- In equation (1), change the boundary conditions to

$$a_x \leq X(t) \leq b_x$$
$$a_y \leq Y(t) \leq b_y$$

  Or change the boundary conditions in equation (3) to make the two equations consistent with each other. DONE

- Delete $0 < t' \leq t$ below equation (2). DONE

- Change equation (3) to

$$\frac{1}{dxdy}\Pr\left(\begin{array}{l} X(t) \in [x, x+dx), Y(t) \in [y, y+dy), \\ \min\limits_{t'\in[0,t]} X(t') \geq a_x, \max\limits_{t'\in[0,t]} X(t') \leq b_x, \\ \min\limits_{t'\in[0,t]} Y(t') \geq a_y, \max\limits_{t'\in[0,t]} Y(t') \leq b_y \end{array} \middle| X(0) = x_0, Y(0) = y_0, \theta \right)$$

  DONE

- In equation (4), change to $\dfrac{\partial}{\partial t'}$ DONE

- Change equation (6) to

$$\frac{\partial^4}{\partial a_x \partial b_x \partial a_y \partial b_y} q(x,y,t) = \text{density}\left(\begin{array}{l} X(t) = x, Y(t) = y, \\ \min\limits_{t'\in[0,t]} X(t') = a_x, \max\limits_{t'\in[0,t]} X(t') = b_x, \\ \min\limits_{t'\in[0,t]} Y(t') = a_y, \max\limits_{t'\in[0,t]} Y(t') = b_y \end{array} \middle| X(0) = x_0, Y(0) = y_0, \theta \right)$$

  DONE

- At the end of paragraph after equation (6), change to
  " ... method necessary for carrying out inferential procedures with ... " DONE

- In the last paragraph of page 1, change to
  " ... each eigenfunction of the differential operator is a product of two sine functions, one in each dimension ..."
  DONE

- In the last paragraph of page 1

  Use a_x, a_y instead of a_1, a_2

  Use b_x, b_y instead of b_1, b_2 DONE

- In the last paragraph of page 1, change to

  "This, however, requires one to solve Fokker-Planck equation (4)-(5) accurately for at least 16 slightly different sets of boundaries and combine the results with numerical differentiation to evaluate the density function for a just single observation ... " DONE

- At the beginning of first paragraph on page 2, add

  "for non-zero correlation, "

  DONE

  Change to

  "... where the eigenfunctions for the differential operator are approximated by the eigenvectors of a linear system obtained using a truncated expansion based on a set of separable basis function, each of which is a product of two sine functions (one in each dimension) satisfying the boundary conditions ..."

  DONE

- In the middle of first paragraph on page , change to

  "... makes the expansion a slow, if not unfeasible, solution."

  Change to

  "... from either using a separable representation for the differential operator that is intrinsically correlated in the two dimensions (trigonometric series) or ... " DONE

- At the beginning of second paragraph on page , change to

  "In this paper, we propose a robust and efficient solution to the general problem (4)-(5). The solution is obtained by combining a small-time analytic solution with a Galerkin discretization based on basis functions that are correlated. " DONE

- Change equation (9) to

$$p(x, y, t) = \sum_{v} h_v \phi_v(x, y) e^{-\lambda_v t}$$

where h_v is the coefficient of $\phi_v(x, y)$ in the eigenfunction expansion of $p(x, y, 0)$. DONE

- 4 lines below equation (9) on page 3, change to
  "... we approximate the eigenfunction using a finite set of orthogonal basis functions satisfying boundary conditions, i.e., a finite sequence of sines, ... " DONE

- The equation below equation (9) on page 3, make changes
  1. Each index should start at 1, not 0.
  2. Make the associated changes in the definition of $\Psi(x, y)$ and c_v. DONE

- In the paragraph below equation (10) on page 3, change to
  " ... Applying L to the basis function expansion of ϕ_v and again approximating the result using the finite set of basis functions yields " DONE

- 4 lines below equation (10) on page 3, **what is $\tilde{\theta}$?** This is the first time this parameter/quantity is mentioned. DONE

