Deep Gaussian Processes

- Gaussian process model
- Deep Gaussian process model
  (and why we can't do it)
- Variational sparse Gaussian processes
- Deep Gaussian process model revisited

Data likelihood:

$$y_i = f(\mathbf{x}_i) + \epsilon, \qquad \epsilon \sim \mathcal{N}(0,\sigma^2)$$

$$p(\mathbf{y} \mid \mathbf{f}) = \mathcal{N}(\mathbf{y} \mid \mathbf{f}, \mathbf{I}\sigma^2),$$

$$\mathbf{f} = [f(x_1), f(x_2), \ldots, f(x_N)]^T = [f_1, f_2, \ldots, f_N]^T$$

Approach: Put prior on function and marginalise out

$$p(\mathbf{y} \mid \mathbf{x}) = \int_{\mathbf{f}} p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{x})$$

Two objectives:

1. Find marginal likelihood:

$$p(\mathbf{y} \mid \mathbf{x}) = \int_{\mathbf{f}} p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{x})$$

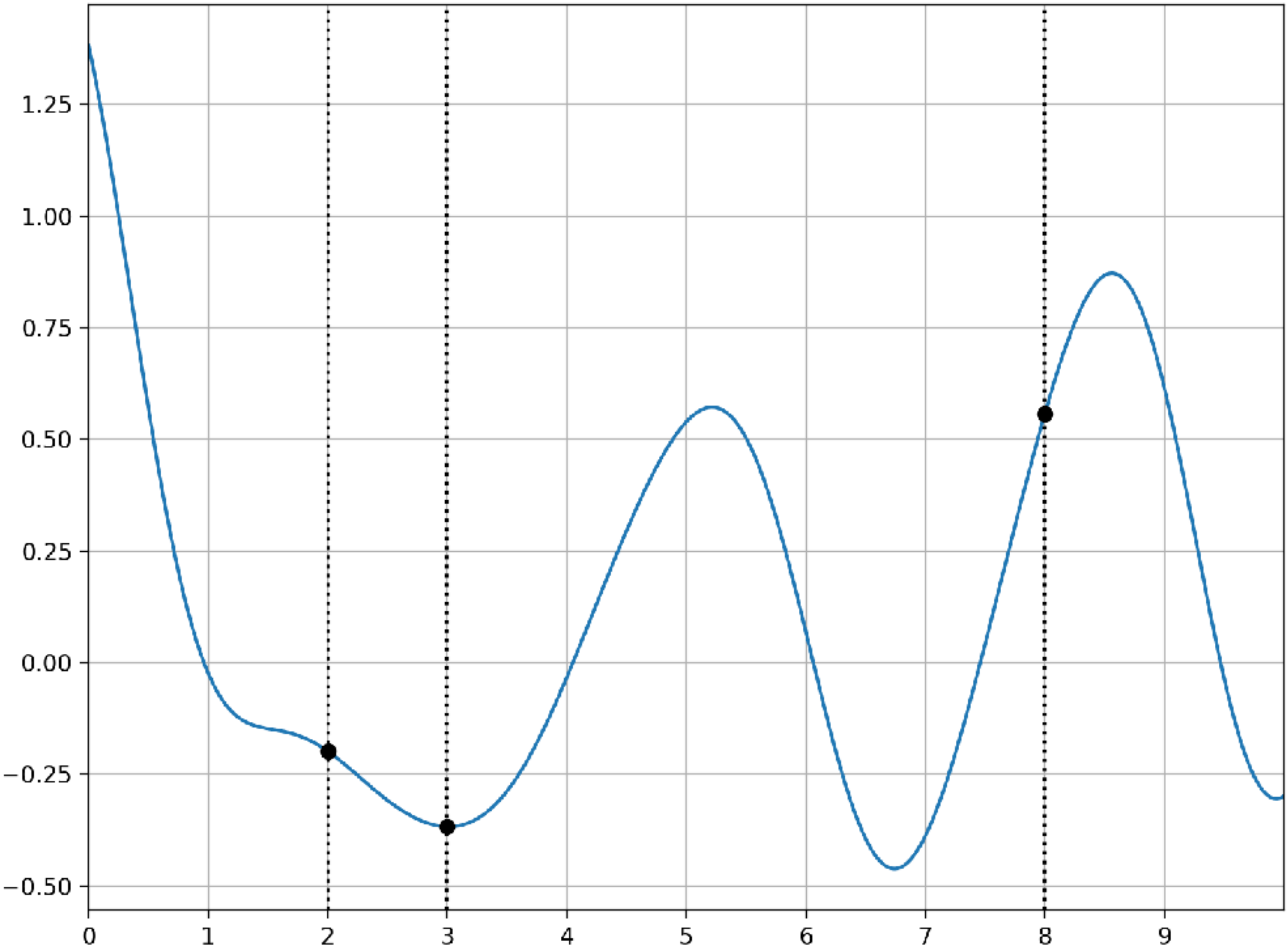
2. Derive posterior for new function evaluations:

$$p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{f})$$

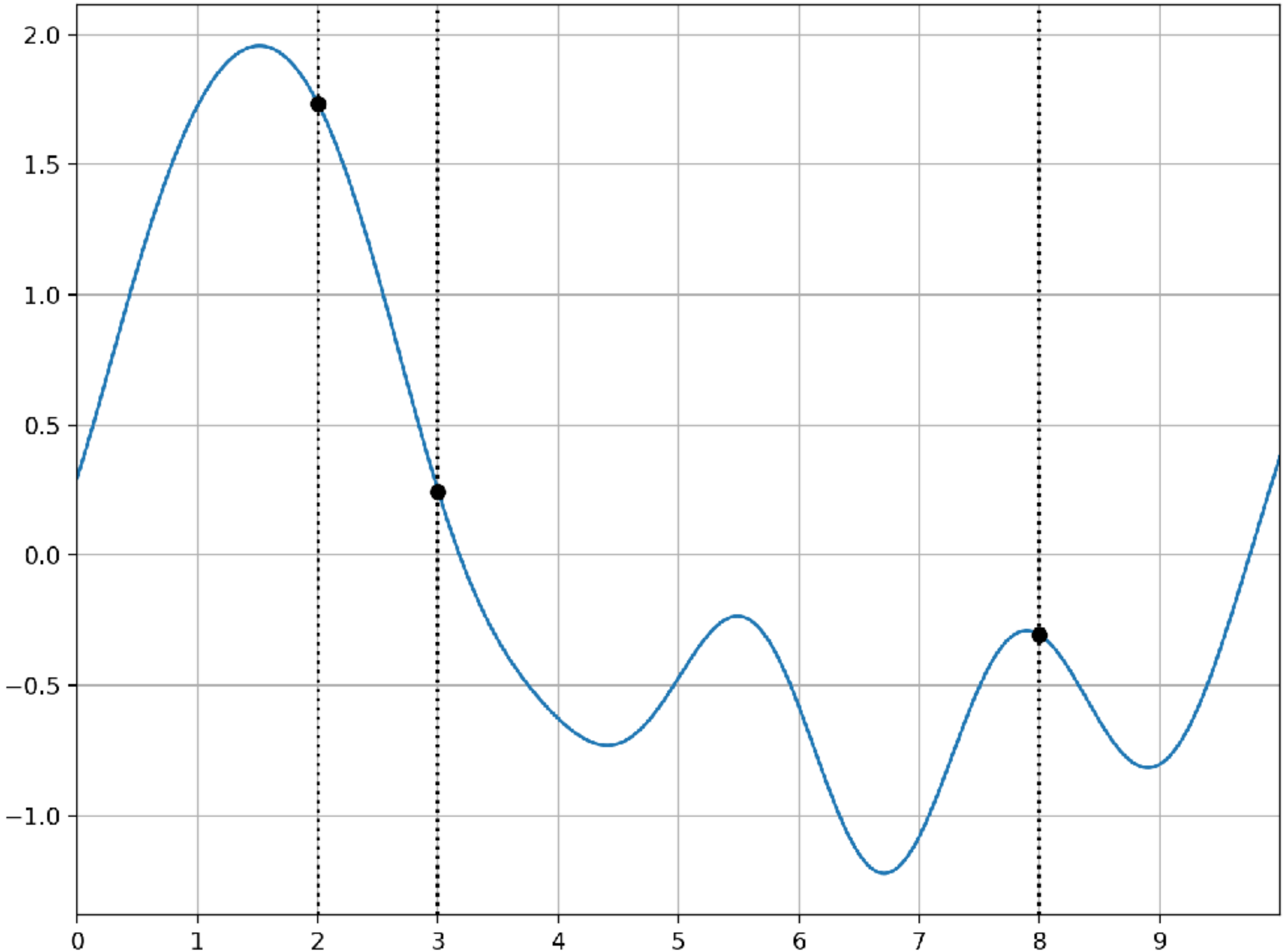
Prior over functions

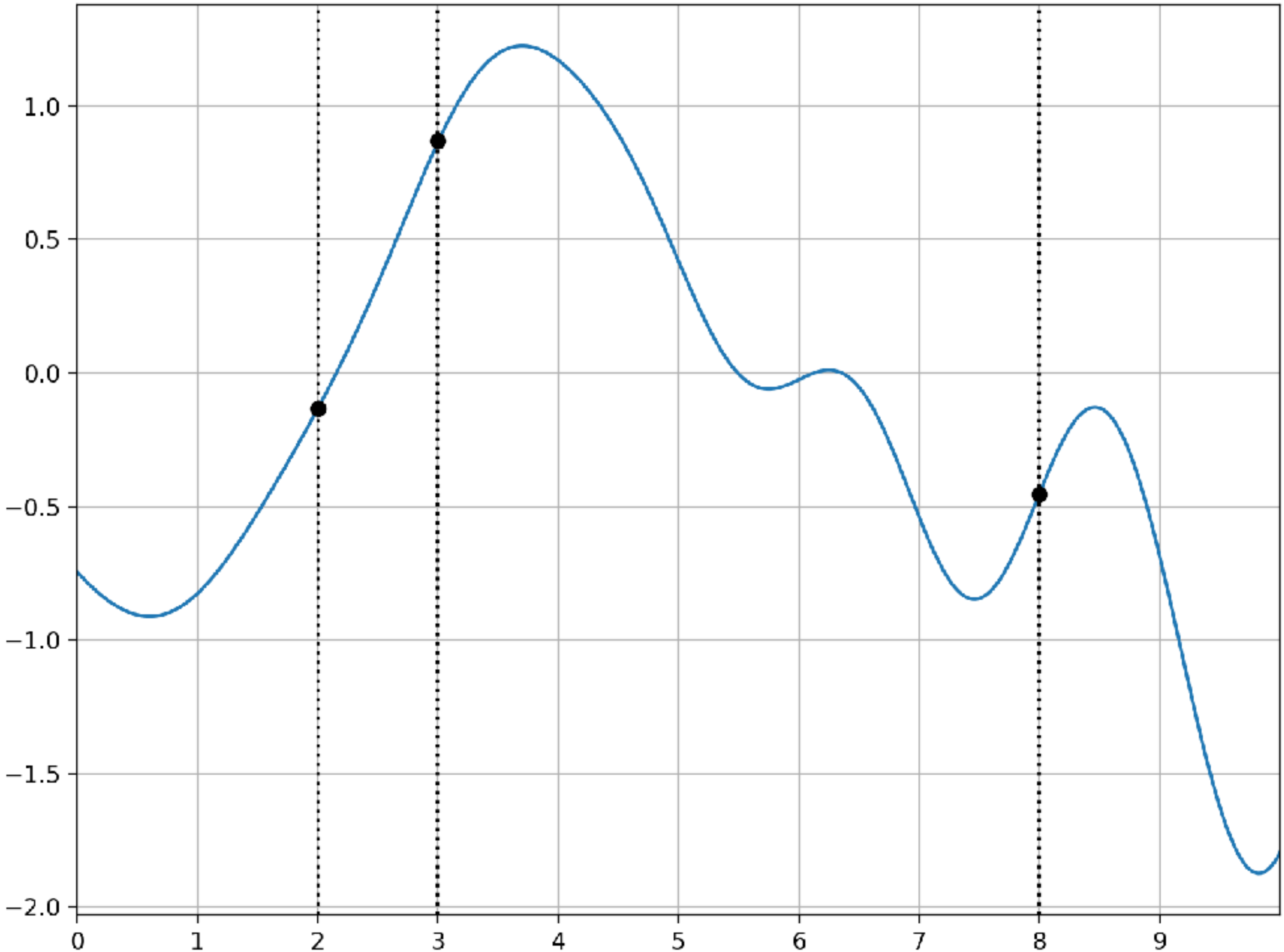
$$p(\mathbf{f} \mid \mathbf{x}) = ?$$

*Assigns probability density to a set of function evaluations independently of **y***

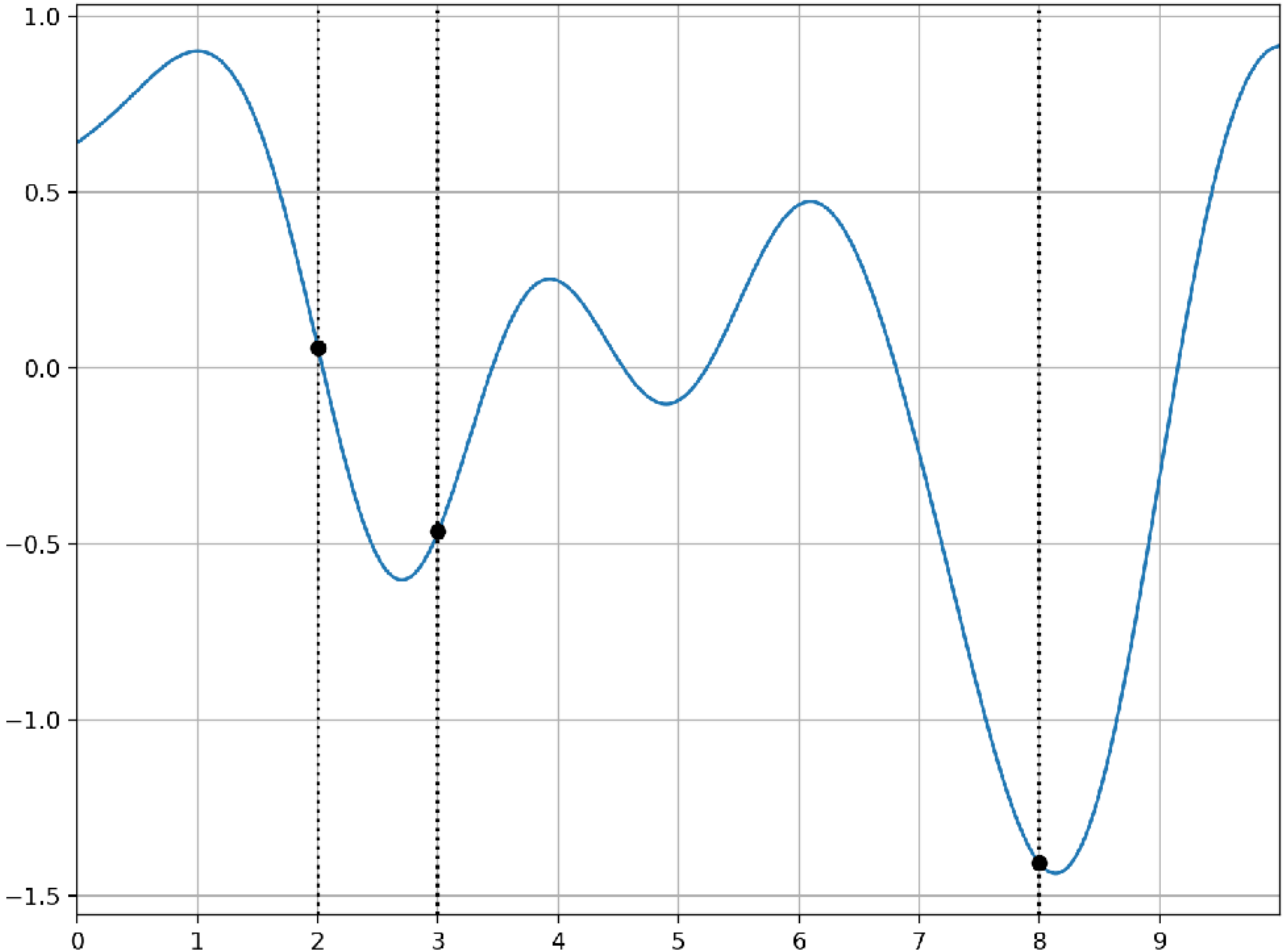


1.25
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
0
1
2
3
4
5
6
7
8
9


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
0
1
2
3
4
5
6
7
8
9



1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
0
1
2
3
4
5
6
7
8
9


1.0
0.5
0.0
−0.5
−1.0
−1.5
0
1
2
3
4
5
6
7
8
9

Gaussian process prior definition:
Given a collection of inputs, **x**, and function evaluations, ***f***, the probability density of ***f*** is given by

$$p(\mathbf{f} \mid \mathbf{x}) = \mathcal{N}(\mathbf{f} \mid \mu, \mathbf{K})$$
$$\mu_i = m(\mathbf{x}_i)$$
$$\mathbf{K}_{i,j} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$$

Gaussian process prior definition:
Given a collection of inputs, **x**, and function evaluations, ***f***, the probability density of ***f*** is given by

$$p(\mathbf{f} \mid \mathbf{x}) = \mathcal{N}(\mathbf{f} \mid \mu, \mathbf{K})$$

$$\mu_i = m(\mathbf{x}_i)$$ **Mean function** (usually $m(\mathbf{X}) = 0$)

$$\mathbf{K}_{i,j} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$$ **Kernel function**

Gaussian process prior definition:
Given a collection of inputs, **x**, and function evaluations, ***f***, the probability density of ***f*** is given by

