

This is a repository copy of *FInCH: Fast statistical inference for continuous-time animal movement*.

White Rose Research Online URL for this paper: <u>https://eprints.whiterose.ac.uk/225922/</u>

Version: Published Version

Article:

Grainger, D. and Blackwell, P.G. orcid.org/0000-0002-3141-4914 (2025) FInCH: Fast statistical inference for continuous-time animal movement. Methods in Ecology and Evolution. ISSN 2041-210X

https://doi.org/10.1111/2041-210X.70052

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk https://eprints.whiterose.ac.uk/ DOI: 10.1111/2041-210X.70052

RESEARCH ARTICLE

FInCH: Fast statistical inference for continuous-time animal movement

Dominic Grainger 💿 📔 Paul G. Blackwell 💿

School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, UK

Correspondence Paul G. Blackwell Email: p.blackwell@sheffield.ac.uk

Funding information EPSRC Grant/Award Number EP/T517835/1

Handling Editor: Timo Adam

Abstract

- 1. It is common for movement ecologists to model individual-level animal movement in discrete time using methods such as hidden Markov models (HMMs). Although often the fitting of HMMs is computationally efficient, the key assumptions required to model in discrete time become limiting when dealing with temporally irregular data or an animal that changes behaviour frequently, or when comparing separate analyses on different timescales.
- 2. Continuous-time models of animal movement, which can be formulated in a scale-invariant way, avoid these complications but typically lack computational efficiency. Most continuous-time methods only allow for inference in a Bayesian Markov chain Monte Carlo (MCMC) framework, sampling from a parameter space of high dimensionality, which has rendered them inaccessible to biologists, inhibiting their uptake.
- 3. In this work, we seek to address this inaccessibility by rigorously approximating existing inference methods for a class of spatially homogeneous continuous-time models. We have developed a methodology that involves limiting the number of switches in behavioural state and then integrating out the times of those switches, using a combination of analytical and numerical methods, known as the fast integrated continuous-time HMM (FInCH) approach. Our method allows for rapid evaluation of the likelihood, permitting direct maximisation of the likelihood or the posterior density, or the use of off-the-shelf fixed-dimension MCMC.
- 4. We demonstrate this approach using a range of simulated and real data, showing that the FInCH approach competes with its discrete-time counterparts in terms of efficiency while improving accuracy. By using spline-based interpolation of terms in the likelihood, the method extends to large datasets while remaining competitive. We include examples with up to 100,000 observations.

KEYWORDS

Brownian motion, diffusion, hidden Markov models, maximum likelihood, optimisation

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). Methods in Ecology and Evolution published by John Wiley & Sons Ltd on behalf of British Ecological Society.

1 | INTRODUCTION

By modelling the movement of animals, we may learn about how inter-species, intra-species and environmental interaction affects their behaviour. This is crucial for garnering ecological understanding and assessing the ecological impact of invasive species and anthropogenic environmental change. In recent years, methods for telemetric data collection have drastically improved, and so-phisticated methods for statistical analysis have been introduced (Patterson et al., 2017). However, the extensive analysis of large data (in particular, data with temporal irregularity) remains a logistical and computational challenge, partly due to a lack of accessible model formulation and software.

It is common for biologists to model individual movement as a discrete-time process, typically a hidden Markov model (HMM). An HMM is a hidden discrete-time Markov process defined by a finite number of discrete states and the animal's location (the state-dependent process). As Blackwell (1997) explains, many behavioural factors may affect the movement patterns of an animal, such as whether the animal is encamped or foraging; Blackwell (1997) introduces a discrete behavioural state in a continuous-time setting, while Morales et al. (2004) apply it in discrete time. Patterson et al. (2017) give the example that if an animal is encamped, we expect it to move relatively short distances but change direction more often; if it is foraging, the opposite is true. Accordingly, we refer to the discrete states of the HMM as the 'behavioural states' of the animal. However, this discrete-time approach is problematic when faced with separate analyses on different timescales or irregular data and loses information when handling missing observations from otherwise regular data. These pitfalls of discrete-time modelling are covered in detail by Harris and Blackwell (2013). Continuous-time models of individual animal movement avoid these issues by formulating behavioural changes independently of the timing of observations. Indeed, Parton and Blackwell (2017) argue that continuous-time approaches are the 'gold standard' of animal movement models. However, movement ecologists have been slow to adopt these methods due mainly to their perceived complexity, the computational cost of algorithms for inference in continuous time and the widespread use of accessible R packages for inference in discrete time such as moveHMM (Michelot et al., 2016). We give an overview of these existing methods for inference in Section 2.

We seek to improve the accessibility of continuous-time models by rigorously approximating the fully Bayesian but computationally slow 'exact approach' given by Blackwell et al. (2015) for a class of switching-diffusion models, improving the efficiency of this method in an MCMC framework and allowing for direct optimisation. To do this, we limit the number of behavioural changes per interval before integrating the potential times of these changes out of the movement density term in the likelihood. We also build on the work of Blackwell (2020) by framing the exact approach as a temporally inhomogeneous HMM, using the forward algorithm to indirectly sum over visited behavioural states (Baum et al., 1970). In doing so, we develop the fast integrated continuous-time HMM (FInCH) approach. We detail the FInCH likelihood formulation in Section 3.

In Section 4, we demonstrate the effectiveness of direct optimisation with the FInCH approach compared to pre-existing approximations of the exact inference formulated in discrete time. In Section 4.1, we compare the FInCH approach with the HMM for regular data, showing that the FInCH approach generally performs better at estimating parameter values than the HMM, especially for small or sparse data sets. In Section 4.2, we show that the FInCH approach performs better on real, temporally irregular kinkajou (Potos flavus) movement data in comparison to a non-directional multiple imputation (MI) technique similar to that of McClintock (2017). In Section 4.3, we illustrate a potential workflow when using FInCH on a real, large, temporally irregular red fox (Vulpes vulpes) data set consisting of 5863 observations. In Section 4.4, we demonstrate the efficiency of FInCH for large regular data sets (up to 100,000 data points), made computationally feasible by spline interpolation of likelihood terms based on a user-defined portion of the data.

2 | EXISTING METHODS FOR INFERENCE

In this section, we give an overview of some of the key existing methods of inference on individual animal movement data. We describe methods in both continuous and discrete time, detailing the model formulation, strengths and weaknesses for each. These methods represent the animal's behaviour as an unobserved discrete-space Markov chain, whose state determines the animal's movement process, using continuous- and discrete-time Markov chains, respectively. Behavioural-switching models formulated in discrete time may, therefore, be considered an approximation of their continuoustime counterparts.

2.1 | Inference in discrete time: The HMM

By far the most common method for inference of individual animal movement data in discrete time is the hidden Markov model (HMM). In this approach, the animal's movement is typically formulated in terms of the spatial step length of the animal and its turning angle between successive observations (often referred to as 'steps and turns'). For example, if an animal is in a 'foraging' state, we may expect to see it turn frequently over short distances, but in a 'migratory' state, we may expect it to move long distances, in a more directed way.

Assume that we have *T* temporally regular observations of an animal which moves according to one of *N* behavioural states, such that S_t is the state that influences the animal's movement between observations at times *t* and *t* + 1. (For ease of exposition, in this section we denote all variables as if they were defined at intervals of length 1.) Let the observed movement process of this animal at time *t* be given by $\{X_t\}_{t=1}^T$ with realisations $x_t = (l_t, \phi_t)$, where l_t is the step length of the animal between times *t* and *t* + 1 and ϕ_t is the turning angle of the animal during the intervals [t - 1, t] and [t, t + 1]. The HMMs fitted in

3



FIGURE 1 The state dependence structure of an HMM. The latent variables, $\{S_t\}_{t=1}^{T}$, represent the (hidden) behavioural state of an animal at each observation time 1, ..., T. The state-dependent process, $\{X_t\}_{t=1}^{T}$, is the observed movement of the animal, often given as a bivariate series of step lengths and turning angles.

Sections 4.1 and 4.2 assume that this turning angle is uniformly distributed. Figure 1 illustrates the structure and dependencies within the HMM.

2.1.1 | The forward algorithm

The probability of movement between states at observation times is summarised by the transition probability matrix (TPM) Γ , an $N \times N$ matrix, such that $\Gamma = \{\gamma_{ij}\}$ where $\gamma_{ij} = \Pr(S_t = j | S_{t-1} = i)$, for i, j = 1, ..., N. Let the movement process of the animal at time t, conditional on state j, be given by $f(\mathbf{x}_t | S_t = j)$, where f denotes either a density or probability mass function. Given data $\mathbf{x} = \{\mathbf{x}_t\}$, it may be shown that the likelihood for the HMM, as constructed above, is given by

$$\mathscr{L}(\theta | \mathbf{x}) = \sum_{s_1 = 1}^{N} \dots \sum_{s_T = 1}^{N} \Pr(S_1 = s_1) \prod_{t=1}^{T} f(\mathbf{x}_t | S_t = s_t) \prod_{t=2}^{T} \Pr(S_t = s_t | S_{t-1} = s_{t-1}),$$
(1)

which has N^{T} summands—infeasible even for moderate amounts of data (Patterson et al., 2017). However, the forward algorithm allows for a much more efficient likelihood calculation (Baum et al., 1970). This recursive scheme is the reason for the tractability and widespread use of the HMM.

First, define the forward probability of state *j* at time *t* to be $\alpha_t(j) = f(\mathbf{x}_1, \ldots, \mathbf{x}_t, S_t = j)$, summarised for every state by the vector $\alpha_t = (\alpha_t(1), \ldots, \alpha_t(N))$. Note that $\alpha_t(j)$ contains information on the probability of state *j* being active at time *t* (conditional on all observations up to time *t*) and the likelihood of all observations up to time *t*; we may write the global likelihood as

$$\mathscr{L}(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_{i=1}^{N} \alpha_{\mathsf{T}}(i) = \boldsymbol{\alpha}_{\mathsf{T}} \mathbf{1}^{\mathsf{T}}$$

where $\mathbf{1} \in \mathbb{R}^{N}$ is a row vector of ones. By exploiting the Markov property, we may write the forward probability of state *j* at time *t* in terms of the previous state at time t - 1,

EXAMPLE Methods in Ecology and Evolution

$$\alpha_t(j) = \sum_{i=1}^{N} \alpha_{t-1}(i) \Pr \left(S_t = j | S_{t-1} = i \right) f(\mathbf{x}_t | S_t = j),$$

or, in matrix notation,

$$\boldsymbol{\alpha}_{t} = \boldsymbol{\alpha}_{t-1} \boldsymbol{\Gamma} \boldsymbol{\mathsf{Q}} \left(\boldsymbol{x}_{t} \right)$$

where $\mathbf{Q}(\mathbf{x}_t) = \text{diag}(f(\mathbf{x}_t | S_t = 1), \dots, f(\mathbf{x}_t | S_t = N))$. The forward algorithm traverses the time series, updating $\boldsymbol{\alpha}_t$ at every observation time, so that we may recursively find the likelihood as

$$\mathscr{L}(\boldsymbol{\theta} | \boldsymbol{x}) = \boldsymbol{\delta} \boldsymbol{Q}(\boldsymbol{x}_1) \boldsymbol{\Gamma} \dots \boldsymbol{\Gamma} \boldsymbol{Q}(\boldsymbol{x}_T) \boldsymbol{1}^{\mathsf{T}},$$
(2)

where $\delta = (\Pr(S_1 = 1), ..., \Pr(S_1 = N))$. This formulation of the likelihood is exceptionally efficient, meaning that direct maximum likelihood estimation is feasible and straightforward in most cases (see, for example, Zucchini et al. (2016) and Patterson et al. (2017)).