- After $\tilde{\theta}$ on page 3, add
  "In the last part of the equation above, we have truncated the infinite sine series expansion of $L\psi_{l,m}(x, y)$ " DONE

- 4 lines below equation (10) on page 3, delete the subscript outside the curly bracket it should be $\{\psi_{l,m}(x, y)\}$. DONE

- 5 lines below equation (10) on page 3, change to
  " The dense structure of matrix A is caused by the mixing terms ... " DONE

- 9 lines below equation (10) on page 3, change to
  " ... , we arrive at the matrix eigenvalue problem " DONE

- Equation at the bottom of page 3, move $\Psi(x, y)^T$ outside the summation and add the coefficient h_v
  It should be
  $$p(x, y, t) \approx \Psi(x, y)^T \sum_v h_v c_v e^{-\lambda_v t}$$
  DONE

- Page 4, line 7, change“such that ...” to

  “The 4-th derivative of p with respect to the 4 boundaries is approximated as ”

$$\frac{\partial^4}{\partial a_x \partial b_x \partial a_y \partial b_y} p(x, y, t) \approx \frac{\sum_{k_1,k_2,k_3,k_4=\pm 1} c_{\{k_1,k_2,k_3,k_4\}} p(x, y, t|\, a_x + k_1\varepsilon, b_x + k_2\varepsilon, a_y + k_3\varepsilon, b_y + k_4\varepsilon)}{(2\varepsilon)^4}$$

  DONE

- Page 4, line 9, change to

  “ ... requires many terms in the basis function expansions of the eigenfunctions, ...”

  DONE

- Paragraph after equation (11) on page 4, change to

  “ ... Here, c(t) is a vector consisting of values of the solution in (8) on a set of grid points over Ω at time t, ... ” DONE

- Line 6 below equation (11) on page 4, change to

  “ For example, using $c_{l,m}(t)$ to denote the approximation of the solution at grid point (x_l, y_m), and assembling vector $c(t)$ using the index scheme $c_{l,m}(t) \to c_k(t)$ with $k = (l-1)M + m$, we can approximate the operator $\frac{\partial^2}{\partial x^2}$ as DONE

- Line 12 below equation (11) on page 4, change to

  “ ... with a constant h independent of parameters is appealing ...” DONE

- last sentence of page 4, delete the sentence since we are not discussing the approach in details. If we want to keep the sentence, we need to describe the approach in some details, including i) keeping $\tau_x = 1$ and $\tau_y = 1$; ii) keeping the 3 boundaries on grid points; iii) having to work with $\Omega \neq$ a square; and iv) one side of boundaries not falling on grid points. DONE

- Page 5, line 3, change to

  “ ... , h cannot be made too small because round-off error and the irregular part of the truncation error become an issue relatively quickly. DONE

- Equations and text from line 6 of page 5 to the end of section 2, change to

$$p_{FD}(x,y,t|\,b) = p(x,y,t|\,b) + h^2\mathrm{F}_{reg}(b) + h^2\mathrm{F}_{irreg}(b) + \varepsilon_{mach}\mathrm{F}_{round}(b)$$

where we have included parameter b explicitly as a simplified notation for $[a_x, b_x]$ and $[a_y, b_y]$ after shifting a_x anda_y to 0.

Note that when expressed using the chain rule, both $\dfrac{\partial}{\partial a_x}$ and $\dfrac{\partial}{\partial b_x}$ contain $\dfrac{\partial}{\partial b}$. As a result, $\dfrac{\partial^2}{\partial a_x \partial b_x}$ leads to $\dfrac{\partial^2}{\partial b^2}$. Although in the discussion below, for simplicity, we only illustrate the numerical differentiation on the first derivative, keep in mind that it is the second derivative that is more relevant in the calculation of $\dfrac{\partial^4}{\partial a_x \partial b_x \partial a_y \partial b_y} p(x,y,t)$.

