

We now focus on the following time-dependent PDE:

$$\mathcal{L}u(x, t) := \partial_t u(x, t) - \nabla \cdot (a(x) \nabla u(x, t)) = f(x, t), \quad x \in \Omega, \quad t \in [0, T] \quad (1)$$

$$u(x, t) = 0, \quad x \in \partial\Omega, \quad t \in [0, T] \quad (2)$$

$$u(x, 0) = u_0(x), \quad x \in \Omega \quad (3)$$

We will set up a prior on the solution u to the above problem. To do so we first let $v_h \in S_h$ be some approximation of the initial condition $u_0(x)$ in the FEM space S_h . To be more specific we will assume that $v_h(x) = \Phi(x)^* \gamma := \sum_{i=1}^J \phi_i(x) \gamma_i$. Note that $\Phi(x) := (\phi_1(x), \dots, \phi_J(x))^T$. We take the prior on u to be:

$$u \sim \mathcal{N}(m_0, V_0) \quad (4)$$

where $m_0(x, t) := v_h(x) = \Phi^*(x) \gamma$ (m_0 is constant in time). The prior covariance operator V_0 is defined as:

$$(V_0 g)(x, t) = \int_{\Omega} \int_0^T k_{x,t,y,s}^{(0)} g(y, s) ds dy \quad (5)$$

where $k_{x,t,y,s}^{(0)}$ is defined as follows:

$$k_{x,t,y,s}^{(0)} := \sum_{i=1}^J \lambda_j \phi_j(x) \phi_j(y) k^{(0)}(t, s) \quad (6)$$

We now introduce a uniform time grid:

$$t_n = n\delta, \quad n = 0, 1, \dots, N$$

where δ is the spacing between consecutive times and $N = \frac{T}{\delta}$ (assume that N is an integer). The time kernel $k^{(0)}(t, s)$ will be taken to be:

$$k^{(0)}(t, s) := \sum_{i=0}^{N-1} (t - t_i)(s - t_i) \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s) \quad (7)$$

We now introduce the following operators $\mathcal{I}_s := (I_1(s), \dots, I_J(s))^T$ where:

$$I_i(s)g := \int_{\Omega} \phi_i(x) g(x, s) dx \quad (8)$$

To move from $t = t_0 = 0$ to $t = t_1 = \delta$ we condition on observing $\mathcal{I}_{\delta} \mathcal{L}u = \mathcal{I}_{\delta} f =: F^{(1)}$. Let $\tilde{\mathcal{A}}_{\delta} := \mathcal{I}_{\delta} \mathcal{L}$. For a fixed realisation of f (and so of $F^{(1)}$) we thus seek the following conditional distribution:

$$u | \{\tilde{\mathcal{A}}_{\delta} u = F^{(1)}, f\} \sim \mathcal{N}(m_1, V_1) \quad (9)$$

That this distribution is itself Gaussian follows from considering the following joint distribution:

$$\begin{pmatrix} u \\ \tilde{\mathcal{A}}_{\delta} u \end{pmatrix} = \begin{pmatrix} I \\ \tilde{\mathcal{A}}_{\delta} \end{pmatrix} u \sim \mathcal{N} \left(\begin{pmatrix} m_0 \\ \tilde{\mathcal{A}}_{\delta} m_0 \end{pmatrix}, \begin{pmatrix} V_0 & V_0 \tilde{\mathcal{A}}_{\delta}^* \\ \tilde{\mathcal{A}}_{\delta} V_0 & \tilde{\mathcal{A}}_{\delta} V_0 \tilde{\mathcal{A}}_{\delta}^* \end{pmatrix} \right)$$

It follows that the conditional distribution is Gaussian and the mean and covariance are given by:

$$m_1 = m_0 + V_0 \tilde{\mathcal{A}}_{\delta}^* (\tilde{\mathcal{A}}_{\delta} V_0 \tilde{\mathcal{A}}_{\delta}^*)^{-1} (F^{(1)} - \tilde{\mathcal{A}}_{\delta} m_0) \quad (10)$$

$$V_1 = V_0 - V_0 \tilde{\mathcal{A}}_{\delta}^* (\tilde{\mathcal{A}}_{\delta} V_0 \tilde{\mathcal{A}}_{\delta}^*)^{-1} \tilde{\mathcal{A}}_{\delta} V_0 \quad (11)$$

We now rewrite the mean update equation as follows:

$$m_1 = \left(1 - V_0 \tilde{\mathcal{A}}_{\delta}^* (\tilde{\mathcal{A}}_{\delta} V_0 \tilde{\mathcal{A}}_{\delta}^*)^{-1} \tilde{\mathcal{A}}_{\delta} \right) m_0 + V_0 \tilde{\mathcal{A}}_{\delta}^* (\tilde{\mathcal{A}}_{\delta} V_0 \tilde{\mathcal{A}}_{\delta}^*)^{-1} F^{(1)} \quad (12)$$

To make progress we must now start computing various terms needed for our mean and covariance update rules. We start with $V_0 \tilde{\mathcal{A}}_{\delta}^*$. We have:

$$V_0 \tilde{\mathcal{A}}_{\delta}^* = V_0 \mathcal{L}^* (I_1(\delta)^*, \dots, I_J(\delta)^*)$$