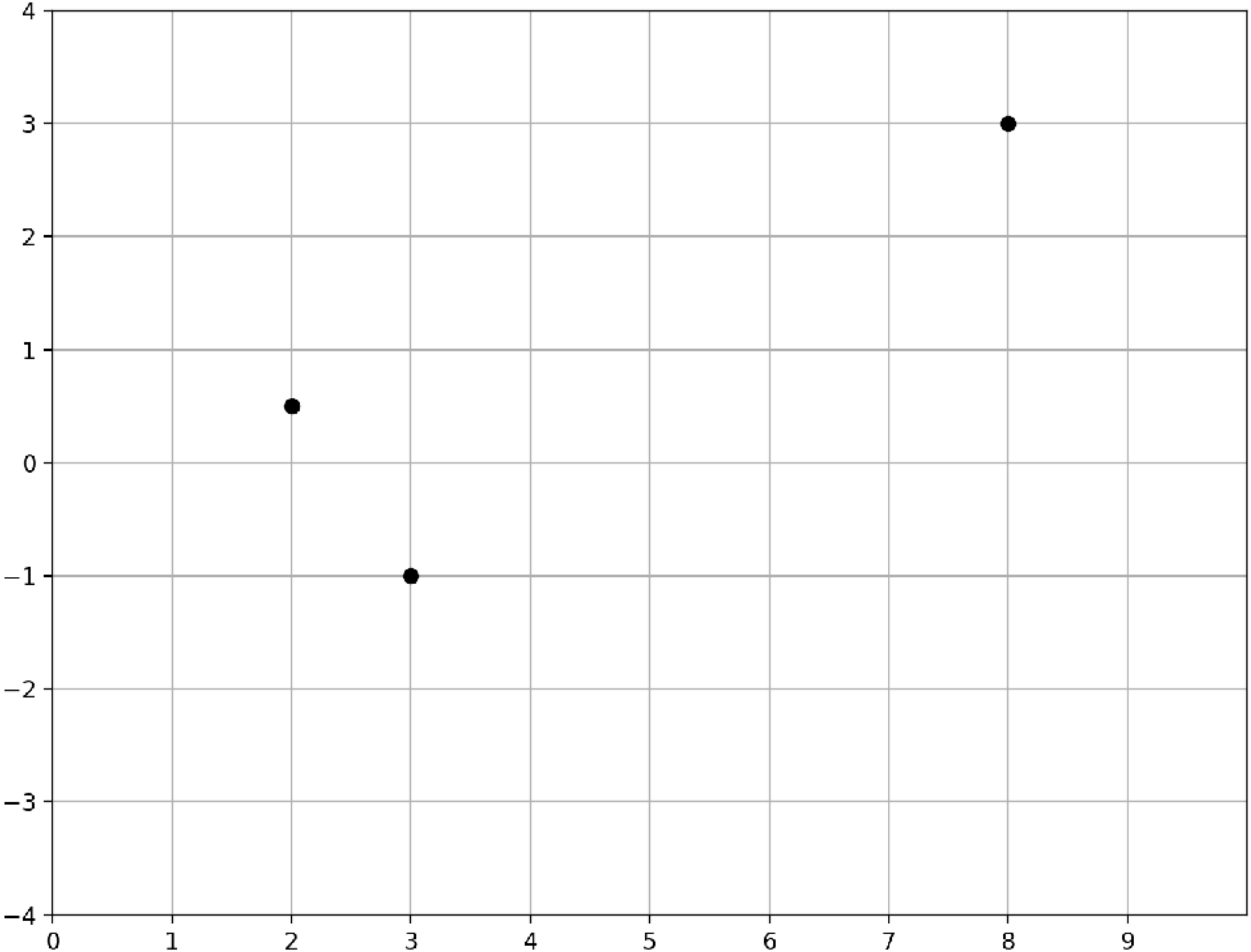
$$p(\mathbf{f} \mid \mathbf{x}) = \mathcal{N}(\mathbf{f} \mid \mu, \mathbf{K})$$

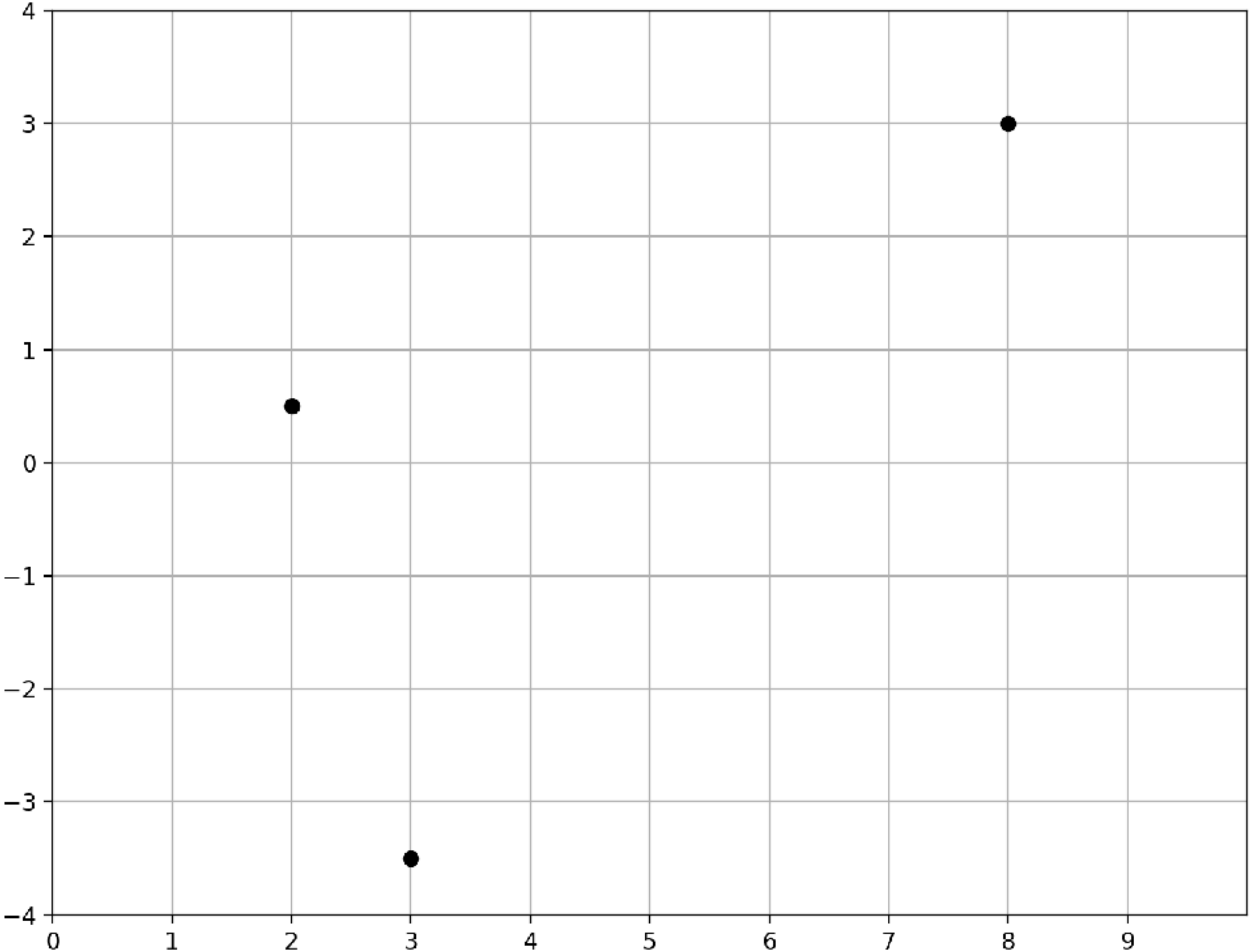
$$\mu_i = m(\mathbf{x}_i)$$ **Mean function** (usually $m(\mathbf{X}) = 0$)

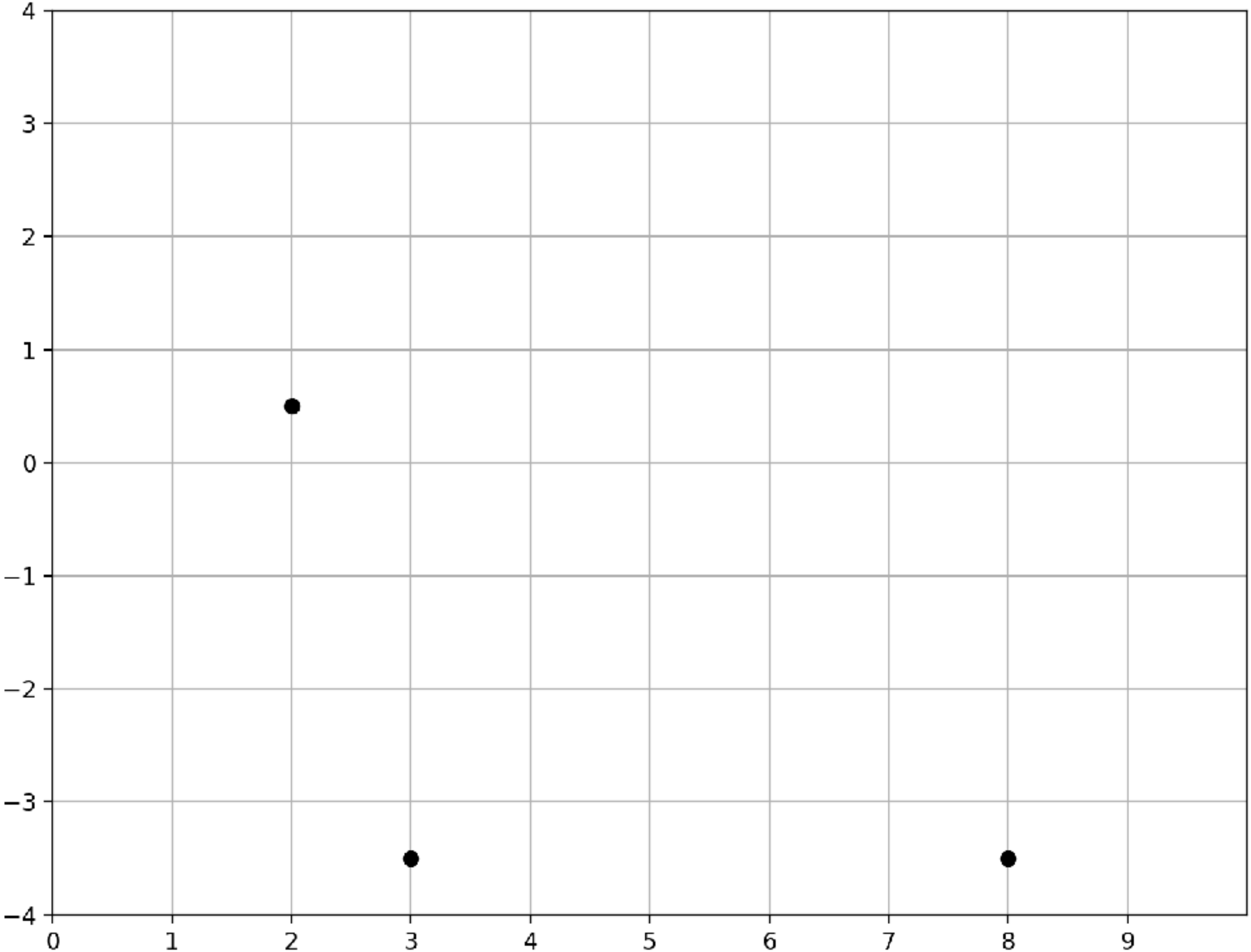
$$\mathbf{K}_{i,j} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$$ **Kernel function**

Common kernel function:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = v \cdot \exp\left(-\frac{1}{2\ell^2}\|\mathbf{x}_i - \mathbf{x}_j\|^2\right)$$


4
3
2
1
0
−1
−2
−3
−4
0
1
2
3
4
5
6
7
8
9


4
3
2
1
0
−1
−2
−3
−4
0
1
2
3
4
5
6
7
8
9


4
3
2
1
0
−1
−2
−3
−4
0
1
2
3
4
5
6
7
8
9

Marginal likelihood

$$p(\mathbf{y} \mid \mathbf{f}) = \mathcal{N}(\mathbf{y} \mid \mathbf{f}, \mathbf{I}\sigma^2)$$

$$p(\mathbf{f} \mid \mathbf{x}) = \mathcal{N}(\mathbf{0}, \mathbf{K}), \qquad \mathbf{K}_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$$

$$\begin{aligned} p(\mathbf{y} \mid \mathbf{x}) &= \int_{\mathbf{f}} p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{x}) \\ &= \int_{\mathbf{f}} \mathcal{N}(\mathbf{y} \mid \mathbf{f}, \mathbf{I}\sigma^2) \mathcal{N}(\mathbf{0}, \mathbf{K}) \\ &= \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{K} + \mathbf{I}\sigma^2) \end{aligned}$$

$$e.g. \qquad \mathbf{K}_{i,j} = v \exp\left(\frac{1}{2\ell^2} \|x_i - x_j\|^2\right)$$

Marginal likelihood

$$p(\mathbf{y} \mid \mathbf{f}) = \mathcal{N}(\mathbf{y} \mid \mathbf{f}, \mathbf{I}\sigma^2)$$

$$p(\mathbf{f} \mid \mathbf{x}) = \mathcal{N}(\mathbf{0}, \mathbf{K}), \qquad \mathbf{K}_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$$

$$p(\mathbf{y} \mid \mathbf{x}) = \int_{\mathbf{f}} p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{x})$$

$$= \int_{\mathbf{f}} \mathcal{N}(\mathbf{y} \mid \mathbf{f}, \mathbf{I}\sigma^2) \mathcal{N}(\mathbf{0}, \mathbf{K})$$

$$= \mathcal{N}(\mathbf{y} \mid \mathbf{0}, \mathbf{K} + \mathbf{I}\sigma^2)$$

hyper-parameters

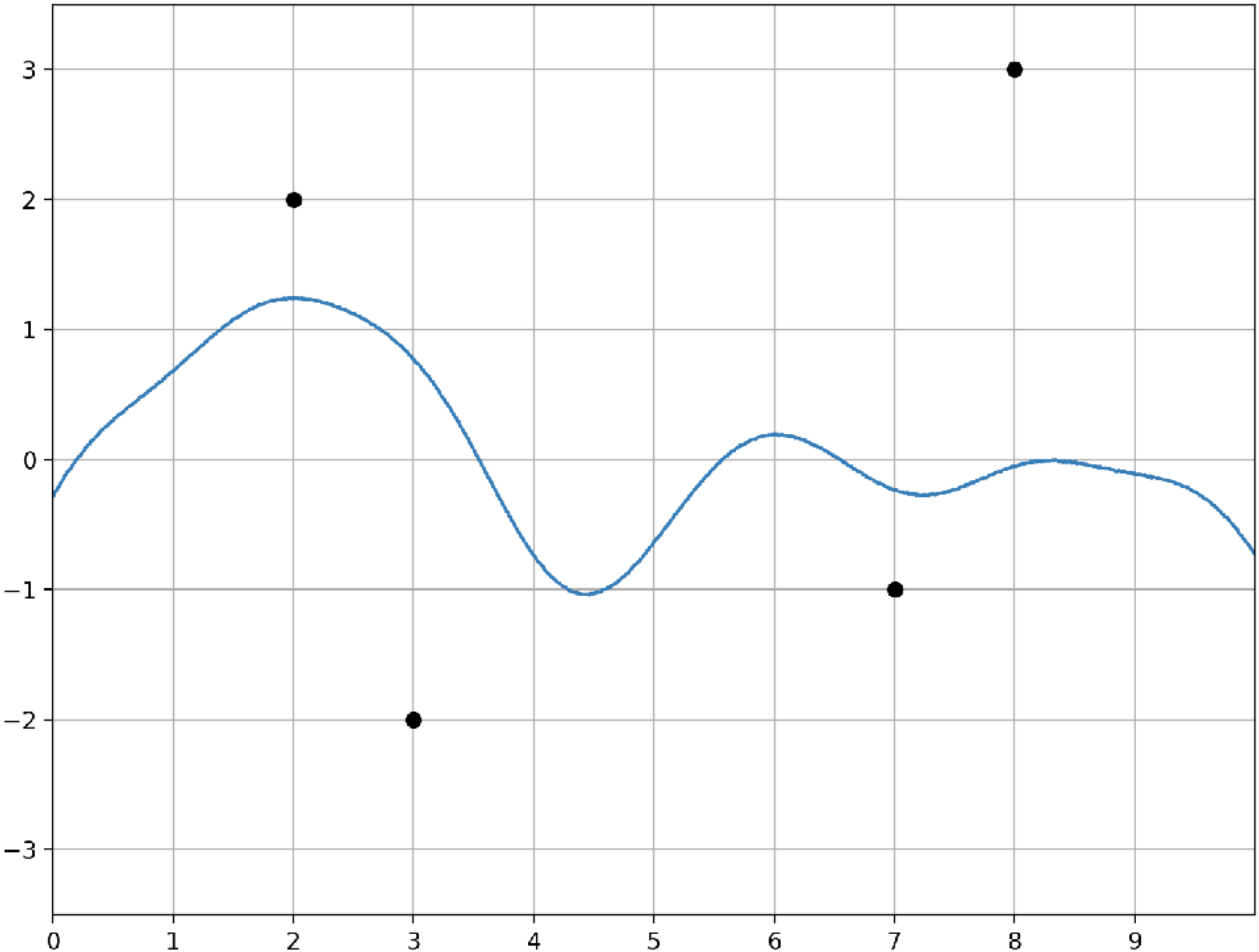
$$e.g. \qquad \mathbf{K}_{i,j} = v \exp\left(\frac{1}{2\ell^2} \|x_i - x_j\|^2\right)$$