2.1.2 | The hidden Markov model in practice

The hidden Markov model is useful and tractable as it separates the movement and behavioural processes of the animal. Due to the forward algorithm described in the previous section, HMMs are exceptionally computationally efficient. This availability of fast direct maximum likelihood estimation has led to their widespread usage in statistical ecology. For example, HMMs have been used over the years to study caribou (Dedeban et al., 2023; Franke et al., 2004), woodpeckers (McKellar et al., 2015), beluga whales (Storrie et al., 2023), petrels (Zhang et al., 2019) and elephants (Taylor et al., 2019). The uptake of this methodology is enabled further by accessible R packages such as moveHMM (Michelot et al., 2016).

The HMM described in this section implicitly treats the underlying state as being constant between observations, so that the observation process at some time t-actually a summary based on either two or three successive locations-depends only on the state at time t. Thus the state, for example the behaviour, is effectively defined only at observation times (Langrock et al., 2012). Given that we formulate the observation process in terms of an animal's step length, these observations are assumed to be regularly spaced in time. This is limiting in the face of temporally irregular or missing data, an animal that changes behaviour frequently (for example, see Figure 4), or separate analyses on different timescales due to issues with interpretability (demonstrated in Section 4.5). A small number of missing observations on an otherwise regular grid may be dealt with by replacing the corresponding likelihood matrix **Q** with the identity matrix (Langrock et al., 2012); this correctly allows for state transition probabilities over time intervals that are a multiple of the usual interval between observations, but neglects the information represented by the locations after the missing observations given those before them. Alternatively, missing data can be handled by treating the data segments before and after as separate time series (Conners et al., 2021); this neglects

Methods in Ecology and Evolution

the state transition information over the longer intervals as well as the movement information. In summary, these approaches, while straightforward to implement, do not permit the full use of the information from observations at intervals other than a single fixed length. More generally, irregular data may be dealt with by taking the multiple imputation (MI) approach of McClintock (2017). When taking the MI approach, the animal movement path is resampled at regular intervals. This method is popular with ecologists due to its inclusion in the R package momentuHMM (McClintock & Michelot, 2018)—we use an approach similar to this in comparison to the FInCH approach in Section 4.2. All these approaches assume that the missingness of intended observations is not in itself informative; Chassan and Concordet (2023) explore the more general case.

In theory, issues with interpretability (but not loss of information) may be dealt with by framing the HMM as a continuous-time model, inferring transition rates (detailed in Section 2.2.1) directly as well as transition probabilities. This step is very rarely done in practice—it does not form part of the workflow when modelling in discrete time in either moveHMM or momentuHMM. Alternatively, we may find transition rates by taking the matrix logarithm of the TPM (inverting Equation 4). Such a step potentially suffers from issues of existence and uniqueness—see Section 5 for further discussion. Here, we do not pursue this approach further.

In this paper, we wish to make the case that, even when the irregularity of the data is negligible, it is better to avoid any loss of information or adverse effect on the inference by using scale-invariant methods for inference.

2.2 | Inference in continuous time

Many models for inference in continuous time are similar to their discrete-time counterparts, with animal movement dependent on a (hidden) underlying state process. However, we formulate behavioural changes in continuous time independently of observation time, representing an animal's behaviour as a continuous-time Markov chain. An animal's movement is typically represented by a switching-diffusion process (Blackwell, 1997, 2003), although recent efforts have been made to incorporate directional persistence akin to the 'step-and-turn' HMM (Michelot & Blackwell, 2019; Parton & Blackwell, 2017). In the switching-diffusion case, an animal's movement process is most often represented by a Brownian motion or Ornstein-Uhlenbeck (OU) process (Uhlenbeck & Ornstein, 1930). The parameters of these movement processes may then change according to the animal's behaviour at any instant (Blackwell, 2003). Consequently, these continuous-time models violate the 'snapshot' property, which would require that each observation depends only on the underlying state at single instant (Glennie et al., 2023). Thus, existing straightforward methods for continuous-time HMMs with the snapshot property are not applicable (Blackwell, 2020).

2.2.1 | Infinitesimal transition rates

As before, consider the case where we have *N* behavioural states and let S_t represent the behavioural state of the animal at some arbitrary time, *t*. When making inference in continuous time, we allow for behavioural changes to be formulated independently of observations by modelling the infinitesimal transition rate between behavioural states, rather than transition probability at observation times as described in Section 2.1. The rate at which the animal changes behaviours is given by the generator matrix $\mathbf{G} = \{g_{ij}\}$, where $g_{ij} \ge 0$ represents the infinitesimal transition rate from state *i* to state *j*, with $i \neq j$, for i, j = 1, ..., N, such that

G =	(-g ₁	g ₁₂		g _{1N}
	g ₂₁	-g ₂		g _{2N}
	:	÷	·.	÷
	g _{N1}	g _{2N}		-g _N)

where $g_i \ge 0$ represents the transition rate out of each state, *i*. The generator matrix **G** determines the latent behavioural process; if an animal is in state *i*, then it remains so for holding time τ_i , where

$$\tau_i \sim \operatorname{Exp}(g_i).$$
 (3)

The probability of an animal moving from state i to state j given some increment in time δt may be written as

$$\Pr\left(S_{t+\delta t}=j|S_t=i\right) = \begin{cases} g_{ij}\delta t + o(\delta t) & \text{for } i \neq j, \\ 1-g_i\delta t + o(\delta t) & \text{otherwise,} \end{cases}$$

where $o(\delta t)$ denotes a term such that

$$\lim_{\delta t \to 0} \frac{o(\delta t)}{\delta t} = 0$$

Hence we may approximate the probability of moving from state i to state j by

$$\Pr\left(S_{t+\delta t}=j|S_t=i\right) \approx \begin{cases} g_{ij}\delta t & \text{for } i\neq j,\\ 1-g_i\delta t & \text{otherwise} \end{cases}$$

when the time increment, δt , is small (Graham et al., 1989). It may be shown that the entries of $\Gamma(\delta t) = \{ \Pr(S_{t+\delta t} = j | S_t = i) \}$ are found by

$$\Gamma(\delta t) = \exp(\mathbf{G}\delta t),\tag{4}$$

which is used in the FInCH likelihood formulation in Equation (17). An explanation of this result is given in Supporting Information S.1.

Finally, we may incorporate spatial or environmental information into this transition rate to account for the way in which an animal's behaviour is affected by its surroundings. Let $\mathbf{x}_t = (\mathbf{x}_t, \mathbf{y}_t)$ be the location of an animal at time *t*. Perhaps the simplest spatially inhomogeneous model accounts for instances where an animal tends to only exhibit each behaviour in a given spatial region. In this case, the infinitesimal transition rate of the animal from state i to state j is given by

$$g_{ij}(\mathbf{x}_t) = \begin{cases} g_{ij} & \text{if } \mathbf{x}_t \in A_j, \\ 0 & \text{otherwise,} \end{cases}$$
(5)

for some constants $g_{ij} \ge 0$ and a partition of the space $\{A_j, j = 1, ..., N\}$. This example, where the transition rate between behaviours is dependent upon the region of the habitat an animal is in, is known as the 'adaptive' case. Examples of this model may be found in Harris and Blackwell (2013), Blackwell et al. (2015) and Alkhezi (2019). More generally, assume we have data for *I* environmental covariate values at time *t*, $\mathbf{z}_t = \{\mathbf{z}_{1t}, \ldots, \mathbf{z}_{lt}\}$. We may write the infinitesimal transition rate from state *i* to state *j* in (for example) the linear form

$$g_{ij}(\mathbf{z}_t) = \exp\left(\beta_0^{ij} + \sum_{m=1}^l \beta_m^{ij} \mathbf{z}_{mt}\right),\tag{6}$$

making inference on coefficient parameters $\beta_{0}^{ij}, \beta_{1}^{ij}, \dots, \beta_{l}^{ij}$

Although we do not focus on the spatially inhomogeneous case, this serves as a natural extension to the novel methods in this paper, as discussed in Section 5.

2.2.2 | A simple example: Observed behaviour

Throughout this paper, we assume that the underlying state process is completely unobserved. However, for this short example, we assume the opposite. Guttorp (1995) parameterises the behavioural state switching process described in Section 2.2.1 in terms of transition rates out of each state, $\lambda = \{g_i\}$, and jump probabilities $P = \{q_{ij}\} = \{g_{ij}/g_i\}$ from each state *i* to each state *j* with $i \neq j$. These are the model parameters we seek to make inference on, which we group by $\theta = (\lambda, P)$. Given an observed underlying state-changing process, we may formulate the likelihood rather straightforwardly in terms of sufficient statistics. Let the number of behavioural changes from each state *i* to each state *j* be given by $\mathbf{n} = \{n_{ij}\}$. Let the time spent in state *i* on the *m*th visit from state *j* be given by $\mathbf{r} = \{\mathbf{\tau}_i\} = \{\sum_{i\neq j} \sum_{m=1}^{n_{ij}} \tau_{ijm}\}$. The likelihood for λ and P is given by

$$\mathscr{L}(\boldsymbol{\theta}|\boldsymbol{\tau}, \mathbf{n}) = \exp\left(-\sum_{i} g_{i}\boldsymbol{\tau}_{i}\right) \prod_{i \neq j} \left(q_{ij}g_{i}\right)^{n_{ij}}.$$
(7)

The review of Patterson et al. (2017) gives a general overview of both model switching in continuous time and simulations of this continuous time model switching process. Carrying out inference in continuous time, allowing for changes in behaviour between observations, has obvious advantages with regard to model accuracy. However, we often do not observe an animal's behaviour directly and therefore treat it as hidden (e.g. see Section 2.1). Inference in this case is much more

Methods in Ecology and Evolution

involved than maximising the likelihood given in Equation (7). In this paper, we focus on and extend a method for inference in continuous time, in which an animal's behaviour is unobserved (or observed only sporadically), introduced by Blackwell (2003) and further developed by Blackwell (2020), exploiting ideas from Blackwell et al. (2015). We call this method the 'exact approach', which allows for exact, fully Bayesian inference.

2.2.3 | Exact inference: Unobserved behaviour

Statistical inference in continuous time is particularly challenging if the underlying behavioural process of the animal is unobserved. This is because the conditional distribution of the animal's position, given a previous position, is dependent on the complete behavioural process between these two times. This problem is naturally bypassed when taking a discrete-time approximation, where behavioural changes only occur at observation times. Models for exact inference completely reconstruct the animal's behaviour, including all states and the times of the state changes, jointly with the estimation of the model's parameters. All of this is done by sampling in a Metropolis-within-Gibbs MCMC algorithm-implicitly a Reversible Jump algorithm (Green, 1995), to accommodate the unknown number of changes-giving a joint posterior distribution for the behavioural states and the parameters. The dimension of the space to be explored by the MCMC is variable but necessarily high, meaning that the algorithm is computationally expensive.

Blackwell (1997, 2003) successfully carries out inference in continuous time but assumes that the location of an animal has no bearing on its behaviour. Harris and Blackwell (2013) describe extensions to the aforementioned models that allow for spatial covariates, whereby the movement and behaviour of the animal may depend on the discrete spatial region in which the animal is located in any given instant. Finally, Blackwell et al. (2015) give a method for exact, fully Bayesian inference, with the potential for both continuous and discrete spatial covariates to be accounted for explicitly. It is this method, the 'exact approach', which we build upon in this paper.

