SUPPLEMENTARY MATERIALS: NUMERICAL GAUSSIAN PROCESSES FOR TIME-DEPENDENT AND NON-LINEAR PARTIAL DIFFERENTIAL EQUATIONS*

MAZIAR RAISSI[†], PARIS PERDIKARIS[‡], AND GEORGE EM KARNIADAKIS[§]

SM1. Wave Equation.

SM1.1. Prior. The covariance functions for the wave equation example are given by

$$\begin{aligned}
&k_{u,u}^{n,n} = k_u + \tfrac{1}{4}\Delta t^2 k_v, && k_{u,v}^{n,n} = \tfrac{1}{2}\Delta t \tfrac{d^2}{dx'^2}k_u + \tfrac{1}{2}\Delta t k_v,\\
&k_{u,u}^{n,n-1} = k_u - \tfrac{1}{4}\Delta t^2 k_v, && k_{u,v}^{n,n-1} = -\tfrac{1}{2}\Delta t \tfrac{d^2}{dx'^2}k_u + \tfrac{1}{2}\Delta t k_v,\\
&k_{v,v}^{n,n} = k_v + \tfrac{1}{4}\Delta t^2 \tfrac{d^2}{dx^2}\tfrac{d^2}{dx'^2}k_u, && k_{v,u}^{n,n-1} = -\tfrac{1}{2}\Delta t k_v + \tfrac{1}{2}\Delta t \tfrac{d^2}{dx^2}k_u,\\
&k_{v,v}^{n,n-1} = k_v - \tfrac{1}{4}\Delta t^2 \tfrac{d^2}{dx^2}\tfrac{d^2}{dx'^2}k_u, && k_{u,u}^{n-1,n-1} = k_u + \tfrac{1}{4}\Delta t^2 k_v,\\
&k_{u,v}^{n-1,n-1} = -\tfrac{1}{2}\Delta t \tfrac{d^2}{dx'^2}k_u - \tfrac{1}{2}\Delta t k_v, && k_{v,v}^{n-1,n-1} = k_v + \tfrac{1}{4}\Delta t^2 \tfrac{d^2}{dx^2}\tfrac{d^2}{dx'^2}k_u.
\end{aligned} \tag{SM1}$$

It is worth highlighting that the only non-trivial but straightforward operations involved in the aforementioned kernel computations are

$$\begin{aligned}
\frac{d^2}{dx'^2}k_u(x,x';\theta_u) &= \frac{d^2}{dx^2}k_u(x,x';\theta_u) \qquad \text{(SM2)}\\
&= \gamma_u^2 w_u e^{-\frac{1}{2}w_u(x-x')^2}\left(w_u(x-x')^2 - 1\right),\\
\frac{d^2}{dx^2}\frac{d^2}{dx'^2}k_u(x,x';\theta_u) &= \gamma_u^2 w_u^2 e^{-\frac{1}{2}w_u(x-x')^2}\left(w_u(x-x')^2\left(w_u(x-x')^2-6\right)+3\right).
\end{aligned}$$

SM1.2. Training. The hyper-parameters θ_u and θ_v can be trained by minimizing the Negative Log Marginal Likelihood resulting from

$$\begin{bmatrix} \boldsymbol{u}_b^n \\ \boldsymbol{u}^{n-1} \\ \boldsymbol{v}^{n-1} \end{bmatrix} \sim \mathcal{N}(0, \boldsymbol{K}), \tag{SM3}$$

where $\{\boldsymbol{x}_b^n, \boldsymbol{u}_b^n\}$ are the data on the boundary, $\{\boldsymbol{x}_u^{n-1}, \boldsymbol{u}^{n-1}\}$, $\{\boldsymbol{x}_v^{n-1}, \boldsymbol{v}^{n-1}\}$ are *artificially generated data*, and

$$\boldsymbol{K} := \begin{bmatrix} \boldsymbol{K}_{u,u}^{n,n} + \sigma_n^2 I & \boldsymbol{K}_{u,u}^{n,n-1} & \boldsymbol{K}_{u,v}^{n,n-1} \\ & \boldsymbol{K}_{u,u}^{n-1,n-1} + \sigma_{u,n-1}^2 I & \boldsymbol{K}_{u,v}^{n-1,n-1} \\ & & \boldsymbol{K}_{v,v}^{n-1,n-1} + \sigma_{v,n-1}^2 I \end{bmatrix},$$

