The appendices to the manuscript "A model for leveraging animal movement to understand spatio-temporal disease dynamics" serve three functions. First, they give a more formal presentation of the ideas presented in the main text. Second, they illustrate extensions of MoveSTIR that highlight its flexibility. Third, they provide details on the implementation of MoveSTIR, with reference to the code that accompanies the manuscript. While all appendices contribute to increased understanding of MoveSTIR, they are not all required for the casual reader/user to interface with MoveSTIR. At the beginning of each Appendix, we provide a one sentence summary of the section and whether it is necessary to understand MoveSTIR as presented in the main text. We also provide multiple worked examples of MoveSTIR in the Jupyter notebooks 'moveSTIR_examples.ipynb' and 'moveSTIR_tutorial.ipynb' available at https://github.com/mqwilber/moveSTIR or https://zenodo.org/badge/latestdoi/409263341.

1 Appendix S1: The contact function $\Phi(\mathbf{s}_j(\tau_d), \mathbf{s}_i(\tau_a))$

Summary: *This appendix describes the units associated with and possible functional forms for the contact function of MoveSTIR. It is recommended for understanding how MoveSTIR is implemented in the main text.*

The function $\delta_{x_j(u)}(x)$ in equation 6 in the main text specifies whether or not the location of individual j at time u is in location x. However, it does not explicitly reference the area of location x. Considering area is important because the density of pathogen encountered by an acquiring host affects the force of infection and this density inherently depends on the area of x. Equation 4 in the main text implicitly accounts for area by defining $\beta' = \tilde{\beta}/(\text{Area of } x)$, where $\tilde{\beta}$ has units $\frac{\text{area units}}{\text{time}}$ (e.g., $\frac{m^2}{\text{hour}}$). This means that increasing the area of x decreases the overall force of infection because it decreases encounters between the acquiring host and deposited pathogen, conditional on hosts being in area x.

We can more generally account for area by writing $\beta'\delta_{x_j(u)}(x)$ as $\tilde{\beta}\Phi(x_j(u), x)$ (Gurarie & Ovaskainen 2013; Martinez-Garcia *et al.* 2020). The function $\Phi(x_j(u), x)$ is the contact function. We can write the contact function as $\Phi(\mathbf{s}_j(\tau_d), \mathbf{s}_i(\tau_a))$, where $\mathbf{s}_i(\tau_a)$ and $\mathbf{s}_j(\tau_d)$ are the locations of the acquiring host i and depositing host j at time τ_a and τ_d, respectively. The contact function is a probability density function that specifies how likely contact is to occur between an acquiring host i at time τ_a and a present or past depositing host j at time $\tau_d < \tau_a$. The contact function depends on the distance between host i and host j (Gurarie & Ovaskainen 2013). There are two types of contact functions that we consider: a top-hat contact function and a Gaussian contact function (Gurarie & Ovaskainen 2013; Martinez-Garcia *et al.* 2020).

The top-hat contact function is defined in two-dimensional space as

$$\Phi(\mathbf{z} = \mathbf{s}_i(\tau_a) - \mathbf{s}_j(\tau_d)) = \begin{cases} \frac{1}{\pi\alpha^2} & \text{if } ||\mathbf{z}|| < \alpha \\ 0 & \text{otherwise} \end{cases} \tag{S1}$$

where $||\mathbf{z}||$ is the Euclidean distance between locations $\mathbf{s}_j(\tau_d)$ and $\mathbf{s}_i(\tau_a)$ (Gurarie & Ovaskainen 2013). The top-hat contact function is uniform within a circular area $\pi\alpha^2$ and zero everywhere else, where α is the radius of the circle.

The Gaussian contact function is defined in two-dimensional space as

$$\Phi(\mathbf{z} = \mathbf{s}_i(\tau_a) - \mathbf{s}_j(\tau_d)) = \frac{1}{4\alpha^2} \exp(-\frac{\pi||\mathbf{z}||^2}{4\alpha^2}) \tag{S2}$$

where α is the mean distance of the half Gaussian (Gurarie & Ovaskainen 2013). The Gaussian contact function decays with distance $||\mathbf{z}||$ between the acquiring host and depositing host.

Conceptually, one can think of the contact function in two equivalent ways. First, the depositing host deposits a packet of pathogen at a point location and the contact function then modifies how the acquiring host encounters and acquires that packet in the local area it searches. Alternatively, one can envision the acquiring host at a point location and the contact function instantaneously redistributing deposited pathogen across a local area, reducing the density of pathogen that the acquiring host encounters at its point location. This latter interpretation has the benefit of highlighting that the contact function could include time-dependence to account for the diffusion of the pathogen through space.

2 Appendix S2: Extending the transmission kernel

Summary: *This appendix describes how the transmission kernel of MoveSTIR can be extended to account for additional biological realism. It is not necessary for understanding MoveSTIR as presented in the main text.*

There are myriad possibilities for adding additional biological realism to the transmission kernel and they will depend on the system under investigation. We consider two extensions here: temporal and spatial behavioral filters and pathogen decay that varies in space in time.

Behavioral and spatial filters: The transmission kernel defined in equation 5 in the main text assumes that infected hosts constantly deposit pathogen as they move. Similarly, the kernel assumes that susceptible hosts constantly acquire pathogen as they move (given pathogen is present at their current location). However, acquisition and deposition will often depend on behavioral and spatial context. For example, acquisition of strongyle nematodes infecting sheep occurs when a host is feeding (Hayward *et al.* 2019).

Deposition of raccoon roundworm occurs during defecation and typically at specific spatial locations (i.e., latrines, Weinstein *et al.* 2018). Adding these temporal or spatial filters to the transmission kernel is as simple as including additional multiplicative terms in $K_{a_i \leftarrow d_j}(\tau_a, \tau_d)$. These terms specify, probabilistically or deterministically, whether a host is engaging in a behavior or is in a spatial location at a given time that is conducive to acquisition or deposition. For example, we might re-write the transmission kernel as

$$K_{a_i \leftarrow d_j}(\tau_a, \tau_d) = \begin{cases} [\tilde{\beta}\pi_{i,\text{foraging}}(\tau_a)][\Phi(\mathbf{s}_j(\tau_d), \mathbf{s}_i(\tau_a))][\lambda\delta_{I_j(\tau_d)}(I)\delta_{\mathbf{s}_j(\tau_d)}(A_{\text{latrine}})][e^{-\nu(\tau_a - \tau_d)}] & \text{for } \tau_d \leq \tau_a \\ 0 & \text{otherwise} \end{cases} \tag{S3}$$

where $\pi_{i,\text{foraging}}(\tau_a)$ is the probability that host i is foraging at time τ_a and $\delta_{\mathbf{s}_j(\tau_d)}(A_{\text{latrine}})$ is an indicator function that specifies whether host j was in the area A_{latrine} at time τ_d. In this example, deposition only occurs in area A_{latrine} and acquisition rate is modified by the probability that a host is foraging. Acquisition rate could also be modified by climatic variables such as temperature and this would be a simple extension of modifying $\tilde{\beta}$ to $\tilde{\beta}(g(\tau_a))$ where $g(\tau_a)$ is a function that returns the temperature at time τ_a. In terms of implementation, allowing acquisition to vary with temperature would increase the runtime of the model by $O(n^2)$, where n is the number of discretized time points of the continuous-time movement trajectory (see Appendix S4 for additional details on implementation).

Empirically informing these filters is more challenging, but possible. For example, recent advances in movement methods and technology allow for probabilistic identification of behavioral states directly from high-resolution movement data (Edelhoff *et al.* 2016), such as whether a host is resting, foraging or engaging in different movement types such as migration- or home range-related movements. This type of information could be used to directly inform the behavioral filter on the transmission kernel. Similarly, if the spatial locations of latrines are known *a priori*, they can be included directly in the kernel as a spatial filter on deposition (e.g., specifying that deposition occurs when a host is at a latrine).

Pathogen decay that varies in space and time: The transmission kernel in equation 5 in the main text also assumes that pathogen decay is constant in space and time. However, different aspects of the spatial environment (e.g., soil type; Saunders *et al.* 2012) and temporal environment (e.g., temperature and humidity; Fine *et al.* 2011) can significantly affect the rate of pathogen decay. For example, *a priori* information on the proportional change in pathogen decay rate in different soils types (Saunders *et al.* 2012) and a spatial map of soil types on the landscape on which hosts are moving could be incorporated into the transmission kernel by modifying the pathogen survival function $e^{-\nu(\tau_a - \tau_d)}$ to $e^{-\nu_{\mathbf{s}(\tau_d)}(\tau_a - \tau_d)}$. The pathogen decay rate now depends on the spatial location of the depositing host at the time of deposition. When

pathogen decay varies in time, for example due to changes in temperature (Fine *et al.* 2011), we could update the pathogen survival function to $e^{-\int_{\tau_d}^{\tau_a} \nu(g(\kappa))d\kappa}$ where $\nu(g(\kappa))$ illustrates that pathogen decay rate changes with some function of time $g(\kappa)$, which may be temperature or humidity for example. By working directly with the transmission kernel, we can add biologically meaningful modifications and still apply the same set of tools to quantify infection risk in space, time, and among individuals (Table 1 in the main text).

3 Appendix S3: Toy examples with the transmission kernel

Summary: *This appendix provides three toy examples that illustrate how the transmission kernel $K_{a_i \leftarrow d_j}$ can be used to understand the maximum infection risk experienced by two individuals. It is intended to augment the conceptual grasp of MoveSTIR in the main text, but is not strictly necessary for understanding MoveSTIR.*

Example 1: Two hosts are in the same place at all times

Consider the simplest situation when two hosts are in the same place over a time period from $t_{\text{start}} = 0$ to t_{end} (Fig. S1A). For conceptual understanding, let us assume that there is no decay of the pathogen in the environment, such that $e^{-\nu(\tau_a - \tau_d)}$ is always one. The transmission kernel is the same for both host 1 experiencing a force of infection from host 2 and vice versa: $K_{a_1 \leftarrow d_2}(\tau_a, \tau_d) = K_{a_2 \leftarrow d_1}(\tau_a, \tau_d) = \beta' \lambda$. We can compute the maximum potential force of infection experienced by host 1 from host 2 at any time point t (i.e., $h_{1 \leftarrow 2}(t)$) by integrating over host 2's deposition rate from $t_{\text{start}} = 0$ to t: $h_{1 \leftarrow 2}(t) = \int_0^t \beta' \lambda du = \beta' \lambda t$ (Table 1 in the main text). Note that β' has units time^{-1} and λt is unitless so $\beta' \lambda t$ is a rate, consistent with force of infection. As shown in Fig. S1B, the maximum potential force of infection increases through time. This is because, if infected, host 2 is continually depositing pathogen into the environment through time. Since we assume in this example that there is no pathogen decay, pathogen will continue to accumulate and the instantaneous infection risk experienced by host 1 from host 2 (or vice versa) will continue to increase.

The potential cumulative infection risk experienced by host 1 from host 2 up to time t ($H_{1 \leftarrow 2}(t)$ and vice versa) can be calculated by integrating the maximum potential force of infection $h_{1 \leftarrow 2}(t)$: $\int_0^t h_{1 \leftarrow 2}(u)du = \int_0^t \int_0^u K_{a_1 \leftarrow d_2}(u, \tau) d\tau du$ (Table 1 in the main text). In this example, we get $\int_0^t \beta' \lambda u du = \beta' \lambda t^2/2$ (Fig. S1C). The cumulative hazard is a monotonically increasing unitless function.

Example 2: Two hosts are never in the same location at the same time

In this example, we consider three locations I, II, and III. Hosts 1 and 2 move among these locations from $t_{\text{start}} = 0$ to t_{end} as shown in Fig. S1D, such that hosts are never in the same place at the same time, but are in the same place at different times. The maximum potential force of infection felt by host 1 from host 2

and host 2 from host 1 are no longer the same. Fig. S1E shows the force of infection functions $h_{1\leftarrow 2}(t)$ and $h_{2\leftarrow 1}(t)$. From time $t_{\text{start}} = 0$ to t_1 host 1 is in location I and and host 2 is in location II. Thus, they do not experience any force of infection from each other (Fig. S1E). From time t_1 to time t_2, host 1 is in location II and is exposed to pathogens previously deposited by host 2. Because host 2 has left location II and we assume no pathogen decay, the force of infection experienced by host 1 from host 2 over t_1 to time t_2 is the constant $\beta'\lambda t_1$ because host 2 is no longer contributing any additional pathogen. In contrast, from time t_1 to t_2, host 2 is in location III where host 1 has not yet visited, so the maximum potential force of infection experienced by host 2 from host 1 is still zero. Finally, from t_2 to t_{end} host 1 is now in location III and host 2 is in location II (Fig. S1E). Host 1 experiences a constant force of infection $\beta'\lambda(t_2 - t_1)$ from host 2 based on the past pathogen deposited when host 2 was in location II. Additionally, host 2 now experiences a force of infection $\beta'\lambda(t_2 - t_1)$ from past pathogen deposited by host 1 when it was in location II from t_1 to t_2. The cumulative force of infection is shown in Fig. S1F.