The 'exact approach' of Blackwell et al. (2015) allows for the inclusion of environmental covariates with regard to switching probability by uniformising the underlying Markov process via the incorporation of the 'potential switch'–setting a uniform rate of switches, and allowing switches from a behavioural state to itself (Blackwell et al., 2015; Jensen, 1953). At time t, for spatial covariates z_t at location x_t , let $g_{ij}(z_t)$ be the infinitesimal transition rate from state i to state j and let $g_i(z_t)$ be the total transition rate out of state i. We take some (fairly arbitrary) uniform upper bound on transition rates between states,

$$x \geq \max_{i,x,t} g_i(\mathbf{z}_t),$$

and allow 'potential' changes in behaviour to occur at this rate. When a potential switch occurs, the probability of an actual switch out of state *i* is given by

$$p_i(\mathbf{z}_t) = g_i(\mathbf{z}_t) / \kappa,$$

and the probability of an actual switch from state *i* to state *j* is given by

$$p_{ij}(\mathbf{z}_t) = g_{ij}(\mathbf{z}_t) / \kappa.$$

In the spatially homogeneous case, this framework remains the same but with constant infinitesimal transition rates, such that $\kappa \ge \max_i \{g_i\}$. This method allows for exact, fully Bayesian inference within an MCMC framework, which accounts for the environmental covariates that affect the transition rates to and from each state. However, because we have to consider the number and time of potential switches between observations, this involves a parameter space of very high dimension, which is computationally expensive due to poor mixing.

One method to improve the efficiency of this approach is the 'Integrated Continuous-time HMM' (InCH), introduced by Blackwell (2020), which frames the exact inference as a spatially inhomogeneous HMM, defined at potential switching times. The InCH approach utilises the forward algorithm and indirectly sums over behavioural states in the likelihood formulation (meaning they no longer need to be sampled), reducing the dimensionality of the parameter space and improving mixing substantially.

2.2.4 | Continuous-time inference in practice

The aforementioned continuous-time methods allow for exact inference and are, therefore, more accurate than the approximate formulation described in Section 2.1. However, inference is possible only by constructing potential changes in behaviour between observations, which in practice requires an MCMC framework. Thus, these methods are rendered somewhat inaccessible compared to discrete-time methods, which allow for rapid direct optimisation.

2.3 | Bounding the numbers of switches

We can rigorously approximate the InCH approach by assuming that there may be no more than two changes in behaviour per interval between observations. The limit of two switches is enough to include intervals in which there are one or zero switches, which we assume to be the majority of intervals in any realistic case, and cases in which there are two switches, that is a 'visit' to a behavioural state which starts and ends between a given pair of observations. Thus, it covers the vast majority of cases, and allowing for the effects of such a short visit to a state captures much of the uncertainty and variability in the details of more complex realisations. This approximation has been shown to be very close in the case with only one switch between observations (Alkhezi, 2019), but in itself does not greatly accelerate the fitting process. However, limiting the number of changes in this way is the starting point for the new methods presented in Section 3.

3 | FINCH LIKELIHOOD FORMULATION & IMPLEMENTATION

Limiting the number of switches to at most 1 or 2 in each interval between observations creates the opportunity for the times of potential changes in behaviour to be 'integrated out', rather than treated as variables to be sampled or imputed. By removing the number and time of potential switches from the parameter space, we can greatly reduce its dimensionality. This allows for direct likelihood maximisation or off-the-shelf MCMC sampling of just a few parameters, a key step towards enabling continuous-time models to compete with their discrete-time counterparts in terms of efficiency and accessibility.

As with the InCH approach, we frame the exact inference as a temporally inhomogeneous HMM defined at potential behaviouralswitching times. We calculate a likelihood locally, conditional on the behavioural state at the beginning and end of each interval, in order to utilise the forward algorithm. Consider successive two observations at times t_1 and t_2 and locations $\mathbf{x}_{t_1} = (\mathbf{x}_{t_1}, \mathbf{y}_{t_1})$ and $\mathbf{x}_{t_2} = (\mathbf{x}_{t_2}, \mathbf{y}_{t_2})$, respectively. Let the length of time between these observations be $\delta t = t_2 - t_1$. Let S_t represent the animal's behavioural state at time t. Here, we focus on the likelihood for parameter values based on only one interval between observations, the 'local likelihood'. Suppressing conditioning on parameter values, the local likelihood is given by some expression

$$\rho(\mathbf{x}_{t_2} | \mathbf{x}_{t_1}, S_{t_1} = i, S_{t_2} = j, \delta t),$$

for some initial state *i* and final state *j*. Crucial to the formulation of this local likelihood is the local movement density of the animal. In this regard, likelihood formulation for the FInCH approach is akin to the formulation of the likelihood for the spatially homogeneous InCH model, as described by Blackwell (2020).

However, the crux of the FInCH approach lies in the formulation of the animal movement density itself. We assume that no more than two changes in behaviour occur between observations, integrating out the number of potential switches in each case. This assumption is discussed in more detail in Section 2.3.

3.1 | An overview of location density

The location density of an animal is given by limiting the number of switches between observations to two and integrating the joint density of its location at the end of the interval and the time of potential switches over time. The overall diffusion rate is calculated as a weighted combination of diffusion rates in each visited state, as in Blackwell (2020). An expression for the overall diffusion rate is given in Equation (9). This paper considers only cases where an animal moves according to isotropic Brownian motion in each behavioural state.

Assume that an animal traverses a two-dimensional space, without loss of generality over a time interval of length δt starting at location **0** = (0, 0). Let *M* be the number of changes in animal behaviour in this interval and assume that a given animal changes behaviour M = m times, given by t_1, \ldots, t_m and visits states with diffusion rates $\sigma_1^2, \ldots, \sigma_{m+1}^2$. Then the distribution of its final location, (x, y), is given by a bivariate normal distribution,

$$\mathbf{x}, \mathbf{y} \mid \mathbf{0}, \eta \sim \mathcal{N}(\mathbf{0}, \eta \mathbf{I}), \tag{8}$$

with density $f(x, y | \mathbf{0}, \eta)$, where

$$\eta = t_1 \sigma_1^2 + \sum_{i=2}^m (t_i - t_{i-1}) \sigma_i^2 + (\delta t - t_m) \sigma_{m+1}^2.$$
(9)

Intuitively, a Brownian motion with diffusion rate σ^2 over an interval of length δt is equivalent to one with diffusion rate 1 over time $\delta t \sigma^2$ and so a mixture of processes as described here can be written as a process with diffusion rate 1 over an effective time η as calculated in Equation (9).

Let $\rho(t_1, \ldots, t_m | M = m)$ be a joint uniform distribution of *m* random variables representing the probability distribution of the time of each potential switch. The marginal movement density over this interval is then given by integrating over the time of each potential switch,

$$\rho(\mathbf{x}, \mathbf{y}|\delta t, \mathbf{M} = \mathbf{m}) = \int_{0 < t_1 < \ldots < t_m < \delta t} f(\mathbf{x}, \mathbf{y}|\mathbf{0}, \eta) \rho(t_1, \ldots, t_m | \mathbf{M} = \mathbf{m}) dt_1, \ldots, dt_m,$$
(10)

for which we suppress conditioning on visited states. For each sequence of possible visited states, we calculate analytic, semi-analytic and numeric solutions for this movement density for M = 0, M = 1 and M = 2 potential changes in state.

3.2 | Movement density solutions

In this section, we discuss density values for a given number of behavioural changes in an arbitrary interval. The methodology for movement density calculations is given in Supporting Information S.2. For the case with no potential behavioural changes in an interval of length δt , in state *i* with diffusion rate σ_i^2 , we have a FInCH movement density given by

$$\mathbf{x}, \mathbf{y} \mid t, i, M = \mathbf{0} \sim N(\mathbf{0}, \sigma_i^2 \delta t \mathbf{I}), \tag{11}$$

which follows directly from Equation (8). In the case with one potential switch between observations, such that we begin in state *i* and move to state *j*, with diffusion rate σ_i^2 , we have

$$\rho(x, y | \delta t, i, j, M = 1) = \frac{1}{2\pi \left(\sigma_i^2 - \sigma_j^2\right)} \int_{R^2/(2\delta t \sigma_i^2)}^{R^2/(2\delta t \sigma_j^2)} \frac{\exp(-z)}{z} dz, \quad (12)$$

where $R = \sqrt{x^2 + y^2}$ represents the Euclidean distance moved by the animal in the interval. The integral in Equation (12) is the exponential

Methods in Ecology and Evolution -

integral, for which fast routines exist as part of the GNU Scientific Library (Galassi et al., 2006; Hankin, 2006).

In the case with two potential switches, such that we move from state *i* to state *j*, before transitioning to state *k* with diffusion rate σ_{ν}^2 , we have

$$\rho(\mathbf{x}, \mathbf{y}|\delta t, i, j, k, M = 2) = \int_{0}^{\delta t} \left\{ \frac{1}{\pi \delta t^2 \left(\sigma_i^2 - \sigma_j^2\right)} \int_{a}^{b} \frac{\exp(-z)}{z} dz \right\} dv (13)$$

where

$$a = \frac{R^2}{2\left(\sigma_i^2 v + \sigma_k^2(\delta t - v)\right)}, b = \frac{R^2}{2\left(\sigma_j^2 v + \sigma_k^2(\delta t - v)\right)}$$

In Equation (13), the inner integral is akin to the one-switch case. We solve the outer integral numerically, using adaptive quadrature via the inbuilt integrate() function in R.

3.3 | Generating a local likelihood

To utilise the local movement densities in a local likelihood, we must find expressions for each number of switches, and all visited states, conditional on start and end states. Naturally, when M = 0 or M = 1, this is straightforward, so that Equations (11) and (12) may be used directly. Recall from Section 2.2.3 that the probability of an 'actual' switch from state *i* to state *j* is given by p_{ij} . In the two-switch case, we must account for every feasible first-visited 'intermediate' state. To do this, we write the movement density, conditional on initial state *i* and final state *k*, as

$$\rho(x, y | \delta t, i, k, M = 2) = \sum_{j=1}^{N} \rho(x, y, j | \delta t, i, k, M = 2)$$
(14)

$$= \frac{\sum_{j=1}^{N} \rho(x, y | \delta t, i, j, k, M = 2) p_{ij} p_{jk}}{\sum_{j=1}^{N} p_{ij} p_{jk}},$$
 (15)

where we sum over all possible first-visited states. This process, with M = 2 behavioural changes and N = 3 possible behavioural states, is demonstrated in Figure 2. Weighting movement densities based on the probability of each number of potential switches, we then have

$$\rho(x, y | \delta t, i, k) = \sum_{m=0}^{2} \Pr(M = m | \delta t, i, k) \rho(x, y | \delta t, i, k, M = m), \quad (16)$$

as a combined movement density conditional on start and end states, where $P(m|\delta t, i, k)$ is obtained using existing algorithms for finding the conditional probability of a number of changes of a uniformised Markov chain (Hobolth & Stone, 2009). For Equation (16) to hold, we assume that each Pr (M = m|t, i, k) term sums to 1. In line with this (due to the two-switch limit assumption), we use the approximation



FIGURE 2 The state-changing process, conditional on beginning and end states, for the case with three possible behaviours. In the upper plot, we demonstrate all potential state sequences for some initial state *i* and final state *j*. In the lower plot, we demonstrate this process as a line graph, assuming that an animal begins and ends an interval in state 3.

$$\Pr(M=2|\delta t,i,k)\approx 1-\sum_{m=0}^{1}\Pr(M=m|\delta t,i,k).$$

3.4 | A spline-based interpolation

To find the full local likelihood, we find the total probability of transition from state *i* to state *k* over an interval of length δt . This is given by $\gamma_{ik}(\delta t)$, the appropriate element of the matrix $\Gamma(\delta t) = \exp(\mathbf{G}\delta t)$, where **G** is the generator matrix of the underlying Markov process, as described in Section 2.2.1. Given the combined local movement density described in Section 3.3, we may obtain the joint density for location and final state,

$$\rho(\mathbf{x}, \mathbf{y}, \mathbf{k} | \delta \mathbf{t}, \mathbf{i}) = \gamma_{ik}(\delta \mathbf{t})\rho(\mathbf{x}, \mathbf{y} | \delta \mathbf{t}, \mathbf{i}, \mathbf{k}).$$
(17)

The process of finding the global likelihood from this expression is detailed in Section 3.5.