where

$$\begin{aligned}
&\boldsymbol{K}_{u,u}^{n,n} = k_{u,u}^{n,n}(\boldsymbol{x}_b^n, \boldsymbol{x}_b^n), && \boldsymbol{K}_{u,u}^{n,n-1} = k_{u,u}^{n,n-1}(\boldsymbol{x}_b^n, \boldsymbol{x}_u^{n-1}),\\
&\boldsymbol{K}_{u,v}^{n,n-1} = k_{u,v}^{n,n-1}(\boldsymbol{x}_b^n, \boldsymbol{x}_v^{n-1}), && \boldsymbol{K}_{u,u}^{n-1,n-1} = k_{u,u}^{n-1,n-1}(\boldsymbol{x}_u^{n-1}, \boldsymbol{x}_u^{n-1}),\\
&\boldsymbol{K}_{u,v}^{n-1,n-1} = k_{u,v}^{n-1,n-1}(\boldsymbol{x}_u^{n-1}, \boldsymbol{x}_v^{n-1}), && \boldsymbol{K}_{v,v}^{n-1,n-1} = k_{v,v}^{n-1,n-1}(\boldsymbol{x}_v^{n-1}, \boldsymbol{x}_v^{n-1}).
\end{aligned}$$

*Supplementary materials for SISC MS#M112076.

Funding: This works received support by the DARPA EQUiPS grant N66001-15-2-4055, and the AFOSR grant FA9550-17-1-0013.

[†]Division of Applied Mathematics, Brown University, Providence, RI, USA (maziar_raissi@brown.edu).

[‡]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA (parisp@mit.edu).

[§]Division of Applied Mathematics, Brown University, Providence, RI, USA (george_karniadakis@brown.edu).

Here, the data on the boundary are given by

$$\boldsymbol{x}_b^n = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{u}_b^n = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which correspond to the Dirichlet boundary conditions (24).

SM1.3. Posterior. In order to predict $u^n(x_{*u}^n)$ and $v^n(x_{*v}^n)$ at new test points x_{*u}^n and x_{*v}^n, respectively, we use the following conditional distribution

$$\begin{bmatrix} u^n(x_{*u}^n) \\ v^n(x_{*v}^n) \end{bmatrix} \mid \begin{bmatrix} \boldsymbol{u}_b^n \\ \boldsymbol{u}^{n-1} \\ \boldsymbol{v}^{n-1} \end{bmatrix} \sim$$
$$\mathcal{N}\left(\boldsymbol{q}^T \boldsymbol{K}^{-1} \begin{bmatrix} \boldsymbol{u}_b^n \\ \boldsymbol{u}^{n-1} \\ \boldsymbol{v}^{n-1} \end{bmatrix}, \begin{bmatrix} k_{u,u}^{n,n}(x_{*u}^n, x_{*u}^n) & k_{u,v}^{n,n}(x_{*u}^n, x_{*v}^n) \\ & k_{v,v}^{n,n}(x_{*v}^n, x_{*v}^n) \end{bmatrix} - \boldsymbol{q}^T \boldsymbol{K}^{-1} \boldsymbol{q} \right),$$

where $\boldsymbol{q} = [\boldsymbol{q}_u \ \boldsymbol{q}_v]$ and

$$\begin{aligned} \boldsymbol{q}_u^T &:= \begin{bmatrix} k_{u,u}^{n,n}(x_{*u}^n, \boldsymbol{x}_b^n) & k_{u,u}^{n,n-1}(x_{*u}^n, \boldsymbol{x}_u^{n-1}) & k_{u,v}^{n,n-1}(x_{*u}^n, \boldsymbol{x}_v^{n-1}) \end{bmatrix}, \\ \boldsymbol{q}_v^T &:= \begin{bmatrix} k_{v,u}^{n,n}(x_{*v}^n, \boldsymbol{x}_b^n) & k_{v,u}^{n,n-1}(x_{*v}^n, \boldsymbol{x}_u^{n-1}) & k_{v,v}^{n,n-1}(x_{*v}^n, \boldsymbol{x}_v^{n-1}) \end{bmatrix}. \end{aligned}$$

SM1.4. Propagating Uncertainty. Since $\{\boldsymbol{x}_u^{n-1}, \boldsymbol{u}^{n-1}\}$ and $\{\boldsymbol{x}_v^{n-1}, \boldsymbol{v}^{n-1}\}$ are *artificially generated data*, to properly propagate the uncertainty associated with the initial data, we have to marginalize them out by employing

$$\begin{bmatrix} \boldsymbol{u}^{n-1} \\ \boldsymbol{v}^{n-1} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_u^{n-1} \\ \boldsymbol{\mu}_v^{n-1} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{u,u}^{n-1,n-1} & \boldsymbol{\Sigma}_{u,v}^{n-1,n-1} \\ & \boldsymbol{\Sigma}_{v,v}^{n-1,n-1} \end{bmatrix} \right),$$

to obtain

$$\begin{bmatrix} u^n(x_{*u}^n) \\ v^n(x_{*v}^n) \end{bmatrix} \mid \boldsymbol{u}_b^n \sim \tag{SM4}$$
$$\mathcal{N}\left(\begin{bmatrix} \mu_u^n(x_{*u}^n) \\ \mu_v^n(x_{*v}^n) \end{bmatrix}, \begin{bmatrix} \Sigma_{u,u}^{n,n}(x_{*u}^n, x_{*u}^n) & \Sigma_{u,v}^{n,n}(x_{*u}^n, x_{*v}^n) \\ & \Sigma_{v,v}^{n,n}(x_{*v}^n, x_{*v}^n) \end{bmatrix} \right),$$