Example 3: Two hosts are always together in space, but their location varies through time
In the third example, two hosts are always in the same location, but that location varies through time (specifically, hosts move from location I to II to III, Fig. S1G). The maximum potential force of infection experienced by host 1 from host 2 (and vice versa) at time t in $t_{\text{start}} = 0$ to t_1 is $\beta'\lambda t$. When hosts move from location I to II at time t_1, there is initially no pathogen in the environment contributed by either host, so the force of infection experienced by host 1 from host 2 (and vice versa) at time t in (t_1, t_2) is $\beta\lambda(t - t_1)$ (Fig. S1H). Similarly, when hosts move from location II to III, there is again no pathogen in the environment so the maximum force of infection experienced by host 1 from host 2 (and vice versa) at time t in (t_2, t_{end}) is $\beta\lambda(t - t_2)$ (Fig. S1H). The cumulative force of infection is shown in Fig. S1I.

4 Appendix S4: Implementing the transmission kernel as a matrix

Summary: *This appendix describes how the transmission kernel of MoveSTIR can be numerically computed. It is not necessary for understanding MoveSTIR as presented in the main text.*

The transmission kernel $K_{a_i\leftarrow d_j}(\tau_a, \tau_d)$ quantifies the transmission weight between host i who is acquiring the pathogen at time τ_a and host j who deposited the pathogen at time τ_d. While the transmission kernel is defined in continuous time, it will often be practical to work with a discretized version of the transmission kernel. We will refer to this discretized version as the transmission matrix $\mathbf{K}_{a_i\leftarrow d_j}$.

To define the transmission matrix $\mathbf{K}_{a_i\leftarrow d_j}$, consider the time interval (t_0, t_n) where we have recorded host movement. We can break the interval into n equally spaced segments with width Δt such that $\mathbf{K}_{a_i\leftarrow d_j}$ is an

$n \times n$ matrix of grid cells with width and height Δt. For a grid in the matrix $\mathbf{K}_{a_i \leftarrow d_j}$ bounded by $(\tau_a, \tau_a + \Delta t)$ and $(\tau_d, \tau_d + \Delta t)$, we evaluate $K_{a_i \leftarrow d_j}(\tau_a, \tau_d)$ at $K_{a_i \leftarrow d_j}(\tau_a + 0.5\Delta t, \tau_d + 0.5\Delta t) = k_{\tau_a, \tau_d}$, consistent with a two-dimensional grid approximation at the grid midpoint. The notation k_{τ_a, τ_d} indicates the transmission weight felt by the acquiring host at time τ_a from the depositing host at time τ_d (approximated at the grid midpoint). The units on k_{τ_a, τ_d} are time^{-2}. The transmission matrix is

$$\mathbf{K}_{a_i \leftarrow d_j} = \begin{array}{c} \\ \tau_0 \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_n \end{array} \begin{array}{c} \begin{array}{ccccc} \tau_0 & \tau_1 & \tau_2 & \cdots & \tau_n \end{array} \\ \begin{bmatrix} k_{0,0} & 0 & 0 & \cdots & 0 \\ k_{1,0} & k_{1,1} & 0 & \cdots & 0 \\ k_{2,0} & k_{2,1} & k_{2,2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ k_{n,0} & k_{n,1} & k_{n,2} & \cdots & k_{n,n} \end{bmatrix} \end{array} \tag{S4}$$

The columns represent host j's time points who is depositing pathogen and the rows represent host i's time points who is acquiring pathogen. The upper triangle is all zeros as host i cannot feel the force of infection of future host j.

To calculate any of the quantities given in Table 1 of the main text, we just need to sum particular grid cells in $\mathbf{K}_{a_i \leftarrow d_j}$. However, we need to be cognizant of the dimensions we are summing over to ensure that we account for the grid approximation. For example, to get the force of infection felt by host i from j at time τ_2 ($h_{i \leftarrow j}(\tau_2)$) we need to sum the row τ_2 and multiply by Δt. Multiplying by Δt accounts for the discretization of host j's trajectory and ensures that $h_{i \leftarrow j}(\tau_2)$ has the correct units, namely time^{-1}. If we want to get the cumulative hazard felt by host 1 due to host 2 up to τ_n, we would sum all of the entries in $\mathbf{K}_{a_i \leftarrow d_j}$ and multiply by Δt^2. This accounts for the fact that we are discretizing over host 1's and host 2's trajectory and ensures that our cumulative hazard is unitless (as it should be). Finally, to calculate the force of infection due to direct transmission, where hosts are in the same place at the same time (within Δt), we could sum the main diagonal of $\mathbf{K}_{a_i \leftarrow d_j}$ and multiply by Δt.

4.1 Combining transmission matrices

When we are considering multiple hosts within a population, we can combine transmission matrices into a larger block matrix that specifies the population-level transmission matrix. For example, consider N interacting hosts. We can define the population-level transmission matrix as

$$\mathbf{F} = \begin{array}{c} \\ \text{host}_1 \\ \text{host}_2 \\ \text{host}_3 \\ \vdots \\ \text{host}_N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{host}_1 & \text{host}_2 & \text{host}_3 & \cdots & \text{host}_N \end{array} \\ \begin{bmatrix} \mathbf{0} & \mathbf{K}_{a_1 \leftarrow d_2} & \mathbf{K}_{a_1 \leftarrow d_3} & \cdots & \mathbf{K}_{a_1 \leftarrow d_N} \\ \mathbf{K}_{a_2 \leftarrow d_1} & \mathbf{0} & \mathbf{K}_{a_2 \leftarrow d_3} & \cdots & \mathbf{K}_{a_2 \leftarrow d_N} \\ \mathbf{K}_{a_3 \leftarrow d_1} & \mathbf{K}_{a_3 \leftarrow d_2} & \mathbf{0} & \cdots & \mathbf{K}_{a_3 \leftarrow d_N} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \mathbf{K}_{a_N \leftarrow d_1} & \mathbf{K}_{a_N \leftarrow d_2} & \mathbf{K}_{a_N \leftarrow d_3} & \cdots & \mathbf{0} \end{bmatrix} \end{array} \tag{S5}$$

Summing over the host_1 row, for example, yields a new transmission matrix $\mathbf{K}_{a_1 \leftarrow \sum_{j \in N_{-1}} d_j}$ which gives the transmission weight (units time^{-2}) felt by host 1 from all other hosts. Taking the row sums of $\mathbf{K}_{a_1 \leftarrow \sum_{j \in N_{-1}} d_j}$ and multiplying by Δt yields a vector giving the force of infection felt by host 1 from all other hosts combined at any given time. The block matrix $\mathbf{F}$ will play an important role in calculating R_0, as we describe in Appendix S7.

5 Appendix S5: Using other forms of spatial and temporal data with MoveSTIR

Summary: *This appendix presents the details for how MoveSTIR can be applied to other commonly collected forms of animal movement data. It is not necessary for understanding MoveSTIR as presented in the main text.*

There are three classes of data describing spatial and temporal interactions and co-occurrences that can be used with MoveSTIR: continuous space-discrete (high resolution) time, discrete (low resolution) space-continuous time, and discrete (low resolution) space-discrete (low resolution) time. Some examples of these data types are as follows

1. **Continuous space-discrete (high-resolution) time**: These data include VHF tags, GPS tags, and Argos tags. Using these devices, continuous spatial data are obtained at specific time fixes (e.g., every 30 minutes). Data of this type are the primary focus of our analyses in the main text.

2. **Discrete (low resolution) space-continuous time**: These data include proximity loggers, radio-frequency identification devices (RFID), and some forms of camera traps. Proximity loggers and RFIDs can be deployed on individual hosts and will activate when two hosts are within a pre-determined separation distance and record the duration – but not the location – of a resulting contact (Stehlé *et al.* 2011b; Drewe *et al.* 2012). These devices can also be placed in spatial locations of interest, where

they will activate when hosts are within a certain minimum distance and record the time a host spends at the location (Lavelle *et al.* 2016). Similarly, camera traps can be set up at specific spatial locations and be activated by motion senors to take photos when hosts are within some detection distance of the camera. They can be programmed to take a nearly continuous series of photos while the animal is in the detection range, allowing for a near continuous-time delineation of when an animal was at a given location. There are, however, a range of different setups for camera traps and some would be better categorized as discrete space-discrete time data as described below (see Hamel *et al.* 2013; Burton *et al.* 2015, for a discussion of camera trap sampling designs).

3. **Discrete (low resolution) space-discrete (low resolution) time**: These data include spatially explicit capture-recapture data (SECR) where marking can be done through pit tags, photo ID, toe-clipping, polymers and pigments, and many other techniques. When individuals are captured or recaptured, their general (e.g., the specific pond where they were captured) or specific (e.g., exact location of hair snare) location and time of capture are recorded (Royle *et al.* 2014). Long-term time series of SECR can provide information on how hosts are moving between discrete locations on the landscape (e.g., between different ponds) (Royle *et al.* 2014; Cayuela *et al.* 2017; Silk *et al.* 2021). In addition, previous analyses that use home range overlap to derive contact networks (e.g., Godfrey *et al.* 2010; Godfrey 2013; Springer *et al.* 2017) can be considered a discrete space-discrete/continuous time application and are thus special cases of MoveSTIR.

In the main text we focused on using continuous space-discrete (high resolution) time GPS data with MoveSTIR. In the following sections we show that MoveSTIR can be applied to discrete (low resolution) space-continuous time data and discrete (low resolution) space-discrete (low resolution) time. The key step for using these data with MoveSTIR is understanding how they can be interpolated into continuous time and then framed as a transmission kernel. From MoveSTIR's perspective, it does not matter if space is discrete or continuous.

5.1 Discrete (low resolution) space-continuous time data

For this data type, we will focus on contact data obtained from proximity loggers/RFIDs. When proximity loggers are deployed on individual hosts, they typically provide only temporal information and no spatial information (i.e., we do not know where in space a contact is occurring, but know when it occurs and for how long; Yang *et al.* 2020). Spatial information can be obtained from proximity loggers if they are placed on environmental resources, such as latrines, water sources, or feeding stations (Lavelle *et al.* 2016; Silk *et al.* 2018; Yang *et al.* 2020). In these situations, contacts are recorded when a host is at or near this

spatial location. This type of information can be used to at least partially inform indirect contact rates (Silk *et al.* 2018; Wilber *et al.* 2019). Following Silk *et al.* (2018), we will refer to host-host contacts recorded by proximity loggers as social contacts and host-environment contacts as spatial contacts. As we will show, social and spatial contacts from proximity logger data can be directly converted into a transmission kernel and modeled with MoveSTIR.

We can visualize pairwise social contact data from proximity loggers as the line plot shown in Fig. S2A. The x-axis in this plot is time and the y-axis specifies whether or not two hosts are in proximity of each other over a given time interval (solid if they are, no line if they are not). From pairwise contact data alone, we do not know how hosts are moving while they are in proximity to each other. At one extreme, they may remain in the same location over the duration of the contact. At the other extreme, they may move together through space over the duration of the contact. We will assume that when in contact, hosts remain in the same location.

We can relate social contact data from proximity loggers or RFIDs directly to the transmission kernel of MoveSTIR. Specifically, consider the term in our transmission kernel $\delta_{x_j(u)}(x_i(t))$ that determines whether the acquiring host i at time t is in the same spatial location as the depositing host j at time u. We break $\delta_{x_j(u)}(x_i(t))$ into two components: $\delta_{\text{social}}(\tau_a, \tau_d)$ and $\delta_{\text{spatial}}(\tau_a, \tau_d)$. To calculate $\delta_{\text{social}}(\tau_a, \tau_d)$ from host to host contacts (e.g., the Host 1 - Host 2 line in Fig. S2A), social contacts occurring over a period of time t_m to t_n are represented by right triangles extending off of the diagonal of the transmission kernel (or transmission matrix when we discretize, Fig. S2B). The reasons for this are two-fold. First, the diagonal of the transmission matrix represents direct contacts that occur when hosts are in the same place at the same time. Second, because we assume that direct contacts occur in the same location, at the end of the contact duration t_n host i will feel force of infection from pathogen deposited by host j prior to t_n, but after t_m. The triangle emerging off the diagonal accounts for this potential *indirect* contact associated with a *direct* contact of non-zero duration. If we instead assumed that hosts were moving in space while in contact with each other, $\delta_{\text{social}}(\tau_a, \tau_d)$ would reduce to only the diagonal of the transmission matrix.