We can thus see that we need to be able to compute terms of the form $V_0 \mathcal{L}^* I_i(\delta)^* = V_0(I_i(\delta) \mathcal{L})^*$. Now since the operator $I_i(\delta) \mathcal{L}$ takes in a function on $\Omega \times [0, T]$ and outputs a real number its adjoint should take in a real number and output a function on $\Omega \times [0, T]$. This adjoint should satisfy the following relation:

$$\alpha(I_i(\delta) \mathcal{L} g) = \int_{\Omega} \int_0^T ((I_i(\delta) \mathcal{L})^* \alpha)(x, t) g(x, t) dt dx \quad \forall g, \quad \forall \alpha \in \mathbb{R} \quad (13)$$

Using this we can now compute:

$$\begin{aligned} (V_0(I_i(\delta) \mathcal{L})^* \alpha)(x, s) &= \int_{\Omega} \int_0^T k_{x,s,y,w}^{(0)} ((I_i(\delta) \mathcal{L})^* \alpha)(y, w) dw dy \\ &= \alpha(I_i(\delta) \mathcal{L} k_{x,s,\cdot,\cdot}^{(0)}) \\ &= \alpha \int_{\Omega} \phi_i(y) (\mathcal{L} k_{x,s,\cdot,\cdot}^{(0)})(y, \delta) dy \end{aligned}$$

We now work out $(\mathcal{L} k_{x,s,\cdot,\cdot}^{(0)})(y, \delta)$ taking care to remember that x, s are fixed and so \mathcal{L} acts on the variables y, δ :

$$\begin{aligned} (\mathcal{L} k_{x,s,\cdot,\cdot}^{(0)})(y, \delta) &= \partial_2 k^{(0)}(s, \delta) \sum_{j=1}^J \lambda_j \phi_j(x) \phi_j(y) - k^{(0)}(s, \delta) \nabla_y \cdot \left(a(y) \nabla_y \sum_{j=1}^J (\lambda_j \phi_j(x) \phi_j(y)) \right) \\ &= \partial_2 k^{(0)}(s, \delta) \sum_{j=1}^J \lambda_j \phi_j(x) \phi_j(y) - k^{(0)}(s, \delta) \sum_{j=1}^J \lambda_j \phi_j(x) \nabla_y \cdot (a(y) \nabla_y \phi_j(y)) \end{aligned}$$

So we can now compute:

$$\begin{aligned} (V_0(I_i(\delta) \mathcal{L})^* \alpha)(x, s) &= \alpha \int_{\Omega} \phi_i(y) \partial_2 k^{(0)}(s, \delta) \sum_{j=1}^J \lambda_j \phi_j(x) \phi_j(y) dy - \alpha \int_{\Omega} \phi_i(y) k^{(0)}(s, \delta) \sum_{j=1}^J \lambda_j \phi_j(x) \nabla_y \cdot (a(y) \nabla_y \phi_j(y)) dy \\ &= \alpha \partial_2 k^{(0)}(s, \delta) \sum_{j=1}^J \lambda_j \phi_j(x) M_{ji} + \alpha k^{(0)}(s, \delta) \sum_{j=1}^J \lambda_j \phi_j(x) A_{ji} \end{aligned}$$

where M is the Galerkin mass matrix and A is the Galerkin stiffness matrix, i.e. the matrices with entries given by:

$$M_{ij} := \int_{\Omega} \phi_i(x) \phi_j(x) dx \quad (14)$$

$$A_{ij} := \int_{\Omega} a(x) \nabla \phi_i(x) \nabla \phi_j(x) dx \quad (15)$$

Using this result we can deduce that:

$$(V_0 \tilde{A}_{\delta}^* \mathbf{v})(x, s) = \partial_2 k^{(0)}(s, \delta) \Phi(x)^* \Lambda M \mathbf{v} + k^{(0)}(s, \delta) \Phi(x)^* \Lambda A \mathbf{v} \quad (16)$$

for any $\mathbf{v} \in \mathbb{R}^J$, where $\Lambda = \text{diag}\{\lambda_i\}_{i=1}^J$.

We now move onto computing:

$$\begin{aligned} \tilde{A}_{\delta} V_0 \tilde{A}_{\delta}^* &= \mathcal{I}_{\delta} \mathcal{L} V_0 \mathcal{L}^* \mathcal{I}_{\delta}^* \\ &= \begin{pmatrix} I_1(\delta) \\ \vdots \\ I_J(\delta) \end{pmatrix} \mathcal{L} V_0 \mathcal{L}^* \begin{pmatrix} I_1(\delta)^* & \dots & I_J(\delta)^* \end{pmatrix} \end{aligned}$$

This operator has ij -th entry which is given by:

$$\begin{aligned}
(\tilde{A}_\delta V_0 \tilde{A}_\delta^*)_{ij} \alpha &= I_i(\delta) \mathcal{L} V_0 \mathcal{L}^* I_j(\delta)^* \alpha \\
&= \int_{\Omega} \phi_i(x) [(\mathcal{L} V_0 (I_j(\delta) \mathcal{L})^* \alpha)(x, \delta)] dx \\
&= \int_{\Omega} \phi_i(x) \left[\alpha \partial_1 \partial_2 k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l \phi_l(x) M_{lj} + \alpha \partial_1 k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l \phi_l(x) A_{lj} \right. \\
&\quad \left. - \alpha \partial_2 k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l M_{lj} \nabla \cdot (a(x) \nabla \phi_l(x)) - \alpha k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l A_{lj} \nabla \cdot (a(x) \nabla \phi_l(x)) \right] dx \\
&= \alpha \partial_1 \partial_2 k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l M_{il} M_{lj} + \alpha \partial_1 k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l M_{il} A_{lj} \\
&\quad + \alpha \partial_2 k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l M_{lj} A_{il} + \alpha k^{(0)}(\delta, \delta) \sum_{l=1}^J \lambda_l A_{il} A_{lj} \\
&= \alpha \partial_1 \partial_2 k^{(0)}(\delta, \delta) (M \Lambda M)_{ij} + \alpha \partial_1 k^{(0)}(\delta, \delta) (M \Lambda A)_{ij} + \alpha \partial_2 k^{(0)}(\delta, \delta) (A \Lambda M)_{ij} + \alpha k^{(0)}(\delta, \delta) (A \Lambda A)_{ij}
\end{aligned}$$

