Supplementary Information 3: Simulations

Clumping of individuals

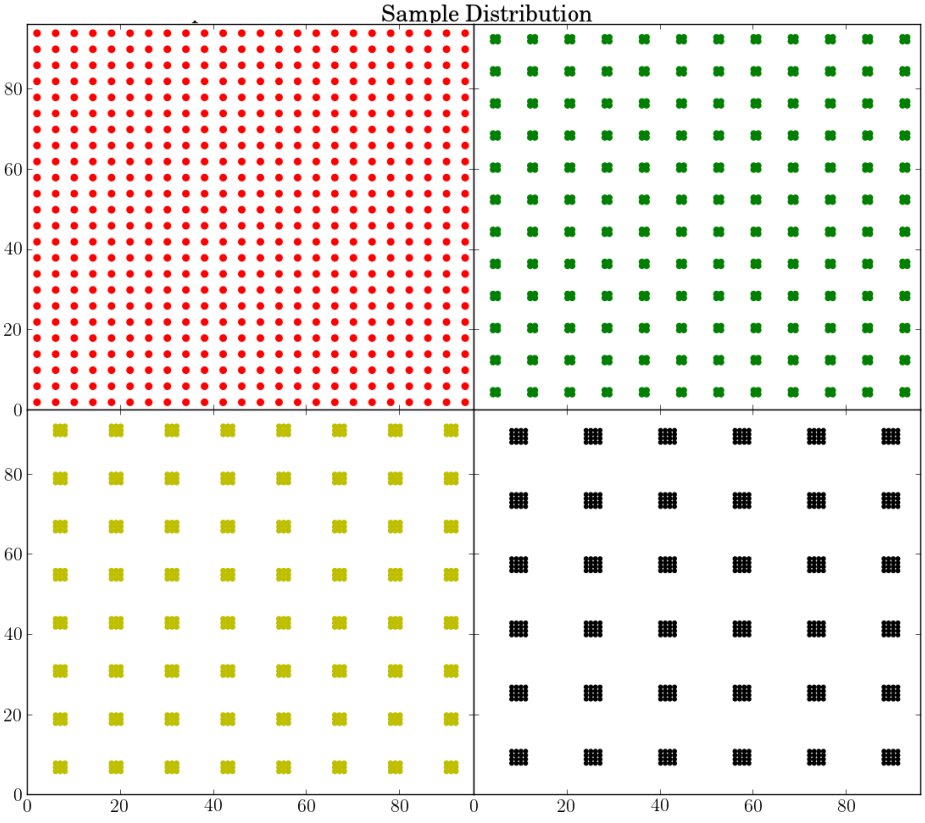
To assess how much our inference method is affected by clumping, we simulated four scenarios of two different kinds of clumping. Simulations were always done on a torus grid with axial length 96, with Laplace dispersal along each axis, dispersal distance set such that $\sigma = 1$ and assuming a constant population size.