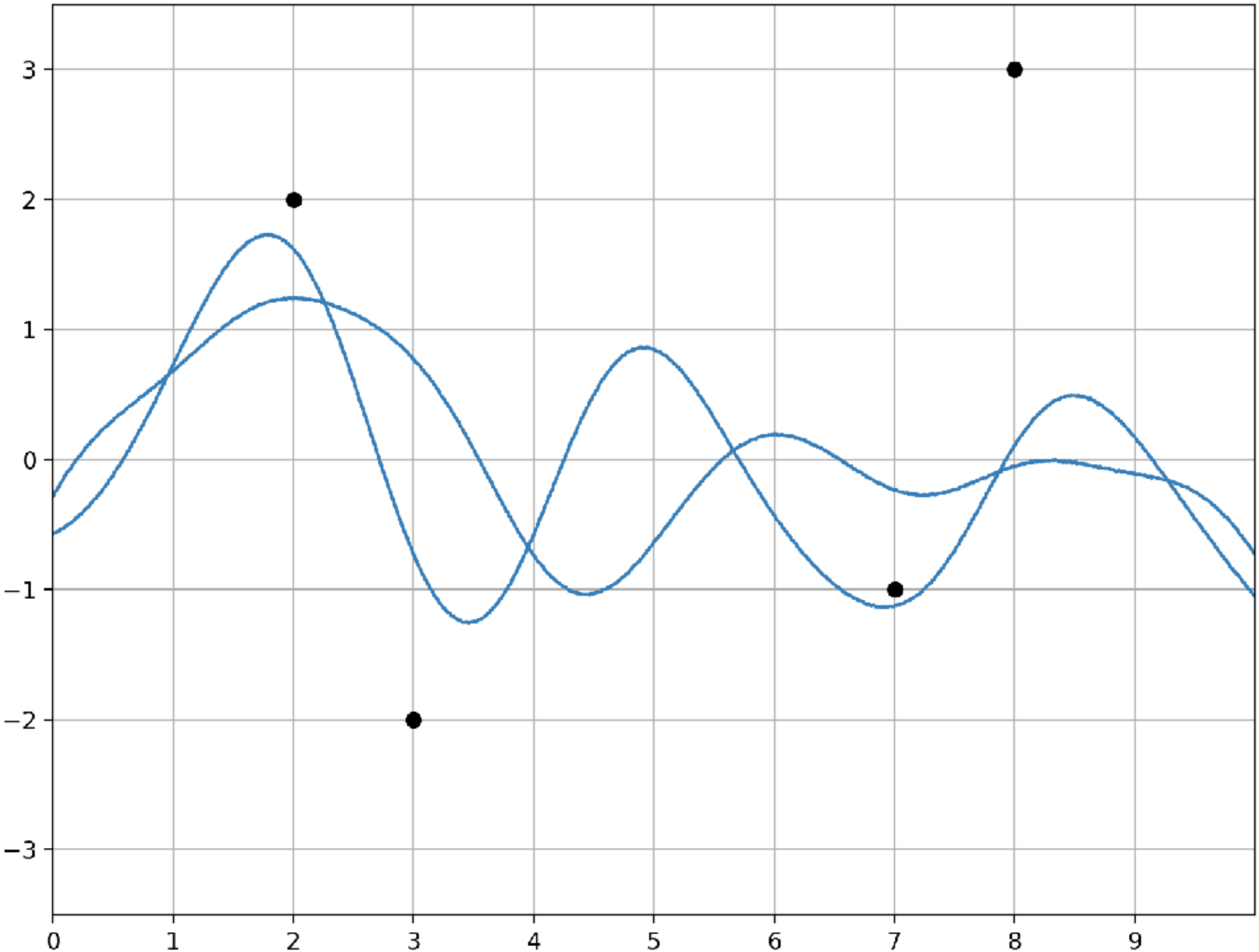
Function posterior

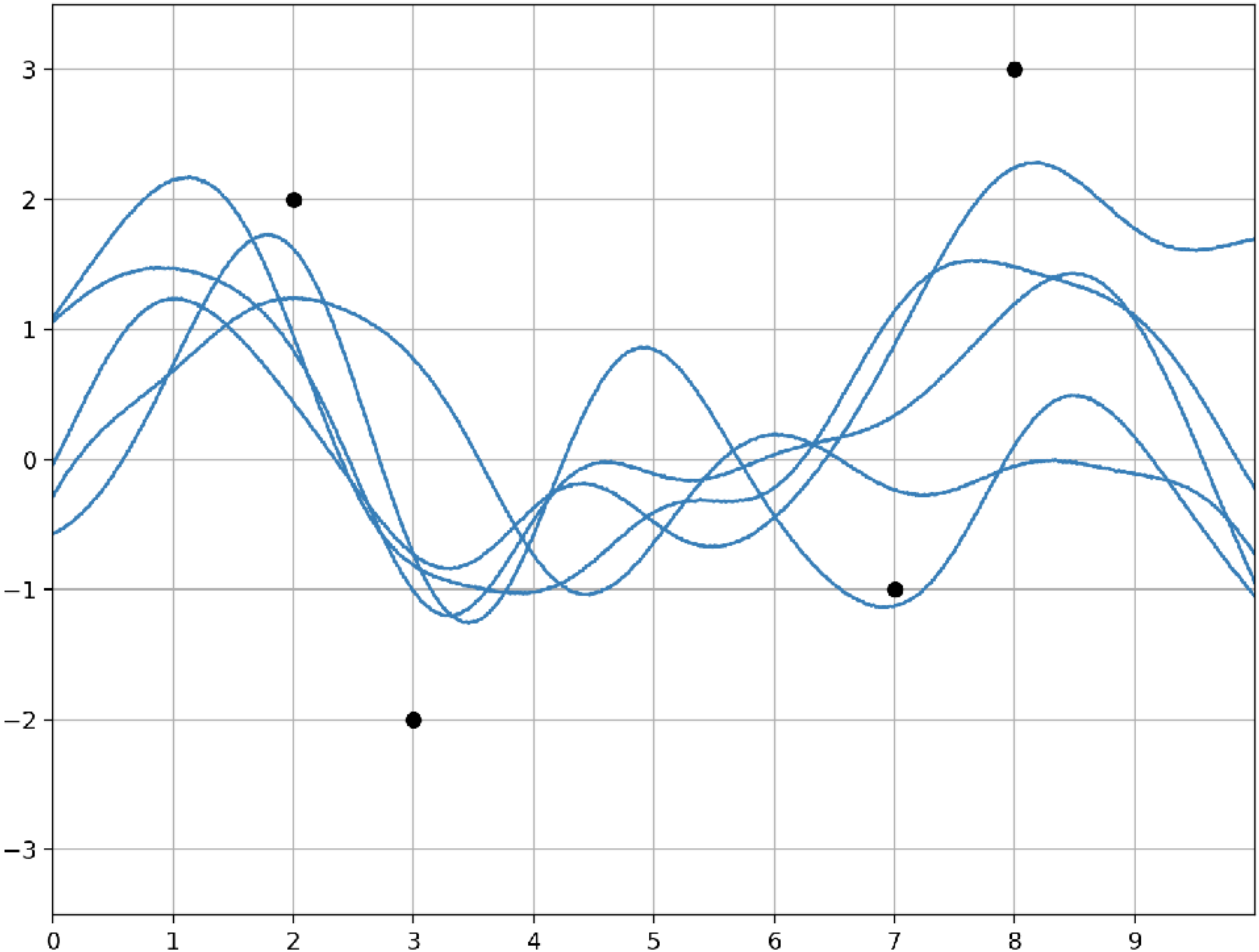
$$p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y}) = \,?$$

Idea: MCMC Sampling

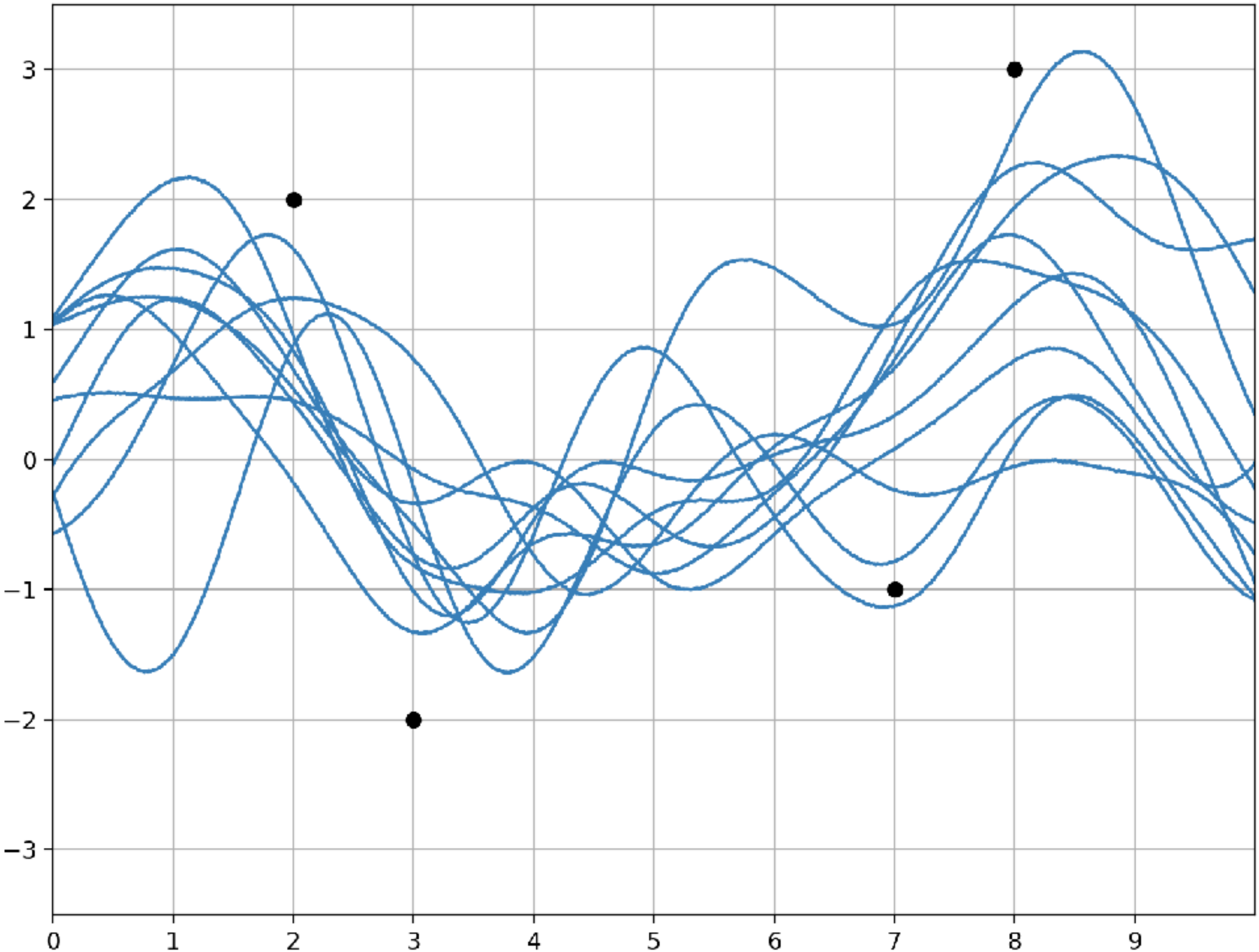
$$\begin{aligned} p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y}) &\propto \int_{\mathbf{f}} p(\mathbf{y}, \mathbf{f}_*, \mathbf{f} \mid \mathbf{x}_*, \mathbf{x}) \\ &\approx \frac{1}{K} \sum p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f}_*, \mathbf{f} \mid \mathbf{x}_*, \mathbf{x}) \end{aligned}$$