To incorporate spatial contacts through shared environmental resources (e.g., host 1 contacting a location following a contact by host 2; $\delta_{\text{spatial}}(\tau_a, \tau_d)$), we can represent the proximity logger data as a simplified movement trajectory for each host. As an example, assume there are two spatial resource locations where we have placed proximity loggers, location A and B. We know when and how long hosts were at locations A and B from the data we collected. We also know when hosts were not at A and B, which we will refer to as O. Two (discretized) movement trajectories might look like (also, see Fig. S2C):

Host 1: $OOOAOOBBBO$

Host 2: $AAOOBBBBBB$

We can interpret host 1's trajectory as follows. Host 1 is somewhere other than location A over the first three time steps (we do not know where) and then arrives in location A at time step four. Following time step 4, it again goes somewhere else. Host 1 arrives in location B by time step 7 and remains there through time step 9. Finally, it leaves location B and goes somewhere else. One can similarly convert the data shown in Fig. S2A to these types of movement trajectories.

For our purposes, these trajectories are conceptually identical to the spatial movement trajectories $\mathbf{s}(t)$ we have been using with MoveSTIR in the main text. We can use them to evaluate whether host 1 is in the same location as past or present host 2 (and vice versa, Fig. S2C). The only nuance is that when both hosts have a value of O, it does not count as the same location because O simply means "not a known spatial location". By combining $\delta_{\text{spatial}}(\tau_a, \tau_d)$ and $\delta_{\text{social}}(\tau_a, \tau_d)$ with a logical "OR", we have fully specified $\delta_{x_{j(u)}}(x_i(t))$ from proximity logger or RFID data (Fig. S2D). Now all of the analyses described in the main text for MoveSTIR apply.

5.2 Discrete space-discrete time data

Spatially explicit capture-recapture (SECR) data, whether from pit tags, camera traps, or other marking strategies, are a common form of (potentially coarse) resolution discrete space-discrete time movement data (Royle & Young 2008; Royle *et al.* 2009; Cayuela *et al.* 2017). For example, we might envision marked amphibians inhabiting a series of ponds and repeated surveys recapturing amphibians at different ponds within a metapopulation (Cayuela *et al.* 2020). There is a rich statistical literature on inference from SECR data and we will not attempt to review that here (e.g., Royle & Young 2008; Royle *et al.* 2009, 2014; Cayuela *et al.* 2017). Importantly, these statistical approaches allow us to infer the expected location of individuals not observed during a given primary/survey period, but observed at later primary periods (Royle & Young 2008; Cayuela *et al.* 2017). What this means is that after fitting the appropriate statistical model to our SECR data, we can obtain predictions in discrete space and time about where hosts are and when they are there (Fig. S3A). To use these data with MoveSTIR, we now need to represent them in continuous time. Recall that MoveSTIR applies equally well to discrete or continuous space. We will focus our example on SECR where animals are captured and recaptured in discrete habitat patches that form a metapopulation.

The simplest option to convert our SECR data to continuous time is to assume that the spatio-temporal data follow a step function. With this assumption, a host resides in a given patch until the time of its next

observation and, if the host is in a new patch at the next observation, the host immediately moves to this new patch from the previous patch. This is consistent with many metapopulation models where transit time between patches is assumed to be instantaneous (Wilber *et al.* 2020). Another step function approximation would be to use a midpoint rule such that if a host transitions from patch I to patch II somewhere between time t_1 and t_2, then we assume the transition occurred at the halfway point of the time interval (Fig. S3B). Moreover, if transit time between patches is expected to be long relative to the duration of infection, then transit time may be critical to consider to capture among-patch epidemiological dynamics (Cross *et al.* 2005). In this situation, one could add information on the host's transit by including a straight-line path between patches and allowing hosts to traverse the path over some pre-specified time determined by the distance between patches and the average speed the animal moves. Whether a step function or something more complicated is used, the key point is that there are reasonable approaches we can use to represent SECR data in continuous time and discrete space.

From MoveSTIR's perspective, once we have made the conversion to continuous time there are no fundamental differences between the SECR data and the GPS movement data we discuss in the main text. Therefore, we can represented the SECR data as pairwise transmission kernels and apply all of the tools of MoveSTIR (Fig. S3C,D).

One subtle point worth mentioning is that we need to consider the areas of the habitat patches among which hosts are moving. As discussed in Appendix S1, acquisition rate scales inversely with area given density-dependent contacts between hosts and pathogen in a local area. Therefore, the force of infection should be explicitly written as

$$h_{i \leftarrow j}(t, x_i) = \int_0^t \frac{\tilde{\beta}}{A_{x_i}} \lambda \delta_{x_j(u)}(x_i) \delta_{I_j(u)}(I) e^{-\nu(t-u)} du \tag{S6}$$

where $x_i \in \{x_1, x_2, \ldots, x_n\}$ is a discrete set of patches, each with a unique area A_{x_i}. The key assumption here is that the deposited pathogen is well-mixed within a habitat patch such that, given the same number of individuals, larger patches have lower densities of pathogen leading to lower acquisition rates. This is consistent with host-pathogen metapopulation models (Wilber *et al.* 2020).

6 Appendix S6: Using MoveSTIR to explore the epidemiological consequences of contact networks

Summary: *This appendix provides an example of how MoveSTIR can be represented as a static and dynamic contact network. It is not required for understanding the results presented in the main text, though supports*

statements therein.

In the main text, we provide empirical examples to show how MoveSTIR can be used to explore the structure of the direct and indirect contact networks defined by movement trajectories. Here, we use a simple simulated example to illustrate how MoveSTIR can be used to ask the question: how do static, weighted networks differ in their epidemiological predictions than dynamic, weighted networks (e.g., Stehlé *et al.* 2011a; Springer *et al.* 2017)? We use simulated data of five hosts moving on a landscape (Fig. S4). We define the following transmission kernel to compute maximum potential infection risk

$$K_{a_i \leftarrow d_j}(\tau_a, \tau_d) = \begin{cases} [\tilde{\beta}\lambda][\Phi(\mathbf{s}_j(\tau_d), \mathbf{s}_i(\tau_a))][e^{-\nu(\tau_a - \tau_d)}] & \text{for } \tau_d \leq \tau_a \\ 0 & \text{otherwise} \end{cases} \tag{S7}$$

where $\Phi(\mathbf{s}_j(\tau_d), \mathbf{s}_i(\tau_a))$ is the top-hat contact function that only allows transmission to occur when an acquiring host is within some minimum distance of the present or past depositing host. In this example, we set that distance to be 0.71 units. We set pathogen decay rate as 0.1 time^{-1}, acquisition rate $\tilde{\beta} = 1.5$ $\frac{\text{area units}}{\text{time}}$, deposition rate $\lambda = 1.5$ time^{-1}, and loss of infection rate $\gamma = 0.11$ time^{-1}.

From these data, we used MoveSTIR to build two contact networks. First, we built a static contact network of maximum potential infection risk. To calculate the static edge weights between individual hosts in this network, we computed the average of the maximum potential force of infection felt by host i from host j over the movement trajectory ($\bar{h}_{i \leftarrow j}(t)$, Table 1 in the main text). The resulting weighted, static network is shown in Figure S5A-B.

Second, we built a dynamic, weighted contact network based on the movement trajectories of the five hosts (Fig. S5C-D). Fig. S5C gives a visual representation of each pairwise transmission kernel $K_{a_i \leftarrow d_j}(\tau_a, \tau_d)$ (a single kernel is a grid in Fig. S5C) that together define the population-level transmission kernel $\mathbf{F}$ (Appendix S4) and the dynamic contact network. Fig. S5D shows a simplified representation of the dynamic contact network, where we discretized host movement into ten equally spaced temporal nodes that together span the entire movement trajectory. The edges represent the average force of infection felt by host i from host j over the given time step (0 through 9, Fig. S5C). Now we can ask: how do our predictions regarding epidemiological dynamics change when we use the weighted, static contact network (Fig. S5A) compared to the dynamic contact network (Fig. S5C-D)?

We first calculated R_0 for both networks and they were nearly identical (see Appendix S7 for R_0 calculations). The dynamic contact network predicted $R_0 = 1.44$ and the static contact network predicted $R_0 = 1.45$. However, we began to see divergence between R_0 in the static and dynamic network as we increased the loss of infection rate γ. Specifically, the static, weighted network began to underestimate R_0

compared to the dynamic network. In this situation, higher-order, time-dependent interactions defined by host movement trajectories, which were not captured by the static network, were increasingly important for transmission. This was because hosts were likely to lose the infection over the course of their movement trajectory and the invasion of the pathogen in the population depended more strongly on where hosts were and when they were there. This brief example shows how MoveSTIR can be used to expand our understanding of when dynamic networks predict different epidemiological dynamics than static, weighted networks.

7 Appendix S7: From transmission risk to infection dynamics

Summary: *This section describes how the transmission kernel can be linked to dynamic epidemiological models. It is necessary reading to understand how to derive epidemiological quantities such as R_0 from MoveSTIR.*

7.1 Specification of individual-level model

To link our transmission kernels $K_{a_i \leftarrow d_j}(\tau_a, \tau_d)$ to population-level disease dynamics, we re-construct our motivating model from equation 1 in the main text as a continuous-time, discrete-state Markov process, where the discrete states are $H_{i,S}, H_{i,I}$ for $i \in N$, representing whether individual host i is **S**usceptible or **I**nfected at time t. Importantly, a fundamental component of this model is the transmission kernel $K_{a_i \leftarrow d_j}(\tau_a, \tau_d)$. The model is

$$\begin{aligned} \frac{dp_{H_{i,S}}(t)}{dt} &= -p_{H_{i,S}}(t) \sum_{j \in N_{-i}} \int_0^t K'_{a_i \leftarrow d_j}(t,u)du + \gamma p_{H_{i,I}}(t) \\ \frac{dp_{H_{i,I}(t)}}{dt} &= p_{H_{i,S}}(t) \sum_{j \in N_{-i}} \int_0^t K'_{a_i \leftarrow d_j}(t,u)du - \gamma p_{H_{i,I}}(t) \end{aligned} \tag{S8}$$

where $p_{H_{i,S}}(t)$ and $p_{H_{i,I}}(t)$ are the probabilities of host i being susceptible or infected at time t, respectively, and $K'_{a_i \leftarrow d_j}(t,u) = [\tilde{\beta}][\Phi(\mathbf{s}_j(\tau_d = u), \mathbf{s}_i(\tau_a = t))][\lambda p_{H_{j,I}}(u)][e^{-\nu(t-u)}]$ for $u < t$. The only difference from the transmission kernel discussed in the main text is that we have replaced the indicator function $\delta_{I_j(\tau_d)}(I)$ with the probability that host j is infected at time $u < t$, $p_{H_{j,I}}(u)$. As before, key components of the transmission kernel, such as contact formation and duration across a direct to indirect continuum, can be directly estimated from movement data. We again assume that hosts who lose infection are immediately susceptible. However, equation S8 could easily be updated to include hosts that permanently recover from infection or die due to natural or disease-induced mortality (e.g., by adding a "Recovered" or "Mortality" class). Similarly, adding an additional "Exposed" class would be a simple extension (though we would need

to update our calculations for R_0 given below).

As an example, consider the case where we know when a host died. This information is often provided by GPS collaring technology. Without changing the structure of equation S8, we could update the transmission kernel to

$$K'_{a_i \leftarrow d_j}(t,u) = [\tilde{\beta}][\Phi(\mathbf{s}_j(\tau_d = u), \mathbf{s}_i(\tau_a = t))][\delta_{A_j(u)}(\text{Alive})][\lambda p_{H_{j,I}}(u)][e^{-\nu(t-u)}] \text{ for } u < t$$

where $\delta_{A_j(u)}(\text{Alive})$ evaluates to one if host j is alive at time u and zero otherwise. This ensures that dead hosts do not contribute to the force of infection felt by host i.

