Fast solutions for the first-passage distribution of diffusion models with space-time-dependent drift functions and time-dependent boundaries

Udo Boehm [a],[*],[1], Sonja Cox [b],[1], Gregor Gantner [b],[1],[2], Rob Stevenson [b],[1]

[a] *Department of Psychology, University of Amsterdam, PO Box 15906, 1001 NK Amsterdam, The Netherlands*
[b] *Korteweg–de Vries (KDV) Institute for Mathematics, University of Amsterdam, PO Box 94248, 1090 GE Amsterdam, The Netherlands*

A R T I C L E   I N F O

A B S T R A C T

Diffusion models with constant boundaries and constant drift function have been successfully applied to model phenomena in a wide range of areas in psychology. In recent years, more complex models with time-dependent boundaries and space-time-dependent drift functions have gained popularity. One obstacle to the empirical and theoretical evaluation of these models is the lack of simple and efficient numerical algorithms for computing their first-passage time distributions. In the present work we use a known series expansion for the first-passage time distribution for models with constant drift function and constant boundaries to simplify the Kolmogorov backward equation for models with time-dependent boundaries and space-time-dependent drift functions. We show how a simple Crank–Nicolson scheme can be used to efficiently solve the simplified equation.

1. Introduction

Diffusion models have been introduced in psychology to account for behavioural data from two-alternative forced choice tasks. Over the last four decades these models have been applied to numerous domains, including perceptual decision-making (Bogacz et al., 2006; Ratcliff, 2002; Smith et al., 2014, 2004), multisensory decision-making (Diederich, 1995; Nidiffer et al., 2018), memory retrieval (McKoon & Ratcliff, 1996; Starns, 2014; White et al., 2014), lexical decision-making (Ratcliff et al., 2004; Wagenmakers et al., 2008; Yap et al., 2015), and neurophysiology (Churchland et al., 2008; Kühn et al., 2011; Philiastides, 2006; Purcell et al., 2010). Much of this success of diffusion models is due to their ability to simultaneously account for the complete distribution of observed response times and accuracies.

Diffusion models describe the process by which a decision maker chooses between two response options as a one-dimensional bounded diffusion process. Each decision boundary corresponds to one of the two response options. The response time and decision outcome are determined by the first-passage of the process at one of the two boundaries. Hence, in order to

fit a diffusion model to data, researchers need to compute the first-passage time distributions at both boundaries.

In many classical models the drift of the diffusion process is assumed to have a simple form and the decision boundaries are assumed to be constant over time. For these models simple numerical approximation schemes or series expansion for first-passage time distribution can be found (e.g., Diederich & Busemeyer, 2003; Ratcliff, 1978; Voss & Voss, 2008). Motivated by misfits between classical models and behavioural data from primates, several authors have suggested diffusion models with more complex drift functions and time-dependent boundaries (e.g., Churchland et al., 2008; Ditterich, 2006a, 2006b; Drugowitsch et al., 2012; see Boehm et al., 2016 for a review). One obstacle to wider application and empirical testing of these models is the difficulty in computing the first-passage time distribution Hawkins et al. (2015).

In this article we extend and improve the approach introduced by Voss and Voss (2008) to calculate the distribution of the first-passage time for general diffusion processes. Voss and Voss (2008) numerically solve a Fokker–Planck partial differential equation (PDE) that arises as the backward Kolmogorov equation for the distribution function of the first-passage time of the Ratcliff (1978) diffusion model, that is, a model involving constant boundaries, drift, and diffusion. This approach can also be applied to diffusion models with time and space-dependent drift and time-dependent boundaries. More specifically, the distribution function of the first passage time in a given (fixed) time point and for a range of starting points is given by the solution to a

* Corresponding author.
E-mail addresses: u.bohm@uva.nl (U. Boehm), s.g.cox@uva.nl (S. Cox), g.gantner@uva.nl (G. Gantner), r.p.stevenson@uva.nl (R. Stevenson).

Fokker–Planck equation. This is explained in detail in Section 2 below.

Voss and Voss (2008) use a Crank–Nicolson/finite difference scheme to numerically solve their Fokker–Planck equation. While this approach is reasonably efficient, it does lead to a large approximation error for small response times. This is the consequence of a discontinuity between the initial data and the boundary conditions in the Fokker–Planck PDE. We overcome this problem by first subtracting a known function that satisfies the same initial and boundary conditions from the solution to the PDE, which results in a PDE for the difference function with homogeneous (i.e., well-behaved) initial- and boundary conditions, see Section 3.2 for details. However, in order to be able to subtract such a known function, we first need to transform the Fokker–Planck PDE to a rectangular domain: indeed, we use that the series expansion of the solution to the Fokker–Planck PDE with constant drift on a rectangular domain and prescribed initial- and boundary conditions is known. The transformation is described in Section 3.1. What remains after the transformation and subtraction is a Fokker–Planck PDE on a rectangular domain with homogeneous boundary- and initial conditions. In Section 3.3, we explain how to solve this PDE numerically by means of a Crank–Nicolson/finite difference scheme (i.e., we apply the same numerical scheme as Voss and Voss (2008), but to a 'nicer' PDE). Finally, in Section 4, we use the approach outlined above to calculate the distribution function of the first passage time for three different models: the Ornstein–Uhlenbeck model, a model with urgency signal, and a model with collapsing boundaries. The numerical experiments confirm that the error for the approximate distribution functions of the first passage time is lower when using our procedure of first reducing to a Fokker–Planck PDE with homogeneous boundary conditions instead of numerically solving the original Fokker–Planck PDE.

Various other approaches for calculating the distribution and density of the first passage time have been considered in the literature. For example, several authors suggested using simulation-based methods (e.g., Brandon & Sederberg, 2014; Fengler et al., 2020; Radev et al., 2020; Wood, 2010). These methods use large synthetic data sets to approximate the first-passage time distribution. Fengler et al. (2020) and Radev et al. (2020), for instance, suggest training deep neural networks on simulated data to learn the functional relationship between data and the corresponding first-passage time distribution. While such simulation-based methods can be applied to a wide range of models, they also come with two disadvantages. Firstly, the computational costs of simulating large data sets are often prohibitive and require extensive computational infrastructure. Secondly, there are often no theoretical results available on the approximation error between the true and estimated first-passage time distribution (Fengler et al., 2020; Turner & Van Zandt, 2018).

A different approach to computing the first-passage time distribution relies on numerically approximating the distribution function. A first method using integral equations was introduced by Smith (2000), based on the work by Buonocore et al. (1990) and Buonocore et al. (1987). By exploiting the Markov property of the diffusion process, the first-passage time *density* can be expressed as the solution to an integral equation, which can be approximated by numerical integration. Buonocore et al. (1987) introduced a transformation for the integral equation for which the numerical integrals can be shown to converge to the first-passage time density as the step size of the numerical integration scheme decreases to zero. In addition to provable convergence, the integral equation method also provides sufficient flexibility for most psychological applications; the method can be applied to models with time-dependent boundaries and time and space-dependent drift functions. One disadvantage of the method is

that the first-passage time density can only be computed for one value of the starting point of the diffusion process at a time. Since in practical applications the starting point is typically treated as a model parameter that needs to be estimated from the data, the integral equation needs to be solved repeatedly, once for each candidate value of the starting point. Moreover, Buonocore et al. (1987) mention that for the models discussed in their work (i.e., a model with constant drift and diffusion function and the Ornstein–Uhlenbeck model), computing time grows quadratically with the number of discretisation points in the numerical integration scheme. In the method we present here, on the other hand, computing time grows linearly with the number of discretisation points.

Yet another method was proposed by Diederich and Busemeyer (2003). Their 'matrix method' is based on approximating the diffusion process that is continuous in space and time by a process that is discrete in time and space. The transition probabilities of this discrete process are chosen based on the drift and diffusion function of the diffusion process such that, as the size of the time and space steps tends to zero, the approximating process converges in distribution to the diffusion process. The density of the first-passage time of the diffusion process is then approximated by the (discrete) density of the first-passage time of the discrete process, which can be computed by a sequence of matrix multiplications. The matrix method also provides a high level of flexibility as it can be applied to a large class of diffusion models that are of interest to psychologists (see, e.g., Shinn et al., 2020). Indeed, the matrices modelling the transition probabilities of the discrete-time finite-state process are strikingly similar to the matrices we consider to discretise the Fokker–Planck PDE. This is no coincidence: as Diederich and Busemeyer (2003) point out, the transition matrix used in their matrix method is equal to the matrix used in the Euler-forward scheme (see Appendix A.4 in their paper). That is, the matrix method is equivalent to a Euler-forward/finite difference approximation of the solution of a Fokker–Planck equation with homogeneous boundary conditions and a Dirac delta function as initial data. This reveals the problems involved in this approach: firstly, the wild behaviour of the initial data leads to large errors in the calculated transition probabilities. Secondly, a drawback of the Euler forward scheme is that it is not unconditionally stable. This means that the size of the time steps in the discretisation needs to be proportional to the square of the size of the space steps. Therefore, the matrix method quickly becomes computationally unfeasible when high accuracy (i.e., small space steps) is required. The problem regarding the lack of stability is overcome in Shinn et al. (2020) by considering a Crank–Nicolson method instead of an Euler forward scheme.

In summary, both the matrix method and the method introduced in Voss and Voss (2008) are suitable for calculating the distribution of the first-passage time for general diffusion models. When we compare the two, we see that the matrix model is based on a Fokker–Planck PDE whose solution is the density of the diffusion process starting in a fixed position, and which is absorbed at the boundaries. By (numerically) keeping track of the portion that is absorbed at either boundary, one obtains an approximation of the density of the first passage time for the diffusion process starting in a fixed position. The method introduced in Voss and Voss (2008) is based on a Fokker–Planck PDE whose solution gives the distribution function of the first passage time in a fixed time point, but for range of starting points and thus also an approximation of the density. By solving multiple Fokker–Planck PDEs we obtain the distribution function of the first passage time for a range of time points. Both Fokker–Planck equations have properties that inhibit efficient numerical approximation: in the matrix method, the initial data is ill-behaved, in the method

of Voss and Voss (2008), one has a discontinuity in the initial-and boundary conditions. However, in the case of Voss and Voss (2008), the problem can be overcome by the method described in this paper, that is, by subtracting a known function and then solving a 'well-behaved' Fokker–Planck equation.

Our exposition is aimed at a psychological audience and is largely self-contained. In a companion paper (Boehm et al., 2021) we provide a rigorous proof that the Fokker–Planck equation with homogeneous initial and boundary conditions is more regular than the original Fokker–Planck equation. We use this regularity result to obtain error bounds for the numerical approximation of the Fokker–Planck PDE when using a space-time minimal residual method. Unfortunately, we are not able to provide (optimal) error bounds for the Crank–Nicolson method for the Fokker–Planck PDE. The minimal residual method is not as straightforward to implement as the Crank–Nicolson method, and the regularity results in Boehm et al. (2021) are beyond the scope of this paper. Thus, for the sake of clarity of the exposition, and as the numerical experiments suggest that both methods behave similarly, we restrict ourselves to the Crank–Nicolson method here and refrain from discussing the results in Boehm et al. (2021) in further detail.