In the expression of finite difference solution above, $h^2\mathrm{F}_{reg}(b)$ is the regular part of the truncation error from discretizing the differential operator on the grid; $h^2\mathrm{F}_{irreg}(b)$ is the irregular part of the truncation error from the interpolations invoked by $\{x_0, y_0, x, y\}$ not being on grid points; and $\varepsilon_{mach}\mathrm{F}_{round}(b)$ is the effect of round-off errors with $\varepsilon_{mach} \sim 10^{-16}$ denoting the machine epsilon for IEEE double precision system. The coefficient, $\mathrm{F}_{reg}(b)$, of the regular part of truncation error is a smooth function of b with derivative $= O(1)$. The coefficient, $\mathrm{F}_{round}(b)$, in the effect of round-off errors, behaves virtually like a random variable, discontinuous in b. For the irregular part of truncation error, the coefficient $\mathrm{F}_{irreg}(b)$ in general is continuous in b but not smooth in b where the derivative has discontinuities of magnitude $O\left(\frac{1}{h}\right)$. For example, the error in a piecewise linear interpolation of a smooth function at position b using step h has the general form of

$$\text{Interpolation error } = O(h^2)(1 - \mathrm{rem}(b/h, 1))\mathrm{rem}(b/h, 1)$$

The coefficient part $\mathrm{F}_{irreg}(b) = (1 - \mathrm{rem}(b/h, 1))\mathrm{rem}(b/h, 1)$ is continuous in b but not differentiable. Its first derivative has the behavior of

$$\frac{\partial}{\partial b}\mathrm{F}_{irreg}(b) = \frac{1}{h}\left(1 - 2\mathrm{rem}(b/h, 1)\right)$$

Based on the expression we wrote out above for the finite difference solution, applying the numerical differentiation on t with step ε yields:

$$\frac{p_{FD}(x,y,t|\,b+\varepsilon) - p_{FD}(x,y,t|\,b-\varepsilon)}{2\varepsilon}$$

$$
\begin{aligned}
&= \frac{\partial}{\partial b} p(x,y,t|\,b) + O(\varepsilon^2) + h^2 \frac{\mathrm{F}_{reg}(b+\varepsilon) - \mathrm{F}_{reg}(b-\varepsilon)}{2\varepsilon} \\
&+ h^2 \frac{\mathrm{F}_{irreg}(b+\varepsilon) - \mathrm{F}_{irreg}(b-\varepsilon)}{2\varepsilon} + \varepsilon_{mach} \frac{\mathrm{F}_{round}(b+\varepsilon) - \mathrm{F}_{round}(b-\varepsilon)}{2\varepsilon}
\end{aligned}
$$

In the equation above, as the step in the numerical differentiation is refined, the first line of the RHS is well behaved, converging to the true value $\frac{\partial}{\partial b} p(x,y,t|\,b)$ as $\varepsilon \to 0$. The second line of RHS, however, is problematic. As $\varepsilon \to 0$, the contribution from round-off error blows up to infinity

$$
\varepsilon_{mach} \frac{\mathrm{F}_{round}(b+\varepsilon) - \mathrm{F}_{round}(b-\varepsilon)}{2\varepsilon} = O\left(\frac{\varepsilon_{mach}}{\varepsilon}\right) \longrightarrow \infty \quad \text{as } \varepsilon \to 0
$$

The contribution from the irregular part of truncation error is

$$
h^2 \frac{\mathrm{F}_{irreg}(b+\varepsilon) - \mathrm{F}_{irreg}(b-\varepsilon)}{2\varepsilon} = O\left(\frac{h^2}{\max(\varepsilon, h)}\right)
$$

In the second order numerical differentiation, however, the contribution from the irregular part of truncation error behaves like

$$
h^2 \frac{\mathrm{F}_{irreg}(b+\varepsilon) - 2\mathrm{F}_{irreg}(b) + \mathrm{F}_{irreg}(b-\varepsilon)}{\varepsilon^2} = O\left(\frac{h^2}{\max(\varepsilon^2, h^2)}\right)
$$

DONE