For some host-pathogen dynamics, such as wild pig and African swine fever (Pepin *et al.* 2020), dead hosts continue to contribute to infection through depositing pathogen via the carcass. We again assume that the time of death is known and occurs at d_τ. We also assume that carcasses are removed from the environment at rate c and deposit pathogen at rate λ_{carcass}. We could then update the transmission kernel to

$$\begin{aligned} K'_{a_i \leftarrow d_j}(t,u) =& [\delta_{A_j(u)}(\text{Alive})][\tilde{\beta}] \underbrace{[\Phi(\mathbf{s}_j(\tau_d = u), \mathbf{s}_i(\tau_a = t))]}_{\text{Contact with living host}} [p_{H_{j,I}}(u)][\lambda e^{-\nu(t-u)}] + \\ & [\delta_{A_j(u)}(\text{Dead})][\tilde{\beta}] \underbrace{[\Phi(\mathbf{s}_j(\tau_d = d_\tau), \mathbf{s}_i(\tau_a = t))]}_{\text{Contact with carcass}} \underbrace{[p_{H_{j,I}}(d_\tau) e^{-c(t-d_\tau)}]}_{\text{Carcass decay}} \underbrace{[\lambda_{\text{carcass}} e^{-\nu(t-u)}]}_{\text{Pathogen deposition and decay}} \\ & \text{for } d_\tau, u < t \end{aligned}$$

Note that $\mathbf{s}_j(\tau_d = d_\tau)$ assumes that host j remains where it was when it died at time d_τ. It is also possible to consider the situations where we do not know the time of death d_τ within the MoveSTIR framework. We will explore these situations in a future study.

7.2 R_0, individual-level $R_{0,i}$, and pairwise $R_{0,i \leftrightarrow j}$: derivation and perturbations

We can use equation S8 to estimate key epidemiological quantities, such as the fundamental reproductive number R_0. In the context of our model, R_0 can be interpreted as the expected number of individuals infected by an average infected individual over its infected lifetime. When $R_0 > 1$, the pathogen can successfully invade the host population. When $R_0 \leq 1$, the pathogen fails to invade the host population. To calculate R_0 from equation S8, we invoke a periodicity assumption where we assume that at the completion of the observed host movement trajectories the hosts immediately begin repeating the same movements. This might be a reasonable assumption if movement data were collected, say, over the course of a year for a migratory

species where we expect similar movements in the following year. However, movement data often are not collected over a biologically meaningful "period". Thus, it may be unreasonable (or completely impossible given the location of the host at the end of the movement trajectory relative to the start) to assume that hosts will immediately "restart" their movement trajectory following the completion of the observed movement trajectory. However, this periodicity assumption is still useful as it allows us to ask: given the movement we have observed, is the resulting dynamic contact network sufficient to allow a small amount of pathogen to on average increase by the end of the observed movement trajectory? Regardless of our ability to extrapolate beyond the observed movement trajectories (which may augment or reduce the overall potential for pathogen invasion), the periodicity assumption allows us to explore the implications of observed movement patterns on disease dynamics. This periodicity assumption has been previously used to compute pathogen invasion thresholds on dynamic networks (Valdano *et al.* 2015; Leitch *et al.* 2019).

To compute R_0 for equation S8, we use the h-state approach described in Diekmann *et al.* (2013) and expand our number of state variables to account for a time-in-the-movement-trajectory h-state for each individual i. In other words, we index each state variable $H_{i,S}$ or $H_{i,I}$ by t and expand our state space. While t is technically continuous leading to an infinite state space, in practice we work with t discretized over n equally spaced intervals of time yielding a finite space (though the state space could potentially be quite large). Note that the expanded state space model does not change anything about the dynamics of equation S8, it just allows us to draw on standard approaches to compute R_0 (Diekmann *et al.* 2013). We then linearize the expanded state space model about the disease free equilibrium for equation S8, resulting in a Jacobian matrix $\mathbf{J}$. Next, we decompose the resulting Jacobian matrix into $\mathbf{J} = \mathbf{F} + \mathbf{U}$. The matrix $\mathbf{F}$ defines how one infected host type produces infected hosts of other types. This matrix is identical to $\mathbf{F}$ defined in equation S5 and is completely defined by our transmission matrices. The matrix $\mathbf{U}$ defines the rate at which infected hosts of all types leave the infected class (Diekmann *et al.* 2013). The next-generation matrix is then given by $\mathbf{R} = \Delta t^2 \mathbf{F}(\mathbf{I} - \mathbf{U})^{-1}$ (Bacaër 2009) where $\mathbf{I}$ is an identity matrix of the same dimension as $\mathbf{U}$. The Δt^2 ensures that $\mathbf{R}$ is unitless and is needed because we discretize time. The dominant eigenvalue of $\mathbf{R}$ gives R_0.

The matrix $\mathbf{U}$ is given by the block diagonal matrix

$$\mathbf{U} = \begin{array}{c} \\ \text{host}_1 \\ \text{host}_2 \\ \text{host}_3 \\ \vdots \\ \text{host}_N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{host}_1 & \text{host}_2 & \text{host}_3 & \cdots & \text{host}_N \end{array} \\ \begin{bmatrix} \mathbf{\Gamma} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Gamma} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{\Gamma} \end{bmatrix} \end{array} \tag{S9}$$

where $\mathbf{0}$ is an $n \times n$ and $\mathbf{\Gamma}$ is the $n \times n$ matrix

$$\mathbf{\Gamma} = \begin{array}{c} \\ \tau_0 \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_n \end{array} \begin{array}{c} \begin{array}{ccccc} \tau_0 & \tau_1 & \tau_2 & \cdots & \tau_n \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & \cdots & 1-\gamma\Delta t \\ 1-\gamma\Delta t & 0 & 0 & \cdots & 0 \\ 0 & 1-\gamma\Delta t & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \end{array} \tag{S10}$$

The term Δt is needed because we discretize time into n equally spaced intervals. The term in the upper right-hand corner reflects the periodicity assumption – infected hosts at the end of the movement trajectory τ_n start again at the beginning τ_0 if they do not lose their infection.

We can also compute additional useful metrics, such as individual $R_{0,i}$ and pairwise $R_{0,i\leftrightarrow j}$. We define individual $R_{0,i}$ as the expected number of new infections produced by infected host i over its infected lifetime when density-dependent processes are absent. We define $R_{0,i\leftrightarrow j}$ as the expected number of host i (or j) infections produced by host i (or j) over the sequence host $i \rightarrow$ host $j \rightarrow$ host i (or host $j \rightarrow$ host $i \rightarrow$ host j) when density-dependent processes are absent. Because we assume there is no self-reinfection, a pathogen has to be able to complete this cycle to persist and invade.

We can understand individual $R_{0,i}$ and pairwise $R_{0,i\leftrightarrow j}$ by using a blocked representation of $\mathbf{R}$

$$\mathbf{R} = \begin{array}{c} \\ \text{host}_1 \\ \text{host}_2 \\ \text{host}_3 \\ \vdots \\ \text{host}_N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{host}_1 & \text{host}_2 & \text{host}_3 & \cdots & \text{host}_N \end{array} \\ \begin{bmatrix} \mathbf{0} & \mathbf{R}_{1,2} & \mathbf{R}_{1,3} & \cdots & \mathbf{R}_{1,N} \\ \mathbf{R}_{2,1} & \mathbf{0} & \mathbf{R}_{2,3} & \cdots & \mathbf{R}_{2,N} \\ \mathbf{R}_{3,1} & \mathbf{R}_{3,2} & \mathbf{0} & \cdots & \mathbf{R}_{3,N} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \mathbf{R}_{N,1} & \mathbf{R}_{N,2} & \mathbf{R}_{N,3} & \cdots & \mathbf{0} \end{bmatrix} \end{array} \tag{S11}$$

A useful property of the $\mathbf{R}$ matrix is that $\mathbf{R}_{i,j}$ and $\mathbf{R}_{j,i}$ are independent of any hosts other than i and j. In other words, if we calculated $\mathbf{R}$ with only hosts i and j, then $\mathbf{R}_{i,j}$ and $\mathbf{R}_{j,i}$ would be the same as if we calculated $\mathbf{R}$ with N hosts and extract $\mathbf{R}_{i,j}$ and $\mathbf{R}_{j,i}$. Thus, the dominant eigenvalue of the sub-matrix

$$\begin{bmatrix} \mathbf{0} & \mathbf{R}_{i,j} \\ \mathbf{R}_{j,i} & \mathbf{0} \end{bmatrix}, \tag{S12}$$

is pairwise $R_{0,i\leftrightarrow j}$ and it is robust to the inclusion or exclusion of individual hosts other than i and j. Examining $R_{0,i\leftrightarrow j}$ allows us to understand how particular pairwise interactions contribute to pathogen persistence.

Individual $R_{0,i}$ provides another useful summary of $\mathbf{R}$. When loss of infection rate is low relative to the time period over which movement was tracked, we can approximate $\mathbf{R}$ by taking the dominant eigenvalues of $\mathbf{R}_{i,j}$ and $\mathbf{R}_{j,i}$ ($R_{i,j}$ and $R_{j,i}$, respectively) and defining a matrix $\mathbf{R}_{\text{reduced}}$

$$\mathbf{R}_{\text{reduced}} = \begin{array}{c} \\ \text{host}_1 \\ \text{host}_2 \\ \text{host}_3 \\ \vdots \\ \text{host}_N \end{array} \begin{array}{c} \begin{array}{ccccc} \text{host}_1 & \text{host}_2 & \text{host}_3 & \cdots & \text{host}_N \end{array} \\ \begin{bmatrix} 0 & R_{1,2} & R_{1,3} & \cdots & R_{1,N} \\ R_{2,1} & 0 & R_{2,3} & \cdots & R_{2,N} \\ R_{3,1} & R_{3,2} & 0 & \cdots & R_{3,N} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ R_{N,1} & R_{N,2} & R_{N,3} & \cdots & 0 \end{bmatrix} \end{array} \tag{S13}$$

Summing the columns of $\mathbf{R}_{\text{reduced}}$ yields an approximation to individual-level $R_{0,i}$, defining the average number of infections in other hosts produced over the infected lifetime of host i. This can be a useful metric for identifying hosts that produce many infections. Note, however, that if $R_{0,i} > 1$ for some i this does not mean that the pathogen can invade. Because we assume there is no self-reinfection, other hosts in the population must also produce infections for pathogen invasion. Similarly, examining $R_{i,j}$ and $R_{j,i}$ is useful as it can highlight key asymmetries in pairwise infection risk (e.g., Fig. S6).

We do repeat, however, that $\mathbf{R}_{\text{reduced}}$ only approximates the dynamics of $\mathbf{R}$ when the loss of infection rate is low relative to the length of the movement trajectory, such that once a host is infected it tends to stay infected over an iteration of the movement trajectory. If this is not the case, $\mathbf{R}_{\text{reduced}}$ can significantly underestimate the capacity of a pathogen to invade. To check the robustness of the approximation, we have found it worthwhile to compare the dominant eigenvalues of $\mathbf{R}_{\text{reduced}}$ and $\mathbf{R}$.

7.3 Perturbations and R_0

The number R_0 has important properties when applying MoveSTIR to empirical data. Specifically, elasticities of R_0 (i.e., proportional changes in R_0 due to proportional changes in underlying individual, spatial, or temporal parameters in the model) are invariant to the exact values of acquisition rate $\tilde{\beta}$ and deposition rate λ. Thus, we do not need to know the absolute values of these parameters – which are often difficult to estimate directly – to test the relative contributions of individual hosts, locations, times, and direct versus indirect transmission to pathogen invasion risk.

We can calculate the elasticity of R_0 (or $R_{0,i\leftrightarrow j}$) to lower level parameters in the same way one would with matrix models and integral projection models (IPMs) (Caswell 2001; Ellner *et al.* 2016). Namely, we can directly perturb entries in our transmission matrix or perturb lower-level parameters contributing to each entry in the transmission matrix (e.g., perturb characteristics of the movement trajectory). As with matrix models and IPMs, the possible permutations of sensitivity analysis are enormous.

7.4 A simulated example with five hosts

We explored a simulated example of five hosts moving on a landscape (Fig. S6A) to demonstrate how MoveSTIR can combine the transmission kernel and dynamic epidemiological models to ask prospective questions regarding the role of individual, spatial, and temporal processes on pathogen invasion dynamics. In this example we assumed that infected hosts that lose infection were immediately susceptible, though this could easily be extended to include a **R**ecovered or Exposed class. We used a transmission kernel similar to equation S7 and we chose our rates of deposition (λ), acquisition ($\tilde{\beta}$) and pathogen decay ν such that $R_0 > 1$ (specifically, $R_0 = 1.42$). This meant that our hypothetical pathogen could invade the five host "population" given the observed movement trajectories and the dynamic direct to indirect contact network they defined (Fig. S6B).

In this example, we could identify which individuals were contributing most to pathogen invasion. Host 1, host 4, and host 5 all infected, on average, greater than one other individual over the time period of the movement trajectory, while host 2 and host 3 infect less than one host (Fig. S6C). However, removing host 5 or 4 lead to significantly larger reductions in pathogen invasion potential than removing host 1, 2, or 3 (Fig. S6C). This was because the symmetry in pairwise interactions between host 4 and 5 meant that pairwise $R_{0,4\leftrightarrow 5} = 1.26 > 1$, while the asymmetry in pairwise interactions between host 1 and 2 lead to $R_{0,1\leftrightarrow 2} = 0.5 < 1$.