We can thus conclude that $\tilde{A}_\delta V_0 \tilde{A}_\delta^*$ is the $J \times J$ matrix given by:

$$\tilde{A}_\delta V_0 \tilde{A}_\delta^* = \partial_1 \partial_2 k^{(0)}(\delta, \delta) M \Lambda M + \partial_1 k^{(0)}(\delta, \delta) M \Lambda A + \partial_2 k^{(0)}(\delta, \delta) A \Lambda M + k^{(0)}(\delta, \delta) A \Lambda A \quad (17)$$

We now give the various derivatives for our time kernel which are required (**!!these should be checked rigorously!!**):

$$\begin{aligned}
\partial_1 k^{(0)}(t, s) &= \sum_{i=0}^{N-1} (s - t_i) \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s) \\
\partial_2 k^{(0)}(t, s) &= \sum_{i=0}^{N-1} (t - t_i) \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s) \\
\partial_1 \partial_2 k^{(0)}(t, s) &= \sum_{i=0}^{N-1} \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s)
\end{aligned}$$

We thus have:

$$\begin{aligned}
k^{(0)}(\delta, \delta) &= \delta^2 \\
\partial_1 k^{(0)}(\delta, \delta) &= \delta \\
\partial_2 k^{(0)}(\delta, \delta) &= \delta \\
\partial_1 \partial_2 k^{(0)}(\delta, \delta) &= 1
\end{aligned}$$

So we have:

$$\begin{aligned}
\tilde{A}_\delta V_0 \tilde{A}_\delta^* &= M \Lambda M + \delta M \Lambda A + \delta A \Lambda M + \delta^2 A \Lambda A \\
&= M \Lambda (M + \delta A) + \delta A \Lambda (M + \delta A) \\
&= (M \Lambda + \delta A \Lambda) (M + \delta A) \\
&= (M + \delta A) \Lambda (M + \delta A) = Q \Lambda Q
\end{aligned}$$

where we have defined $Q := M + \delta A$. To make progress we now simplify equation (16) using the derivatives noting the following:

$$\begin{aligned}
k^{(0)}(s, \delta) &= s \delta \mathbb{1}_{(0, \delta]}(s) \\
\partial_2 k^{(0)}(s, \delta) &= s \mathbb{1}_{(0, \delta]}(s)
\end{aligned}$$

We thus have:

$$\begin{aligned}
(V_0 \tilde{A}_\delta^* \mathbf{v})(x, s) &= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \Lambda M \mathbf{v} + s \delta \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \Lambda A \mathbf{v} \\
&= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \Lambda (M + \delta A) \mathbf{v} \\
&= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \Lambda Q \mathbf{v}
\end{aligned} \tag{18}$$

We can now make progress with the mean update equation. We first work out the following term using (18):

$$\begin{aligned}
(V_0 \tilde{A}_\delta^* (\tilde{A}_\delta V_0 \tilde{A}_\delta^*)^{-1} F^{(1)})(x, s) &= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \Lambda Q (Q \Lambda Q)^{-1} F^{(1)} \\
&= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \Lambda Q Q^{-1} \Lambda^{-1} Q^{-1} F^{(1)} \\
&= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* Q^{-1} F^{(1)}
\end{aligned}$$

For the other term involving m_0 in the mean update equation we must first work out $\tilde{A}_\delta m_0 = \mathcal{I}_\delta \mathcal{L} m_0$. To do this we compute:

$$\begin{aligned}
(\mathcal{L} m_0)(x, t) &= \partial_t m_0 - \nabla \cdot (a(x) \nabla m_0(x, t)) \\
&= -\nabla(a(x) \nabla \Phi(x)^* \gamma) \\
&= -\sum_{j=1}^J \gamma_j \nabla \cdot (a(x) \nabla \phi_j(x))
\end{aligned}$$

Thus, the i -th entry of $\tilde{A}_\delta m_0$ can be computed as:

$$\begin{aligned}
(\tilde{A}_\delta m_0)_i &= I_i(\delta) \mathcal{L} m_0 \\
&= \int_{\Omega} \phi_i(x) \left(-\sum_{j=1}^J \gamma_j \nabla \cdot (a(x) \nabla \phi_j(x)) \right) dx \\
&= \sum_{j=1}^J \gamma_j A_{ij} = (A \gamma)_i
\end{aligned}$$

So $\tilde{A}_\delta m_0 = A \gamma$. We can now compute the second term involving m_0 as follows (again using (18)):

$$\begin{aligned}
((V_0 \tilde{A}_\delta^* (\tilde{A}_\delta V_0 \tilde{A}_\delta^*)^{-1} \tilde{A}_\delta) m_0)(x, s) &= (V_0 \tilde{A}_\delta^* (Q \Lambda Q)^{-1} A \gamma)(x, s) \\
&= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \Lambda Q Q^{-1} \Lambda^{-1} Q^{-1} A \gamma \\
&= s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* Q^{-1} A \gamma
\end{aligned}$$