where

$$\begin{bmatrix} \mu_u^n(x_{*u}^n) \\ \mu_v^n(x_{*v}^n) \end{bmatrix} = \boldsymbol{q}^T \boldsymbol{K}^{-1} \begin{bmatrix} \boldsymbol{u}_b^n \\ \boldsymbol{\mu}_u^{n-1} \\ \boldsymbol{\mu}_v^{n-1} \end{bmatrix},$$

and

$$\begin{bmatrix} \Sigma_{u,u}^{n,n}(x_{*u}^n, x_{*u}^n) & \Sigma_{u,v}^{n,n}(x_{*u}^n, x_{*v}^n) \\ & \Sigma_{v,v}^{n,n}(x_{*v}^n, x_{*v}^n) \end{bmatrix} = \begin{bmatrix} k_{u,u}^{n,n}(x_{*u}^n, x_{*u}^n) & k_{u,v}^{n,n}(x_{*u}^n, x_{*v}^n) \\ & k_{v,v}^{n,n}(x_{*v}^n, x_{*v}^n) \end{bmatrix} -$$
$$\boldsymbol{q}^T \boldsymbol{K}^{-1} \boldsymbol{q} + \boldsymbol{q}^T \boldsymbol{K}^{-1} \begin{bmatrix} 0 & 0 & 0 \\ & \boldsymbol{\Sigma}_{u,u}^{n-1,n-1} & \boldsymbol{\Sigma}_{u,v}^{n-1,n-1} \\ & & \boldsymbol{\Sigma}_{v,v}^{n-1,n-1} \end{bmatrix} \boldsymbol{K}^{-1} \boldsymbol{q}.$$

Now, we can use the resulting posterior distribution to obtain the artificially generated data $\{\boldsymbol{x}_u^n, \boldsymbol{u}^n\}$ and $\{\boldsymbol{x}_v^n, \boldsymbol{v}^n\}$ for the next time step with

$$\begin{bmatrix} \boldsymbol{u}^n \\ \boldsymbol{v}^n \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_u^n \\ \boldsymbol{\mu}_v^n \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{u,u}^{n,n} & \boldsymbol{\Sigma}_{u,v}^{n,n} \\ & \boldsymbol{\Sigma}_{v,v}^{n,n} \end{bmatrix} \right). \tag{SM5}$$

SM2. Advection Equation.

SM2.1. Prior. The covariance functions for the advection equation example are given by

$$
\begin{aligned}
k_{u,3}^{n+1,n} &= k_{u,u}^{n+1,n+1}, & k_{u,3}^{n+\tau_2,n} &= b_2 \Delta t \tfrac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2},\\
k_{u,2}^{n+\tau_2,n} &= k_{u,u}^{n+\tau_2,n+\tau_2} + a_{22}\Delta t \tfrac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2}, & k_{u,1}^{n+\tau_2,n} &= a_{12}\Delta t \tfrac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2},\\
k_{u,3}^{n+\tau_1,n} &= b_1 \Delta t \tfrac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1}, & k_{u,2}^{n+\tau_1,n} &= a_{21}\Delta t \tfrac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1},\\
k_{u,1}^{n+\tau_1,n} &= k_{u,u}^{n+\tau_1,n+\tau_1} + a_{11}\Delta t \tfrac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1},
\end{aligned}
$$

and

$$
\begin{aligned}
k_{3,3}^{n,n} &= k_{u,u}^{n+1,n+1} + b_1^2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1} + b_2^2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2},\\
k_{3,2}^{n,n} &= b_2 \Delta t \frac{d}{dx} k_{u,u}^{n+\tau_2,n+\tau_2} + a_{21} b_1 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1}\\
&\quad + a_{22} b_2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2},\\
k_{3,1}^{n,n} &= b_1 \Delta t \frac{d}{dx} k_{u,u}^{n+\tau_1,n+\tau_1} + a_{11} b_1 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1}\\
&\quad + a_{12} b_2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2},
\end{aligned}
$$

$$
\begin{aligned}
k_{2,2}^{n,n} &= k_{u,u}^{n+\tau_2,n+\tau_2} + a_{21}^2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1} + a_{22}^2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2},\\
k_{2,1}^{n,n} &= a_{12}\Delta t \frac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2} + a_{21}\Delta t \frac{d}{dx} k_{u,u}^{n+\tau_1,n+\tau_1}\\
&\quad + a_{21} a_{11} \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1} + a_{22} a_{12} \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2},
\end{aligned}
$$

$$
k_{1,1}^{n,n} = k_{u,u}^{n+\tau_1,n+\tau_1} + a_{11}^2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_1,n+\tau_1} + a_{12}^2 \Delta t^2 \frac{d}{dx}\frac{d}{dx'} k_{u,u}^{n+\tau_2,n+\tau_2}.
$$