2. Models with space-time-dependent drift functions

Diffusion models can be expressed formally in terms of a stochastic differential equation

$$X(t) = x_0 + \int_0^t v(X(s), s)\,ds + \int_0^t \sigma(X(s), s)\,dB(s), \quad t \in [0, \infty),$$

(1)

where $v : \mathbb{R} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is the drift function, and $\sigma : \mathbb{R} \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is the diffusion function, B is a standard Brownian motion, and $x_0 \in \mathbb{R}$ is the initial value. As diffusion models used in psychology typically assume that the diffusion function is constant, we will assume from now on that $\sigma(x, t) = \sigma$ is constant. The, possibly time-dependent, lower and upper boundaries are $l, u : \mathbb{R}_{\geq 0} \to \mathbb{R}$.

We denote the probability that $X(t)$ hits boundary u before time $T \in \mathbb{R}_{\geq 0}$ *and before hitting boundary l* by $\phi_{x_0}(T)$. Using well-known results from stochastic analysis it can be shown that this probability satisfies

$$\phi_{x_0}(T) = F(x_0, T),$$

(2)

where F is the solution to the following Fokker–Planck equation (see Appendix A in Voss and Voss (2008) for the case that l and u are constant and v does not depend on time or Chapter 7 in Øksendal (1998) for general Fokker–Planck equations)

$$\partial_t F(x, t) = \frac{\sigma^2}{2} \partial_x^2 F(x, t) + v(x, T - t)\partial_x F(x, t),$$
$$\forall\, 0 < t < T,\ l(T - t) < x < u(T - t),$$

(3a)

with initial condition

$$F(x, 0) = 0, \quad \forall\, l(T) < x < u(T)$$

(3b)

and boundary conditions

$$F(l(T - t), t) = 0, \quad F(u(T - t), t) = 1, \quad \forall\, 0 < t < T.$$

(3c)

Thus, as observed in Voss and Voss (2008), if the boundaries l and u are constant and the drift v does not depend on time, then the first passage time probability $\phi_{x_0}(t)$ is obtained *for all* $x_0 \in (l(0), u(0))$ *and all* $t \in (0, T)$ *simultaneously* by solving Eq. (3). This is illustrated in the top panel of Fig. 1. Since the boundaries and drift function do not depend on time, the probability that a process that starts at x_0 at time 0 passes the upper boundary first and before time $T - t_0$ is equal to the probability that a
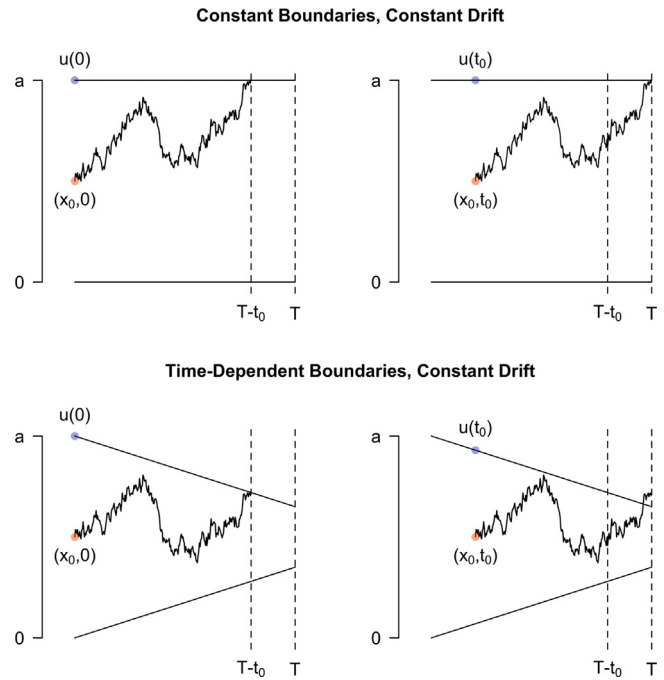

Fig. 1. Diffusion model with constant vs. time-dependent boundaries. Due to the time-homogeneity of models with constant boundaries and constant drift function, the first-passage time distribution can be obtained by solving a single Fokker–Planck equation. For models with time-dependent boundaries, a different Fokker–Planck equation needs to be solved for each first-passage time probability of interest.

process that starts at x_0 at time t_0 passes the upper boundary first and before time T. The latter probability is equal to the value $F(x_0, T - t_0)$ of the solution of the corresponding Fokker–Planck equation. Hence, the first-passage time distribution for all $t \in (0, T)$ can be obtained by evaluating the solution to Eq. (3) at different values of t with constant boundary conditions.

However, if l and/or u depend on time, then *for each choice of T* the first passage time probability $\phi_{x_0}(T)$ involves a *different* Fokker–Planck equation, where the difference lies not only in the final time, but also, and more importantly, in the position of the boundaries. This is illustrated in the bottom panel of Fig. 1. Since the boundaries depend on time, the height of the boundaries differs between a process that starts at x_0 at time 0 and a process that starts at x_0 at time t_0. Hence, the first-passage time distribution for models with time-dependent boundaries cannot be obtained by evaluating the solution to Eq. (3) with the same boundary conditions at different values of t. Instead, due to the time inversion $T - t$, the Fokker–Planck equation must be solved on different time domains $(0, T)$ and the solution evaluated at the terminal time T, where T is chosen to equal the times for which the first-passage time distribution function is desired.

Similarly, if the drift v depends on time, then $\phi_{x_0}(T)$ also involves a different Fokker–Planck equation for each choice of T, where the difference lies not only in the final time, but also in the values of the drift. As a consequence, we think of T as a parameter indexing a family of Fokker–Planck equations that we wish to solve. Note that we do not think of x_0, the initial value of X, as a parameter in this context as we always obtain $\phi_{x_0}(T)$ for all $x_0 \in (l(0), u(0))$ simultaneously when we work via the Fokker–Planck equation.

For illustration, we will consider three concrete example models, two with space-time-dependent drift functions and one with time-dependent boundaries. However, we stress that the results presented in this work apply to general space-time-dependent

drift functions and time-dependent boundaries. As we will explain in Section 3.2, models with time-dependent boundaries can be reduced to the case of a space-time-dependent drift function.

2.1. Ornstein–Uhlenbeck model

The Ornstein–Uhlenbeck model has been considered as a model for response time and accuracy data in two-alternative forced choice tasks (Busemeyer & Townsend, 1992; Heath, 2000; Smith, 1995, 2010). More recently the model has also been applied to event sampling data de Haan-Rietdijk et al. (2017). In this model the drift function depends on the space variable x, which represents the current state of the diffusion process

$$v(x) = v_0 - \beta x \tag{4}$$

The boundaries are assumed to be constant and located at $l(t) = 0$ and $u(t) = a > 0$. The three parameters of this model are a, v_0 and β.

In psychological terms, the boundary separation a represents the decision-maker's response caution. The drift rate v_0 represents the rate at which information in favour of one response option over the other becomes available, and the decay term β determines how quickly information that has been accumulated earlier is pulled back to 0.

2.2. Models with urgency signal

Various models with time-dependent drift functions have been considered to account for behavioural and physiological data from primates (e.g., Churchland et al., 2008; Ditterich, 2006a; Hanks et al., 2014; Hawkins et al., 2015). As an example, we consider the drift function suggested by Churchland et al. (2008) and Hanks et al. (2014) as a representation of response urgency. The hyperbolic drift function

$$v(t) = v_0 + v_1 \frac{t}{t + \tau} \tag{5}$$

has parameters v_0, v_1, τ. Here, we will consider the hyperbolic drift function in combination with simple constant boundaries located at $l(t) = 0$ and $u(t) = a > 0$. Hence, the four parameters of this model are a, v_0, v_1 and τ.

In psychological terms, the boundary separation a again represents the decision-maker's response caution. The drift rate v_0 represents the rate at which information in favour of one response option over the other becomes available, v_1 is the maximum gain due to response urgency, and τ is the time at which half of the maximum gain is reached.

2.3. Models with collapsing boundaries

Several authors have suggested models with time-dependent decision boundaries of varying complexity (e.g., Bowman et al., 2012; Churchland et al., 2008; Ditterich, 2006a; Drugowitsch et al., 2012; Hanks et al., 2014; Hawkins et al., 2015; Milosavljevic et al., 2010). Here, we will focus on the simplest model with a constant drift function $v(x, t) = v_0$ and linear collapsing bounds (Bowman et al., 2012; Evans et al., 2020; Milosavljevic et al., 2010). The lower bound $l(t)$ and upper bound $u(t)$ in this model are

$$l(t) = a \frac{t}{2T_\infty}, \quad u(t) = a \left(1 - \frac{t}{2T_\infty}\right). \tag{6}$$

The three parameters of this model are v_0, a, and T_∞.

In psychological terms, a is the boundary separation at time $t = 0$ and T_∞ is the time at which the boundaries have collapsed completely. The drift rate v_0 represents the rate at which information becomes available.

3. Numerical approximation

Developing a numerical solution scheme for the Fokker–Planck Eq. (3) comes with two problems. The first problem is the singularity induced by the discontinuity between initial and boundary conditions, which leads to large approximation errors. Indeed, the initial condition (3b) means that $\lim_{x \uparrow u(T)} F(x, 0) = 0$ while the boundary condition (3c) means that $\lim_{t \downarrow 0} F(u(T - t), t) = 1$. Numerical methods often approximate the solution F to Eq. (3) locally by smooth functions, which cannot accommodate the jump at $(x, t) = (u(T), 0)$ and therefore provide a poor approximation of the solution near the singularity.

The second problem is the fact that the domain of the solution F to Eq. (3) is not rectangular when l and u are time-dependent. The construction and analysis of numerical schemes for evolutionary PDEs with time-dependent spatial domains is complicated. To keep our numerical approximation simple and accessible we transform Eq. (3) to a Fokker–Planck equation on the space-time domain $(0, 1) \times (0, 1)$. This has the additional benefit that it allows for easy interpolation between solutions for different parameter values, which is particularly relevant in this context as the final time T is also viewed as a parameter. More importantly, it also paves the path to solving our first problem: the series expansion of the solution to the Fokker–Planck equation on a rectangular domain with above-mentioned discontinuity in initial- and boundary conditions and constant drift is known. By subtracting this function, we obtain a Fokker–Planck PDE for which numerical approximation is more efficient.

In summary, the calculation of the solution F to Eq. (3) proceeds in three steps, each corresponding to a section below:

1. Transformation of Eq. (3) to a Fokker–Planck equation on space-time domain $(0, 1) \times (0, 1)$, see Eq. (7);
2. Calculation of $\hat{F}_{\hat{v}_0}$ by means of a series expansion, where $\hat{F}_{\hat{v}_0}$ is the solution to a Fokker–Planck equation with constant drift $\hat{v}_0$ and the same initial and boundary conditions as Eq. (3);
3. Calculation of $\hat{E} := \hat{F} - \hat{F}_{\hat{v}_0}$ by means of a Crank–Nicolson discretisation scheme, where $\hat{F}$ is the solution to Eq. (7). We use that $\hat{E}$ is the solution to a Fokker–Planck equation with homogeneous initial and boundary conditions, see Eq. (8).