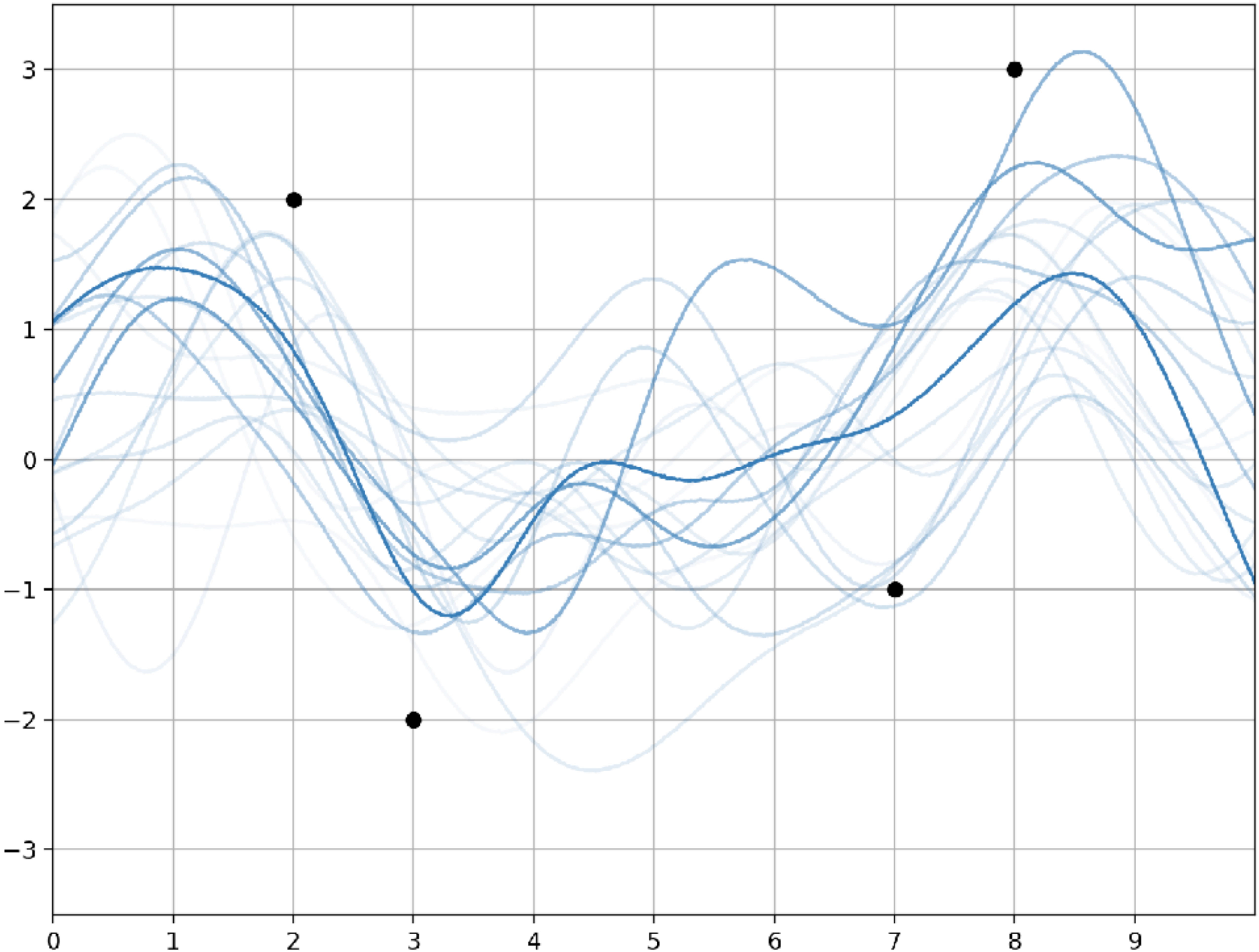

3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9

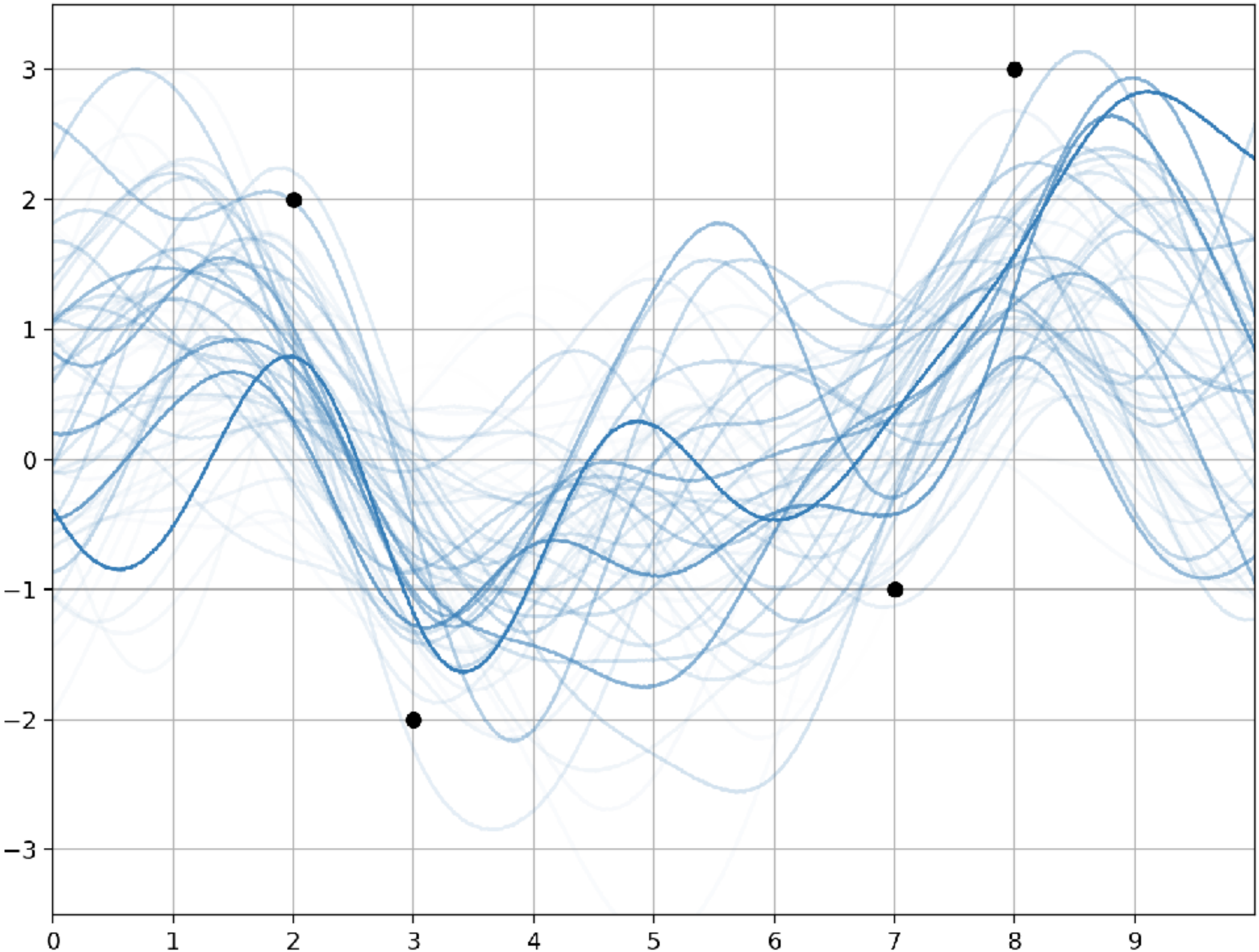

3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9

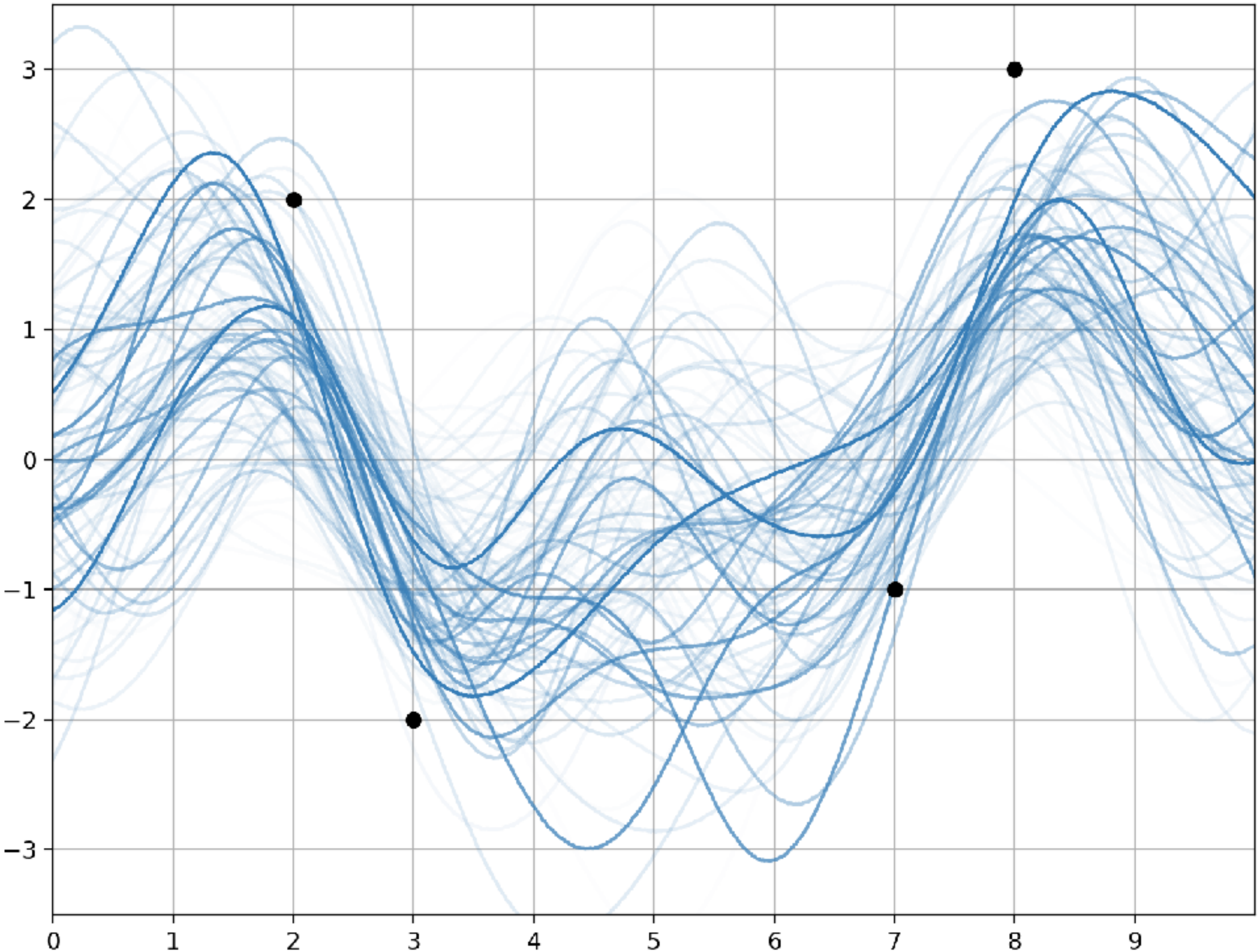


3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9

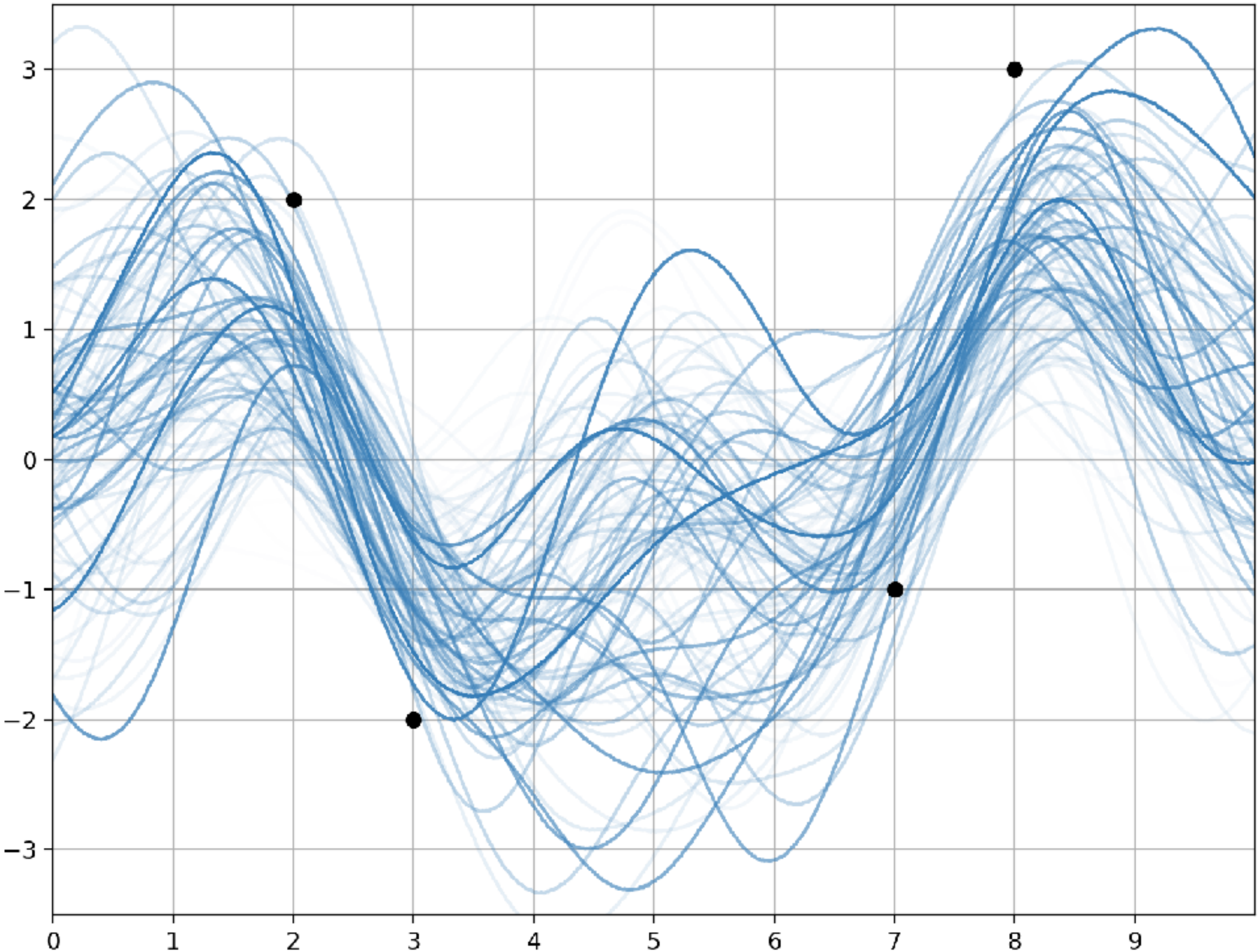

3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9


3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9

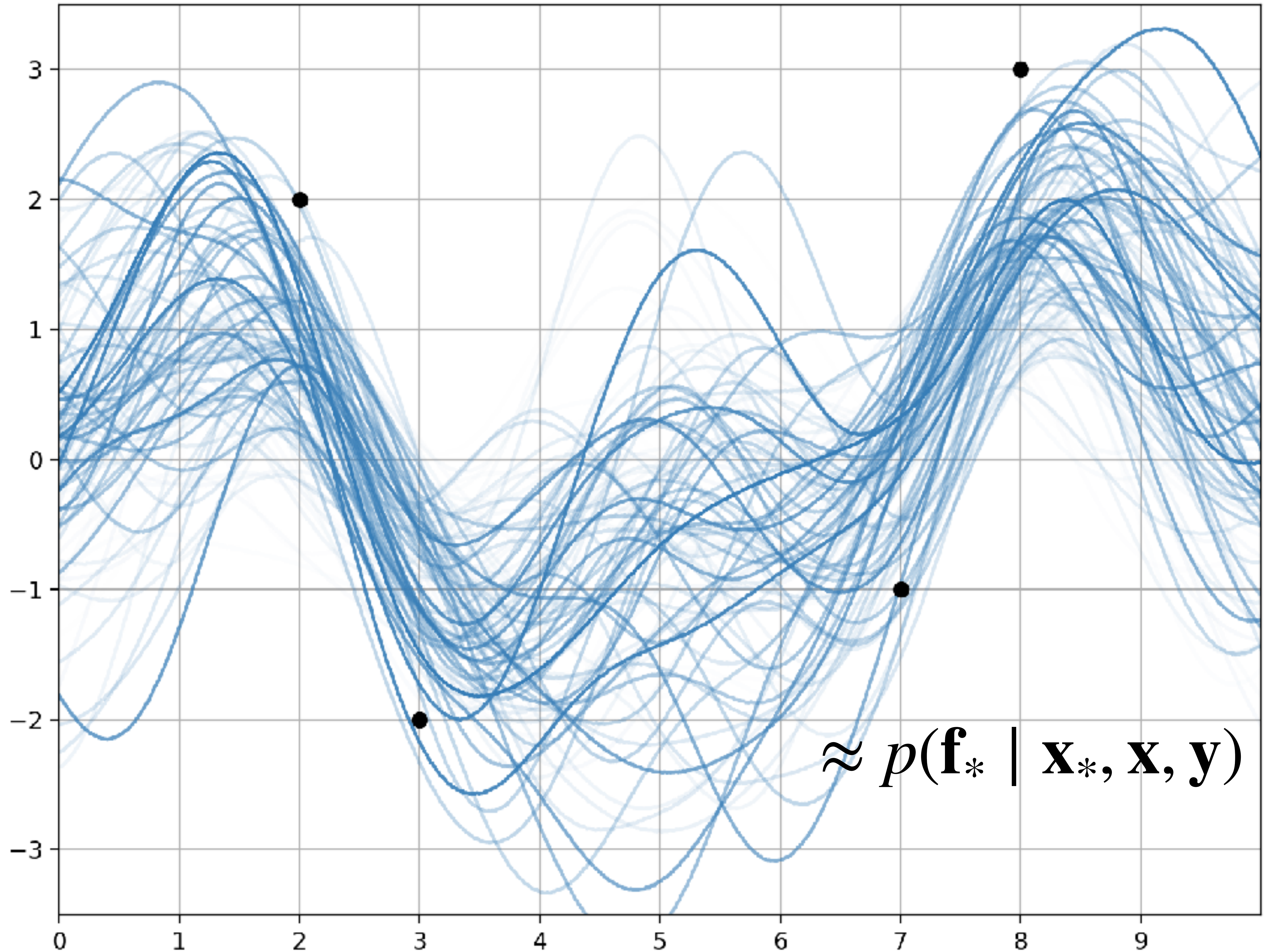

3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9



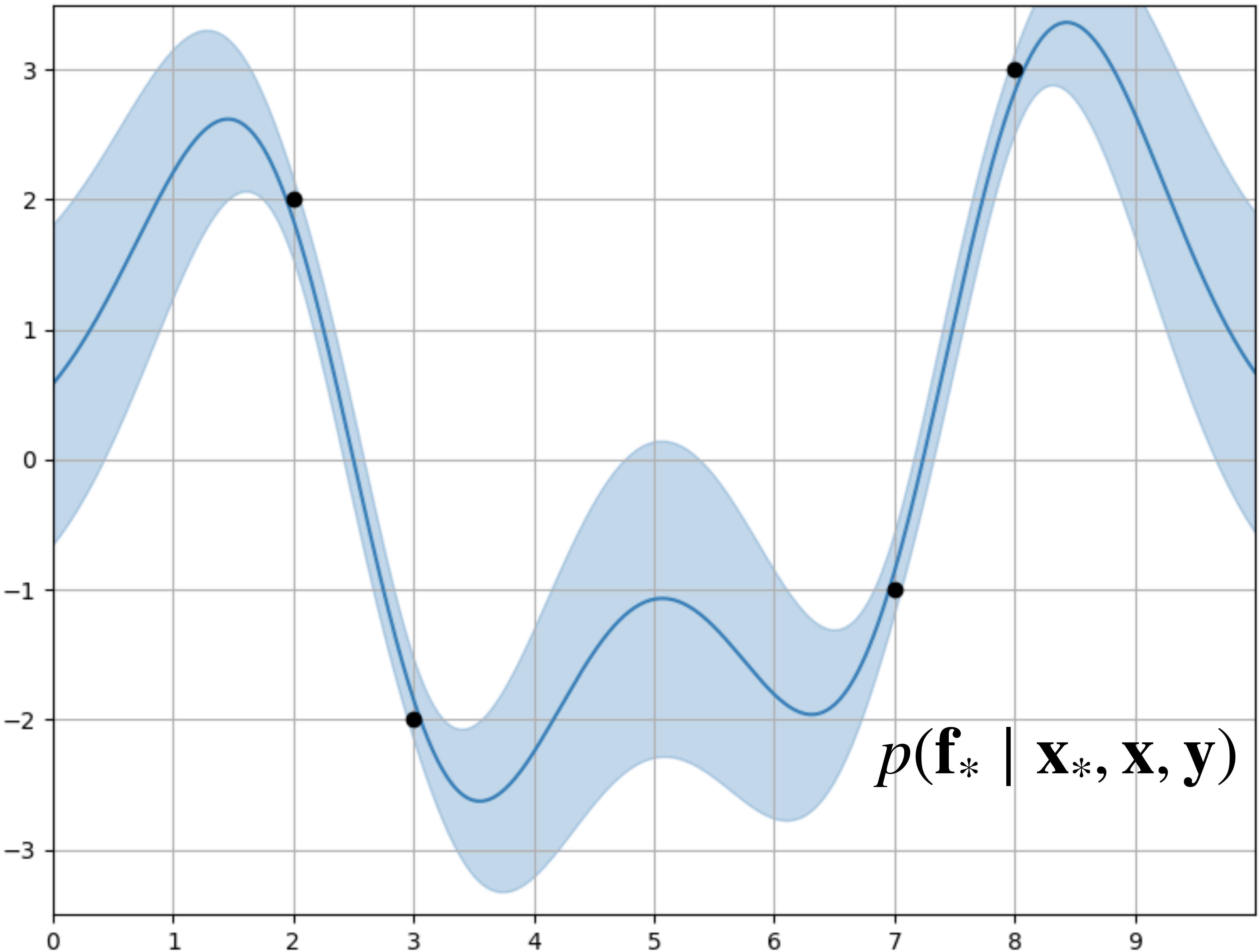
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9



3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9



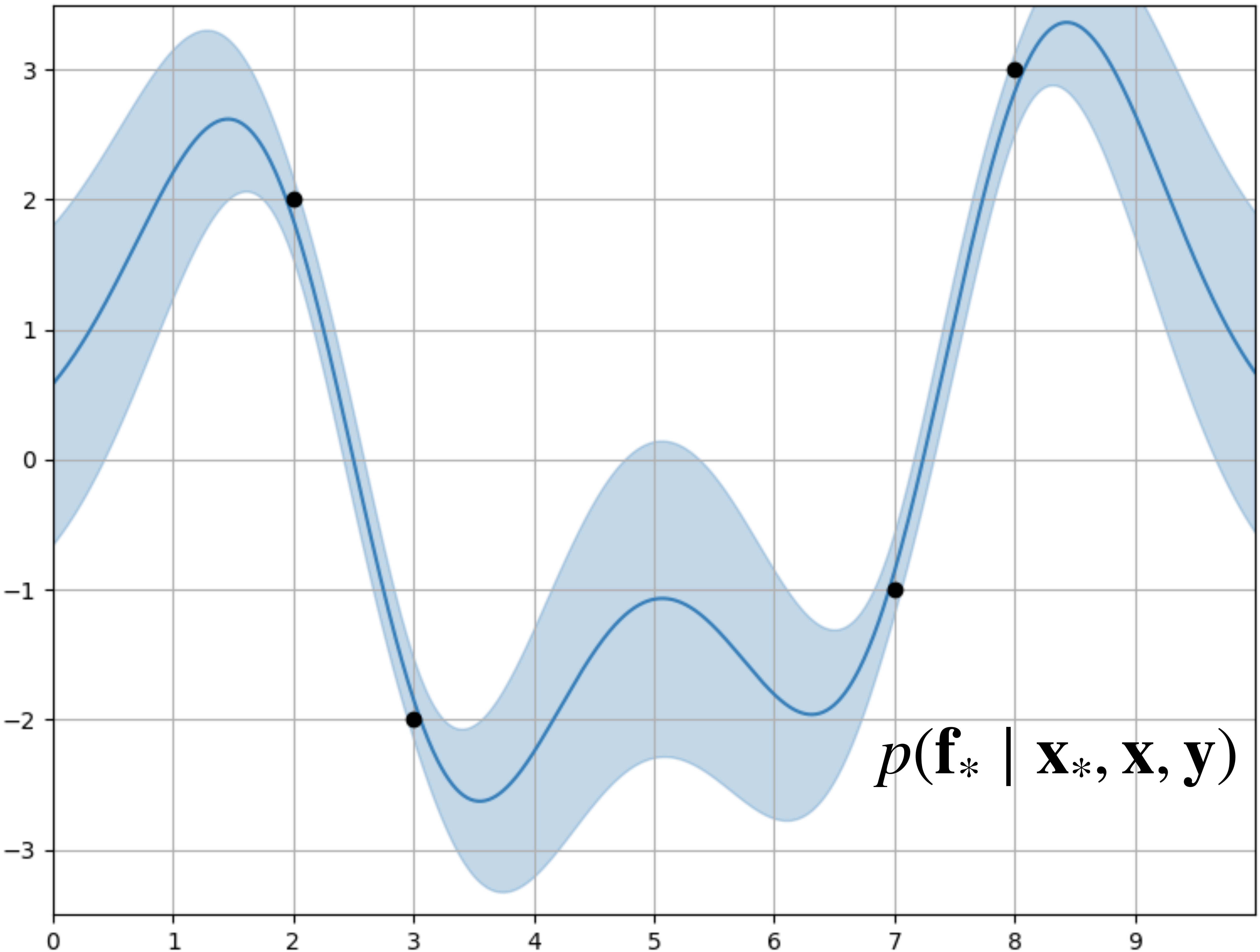
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
$\approx p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y})$


$p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y})$
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9

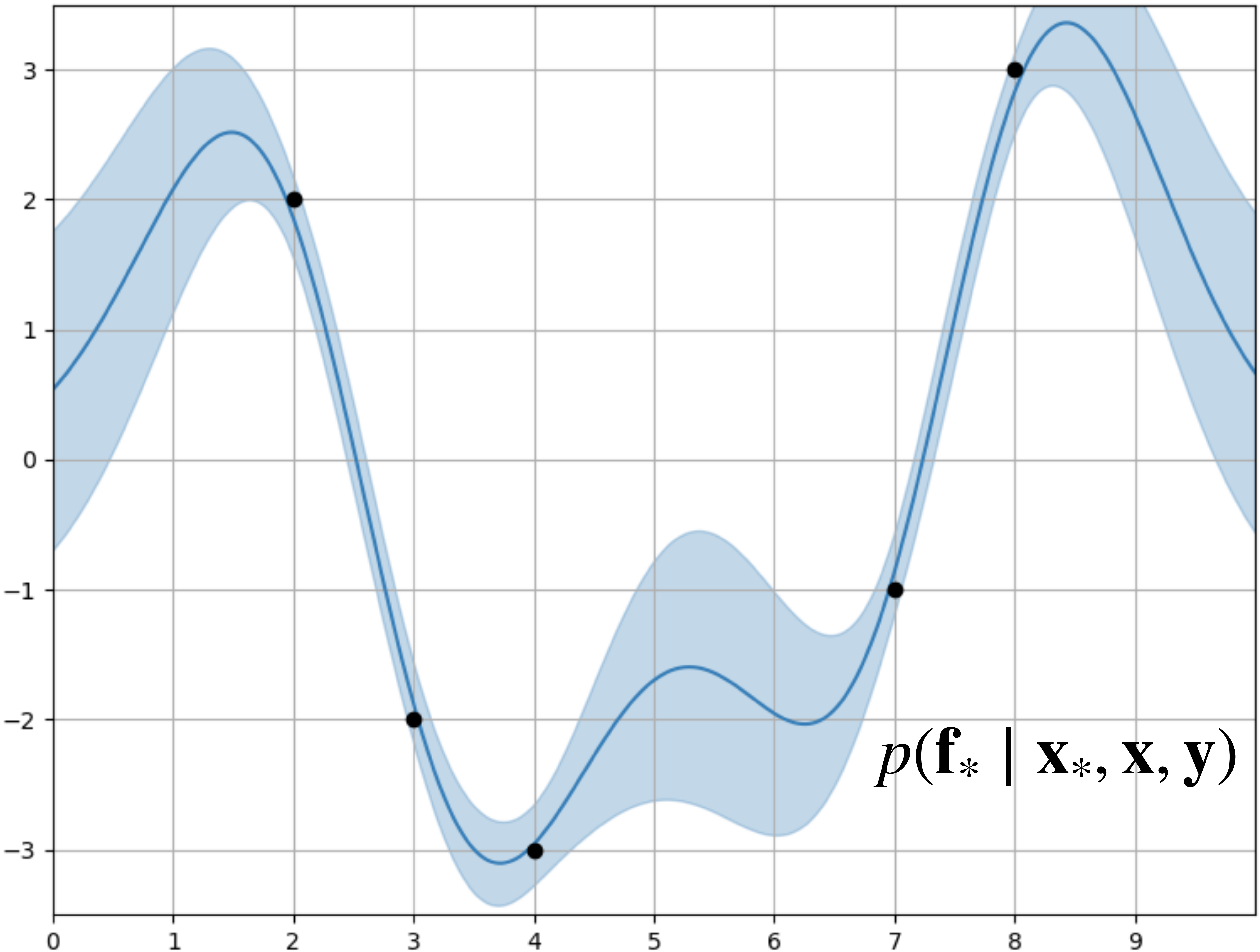
$$p(\mathbf{f}, \mathbf{f}_* \mid \mathbf{x}, \mathbf{x}_*) = \mathcal{N}\left(\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{K}_* \\ \mathbf{K}_*^\top & \mathbf{K}_{**} \end{bmatrix}, \right)$$