Thus we can compute:

$$\begin{aligned}
(1 - (V_0 \tilde{A}_\delta^* (\tilde{A}_\delta V_0 \tilde{A}_\delta^*)^{-1} \tilde{A}_\delta) m_0)(x, s) &= \Phi(x)^* \gamma - s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* Q^{-1} A \gamma \\
&= \Phi(x)^* Q^{-1} [Q - s \mathbb{1}_{(0, \delta]}(s) A] \gamma \\
&= \Phi(x)^* Q^{-1} [M + \delta A - s \mathbb{1}_{(0, \delta]}(s) A] \gamma \\
&= \Phi(x)^* Q^{-1} [M + (\delta - s \mathbb{1}_{(0, \delta]}(s)) A] \gamma
\end{aligned}$$

Putting this all together we obtain:

$$m_1(x, s) = \Phi(x)^* Q^{-1} [M + (\delta - s \mathbb{1}_{(0, \delta]}(s)) A] \gamma + s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* Q^{-1} F^{(1)} \tag{19}$$

Thus we see that performing this mean update and then evaluating at the time $t_1 = \delta$ we obtain the following:

$$m_1(x, t_1) = m_1(x, \delta) = \Phi(x)^* \left(Q^{-1} M \gamma + \delta Q^{-1} F^{(1)} \right) \tag{20}$$

Note that this is the same as the update equation for the coefficients in the backward-Euler Galerkin method:

$$\gamma \mapsto (M + \delta A)^{-1} M \gamma + \delta (M + \delta A)^{-1} F^{(1)} \tag{21}$$

!!! Before proceeding to compute the covariance V_1 we first note that m_1 can be expressed as follows:

$$\begin{aligned} m_1(x, s) &= c(x, s) + \Psi^* Q^{-1} \mathcal{I}_\delta f \\ \text{i.e. } m_1 &= Lf + c \end{aligned}$$

where $c(x, s) = \Phi(x)^* Q^{-1} [M + (\delta - s \mathbb{1}_{(0, \delta]}(s))A] \gamma$, Ψ^* is the following function from $\mathbb{R}^J \rightarrow L^2(\Omega \times [0, T])$ given by $(\Psi^* \mathbf{v})(x, s) := s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \mathbf{v}$ and $L := \Psi^* Q^{-1} \mathcal{I}_\delta$. This observation will prove useful later when we marginalise over the noise term. !!!

We can now move on to computing the covariance V_1 . We start by computing $\tilde{A}_\delta V_0$. Computing this involves determining how $I_j(\delta) \mathcal{L} V_0$ acts on functions g for $j = 1, \dots, J$. We have:

$$I_j(\delta) \mathcal{L} V_0 g = \int_{\Omega} \phi_j(x) (\mathcal{L} V_0 g)(x, \delta) dx$$

Now recalling that $V_0 g(x, \delta) = \int_{\Omega} \int_0^T k_{x\delta, ys}^{(0)} g(y, s) ds dy$ we deduce:

$$\begin{aligned} (\mathcal{L} V_0 g)(x, \delta) &= \int_{\Omega} \int_0^T (\mathcal{L} k_{\cdot, y, s}^{(0)})(x, \delta) g(y, s) ds dy \\ &= \int_{\Omega} \int_0^T \left(\partial_1 k^{(0)}(\delta, s) \sum_{i=1}^J \lambda_i \phi_i(x) \phi_i(y) - k^{(0)}(\delta, s) \sum_{i=1}^J \lambda_i \nabla_x \cdot (a(x) \nabla_x \phi_i(x)) \phi_i(y) \right) g(y, s) ds dy \end{aligned}$$

We can now perform the integration to obtain:

$$\begin{aligned} I_j(\delta) \mathcal{L} V_0 g &= \int_{\Omega} \phi_j(x) \left(\int_{\Omega} \int_0^T \left(\partial_1 k^{(0)}(\delta, s) \sum_{i=1}^J \lambda_i \phi_i(x) \phi_i(y) - k^{(0)}(\delta, s) \sum_{i=1}^J \lambda_i \nabla_x \cdot (a(x) \nabla_x \phi_i(x)) \phi_i(y) \right) g(y, s) ds dy \right) dx \\ &= \int_{\Omega} \int_0^T \left(\partial_1 k^{(0)}(\delta, s) \sum_{i=1}^J \lambda_i M_{ij} \phi_i(y) g(y, s) + k^{(0)}(\delta, s) \sum_{i=1}^J \lambda_i A_{ij} \phi_i(y) g(y, s) \right) ds dy \\ &= \sum_{i=1}^J \lambda_i \int_{\Omega} \int_0^T (\partial_1 k^{(0)}(\delta, s) M_{ij} + k^{(0)}(\delta, s) A_{ij}) \phi_i(y) g(y, s) ds dy \\ &= \sum_{i=1}^J \lambda_i \left[\int_0^T M_{ij} \partial_1 k^{(0)}(\delta, s) (I_i(s) g) ds + \int_0^T A_{ij} k^{(0)}(\delta, s) (I_i(s) g) ds \right] \\ &= \sum_{i=1}^J \lambda_i \left[M_{ij} \left(\int_0^T \partial_1 k^{(0)}(\delta, s) (\mathcal{I}_s g) ds \right)_i + A_{ij} \left(\int_0^T k^{(0)}(\delta, s) (\mathcal{I}_s g) ds \right)_i \right] \\ &= \left(M \Lambda \int_0^T \partial_1 k^{(0)}(\delta, s) (\mathcal{I}_s g) ds + A \Lambda \int_0^T k^{(0)}(\delta, s) (\mathcal{I}_s g) ds \right)_j \end{aligned}$$

