. Physical Review Research, 7:L032025, 2025.
- [47] D. Spivakovskaya, J. M. Redondo, and J. Chernich. *Lagrangian modelling of multi-dimensional advection-diffusion with space-varying diffusivities: theory and idealized test cases*. Ocean Dynamics, 57:189–203, 2007.
- [48] R. L. Stratonovich. *A new representation for stochastic integrals and equations*. SIAM Journal on Control, 4:362–371, 1966.
- [49] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. Springer-Verlag, New York, 1979.
- [50] S. Watanabe and T. Yamada. *On the uniqueness of solutions of stochastic differential equations*. Journal of Mathematics of Kyoto University, 11:155–167, 1971.
- [51] R. Zwanzig. *Nonequilibrium Statistical Mechanics*. Oxford University Press, New York, 2001.

Carlos Escudero
Departamento de Matemáticas Fundamentales
Universidad Nacional de Educación a Distancia
`cescudero@mat.uned.es`

Helder Rojas
Departamento de Matemáticas Fundamentales
Universidad Nacional de Educación a Distancia
`hrojas36@alumno.uned.es`