Regular Clumping

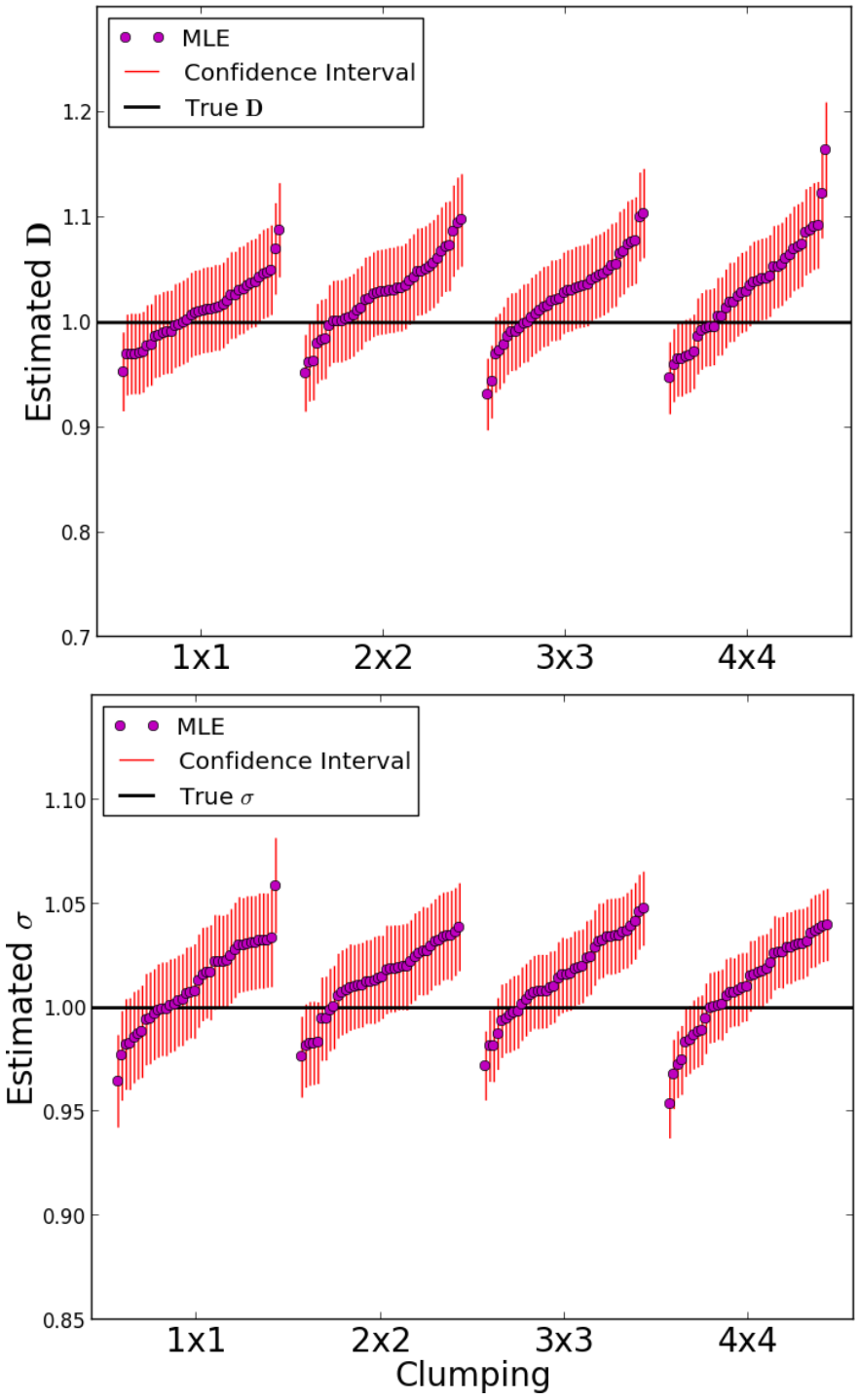
First, in four scenarios, a total number of 576 samples (150 cM chromosomes) were clustered on regular grids:

- Scenario 1: Samples are evenly spaced 4 distance units apart.
- Scenario 2: Samples are clustered into 2×2 clusters 8 distance units apart (along each axis).
- Scenario 3: Samples are clustered into 3×3 clusters 12 distance units apart (along each axis).
- Scenario 4: Samples are clustered into 4×4 clusters 16 distance units apart (along each axis).

This sampling scheme is visualized in the following picture:



We simulated block sharing until $t = 200$ generations back. Our inference scheme applied to blocks 4–20 cM simulated under 20 replicates of each of these scenarios yielded the following parameter estimates and 95% confidence intervals:



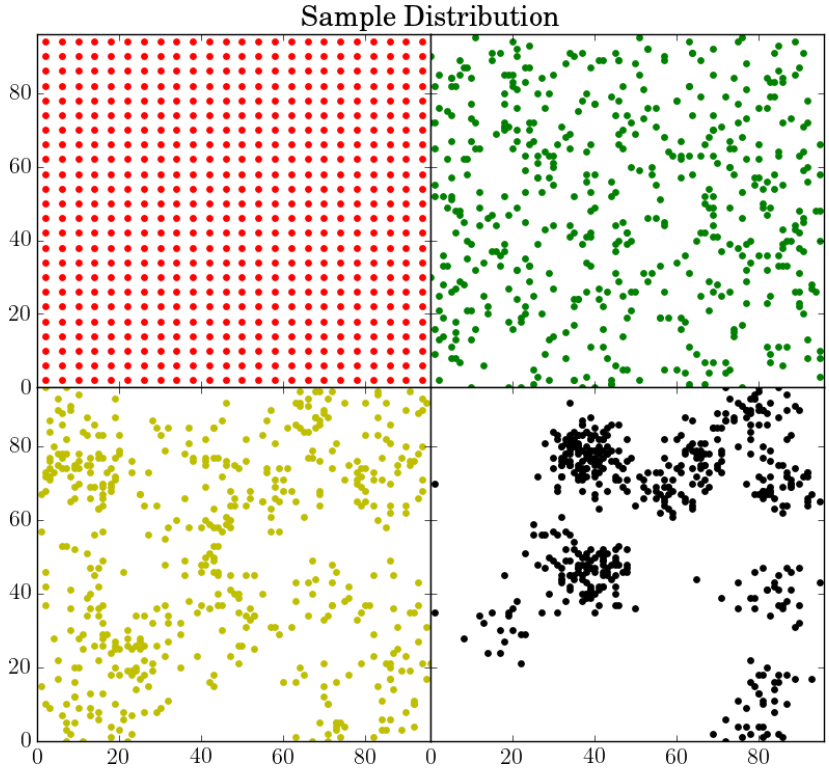
The inference method was very robust with respect to clumping, neither the bias nor the uncertainty of estimates differed substantially between the sampling scenarios considered here.

Irregular Clumping

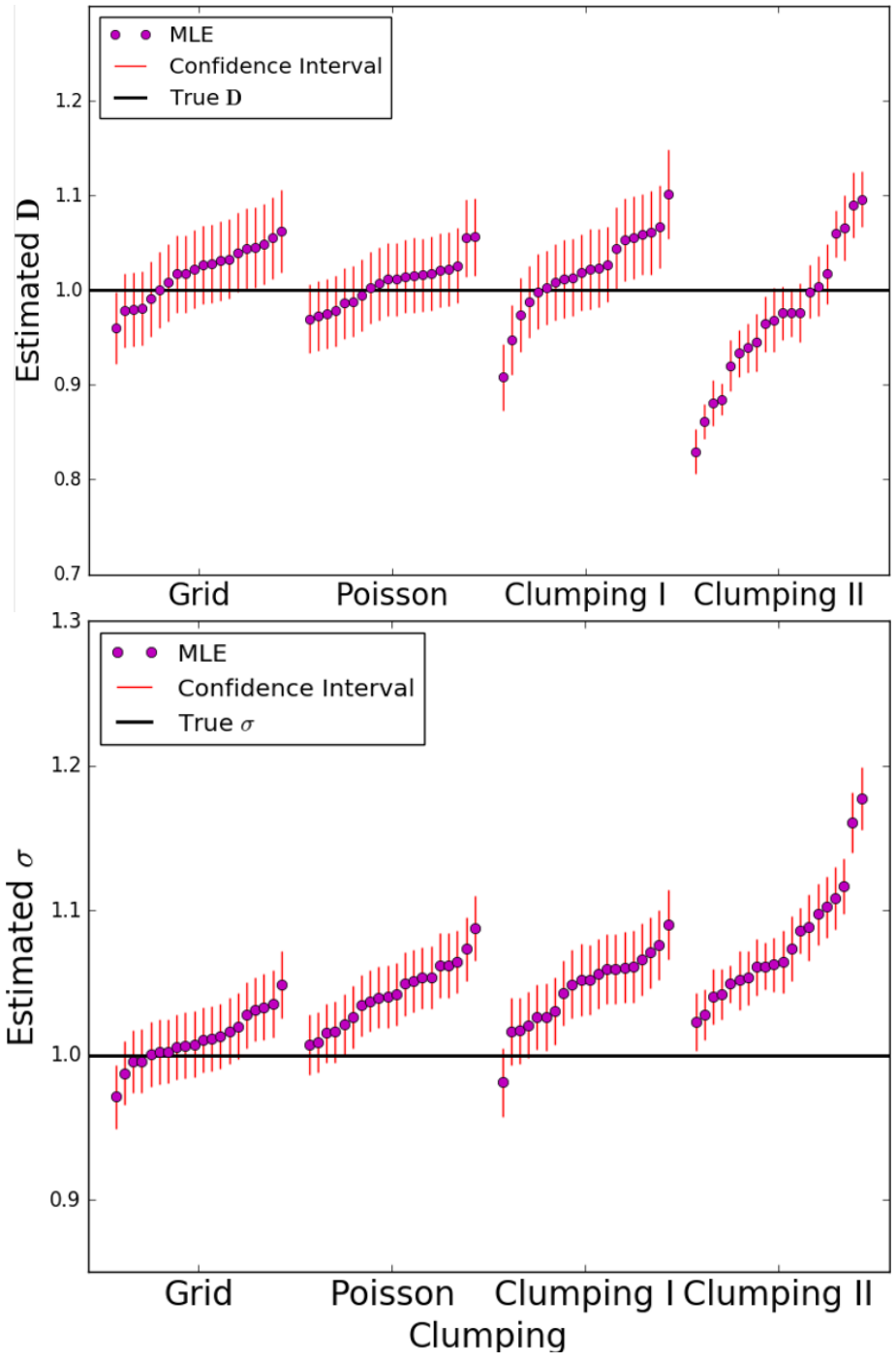
We then also assessed the effect of more irregular and asymmetric clumping. In four scenarios, a total number of 576 samples (150 cM chromosomes) were clustered to different degrees:

- Scenario 1: Samples are evenly spaced 4 distance units apart on a regular grid.
- Scenario 2: Every sample is independently drawn from all grid positions.
- Scenario 3: Randomly clustered samples are simulated in three steps. First, centers of clusters are drawn at random. Then, for every center a random number is drawn from a geometric distribution with mean 5 to determine the number of samples per cluster. Last, for each individual sample, the offset by the mean is determined by a discretized Gaussian with standard deviation $\sigma = 5$ along each axis. Individual samples are drawn until the total number of 576 samples is reached.
- Scenario 4: Same as in Scenario 3, but now the number of samples in every cluster is geometrically distributed with mean 50.

Single realizations of the random clustering schemes described above are visualized in the following picture:



For each scenarios, we did 20 replicates. In Scenario $2-4$, the sample positions were generated independently for every replicate, and in each case we simulated block sharing until $t = 200$ generations back. Our inference scheme applied to blocks of length 4–20 cM yielded the following parameter estimates and 95% confidence intervals:



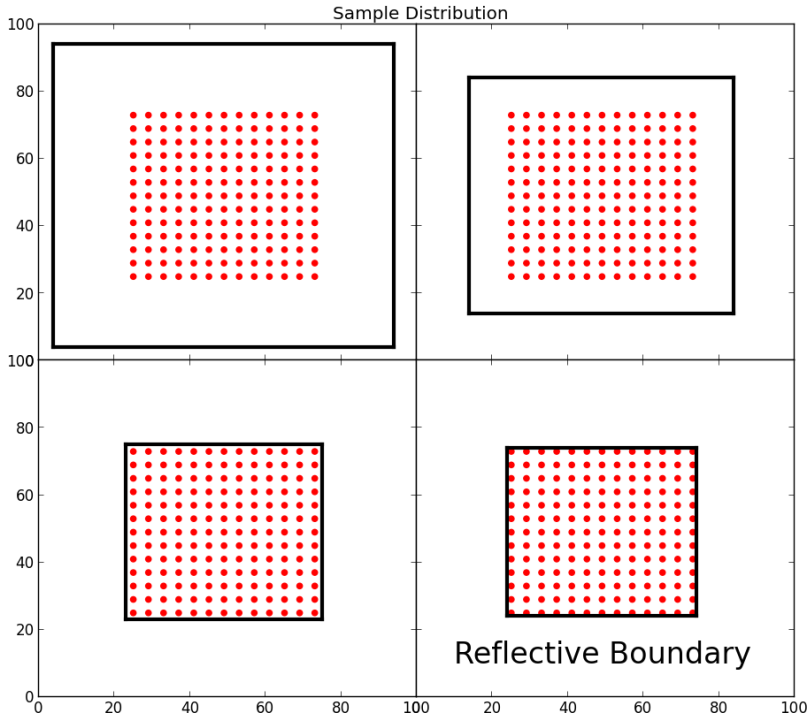
Overall, irregular clumping did not severely affect estimates. However, dispersal estimates became slightly biased upwards with increasing degree of clumping. This weak effect is likely due to clusters of individuals on very small spatial scales; since for geographically close samples the diffusion approximation is not expected to accurately predict block sharing for the here simulated leptokurtic single generation dispersal kernel. This hypothesis is supported by the fact that the bias vanishes if spatial size of the sample clusters is increased (data not shown).