Thus we can deduce:

$$\tilde{A}_\delta V_0 g = M \Lambda \int_0^T \partial_1 k^{(0)}(\delta, s) (\mathcal{I}_s g) ds + A \Lambda \int_0^T k^{(0)}(\delta, s) (\mathcal{I}_s g) ds \quad (22)$$

We now utilise the specific form of the time kernel:

$$\begin{aligned} k^{(0)}(\delta, s) &= s \delta \mathbb{1}_{(0, \delta]}(s) \\ \partial_1 k^{(0)}(\delta, s) &= s \mathbb{1}_{(0, \delta]}(s) \end{aligned}$$

to further deduce:

$$\begin{aligned} \tilde{A}_\delta V_0 g &= M \Lambda \int_0^T s \mathbb{1}_{(0, \delta]}(s) (\mathcal{I}_s g) ds + A \Lambda \int_0^T s \delta \mathbb{1}_{(0, \delta]}(s) (\mathcal{I}_s g) ds \\ &= Q \Lambda \nu_g^{(0)} \end{aligned}$$

where:

$$\boldsymbol{\nu}_g^{(i)} := \int_{t_i}^{t_{i+1}} (s - t_i) (\mathcal{I}_s g) ds \text{ for } i = 0, \dots, N-1 \quad (23)$$

Having worked this out we can now compute the second term in the formula for V_1 as follows:

$$\begin{aligned} (V_0 \tilde{A}_\delta^* (\tilde{A}_\delta V_0 \tilde{A}_\delta^*)^{-1} \tilde{A}_\delta V_0 g)(x, t) &= (V_0 \tilde{A}_\delta^* (Q \Lambda Q)^{-1} Q \Lambda \boldsymbol{\nu}_g^{(0)})(x, t) \\ &= t \mathbb{1}_{(0, \delta]}(t) \Phi(x)^* \Lambda Q Q^{-1} \Lambda^{-1} Q^{-1} Q \Lambda \boldsymbol{\nu}_g^{(0)} \\ &= t \mathbb{1}_{(0, \delta]}(t) \Phi(x)^* \Lambda \boldsymbol{\nu}_g^{(0)} \end{aligned}$$

One can also easily show that the action of V_0 on functions can be rewritten as follows:

$$(V_0 g)(x, t) = \sum_{i=0}^{N-1} (t - t_i) \mathbb{1}_{(t_i, t_{i+1}]}(t) \Phi(x)^* \Lambda \boldsymbol{\nu}_g^{(i)} \quad (24)$$

and so we can deduce that V_1 is given by:

$$(V_1 g)(x, t) = \sum_{i=1}^{N-1} (t - t_i) \mathbb{1}_{(t_i, t_{i+1}]}(t) \Phi(x)^* \Lambda \boldsymbol{\nu}_g^{(i)}$$

This can be rewritten as follows:

$$(V_1 g)(x, t) = \int_{\Omega} \int_0^T k_{x,t,y,s}^{(1)} g(y, s) ds dy \quad (25)$$

where the kernel is now:

$$k_{x,t,y,s}^{(1)} := \sum_{j=1}^J \lambda_j \phi_j(x) \phi_j(y) k^{(1)}(t, s) \quad (26)$$

with time kernel:

$$k^{(1)}(t, s) := \sum_{i=1}^{N-1} (t - t_i)(s - t_i) \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s) \quad (27)$$

We can thus see that the new covariance operator is of the same form as the prior covariance; the only difference is the time kernel has changed by simply removing the term corresponding to the first time interval $(t_0, t_1]$. This observation will allow us to very easily move to the next time step $t = t_2 = 2\delta$. To do so we now condition on observing $\mathcal{I}_{2\delta} \mathcal{L} u = \mathcal{I}_{2\delta} f =: F^{(2)}$. As before let $\tilde{A}_{2\delta} = \mathcal{I}_{2\delta} \mathcal{L}$. For the same fixed realisation of f we seek the following conditional distribution:

$$u | \{ \tilde{A}_\delta u = F^{(1)}, \tilde{A}_{2\delta} u = F^{(2)}, f \} \sim \mathcal{N}(m_2, V_2) \quad (28)$$

where we have the update equations:

$$m_2 = m_1 + V_1 \tilde{A}_{2\delta}^* (\tilde{A}_{2\delta} V_1 \tilde{A}_{2\delta}^*)^{-1} (F^{(2)} - \tilde{A}_{2\delta} m_1) \quad (29)$$

$$V_2 = V_1 - V_1 \tilde{A}_{2\delta}^* (\tilde{A}_{2\delta} V_1 \tilde{A}_{2\delta}^*)^{-1} \tilde{A}_{2\delta} V_1 \quad (30)$$