In addition, we could use MoveSTIR to identify spatial locations contributing to pathogen invasion risk without extensive spatial simulation. Rather, we could simply perturb entries in the transmission kernel

that were within some spatial area of interest A and ask how these perturbations change R_0. Fig. S6D shows a map of the spatial areas that were most important to overall R_0. Consistent with our analysis of pairwise $R_{0,i\leftrightarrow j}$, we saw that the most important area for population-level R_0 occurred where there were interactions between host 4 and host 5 (Fig. S6D). In contrast, areas where hosts other than host 4 and 5 interacted were significantly less important for overall R_0. Spatial elasticity of R_0 can directly identify regions on the landscape that might be optimal management targets for reducing pathogen persistence or linked to resource covariates on the landscape to predict such locations. Using MoveSTIR, we could further identify the sensitivity of R_0 to moments in time, direct and indirect transmission, or any other dimensions and lower-level parameters of the transmission kernel.

8 Appendix S8: Application of MoveSTIR to wild pig movement trajectories

Summary: *This appendix provides the details for how we analyzed the pig data in the main text.*

8.1 Fitting continuous-time movement models to movement trajectories

We used the R package `ctmm` (Calabrese *et al.* 2016) to fit continuous-time movement models to the observed pig movement trajectories. As our purpose here was to demonstrate the type of inference we can make using MoveSTIR given continuous-time movement trajectories, we did not perform any model comparison with the `ctmm` package. Rather, we used the function `ctmm.guess` to generate initial parameters for the CTMM model and fit the model with these initial parameters (see the script `fit_and_predict_ctmm.R`). We then used the fitted CTMM model to interpolate the host trajectories to five minute intervals.

8.2 Analyzing spatio-temporal infection risk of wild pigs from ASFV using MoveSTIR

Given the continuous-time movement trajectories fit using `ctmm` (discretized to five minute intervals), we then applied MoveSTIR to estimate pairwise transmission kernels $K_{a_i\leftarrow d_j}(\tau_a, \tau_d)$ for all pairs of pigs (19×19 - 19 = 342 transmission kernels). To estimate the transmission kernels, we needed the following parameters: the acquisition rate β', the deposition rate λ, the pathogen decay rate ν, and the contact function with parameter α. As discussed in the main text, we used a top-hat contact function (Appendix S1) with distance α between 1-10m and a pathogen decay rate of $\nu = 1/5$ days$^{-1} = 1/(24 \times 60 \times 5)$ minutes^{-1}. As is often the case with wildlife pathogens, we did not have estimates of β' or λ. However, we can still make inference on relative

force of infection among individuals, time, and space using MoveSTIR as β' and λ are multipliers in our transmission kernel (equation 5 in the main text) and cancel when computing relative infection risk. Thus, we set β' and λ to be unity and only made inference on relative infection risk in time, space, and across individuals. All of our analyses can be reproduced using the script `movestir_pig_movements.ipynb`.

8.3 A SIR model for African Swine Fever

To make inference on R_0 for the ASFV-wild pigs system, we defined a SIR model given by

$$\begin{aligned}
\frac{dp_{H_{i,S}}(t)}{dt} &= -p_{H_{i,S}}(t) \sum_{j \in N_{-i}} \int_0^t K'_{a_i \leftarrow d_j}(t,u)du \\
\frac{dp_{H_{i,I}(t)}}{dt} &= p_{H_{i,S}}(t) \sum_{j \in N_{-i}} \int_0^t K'_{a_i \leftarrow d_j}(t,u)du - \theta p_{H_{i,I}}(t) \\
\frac{dp_{H_{i,R}(t)}}{dt} &= \theta p_{H_{i,I}}(t)
\end{aligned} \tag{S14}$$

where θ is the recovery rate and $p_{H_{i,R}(t)}$ is the probability that a host is in the recovered class at time t. Previous models have included an exposed class when modeling ASFV dynamics (e.g., Pepin *et al.* 2020). However, because this exposed class is relatively short (e.g., 4 days; Pepin *et al.* 2020) compared to the life span of a pig (on the order of years), we did not expect this exclusion to significantly affect our calculations of R_0 (Keeling & Rohani 2008). Because recovery rate θ plays the same role as the loss of infection rate γ in equation S8, calculating R_0 for equation S14 is the same as for equation S8, replacing γ with θ. Moreover, because we were interested in relative R_0 values (i.e., ratios of R_0 values), the exact value of θ does not affect any of our conclusions as it cancels out in the ratio.

9 Appendix S9: Home range overlap analyses as a special case of MoveSTIR

Summary: *This section provides a derivation of why home range overlap analyses can be considered as a special case of MoveSTIR. It is not required for understanding the results presented in the main text, but supports statements therein.*

An often-used approach to derive contact networks from movement data relies on defining edge weights between individuals based on some metric of home range overlap (e.g., Godfrey *et al.* 2010; Springer *et al.* 2017; Noonan *et al.* 2021). Home range overlap analyses are a special case of MoveSTIR. Here, we demonstrate the workflow needed to analyze home range overlap within the MoveSTIR framework, referencing the

wild pig movement data we use in the main text.

Area of home range overlap

We assume that we start with movement data (e.g., GPS fixes) to which we can apply previously developed software to estimate home ranges and utilization distributions (UD) (e.g., using the R package `adehabitatHR`, Calenge 2006). For example, for the 19 pigs analyzed in the main text we had GPS fixes every 30 minutes for each individual for up to three months. We calculated the 95% UD of each pig using the `kernelUD` function in the `adehabitatHR` package (Calenge 2006). In our analysis, we compute 95% UDs based on the entire collaring period of the pigs.

We then calculated the home range overlap between two individuals. There are multiple metrics with which to compute overlap (Fieberg & Kochanny 2005; Winner *et al.* 2018). To directly connect to MoveSTIR, we define home range overlap as the area of overlap given by the boundaries of two individual's 95% UDs. We used this approach when calculating the area of overlap for the home ranges of pigs used in the main text. However, we also explored different metrics of home range overlap, such as the Bhattacharyya Coefficient, the volume of intersection of the utilization distributions, and the proportion of home range overlap (Fig. S7).

Given estimated home ranges and home range overlap for each pair of individuals, we then estimated potential contact and transmission. For potential contact to occur in the area of overlap, two events need to happen. First, both individuals need to have previously been in or currently be in the area of overlap. Based on an individual's UD and the assumption that individuals move independently (an assumption which we relax below), the probability that both individual i and j are in the area of overlap AO is $p_{i,AO} \times p_{j,AO}$, where $p_{i,AO} = \int_{m=(x,y)\in AO} UD_i(m)dm$ – the integral of the UD of individual i within the area of overlap, where $m = (x, y)$ are spatial coordinates within the area of overlap. In the simplest case where the UD is uniform we obtain $p_{i,AO} = \frac{\text{Area of overlap}}{\text{Area of 95\% UD for host } i}$. We used the uniform assumption for the pig data in the main text in order to specifically test how ignoring fine-scale movements within an individual's home range could affect emergent contact networks and epidemiological dynamics.

The second event that needs to occur is that current or past individuals within the area of overlap need to come into contact with each other and potentially transmit infection. Assuming random walks within the area of overlap (but not necessarily within the host's home range), increasing the area of overlap makes it less likely that two hosts will contact each other. This is for the same reason described in equation S6 – the acquisition rate $\beta' = \frac{\tilde{\beta}}{\text{Area of overlap}}$ scales inversely with the area of overlap.

We now have all the information needed to apply MoveSTIR to our analysis of home range overlap.

Conceptually, we can think of our analysis as a discrete-space metapopulation model (as in equation S6) where contact and transmission can only occur when both hosts are in the area of overlap defined by the home ranges. We can write our force of infection equation as

$$h_{i\leftarrow j}(t, x_i) = \int_{-\infty}^{t} \frac{\tilde{\beta}}{A_{x_{\text{Area of overlap}}}} \lambda \delta_{x_i(t)}(x_{\text{Area of overlap}}) \delta_{x_j(u)}(x_{\text{Area of overlap}}) \delta_{I_j(u)}(I) e^{-\nu(t-u)} du \tag{S15}$$

where $\delta_{x_i(t)}(x_{\text{Area of overlap}})\delta_{x_j(u)}(x_{\text{Area of overlap}})$ are two indicator variables that ensure both host i at time t and host j at time u are in the area of overlap such that transmission can potentially occur. The spatial trajectories $x_i(t)$ and $x_j(t)$ alternate between two locations: the area of overlap and the remaining area of the home range for the individual not contained in the area of overlap. The term $A_{x_{\text{Area of overlap}}}$ gives the area measure (e.g., in m^2) of the area of overlap $x_{\text{Area of overlap}}$.

Individual movement occurs probabilistically between the area of overlap and the rest of the home range and we assume that it has obtained a stationary distribution. The variables $\delta_{x_i(t)}(x_{\text{Area of overlap}})$ and $\delta_{x_j(u)}(x_{\text{Area of overlap}})$ are random variables that take the values 1 (host in area of overlap) or 0 (host not in area of overlap) and given a stationary assumption are independent of t and u. In this situation, taking the expectation of equation S15 with respect to time leads to

$$\bar{h}_{i\leftarrow j} = \left(\frac{\tilde{\beta}}{A_{x_{\text{Area of overlap}}}} \right) \left(\frac{\lambda}{\nu} \right) \left[p_{i,AO} p_{j,AO} + \text{Cov}(\delta_{x_i}(x_{\text{Area of overlap}}), \delta_{x_j}(x_{\text{Area of overlap}})) \right] \tag{S16}$$

where we set $\delta_{I_j(u)}(I) = 1$. The variables $p_{i,AO}$ and $p_{j,AO}$ are the probabilities that i and j are in the area of overlap, respectively. The term $\text{Cov}(\delta_{x_i}(x_{\text{Area of overlap}}), \delta_{x_j}(x_{\text{Area of overlap}}))$ accounts for the covariance in the use of the area of overlap and is zero if hosts move independently. Thus, using the area of home range overlap and MoveSTIR we can directly compute the edge weights for a static contact network defining maximum potential infection risk.

References

Bacaër, N. (2009). Periodic matrix population models: Growth rate, basic reproduction number, and entropy. *Bulletin of Mathematical Biology*, 71, 1781–1792.

Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E. & Boutin,

S. (2015). Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. *Journal of Applied Ecology*, 52, 675–685.

Calabrese, J. M., Fleming, C. H. & Gurarie, E. (2016). CTMM: an R Package for analyzing animal relocation data as a continuous-time stochastic process. *Methods in Ecology and Evolution*, 7, 1124–1132.

Calenge, C. (2006). The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. *Ecological Modelling*, 197, 1035.

Caswell, H. (2001). *Matrix Population Models: Construction, Analysis, and Interpretation*. 2nd edn. Sinauer, Sunderland, Massachusetts.

Cayuela, H., Griffiths, R. A., Zakaria, N., Arntzen, J. W., Priol, P., Léna, J., Besnard, A. & Joly, P. (2020). Drivers of amphibian population dynamics and asynchrony at local and regional scales. *Journal of Animal Ecology*, 89, 1350–1364.

Cayuela, H., Pradel, R., Joly, P. & Besnard, A. (2017). Analysing movement behaviour and dynamic space-use strategies among habitats using multi-event capture-recapture modelling. *Methods in Ecology and Evolution*, 8, 1124–1132.

Cross, P. C., Lloyd-Smith, J. O., Johnson, P. L. F. & Getz, W. M. (2005). Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. *Ecology Letters*, 8, 587–595.

Diekmann, O., Heesterbeek, J. A. P. & Britton, T. (2013). *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton University Press, Princeton.

Drewe, J. A., Weber, N., Carter, S. P., Bearhop, S., Harrison, X. A., Dall, S. R., McDonald, R. A. & Delahay, R. J. (2012). Performance of proximity loggers in recording intra- and inter-species interactions: A laboratory and field-based validation study. *PLoS ONE*, 7, e39068.

Edelhoff, H., Signer, J. & Balkenhol, N. (2016). Path segmentation for beginners: An overview of current methods for detecting changes in animal movement patterns. *Movement Ecology*, 4. URL `http://dx.doi.org/10.1186/s40462-016-0086-5`.

Ellner, S. P., Childs, D. Z. & Rees, M. (2016). *Data-driven Modelling of Structured Populations*. Springer International Publishing, Swizterland. ISBN 9783319288918.