Edge effects

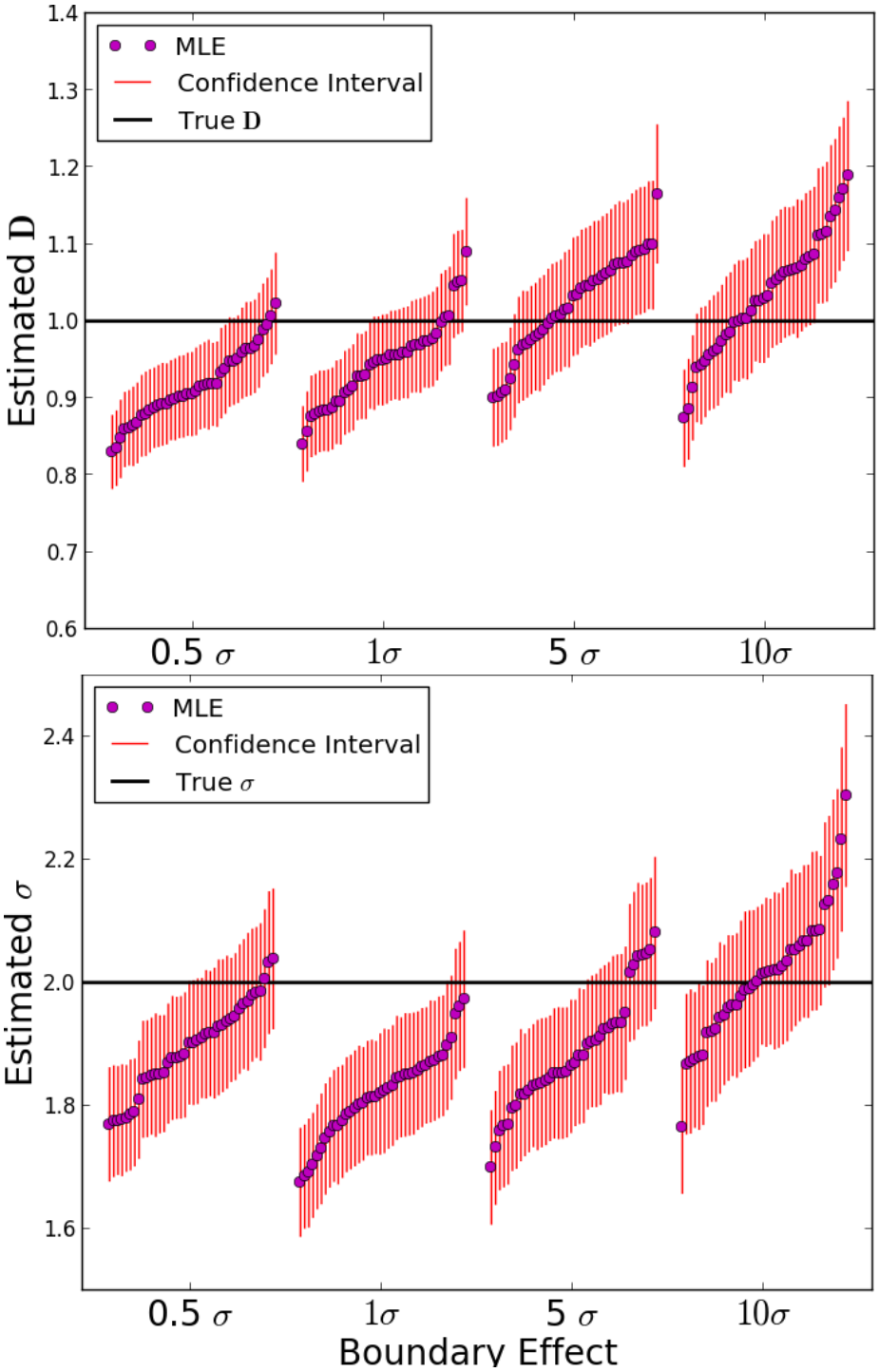
To assess the effect of nearby habitat edges on our inference method, we simulated finite rectangular habitats. For simulations, we utilized our node model, as above with axial Laplace dispersal such that $\sigma = 2$. We assumed that lineages get reflected each time they would trace back beyond an edge. That is, if in a single generation back a lineage would have migrated a certain distance beyond the habitat boundary, that lineage moves to a location an equally far distance from the boundary, but inside the habitat. We simulated four scenarios, where reflective boundaries were surrounding a 12×12 array of samples (150 cM chromosomes) spaced two σ apart.