Since the covariance operator V_1 has the same form as V_0 but with a different time kernel most of the computations from before are the same. We give the results, and note where there are differences. First we have the analogue of (16):

$$(V_1 \tilde{A}_{2\delta}^* \mathbf{v})(x, s) = \partial_2 k^{(1)}(s, 2\delta) \Phi(x)^* \Lambda M \mathbf{v} + k^{(1)}(s, 2\delta) \Phi(x)^* \Lambda A \mathbf{v} \quad (31)$$

The analogue of (17) is given by:

$$\tilde{A}_{2\delta} V_1 \tilde{A}_{2\delta}^* = \partial_1 \partial_2 k^{(1)}(2\delta, 2\delta) M \Lambda M + \partial_1 k^{(1)}(2\delta, 2\delta) M \Lambda A + \partial_2 k^{(1)}(2\delta, 2\delta) A \Lambda M + k^{(1)}(2\delta, 2\delta) A \Lambda A \quad (32)$$

The various derivatives for our new time kernel follow almost exactly the same as before:

$$\begin{aligned}\partial_1 k^{(1)}(t, s) &= \sum_{i=1}^{N-1} (s - t_i) \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s) \\ \partial_2 k^{(1)}(t, s) &= \sum_{i=1}^{N-1} (t - t_i) \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s) \\ \partial_1 \partial_2 k^{(1)}(t, s) &= \sum_{i=1}^{N-1} \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s)\end{aligned}$$

We thus have:

$$\begin{aligned}k^{(1)}(\delta, \delta) &= \delta^2 \\ \partial_1 k^{(1)}(\delta, \delta) &= \delta \\ \partial_2 k^{(1)}(\delta, \delta) &= \delta \\ \partial_1 \partial_2 k^{(1)}(\delta, \delta) &= 1 \\ k^{(1)}(s, 2\delta) &= \delta(s - \delta) \mathbb{1}_{(\delta, 2\delta]}(s) \\ \partial_2 k^{(1)}(s, 2\delta) &= (s - \delta) \mathbb{1}_{(\delta, 2\delta]}(s)\end{aligned}$$

Using this we can now simplify (31) to:

$$(V_1 \tilde{A}_{2\delta}^* \mathbf{v})(x, s) = (s - \delta) \mathbb{1}_{(\delta, 2\delta]}(s) \Phi(x)^* \Lambda Q \mathbf{v} \quad (33)$$

and (32) to:

$$\tilde{A}_{2\delta} V_1 \tilde{A}_{2\delta}^* = Q \Lambda Q \quad (34)$$

We now need to compute the term $\tilde{A}_{2\delta} m_1$. Doing this will involve first computing $\mathcal{L}m_1$. Rewriting m_1 as follows will help with this:

$$\begin{aligned}m_1(x, s) &= \Phi(x)^* Q^{-1} [M + \delta A] \gamma + s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* Q^{-1} [F^{(1)} - A\gamma] \\ &= \Phi(x)^* \gamma + s \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \mathbf{c}^{(1)}\end{aligned}$$

where $\mathbf{c}^{(1)} := Q^{-1} [F^{(1)} - A\gamma]$. We can now compute more easily:

$$(\mathcal{L}m_1)(x, s) = \mathbb{1}_{(0, \delta]}(s) \Phi(x)^* \mathbf{c}^{(1)} - \sum_{j=1}^J (\gamma + s \mathbb{1}_{(0, \delta]}(s) \mathbf{c}^{(1)})_j \nabla \cdot (a(x) \nabla \phi_j(x))$$

This allows us to work out the i -th entry of $\tilde{A}_{2\delta} m_1$ as follows:

$$\begin{aligned}(\tilde{A}_{2\delta} m_1)_i &= I_i(2\delta) \mathcal{L}m_1 \\ &= \int_{\Omega} \phi_i(x) (\mathcal{L}m_1)(x, 2\delta) dx \\ &= - \int_{\Omega} \phi_i(x) \sum_{j=1}^J \gamma_j \nabla \cdot (a(x) \nabla \phi_j(x)) dx \\ &= \sum_{j=1}^J \gamma_j A_{ij} \\ &= (A\gamma)_i\end{aligned}$$

So $\tilde{A}_{2\delta} m_1 = A\gamma$. These results enable us to work out the following:

$$\begin{aligned}(V_1 \tilde{A}_{2\delta}^* (\tilde{A}_{2\delta} V_1 \tilde{A}_{2\delta}^*)^{-1} F^{(2)})(x, s) &= (s - t_1) \mathbb{1}_{(t_1, t_2]}(s) \Phi(x)^* Q^{-1} F^{(2)} \\ (V_1 \tilde{A}_{2\delta}^* (\tilde{A}_{2\delta} V_1 \tilde{A}_{2\delta}^*)^{-1} \tilde{A}_{2\delta} m_1)(x, s) &= (s - t_1) \mathbb{1}_{(t_1, t_2]}(s) \Phi(x)^* Q^{-1} A\gamma\end{aligned}$$

Combining these results yields the new mean:

$$\begin{aligned}
m_2(x, s) &= m_1(x, s) + (s - t_1)\mathbb{1}_{(t_1, t_2]} \Phi(x)^* Q^{-1} [F^{(2)} - A\gamma] \\
&= m_1(x, s) + (s - t_1)\mathbb{1}_{(t_1, t_2]}(s) \Phi(x)^* \mathbf{c}^{(2)} \\
&= \Phi(x)^* \gamma + (s - t_0)\mathbb{1}_{(t_0, t_1]}(s) \Phi(x)^* \mathbf{c}^{(1)} + (s - t_1)\mathbb{1}_{(t_1, t_2]}(s) \Phi(x)^* \mathbf{c}^{(2)} \\
&= \Phi(x)^* \gamma + \sum_{i=1}^2 (s - t_{i-1})\mathbb{1}_{(t_{i-1}, t_i]}(s) \Phi(x)^* \mathbf{c}^{(i)}
\end{aligned}$$