Fieberg, J. & Kochanny, C. O. (2005). Quantifying home-range overlap: the importance of the utilization distribution. *Journal of Wildlife Management*, 69, 1346–1359.

Fine, A. E., Bolin, C. A., Gardiner, J. C. & Kaneene, J. B. (2011). A study of the persistence of Mycobacterium bovis in the environment under Natural weather conditions in Michigan, USA. *Veterinary Medicine International*, 2011, 1–12. URL `https://www.hindawi.com/journals/vmi/2011/765430/`.

Godfrey, S. S. (2013). Networks and the ecology of parasite transmission: A framework for wildlife parasitology. *International Journal for Parasitology: Parasites and Wildlife*, 2, 235–245. URL `http://dx.doi.org/10.1016/j.ijppaw.2013.09.001`.

Godfrey, S. S., Moore, J. A., Nelson, N. J. & Bull, C. M. (2010). Social network structure and parasite infection patterns in a territorial reptile, the tuatara (*Sphenodon punctatus*). *International Journal for Parasitology*, 40, 1575–1585.

Gurarie, E. & Ovaskainen, O. (2013). Towards a general formalization of encounter rates in ecology. *Theoretical Ecology*, 6, 189–202.

Hamel, S., Killengreen, S. T., Henden, J. A., Eide, N. E., Roed-Eriksen, L., Ims, R. A. & Yoccoz, N. G. (2013). Towards good practice guidance in using camera-traps in ecology: Influence of sampling design on validity of ecological inferences. *Methods in Ecology and Evolution*, 4, 105–113.

Hayward, A. D., Garnier, R., Childs, D. Z., Grenfell, B. T., Watt, K. A., Pilkington, J. G., Pemberton, J. M. & Graham, A. L. (2019). From population to individual host scale and back again: testing theories of infection and defence in the Soay sheep of St Kilda. In: *Wildlife Conservation in a Changing Climate* (eds. Wilson, K. R., Fenton, A. & Tompkins, D. M.), chap. 4. Cambridge University Press, Cambridge.

Keeling, M. & Rohani, P. (2008). *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, Princeton, New Jersey.

Lavelle, M. J., Kay, S. L., Pepin, K. M., Grear, D. A., Campa, H. & VerCauteren, K. C. (2016). Evaluating wildlife-cattle contact rates to improve the understanding of dynamics of bovine tuberculosis transmission in Michigan, USA. *Preventive Veterinary Medicine*, 135, 28–36.

Leitch, J., Alexander, K. A. & Sengupta, S. (2019). Toward epidemic thresholds on temporal networks: a review and open questions. *Applied Network Science*, 4.

Martinez-Garcia, R., Fleming, C. H., Seppelt, R., Fagan, W. F. & Calabrese, J. M. (2020). How range residency and long-range perception change encounter rates. *Journal of Theoretical Biology*, 498, 110267.

Merow, C., Latimer, A. M., Wilson, A. M., McMahon, S. M., Rebelo, A. G. & Silander, J. A. (2014). On using integral projection models to generate demographically driven predictions of species' distributions:

development and validation using sparse data. *Ecography*, 37, 1167–1183. URL http://doi.wiley.com/10.1111/ecog.00839.

Noonan, M. J., Martinez-Garcia, R., Davis, G. H., Crofoot, M. C., Kays, R., Hirsch, B. T., Caillaud, D., Payne, E., Sih, A., Sinn, D. L., Spiegel, O., Fagan, W. F., Fleming, C. H. & Calabrese, J. M. (2021). Estimating encounter location distributions from animal tracking data. *Methods in Ecology and Evolution*, 2021, 1–16.

Pepin, K. M., Golnar, A. J., Abdo, Z. & Podgórski, T. (2020). Ecological drivers of African swine fever virus persistence in wild boar populations: insight for control. *Ecology and Evolution*, 10, 2846–2859.

Royle, J. A., Chandler, R. B., Sollmann, R. & Gardner, B. (2014). *Spatial Capture-Recapture*. Elsevier, Oxford.

Royle, J. A., Karanth, K. U., Gopalaswamy, A. M. & Kumar, N. S. (2009). Bayesian inference in camera trapping studies for a class of spatial capture-recapture models. *Ecology*, 90, 3233–3244.

Royle, J. A. & Young, K. V. (2008). A hierarchical model for spatial capture-recapture data. *Ecology*, 92, 526–528.

Saunders, S. E., Bartz, J. C. & Bartelt-Hunt, S. L. (2012). Soil-mediated prion transmission: Is local soil-type a key determinant of prion disease incidence? *Chemosphere*, 87, 661–667. URL http://dx.doi.org/10.1016/j.chemosphere.2011.12.076.

Silk, M. J., Drewe, J. A., Delahay, R. J., Weber, N., Steward, L. C., Wilson-Aggarwal, J., Boots, M., Hodgson, D. J., Croft, D. P. & McDonald, R. A. (2018). Quantifying direct and indirect contacts for the potential transmission of infection between species using a multilayer contact network. *Behaviour*, 155, 731–757.

Silk, M. J., McDonald, R. A., Delahay, R. J., Padfield, D. & Hodgson, D. J. (2021). CMRnet: An r package to derive networks of social interactions and movement from mark–recapture data. *Methods in Ecology and Evolution*, 12, 70–75.

Springer, A., Kappeler, P. M. & Nunn, C. L. (2017). Dynamic vs. static social networks in models of parasite transmission: predicting *Cryptosporidium* spread in wild lemurs. *Journal of Animal Ecology*, 86, 419–433.

Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Colizza, V., Isella, L., Régis, C., Pinton, J.-f., Khanafer, N., Broeck, V. D. & Vanhems, P. (2011a). Simulation of a SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Medicine*, 9, 1–4.

Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J. F., Quaggiotto, M., van den Broeck, W., Régis, C., Lina, B. & Vanhems, P. (2011b). High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS ONE*, 6, e23176.

Valdano, E., Ferreri, L., Poletto, C. & Colizza, V. (2015). Analytical computation of the epidemic threshold on temporal networks. *Physical Review X*, 5, 1–9.

Weinstein, S. B., Moura, C. W., Mendez, J. F. & Lafferty, K. D. (2018). Fear of feces? Tradeoffs between disease risk and foraging drive animal activity around raccoon latrines. *Oikos*, 127, 927–934.

Wilber, M. Q., Johnson, P. T. J. & Briggs, C. J. (2020). Disease hotspots or hot species? Infection dynamics in multi-host metacommunities controlled by species identity, not source location. *Ecology Letters*, 23, 1201–1211.

Wilber, M. Q., Pepin, K. M., Campa Iii, H., Hygnstrom, S. E., Lavelle, M. J., Xifara, T., Vercauteren, K. C. & Webb, C. T. (2019). Modeling multi-species and multi-mode contact networks: implications for persistence of bovine tuberculosis at the wildlife-livestock interface. *Journal of Applied Ecology*, 56, 1471–1481.

Winner, K., Noonan, M. J., Fleming, C. H., Olson, K. A., Mueller, T., Sheldon, D. & Calabrese, J. M. (2018). Statistical inference for home range overlap. *Methods in Ecology and Evolution*, 9, 1679–1691.

Yang, A., Schlichting, P., Wight, B., Anderson, W. M., Chinn, S. M., Wilber, M. Q., Miller, R. S., Beasley, J. C., Boughton, R. K., Vercauteren, K. C., Wittemyer, G., Pepin, K. M. & Yang, A. (2020). Effects of social structure and management on risk of disease establishment in wild pigs. *Journal of Animal Ecology*, 1–14.

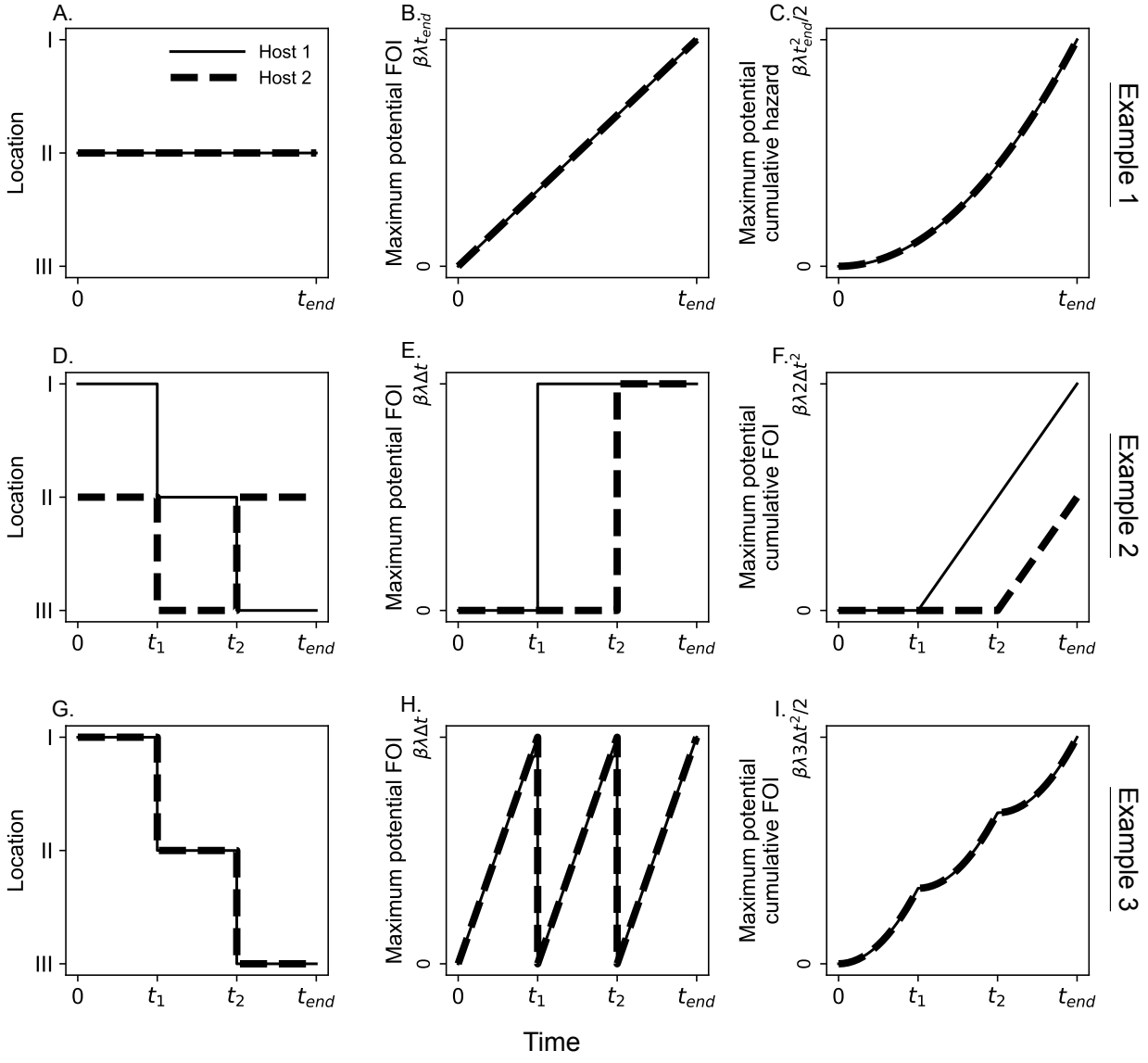

Figure S1: Three examples of how the force of infection (FOI) experienced by host 1 from host 2 (and vice versa) and the cumulative force of infection can be inferred from host movement trajectories using the transmission kernel $K_{a_2 \leftarrow d_1}(\tau_a, \tau_d)$. All examples assume that there is no pathogen decay. **A.-C.** Example 1 assumes that hosts 1 and 2 are in the same location (location II) from time $t_{\text{start}} = 0$ to t_{end}. **D.-F.** Example 2 assumes that hosts are never in the same place at the same time. **G.-I.** Example 3 assumes that hosts are always in the same location as each other, but are moving to different locations through time. For all plots, the interval Δt is the same between 0 to t_1, t_1 to t_2, and t_2 to t_{end}.