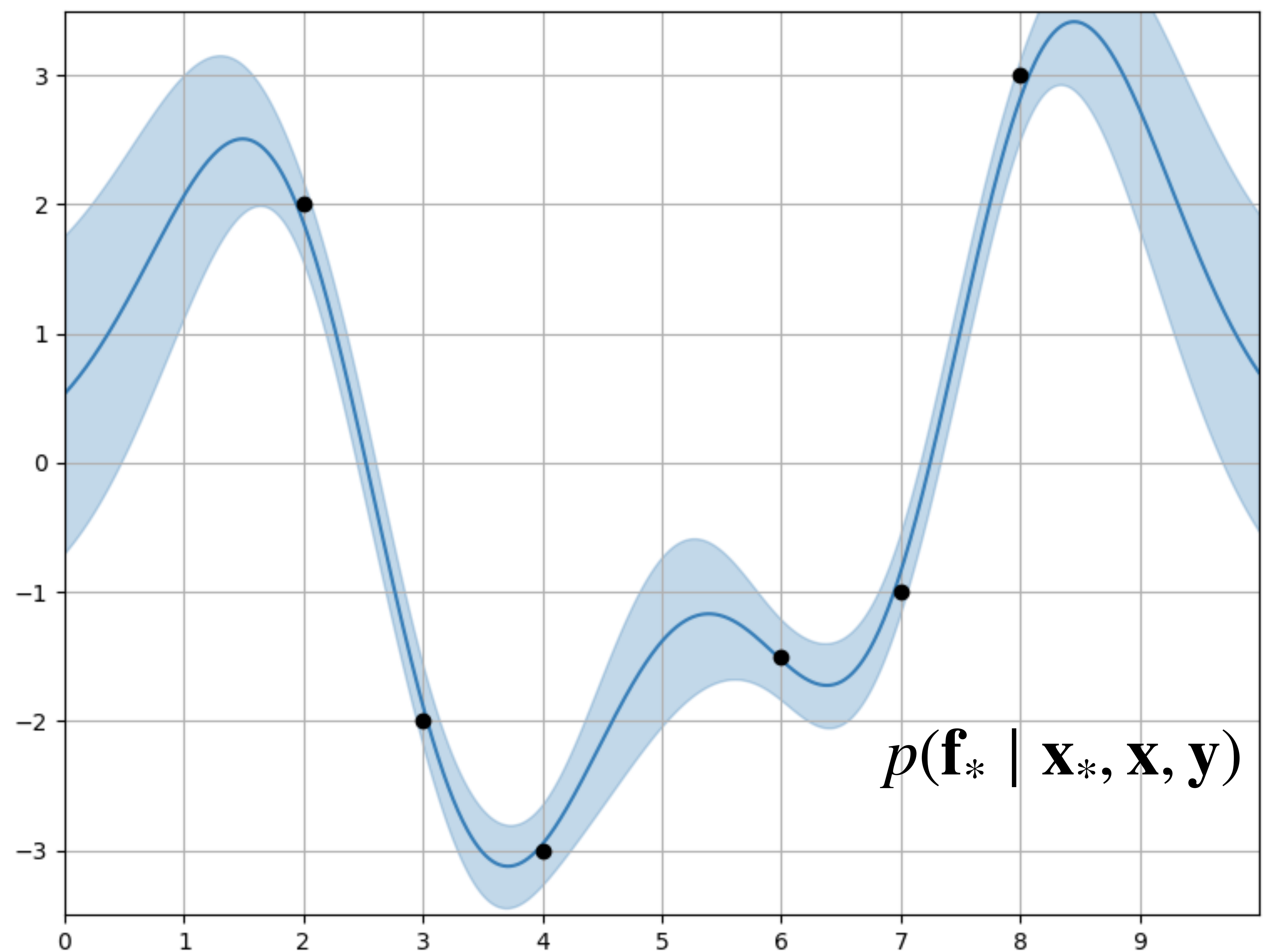
$$\Rightarrow$$

$$p(\mathbf{f}_* \mid \mathbf{x}, \mathbf{x}_*, \mathbf{f}) = \mathcal{N}(\mathbf{f}_* \mid \mathbf{K}_*^\top \mathbf{K}^{-1}\mathbf{f}, \mathbf{K}_{**} - \mathbf{K}_*^\top \mathbf{K}^{-1}\mathbf{K}_*)$$

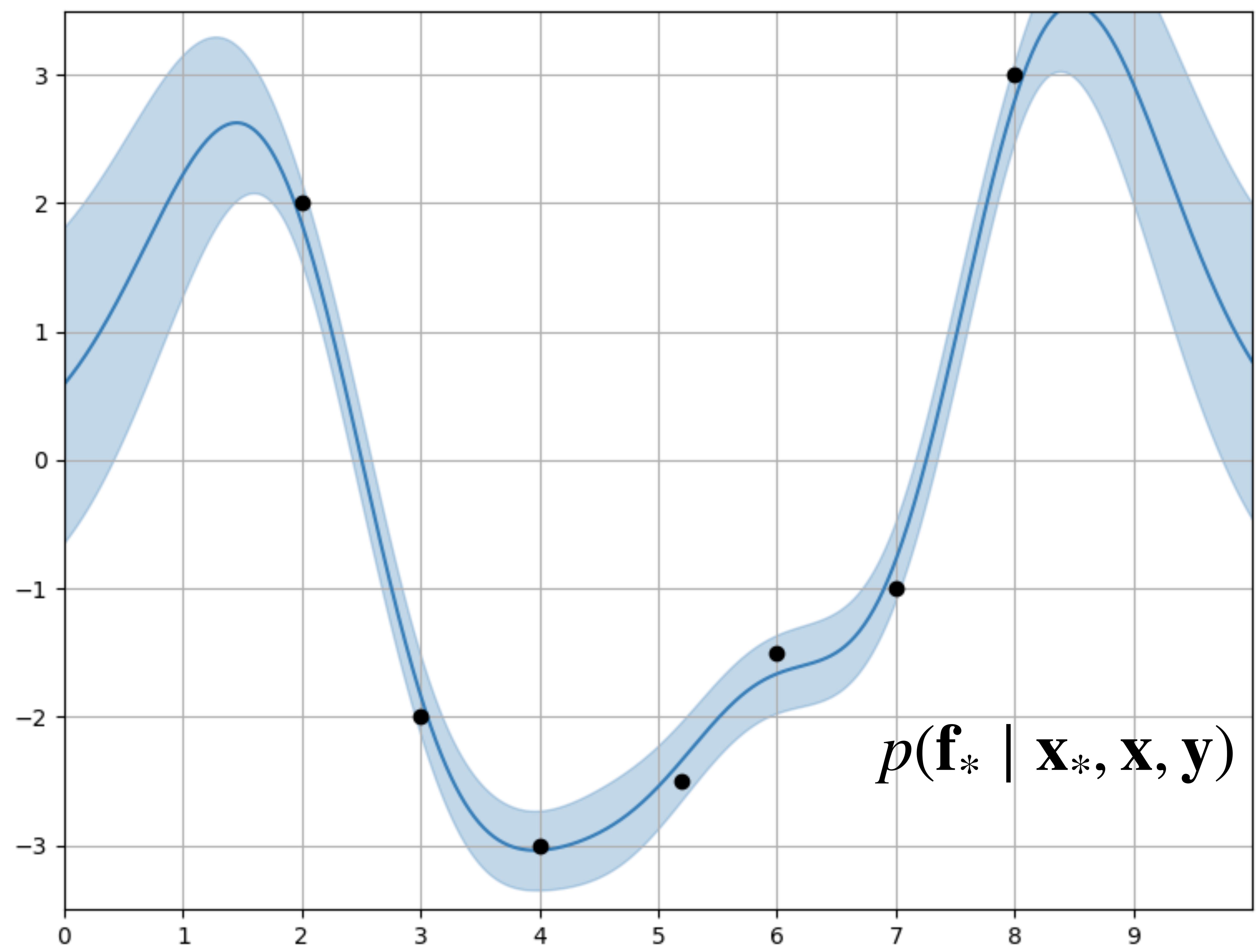


3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
$p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y})$


3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
$p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y})$



3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
$p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y})$

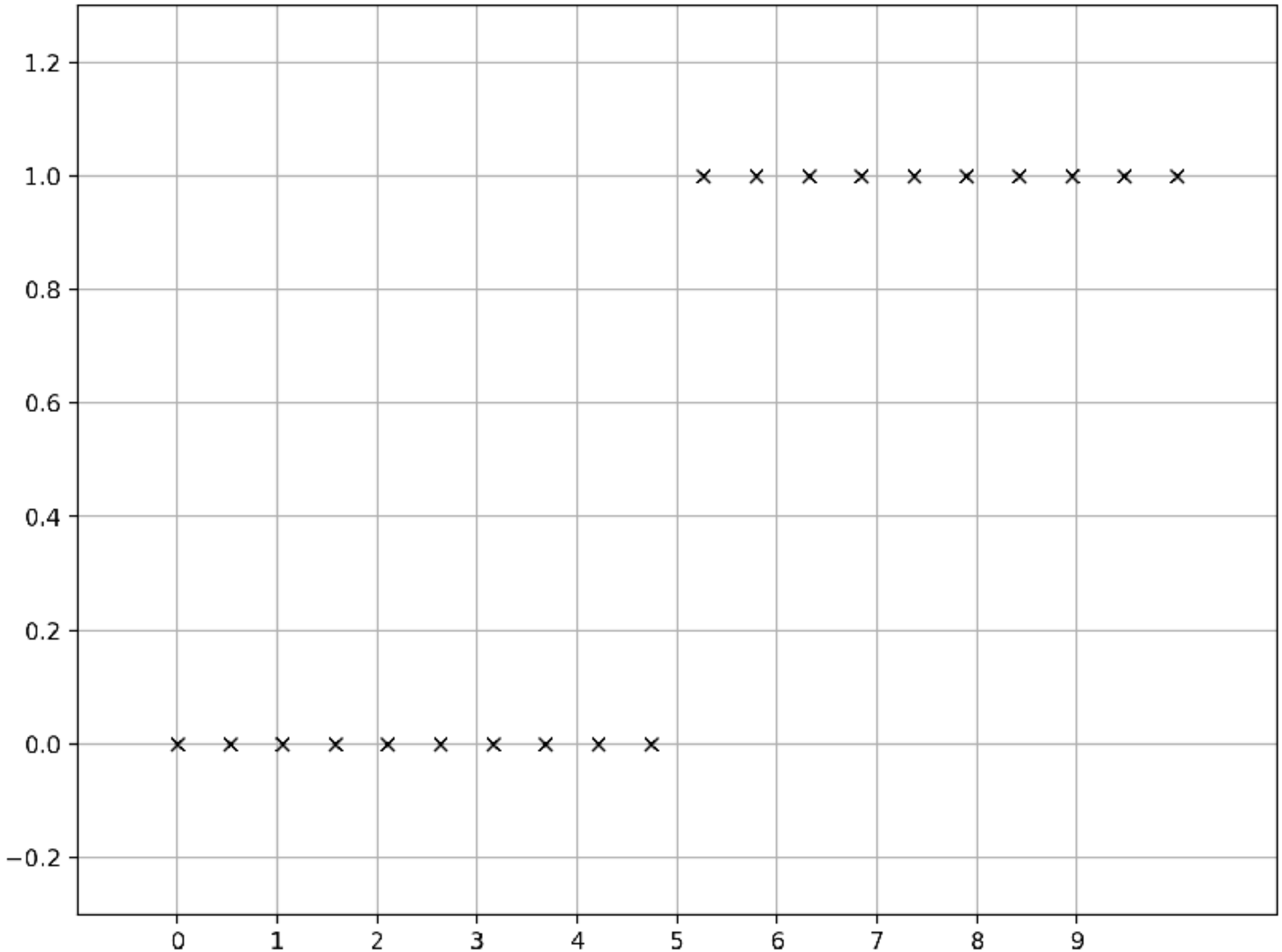


3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
7
8
9
$p(\mathbf{f}_* \mid \mathbf{x}_*, \mathbf{x}, \mathbf{y})$

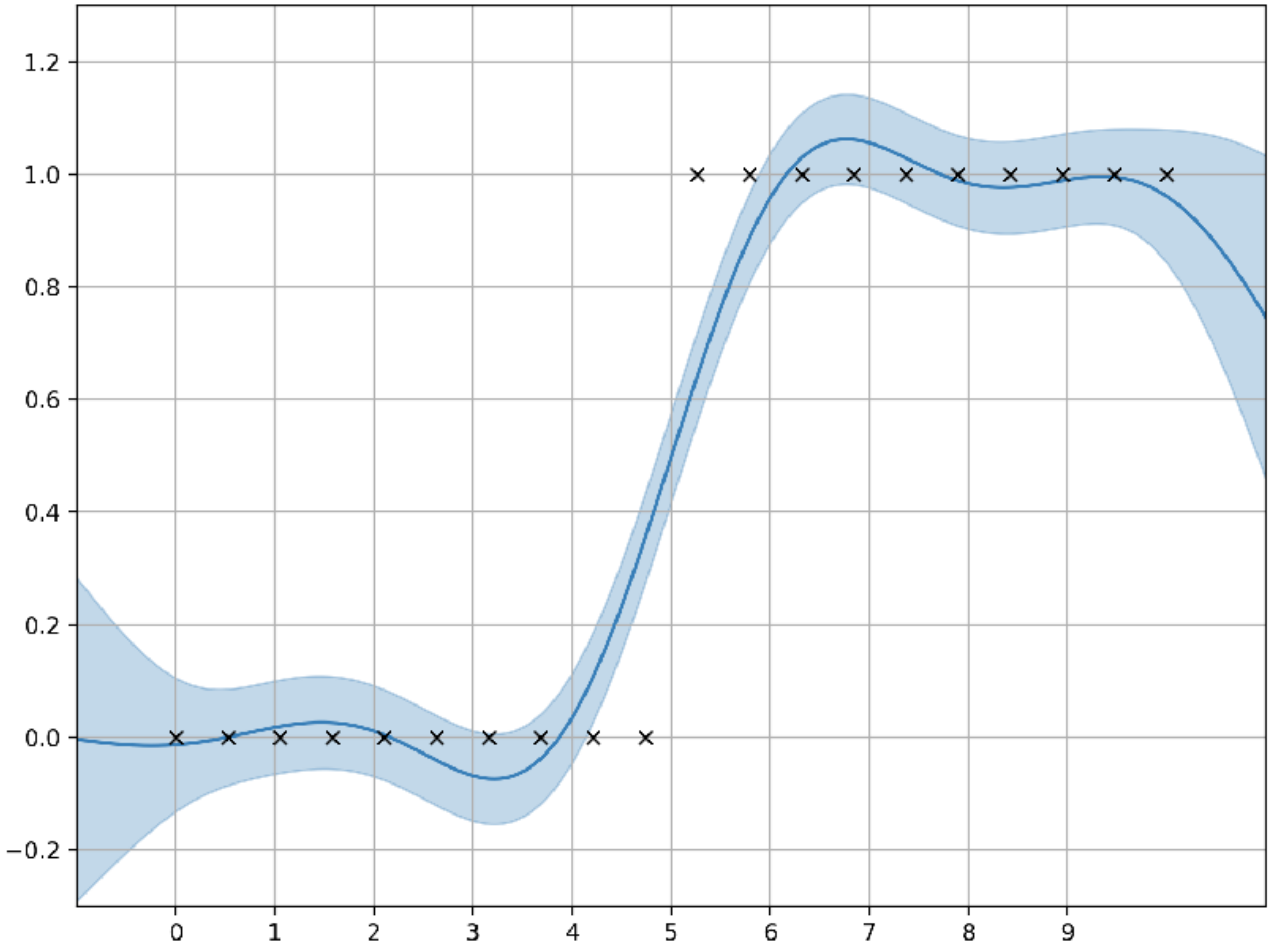
Gaussian process limitation?