SM2.2. Training. The matrix $\boldsymbol{K}$ used in the distribution (39) is given by

$$
\boldsymbol{K} = \begin{bmatrix}
K_{b,b}^{n+1,n+1} & 0 & 0 & \boldsymbol{K}_{b,3}^{n+1,n} & 0 & 0\\
 & K_{b,b}^{n+\tau_2,n+\tau_2} & 0 & \boldsymbol{K}_{b,3}^{n+\tau_2,n} & \boldsymbol{K}_{b,2}^{n+\tau_2,n} & \boldsymbol{K}_{b,1}^{n+\tau_2,n}\\
 & & K_{b,b}^{n+\tau_1,n+\tau_1} & \boldsymbol{K}_{b,3}^{n+\tau_1,n} & \boldsymbol{K}_{b,2}^{n+\tau_1,n} & \boldsymbol{K}_{b,1}^{n+\tau_1,n}\\
 & & & \boldsymbol{K}_{3,3}^{n,n} + \sigma_n^2 I & \boldsymbol{K}_{3,2}^{n,n} & \boldsymbol{K}_{3,1}^{n,n}\\
 & & & & \boldsymbol{K}_{2,2}^{n,n} + \sigma_n^2 I & \boldsymbol{K}_{2,1}^{n,n}\\
 & & & & & \boldsymbol{K}_{1,1}^{n,n} + \sigma_n^2 I
\end{bmatrix},
$$

where

$$
\begin{aligned}
K_{b,b}^{n+1,n+1} &= k_{u,u}^{n+1,n+1}(1,1) - k_{u,u}^{n+1,n+1}(1,0)\\
&\quad - k_{u,u}^{n+1,n+1}(0,1) + k_{u,u}^{n+1,n+1}(0,0),\\
K_{b,b}^{n+\tau_2,n+\tau_2} &= k_{u,u}^{n+\tau_2,n+\tau_2}(1,1) - k_{u,u}^{n+\tau_2,n+\tau_2}(1,0)\\
&\quad - k_{u,u}^{n+\tau_2,n+\tau_2}(0,1) + k_{u,u}^{n+\tau_2,n+\tau_2}(0,0),\\
K_{b,b}^{n+\tau_1,n+\tau_1} &= k_{u,u}^{n+\tau_1,n+\tau_1}(1,1) - k_{u,u}^{n+\tau_1,n+\tau_1}(1,0)\\
&\quad - k_{u,u}^{n+\tau_1,n+\tau_1}(0,1) + k_{u,u}^{n+\tau_1,n+\tau_1}(0,0),
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{K}_{b,i}^{n+1,n} &= k_{u,i}^{n+1,n}(1,\boldsymbol{x}^n) - k_{u,i}^{n+1,n}(0,\boldsymbol{x}^n),\ i=3,2,1,\\
\boldsymbol{K}_{b,i}^{n+\tau_2,n} &= k_{u,i}^{n+\tau_2,n}(1,\boldsymbol{x}^n) - k_{u,i}^{n+\tau_2,n}(0,\boldsymbol{x}^n),\ i=3,2,1,\\
\boldsymbol{K}_{b,i}^{n+\tau_1,n} &= k_{u,i}^{n+\tau_1,n}(1,\boldsymbol{x}^n) - k_{u,i}^{n+\tau_1,n}(0,\boldsymbol{x}^n),\ i=3,2,1,\\
\boldsymbol{K}_{i,j}^{n,n} &= k_{i,j}^{n,n}(\boldsymbol{x}^n,\boldsymbol{x}^n),\quad i,j=3,2,1,\ j\le i.
\end{aligned}
$$

SM2.3. Posterior. In order to predict $u^{n+1}(x_*^{n+1})$ at a new test point x_*^{n+1}, we use

$$
u^{n+1}(x_*^{n+1}) \mid \begin{bmatrix} u^{n+1}(1)-u^{n+1}(0)=0 \\ u^{n+\tau_2}(1)-u^{n+\tau_2}(0)=0 \\ u^{n+\tau_1}(1)-u^{n+\tau_1}(0)=0 \\ \boldsymbol{u}^n \\ \boldsymbol{u}^n \\ \boldsymbol{u}^n \end{bmatrix} \sim
\mathcal{N}\left(\boldsymbol{q}^T\boldsymbol{K}^{-1}\begin{bmatrix} 0\\0\\0\\ \boldsymbol{u}^n\\ \boldsymbol{u}^n\\ \boldsymbol{u}^n \end{bmatrix}, k_{u,u}^{n+1,n+1}(x_*^{n+1},x_*^{n+1}) - \boldsymbol{q}^T\boldsymbol{K}^{-1}\boldsymbol{q}\right),
$$

where

$$
\boldsymbol{q} := \begin{bmatrix} k_{u,u}^{n+1,n+1}(1,x_*^{n+1}) - k_{u,u}^{n+1,n+1}(0,x_*^{n+1}) \\ 0 \\ 0 \\ k_{3,u}^{n,n+1}(\boldsymbol{x}^n, x_*^{n+1}) \\ k_{2,u}^{n,n+1}(\boldsymbol{x}^n, x_*^{n+1}) \\ k_{1,u}^{n,n+1}(\boldsymbol{x}^n, x_*^{n+1}) \end{bmatrix}.
$$