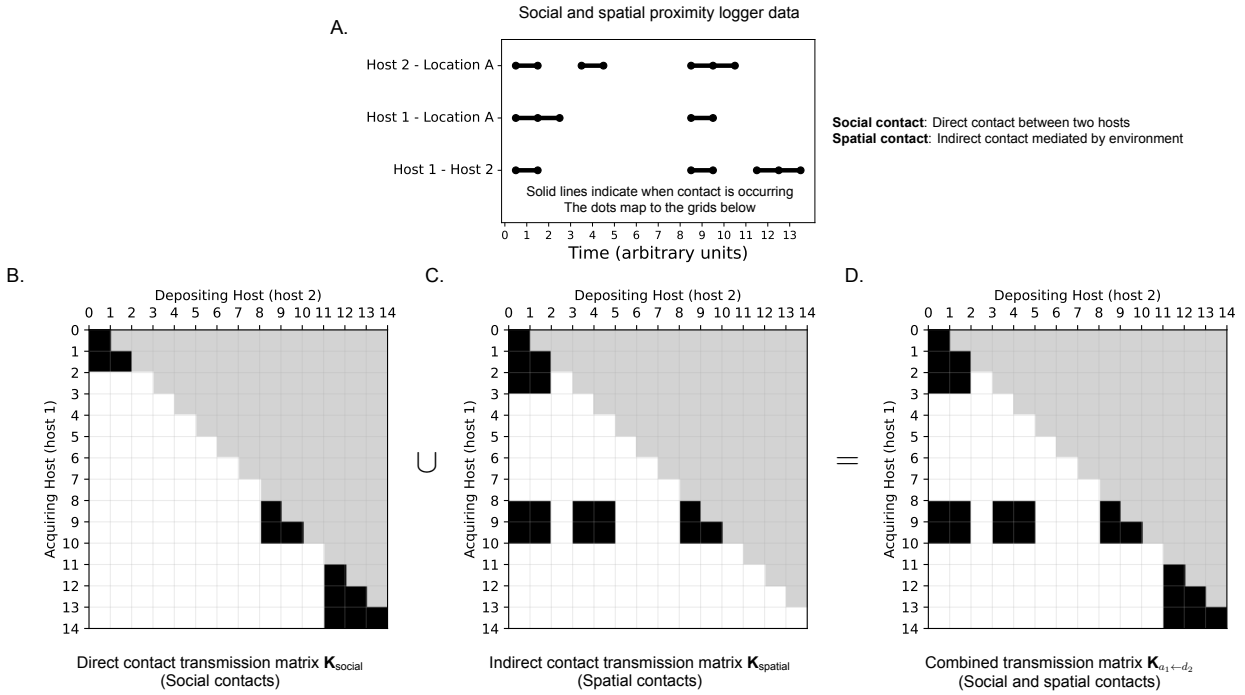

Figure S2: Example of how proximity logger data can be represented as a transmission kernel (or a discretized transmission matrix). This means that proximity logger data can be used directly within the MoveSTIR framework, regardless of whether there is spatial information associated with the proximity logger data. **A.** A visual representation of proximity logger data that contains information on social and spatial contacts. Social contacts occur when two hosts are in direct contact (e.g., Host 1 - Host 2). The solid lines indicate when Host 1 and Host 2 are in proximity of each other and are experiencing a direct contact. Typically, the spatial location of these contacts is unknown. We just know when the contact starts and how long it lasts (the length of the black lines). Proximity logger data can also provide information on spatial contacts. If loggers are placed on both hosts and specific locations of interest (e.g., Location A), then we obtain information on when hosts contact specific locations and for how long (e.g., Host 1 - Location A). This allows for specification of indirect, spatial contacts occurring between two hosts. **B.-D.** We can represent the proximity logger data shown in A. as a transmission matrix. We first consider social contacts (B.), which are given along the diagonal of the transmission matrix (black grids indicate contact, white grids indicate no contact, and the shaded upper triangle indicates that the current acquiring host cannot contact the future trajectory of the depositing host). Here, we assume that over the duration when hosts experience a social contact they remain in the same location, though we do not know where that location is. This assumption means that there is also the potential for indirect contact to occur over the duration of the direct contact, leading to the filled off-diagonals in B. We can also consider spatial contacts that occur when individuals are in the same place at the same or different times (C., contacts are shown by black grid cells). For example, when host 1 is in location A at time 8, it experiences indirect contact with pathogen deposited by host 2 when host 2 was in location A at time 0 and 1. Taking the union (logical "OR") of plots A. and B. yields the transmission matrix in C. that accounts for both social and spatial contacts. For clarity, we do not account for the decay of the pathogen in the environment in this example.

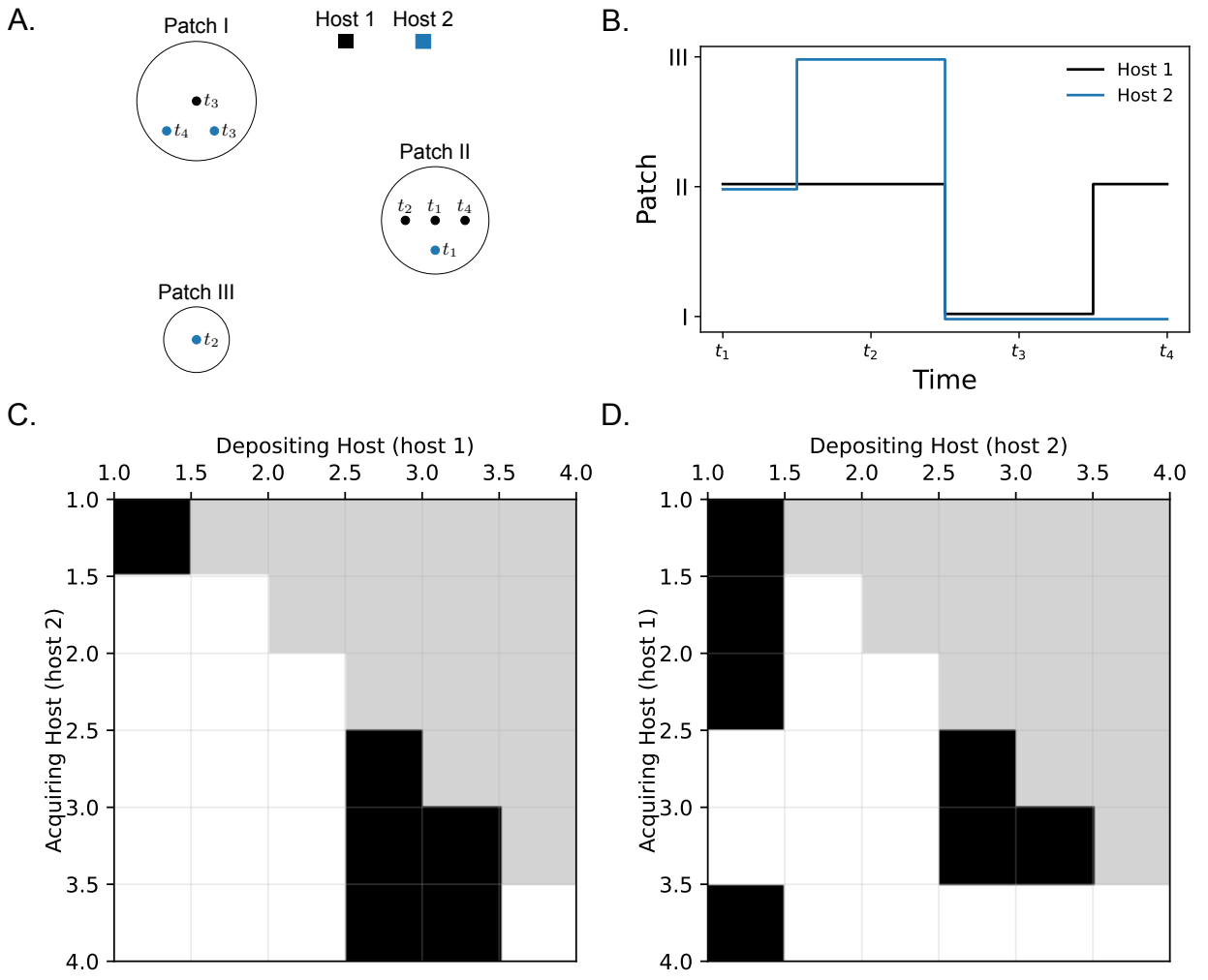

Figure S3: An illustration of how to convert spatially explicit capture-recapture (SECR) data into a transmission kernel. **A.** A spatial representation of SECR data from coarse spatial locations (i.e., habitat patches). Marked hosts are recaptured in distinct habitat patches (e.g., ponds) at relatively coarse time resolutions. For example, the time between primary periods (and thus potential captures) might be weeks or months. While we show points within a patch in different locations, this is strictly for visual purposes. **B.** We can assign a simple continuous-time representation to the SECR data by allowing the spatio-temporal dynamics to follow a step function. Here, we use a midpoint approximation where we assume that if a host moves patches between primary period t_1 and t_2, for example, it does so at the midpoint of the time interval. In B., we slightly stagger the colored lines for Host 1 and Host 2 so they can be visualized. **C.-D.** Once we have the continuous time-discrete space representation of our SECR data, we can convert the data into transmission kernels. Black grids show intervals when and where potential contacts occurred and white grids show when and where contacts did not occur. Gray grids indicate where direct and indirect contact could not possibly occur because the acquiring host cannot contact the future trajectory of the depositing host. Plots C. and D. show the discretized transmission matrices (discretized every 0.5 units of a time step) for host 2 acquiring infection from host 1 (C.) and for host 1 acquiring infection from host 2. For example, in D. we see that when host 1 returns to location I near time 4, it can still experience an indirect contact from host 2 that potentially deposited pathogen in location I at time 1. For simplicity, we do not account for pathogen decay in this example.

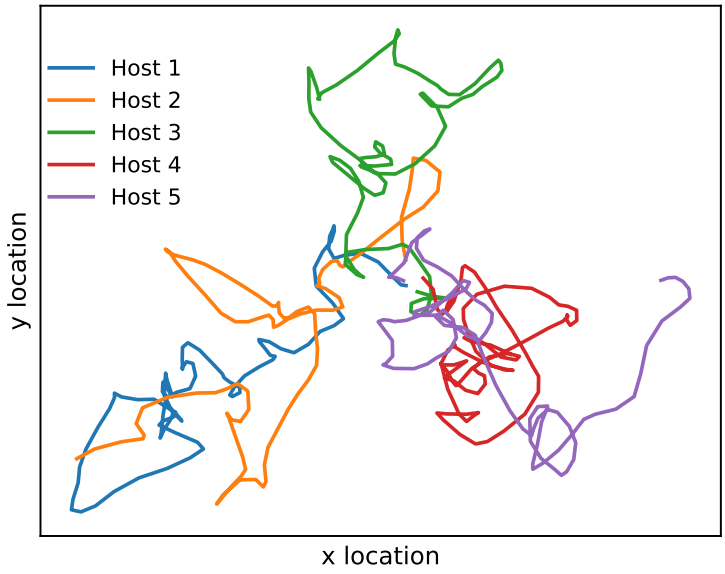

Figure S4: Simulated trajectories of five hosts moving on a landscape. We use these simulated data in Fig. S5.