Requires everything to be Gaussian

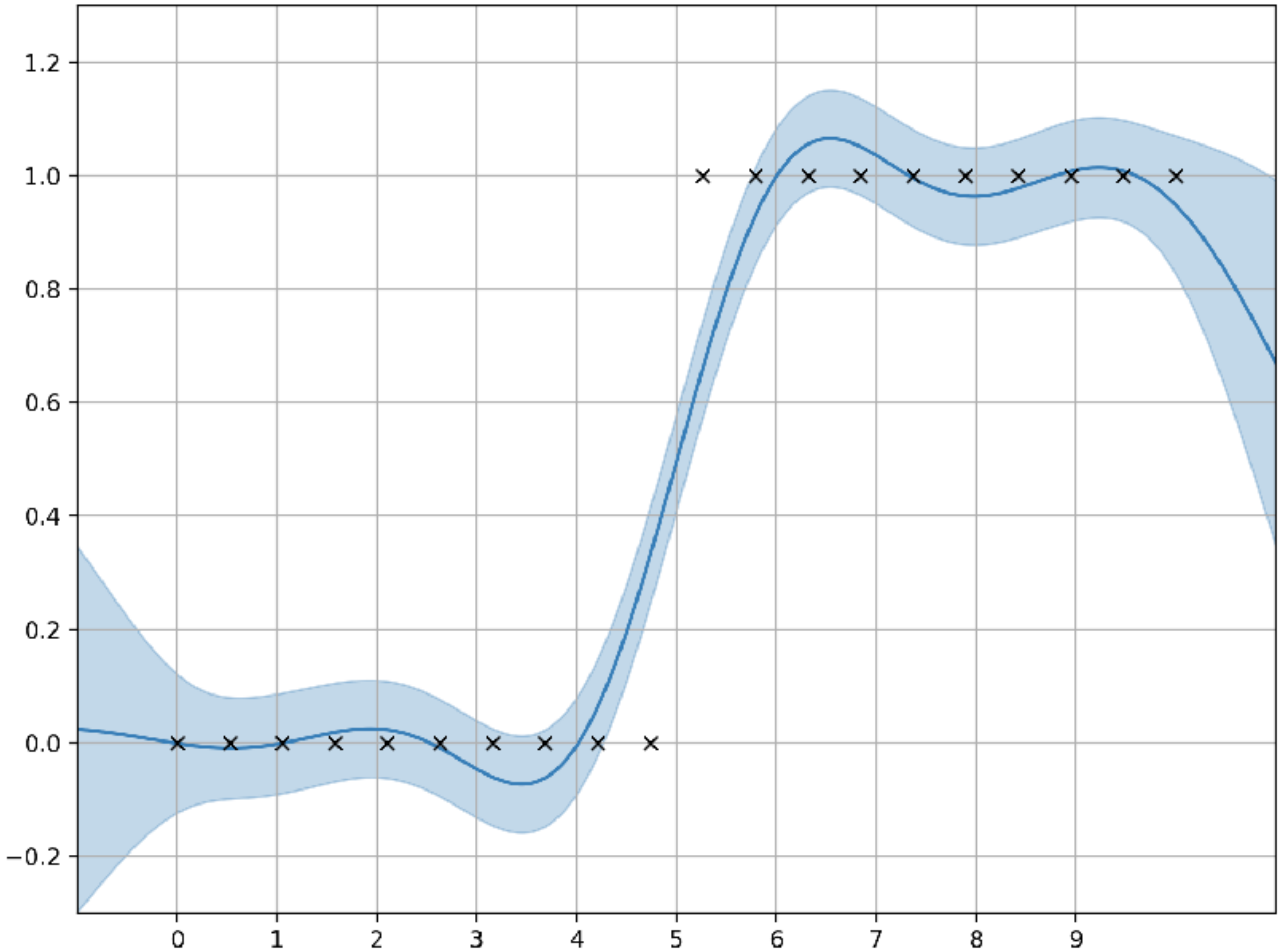
Gaussian process limitations



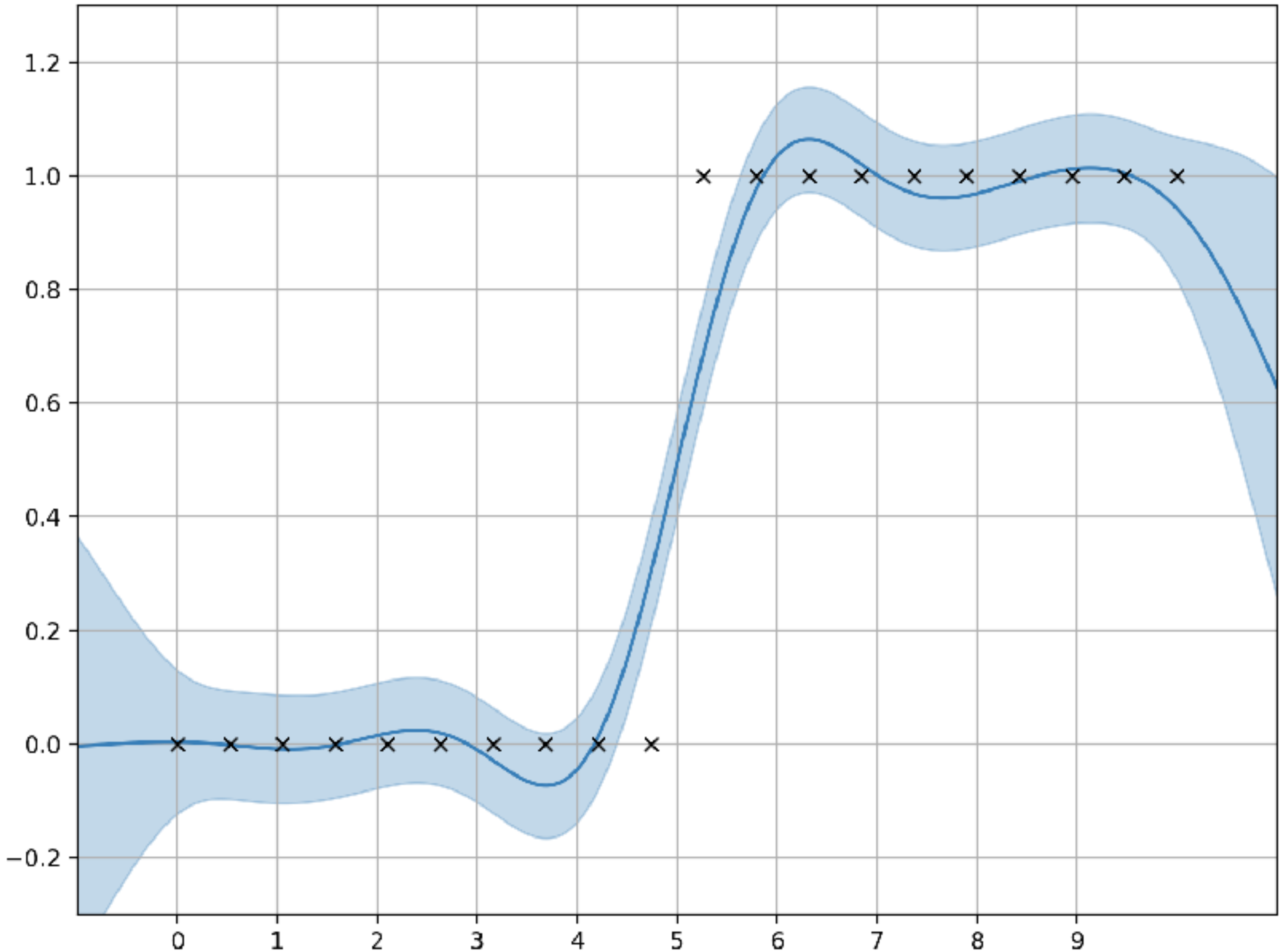
Gaussian process limitations



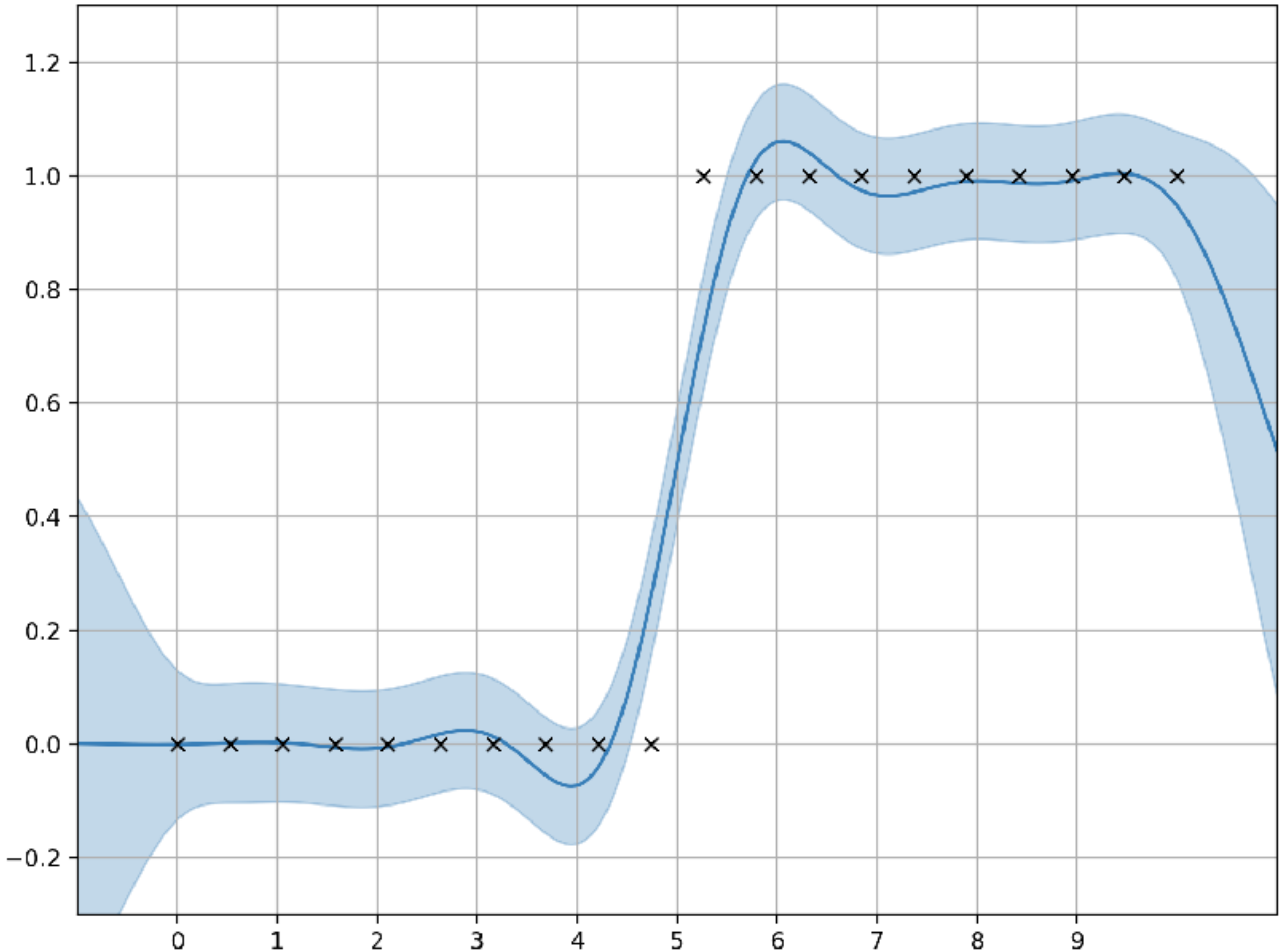
Gaussian process limitations



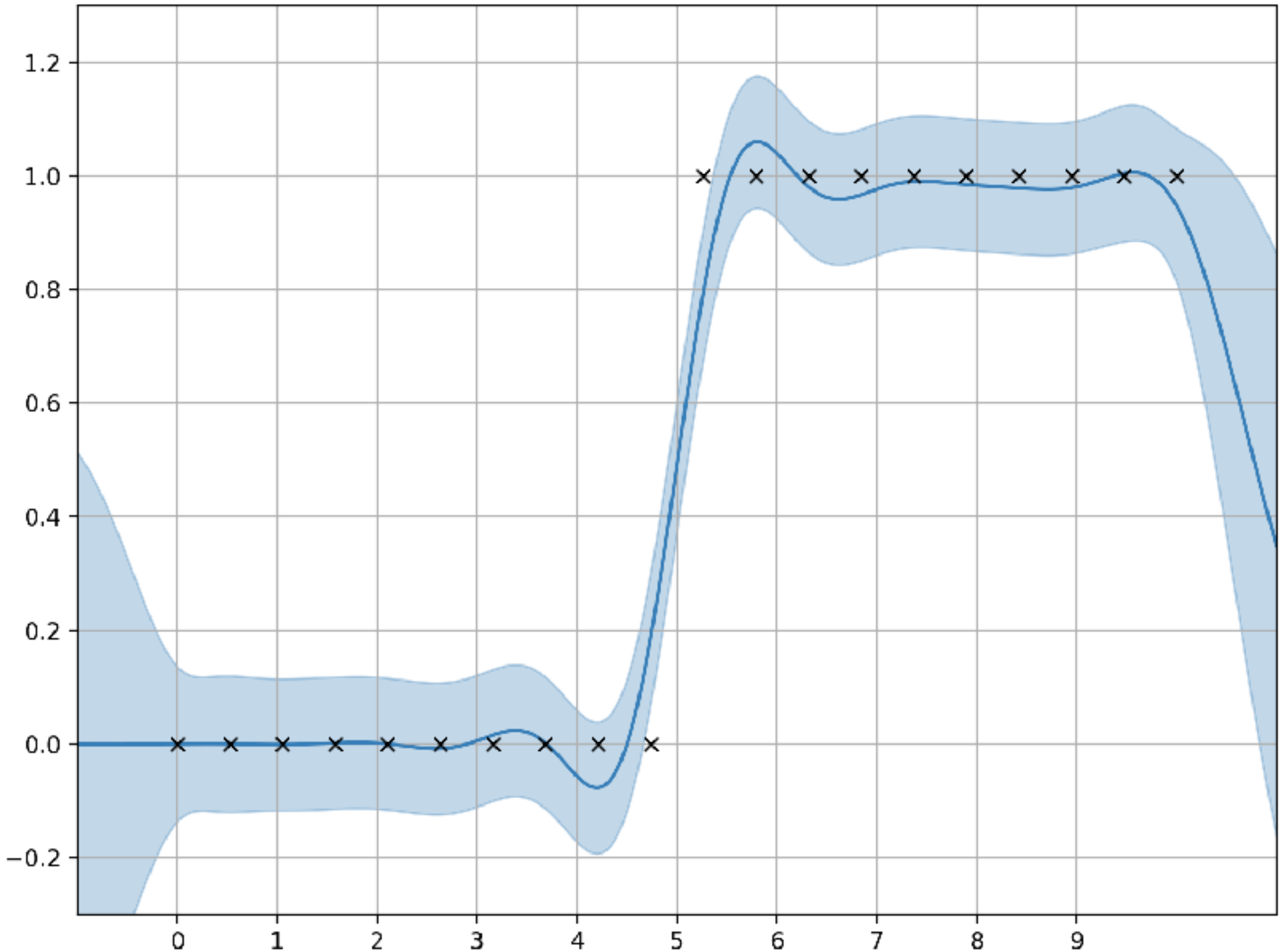
Gaussian process limitations



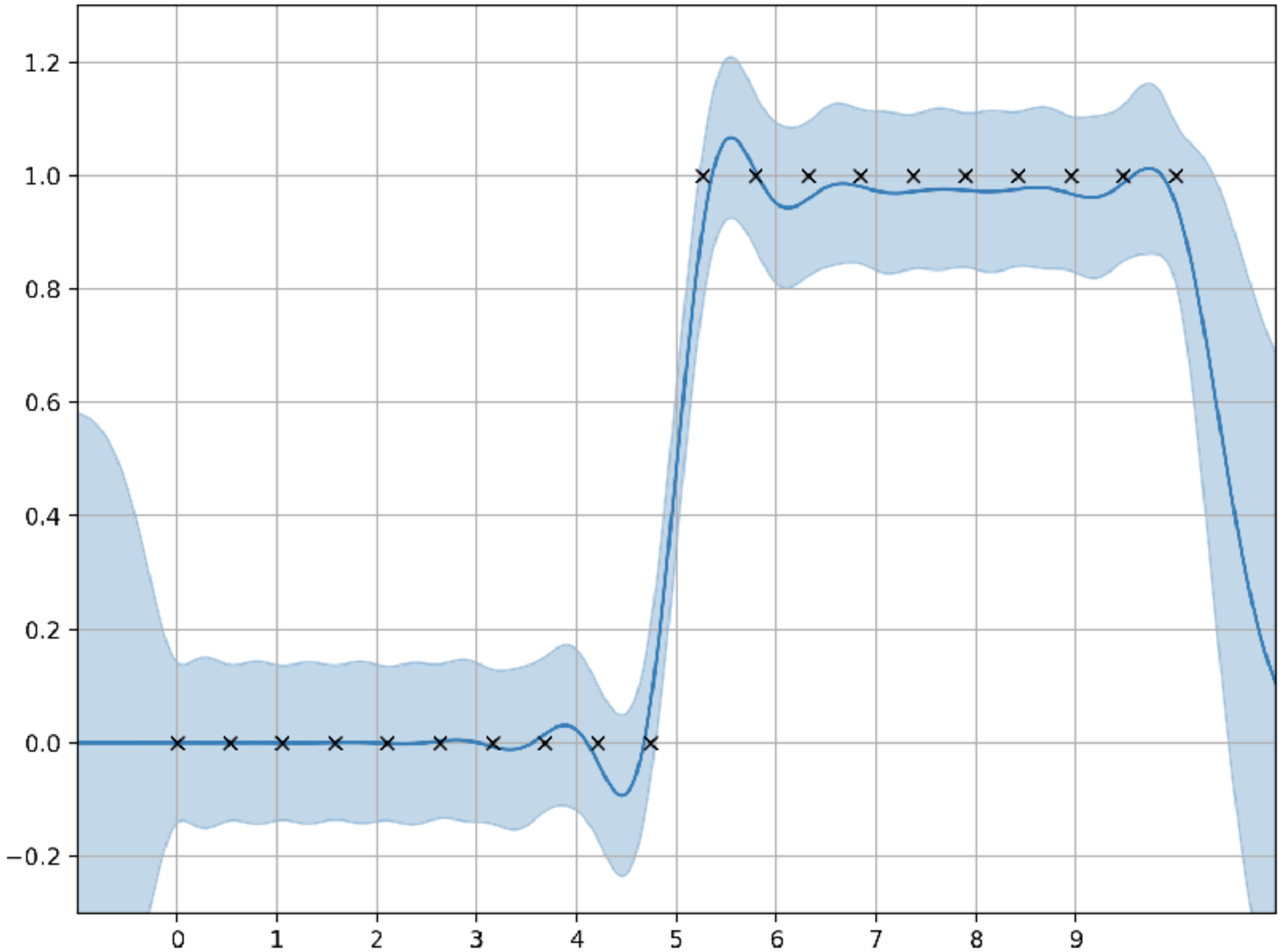
Gaussian process limitations



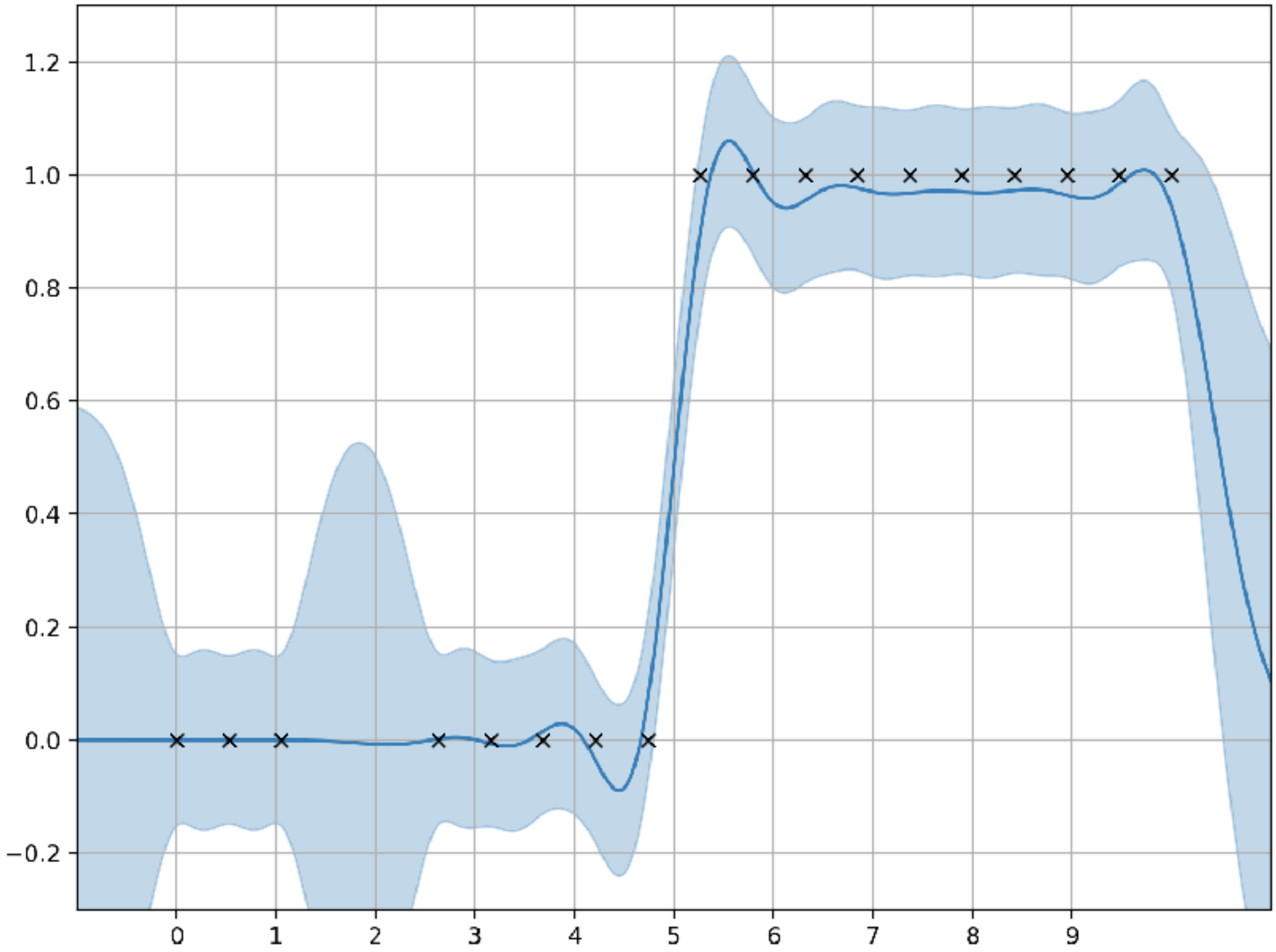
Gaussian process limitations



Gaussian process limitations



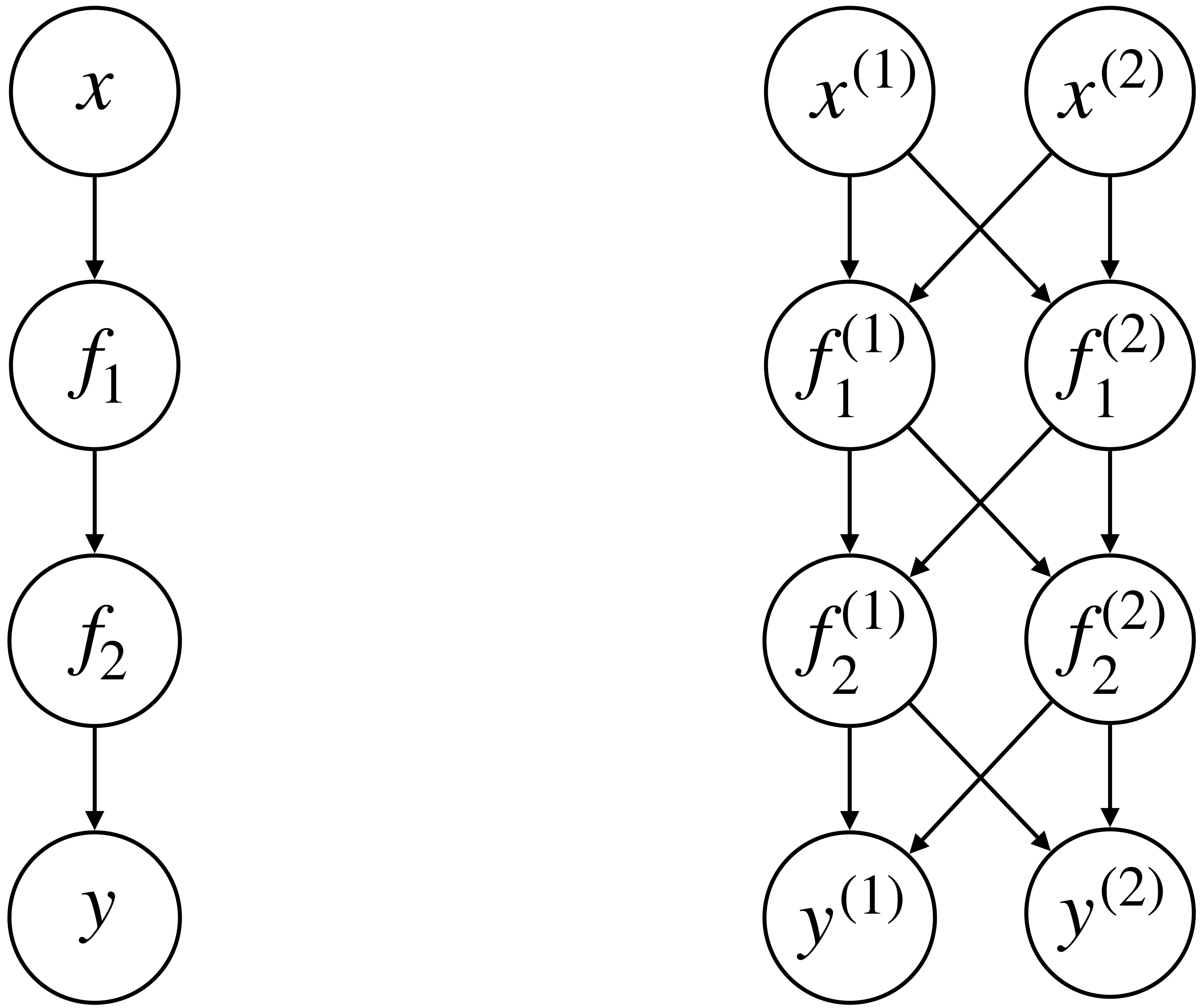
Gaussian process limitations



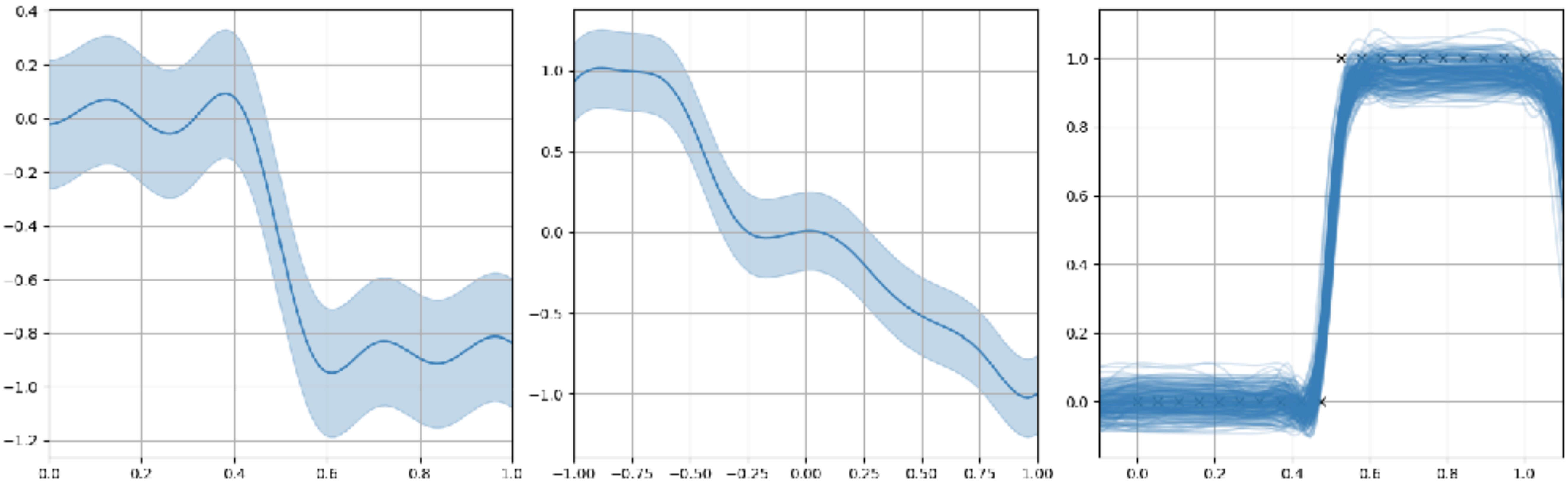
Gaussian process limitations

Remedies:

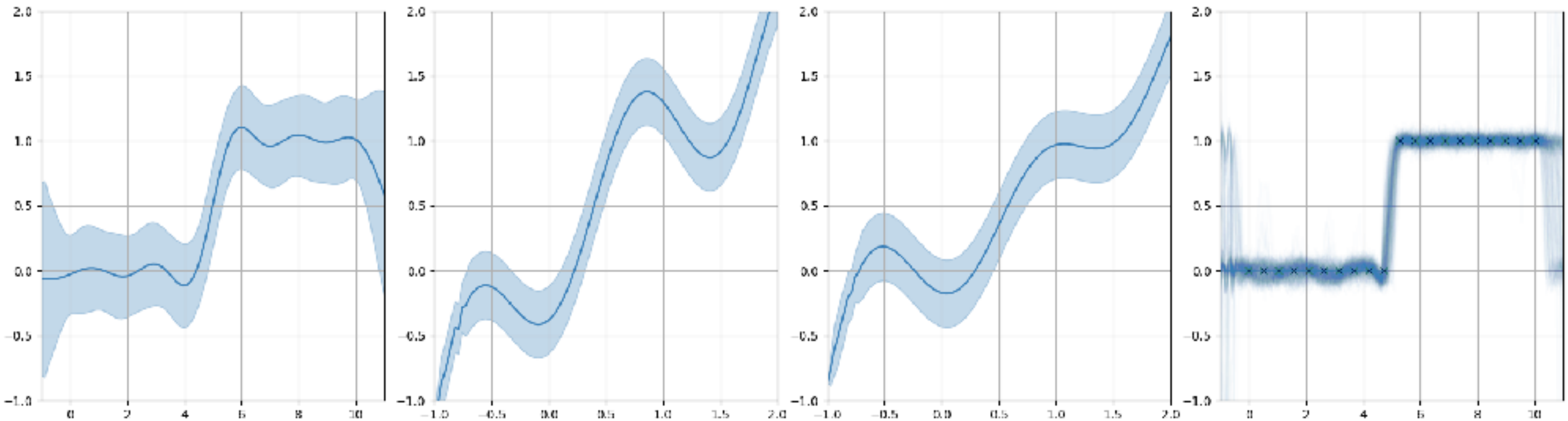
