parameters is challenging. The main reason is the restriction imposed on the support of the posterior distribution by the monotonicity constraint; an MCMC approach based on accept/reject steps is likely to be very inefficient when the support of the proposal distribution is very different from that of the target distribution. While an explicit soft constraint is defined on the derivative function space, the effect of this constraint on the covariance parameters' distribution is not obvious and therefore it is not trivial to define a proposal distribution that is likely to generate values for these parameters with high posterior probability. Full Gibbs sampling is not possible since the full conditional distributions cannot be obtained for the correlation parameters $\boldsymbol{l}$.

To overcome these difficulties we use a variant of SMC samplers [4]. SMC samplers take advantage of a sequence of distributions $\{\pi_t\}_{t=0}^{T}$ that bridge between a distribution that is straightforward to generate from (e.g., the prior) and the target distribution. By iterative weighting and sampling steps an initial sample generated from π_0 is filtered and evolved through the sequence of densities to obtain a sample from the target, π_T. Therefore, the required tools of an SMC sampler are a sequence of densities leading to the target distribution and proper weighting and sampling steps that guarantee sampling from the correct distribution at each time step.

Reference [8] introduced a variant of SMC, referred to as the SCMC, for the case that sampling from the target distribution is challenging due to imposition of constraints. The sequence of densities is defined based on the strictness of constraints. Starting from an initial distribution under which the constraints are relaxed fully or to an extent that sampling is feasible, the restriction is imposed gradually on the distributions. In the following we describe the SCMC algorithm tailored for sampling from (3.6).

As mentioned in section 3.3, the rigidity of the monotonicity constraint is controlled by the parameter τ. Larger values of τ more strictly constrain the partial derivatives to be positive at selected points. We use this property to define the filtering sequence of distributions. By specifying an increasing schedule over the monotonicity parameter, we move particles sampled from an unconstrained GP posterior towards the target model that has a large enough monotonicity parameter, say τ_T. Let the vector of parameters defining each particle be denoted by $\boldsymbol{\eta} = (\boldsymbol{l}, \sigma^2, \mathbf{y}^*, \mathbf{y}'_k)$. The tth posterior in the sequence is given by

$$\pi_t(\boldsymbol{\eta}) \propto [\boldsymbol{l}, \sigma^2][\mathbf{y}^*, \mathbf{y}'_k | \mathbf{y}, \boldsymbol{l}, \sigma^2][\mathbf{y}|\boldsymbol{l}, \sigma^2] \prod_i \Phi(\tau_t y'_k(\mathbf{x}_i^\delta)),$$

where

$$0 = \tau_0 < \tau_1 < \cdots < \tau_T \to \infty.$$

The SMC algorithm tailored for monotone interpolation is given in Algorithm 1. In step 1 of Algorithm 1, π_0 is chosen to be an unconstrained GP model corresponding to $\tau = 0$, which fully relaxes the positivity constraint on the derivatives. However, MCMC is needed to generate a sample of N draws from the unconstrained model. Typical Metropolis within Gibbs algorithms used to sample from a GP posterior can be found in the literature; see, for example, [2].

At time t, the weights are updated according to the current distribution using the incremental weights, $\tilde{w}_t^i$,

$$W_t^i = W_{t-1}^i \tilde{w}_t^i.$$

Algorithm 1. SMC for monotone emulation.

Input: a sequence of constraint parameters τ_t, $t = 1, \dots, T$,
Forward transition kernels K_t,
proposal distributions Q_{1t} and Q_{2t} for $\boldsymbol{l}$ and σ^2,
proposal step adjustment parameter q_t.

Generate an initial sample $\left(\boldsymbol{l}, \sigma^2, \mathbf{y}^*, \mathbf{y}'_k\right)_0^{1:N} \sim \pi_0$