A possible fourth step involves solving Eq. (3) for multiple parameter values and using interpolation to obtain the solution for a complete parameter range. This is explained in detail in Section 5 of Boehm et al. (2021).

3.1. Transformation of space-time domain to $(0, 1) \times (0, 1)$

We will carry out a transformation of the actual space-time domain $\Omega_T := \{(x, t) \in \mathbb{R} \times (0, T) : l(T - t) < x < u(T - t)\}$ of Eq. (3) to the fixed domain $\hat{\Omega} = (0, 1) \times (0, 1)$. In this transformation, we also change the spatial orientation so the inhomogeneous boundary condition ends up on the lower boundary. This change of orientation is not necessary but having the inhomogeneous boundary condition on the lower boundary simplifies the notation for the series expansion discussed in the next section. In order for this transformation to work, we need that $l < u$ on $[0, T]$.

We first choose a transformation of the space variable. To map the time-dependent space domain at each t to the fixed domain $(0, 1)$, we define

$$\hat{x} = \frac{u(T - t) - x}{u(T - t) - l(T - t)}.$$

Intuitively, this transformation inverts the space-direction and stretches the space domain at each time point t so that the

lower boundary becomes 1 and the upper boundary becomes 0. Unfortunately, this transformation of the space variable is time-dependent, and thus results in a Fokker–Planck equation with a time-dependent diffusion. This is problematic, as the trick of subtracting a known function to obtain a nicer PDE requires a constant diffusion coefficient. (Indeed, if the diffusion term is time-dependent, then the forcing function $\hat{g}$ introduced in Section 3.2 involves not only $\partial_x \hat{F}_{\hat{v}_0}$, with $\hat{F}_{\hat{v}_0}$ as introduced in Section 3.2, but also $\partial_{xx}\hat{F}_{\hat{v}_0}$. This second order derivative would render the forcing function too irregular for efficient numerical approximation.) We overcome this problem by also performing a transformation in the time variable which precisely compensates for the introduced time dependence in the diffusion coefficient. The derivations below clarify this.

We define $\tilde{T} = \int_0^T |u(T-s) - l(T-s)|^{-2}\, ds$. Then, as we will see below, a suitable choice for the transformation of the time variable is

$$\hat{t} = \frac{1}{\tilde{T}} \int_0^t |u(T-s) - l(T-s)|^{-2}\, ds,$$

where the factor $\frac{1}{\tilde{T}}$ ensures that the transformed time variable has domain $(0,1)$. The transformations of both variables are invertible, indeed,

$$t = h(\hat{t}) := \inf\left\{ r \in [0,T] \colon \frac{1}{\tilde{T}} \int_0^r |u(T-s) - l(T-s)|^{-2}\, ds > \hat{t} \right\},$$

$$x = g(\hat{x}, \hat{t}) := l(T-h(\hat{t}))\hat{x} + u(T-h(\hat{t}))(1-\hat{x}).$$

Note further that (using that for any invertible mapping ϕ we have the identity $(\phi^{-1})'(y) = [\phi'(\phi^{-1}(y))]^{-1}$)

$$h'(\hat{t}) = \tilde{T}\big|u(T-h(\hat{t})) - l(T-h(\hat{t}))\big|^2,$$

$$\partial_{\hat{t}} g(\hat{x}, \hat{t}) = -h'(\hat{t})\big(l'(T-h(\hat{t}))\hat{x} + u'(T-h(\hat{t}))(1-\hat{x})\big),$$

$$\partial_{\hat{x}} g(\hat{x}, \hat{t}) = -\big(u(T-h(\hat{t})) - l(T-h(\hat{t}))\big),$$

and

$$\partial_{\hat{x}}^2 g(\hat{x}, \hat{t}) = 0.$$

We define $\hat{F}(\hat{x}, \hat{t}) = F(g(\hat{x}, \hat{t}), h(\hat{t}))$ with domain $\hat{\Omega} = (0,1) \times (0,1)$. Using that

$$\partial_{\hat{t}}\hat{F}(\hat{x}, \hat{t}) = \partial_{\hat{t}} g(\hat{t}, \hat{x}) \partial_x F(g(\hat{x}, \hat{t}), h(\hat{t})) + h'(\hat{t})\partial_t F(g(\hat{x}, \hat{t}), h(\hat{t})),$$

$$\partial_{\hat{x}}\hat{F}(\hat{x}, \hat{t}) = \partial_{\hat{x}} g(\hat{x}, \hat{t}) \partial_x F(g(\hat{x}, \hat{t}), h(\hat{t})),$$

$$\partial_{\hat{x}}^2 \hat{F}(\hat{x}, \hat{t}) = (\partial_{\hat{x}} g(\hat{x}, \hat{t}))^2 \partial_x^2 F(g(\hat{x}, \hat{t}), h(\hat{t})),$$

we compute

$$\partial_{\hat{t}}\hat{F}(\hat{x}, \hat{t}) = \partial_{\hat{t}} g(\hat{x}, \hat{t}) \partial_x F(g(\hat{x}, \hat{t}), h(\hat{t})) + h'(\hat{t})\partial_t F(g(\hat{x}, \hat{t}), h(\hat{t}))$$

$$= \partial_{\hat{t}} g(\hat{x}, \hat{t}) \partial_x F(g(\hat{x}, \hat{t}), h(\hat{t}))$$

$$\quad + h'(\hat{t})\left[\frac{\sigma^2}{2} \partial_x^2 F(g(\hat{x}, \hat{t}), h(\hat{t})) + v(g(\hat{x}, \hat{t}), T \right.$$

$$\left. -h(\hat{t}))\partial_x F(g(\hat{x}, \hat{t}), h(\hat{t})) \right]$$

$$= \left[h'(\hat{t})(\partial_{\hat{x}} g(\hat{x}, \hat{t}))^{-2} \frac{\sigma^2}{2} \right] \partial_{\hat{x}}^2 \hat{F}(\hat{x}, \hat{t})$$

$$\quad + (\partial_{\hat{x}} g(\hat{x}, \hat{t}))^{-1} \left[\partial_{\hat{t}} g(\hat{x}, \hat{t}) \right.$$

$$\left. +h'(\hat{t})v(g(\hat{x}, \hat{t}), T - h(\hat{t})) \right] \partial_{\hat{x}}\hat{F}(\hat{x}, \hat{t})$$

$$= \frac{\sigma^2 \tilde{T}}{2} \partial_{\hat{x}}^2 \hat{F}(\hat{x}, \hat{t}) + (\partial_{\hat{x}} g(\hat{x}, \hat{t}))^{-1}$$

$$\quad \times \left[\partial_{\hat{t}} g(\hat{x}, \hat{t}) + h'(\hat{t})v(g(\hat{x}, \hat{t}), T - h(\hat{t})) \right] \partial_{\hat{x}}\hat{F}(\hat{x}, \hat{t}),$$

where our choice of the transformation of the time variable ensures that $h'(\hat{t})$ cancels the, possibly time-dependent, factor $(\partial_{\hat{x}} g(\hat{x}, \hat{t}))^{-2}$ to yield a constant diffusion coefficient. We thus obtain that $\hat{F}$ satisfies the simplified Fokker–Planck equation

$$\partial_{\hat{t}}\hat{F}(\hat{x}, \hat{t}) = \hat{\sigma}\, \partial_{\hat{x}}^2 \hat{F}(\hat{x}, \hat{t}) + \hat{v}(\hat{x}, \hat{t})\partial_{\hat{x}}\hat{F}(\hat{x}, \hat{t}), \quad \forall (\hat{x}, \hat{t}) \in \hat{\Omega}, \tag{7a}$$

where $\hat{\sigma} = \frac{\sigma^2 \int_0^T |u(T-s) - l(T-s)|^{-2}\, ds}{2}$,

$$\hat{v}(\hat{x}, \hat{t}) = (\partial_{\hat{x}} g(\hat{x}, \hat{t}))^{-1} \left[\partial_{\hat{t}} g(\hat{x}, \hat{t}) + h'(\hat{t})v(g(\hat{x}, \hat{t}), T - h(\hat{t})) \right]$$

$$= (u(T-h(\hat{t})) - l(T-h(\hat{t}))) \int_0^T |u(T-s) - l(T-s)|^{-2}\, ds$$

$$\quad \times \big[l'(T-h(\hat{t}))\hat{x} + u'(T-h(\hat{t}))(1-\hat{x})$$

$$\quad - v(l(T-h(\hat{t}))\hat{x} + u(T-h(\hat{t}))(1-\hat{x}), T - h(\hat{t})) \big], \tag{7b}$$

and we have initial condition

$$\hat{F}(\hat{x}, 0) = 0, \quad \forall\, 0 < \hat{x} < 1, \tag{7c}$$

and boundary conditions

$$\hat{F}(0, \hat{t}) = 1, \quad \hat{F}(1, \hat{t}) = 0, \quad \forall\, 0 < \hat{t} < 1. \tag{7d}$$

Applying this transformation to our example models, we obtain the following diffusion coefficients and drift functions. For the Ornstein–Uhlenbeck model, the diffusion coefficient is $\hat{\sigma} = \frac{T\sigma^2}{2a^2}$ and the drift function becomes

$$\hat{v}(\hat{x}, \hat{t}) = \hat{v}_0 - \hat{\beta}\hat{x},$$

where $\hat{v}_0 = \beta T - \frac{Tv_0}{a}$ and $\hat{\beta} = \beta T$. For the model with hyperbolic drift the diffusion coefficient is $\hat{\sigma} = \frac{T\sigma^2}{2a^2}$ and the drift function becomes

$$\hat{v}(\hat{x}, \hat{t}) = \hat{v}_0 + \hat{v}_1 \frac{1 - \hat{t}}{1 - \hat{t} + \hat{\tau}},$$

where $\hat{v}_0 = -\frac{Tv_0}{a}$, $\hat{v}_1 = -\frac{Tv_1}{a}$, and $\hat{\tau} = \frac{\tau}{T}$. Finally, for the model with linear collapsing bounds, we use that

$$\tilde{T} = \frac{TT_\infty}{a^2(T_\infty - T)}, \quad h(\hat{t}) = \frac{T(T_\infty - T)\hat{t}}{T_\infty - T\hat{t}}$$

to obtain the transformed diffusion coefficient $\hat{\sigma} = \frac{TT_\infty \sigma^2}{2a^2(T_\infty - T)}$ and the transformed drift function

$$\hat{v}(\hat{x}, \hat{t}) = \frac{TT_\infty}{a(T_\infty - T\hat{t})} \left(\frac{a}{2T_\infty}(2\hat{x} - 1) - v_0 \right).$$

3.2. Fokker–Planck equation with homogeneous boundary conditions

In the PDE (7), the initial and boundary conditions still introduce a singularity at $(\hat{x}, \hat{t}) = (0, 0)$, which causes large errors in numerical approximations to the solution. The effect of this singularity on the approximation error can be considerably weakened by subtracting the known solution for the case of a constant drift function. For a model with diffusion coefficient $\hat{\sigma}$, constant drift function $\hat{v}(\hat{x}, \hat{t}) = \hat{v}_0$ and constant boundaries located at 0 and 1, the solution to the PDE (7) can be expressed as the fast-converging series (Gondan et al., 2014)

$$\hat{F}_{\hat{v}_0}(\hat{x}, \hat{t}) = e^{-\frac{\hat{v}_0\hat{x}}{2\hat{\sigma}} - \frac{\hat{v}_0^2\hat{t}}{4\hat{\sigma}}} \sum_{k=0}^\infty (-1)^k \varphi\left(\frac{r_k}{\sqrt{2\hat{\sigma}\hat{t}}} \right)$$

$$\times \left(M\left(\frac{r_k - \hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}\hat{t}}} \right) + M\left(\frac{r_k + \hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}\hat{t}}} \right) \right),$$

where $r_k = k + \hat{x}$ if k is even and $r_k = k + 1 - \hat{x}$ if k is odd, and $M(y) = \frac{1 - \Phi(y)}{\varphi(y)}$ is the Mills ratio, with Φ the standard normal distribution function and φ the standard normal density.