Evaluating this new mean at time $t = t_2 = 2\delta$ we obtain:

$$\begin{aligned}
m_2(x, t_2) &= m_2(x, 2\delta) = \Phi(x)^* \gamma + (2\delta - \delta)\Phi(x)^* \mathbf{c}^{(2)} \\
&= \Phi(x)^* \gamma + \delta \Phi(x)^* Q^{-1} [F^{(2)} - A\gamma] \\
&= \Phi(x)^* Q^{-1} [Q\gamma + \delta F^{(2)} - \delta A\gamma] \\
&= \Phi(x)^* Q^{-1} [(Q - \delta A)\gamma + \delta F^{(2)}] \\
&= \Phi(x)^* Q^{-1} [M\gamma + \delta F^{(2)}]
\end{aligned}$$

We now move onto computing V_2 . Since V_1 is of the same form as V_0 with a different time kernel, the computations for V_1 can be reused almost identically to compute V_2 . We have:

$$\begin{aligned}
\tilde{A}_{2\delta} V_0 g &= M\Lambda \int_0^T \partial_1 k^{(1)}(2\delta, s)(\mathcal{I}_s g) ds + A\Lambda \int_0^T k^{(1)}(2\delta, s)(\mathcal{I}_s g) ds \\
&= M\Lambda \int_0^T (s - t_1)\mathbb{1}_{(t_1, t_2]}(s)(\mathcal{I}_s g) ds + \delta A\Lambda \int_0^T (s - t_1)\mathbb{1}_{(t_1, t_2]}(s)(\mathcal{I}_s g) ds \\
&= M\Lambda \int_{t_1}^{t_2} (s - t_1)(\mathcal{I}_s g) ds + \delta A\Lambda \int_{t_1}^{t_2} (s - t_1)(\mathcal{I}_s g) ds \\
&= M\Lambda \boldsymbol{\nu}_g^{(1)} + \delta A\Lambda \boldsymbol{\nu}_g^{(1)} \\
&= Q\Lambda \boldsymbol{\nu}_g^{(1)}
\end{aligned}$$

Having worked this out we can now compute the second term in the formula for V_2 as follows:

$$\begin{aligned}
(V_1 \tilde{A}_{2\delta}^* (\tilde{A}_{2\delta} V_1 \tilde{A}_{2\delta}^*)^{-1} \tilde{A}_{2\delta} V_1 g)(x, t) &= (V_1 \tilde{A}_{2\delta}^* (Q\Lambda Q)^{-1} Q\Lambda \boldsymbol{\nu}_g^{(1)})(x, t) \\
&= (t - t_1)\mathbb{1}_{(t_1, t_2]}(t) \Phi(x)^* \Lambda Q Q^{-1} \Lambda^{-1} Q^{-1} Q\Lambda \boldsymbol{\nu}_g^{(1)} \\
&= (t - t_1)\mathbb{1}_{(t_1, t_2]}(t) \Phi(x)^* \Lambda \boldsymbol{\nu}_g^{(1)}
\end{aligned}$$

Thus, we can conclude that V_2 acts on functions g as follows:

$$(V_2 g)(x, t) = \sum_{i=2}^{N-1} (t - t_i)\mathbb{1}_{(t_i, t_{i+1}]}(t) \Phi(x)^* \Lambda \boldsymbol{\nu}_g^{(i)} = \int_{\Omega} \int_0^T k_{x,t,y,s}^{(2)}(y, s) ds dy \quad (35)$$

where the kernel is now:

$$k_{x,t,y,s}^{(2)} := \sum_{i=2}^J \lambda_j \phi_j(x) \phi_j(y) k^{(2)}(t, s) \quad (36)$$

with time kernel:

$$k^{(2)}(t, s) := \sum_{i=2}^{N-1} (t - t_i)(s - t_i)\mathbb{1}_{(t_i, t_{i+1}]}(t, s) \quad (37)$$

We again can see that the form of the covariance operator remains the same as the previous one. We can now hypothesize a formula for conditioning up to time t_p for $p \in \{1, \dots, N\}$:

$$u | \{ \tilde{A}_{t_1} u = F^{(1)}, \dots, \tilde{A}_{t_p} u = F^{(p)}, f \} \sim \mathcal{N}(m_p, V_p) \quad (38)$$

where $F^{(p)} := \mathcal{I}_{t_p} f = \mathcal{I}_{p\delta} f$ and where the mean and covariance are given by:

$$m_p(x, t) := \Phi(x)^* \gamma + \sum_{i=1}^p (t - t_{i-1}) \mathbb{1}_{(t_{i-1}, t_i]}(t) \Phi(x)^* \mathbf{c}^{(i)} \quad (39)$$

$$(V_p g)(x, t) := \int_{\Omega} \int_0^T k_{x,t,y,s}^{(p)} g(y, s) ds dy \quad (40)$$

$$\mathbf{c}^{(i)} := Q^{-1} [F^{(i)} - A\gamma] \quad (41)$$

$$k_{x,t,y,s}^{(p)} := \sum_{i=p}^{N-1} \lambda_j \phi_j(x) \phi_j(y) k^{(p)}(t, s) \quad (42)$$

$$k^{(p)}(t, s) := \sum_{i=p}^{N-1} (t - t_i)(s - t_i) \mathbb{1}_{(t_i, t_{i+1}]^2}(t, s) \quad (43)$$