$W_1^{1:N} \leftarrow \frac{1}{N}$

for $t := 1, \dots, T-1$ **do**

- **Weight calculation.** $W_t^i \leftarrow W_{t-1}^i \tilde{w}_t^i$, where $\tilde{w}_t^i = \frac{\prod_i \Phi\left(\tau_t y'_k(\mathbf{x}_i^\delta)\right)}{\prod_i \Phi\left(\tau_{t-1} y'_k(\mathbf{x}_i^\delta)\right)}$, $i = 1, \dots, N$, and normalize $W_t^{1:N}$.
- **Resampling.** If $\text{ESS}_t <= \frac{N}{2}$, then resample the particles $\left(\boldsymbol{l}, \sigma^2, \mathbf{y}^*, \mathbf{y}'_k\right)_t^{1:N}$ with weights $W_t^{1:N}$ and update $W_t^{1:N} \leftarrow \frac{1}{N}$.
- **Sampling.** Sample $\left(\boldsymbol{l}, \sigma^2, \mathbf{y}^*, \mathbf{y}'_k\right)_{t+1}^{1:N} \sim K_t\left(\left(\boldsymbol{l}, \sigma^2, \mathbf{y}^*, \mathbf{y}'_k\right)_t^{1:N}, .\right)$ through the following steps:
  - **for** $i := 1 \dots, N$ **do**
    - $\left(\boldsymbol{l}_t^i, \sigma_t^{2i}, \mathbf{y}_t^{*i}, \mathbf{y}_t'^i\right) \leftarrow \left(\boldsymbol{l}_{t-1}^i, \sigma_{t-1}^{2i}, \mathbf{y}_{t-1}^{*i}, \mathbf{y}_{t-1}'^i\right)$;
    - propose $\boldsymbol{l}^{\text{new}} \sim Q_{1t}\left(.|\boldsymbol{l}_t^i\right)$ and
      $\boldsymbol{l}_t^i \leftarrow \boldsymbol{l}^{\text{new}}$ with probability $p = \min\left\{1, \frac{\pi_t\left(\boldsymbol{l}^{\text{new}}, \sigma_t^{2i}, \mathbf{y}_t^{*i}, \mathbf{y}_t'^i\right)}{\pi_t\left(\boldsymbol{l}_t^i, \sigma_t^{2i}, \mathbf{y}_t^{*i}, \mathbf{y}_t'^i\right)}\right\}$;
    - propose $\sigma^{2\text{new}} \sim Q_{2t}\left(.|\sigma_t^{2i}\right)$ and
      $\sigma_t^{2i} \leftarrow \sigma^{2\text{new}}$ with probability $p = \min\left\{1, \frac{\pi_t\left(\boldsymbol{l}_t^i, \sigma^{2\text{new}}, \mathbf{y}_t^{*i}, \mathbf{y}_t'^i\right)}{\pi_t\left(\boldsymbol{l}_t^i, \sigma_t^{2i}, \mathbf{y}_t^{*i}, \mathbf{y}_t'^i\right)}\right\}$;
    - propose $(\mathbf{y}^*, \mathbf{y}')^{\text{new}} \sim \mathcal{N}\left((\mathbf{y}^*, \mathbf{y}')_t^i, q_t \Lambda_{\boldsymbol{l}_t^i}\right)$, where $\Lambda_{\boldsymbol{l}_t^i}$ is the correlation matrix with correlation parameter vector $\boldsymbol{l}_t^i$ and
      $(\mathbf{y}^*, \mathbf{y}')_t^i \leftarrow (\mathbf{y}^*, \mathbf{y}')^{\text{new}}$ with probability $p = \min\left\{1, \frac{\pi_t\left(\boldsymbol{l}_t^i, \sigma_t^{2i}, \mathbf{y}^{*\text{new}}, \mathbf{y}'^{\text{new}}\right)}{\pi_t\left(\boldsymbol{l}_t^i, \sigma_t^{2i}, \mathbf{y}_t^{*i}, \mathbf{y}_t'^i\right)}\right\}$
  - **end for**

end for

Return: Particles $\left(\boldsymbol{l}, \sigma^2, \mathbf{y}^*, \mathbf{y}'_k\right)_T^{1:N}$.

The simplified form of $\tilde{w}_t^i$ in Algorithm 1 is the result of the choice of the forward transition kernels K_t as an MCMC kernel of the invariant distribution π_t in the sampling step [4]. The proposal distributions, Q_{1t} and Q_{2t}, used in the sampling step are chosen to generate adequate values under π_t and depend on t in the generic SCMC algorithm. In the current application the same proposal distribution can be used at all time points, t. However, the proposal step size (the variance of the proposal distribution) is adjusted when progressing through the filtering sequence to prevent particle degeneracy by keeping the acceptance rate of the sampling moves above a lower threshold. Existing guidelines for adjusting proposal moves in the MCMC literature can be used. The proposal step size parameters q_t are chosen in the same manner. We discuss the specific choices made for our examples in section 6.