The series solution for the problem with constant drift function $\hat{v}_0 = \hat{v}(0, 0)$ and the solution to the PDE (7) with a general drift function both satisfy the same initial and boundary conditions. Hence, by subtracting the series solution from the solution

to the general problem, we obtain a PDE with homogeneous initial and boundary conditions. For the difference $\hat{E}(\hat{x}, \hat{t}) = \hat{F}(\hat{x}, \hat{t}) - \hat{F}_{\hat{v}_0}(\hat{x}, \hat{t})$ we compute

$$\partial_t \hat{E}(\hat{x}, \hat{t}) = \partial_t \hat{F}(\hat{x}, \hat{t}) - \partial_t \hat{F}_{\hat{v}_0}(\hat{x}, \hat{t})$$
$$= \hat{\sigma}[\partial_{\hat{x}}^2 \hat{F}(\hat{x}, \hat{t}) - \partial_{\hat{x}}^2 \hat{F}_{\hat{v}_0}(\hat{x}, \hat{t})] + \hat{v}(\hat{x}, \hat{t})\partial_{\hat{x}}\hat{F}(\hat{x}, \hat{t}) - \hat{v}(0, 0)\partial_{\hat{x}}\hat{F}_{\hat{v}_0}(\hat{x}, \hat{t})$$
$$= \hat{\sigma}[\partial_{\hat{x}}^2 \hat{F}(\hat{x}, \hat{t}) - \partial_{\hat{x}}^2 \hat{F}_{\hat{v}_0}(\hat{x}, \hat{t})] + \hat{v}(\hat{x}, \hat{t})[\partial_{\hat{x}}\hat{F}(\hat{x}, \hat{t}) - \partial_{\hat{x}}\hat{F}_{\hat{v}_0}(\hat{x}, \hat{t})]$$
$$+ (\hat{v}(\hat{x}, \hat{t}) - \hat{v}(0, 0))\partial_{\hat{x}}\hat{F}_{\hat{v}_0}(\hat{x}, \hat{t}).$$

In summary, we obtain the following problem for the difference $\hat{E}(\hat{x}, \hat{t}) = \hat{F}(\hat{x}, \hat{t}) - \hat{F}_{\hat{v}_0}(\hat{x}, \hat{t})$:

$$\partial_t \hat{E}(\hat{x}, \hat{t}) = \hat{\sigma}\,\partial_{\hat{x}}^2 \hat{E}(\hat{x}, \hat{t}) + \hat{v}(\hat{x}, \hat{t})\partial_{\hat{x}}\hat{E}(\hat{x}, \hat{t}) + \hat{g}(\hat{x}, \hat{t}), \tag{8a}$$

with forcing function $\hat{g}(\hat{x}, \hat{t}) = (\hat{v}(\hat{x}, \hat{t}) - \hat{v}(0, 0))\partial_{\hat{x}}\hat{F}_{\hat{v}_0}$, initial condition

$$\hat{E}(\hat{x}, 0) = 0, \quad \forall\, 0 < \hat{x} < 1, \tag{8b}$$

and homogeneous boundary conditions

$$\hat{E}(0, \hat{t}) = \hat{E}(1, \hat{t}) = 0, \quad \forall\, 0 < \hat{t} < 1. \tag{8c}$$

Although the factor $\partial_{\hat{x}}\hat{F}_{\hat{v}_0}$ in the definition of the forcing function has a singularity at $(0, 0)$, the factor $(\hat{v}(\hat{x}, \hat{t}) - \hat{v}(0, 0))$ is zero at $(0, 0)$ and thus ensures that the singularity vanishes. In fact, it can be proved that this problem has a smoother solution than the original problem (7) and can therefore be approximated more accurately (see Theorems 3.1 and 4.1, as well as the final remark of Section 3 in Boehm et al. (2021)).

The forcing function $\hat{g}(\hat{x}, \hat{t})$ can be computed efficiently as the space-derivative $\partial_{\hat{x}}\hat{F}_{\hat{v}_0}$ can be expressed as a fast converging series

$$\partial_{\hat{x}}\hat{F}_{\hat{v}_0} = 2e^{-\frac{\hat{v}_0 \hat{x}}{2\hat{\sigma}} - \frac{\hat{v}_0^2 \hat{t}}{2\hat{\sigma}}} \sum_{k=0}^{\infty} (-1)^{k+1} \varphi\left(\frac{r_k}{\sqrt{2\hat{\sigma}\hat{t}}}\right)$$
$$\times \left[\frac{\hat{v}_0}{2\hat{\sigma}} M\left(\frac{r_k + (-1)^{k+1}\hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) + \frac{(-1)^k}{\sqrt{2\hat{\sigma}\hat{t}}}\right]. \tag{9}$$

If the terms $k = 0, \ldots, 2K - 1$ of the series are computed, an upper bound for the truncation error is given by

$$\epsilon = e^{-\frac{\hat{v}_0\hat{x}}{2\hat{\sigma}} - \frac{\hat{v}_0^2\hat{t}}{2\hat{\sigma}}} \left(\frac{|\hat{v}_0|\sqrt{\pi\hat{t}}}{\hat{\sigma}} + 1\right)$$
$$\times \left(2 - \Phi\left(\frac{2(K-1) + \hat{x}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) - \Phi\left(\frac{2K - \hat{x}}{\sqrt{2\hat{\sigma}\hat{t}}}\right)\right).$$

For given values of $\hat{\sigma}, \hat{x}, \hat{t}$, and $\hat{v}_0$, this expression can be evaluated for a small number of values of K and compared to the desired accuracy to determine the number of terms to which the series needs to be computed. In practice it is often sufficient to compute no more than 10 terms to achieve acceptable accuracy (e.g., for $\hat{\sigma} = 2$, $(\hat{x}, \hat{t}) \in [0, 1] \times [0, 1]$, and $-50 \leq \hat{v}_0 \leq 50$, computing 10 terms yields $\epsilon \leq 10^{-16}$). Details of the truncation criterion for the series for $\partial_{\hat{x}}\hat{F}_{\hat{v}_0}$ are given in Appendix A.

3.3. Crank–Nicolson approximation

Our numerical solution method for the PDE (8) is based on a finite difference approximation to the partial derivatives of $\hat{E}$. To ease readability, we write from now on x and t instead of $\hat{x}$ and $\hat{t}$ for space and time variables. Let $N \in \mathbb{N}$. We discretise the closure $[0, 1]$ of the space and time domains into N intervals of length $\Delta x = \Delta t = \frac{1}{N}$. We write $x_i = i\Delta x$ for $0 \leq i \leq N$ and $t_j = j\Delta t$ for $0 \leq j \leq N$. To approximate the second-order partial derivative $\partial_x^2 \hat{E}$, we develop $\hat{E}(x \pm \Delta x, t)$ in a Taylor series:

$$\hat{E}(x + \Delta x, t) = \hat{E}(x, t) + \partial_x \hat{E}(x, t)\Delta x$$

$$+ \frac{1}{2}\partial_x^2 \hat{E}(x, t)(\Delta x)^2 + \mathcal{O}((\Delta x)^3),$$
$$\hat{E}(x - \Delta x, t) = \hat{E}(x, t) - \partial_x \hat{E}(x, t)\Delta x$$
$$+ \frac{1}{2}\partial_x^2 \hat{E}(x, t)(\Delta x)^2 + \mathcal{O}((\Delta x)^3).$$

Adding the two expansions, solving for $\partial_x^2 \hat{E}$ and neglecting the terms of order $\mathcal{O}((\Delta x)^3)$ yields the so-called *centred difference*:

$$\partial_x^2 \hat{E}(x, t) \approx \frac{\hat{E}(x + \Delta x, t) - 2E(x, t) + \hat{E}(x - \Delta x, t)}{(\Delta x)^2}.$$

Similar computations yield the centred difference approximation to $\partial_x \hat{E}$

$$\partial_x \hat{E}(x, t) \approx \frac{\hat{E}(x + \Delta x, t) - \hat{E}(x - \Delta x, t)}{2\Delta x},$$

and the *forward* and *backward difference* approximations to $\partial_t \hat{E}$:

$$\partial_t \hat{E}(x, t) \approx \frac{\hat{E}(x, t + \Delta t) - \hat{E}(x, t)}{\Delta t},$$
$$\partial_t \hat{E}(x, t) \approx \frac{\hat{E}(x, t) - \hat{E}(x, t - \Delta t)}{\Delta t}.$$

To approximate the solution to Eq. (7) at (x_i, t_j), we replace the partial derivatives $\partial_x \hat{E}, \partial_x^2 \hat{E}$ in Eq. (7) by their finite difference approximations and $\partial_t \hat{E}$ by the forward difference approximation. This gives the approximation

$$\frac{\hat{E}(x_i, t_j + \Delta t) - \hat{E}(x_i, t_j)}{\Delta t}$$
$$\approx \hat{\sigma} \frac{\hat{E}(x_i + \Delta x, t_j) - 2E(x_i, t_j) + \hat{E}(x_i - \Delta x, t_j)}{(\Delta x)^2}$$
$$+ \hat{v}(x_i, t_j)\frac{\hat{E}(x_i + \Delta x, t_j) - \hat{E}(x_i - \Delta x, t_j)}{2\Delta x} + \hat{g}(x_i, t_j).$$

Using the more convenient notation $E_{i,j} = \hat{E}(x_i, t_j)$, $v_{i,j} = \hat{v}(x_i, t_j)$, $g_{i,j} = \hat{g}(x_i, t_j)$ and neglecting that the equality is only approximate, we obtain

$$\frac{E_{i,j+1} - E_{i,j}}{\Delta t} = \hat{\sigma}\frac{E_{i+1,j} - 2E_{i,j} + E_{i-1,j}}{(\Delta x)^2} + v_{i,j}\frac{E_{i+1,j} - E_{i-1,j}}{2\Delta x} + g_{i,j}.$$

If we use the backward difference approximation for $\partial_t \hat{E}$ instead, we obtain the approximation to the solution at (x_i, t_{j+1})

$$\frac{E_{i,j+1} - E_{i,j}}{\Delta t} = \hat{\sigma}\frac{E_{i+1,j+1} - 2E_{i,j+1} + E_{i-1,j+1}}{(\Delta x)^2}$$
$$+ v_{i,j+1}\frac{E_{i+1,j+1} - E_{i-1,j+1}}{2\Delta x} + g_{i,j+1}.$$

Averaging these approximations and rearranging terms, we get the *Crank–Nicolson* approximation

$$\left(-\hat{\sigma}\frac{\Delta t}{2(\Delta x)^2} + v_{i,j+1}\frac{\Delta t}{4\Delta x}\right)E_{i-1,j+1} + \left(1 + \hat{\sigma}\frac{\Delta t}{(\Delta x)^2}\right)E_{i,j+1}$$
$$- \left(\hat{\sigma}\frac{\Delta t}{2(\Delta x)^2} + v_{i,j+1}\frac{\Delta t}{4\Delta x}\right)E_{i+1,j+1} \tag{10}$$
$$= \left(\hat{\sigma}\frac{\Delta t}{2(\Delta x)^2} - v_{i,j}\frac{\Delta t}{4\Delta x}\right)E_{i-1,j} + \left(1 - \hat{\sigma}\frac{\Delta t}{(\Delta x)^2}\right)E_{i,j}$$
$$+ \left(\hat{\sigma}\frac{\Delta t}{2(\Delta x)^2} + v_{i,j}\frac{\Delta t}{4\Delta x}\right)E_{i+1,j} + G_{i,j},$$

where $G_{i,j} = \frac{\Delta t}{2}(g_{i,j+1} + g_{i,j})$.