When finding the expression in Equation (17), we may avoid calculating the movement density values in Equations (12) and (13) for every possible combination of behaviours between all successive observations by implementing spline-based interpolation. This is done by separating the data into groups that are similarly temporally spaced, that is with similar values of δt .

For the models that we are considering, the required term $\rho(x, y, k|\delta t, i)$ (Equation 17) depends on x and y only through the distance moved $R = \sqrt{x^2 + y^2}$, so we can write it as $\rho(R, k|\delta t, i)$. In outline, having grouped together cases with similar values of δt , we can find the values of $\rho(R, k|\delta t, i)$ for given *i*, *k* by evaluating it exactly for selected values of *R* and then using spline interpolation for the remaining *Rs*.

Let the data **x** consist of *n* data points, taken at times $\mathbf{t} = \{t_1, \ldots, t_n\}$. Now let $\tilde{t}_i = t_{i+1} - t_i$ represent the temporal length of interval *i*, such that $\tilde{\mathbf{t}} = \{\tilde{t}_1, \ldots, \tilde{t}_{n-1}\}$. We group together all similarly spaced interval times according to some user-defined difference threshold (throughout this paper we use 5%). Given that we expect some variation (albeit within the threshold of 5%), it is worth noting that in this paper we use the first time interval occurrence as the time interval for all occurrences in the group. From here, we consider each group of 'unique' time steps individually.

Consider group *j*, corresponding to each interval with time step (approximately) \tilde{t}_i , containing n_i elements. First, we check that this group contains enough elements for a spline interpolation to come to effect. This, again, is user defined-in calculations throughout this paper we use 20 elements as the cut-off point, so that any group with fewer than 20 elements has Equation (17) calculated explicitly. If this group contains enough elements, for example, $n_i \ge 20$ then we consider the spatial Euclidean step lengths moved by the animal for each interval in the group. Let these step lengths be given by $\mathbf{R} = \{R_1, \dots, R_{n_i}\}$. Suppose we wish to use 10% of the data within each group to find the interpolating function. We now calculate $\rho(R, k|t, i)$ explicitly, for some initial and final state, for 10% of the intervals in this group, including cases with step lengths given by $R_{\min} = \min(\mathbf{R})$, $R_{\max} = \max(\mathbf{R})$ and $R_{\text{median}} = \text{median}(\mathbf{R})$, with the rest of the entries being a random sample. Assume we use explicit calculations for m_i entries in total, given by $\widetilde{\mathbf{R}} = \{\widetilde{R}_1, \dots, \widetilde{R}_{m_i}\}.$

Finally, for group *j*, assume we now have values for $\rho(R, k|t, i)$ given by $\rho_{\widetilde{R}_1}, \ldots, \rho_{\widetilde{R}_{m_j}}$. We may fit a cubic spline to these values, obtaining coefficients a_l, b_l, c_l, d_l for $l = 1, \ldots, m_j - 1$. Then we approximate $\rho(R, k|\delta t, i)$ for every step entry not yet calculated, $R \setminus \widetilde{R}$: for $R \in (\widetilde{R}_l, \widetilde{R}_{l+1})$,

$$\rho(R,k|\delta t,i) \approx a_l + b_l \left(R - \widetilde{R}_l\right) + c_l \left(R - \widetilde{R}_l\right)^2 + d_l \left(R - \widetilde{R}_l\right)^3, \quad (18)$$

which is far more efficient than repeatedly calculating Equation (17) explicitly. This process is repeated for every combination of beginning and end states, for each unique temporal step length. In this paper, we use the method of Forsythe et al. (1977), which is the welltested default in the R function stats::splinefun, but different variations of spline-based interpolation are available. In each of these methods, however, we expect the proportion of data required for a given level of accuracy to be somewhat inversely proportional to the size of the data set. Thus, within the FInCH approach, we base the spline interpolation on a fixed proportion of the data.

3.5 | A temporally inhomogeneous forward algorithm

The joint density in Equation (17) is used in the recursive forward algorithm (described for the HMM in Section 2.1.1) to find the full global likelihood conditional only on movement and transition rate parameters. For instance, let $\delta_i = P(S_0 = i)$ be the initial state probability of some state *i*, which may be inferred or deterministic. As of the second observation, after some time δt , where the animal is observed to be in location $x_1 = (x_1, y_1)$, we may find the joint probability of the animal's location, initial state *i*, and final state *j* by

$$\rho(\mathbf{x}_1, \mathbf{y}_1, i, k | \delta t) = \delta_i \gamma_{ik}(\delta t) \rho(\mathbf{x}_1 | \delta t, i, k).$$
(19)

The combined joint distribution of the animal's state at time *t* and the data up to the second observation is then given by summing over (or integrating out) initial state values to obtain

EXAMPLE A Methods in Ecology and Evolution

$$\rho(\mathbf{x}_1, k | \delta t) = \sum_{i=1}^{N} \rho(\mathbf{x}_1, i, k | \delta t), \qquad (20)$$

where k represents each final state. The full likelihood at this point is given by

$$\mathscr{L}^{\text{two obs}}(\boldsymbol{\theta}|\boldsymbol{x}_1) = \sum_{k=1}^{N} \rho(\boldsymbol{x}_1|k, \delta t), \qquad (21)$$

where θ represents all model parameters on which we make inference. We use the full likelihood value to normalise Equation (20), before treating this value similarly to the initial state distribution (but with state *k* as the new 'initial state') in Equation (19). This process is repeated, with each 'full' likelihood term contributing to the total, global likelihood until each observation has been accounted for.

3.6 | FInCH likelihood summary

As with the work of Blackwell (2020), the process of integrating over behavioural states in Equations (20) and (21) means that visited states do not need to be sampled. Furthermore, because we have integrated out the time until potential changes in behaviour occur, we are no longer required to sample the times or frequency of these. This method is advantageous within an MCMC framework, as it allows for a simple Metropolis-Hastings algorithm and dramatically reduces the dimensionality of the parameter space from which we sample, which we expect to improve mixing substantially. Alternatively, because we have no parameters left to sample outside of the likelihood itself, we allow for direct optimisation, permitting inference in either a Bayesian or frequentist setting.

3.7 | Implementation

Coding is in R for the sake of ease. Whereas a fully compiled language would result in quicker computation, the focus here is relative computation time. Furthermore, this allows for a conservative estimate for the absolute computation time for FInCH. In each of the examples below, all computation was carried out on the same laptop PC (2.3 GHz, 32 GB).

For direct optimisation, the R function optim() was used to maximise log-likelihood for both the FInCH approach and the HMM. In doing so, we use the method of Nelder and Mead (1965), which is robust but reasonably slow compared to other methods, as it utilises only function values. Of course, absolute speed does not matter so much here, as we only concern ourselves with the accuracy and computational efficiency of these techniques relative to one another. In the spirit of fairness with regard to comparisons in terms of efficiency, we use initial log diffusion rate values of $\log \sigma_1^2 = -3$ for the slow state and $\log \sigma_2^2 = 2$ for the fast state in all comparative cases, for all methods. When using the multiple

imputation (MI) method, which is described in Section 4.2.1, we use initial probability values gleaned from a previous arbitrary run of optim(). When comparing FInCH to MI using the real data, we use initial transition rates given by log (0.01). When comparing FInCH to the HMM, we use initial transition rates given by 0.001 or its equivalent discrete probability.

Where interval estimates are given after direct optimisation, they are based on a calculation of the Hessian matrix at the estimated values (again using optim()). They can be interpreted as confidence intervals or as credible intervals using a flat prior distribution based on a locally multivariate normal approximation to the likelihood.

The computation time for direct optimisation with each method represents the expected time for an individual run of optim(). In the simulation study in Section 4.1, this is based on the average of 100 separate pooled runs, whereas in the real data Sections 4.2 and 4.3, this is based on one. The idea here is to show that the FInCH approach is not only computationally feasible but that it competes with its discrete-time counterparts.

With regard to MCMC for exact inference, in order to obtain the shape of the posterior distribution for parameter values using InCH, for comparison in Section 4.2, we carry out 10,000 iterations of MCMC, with a flat prior distribution. To find 95% intervals for posterior mean values, we use the quantile() function to account for lack of symmetry (for an example of this asymmetry, see Figure 6).

4 | RESULTS AND COMPARISONS

In this section, we observe the effectiveness of the FInCH approach in terms of efficiency and accuracy in comparison with existing approaches in continuous and discrete time. This is done for a range of simulated and real data. Section 4.3 details a case study on a large red fox (*Vulpes vulpes*) data set which serves as an example workflow for an ecologist using the FInCH approach.

4.1 | Simulated regular data: FInCH versus HMM

The discrete-time model that is analogous to our switching diffusion is an HMM, as described in Section 2.1, with a uniform turning angle distribution. To observe the accuracy of the FInCH approach in comparison to the HMM, we fit both approaches to spatially homogeneous, temporally regular, simulated data. This simulated animal follows two behavioural states (which we call states 1 and 2) with (log) diffusion rates given by $\log \sigma_1^2 = -2$ and $\log \sigma_2^2 = 1$, respectively. The infinitesimal transition rate out of both states is given by $g_{12} = g_{21} = 0.001$, where g_{12} and g_{21} are the transition rates out of state 1 and state 2, respectively.

To compare approaches, we assess the computation time and accuracy of the FInCH and HMM approaches on data with a varying amount of observations, n (with $n \in \{100, 200, 500, 1000\}$). For each

value of *n*, we generate 100 data sets and carry out inference on each. This accounts for the effect of potential artefacts of individual data sets. In each case, we pool point estimate and uncertainty values for parameters by finding the mean and relevant quantiles of the 100 separate maximum likelihood estimates for each parameter. In every simulated data set, the time between observations is set to be 100s, which corresponds to around 0.1 changes in behaviour per interval. This was deemed to be reasonable as the results of Glennie et al. (2023) give max $(\delta t) = (\max_i (g_i))^{-1} = (0.001)^{-1} = 1000$ to be an intuitive rule of thumb for the maximum time step one would consider using an HMM for when faced with an animal moving according to these parameters.

When using the FInCH approach, we can vary the amount of data used to carry out a spline interpolation, described in Section 3.4, on the local likelihood. Generally, the more (regular) data we have, the lower the proportion of the data we use. This manually assigned trade-off between computation time and accuracy allows for computational feasibility when dealing with large data sets (see Table 4 for an example of this). The results of this simulation study-based comparison of inference techniques are shown in Table 1.

By Table 1, the FInCH approach appears to be far more accurate than its discrete-time counterpart. As the number of observations decreases, there appears to be a greater disparity between the two methods, although point estimates for the log diffusion rate in each state appear to have stabilised at around 200 observations. This is evident in Figure 3, which demonstrates how estimates for the log diffusion rates in each state vary based on the number of observations. The MLE log diffusion rate values inferred when using the HMM approach exhibit less separation between the diffusion rates in the two states than the FInCH method, although uncertainty around point estimates is comparable and decreases as the number of observations increases throughout.