This can probably be shown by induction. The above implies that conditioning at all N time points will yield the following degenerate Gaussian distribution:

$$u | \{\tilde{A}_{t_1} u = F^{(1)}, \dots, \tilde{A}_{t_N} u = F^{(N)}, f\} \sim \mathcal{N}(m_N, 0) = \delta_{m_N} \quad (44)$$

since $V_N = 0$. This is a Dirac point mass located at the function:

$$m_N(x, t) = \Phi(x)^* \gamma + \sum_{i=1}^N (t - t_{i-1}) \mathbb{1}_{(t_{i-1}, t_i]}(t) \Phi(x)^* \mathbf{c}^{(i)} \quad (45)$$

We will now marginalise over f in order to obtain the averaged conditional distribution. To do this we will need the following Lemma (which we prove below):

Lemma 0.1. Let $f \sim \mathcal{N}(\bar{f}, K)$ where we assume that this Gaussian measure is on a Hilbert space of functions $\mathcal{H}_1 \subset \mathbb{R}^{\mathcal{X}}$. Suppose that for a fixed realisation of f we have

$$y|f \sim \mathcal{N}(Lf + c, V)$$

where L is a bounded linear operator from \mathcal{H}_1 to another Hilbert space $\mathcal{H}_2 \subset \mathbb{R}^{\mathcal{X}}$ (so y lies in \mathcal{H}_2), c is a deterministic function and the covariance operator V does not depend on f . Then marginalizing over f yields:

$$y \sim \mathcal{N}(L\bar{f} + c, LKL^* + V)$$

as the averaged distribution of y .

Proof: The fact that $y|f \sim \mathcal{N}(Lf + c, V)$ is equivalent to saying that:

$$y = Lf + c + \tilde{y}$$

where $\tilde{y} \sim \mathcal{N}(0, V)$ is independent of f . Thus we have:

$$\begin{pmatrix} f \\ \tilde{y} \end{pmatrix} = \mathcal{N}\left(\begin{pmatrix} \bar{f} \\ 0 \end{pmatrix}, \begin{pmatrix} K & 0 \\ 0 & V \end{pmatrix}\right)$$

Since we can write:

$$y = \begin{pmatrix} L & 1 \end{pmatrix} \begin{pmatrix} f \\ \tilde{y} \end{pmatrix} + c$$

we deduce:

$$y \sim \mathcal{N}\left(L\bar{f} + c, \begin{pmatrix} L & 1 \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} L^* \\ 1 \end{pmatrix}\right) = \mathcal{N}(L\bar{f} + c, LKL^* + V)$$

as required. ■

We can now perform the marginalisation over f noting that V_1 does not depend on f to obtain:

$$\int u|\{\tilde{A}_\delta u = F^{(1)}, f\}df \sim \mathcal{N}(L\bar{f} + c, LKL^* + V_1) =: \mathcal{N}(\bar{m}_1, \bar{V}_1) \quad (46)$$

We now compute \bar{m}_1 and \bar{V}_1 . It is easy to see that \bar{m}_1 is the same as m_1 except F^1 is replaced by $\bar{F}^1 := \mathcal{I}_\delta \bar{f}$. For the covariance we must figure out how L^* acts. This will require the adjoint Ψ of Ψ^* which can be shown to be given by $\Psi g = \nu_g^{(0)}$. We thus can compute:

$$LKL^* = \Psi^* Q^{-1} \mathcal{I}_\delta K \mathcal{I}_\delta^* Q^{-1} \Psi = \Psi^* H^{(1)} \Psi \quad (47)$$

where $H^{(1)} := Q^{-1} \mathcal{I}_\delta K \mathcal{I}_\delta^* Q^{-1}$ can be shown to be positive definite. We can now compute how LKL^* acts on a function g as follows:

$$\begin{aligned} (LKL^*g)(x, t) &= (\Psi^* H^{(1)} \nu_g^{(0)})(x, t) \\ &= t \mathbb{1}_{(0, \delta]}(t) \Phi(x)^* H^{(1)} \nu_g^{(0)} \end{aligned}$$

Thus, \bar{V}_1 acts on functions as:

$$\begin{aligned} (\bar{V}_1 g)(x, t) &= t \mathbb{1}_{(0, \delta]}(t) \Phi(x)^* H^{(1)} \nu_g^{(0)} + \sum_{i=1}^{N-1} (t - t_i) \mathbb{1}_{(t_i, t_{i+1}]}(t) \Phi(x)^* \Lambda \nu_g^{(i)} \\ &= \sum_{i=0}^{N-1} (t - t_i) \mathbb{1}_{(t_i, t_{i+1}]}(t) \Phi(x)^* H^{(1, i)} \nu_g^{(i)} \end{aligned}$$

where:

$$H^{(1, i)} := \begin{cases} H^{(1)} & i = 0 \\ \Lambda & i = 1, \dots, N-1 \end{cases}$$

We can now repeat this computation to move from time $t_1 = \delta$ to $t_2 = 2\delta$ by conditioning $\mathcal{N}(\bar{m}_1, \bar{V}_1)$ on $\tilde{A}_{2\delta} u = F^2 =: \mathcal{I}_{2\delta} f$.