The recursive relation (10) can be more conveniently expressed in matrix notation. Let

$$v^+ = 1 + \hat{\sigma}\frac{\Delta t}{(\Delta x)^2}, \qquad v^- = 1 - \hat{\sigma}\frac{\Delta t}{(\Delta x)^2},$$
$$\mu_{i,j}^+ = \hat{\sigma}\frac{\Delta t}{2(\Delta x)^2} + v_{i,j}\frac{\Delta t}{4\Delta x}, \qquad \mu_{i,j}^- = \hat{\sigma}\frac{\Delta t}{2(\Delta x)^2} - v_{i,j}\frac{\Delta t}{4\Delta x}.$$

We define the matrices

$$\mathbf{A}_j^{(1)} = \begin{pmatrix} \nu^+ & -\mu_{1,j}^+ & 0 & \dots & \dots & 0 \\ -\mu_{2,j}^- & \nu^+ & -\mu_{2,j}^+ & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -\mu_{N-2,j}^- & \nu^+ & -\mu_{N-2,j}^+ \\ 0 & \dots & \dots & 0 & -\mu_{N-1,j}^- & \nu^+ \end{pmatrix},$$

and

$$\mathbf{A}_j^{(2)} = \begin{pmatrix} \nu^- & \mu_{1,j}^+ & 0 & \dots & \dots & 0 \\ \mu_{2,j}^- & \nu^- & \mu_{2,j}^+ & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \mu_{N-2,j}^- & \nu^- & \mu_{N-2,j}^+ \\ 0 & \dots & \dots & 0 & \mu_{N-1,j}^- & \nu^- \end{pmatrix}.$$

Note that $\mathbf{A}_j^{(2)}$ is equal to $\mathbf{A}_j^{(1)}$ with ν^+ replaced by ν^- and the sign of $\mu_{i,j}^\pm$ reversed. Moreover, we define the vectors

$$\mathbf{E}_j = (E_{1,j} \; E_{2,j} \; \dots \; E_{N-2,j} \; E_{N-1,j})^\top \quad \text{and}$$
$$\mathbf{G}_j = (G_{1,j} \; G_{2,j} \; \dots \; G_{N-2,j} \; G_{N-1,j})^\top.$$

The recursive relation (10) can now be solved by iteratively solving the equation

$$\mathbf{A}_{j+1}^{(1)} \mathbf{E}_{j+1} = \mathbf{A}_j^{(2)} \mathbf{E}_j + \mathbf{B}_j + \mathbf{G}_j \tag{11}$$

with initial condition $\mathbf{E}_0 = (0 \; \dots \; 0)^\top$ and homogeneous boundary condition $\mathbf{B}_j = (0 \; \dots \; 0)^\top$.

4. Numerical experiments

In this section we present a numerical illustration of our method. MATLAB code that implements our method is available at https://osf.io/xv674/ and we provide a brief illustration of the code in Appendix B. As outlined in the previous section, we apply the Crank–Nicolson method to approximate $\hat{E} = \hat{F} - \hat{F}_{\hat{v}_0}$ resulting from the models (4), (5), or (6) for different parameter choices. For simplicity we assume $\sigma = 1$ throughout. We consider five equidistant values for the parameters determining the boundaries and the drift, for instance, for $v_0 \in [-1, 1]$, we consider $v_0 = -1, -0.5, 0, 0.5, 1$. Recall moreover from Eq. (3) that the evaluation time T is also a model parameter if l, u, or v depends on time, that is, it is a parameter in the models (5) and (6). For these models we consider five equidistant values for T spanning $[0.1, 2.5]$, which is a typical range of response times in diffusion model applications Ratcliff and McKoon (e.g., 2008). For comparison, we also use the Crank–Nicolson method (with adapted boundary condition $\mathbf{B}$) to directly approximate $\hat{F}$ as solution of the PDE (7). We will evaluate the Crank–Nicolson approximations to $\hat{E}$ and $\hat{F}$ at 100 equidistant values of $\hat{x} \in [0, 1]$ and $\hat{t} \in [0, 1]$. As these evaluation points do not necessarily coincide with the points at which the Crank–Nicolson approximations $E_{i,j}$ and $F_{i,j}$ are computed, we use bilinear interpolation of the adjacent Crank–Nicolson points to obtain approximations $\hat{E}_N$ and $\hat{F}_N$ to $\hat{E}$ and $\hat{F}$ at the equidistant evaluation points. Here, N denotes the number of space/time intervals.

To obtain an estimate for the error $\hat{E} - \hat{E}_N$ and $\hat{F} - \hat{F}_N$, we replace the exact solutions $\hat{E}$ and $\hat{F}$ by the improved approximations $\hat{E}_{2N}$

and $\hat{F}_{2N}$, respectively. We measure the errors in the ℓ_2- and in the ℓ_∞-norm, that is,

$$\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_2} = \left(10^{-4} \sum_{\hat{x}, \hat{t} = 0, \frac{1}{99}, \dots, 1} \left(\hat{E}_{2N}(\hat{x}, \hat{t}) - \hat{E}_N(\hat{x}, \hat{t}) \right)^2 \right)^{\frac{1}{2}}$$

and

$$\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_\infty} = \max_{\hat{x}, \hat{t} = 0, \frac{1}{99}, \dots, 1} \left| \hat{E}_{2N}(\hat{x}, \hat{t}) - \hat{E}_N(\hat{x}, \hat{t}) \right|$$

with analogous definitions for $\hat{F}$. For each considered model, we plot these quantities over the computational time. By computational time we mean the average time to compute the terms $\hat{F}_N$, $\hat{F}_{2N}$, $\hat{E}_N$, and $\hat{E}_{2N}$ and to evaluate the differences $\hat{F}_{2N} - \hat{F}_N$ and $\hat{E}_{2N} - \hat{E}_N$, respectively, for the 100×100 equidistant space and time points in $[0, 1] \times [0, 1]$. All computations were carried out on a MacBook Pro 2016, with a 2 GHz Dual-Core Intel Core i5 processor, using MATLAB version R2019a.

We use a double-logarithmic convergence plot so that the slopes of the curves coincide with the corresponding algebraic convergence rates. One says that an error converges at algebraic rate $s > 0$ with respect to computing time if, at least asymptotically, it behaves as $C \times \text{computing time}^{-s}$ with some fixed constant $C > 0$. For instance, an algebraic convergence rate 1 indicates that doubling the computing time halves the approximation error. The Crank–Nicolson method applied to the heat equation with sufficiently smooth solution yields convergence rate 1 (see, e.g., Larsson & Thomée, 2008). We also mention that the computational time is proportional to the number of Crank–Nicolson points N^2.

4.1. Ornstein–Uhlenbeck model (4)

Based on the parameter estimates reported in Matzke and Wagenmakers (2009) and Smith (2010), we consider the parameter ranges $v_0 \in [-2, 2]$, $\beta \in [-4, 4]$, $a \in [0.5, 2]$, and $T = 2.5$. In Fig. 2 we plot the errors $\hat{F}_{2N} - \hat{F}_N \approx \hat{F} - \hat{F}_N$ and $\hat{E}_{2N} - \hat{E}_N \approx \hat{E} - \hat{E}_N$ over the computational time. It is clearly visible that the naive approach of directly applying Crank–Nicolson for the singular solution $\hat{F}$ yields a much worse convergence behaviour than our new approach of applying Crank–Nicolson for the smoother solution $\hat{E}$. Indeed, our approach even yields rate 1, which one would expect for smooth functions.

Fig. 3 shows an example of the approximation to the first-passage time distribution $\phi_{x_0}(T)$ for the absolute starting point $x_0 = 0.42$ based on $\hat{E}_N$ (blue lines) and $\hat{F}_N$ (red lines). The approximation for different numbers of space/time intervals N is shown as dashed lines. As comparison standard, a high-precision approximation with estimated absolute error tolerance $\epsilon = 10^{-7}$ is shown as a solid grey line. As can be seen, the approximation $\hat{E}_N$ is visually indistinguishable from the high-precision approximation for $N > 2$. The approximation $\hat{F}_N$, on the other hand, deviates clearly from the high-precision approximation and converges slowly as N increases.