The FInCH approach has the key advantage of allowing for inference on the infinitesimal transition rate rather than a behavioural change transition probability at observation times. This not only aids with point estimate accuracy but allows for more precise predictive distribution calculation. See Figure 4—the disparity between the HMM approximation and the true predictive distribution for the animal step length, over a typical gap between observations, becomes clearer with more sparse data as we compare the case with 0.1 expected behavioural switches between observations to 0.5. An HMM may intuitively be used on a data set with 0.5 expected observations in practice, based on results from Glennie et al. (2023). In comparison, the FInCH approach performs very well in both instances as it allows for intermediate behavioural changes.

Regarding computation time, the FInCH approach, although slower than its discrete-time counterpart, demonstrates computational feasibility in all cases. The aforementioned user-defined trade-off between computation time and precision (of Section 3.4) allows this method to scale very well with the data. For example, in the case with n = 200 observations, we find that the FInCH method takes 3.73s on average. In the case with n = 1000 observations, FInCH only takes 5.56s on average. This is not a significant increase

TABLE 1 Pooled point estimates and uncertainties for parameter values obtained using direct optimisation via the HMM and FInCH approach on simulated data, compared to actual values.

Approach	Observations	$\log \sigma_1^2$	$\log \sigma_2^2$	γ ₁₂	γ ₂₁	Time (s)
True values	_	-2.00	1.00	0.091	0.091	_
FInCH	1000	-2.01	0.99	0.092	0.091	5.56
		[-2.07, -1.92]	[0.90, 1.09]	[0.071, 0.13]	[0.060, 0.13]	
НММ	1000	-1.97	0.95	0.089	0.082	0.59
		[-2.04, -1.89]	[0.86, 1.04]	[0.070, 0.12]	[0.055, 0.11]	
FInCH	500	-2.00	1.00	0.089	0.093	5.53
		[-2.11, -1.88]	[0.88, 1.12]	[0.061, 0.14]	[0.050, 0.15]	
НММ	500	-1.97	0.95	0.086	0.085	0.31
		[-2.09, -1.83]	[0.83, 1.07]	[0.059, 0.14]	[0.050, 0.13]	
FInCH	200	-2.00	0.99	0.089	0.092	3.73
		[-2.27, -1.72]	[0.78, 1.18]	[0.040, 0.16]	[0.037, 0.20]	
НММ	200	-1.97	0.94	0.083	0.085	0.11
		[-2.22, -1.70]	[0.73, 1.16]	[0.039, 0.15]	[0.038, 0.16]	
FInCH	100	-1.98	0.99	0.090	0.092	2.02
		[-2.35, -1.64]	[0.71, 1.33]	[0.032, 0.25]	[0.037, 0.28]	
НММ	100	-1.94	0.94	0.079	0.089	0.063
		[-2.27, -1.57]	[0.63, 1.25]	[0.028, 0.21]	[0.032, 0.20]	

Note: In each case, we use 100 simulated data sets, each consisting of observations every 100s. To ensure that the FInCH approach is computationally feasible in all cases, we can manually alter the amount of data we use to carry out a spline interpolation to find the 'local' likelihood as a function of spatial step length. For the cases with 1000 and 500 observations, we use 5% and 10% of the data, respectively. We use 20% of the data for the datasets with 100 and 200 observations. In all cases, the FInCH approach correctly predicts the infinitesimal transition rate to and from each state to be $g_{12} = g_{21} = 0.001$ to two significant figures, whereas the HMM is unable to infer this parameter directly.

in computation time because, with more data, we see a more marked gain from the spline-based interpolation.

4.2 | A comparative case study: Kinkajou data

The kinkajou (Potos flavus) is a small mammal that exists in closedcanopy forest environments throughout central and southern America. The kinkajou is an example of a species threatened by not only habitat loss due to deforestation, but also the exotic pet trade (Harrington, 2015). Understanding animal behaviour can inform conservation efforts in the face of such threats. However, this species is nocturnal and, as an arboreal species, it spends most of its time in trees (Ford & Hoffman, 1988). This is a primary limitation for the study of this species, as it is very challenging to directly observe (Galvis et al., 2024; Sanderson & Trolle, 2005). Suppose instead we wish to make inferences on the behaviour of individual kinkajous by using GPS data, collected by tagging an individual specimen. We encounter potential further problems here due to the thickness of the forest canopy causing GPS signal loss, which may result in missing or irregular data (Wright et al., 2017). We therefore require a method of analysis that is accurate, fast and handles temporal irregularity in a straightforward manner.

To compare discrete-time and FInCH approaches on temporally irregular data, we utilise a relatively small individual kinkajou movement dataset obtained from Movebank (Kays & Hirsch, 2015; Powell et al., 2017). This dataset consists of no environmental covariates and 61 separate spatiotemporal locations, around 600s apart, with some observations missing (see Figure 5). Because we may view both FInCH and discrete-time methods as approximations of the exact inference, we assess model performance by comparing the results from both methods with those of the exact, fully Bayesian InCH approach (Blackwell, 2020). We use the closeness of each approximate method to the results found via this exact approach as a proxy for the overall accuracy of each method.

When demonstrating each model, we fit two behavioural states to the kinkajou data. We assume that the kinkajou moves according to a Brownian motion, making inference on its diffusion rate in each state and the rate at which it changes behaviour. From the estimated diffusion rates in each state, given in Figure 6, we can see that the animal moves with a far lower diffusion rate in one behavioural state than the other, which indicates that the kinkajou has both a 'slow' and a 'fast' or 'busy' behaviour. The slower behaviour may represent periods in which the kinkajou is asleep, or otherwise at rest, and the busier state may describe the movement of the kinkajou when it is foraging.

4.2.1 | Multiple imputation

Observing Figure 5, we see that the data set has 6 missing observations, which accounts for around 10% of the data set. We cannot

11



FIGURE 3 Point estimates and uncertainty for the log diffusion rate in each behavioural state, by the size of the data set, found using FInCH and HMM approaches. Estimates for the log diffusion rate of the animal in both behavioural states display less accuracy and more uncertainty as the number of observations decreases. In each case, however, the FInCH approach performs better than its discrete-time counterpart.

justify fitting an HMM directly to the data when dealing with nonnegligible temporal irregularity (see Section 2.1.2). Instead, we take a multiple imputation approach inspired by the widely used R package momentuHMM (McClintock & Michelot, 2018). This method involves resampling the movement path at regular intervals, using a single-state continuous-time model and then fitting an HMM to the resampled observations. This process of imputation and estimation is repeated multiple times in order to account for the uncertainty in the animal's path (see Equation 23). However, whereas this package uses continuous-time correlated random walks (CTCRWs), introduced by Johnson et al. (2008), to reconstruct the movement path, we use a single-state diffusion model instead. In essence, this is the same, but without accounting for directionality (which is consistent with other examples in this paper). Having resampled the movement path, we fit a two-state HMM, finding MLE and uncertainty values using pooled estimates from each iteration. This involves a step-bystep iterative process:

- 1. Fit a single-state diffusion model to the data, using optim() to glean an 'average' diffusion rate.
- Sample a path at regular observation times (in this case, 600s), using this MLE diffusion rate, utilising Brownian bridges (e.g. Bullard, 1999; Horne et al., 2007) to ensure that observed locations are visited by the animal.
- Fit a non-directional two-state HMM to this regular, augmented path.
- In total, we repeat steps 2 and 3 a total of n = 30 times, storing MLE parameter values and parameter uncertainties each

FIGURE 4 A key advantage of continuous-time methods such as the FInCH approach is the ability to infer transition rates between states, rather than the transition probability at observation times. We demonstrate this here by observing the predictive distribution for the step length of an animal between observations, for when we start and end in either the slow or fast state. Both plots are based on the results of inference on the data set consisting of 1000 data points. In the upper plot, observations are 100s apart. In the lower plot, we see the same but for the case where observations are taken 500s apart. In both cases, the predictive distribution of the HMM is far less accurate than that of the FInCH approach. However, the lower plot demonstrates that, as we increase the interval between observations, and hence the number of expected behavioural changes between observations, the HMM sees a far greater reduction in accuracy than FInCH.



time. Let θ represent any of the parameters we wish to estimate. If we let θ^i be the *i*th MLE parameter value, $var(\theta^i)$ be the variance estimate of the *i*th parameter, and $\overline{\theta}$ be the pooled mean MLE parameter value, we may find pooled point and variance estimates for parameter θ by as described by Rubin and Schenker (1986).

$$\overline{\theta} = \frac{1}{n} \sum_{i=1}^{n} \theta^{i}, \qquad (22)$$

$$\operatorname{var}(\overline{\theta}) = \frac{1}{n} \sum_{i=1}^{n} \operatorname{var}(\theta^{i}) + \left(1 + \frac{1}{n}\right) \left[\frac{1}{n-1} \sum_{i=1}^{n} \left(\theta^{i} - \overline{\theta}\right)^{2}\right], \quad (23)$$

4.2.2 | Comparing approaches

In this case study, observing Table 2, the FInCH approach appears much more accurate than its discrete-time counterpart. This is evidenced by the fact that, when using the MI approach, the 95% confidence interval for the log diffusion rate in the slow behavioural state, $\log \sigma_{11}^2$ does not include the mean posterior value found by InCH. This disparity is shown clearly in Figure 6. As with the case study on regular data in Section 4.1, we usually do not make direct inference on the infinitesimal transition rate between states, g_{21} and g_{21} , when using discrete-time methods. However, as before, we can find the transition probabilities at each observation, γ_{12} and γ_{24} , using all approaches. The



FIGURE 5 The time between successive observations in the kinkajou (*Potos flavus*) data set. Clearly, there are occasional missed observations (in one case, two missed observations in succession). Most 'regular' observations are recorded around 600s apart. 'Partly regular' temporal structure such as this is common amongst temporally irregular data sets.

FInCH approach performs far better than the MI method in this regard, too. The results from the exact inference suggest that the kinkajou is in the slow state 51.4% of the time, whereas FInCH and MI suggest that this value is 54.3% and 59.4%, respectively.

The FInCH approach took 5.46s to compute compared to 2.36s for the MI, demonstrating that the FInCH approach can undoubtedly compete with its discrete-time counterparts when used on temporally irregular data. Furthermore, as shown in Section 4.1, the FInCH approach scales fairly well with the data, whereas McClintock (2017) demonstrates that this may not always be true with the MI approach.

For direct optimisation with FInCH, we choose a potential switching rate of $\kappa = 0.003$, which corresponds to around 1.8 potential changes in behaviour between intervals. To account for the low quantity of data, we use a spline interpolation technique, which utilises 50% of similarly temporally spaced data to approximate local likelihood values (detailed in Section 3.4). The choice of 30 iterations of the MI was largely based on the results of McClintock (2017), which demonstrated that more imputations at the expense of more computational effort does not always yield more accurate results. This fact is also evident here, with 400 iterations taking 31.05 s to compute and yielding point estimates $\log \hat{\sigma}_1^2 = -0.72$, $\log \hat{\sigma}_2^2 = 1.60$, $\hat{\gamma}_{12} = 0.32$ and $\hat{\gamma}_{21} = 0.19$. This does not give a substantial improvement on diffusion rate point estimates and suggests that the kinkajou spends 62.7% of its time in the slow behaviour, in contrast with the 51.4% found in the exact inference.