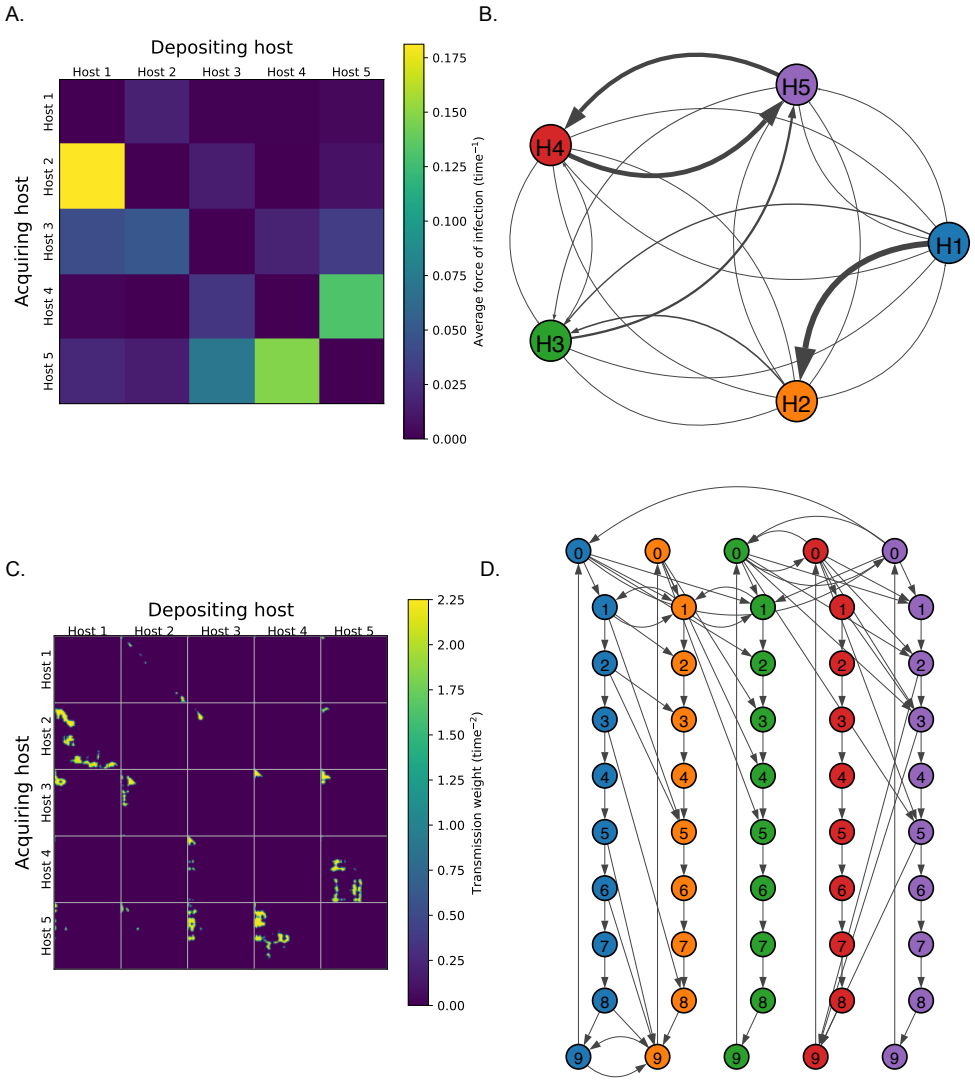

Figure S5: **A.** The static representation of the host movement trajectories shown in Fig. S4. The colors of the grid cells are the average force of infection experience by host i from host j over the time period of the movement trajectory. **B.** A static, weighted network representation of the matrix shown in A. The thickness of the edges represent the directional force of infection shown in A. H1 = host 1, H2 = host 2, H3 = host 3, H4 = host 4, and H5 = host 5 as shown in Fig. S4. **C.** Each grid cell shows a pairwise transmission kernel of the host movement trajectories shown in Fig. S4. A. can be recovered by integrating each grid cell in C. over the acquisition and deposition dimensions and dividing by the length of the movement trajectory. This yields the average force of infection felt by host i from host j over the length of the movement trajectory (see Table 1 in the main text for calculation). **D.** A simplified dynamic network representation of the transmission kernel shown in C. Each colored node represents a host as shown in B., and numbers represent a discrete time interval along the movement trajectory (ten total intervals). Arrows indicate how hosts at one time point contribute to the force of infection of other hosts at different time points. Edges between nodes of the same color indicate that once infected, an infected host has the potential to remain infected through time. Finally, the edge from 9 to 0 indicates the periodicity assumption necessary to compute R_0 (see Appendix S4; Valdano *et al.* 2015) – at the end of the trajectory, hosts start again at the beginning.

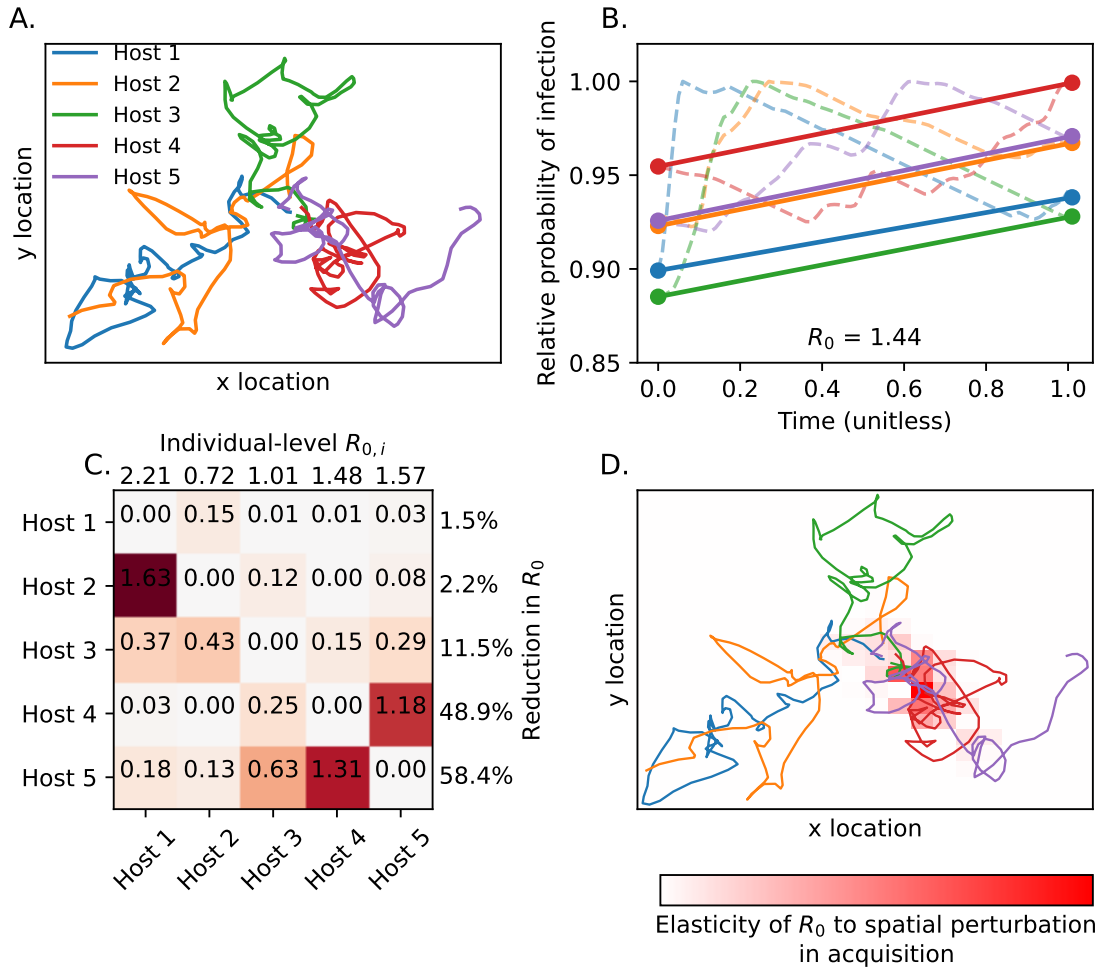

Figure S6: **A.** Simulated movement of five hosts on a landscape. **B.** An example simulation of infection dynamics happening on the observed movement trajectories in A. The simulation uses the individual-level epidemiological model described in Appendix S7. Because $R_0 > 1$, during the growth phase of the pathogen in the population we see an increase in the probability that a host is infected from the start to the end of the movement trajectory (straight colored lines). However, the probability that a host is infected may increase or decrease over this time period depending on movement (colored dashed line). For this example, the transmission kernel had similar form given in equation S7. We let acquisition rate $\tilde{\beta} = 1.5 \frac{\text{spatial units}}{\text{time}}$, $\lambda = 1.5$ time^{-1}, pathogen decay rate $\nu = 0.1$ time^{-1}, and recovery rate $\gamma = 0.11$ time^{-1}. The contact function $\Phi(\mathbf{s}_j(\tau_d), \mathbf{s}_i(\tau_a))$ follows a top-hat function where the acquiring host must be within 0.71 units of the past or present depositing host for a contact to occur. **C.** Each matrix entry describes the average number of infected hosts of type i (rows) produced by host j (columns) over host j's time infected. The numbers above the columns give the column sums which are individual-level $R_{0,i}$: the average number of hosts of all types infected by host j over its time infected (see Appendix S7 for calculations). The numbers to the right give the percent reduction in R_0 when the given individual is removed from the landscape. **D.** The elasticity of R_0 to perturbations of the transmission kernel in different spatial locations on the landscape. For each grid cell in the plot, we perturbed the acquisition experienced by all hosts at any time in this area by $\delta = 0.001$ and re-calculated $R_{0,\text{perturbed}}$. We calculated elasticity as $\frac{R_{0,\text{perturbed}} - R_0}{\delta R_0}$ (Merow *et al.* 2014). Higher values (darker red colors) correspond with a larger proportional change in R_0 given changes in force of infection in the focal area.

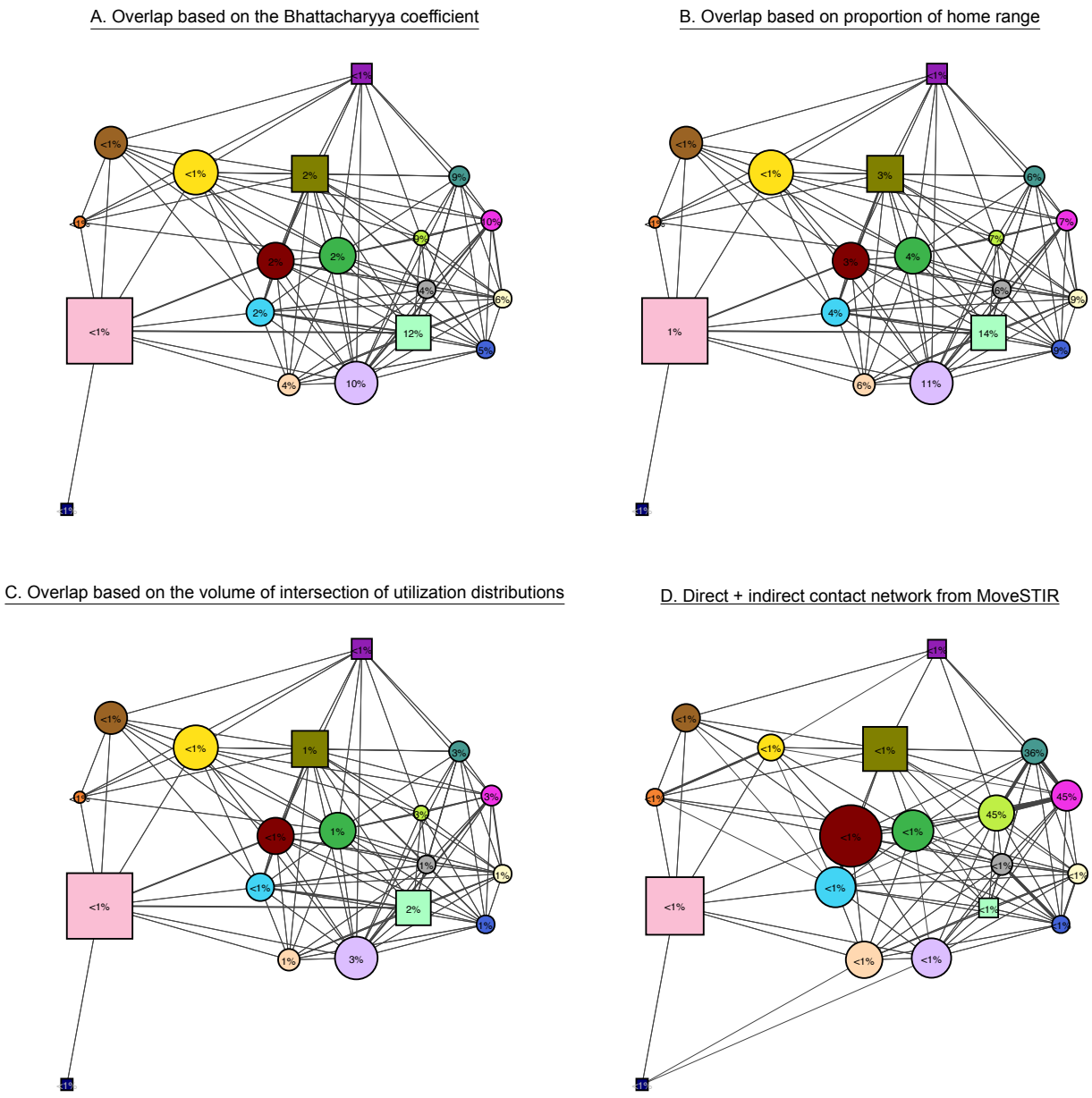

Figure S7: **A.-C.** Contact networks built using different metrics of home range overlap. For all metrics used in A.-C., the unweighted network structures are identical. However, different overlap metrics put different weights on the edges between nodes, where nodes represent individual pigs. All of these metrics of home range overlap are defined in the function `kerneloverlaphr` in the R package `adehabitatHR` (A. BA, B. HR, and C. VI Calenge 2006). **D.** The direct and indirect contact network predicted by MoveSTIR and shown in Figure 5C in the main text. The numbers within each node indicate how much removing that individual pig reduces R_0 defined by the network. Regardless of the metric of home range overlap used in A.-C., accounting for fine-scale host movements and contacts with MoveSTIR (D.) led to significantly different predictions regarding how individual pigs contributed to pathogen invasion (as seen by proportional changes in R_0). Note that unlike in the main text, we could not directly compare overall R_0 among A.-D. as the units were not the same across the different overlap metrics.