4.2. Model with urgency signal (5)

Based on the parameter estimates reported in Churchland et al. (2008) and Hanks et al. (2014), we consider the parameter ranges $v_0 \in [1.64, 1.97]$, $v_1 \in [0.99, 2.31]$, $\tau \in [0.13, 0.40]$, $a \in [1.38, 2.26]$, and $T \in [0.1, 2.5]$. In Fig. 4 we again plot the errors $\hat{F}_{2N} - \hat{F}_N \approx \hat{F} - \hat{F}_N$ and $\hat{E}_{2N} - \hat{E}_N \approx \hat{E} - \hat{E}_N$ over the computational time. Again, direct approximation of $\hat{F}$ yields a considerably worse rate than our approach, which yields rate 1.
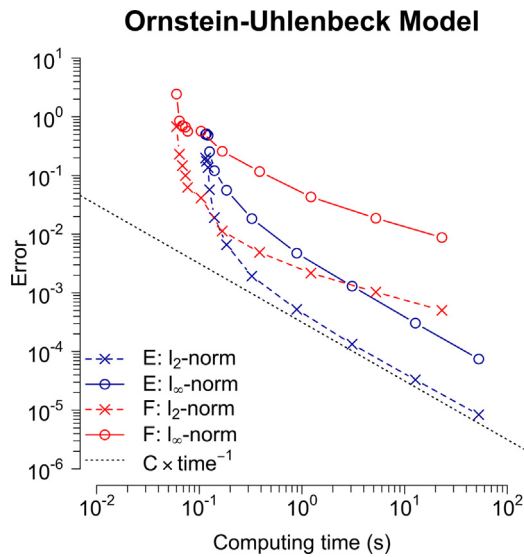
Fig. 2. Comparison of maximal errors for different approximation methods for the Ornstein–Uhlenbeck model (4). Maximal errors in the ℓ_2-norm and the ℓ_∞-norm are shown for the Crank–Nicolson method applied to the original Fokker–Planck equation, $\|\hat{F}_{2N} - \hat{F}_N\|_{\ell_2}$ and $\|\hat{F}_{2N} - \hat{F}_N\|_{\ell_\infty}$, and applied to the simplified Fokker–Planck equation, $\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_2}$ and $\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_\infty}$. Points indicate the maximal error across all considered $5^3 = 125$ parameter combinations for $N = 2^1, 2^2, \ldots, 2^{11}$ space/time intervals. The axes are scaled double-logarithmically and errors are plotted against the average computational time for one parameter combination.
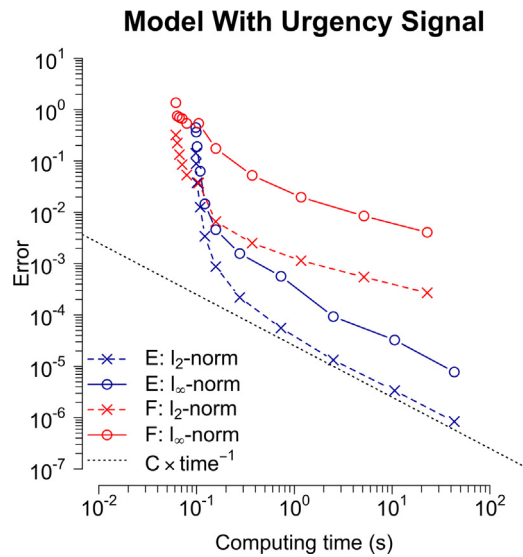

Fig. 4. Comparison of maximal errors for different approximation methods for the model with hyperbolic urgency signal (5). Maximal errors in the ℓ_2-norm and the ℓ_∞-norm are shown for the Crank–Nicolson method applied to the original Fokker–Planck equation, $\|\hat{F}_{2N} - \hat{F}_N\|_{\ell_2}$ and $\|\hat{F}_{2N} - \hat{F}_N\|_{\ell_\infty}$, and applied to the simplified Fokker–Planck equation, $\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_2}$ and $\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_\infty}$. Points indicate the maximal error across all considered $5^5 = 3125$ parameter combinations for $N = 2^1, 2^2, \ldots, 2^{11}$ space/time intervals. The axes are scaled double-logarithmically and errors are plotted against the average computational time for one parameter combination.
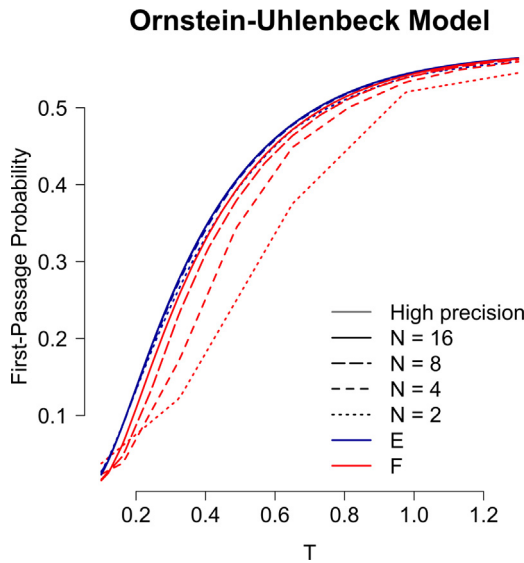

Fig. 3. Example first-passage time distribution based on different approximation methods and different numbers of space/time intervals N for the Ornstein–Uhlenbeck model (4). The first-passage time distribution function $\phi_{x_0}(T)$ is shown for $x_0 = 0.42$ and parameters $(v_0, \beta, a) = (1, 0.4, 1.2)$. The solid grey line shows a high-precision approximation with estimated maximum error tolerance $\epsilon = 10^{-7}$.
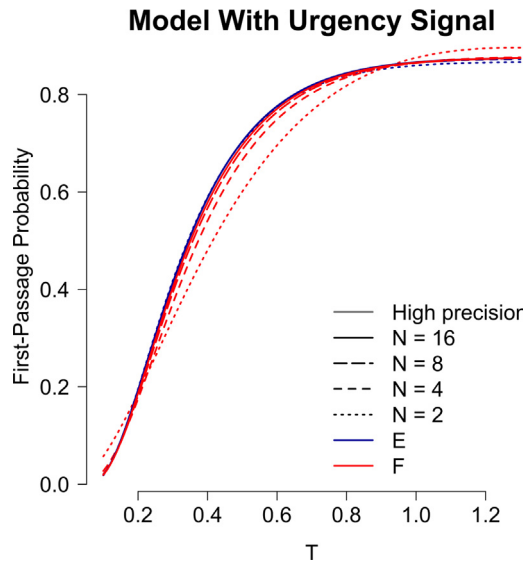

Fig. 5. Example first-passage time distribution based on different approximation methods and different numbers of space/time intervals N for the model with hyperbolic urgency signal (5). The first-passage time distribution $\phi_{x_0}(T)$ is shown for $x_0 = 0.49$ and parameters $(v_0, v_1, \tau, a) = (1.7, 1.1, 0.2, 1.4)$. The solid grey line shows a high-precision approximation with estimated maximum error tolerance $\epsilon = 10^{-7}$.

Fig. 5 shows an example of the approximation to the first-passage time distribution $\phi_{x_0}(T)$ for the absolute starting point $x_0 = 0.49$ based on $\hat{E}_N$ and $\hat{F}_N$ for different N. Similar to the Ornstein–Uhlenbeck model, the approximation $\hat{E}_N$ is visually indistinguishable from the high-precision approximation for $N > 2$. The approximation $\hat{F}_N$, on the other hand, deviates clearly from the high-precision approximation and converges slowly as N increases.

4.3. Model with collapsing boundaries (6)

Based on the parameter estimates reported in Evans et al. (2020), we consider the ranges $v_0 \in [0, 5.86]$, $a \in [0.56, 3.93]$, $T_\infty \in [3, 20]$, and $T \in [0.1, 2.5]$. In Fig. 6 we again plot the errors $\hat{F}_{2N} - \hat{F}_N \approx \hat{F} - \hat{F}_N$ and $\hat{E}_{2N} - \hat{E}_N \approx \hat{E} - \hat{E}_N$ over the computational time. Again, direct approximation of $\hat{F}$ yields a considerably worse rate than our approach, which yields the optimal rate 1.
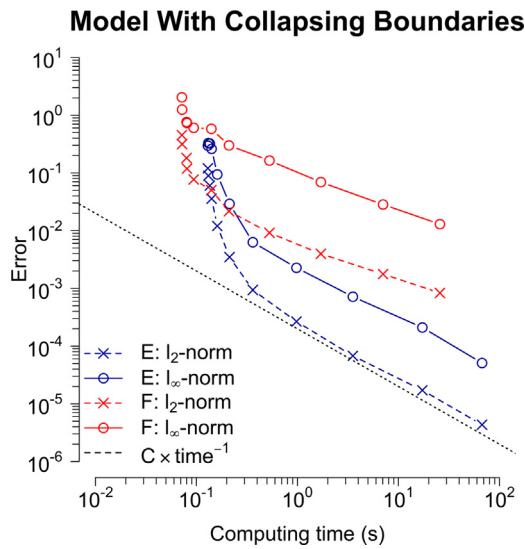
Fig. 6. Comparison of maximum errors for different approximation methods for the model with linear collapsing boundaries (6). Maximal errors in the ℓ_2-norm and the ℓ_∞-norm are shown for Crank–Nicolson method applied to the original Fokker–Planck equation, $\|\hat{F}_{2N} - \hat{F}_N\|_{\ell_2}$ and $\|\hat{F}_{2N} - \hat{F}_N\|_{\ell_\infty}$, and applied to the simplified Fokker–Planck equation, $\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_2}$ and $\|\hat{E}_{2N} - \hat{E}_N\|_{\ell_\infty}$. Points indicate the maximal error across all considered $5^4 = 625$ parameter combinations for $N = 2^1, 2^2, \ldots, 2^{11}$ space/time intervals. The axes are scaled double-logarithmically and errors are plotted against the average computational time for one parameter combination.
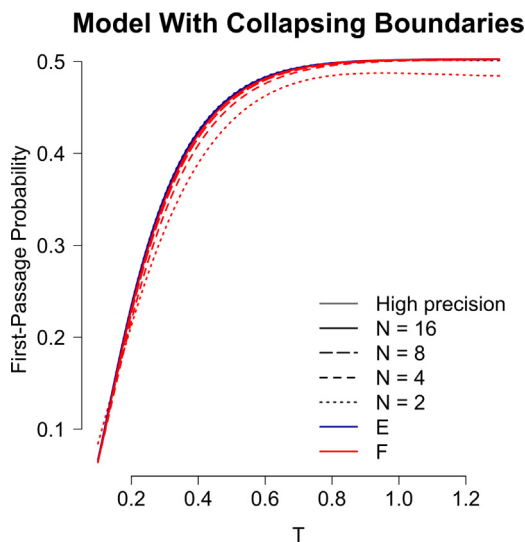

Fig. 7. Example first-passage time distribution based on different approximation methods and different numbers of space/time intervals N for the model with linear collapsing boundaries (6). The first-passage time distribution $\phi_{x_0}(T)$ is shown for $x_0 = 0.35$ and parameters $(v_0, a, T_\infty) = (0.7, 1, 4)$. The solid grey line shows a high-precision approximation with estimated maximum error tolerance $\epsilon = 10^{-7}$.

Fig. 7 shows an example of the approximation to the first-passage time distribution $\phi_{x_0}(T)$ for the absolute starting point $x_0 = 0.35$ based on $\hat{E}_N$ and $\hat{F}_N$ for different N. Similar to the previous two example models, the approximation $\hat{E}_N$ is visually indistinguishable from the high-precision approximation for all N. The approximation $\hat{F}_N$, on the other hand, deviates clearly from the high-precision approximation and converges slowly as N increases.

5. Discussion

In the present work we developed a fast numerical method for computing the first-passage time distribution function for diffusion models with space-time-dependent drift functions and time-dependent boundaries. We solve the Fokker–Planck equation in two steps. In a first step, we transform the model's domain to the unit square. In a second step, we exploit a known fast-converging series solution for models with constant boundaries and constant drift function to obtain a problem with homogeneous boundary conditions. The resulting smoother solution is then approximated numerically using the Crank–Nicolson method. Our numerical examples illustrate the high accuracy and computational efficiency of our method.

Our method extends the earlier work by Voss and Voss (2008) for the Ratcliff diffusion model to models with time-dependent boundaries and space-time-dependent drift function. Instead of applying the Crank–Nicolson scheme to the original Fokker–Planck equation, we apply the Crank–Nicolson scheme to a simplified Fokker–Planck equation that has a smoother solution, which results in a smaller approximation error.