- Non-stationary kernels
- Heteroscedastic observation noise
- *Deep Gaussian processes*



x
f_1
f_2
y
$x^{(1)}$
$x^{(2)}$
$f_1^{(1)}$
$f_1^{(2)}$
$f_2^{(1)}$
$f_2^{(2)}$
$y^{(1)}$
$y^{(2)}$



0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
−1.2
0.0
0.2
0.4
0.6
0.8
1.0
0.5
−0.5
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00


2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
0
2
4
6
8
10
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0

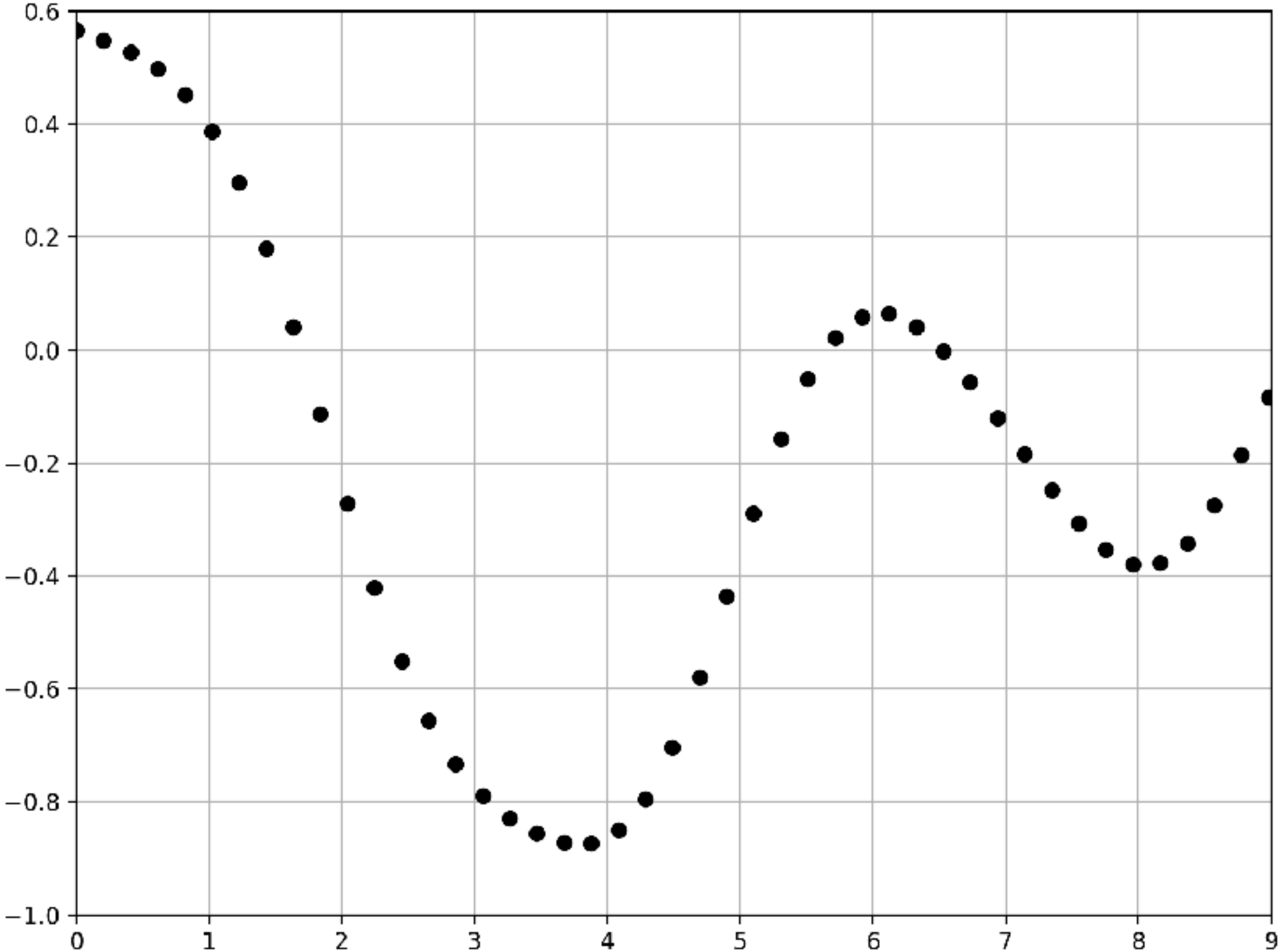
Why we cannot do it:

$$p(\mathbf{y} \mid \mathbf{x}) = \int_{\mathbf{f}_1, \mathbf{f}_2} p(\mathbf{y} \mid \mathbf{f}_2) p(\mathbf{f}_2 \mid \mathbf{f}_1) p(\mathbf{f}_1 \mid \mathbf{x})$$