SM2.4. Propagating Uncertainty. To propagate the uncertainty associate with the noisy initial data through time we have to marginalize out the artificially generated data $\{\boldsymbol{x}^n, \boldsymbol{u}^n\}$ by employing

$$
\boldsymbol{u}^n \sim \mathcal{N}\left(\boldsymbol{\mu}^n, \boldsymbol{\Sigma}^{n,n}\right),
$$

to obtain

$$
\begin{aligned}
u^{n+1}(x_*^{n+1}) \mid &\begin{bmatrix} u^{n+1}(1)-u^{n+1}(0)=0 \\ u^{n+\tau_2}(1)-u^{n+\tau_2}(0)=0 \\ u^{n+\tau_1}(1)-u^{n+\tau_1}(0)=0 \end{bmatrix} \\
&\sim \mathcal{N}\left(\mu^{n+1}(x_*^{n+1}), \Sigma^{n+1,n+1}(x_*^{n+1}, x_*^{n+1})\right),
\end{aligned} \tag{SM6}
$$

where

$$
\mu^{n+1}(x_*^{n+1}) = \boldsymbol{q}^T\boldsymbol{K}^{-1}\begin{bmatrix} 0\\0\\0\\ \boldsymbol{\mu}^n\\ \boldsymbol{\mu}^n\\ \boldsymbol{\mu}^n \end{bmatrix},
$$

and

$$\Sigma^{n+1,n+1}(x_*^{n+1}, x_*^{n+1}) = k_{u,u}^{n+1,n+1}(x_*^{n+1}, x_*^{n+1}) - \boldsymbol{q}^T \boldsymbol{K}^{-1} \boldsymbol{q}$$
$$+\boldsymbol{q}^T \boldsymbol{K}^{-1} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ & & & \boldsymbol{\Sigma}^{n,n} & \boldsymbol{\Sigma}^{n,n} & \boldsymbol{\Sigma}^{n,n} \\ & & & & \boldsymbol{\Sigma}^{n,n} & \boldsymbol{\Sigma}^{n,n} \\ & & & & & \boldsymbol{\Sigma}^{n,n} \end{bmatrix} \boldsymbol{K}^{-1} \boldsymbol{q}.$$

Now, we can use the resulting posterior distribution (SM6) to obtain the artificially generated data $\{\boldsymbol{x}^{n+1}, \boldsymbol{u}^{n+1}\}$ with

$$\boldsymbol{u}^{n+1} \sim \mathcal{N}\left(\boldsymbol{\mu}^{n+1}, \boldsymbol{\Sigma}^{n+1,n+1}\right). \tag{SM7}$$

SM3. Heat equation.

SM3.1. Prior.

The covariance functions for the Heat equation are given by

$$\begin{aligned}
k_{u,v}^{n+1,n+1} &= \frac{d}{dx_2'} k_{u,u}^{n+1,n+1}, \\
k_{u,3}^{n+1,n} &= k_{u,u}^{n+1,n+1} - \frac{1}{2}\Delta t \frac{d^2}{dx_1'^2} k_{u,u}^{n+1,n+1} - \frac{1}{2}\Delta t \frac{d^2}{dx_2'^2} k_{u,u}^{n+1,n+1}, \\
k_{v,v}^{n+1,n+1} &= \frac{d}{dx_2} \frac{d}{dx_2'} k_{u,u}^{n+1,n+1}, \\
k_{v,3}^{n+1,n} &= \frac{d}{dx_2} k_{u,u}^{n+1,n+1} - \frac{1}{2}\Delta t \frac{d}{dx_2} \frac{d^2}{dx_1'^2} k_{u,u}^{n+1,n+1} - \frac{1}{2}\Delta t \frac{d}{dx_2} \frac{d^2}{dx_2'^2} k_{u,u}^{n+1,n+1}, \\
k_{u,v}^{n,n} &= \frac{d}{dx_2'} k_{u,u}^{n,n}, \\
k_{u,3}^{n,n} &= -\frac{1}{2}\Delta t \frac{d^2}{dx_1'^2} k_{u,u}^{n,n} - \frac{1}{2}\Delta t \frac{d^2}{dx_2'^2} k_{u,u}^{n,n}, \\
k_{u,1}^{n,n} &= k_{u,u}^{n,n} \\
k_{v,v}^{n,n} &= \frac{d}{dx_2} \frac{d}{dx_2'} k_{u,u}^{n,n}, \\
k_{v,3}^{n,n} &= -\frac{1}{2}\Delta t \frac{d}{dx_2} \frac{d^2}{dx_1'^2} k_{u,u}^{n,n} - \frac{1}{2}\Delta t \frac{d}{dx_2} \frac{d^2}{dx_2'^2} k_{u,u}^{n,n}, \\
k_{v,1}^{n,n} &= \frac{d}{dx_2} k_{u,u}^{n,n},
\end{aligned} \tag{SM8}$$