One limitation of our method is that the Fokker–Planck equation needs to be solved repeatedly for different values of T that correspond to the observed response times. Moreover, varying the parameter values that determine the drift and the boundaries also requires repeated solution of the Fokker–Planck equation. However, since the Fokker–Planck equation can be solved independently for different values of the parameters, the computations can be fully parallelised. Our numerical results for the example models considered here suggest that an absolute accuracy of three decimals for any T can be achieved at a computing time of 1 second. Moreover, the solution to the Fokker–Planck equation can be shown to depend smoothly on the model parameters (a proof is given in Section 5 of Boehm et al. (2021)). Hence, if the first-passage time distribution function needs to be evaluated at a large number of response times and/or different parameter values, the Fokker–Planck equation might be solved for a suitably chosen subset of parameter values and the value of the distribution function can subsequently be obtained by interpolation.

A further limitation of our method is that it cannot be applied to models with non-differentiable time-dependent boundaries, as the transformation to a square domain involves the derivatives of the boundaries (see Eq. (7b)). Moreover, regardless of the transformation, non-smooth boundaries would pose challenges as they would translate to a lack of smoothness in the solution to the Fokker–Planck equation. Consequently, other methods for approximating the first-passage time distribution, even if they produce a solution in the case of non-differentiable boundaries, will likely result in large approximation errors. Fortunately, many models that are used in practice do have smooth boundaries (e.g., Bowman et al., 2012; Ditterich, 2006b; Hawkins et al., 2015).

Our method enables the fast computation of the complete distribution function that is required for many of the statistics regularly used in model fitting. The chi-squared statistic, for instance, requires evaluating the distribution function at a given set of, typically six, quantile response times (Ratcliff & Tuerlinckx, 2002). The Kolmogorov–Smirnov statistic requires evaluating the distribution function at all observed response times (Voss et al., 2004). Extension of our method to the computation of the first-passage density might also be possible for some models, which would support maximum-likelihood and Bayesian model fitting methods. The first-passage time density is the time-derivative of the first-passage time distribution and can, in principle, be obtained by differentiating the Fokker–Planck equation. For models with time-independent drift function and time-independent

boundaries the resulting PDE is a Fokker–Planck equation with a Dirac-delta function as initial condition. As mentioned before, numerical methods for these types of problems are notoriously difficult to implement due to the singular initial condition. However, a fast-converging series solution is known for simple models with constant drift function and boundaries (Gondan et al., 2014). Hence, the difference between the target density and the density of the simple model satisfies a simplified Fokker–Planck equation with homogeneous initial and boundary conditions. If the forcing function in this simplified Fokker–Planck equation is sufficiently well-behaved and if a fast-converging series expansion can be found for the space-derivative of the density of the simple model, the solution to this PDE could then be approximated with high precision using a Crank–Nicolson scheme. Future work should explore this possibility further.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Truncation criterion for space-derivative of $\hat{F}_{v_0}(t, x)$

The solution to the PDE (7) with constant drift function $v(x, t) = \hat{v}_0$ and diffusion coefficient $\hat{\sigma}$ is given by the series Gondan et al. (2014)

$$\hat{F}_{\hat{v}_0}(\hat{x}, \hat{t}) = e^{-\frac{\hat{v}_0\hat{x}}{2\hat{\sigma}} - \frac{\hat{v}_0^2\hat{t}}{4\hat{\sigma}}} \sum_{k=0}^{\infty} (-1)^k \varphi\left(\frac{r_k}{\sqrt{2\hat{\sigma}\hat{t}}}\right)$$
$$\times \left(M\left(\frac{r_k - \hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) + M\left(\frac{r_k + \hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) \right).$$

To obtain a series expansion for the space-derivative $\partial_{\hat{x}}\hat{F}_{\hat{v}_0}$, we differentiate the series term-by-term and simplify

$$\partial_{\hat{x}}\hat{F}_{\hat{v}_0} = 2e^{-\frac{\hat{v}_0\hat{x}}{2\hat{\sigma}} - \frac{\hat{v}_0^2\hat{t}}{4\hat{\sigma}}} \sum_{k=0}^{\infty} (-1)^{k+1} \varphi\left(\frac{r_k}{\sqrt{2\hat{\sigma}\hat{t}}}\right)$$
$$\times \left[\frac{\hat{v}_0}{2\hat{\sigma}} M\left(\frac{r_k + (-1)^{k+1}\hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) + \frac{(-1)^k}{\sqrt{2\hat{\sigma}\hat{t}}} \right]. \quad \text{(A.1)}$$

To derive a truncation criterion for the series (A.1), we first bound the product term

$$\left| (-1)^{k+1} \frac{\hat{v}_0}{2\hat{\sigma}} M\left(\frac{r_k + (-1)^{k+1}\hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}t}}\right) - \frac{1}{\sqrt{2\hat{\sigma}t}} \right|$$
$$\leq \frac{|\hat{v}_0|}{2\hat{\sigma}}\sqrt{2\pi} + \frac{1}{\sqrt{2\hat{\sigma}\hat{t}}}. \quad \text{(A.2)}$$

This follows from an application of Markov's inequality, $1 - \Phi(y) \leq e^{\frac{y^2}{2}}$, which gives the upper bound $\sqrt{2\pi} \geq M(y)$ for the Mills ratio. It now suffices to find a truncation criterion for $\sum_{k=0}^{\infty} \varphi\left(\frac{r_k}{\sqrt{2\hat{\sigma}t}}\right)$. We split the series into odd and even terms

$$\sum_{k=0}^{\infty} \varphi\left(\frac{r_k}{\sqrt{2\hat{\sigma}t}}\right) = \sum_{k=0}^{\infty} \varphi\left(\frac{r_{2k}}{\sqrt{2\hat{\sigma}t}}\right) + \sum_{k=0}^{\infty} \varphi\left(\frac{r_{2k+1}}{\sqrt{2\hat{\sigma}t}}\right)$$

and, using that

$$e^{-\frac{(2k-\mu)^2}{2\sigma^2}} \leq \int_{k-1}^{k} e^{-\frac{(2s-\mu)^2}{2\sigma^2}} \, ds,$$

we get for $K > 0$

$$\sum_{k=K}^{\infty} \varphi\left(\frac{r_{2k}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) \leq \int_{K-1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(2s+\hat{x})^2}{4\hat{\sigma}\hat{t}}} \, ds$$
$$= \frac{\sqrt{2\hat{\sigma}\hat{t}}}{2} \left(1 - \Phi\left(\frac{2(K-1) + \hat{x}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) \right) \quad \text{(A.3)}$$

and

$$\sum_{k=K}^{\infty} \varphi\left(\frac{r_{2k+1}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) \leq \int_{K}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(2s-\hat{x})^2}{4\hat{\sigma}\hat{t}}} \, ds$$
$$= \frac{\sqrt{2\hat{\sigma}\hat{t}}}{2} \left(1 - \Phi\left(\frac{2K - \hat{x}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) \right). \quad \text{(A.4)}$$

Combining the inequalities (A.3) and (A.4) with (A.2), we have the truncation criterion for a given error tolerance ϵ:

$$\left| 2e^{-\frac{\hat{v}_0\hat{x}}{2\hat{\sigma}} - \frac{\hat{v}_0^2\hat{t}}{4\hat{\sigma}}} \sum_{k=2K}^{\infty} (-1)^{k+1} \varphi\left(\frac{r_k}{\sqrt{2\hat{\sigma}\hat{t}}}\right) \right.$$
$$\left. \times \left[\frac{\hat{v}_0}{2\hat{\sigma}} M\left(\frac{r_k + (-1)^{k+1}\hat{v}_0\hat{t}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) + \frac{(-1)^k}{\sqrt{2\hat{\sigma}\hat{t}}} \right] \right|$$
$$\leq e^{-\frac{\hat{v}_0\hat{x}}{2\hat{\sigma}} - \frac{\hat{v}_0^2\hat{t}}{2\hat{\sigma}}} \left(\frac{|\hat{v}_0|\sqrt{\pi\hat{t}}}{\hat{\sigma}} + 1 \right)$$
$$\times \left(2 - \Phi\left(\frac{2(K-1) + \hat{x}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) - \Phi\left(\frac{2K - \hat{x}}{\sqrt{2\hat{\sigma}\hat{t}}}\right) \right) \quad \text{(A.5)}$$
$$\leq \epsilon.$$

In this latter estimate of the truncation error, the largest computational costs are incurred by the approximation of the normal distribution function Φ. If the time and space domains are discretised as $t_0 < t_1 < \cdots < t_N$ and $x_0 < x_1 < \cdots < x_N$, respectively, the truncation error will be largest at $t = t_N$ and $x = x_0$. Eq. (A.5) cannot be solved for K explicitly. However, since $\Phi(z)$ tapers off quickly as $z \to \infty$, K can usually be found by computing (A.5) for a small number of values of K.

Appendix B

MATLAB code

The MATLAB code consists of three high-level functions that serve as user interface. The function `FokkerPlanck` computes an approximation to the solution $\hat{F}$ of Eq. (7). This function takes five input arguments:

- an $m \times n$ matrix `x_hat` that specifies the space values $\hat{x}$ at which the approximate solution should be evaluated,
- an $m \times n$ matrix `t_hat` that specifies the time values $\hat{t}$ at which the approximate solution should be evaluated,
- a function handle `v_hat` for the space-time-dependent transformed drift function $\hat{v}(\hat{x}, \hat{t})$,
- a positive real number `sigma_hat` that specifies the value of $\hat{\sigma}$,
- a positive real number `tol` that specifies the desired approximate error tolerance for the Crank–Nicolson approximation,
- and an optional positive integer `iter_max` that specifies the maximum number of iteration for which the Crank–Nicolson scheme should be run (default = 11).

The function `FokkerPlanck` is general in the sense that it can compute the approximation to $\hat{F}$ for any model with space-time-dependent drift function once the transformations in Section Section 3.1 have been carried out.

The function `Examples` implements the necessary transformations for the three example models discussed in the main text and approximates the solution F of Eq. (3). Its input arguments are:

- an $m \times n$ matrix `x_rel` that specifies the *relative* starting point values $x_{\text{rel}} = \frac{x - l(0)}{u(0) + l(0)}$ for which the approximate solution should be evaluated,
- an $m \times n$ matrix `t` that specifies the time values t at which the approximate solution should be evaluated,
- a positive real value `T` that specifies the boundary of the time domain T,
- a positive real value `sigma` that specifies the diffusion coefficient σ,
- a positive real number `tol` that specifies the desired approximate error tolerance for the Crank–Nicolson approximation,
- an integer `modID` taking the value

    - 1 for the Ornstein–Uhlenbeck model,
    - 2 for the model with hyperbolic urgency signal,
    - 3 for the model with linear collapsing boundaries,

- and a vector `parameters` with the parameters of the example model,

    - (v_0, β, a) for the Ornstein–Uhlenbeck model,
    - (v_0, v_1, τ, a) for the model with hyperbolic urgency signal,
    - (v_0, T_∞, a) for the model with linear collapsing boundaries.