When carrying out exact inference, we use the Integrated Continuous-time HMM (InCH) approach of Blackwell (2020), calculating posterior distribution parameters within an MCMC framework with flat priors. We do this for 10,000 iterations. To find 95% confidence intervals with InCH, we use the quantile() function in R in order to account for a lack of symmetry (which is evident in Figure 6).

4.3 | A big data case study: Red fox

In this section, we consider a large data set with missing data (n = 5863) concerning the GPS location of an individual red fox (*Vulpes vulpes*) around UK grasslands (Porteus et al., 2024a, 2024b). We assume that this fox moves according to Brownian motion with a diffusion rate, σ_i^2 , determined by three separate behaviours, $i \in \{1,2,3\}$. Unlike the comparative studies in Section 4.1 and Section 4.2, we demonstrate only the FInCH approach. This section serves as both an example of this approach for big data and an example workflow for an ecologist working in a FInCH framework.

We choose (fairly arbitrary) starting parameters given by $\log \sigma_{1:0}^2 = -3$, $\log \sigma_{2:0}^2 = 0$, $\log \sigma_{3:0}^3 = 2$ and $g_{ij} = 0.001$ for $i \neq j$ with $i, j \in \{1, 2, 3\}$. This transition rate value corresponds to around 0.6 changes in behaviour per 10 min interval. We interpolate based on 1% of similarly temporally spaced data, and set a very conservative potential switching rate of $\kappa = 0.2$. The time taken for inference via direct optimisation was around 20 min and involved 2978 function calls, largely due to the increased number of model parameters and MLE values $\hat{g}_{13} \approx \hat{g}_{31} \approx 0$ taking a while to converge. In contrast, all runs of optim() elsewhere in this paper converge in less than 500 iterations. Fixing $g_{13} = g_{31} = 0$ ahead of running FInCH and setting $\kappa = 0.02$ yields similar MLE values for parameters in around 8 min, requiring 934 iterations. Full MLE parameter values and uncertainties are given in Table 3. Diffusion rate point estimates are given by $\log \hat{\sigma}_1^2 \approx -3.78$, $\log \hat{\sigma}_2^2 \approx -1.30$ and $\log \hat{\sigma}_3^2 \approx 3.52$. We therefore have a very slow 'resting' state 1, an active 'running' state 3 and a moderately active 'walking' state, 2. The estimated generator matrix, $\hat{\mathbf{G}}$, is given by

FIGURE 6 The uncertainty around point estimates for the kinkajou diffusion rate, in each state, using each method, where the HMM is fitted by multiple imputation (MI). The histogram shows the posterior distribution for each log diffusion rate obtained using InCH. Note here that by taking a direct optimisation approach, we have assumed a multivariate normal distribution for approximate (FInCH and MI) parameter values. The MLE values found through FInCH are a close estimate of the posterior mean values and an almost exact estimate of the posterior modes. In contrast, the MI approach is somewhat inaccurateespecially in the case for the slow state.



$$\widehat{\mathbf{G}} = \begin{pmatrix} -\widehat{g}_{11} & \widehat{g}_{12} & \widehat{g}_{13} \\ \widehat{g}_{21} & -\widehat{g}_{22} & \widehat{g}_{23} \\ \widehat{g}_{31} & \widehat{g}_{32} & -\widehat{g}_{33} \end{pmatrix} \approx \begin{pmatrix} -0.00025 & 0.00025 & 0 \\ 0.00049 & -0.001 & 0.00052 \\ 0 & 0.00041 & -0.00041 \end{pmatrix},$$
(24)

(24) top speed of this red fox to be arou supported by current literature (Fo which indicates that the fox never moves directly between state 1 and

Based on the point estimates for diffusion rates found in this inference, we may calculate the distribution for the animal's speed in each behaviour. For a diffusion rate σ_{i}^{2} , the Euclidean distance

state 3.

moved by the animal in a given unit time is given by a $I_i \sim \text{Rayleigh}(\sigma_i)$ (Chattamvelli & Shanmugam, 2021). This is demonstrated, along with point estimates and uncertainties for diffusion rates, in Figure 7. Based on the 95th quantile of the distribution for the fox's speed in its 'running' behaviour, $I_3 \sim \text{Rayleigh}(\exp(3.52))$, we may expect the top speed of this red fox to be around 14.29 m/s (51.4 km/h), which is supported by current literature (Fothergill & Cordey, 2015).

4.3.1 | State reconstruction

By formulating its likelihood using the temporally inhomogeneous forward algorithm, the FInCH approach indirectly sums

Approach	$\log \sigma_1^2$	$\log \sigma_2^2$	γ_{12}	γ_{21}	g ₁₂	g ₂₁	Time (s)
Exact	-2.29	1.80	0.38	0.36	0.0011	0.0011	33.82
	[-3.49, -0.95]	[1.40, 2.30]					
FInCH	-2.67	1.85	0.43	0.38	0.0015	0.0013	5.46
	[-3.64, -1.70]	[1.38, 2.32]					
MI	-0.62	1.64	0.29	0.20	-	-	2.36
	[-1.82, 0.59]	[1.13, 2.16]					

Note: For the InCH approach, we have 95% confidence intervals for parameter values given by [0.00053, 0.0016] and [0.00041, 0.0016] for g_{12} and g_{21} , respectively. The FInCH approach yield 95% confidence intervals of [0.00066, 0.0033] and [0.00032, 0.0052] for g_{12} and g_{21} , respectively. In discrete time, we estimate the transition probabilities to and from each state in the HMM case, yielding 95% confidence intervals of [0.11, 0.79] for γ_{12} and [0.058, 0.69] for γ_{21} . In the FInCH case, we obtain 95% confidence intervals of [0.26, 0.57] for γ_{12} and [0.14, 0.67] for γ_{21} . In the InCH case, we have 95% confidence intervals of [0.21, 0.49] for γ_{12} and [0.17, 0.48] for γ_{21} . This demonstrates far more uncertainty when using the discrete-time approach than the two continuous-time methods. The general high uncertainty in each case highlights the importance of making inference on the transition probabilities out of each state at observations, in each case, are based on intervals of 740005.

TABLE 3Point estimates and uncertainty values for modelparameters obtained from running the FInCH model on the foxdata.

$\log \sigma_1^2$	$\log \sigma_2^2$	$\log \sigma_3^2$
-3.78	-1.30	3.52
[-3.85, -3.70]	[-1.46, -1.15]	[3.47, 3.58]
g ₁₂	g ₁₃	g ₂₁
0.00025	≈0	0.00049
[0.00021, 0.00030]		[0.00040, 0.00060]
g ₂₃	g ₃₁	g ₃₂
0.00052	≈0	0.00041
[0.00044, 0.00062]		[0.00034, 0.00049]

over the animal's behavioural states, meaning we do not have to sample them. However, in practice an ecologist may be interested in reconstructing the animal's behavioural states to gain a more in-depth interpretation of the data and of the behaviours themselves. In a standard discrete-time HMM setting the Viterbi algorithm, introduced by Viterbi (1967), may be used to find an animal's most likely behaviour at each observation time (Zucchini et al., 2016). This is a readily available option in popular software for implementing HMMs, such as moveHMM (Michelot et al., 2016).

The Viterbi algorithm may also be used directly in a FInCH framework. The probability of the fox being in each behaviour over the first 150h of data is given in Figure 8, in which there is a clear cyclic nature to the fox's most and least active behaviours over a 24h period.

TABLE 2Point estimates, with 95%confidence intervals for parameter valuesgiven for diffusion rate values, derivedusing the three methods.

Given the results in Figure 8, an ecologist may wish to make inferences on the cyclic nature of the fox's activity. In Figure 9, we fit the probability of the fox being in each behaviour found via the Viterbi algorithm to the hour of the day using a generalised additive model (GAM). The fox exhibits a high likelihood of being at rest between 12:00 and 21:30, whereas fox activity is predicted to peak between 00:00 and 08:00. This crepuscular and nocturnal cyclic activity is supported by the findings of Díaz-Ruiz et al. (2016).

4.4 | The FInCH and the HMM for temporally regular large datasets

As described in Section 3.4, we may use a spline-based interpolation method to reduce the computation time for inference via the FInCH method, with an effect that increases based on the regularity of the data. When faced with entirely regular data, the FInCH approach scales very well with the size of the data in a way that competes with the HMM in terms of computation time. An example of this, when we fit two behavioural states to regular data sets of size n = 100,000 and n = 10,000, is given in Table 4.

4.5 | The importance of continuous-time inference

In Section 4.2, we demonstrated the importance of inferring the generator matrix, with transition rates, rather than the discrete-time transition probability matrix (TPM) for ecological interpretability, based on the potential uncertainty of TPM values. This section offers a stronger example—the generator matrix in Equation (24) yields a TPM (over 600s intervals) given by

FIGURE 7 In the upper plot, we have point estimates and uncertainty for log diffusion rates in the three behavioural states. Because this is a very large data set, we see very little uncertainty around point estimates. In the lower plot, we see the distribution for the fox's speed (in metres per second) in these behaviours.



$$\widehat{\Gamma} = \{\widehat{\gamma}_{ij}\} = \{\Pr\left(S_{t+600} = j | S_t = i\right)\} \approx \begin{pmatrix} 0.88 & 0.11 & 0.02\\ 0.21 & 0.58 & 0.21\\ 0.03 & 0.17 & 0.81 \end{pmatrix}, (25)$$

with $\hat{\gamma}_{13}$ and $\hat{\gamma}_{31}$ values accounting for a visit to state 2 in the interval between observations. However, $\hat{\Gamma}$ alone, without knowledge of transition rates, does not demonstrate that the animal may never move directly between state 1 and state 3. Thus, this key structural feature would likely have been missed by an ecologist if inference had taken place in discrete time.

A further advantage of making inference in continuous time and inferring transition rate values may be seen when reconstructing an animal's behavioural state sequence. In this example, due to the high number of data used, we viewed it appropriate to use the Viterbi algorithm to find the probability of the fox being in each state at each observation time. However, when the transition rates from and to each state are inferred, we may *sample* the behavioural sequence of the animal on any timescale. For example, Blackwell (2020) demonstrates this type of approach by using the Forward-Filtering Backward-Sampling method of Früwirth-Schantter (1994). This is especially useful with particularly sparse data or when comparing different analyses on contrasting timescales.



 24
 48
 72
 96
 120
 144

 Observation time (hours after initial observation)

FIGURE 8 We may use the Viterbi algorithm to reconstruct an animal's state sequence when using FInCH, just as we may when using the HMM. In this example, which looks at the first 150h of data, it seems as if the period in which the fox is most active (that is, in the 'running' behavioural state) is fairly consistent throughout the day.

4.6 | Choosing a method

In this paper, we have demonstrated that the FInCH approach is highly effective in comparison to its discrete-time counterparts, regarding both accuracy and the ability to compete in terms of computational expense, in a spatially homogeneous switching-diffusion setting. However, in cases with regular large data sets recorded at a high sampling frequency relative to the rate at which an animal changes behaviour, the disparity between the HMM and FInCH performance may not be too great, assuming that the HMM is formulated in a continuous-time framework. Existing R packages do not formulate the HMM in this way, however, and Table 4 demonstrates that the FInCH approach scales very well with the size of regular data sets. Thus, the FInCH approach may still be preferable even in this special, albeit fairly common, case.