and

$$
\begin{aligned}
k_{3,3}^{n,n} &= k_{u,u}^{n+1,n+1} - \frac{1}{2}\Delta t \frac{d^2}{dx_1'^2} k_{u,u}^{n+1,n+1} - \frac{1}{2}\Delta t \frac{d^2}{dx_2'^2} k_{u,u}^{n+1,n+1} \\
&+ \frac{1}{4}\Delta t^2 \frac{d^2}{dx_1^2}\frac{d^2}{dx_1'^2} k_{u,u}^{n,n} + \frac{1}{4}\Delta t^2 \frac{d^2}{dx_1^2}\frac{d^2}{dx_2'^2} k_{u,u}^{n,n} \\
&- \frac{1}{2}\Delta t \frac{d^2}{dx_1^2} k_{u,u}^{n+1,n+1} + \frac{1}{4}\Delta t^2 \frac{d^2}{dx_1^2}\frac{d^2}{dx_1'^2} k_{u,u}^{n+1,n+1} + \frac{1}{4}\Delta t^2 \frac{d^2}{dx_1^2}\frac{d^2}{dx_2'^2} k_{u,u}^{n+1,n+1} \\
&+ \frac{1}{4}\Delta t^2 \frac{d^2}{dx_2^2}\frac{d^2}{dx_1'^2} k_{u,u}^{n,n} + \frac{1}{4}\Delta t^2 \frac{d^2}{dx_2^2}\frac{d^2}{dx_2'^2} k_{u,u}^{n,n} \\
&- \frac{1}{2}\Delta t \frac{d^2}{dx_2^2} k_{u,u}^{n+1,n+1} + \frac{1}{4}\Delta t^2 \frac{d^2}{dx_2^2}\frac{d^2}{dx_1'^2} k_{u,u}^{n+1,n+1} + \frac{1}{4}\Delta t^2 \frac{d^2}{dx_2^2}\frac{d^2}{dx_2'^2} k_{u,u}^{n+1,n+1},
\end{aligned}
\tag{SM9}
$$

$$
\begin{aligned}
k_{3,1}^{n,n} &= -\frac{1}{2}\Delta t \frac{d^2}{dx_1^2} k_{u,u}^{n,n} - \frac{1}{2}\Delta t \frac{d^2}{dx_2^2} k_{u,u}^{n,n}, \\
k_{1,1}^{n,n} &= k_{u,u}^{n,n}.
\end{aligned}
$$

SM3.2. Training. The matrix $\boldsymbol{K}$ used in the distribution (46) is given by

$$
\boldsymbol{K} = \begin{bmatrix}
\boldsymbol{K}_{D,D}^{n+1,n+1} & \boldsymbol{K}_{D,N}^{n+1,n+1} & 0 & 0 & \boldsymbol{K}_{D,3}^{n+1,n} & 0 \\
 & \boldsymbol{K}_{N,N}^{n+1,n+1} & 0 & 0 & \boldsymbol{K}_{N,3}^{n+1,n} & 0 \\
 & & \boldsymbol{K}_{D,D}^{n,n} & \boldsymbol{K}_{D,N}^{n,n} & \boldsymbol{K}_{D,3}^{n,n} & \boldsymbol{K}_{D,1}^{n,n} \\
 & & & \boldsymbol{K}_{N,N}^{n,n} & \boldsymbol{K}_{N,3}^{n,n} & \boldsymbol{K}_{N,1}^{n,n} \\
 & & & & \boldsymbol{K}_{3,3}^{n,n} & \boldsymbol{K}_{3,1}^{n,n} \\
 & & & & & \boldsymbol{K}_{1,1}^{n,n}
\end{bmatrix}.
$$