- Scenario 1: The reflective boundary was at a distance 10 σ.
- Scenario 2: The reflective boundary was at a distance 5 σ.
- Scenario 3: The reflective boundary was at a distance 1 σ.
- Scenario 4: The reflective boundary was at a distance 0.5 σ.

This sampling scheme is illustrated in the following figure:



We simulated block sharing until $t = 200$ generations back. Our inference scheme applied to blocks 4–20 cM simulated under 20 replicates of each of these scenarios yielded the following parameter estimates and 95% confidence intervals:



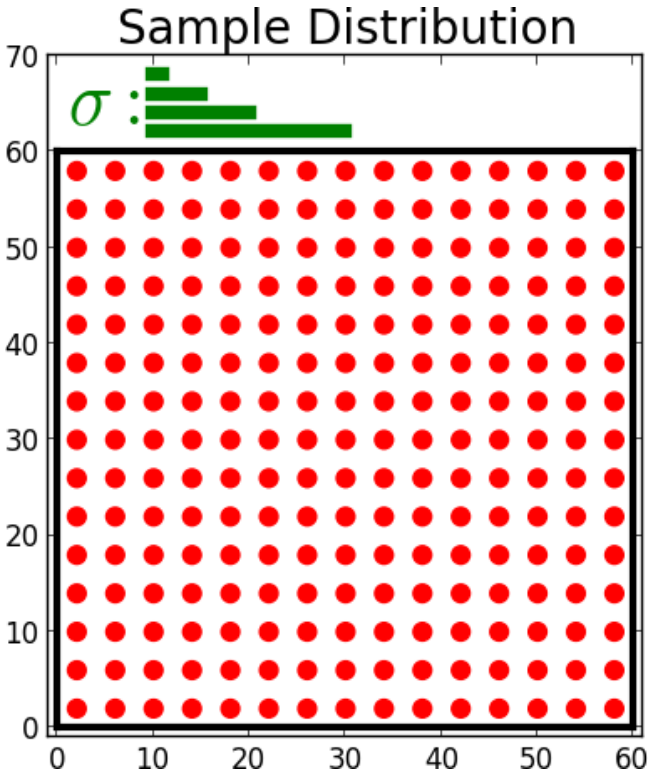
Both dispersal and density estimates got noticeably biased downwards as soon as range boundaries got closer, but remained right on orders of magnitude.