The 'two-switch' FInCH approach in this paper is efficient due to the semi-analytical solutions for the location density in Section 3.2, which are obtained under the assumption of a fairly simplistic animal movement process. One advantage of the HMM is the easy incorporation of directional persistence in an animal's movement (see Section 2.1 or the discussion in Section 5 for details). The existing implementation of the FInCH methodology does not support the inclusion of velocity, so if this is expected to be an important aspect of animal behaviour in a particular study, then an ecologist may not wish to use this approach until it has been suitably extended. Viable alternatives exist that allow for exact inference, such as that of Michelot and Blackwell (2019), or for finerscale approximation, as in Parton and Blackwell (2017). However, if an ecologist wishes to carry out inference on regular data without imputation of times, locations and velocities between observations then currently an HMM approximation may be preferred, despite the limitations described above. Generally, however, we claim that, provided that there is a reasonable tolerance for computation time, the FInCH approach is a favourable alternative to the HMM whenever the data and ecological assumptions made fit within a FInCH framework.

In Section 4.2, the FInCH method is shown to be markedly faster than the exact inference, while producing similar results. This approach offers a viable, more accessible alternative to its exact counterpart, allowing for inference via direct maximisation of the likelihood, as illustrated here, or of the posterior density, if a penalty term representing the prior distribution is included. The user may still wish to utilise the exact inference if they expect many behavioural changes between observations, accept the sampling-based Metropolis-within-Gibbs mechanisms behind the InCH approach and can accommodate the computational expense. However, even in this case, we are able to simply increase the number of behavioural changes possible when taking the FInCH approach between observations to allow for relatively straightforward direct inference. As previously discussed, alternative methods exist for exact inference that allow for the incorporation of directionality in inference without

19





FIGURE 9 Upper plot: The estimated probability of the fox being in its resting state, with uncertainty, throughout a given day. Lower plot: A stacked area plot showing the probability of the fox being in either of its three behavioural states throughout a given 24 h day. These results were calculated by normalising results from a generalised additive model (GAM) which fits a cyclic cubic spline, with 7 knots, to the hour of the day. Here, we have overwhelming evidence against there being no relationship between fox activity and the hour of the day ($p < 10^{-15}$).

time-discretisation error. These methods may be preferred by an ecologist if there is a wish to account for directional persistence.

5 | DISCUSSION AND NEXT STEPS

The FInCH approach has been demonstrated to be very efficient when used on regular data due to a spline-based interpolation of local likelihood terms. As stated in Section 3.4, we expect that the proportion of data required for a given level of accuracy will be inversely proportional to the size of the data set. Furthermore, increasing the amount of data used for this interpolation will increase the accuracy of the inference at greater computational expense. In this paper, we have ensured that no fewer than 20 similarly spaced data points have been used for this interpolation, lowering the proportion of data used for larger data sets (for example, see Section 4.1).

Methods in Ecology and Evolution

TABLE 4An example of the time taken for the FInCH and HMMapproach on regular data of size 100,000 and 10,000, respectively.

Approach	Observations	Time (s)
FInCH	100,000	95.4
HMM	100,000	64.9
FInCH	10,000	16.4
HMM	10,000	9.38

Note: Two behavioural states were fitted in each case. To carry out inference in the n = 100,000 case, a spline interpolation utilising 0.1% of the data has been used. In the n = 10,000 case, we incorporate 0.5% of the data. Point estimates for parameter values inferred in these examples are given in Supporting Information S.3.

We have used a higher proportion of temporally irregular data (for example, we used 50% of similarly spaced data in the kinkajou case study in Section 4.2). Inferring the amount of data that should be used on a case-by-case basis falls beyond the remit of this paper but is undoubtedly worthy of investigation.

Throughout this paper, we have considered infinitesimal transition rates found directly when using methods formulated in continuous time. In theory, one may find the matrix logarithm of the transition probability matrix to find the generator matrix after making inference in discrete time. However, this process is not as straightforward as taking the matrix exponential—the matrix logarithm of a given TPM may not exist, and if it does exist, may not be unique. This is not merely a mathematical technicality; it reflects the fact that a given discrete-time model may not have any reasonable interpretation in continuous time, which in turn casts some doubt on the interpretation of that model. Albeit a feasible step in practice, software for inference in discrete time such as that of Michelot et al. (2016) and McClintock and Michelot (2018) tends not to include finding the infinitesimal transition rate as part of their workflow. Thus, we have not accounted for this in our analysis.

Although here we have only demonstrated up to three behavioural states, we may easily incorporate more using the same method. We expect this to result in a slight increase in computation time as we need to account for the local likelihood by considering every possible visited state, although of course any method will experience some increase in computational cost with more states. Likewise, we may readily apply this approach to data in any number of dimensions, though the specific analytic integrals (shown in Supporting Information S.2) will differ. The choice of two dimensions here reflects most animal movement data. Again, increasing the number of dimensions used in the data is liable to increase computation time as we are more likely to require adaptive quadrature to acquire an expression for the animal's movement density.

In the calculation for the local likelihood, as given previously, we take the probability of two potential switches to be given by the probability of having two or more potential switches in the exact model. In theory, this overinflates the probability of having two potential switches. However, this is compatible with the idea that the case of two switches partly acts as a proxy for more complex realisations with multiple changes in behaviour. Furthermore, we found that a variation of the likelihood approximation, such that we explicitly calculate the probability of two switches, gave only a negligible improvement in model accuracy at the expense of increased computation time.

In the homogeneous FInCH approach described here, it is perfectly reasonable to assume an upper bound exists on transition rates between behavioural states, which may serve as the 'potential' switching rate as earlier described. Allowing a 'switch' from a state to itself—that is, a potential switch that is not an actual switch—is an aspect of 'uniformisation' (Blackwell et al., 2015; Jensen, 1953). We considered removing the potential switching rate, κ , as far as possible, moving somewhat further from the InCH approach (Blackwell, 2020), by assuming that two potential changes in behaviour always occur between observations. This seemed to improve computation time, but its effect on model accuracy is less clear; further experimentation is in progress.

Although a key strength of the FInCH approach is that it allows for fast direct optimisation, with associated approximate interval estimates, the user may still wish to make fully Bayesian inference. Our method reduces the dimensionality of the space from which we need to sample, from being large and variable (Blackwell, 2003, 2020), or bounded but proportional to the number of observations (Alkhezi, 2019), to being fixed and small—just the number of parameters in the model. Using the likelihood defined here, inference within a Metropolis-Hastings framework will be straightforward and computationally feasible.

A natural next step for the FInCH approach in its current form is the incorporation of environmental covariates. For example, an ecologist may be interested in how an animal's proximity to an environmental feature (such as a river or forest) affects its probability of changing behaviour. Indeed, this is the original purpose behind the introduction of the 'potential switch', which differentiates the exact approach of Blackwell et al. (2015) from that of Blackwell (2003). To do this and maintain an upper bound on behavioural transition rates, we assume that the FInCH approach represents a 'separable' model, as described by Blackwell et al. (2015) and Harris and Blackwell (2013). In such models, we assume that movement into a new environment only directly affects an animal's transition rate between behaviours, such that its movement patterns do not directly depend on its location. Thus, we do not have to deal with instantaneous behavioural switches. As discussed in Harris and Blackwell (2013), this is not a particularly limiting assumption. Separability is also assumed in the spatially heterogeneous exact inference method of Blackwell et al. (2015) and the InCH approach of Blackwell (2020) in order to allow for the assessment of an animal's environment as a prerequisite for finding the probability of an actual change in behaviour. We may 'integrate over' an animal's environment, in effect integrating with respect to space as well as time when calculating its movement density.

Existing methods for inference in discrete time, such as the stepand-turn HMM, account for the turning angle of an animal between observations when inferring its behavioural state. For example, an animal which is migrating, therefore moving a more directed way, is likely to 'turn' less than an animal which is foraging. There are clear advantages to considering turning angle in continuous time, which would also allow for a fairer comparison to the HMM. To do this, we believe the FInCH approach may be amenable to 'continuoustime correlated random walks' (CTCRWs), introduced by Johnson et al. (2008). These allow for directionality by assigning a diffusion process to the animal's velocity rather than directly to its position. We may then integrate velocity to produce a smooth movement process (which is also more realistic over a fine timescale). Future work may serve as an extension of Michelot and Blackwell (2019), which incorporates behavioural switching into CTCRWs.

This methodology is novel in allowing for inference via direct optimisation when behavioural changes are formulated independently of observation times. However, it is not the only example of scaleinvariant inference that allows for direct likelihood maximisation. For example, recent unpublished work (Blackwell, in review) builds on earlier work of Michelot et al. (2019) to incorporate behavioural switching into the Langevin diffusion model while maintaining the interpretation of its movement parameters as describing habitat selection within each behavioural state. This model could be fitted using existing Langevin methods (Blackwell & Matthiopoulos, 2024; Michelot et al., 2019) combined with standard HMM techniques, by assuming that changes in behaviour occur only at observation times and that spatial covariates (and therefore both movement rates and transition rates between behaviours) depend only on locations at the times of observations and remain piecewise constant over time. It would be interesting, however, to see if this model framework is amenable to a FInCH-type approach.

6 | CONCLUSIONS

We have shown that the spatially homogeneous FInCH approach is more accurate than its discrete-time counterparts for a range of real and simulated data. This is especially true for irregular data or data for an animal that changes behaviour frequently. The FInCH approach is also far more accessible than pre-existing continuous-time methods for statistical inference, such as the InCH model, in that it allows for full likelihood calculation, and therefore direct maximisation, at a computational cost that competes with models formulated in discrete time. Furthermore, the spline interpolation between the likelihood terms of the FInCH approach reduces computation time, allowing for computationally feasible inference when faced with large data sets. One advantage of this method is that the amount of data used in the interpolation is user-defined in response to the size of the dataset, so in practice the computational cost scales well.

AUTHOR CONTRIBUTIONS

Paul G. Blackwell initially conceived the ideas; Dominic Grainger and Paul G. Blackwell designed the methodology; Dominic Grainger implemented the methodology, planned and carried out the investigation of its properties and the comparison with other methods and

Methods in Ecology and Evolution

led the writing of the manuscript. Both authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

Dominic Grainger was supported by an EPSRC DTP Studentship, grant reference EP/T517835/1. The authors are grateful for useful comments and questions from two anonymous reviewers.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www. webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.70052.

DATA AVAILABILITY STATEMENT

The data are available from the Movebank repository via https://doi. org/10.5441/001/1.41076dq1 (Kays & Hirsch, 2015) and https:// doi.org/10.5441/001/1.304 (Porteus et al., 2024b). The code is available from the online research data repository (ORDA) at the University of Sheffield via https://doi.org/10.15131/shef.data. 28675073 (Grainger & Blackwell, 2025).