Here,

$$
\begin{aligned}
\boldsymbol{K}_{D,D}^{n+1,n+1} &= k_{u,u}^{n+1,n+1}\left((\boldsymbol{x}_{1,D}^{n+1}, \boldsymbol{x}_{2,D}^{n+1}), (\boldsymbol{x}_{1,D}^{n+1}, \boldsymbol{x}_{2,D}^{n+1})\right) + \sigma_{D,n+1}^2 I, \\
\boldsymbol{K}_{D,N}^{n+1,n+1} &= k_{u,v}^{n+1,n+1}\left((\boldsymbol{x}_{1,D}^{n+1}, \boldsymbol{x}_{2,D}^{n+1}), (\boldsymbol{x}_{1,N}^{n+1}, \boldsymbol{x}_{2,N}^{n+1})\right), \\
\boldsymbol{K}_{D,3}^{n+1,n} &= k_{u,3}^{n+1,n}\left((\boldsymbol{x}_{1,D}^{n+1}, \boldsymbol{x}_{2,D}^{n+1}), (\boldsymbol{x}_1^{n}, \boldsymbol{x}_2^{n})\right), \\
\boldsymbol{K}_{N,N}^{n+1,n+1} &= k_{v,v}^{n+1,n+1}\left((\boldsymbol{x}_{1,N}^{n+1}, \boldsymbol{x}_{2,N}^{n+1}), (\boldsymbol{x}_{1,N}^{n+1}, \boldsymbol{x}_{2,N}^{n+1})\right) + \sigma_{N,n+1}^2 I, \\
\boldsymbol{K}_{N,3}^{n+1,n} &= k_{v,3}^{n+1,n}\left((\boldsymbol{x}_{1,N}^{n+1}, \boldsymbol{x}_{2,N}^{n+1}), (\boldsymbol{x}_1^{n}, \boldsymbol{x}_2^{n})\right),
\end{aligned}
\tag{SM10}
$$

$$
\begin{aligned}
\boldsymbol{K}_{D,D}^{n,n} &= k_{u,u}^{n,n}\left((\boldsymbol{x}_{1,D}^{n}, \boldsymbol{x}_{2,D}^{n}), (\boldsymbol{x}_{1,D}^{n}, \boldsymbol{x}_{2,D}^{n})\right) + \sigma_{D,n}^2 I, \\
\boldsymbol{K}_{D,N}^{n,n} &= k_{u,v}^{n,n}\left((\boldsymbol{x}_{1,D}^{n}, \boldsymbol{x}_{2,D}^{n}), (\boldsymbol{x}_{1,N}^{n}, \boldsymbol{x}_{2,N}^{n})\right), \\
\boldsymbol{K}_{D,3}^{n,n} &= k_{u,3}^{n,n}\left((\boldsymbol{x}_{1,D}^{n}, \boldsymbol{x}_{2,D}^{n}), (\boldsymbol{x}_1^{n}, \boldsymbol{x}_2^{n})\right), \\
\boldsymbol{K}_{D,1}^{n,n} &= k_{u,1}^{n,n}\left((\boldsymbol{x}_{1,D}^{n}, \boldsymbol{x}_{2,D}^{n}), (\boldsymbol{x}_1^{n}, \boldsymbol{x}_2^{n})\right), \\
\boldsymbol{K}_{N,N}^{n,n} &= k_{v,v}^{n,n}\left((\boldsymbol{x}_{1,N}^{n}, \boldsymbol{x}_{2,N}^{n}), (\boldsymbol{x}_{1,N}^{n}, \boldsymbol{x}_{2,N}^{n})\right) + \sigma_{N,n}^2 I, \\
\boldsymbol{K}_{N,3}^{n,n} &= k_{v,3}^{n,n}\left((\boldsymbol{x}_{1,N}^{n}, \boldsymbol{x}_{2,N}^{n}), (\boldsymbol{x}_1^{n}, \boldsymbol{x}_2^{n})\right), \\
\boldsymbol{K}_{N,1}^{n,n} &= k_{v,1}^{n,n}\left((\boldsymbol{x}_{1,N}^{n}, \boldsymbol{x}_{2,N}^{n}), (\boldsymbol{x}_1^{n}, \boldsymbol{x}_2^{n})\right),
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{K}_{3,3}^{n,n} &= k_{3,3}^{n,n}\left((\boldsymbol{x}_1^n, \boldsymbol{x}_2^n), (\boldsymbol{x}_1^n, \boldsymbol{x}_2^n)\right) + \sigma_n^2 I,\\
\boldsymbol{K}_{3,1}^{n,n} &= k_{3,1}^{n,n}\left((\boldsymbol{x}_1^n, \boldsymbol{x}_2^n), (\boldsymbol{x}_1^n, \boldsymbol{x}_2^n)\right),\\
\boldsymbol{K}_{1,1}^{n,n} &= k_{1,1}^{n,n}\left((\boldsymbol{x}_1^n, \boldsymbol{x}_2^n), (\boldsymbol{x}_1^n, \boldsymbol{x}_2^n)\right) + \sigma_n^2 I.
\end{aligned}
$$