If we want to compute the values of the first-passage time distribution for the Ornstein–Uhlenbeck model at $\vec{x}_{\text{rel}} = (0.1, 0.4)$ and $\vec{t} = (0.2, 0.35, 0.5)$ with parameters $(v_0, \beta, a) = (1, 0.4, 1.2)$, $\sigma = 1$, and error tolerance $\epsilon = 10^{-4}$, for instance, our function call is

```
[x_rel, t] = meshgrid([0.1, 0.4], [0.2, 0.35, 0.5]);
T = max(t, [], 'all');
sigma = 1;
tol = 1e-4;
modID = 1;
parameters = [1, 0.4, 1.2];
Examples(x_rel, t, T, sigma, tol, modID, parameters)
```

and we get as output

```
ans =
          0.0274    0.1718
          0.0915    0.3493
          0.1394    0.4620
```

Next, assume we want to compute the values of the first-passage time distribution for the model with the hyperbolic urgency signal at $\vec{x}_{\text{rel}} = (0.1, 0.4)$ and $\vec{t} = (0.2, 0.35, 0.5)$ with parameters $(v_0, v_1, \tau, a) = (1, 0.4, 0.7, 1.3)$, $\sigma = 1$, and error tolerance $\epsilon = 10^{-4}$. Because the drift function of this model is time-dependent, we need to solve Eq. (3) on the domain Ω_t once for each entry t of $\vec{t}$ and evaluate the approximate solution at the right-hand boundary of the time domain. Hence, our function call now is

```
[x_rel, t] = meshgrid([0.1, 0.4], [0.2, 0.35, 0.5]);
sigma = 1;
tol = 1e-4;
modID = 1;
parameters = [1, 0.4, 0.7, 1.3];
F = zeros(size(t));
```

```
for i = 1:3
        F(i,:) = Examples(x_rel(i,:), t(i,:),...
    t(i,1), sigma, tol, modID, parameters);
end
```

and we get as output

```
F =
          0.0233    0.1694
          0.0972    0.3767
          0.1601    0.5157
```

Finally, assume we want to compute the first-passage time distribution of the model with linear collapsing bounds at $\vec{x}_{\text{rel}} = (0.3, 0.8)$ and $\vec{t} = (0.2, 0.35, 0.39)$ with parameters $(v_0, T_\infty, a) = (1.2, 1, 1)$, $\sigma = 1$, and error tolerance $\epsilon = 10^{-4}$. Note that for $x_{\text{rel}} = 0.8$, the slowest response time $t = 0.39$ falls just within the upper decision boundary. Because the boundaries of this model are time-dependent, we again need to solve Eq. (3) on the domain Ω_t once for each entry t of $\vec{t}$ and evaluate the approximate solution at the right-hand boundary of the time domain. In this case our function call is

```
[x_rel, t] = meshgrid([0.3, 0.8], [0.2, 0.35, 0.39]);
sigma = 1;
tol = 1e-4;
modID = 3;
parameters = [1.2, 1, 1];
F = zeros(size(t));
for i = 1:3
        F(:,i) = Examples(x_rel(:,i), t(:,i),...
    t(1,i), sigma, tol, modID, parameters);
end
```

This gives as output

```
F =
          0.3016    0.8469
          0.4706    0.9158
          0.4876    0.9224
```

References

Boehm, U., Cox, S., Gantner, G., & Stevenson, R. (2021). Efficient numerical approximation of a non-regular fokker–Planck equation associated with first-passage time distributions. ArXiv E-Prints, arXiv:2103.04839.

Boehm, U., Hawkins, G. E., Brown, S., van Rijn, H., & Wagenmakers, E.-J. (2016). Of monkeys and men: Impatience in perceptual decision-making. *Psychonomic Bulletin & Review*, *23*(3), 738–749. http://dx.doi.org/10.3758/s13423-015-0958-5.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. http://dx.doi.org/10.1037/0033-295X.113.4.700.

Bowman, N. E., Kording, K. P., & Gottfried, J. A. (2012). Temporal integration of olfactory perceptual evidence in human orbitofrontal cortex. *Neuron, 75*(5), 916–927.

Brandon, B. M., & Sederberg, P. B. (2014). A generalized, likelihood-free method for posterior estimation. *Psychonomic Bulleting & Review*, *21*(2), 227–250. http://dx.doi.org/10.3758/s13423-013-0530-0.

Buonocore, A., Giorno, V., Nobile, A. G., & Ricciardi, L. M. (1990). On the two-boundary first-crossing-time problem for diffusion processes. *Journal of Applied Probability, 27*(1), 102–114.

Buonocore, A., Nobile, A. G., & Ricciardi, L. M. (1987). A new integral equation for the evaluation of first-passage-time probability densities. *Advances in Applied Probability, 19*(4), 784–800.

Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations from decision field theory. *Mathematical Social Sciences, 23*, 255–282.

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience, 11*(6), 693–702.

de Haan-Rietdijk, S., Voelkle, M. C., Keijsers, L., & Hamaker, E. L. (2017). Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data. *Frontiers in Psychology*, 8, 1849. http://dx.doi.org/10.3389/fpsyg.2017.01849.

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology*, 39(2), 197–215. http://dx.doi.org/10.1006/jmps.1995.1020.

Diederich, A., & Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *Journal of Mathematical Psychology*, 47(3), 304–322. http://dx.doi.org/10.1016/S0022-2496(03)00003-8.

Ditterich, J. (2006). Evidence for time-variant decision making. *European Journal of Neuroscience*, 24, 3628–3641.

Ditterich, J. (2006). Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks*, 19, 981–1012.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, 32(11), 3612–3628. http://dx.doi.org/10.1523/JNEUROSCI.4010-11.2012.

Evans, N. J., Trueblood, J. S., & Holmes, W. R. (2020). A parameter recovery assessment of time-variant models of decision-making. *Behavior Research Methods*, 52, 193–206.

Fengler, S. N., Govindarajan, L. N., Chen, T., & Frank, M. J. (2020). *Likelihood approximation networks (LANs) for fast inference of simulation models in cognitive neuroscience (BioRxiv)*. http://dx.doi.org/10.1101/2020.11.20.392274.

Gondan, M., Blurton, S. P., & Kesselmeier, M. (2014). Even faster and even more accurate first-passage time densitites and distributions for the Wiener diffusion model. *Journal of Mathematical Psychology*, 60, 20–22.

Hanks, T., Kiani, R., & Shadlen, M. N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *ELife*, 3, Article e02260.

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience*, 35(6), 2476–2484.

Heath, R. A. (2000). The Ornstein–Uhlenbeck model for decision time in cognitive tasks: An example of control of nonlinear network dynamics. *Psychological Research*, 63, 183–191. http://dx.doi.org/10.1007/PL00008177.

Kühn, S., Schmiedek, F., Schott, B., Ratcliff, R., Heinze, H.-J., Düzel, E., Lindenberger, U., & Lövden, M. (2011). Brain areas consistently linked to individual differences in perceptual decision-making in younger as well as older adults before and after training. *Journal of Cognitive Neuroscience*, 23(9), 2147–2158. http://dx.doi.org/10.1162/jocn.2010.21564.

Larsson, S., & Thomée, V. (2008). *Partial differential equations with numerical methods (Vol. 45)*. Berlin: Springer Science & Business Media.

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, 16(5), 798–817. http://dx.doi.org/10.3758/PBR.16.5.798.

McKoon, G., & Ratcliff, R. (1996). Separating implicit from explicit retrieval processes in perceptual identification. *Consciousness and Cognition*, 5(4), 500–511. http://dx.doi.org/10.1006/ccog.1996.0029.

Milosavljevic, M., Malmaud, J., & Huth, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making, 5*(6), 437–449.

Nidiffer, A. R., Diederich, A., Ramachandran, R., & Wallace, M. T. (2018). Multisensory perception reflects individual differences in processing temporal correlations. *Scientific Reports*, 8, 14483. http://dx.doi.org/10.1038/s41598-018-32673-y.

Øksendal, B. (1998). *Stochastic differential equations (5th ed.)*. Berlin: Springer.

Philiastides, M. G. (2006). Neural representation of task difficulty and decision making during perceptual categorization: A timing diagram. *Journal of Neuroscience*, 26(35), 8965–8975. http://dx.doi.org/10.1523/JNEUROSCI.1655-06.2006.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review*, 117(4), 1113–1143. http://dx.doi.org/10.1037/a0020311.

Radev, S., Mertens, U., Voss, A., & Köthe, U. (2020). Towards end-to-end likelihood-free inference with convolutional neural networks. *British Journal of Mathematical and Statistical Psychology*, 73, 23–43. http://dx.doi.org/10.1111/bmsp.12159.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108.

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, 9(2), 278–291. http://dx.doi.org/10.3758/BF03196283.

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, 111(1), 159–182. http://dx.doi.org/10.1038/nature13314.A.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. http://dx.doi.org/10.1162/neco.2008.12-06-420.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9(3), 438–481. http://dx.doi.org/10.3758/BF03196302.

Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *ELife*, 9, Article e56938. http://dx.doi.org/10.7554/eLife.56938.

Smith, P. L. (1995). Psychophysically principled models of visual simple reaction time. *Psychological Review, 102*(3), 567–593.

Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology, 44*, 408–463.

Smith, P. L. (2010). From Poisson shot noise to the integrated Ornstein–Uhlenbeck process: Neurally principled models of information accumulation in decision-making and response time. *Journal of Mathematical Psychology, 54*, 266–283.

Smith, P. L., Ratcliff, R., & Sewell, D. K. (2014). Modeling perceptual discrimination in dynamic noise: Time-changed diffusion and release from inhibition. *Journal of Mathematical Psychology*, 59(1), 95–113. http://dx.doi.org/10.1016/j.jmp.2013.05.007.

Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision Research, 44*, 1297–1320.

Starns, J. J. (2014). Using response time modeling to distinguish memory and decision processes in recognition and source tasks. *Memory and Cognition*, 42(8), 1357–1372. http://dx.doi.org/10.3758/s13421-014-0432-z.

Turner, B. M., & Van Zandt, T. (2018). Approximating Bayesian inference through model simulation. *Trends in Cognitive Science*, 22(9), 826–840. http://dx.doi.org/10.1016/j.tics.2018.06.003.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32*(7), 1206–1220.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52*(52), 1–9.

Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, 58, 140–159. http://dx.doi.org/10.1016/j.jml.2007.04.006.

White, C. N., Kapucu, A., Bruno, D., Rotello, C. M., & Ratcliff, R. (2014). Memory bias for negative emotional words in recognition memory is driven by effects of category membership. *Cognition & Emotion*, 28(5), 867–880. http://dx.doi.org/10.1080/02699931.2013.858028.

Wood, S. N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, 446, 1102–1104. http://dx.doi.org/10.1038/nature0931.

Yap, M. J., Sibley, D. E., Balota, D. a., Ratcliff, R., & Rueckl, J. (2015). Responding to nonwords in the lexical decision task: Insights from the english lexicon project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 597–613. http://dx.doi.org/10.1037/xlm0000064.