ORCID

Dominic Grainger D https://orcid.org/0009-0000-8194-297X Paul G. Blackwell D https://orcid.org/0000-0002-3141-4914

REFERENCES

- Alkhezi, H. (2019). Bayesian inference for animal movement in continuous time. PhD thesis, University of Sheffield, UK. https://etheses.white rose.ac.uk/25274/1/thesis.pdf
- Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41(1), 164–171. https://api.semanticscholar.org/CorpusID:122568650
- Blackwell, P. G. (1997). Random diffusion models for animal movement. Ecological Modelling, 100, 87-102. https://doi.org/10.1016/S0304 -3800(97)00153-1
- Blackwell, P. G. (2003). Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika*, 90(3), 613–627. https://doi.org/10.1093/biomet/90.3.613
- Blackwell, P. G. (2020). Integrated continuous-time hidden Markov models. arXiv preprint, arXiv:1807.11907. https://doi.org/10.48550/ arXiv.1807.11907
- Blackwell, P. G., & Matthiopoulos, J. (2024). Joint inference for telemetry and spatial survey data. Ecology, 105, e4457. https://doi.org/10. 1002/ecy.4457
- Blackwell, P. G., Niu, M., Lambert, M. S., & LaPoint, S. D. (2015). Exact Bayesian inference for animal movement in continuous time. *Methods in Ecology and Evolution*, 7, 184–195. https://doi.org/10. 1111/2041-210X.12460
- Bullard, F. (1999). Estimating the home range of an animal: a Brownian bridge approach. PhD thesis, University of North Carolina, Chapel Hill, North Carolina, USA.
- Chassan, M., & Concordet, D. (2023). How to test the missing data mechanism in a hidden Markov model. *Computational Statistics* &

Data Analysis, 182, 107723. https://doi.org/10.1016/j.csda.2023. 107723

- Chattamvelli, R., & Shanmugam, R. (2021). The Rayleigh distribution. In Continuous distributions in engineering and the applied sciences – Part II (pp. 245–252). Springer International Publishing. https://doi.org/ 10.1007/978-3-031-02435-1_9
- Conners, M., Michelot, T., Heywood, E., Orben, R., Philips, R., Vyssotski, A., Shaffer, S., & Thorne, L. (2021). Hidden Markov models identify major movement modes in accelerometer and magnetometer data from four albatross species. *Movement Ecology*, *9*(1), 7. https://doi. org/10.1186/s40462-021-00243-z
- Dedeban, E., Webber, Q., Laforge, M., Robitaille, A., & Wal, E. (2023). Uncertainty during migration: The role of predictable resources in unpredictable conditions in determining movement in woodland caribou (*Rangifer tarandus*). Journal of Mammalogy, 104(6), 1290– 1301. https://doi.org/10.1093/jmammal/gyad081
- Díaz-Ruiz, F., Caro, J., Delibes-Mateos, M., Arroyo, B., & Ferreras, P. (2016). Drivers of red fox (Vulpes vulpes) daily activity: Prey availability, human disturbance or habitat structure? Journal of Zoology, 298(2), 128–138. https://doi.org/10.1111/jzo.12294
- Ford, L. S., & Hoffman, R. S. (1988). Potos flavus. Mammalian Species, 321, 1–9.
- Forsythe, G., Malcolm, M., & Moler, C. (1977). Computer methods for mathematical computations. Wiley.
- Fothergill, A., & Cordey, H. (2015). The hunt. Random House.
- Franke, A., Caelli, T., & Hudson, R. (2004). Analysis of movements and behaviour of caribou (*Rangifer tarandus*) using hidden Markov models. *Ecological Modelling*, 173(2–3), 259–270. https://doi.org/10. 1016/j.ecolmodel.2003.06.004
- Früwirth-Schantter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2), 183–202. https://doi. org/10.1111/j.1467-9892.1994.tb00184.x
- Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Booth, M., & Rossi, F. (2006). GNU scientific library reference manual. Network Theory Limited.
- Galvis, N., Link, A., Mosquera, D., Vinueza, G., Carrillo, L., Mopán-Chilito, A., & Montillla, S. O. (2024). Notes on the ecology, activity patterns and behavior of the kinkajou (*Potos flavus*). *Mammalia*, 88(4), 292– 298. https://doi.org/10.1515/mammalia-2023-0055
- Glennie, R., Adam, T., Leos-Barajas, V., Michelot, T., Photopoulou, T., & McClintock, B. T. (2023). Hidden Markov models: Pitfalls and opportunities in ecology. *Methods in Ecology and Evolution*, 14, 43–56. https://doi.org/10.1111/2041-210X.13801
- Graham, R. L., Knuth, D. E., & Patashnik, O. (1989). Concrete mathematics: A foundation for computer science. Addison-Wesley Publishing Company.
- Grainger, D., & Blackwell, P. G. (2025). R code for FInCH: Fast statistical inference for continuous-time animal movement. The University of Sheffield. Software. https://doi.org/10.15131/shef.data.28675073.v1
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4), 711–732. https://doi.org/10.1093/biomet/82.4.711
- Guttorp, P. (1995). Stochastic modelling of scientific data. Chapman and Hall/CRC.
- Hankin, R. K. S. (2006). Special functions in R: Introducing the gsl package. R News, 6, 1–7.
- Harrington, L. (2015). International commercial trade in live carnivores and primates 2006–2012: Response to Bush et al. 2014. *Conservation Biology*, *29*(1), 293–296. https://doi.org/10.1111/cobi. 12448
- Harris, K. J., & Blackwell, P. G. (2013). Flexible continuous-time modelling for heterogeneous animal movement. *Ecological Modelling*, 255, 29–37. https://doi.org/10.1016/j.ecolmodel.2013.01.020
- Hobolth, A., & Stone, E. (2009). Simulation from endpoint-conditioned, continuous-time Markov chains on a finite state space, with

applications to molecular evolution. *Annals of Applied Statistics*, 3(3), 1204–1231.

- Horne, J. S., Garton, E. O., Krone, S. M., & Lewis, J. S. (2007). Analyzing animal movements using Brownian bridges. *Ecology*, 88(9), 2354– 2363. https://doi.org/10.1890/06-0957.1
- Jensen, A. (1953). Markoff chains as an aid in the study of Markoff processes. Skandinavisk Aktuarietidskrift, 36, 87–91. https://doi.org/10. 1080/03461238.1953.10419459
- Johnson, D. S., London, J. M., Lea, M.-A., & Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89(5), 1208–1215. https://doi.org/10.1890/07-1032.1
- Kays, R., & Hirsch, B. (2015). Data from: Stink or swim: techniques to meet the challenges for the study and conservation of small critters that hide, swim or climb and may otherwise make themselves unpleasant. *Movebank Data Repository*, https://doi.org/10.5441/ 001/1.41076dq1
- Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., & Morales, J. M. (2012). Flexible and practical modeling of animal telemetry data: Hidden Markov models and extensions. *Ecology*, 93(11), 2336–2342. https://doi.org/10.1890/11-2241.1
- McClintock, B. T. (2017). Incorporating telemetry error into hidden Markov models of animal movement using multiple imputation. Journal of Agricultural, Biological, and Environmental Statistics, 22, 249–269. https://doi.org/10.1007/s13253-017-0285-6
- McClintock, B. T., & Michelot, T. (2018). momentuHMM: R package for generalized hidden Markov models of animal movement. *Methods* in Ecology and Evolution, 9(6), 1518–1530. https://doi.org/10.1111/ 2041-210X.12995
- McKellar, A., Langrock, R., Walters, J., & Kesler, D. (2015). Using mixed hidden Markov models to examine behavioural states in a cooperatively breeding bird. *Behavioral Ecology*, 26(1), 148–157. https://doi. org/10.1093/beheco/aru171
- Michelot, T., & Blackwell, P. G. (2019). State-switching continuous-time correlated random walks. *Methods in Ecology and Evolution*, 10(5), 637–649. https://doi.org/10.1111/2041-210X.13154
- Michelot, T., Gloaguen, P., Blackwell, P. G., & Étienne, M. (2019). The Langevin diffusion as a continuous-time model of animal movement and habitat selection. *Methods in Ecology and Evolution*, 10(1), 1894–1907. https://doi.org/10.1111/2041-210X.13275
- Michelot, T., Langrock, R., & Patterson, T. A. (2016). moveHMM: An R package for the statistical modelling of animal movement data using hidden Markov models. *Methods in Ecology and Evolution*, 7(11), 1308–1315. https://doi.org/10.1111/2041-210X.12578
- Morales, J. M., Haydon, D. T., Frair, J., Holsiner, K. E., & Fryxell, J. M. (2004). Extracting more out of relocation data: Building movement models as mixtures of random walks. *Ecology*, 85(9), 2436–2445. https://doi.org/10.1890/03-0269
- Nelder, J., & Mead, R. (1965). A simplex method for function minimization. The Computer Journal, 7(4), 308–313. https://doi.org/10.1093/ comjnl/7.4.308
- Parton, A., & Blackwell, P. G. (2017). Bayesian inference for multistate 'step and turn' animal movement in continuous time. *Journal of Agricultural*, *Biological*, and *Environmental Statistics*, 22(3), 373–392. https://doi.org/10.1007/s13253-017-0286-5
- Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., & King, R. (2017). Statistical modelling of individual animal movement: An overview of key methods and a discussion of practical challenges. AStA Advances in Statistical Analysis, 101(4), 399–438. https://doi.org/10.1007/s10182-017-0302-7
- Porteus, T. A., Short, M. J., Hoodless, A. N., & Reynolds, J. C. (2024a). Movement ecology and minimum density estimates of red foxes in wet grassland habitats used by breeding wading birds. *European Journal of Wildlife Research*, 70(1), 8. https://doi.org/10.1007/s1034 4-023-01759-y

- Porteus, T. A., Short, M. J., Hoodless, A. N., & Reynolds, J. C. (2024b). Data from: Study 'Red Fox (Vulpes vulpes) in UK wet grasslands'. https://doi.org/10.5441/001/1.304
- Powell, R. A., Ellwood, S., Kays, R., & Maran, T. (2017). Stink or swim: Techniques to meet the challenges for the study and conservation of small critters that hide, swim, or climb, and may otherwise make themselves unpleasant. In D. Macdonald, C. Newman, & L. Harrington (Eds.), *Biology and conservation of musteloids* (pp. 216– 230). Oxford University Press.
- Rubin, D., & Schenker, N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association*, 81(394), 366–374. https://doi.org/10.2307/2289225
- Sanderson, J., & Trolle, M. (2005). Monitoring elusive mammals: Unattended cameras reveal secrets of some of the world's wildest places. American Journal of Science, 93, 148–155.
- Storrie, L., Loseto, L., Sutherland, E., MacPhee, S., O'Corry-Crowe, G., & Hussey, N. (2023). Do beluga whales truly migrate? Testing a key trait of the classical migration syndrome. *Methods in Ecology and Evolution*, 11(1), 53. https://doi.org/10.1186/s40462-023-00416-y
- Taylor, L., Vollrath, F., Lambert, B., Lunn, D., Douglas-Hamilton, I., & Wittemyer, G. (2019). Movement reveals reproductive tactics in male elephants. *Journal of Animal Ecology*, 89, 57–67. https://doi. org/10.1111/1365-2656.13035
- Uhlenbeck, G., & Ornstein, L. (1930). On the theory of the Brownian motion. The Physical Review, 36(5), 823–841. https://doi.org/10.1103/ PhysRev.36.823
- Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on*

Methods in Ecology and Evolution -

Information Theory, 13, 260–269. https://doi.org/10.1109/TIT.1967. 1054010

- Wright, W. C., Wilkinson, B. E., & Cropper, W. P. (2017). Estimating GPS signal loss in a natural deciduous forest using sky photography. *Papers in Applied Geography*, 3(2), 119–128. https://doi.org/10. 1080/23754931.2016.1264990
- Zhang, J., Rayner, M., Vickers, S., Landers, T., Sagar, R., Stewart, J., & Dunphy, B. (2019). GPS telemetry for small seabirds: Using hidden Markov models to infer foraging behaviour of common diving petrels (Pelecanoides urinatrix urinatrix). Emu - Austral Ornithology, 119(2), 126-137. https://doi.org/10.1080/01584197.2018.1558997
- Zucchini, W., MacDonald, I., & Langrock, R. (2016). Hidden Markov models for time series: An introduction using R (2nd ed.). Chapman and Hall/CRC.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article. **Appendix S1:** Derivations and large-sample simulations.

How to cite this article: Grainger, D., & Blackwell, P. G. (2025). FInCH: Fast statistical inference for continuous-time animal movement. *Methods in Ecology and Evolution*, 00, 1–23. https://doi.org/10.1111/2041-210X.70052