SM3.3. Posterior. In order to predict $u^{n+1}(x_{1,*}^{n+1}, x_{2,*}^{n+1})$ at a new test point $(x_{1,*}^{n+1}, x_{2,*}^{n+1})$, we use

$$
u^{n+1}(x_{1,*}^{n+1}, x_{2,*}^{n+1}) \mid \begin{bmatrix} \boldsymbol{u}_D^{n+1} \\ \boldsymbol{v}_N^{n+1} \\ \boldsymbol{u}_D^n \\ \boldsymbol{v}_N^n \\ \boldsymbol{u}^n \\ \boldsymbol{u}^n \end{bmatrix} \sim
$$

$$
\mathcal{N}\left(\boldsymbol{q}^T \boldsymbol{K}^{-1} \begin{bmatrix} \boldsymbol{u}_D^{n+1} \\ \boldsymbol{v}_N^{n+1} \\ \boldsymbol{u}_D^n \\ \boldsymbol{v}_N^n \\ \boldsymbol{u}^n \\ \boldsymbol{u}^n \end{bmatrix}, k_{u,u}^{n+1,n+1}\left((x_{1,*}^{n+1}, x_{2,*}^{n+1}), (x_{1,*}^{n+1}, x_{2,*}^{n+1})\right) - \boldsymbol{q}^T \boldsymbol{K}^{-1} \boldsymbol{q}\right),
$$

where

$$
\boldsymbol{q} := \begin{bmatrix} k_{u,u}^{n+1,n+1}\left((\boldsymbol{x}_{1,D}^{n+1}, \boldsymbol{x}_{2,D}^{n+1}), (x_{1,*}^{n+1}, x_{2,*}^{n+1})\right) \\ k_{v,u}^{n+1,n+1}\left((\boldsymbol{x}_{1,N}^{n+1}, \boldsymbol{x}_{2,N}^{n+1}), (x_{1,*}^{n+1}, x_{2,*}^{n+1})\right) \\ 0 \\ 0 \\ k_{3,u}^{n,n+1}\left((\boldsymbol{x}_1^n, \boldsymbol{x}_2^n), (x_{1,*}^{n+1}, x_{2,*}^{n+1})\right) \\ 0 \end{bmatrix}.
$$

SM3.4. Propagating Uncertainty. To propagate the uncertainty associate with the noisy initial data through time we have to marginalize out the artificially generated data $\{(\boldsymbol{x}_1^n, \boldsymbol{x}_2^n), \boldsymbol{u}^n\}$ by employing

$$
\boldsymbol{u}^n \sim \mathcal{N}\left(\boldsymbol{\mu}^n, \boldsymbol{\Sigma}^{n,n}\right),
$$

to obtain

$$
\begin{aligned}
&u^{n+1}(x_{1,*}^{n+1}, x_{2,*}^{n+1}) \mid \begin{bmatrix} \boldsymbol{u}_D^{n+1} \\ \boldsymbol{v}_N^{n+1} \\ \boldsymbol{u}_D^n \\ \boldsymbol{v}_N^n \end{bmatrix} \\
&\sim \mathcal{N}\left(\mu^{n+1}(x_{1,*}^{n+1}, x_{2,*}^{n+1}), \Sigma^{n+1,n+1}((x_{1,*}^{n+1}, x_{2,*}^{n+1}), (x_{1,*}^{n+1}, x_{2,*}^{n+1}))\right),
\end{aligned} \tag{SM11}
$$

where

$$
\mu^{n+1}(x_{1,*}^{n+1}, x_{2,*}^{n+1}) = \boldsymbol{q}^T \boldsymbol{K}^{-1} \begin{bmatrix} \boldsymbol{u}_D^{n+1} \\ \boldsymbol{v}_N^{n+1} \\ \boldsymbol{u}_D^n \\ \boldsymbol{v}_N^n \\ \boldsymbol{\mu}^n \\ \boldsymbol{\mu}^n \end{bmatrix},
$$

and

$$\Sigma^{n+1,n+1}(x_*^{n+1}, x_*^{n+1}) = k_{u,u}^{n+1,n+1}(x_*^{n+1}, x_*^{n+1}) - \boldsymbol{q}^T \boldsymbol{K}^{-1} \boldsymbol{q}$$
$$+\boldsymbol{q}^T \boldsymbol{K}^{-1} \begin{bmatrix} 0 & 0 & 0 \\ & \boldsymbol{\Sigma}^{n,n} & \boldsymbol{\Sigma}^{n,n} \\ & & \boldsymbol{\Sigma}^{n,n} \end{bmatrix} \boldsymbol{K}^{-1} \boldsymbol{q}.$$

Now, we can use the resulting posterior distribution (SM11) to obtain the artificially generated data $\{(\boldsymbol{x}_1^{n+1}, \boldsymbol{x}_2^{n+1}), \boldsymbol{u}^{n+1}\}$ with

$$\boldsymbol{u}^{n+1} \sim \mathcal{N}\left(\boldsymbol{\mu}^{n+1}, \boldsymbol{\Sigma}^{n+1,n+1}\right). \tag{SM12}$$