Limited habitat size

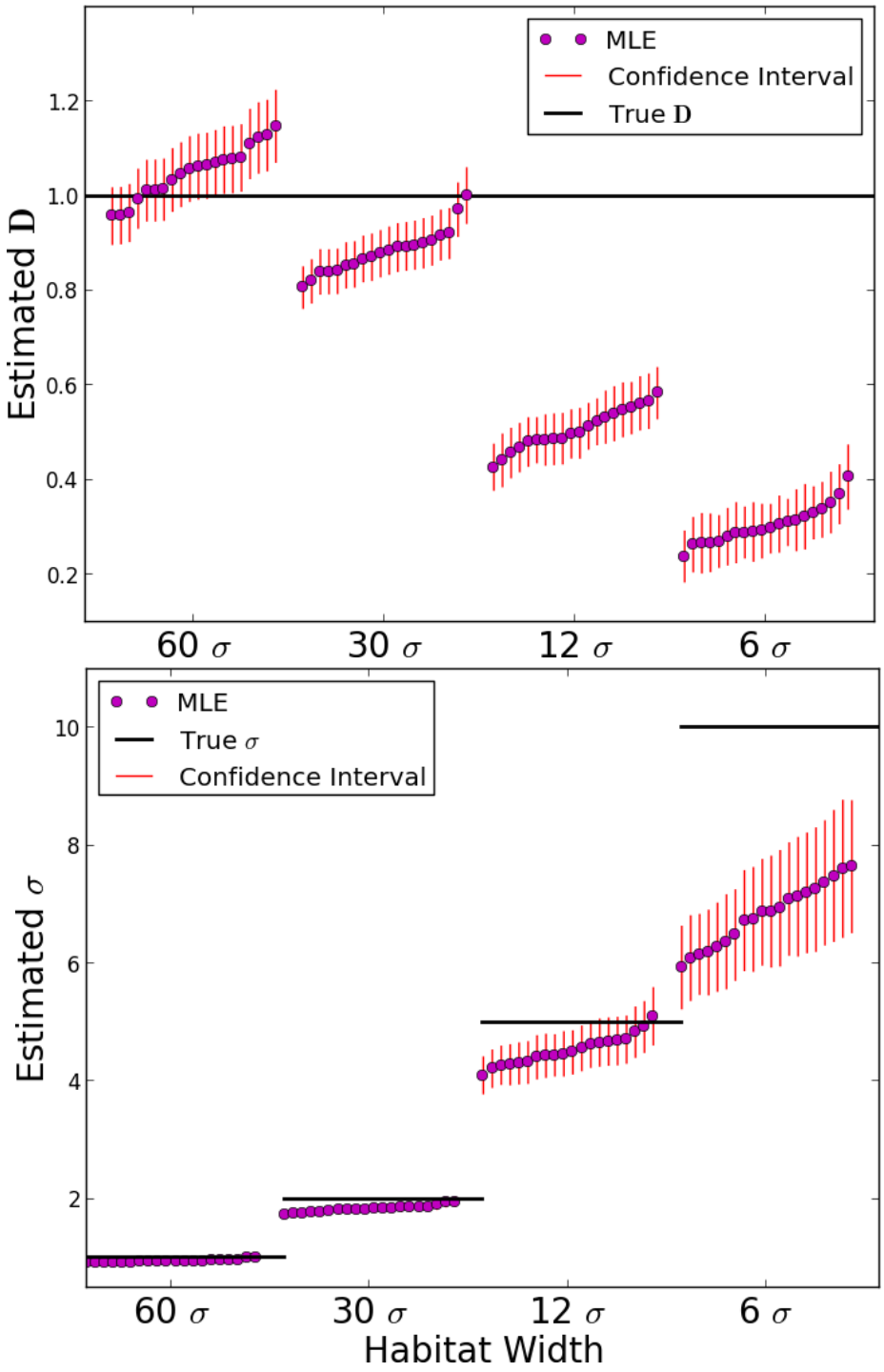
To assess the effect of limited habitat size compared to dispersal width σ, we simulated a finite rectangular habitat. Similarly, we utilized our grid model with reflective boundaries, and a Laplace dispersal kernel along each axis. We simulated four scenarios, where the samples were distributed on a fixed 15×15 grid two distance units apart from each other within a total habitat of axial width 60. Dispersal was set to various values:

- Scenario 1: $\sigma = 1$ (habitat width is 60 σ).
- Scenario 2: $\sigma = 2$ (habitat width is 30 σ).
- Scenario 3: $\sigma = 5$ (habitat width is 12 σ).
- Scenario 4: $\sigma = 10$ (habitat width is 6 σ).

These sampling schemes and width of σ are illustrated in the following figure:



We simulated block sharing until $t = 200$ generations back. Our inference scheme applied to blocks 4–20 cM simulated under 20 replicates of each of these scenarios yielded the following parameter estimates and 95% confidence intervals:



For habitats of width $\approx 10\,\sigma$ and smaller the inference method gave markedly downward biased estimates for σ and D_e, with particularly large bias for D_e.