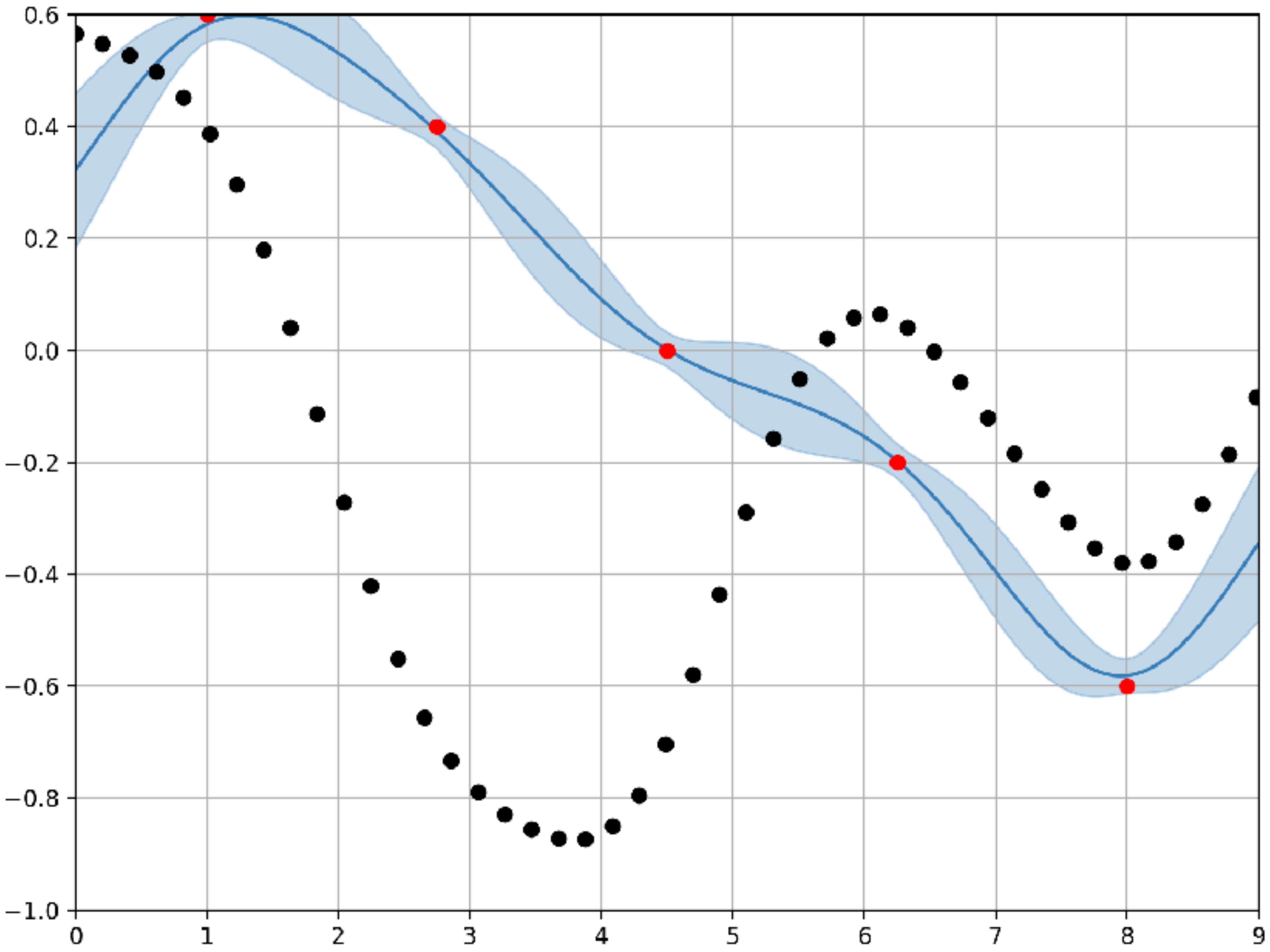
$$\mathcal{N}(\mathbf{f}_2 \mid \mathbf{0}, \mathbf{K}_{\mathbf{f}_1}) = (2\pi)^{-\frac{N}{2}} \cdot |\mathbf{K}_{\mathbf{f}_1}|^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2}\mathbf{f}_2^T \mathbf{K}_{\mathbf{f}_1}^{-1} \mathbf{f}_2\right)$$

$$(\mathbf{K}_{\mathbf{f}_1})_{i,j} = \kappa((f_1)_i, (f_1)_j)$$

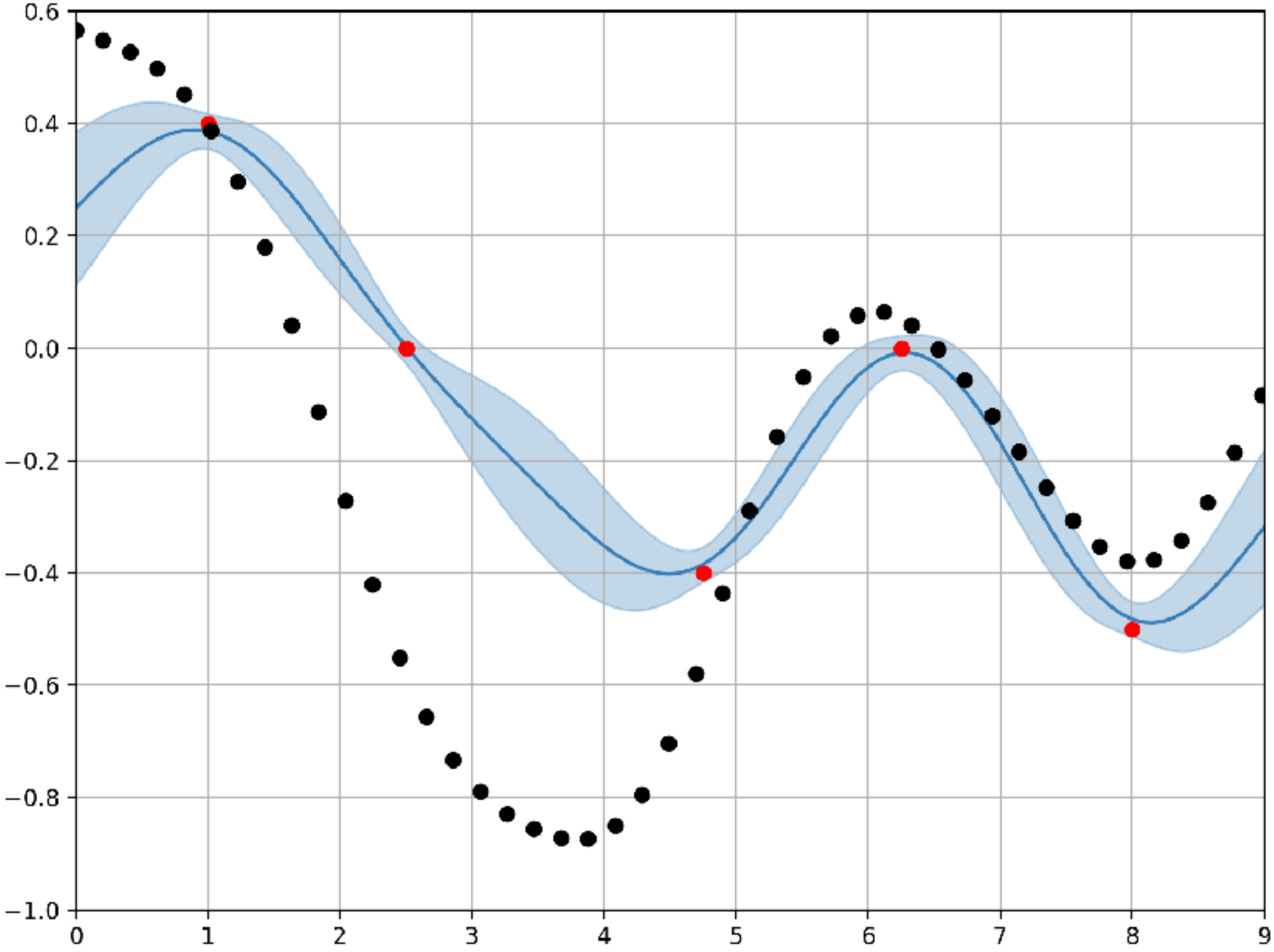
Variational inference of auxiliary data for sparse Gaussian processes



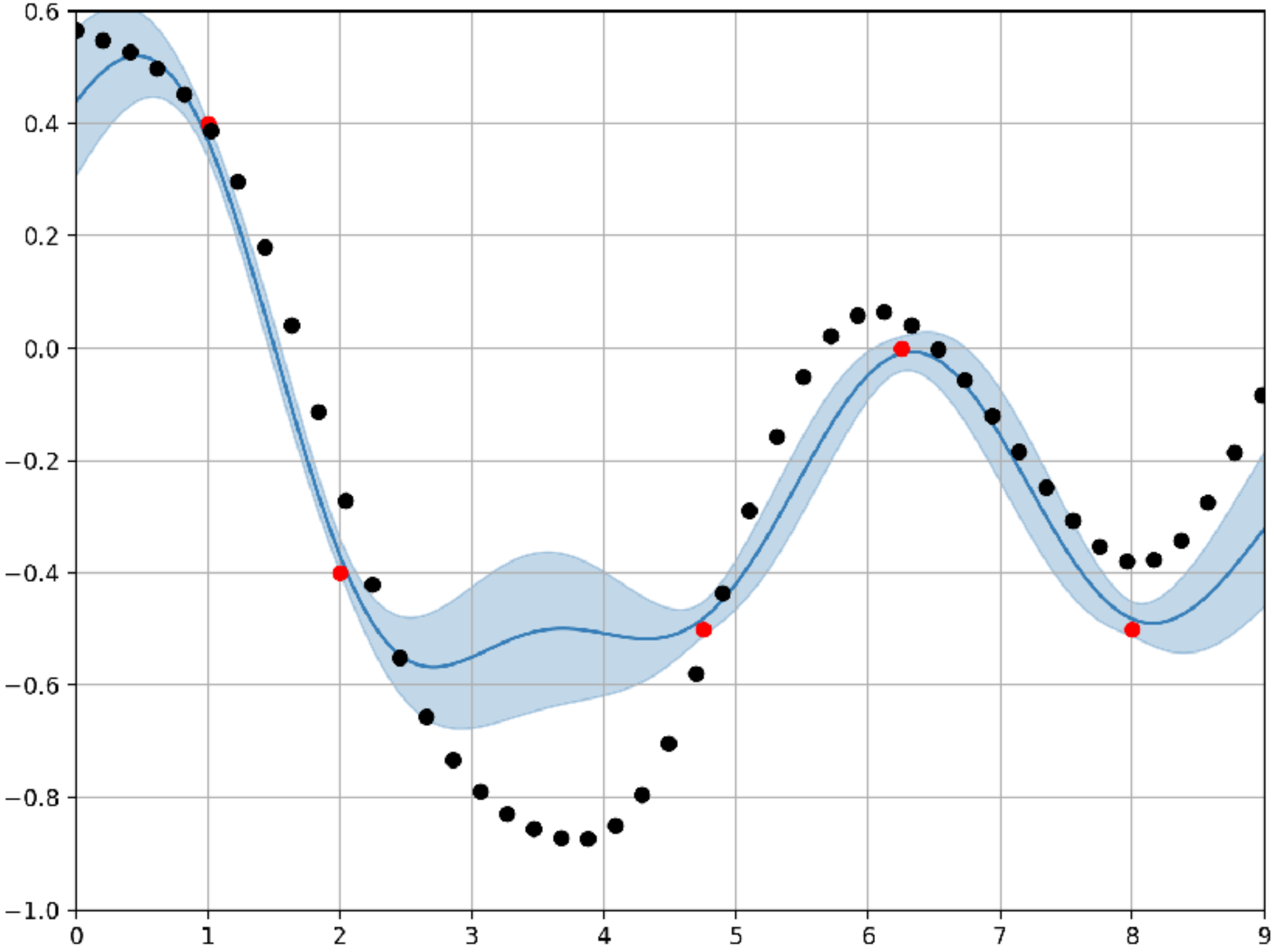
Variational inference of auxiliary data for sparse Gaussian processes



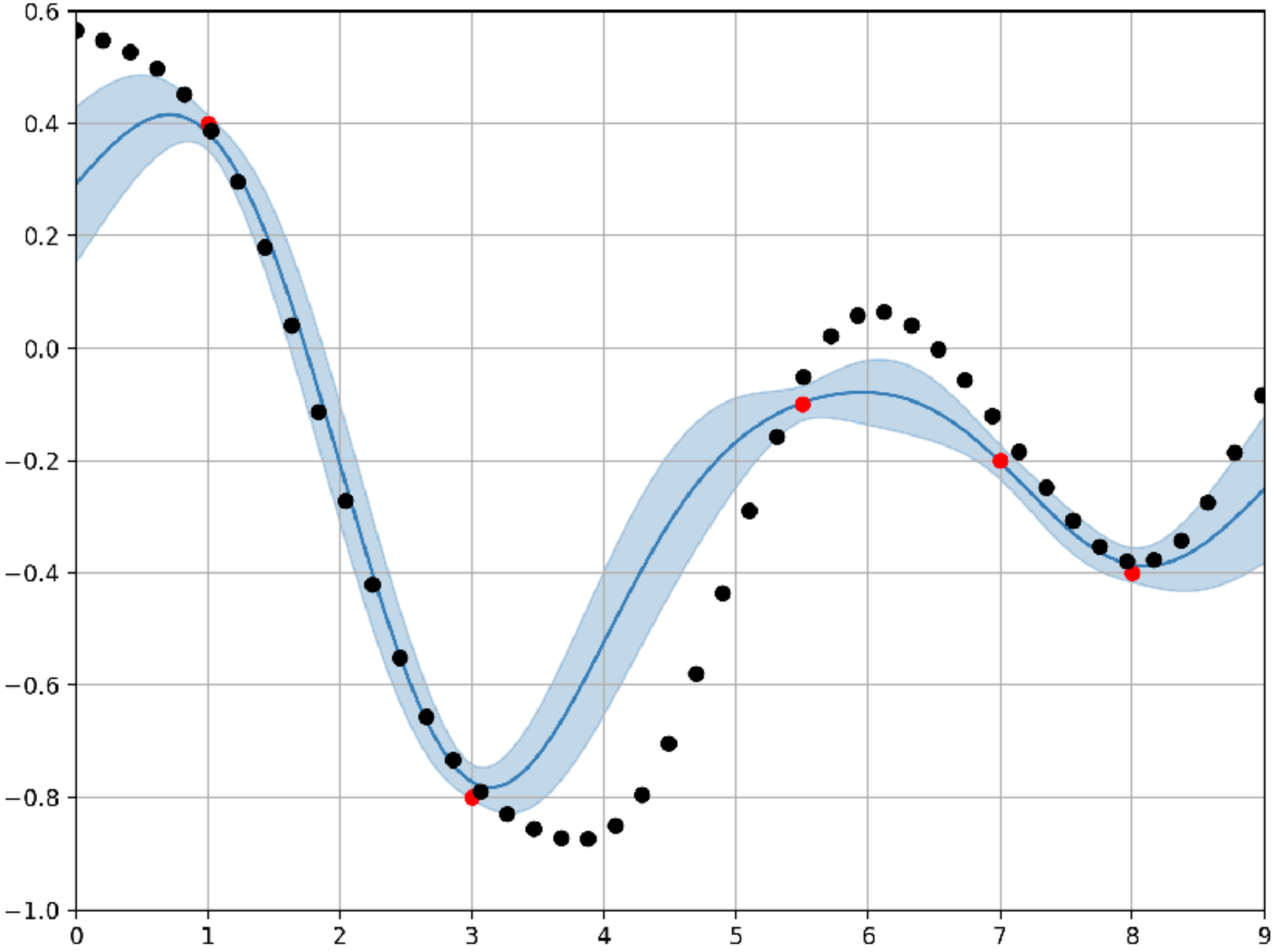
Variational inference of auxiliary data for sparse Gaussian processes



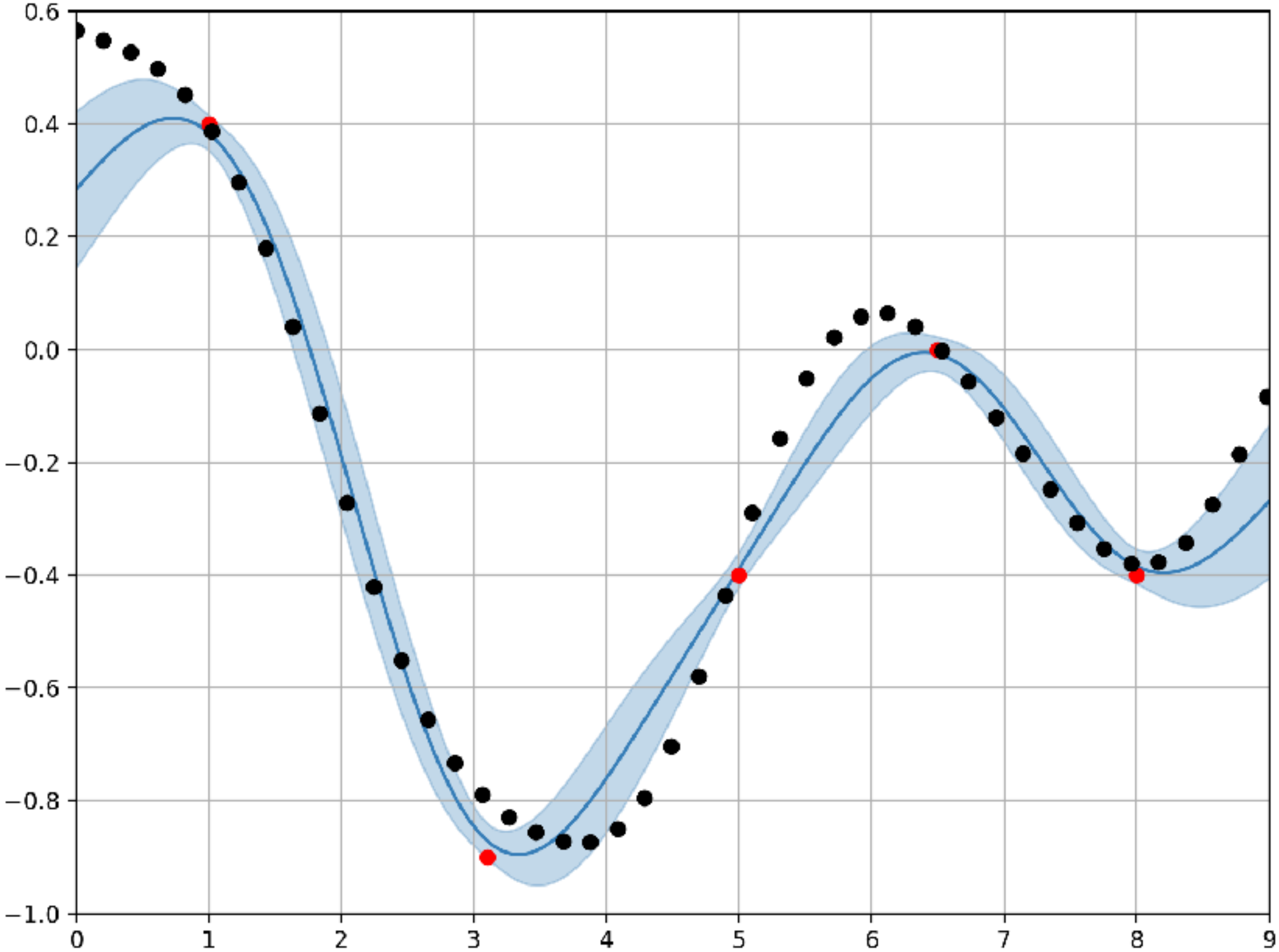
Variational inference of auxiliary data for sparse Gaussian processes



Variational inference of auxiliary data for sparse Gaussian processes



Variational inference of auxiliary data for sparse Gaussian processes



Assume **u** and **z** to be drawn from the same GP as **f** and **x**

$$p(\mathbf{f} \mid \mathbf{x}) = \mathcal{N}(\mathbf{f} \mid \mathbf{0}, \mathbf{K}_{NN})$$
$$p(\mathbf{u} \mid \mathbf{z}) = \mathcal{N}(\mathbf{u} \mid \mathbf{0}, \mathbf{K}_{MM}) \qquad M \ll N$$

Assume **u** and **z** to be drawn from the same GP as **f** and **x**

$$p(\mathbf{f} \mid \mathbf{x}) = \mathcal{N}(\mathbf{f} \mid \mathbf{0}, \mathbf{K}_{NN})$$

$$p(\mathbf{u} \mid \mathbf{z}) = \mathcal{N}(\mathbf{u} \mid \mathbf{0}, \mathbf{K}_{MM}) \qquad M \ll N$$

$$\begin{aligned} p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}) &= \mathcal{N}\left(\begin{bmatrix} \mathbf{f} \\ \mathbf{u} \end{bmatrix} \mid \mathbf{0}, \begin{bmatrix} \mathbf{K}_{NN} & \mathbf{K}_{NM} \\ \mathbf{K}_{MN} & \mathbf{K}_{MM} \end{bmatrix} \right) \\ &= p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) p(\mathbf{u} \mid \mathbf{z}) \end{aligned}$$

$$p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) = \mathcal{N}\left(\mathbf{f} \mid \mathbf{K}_{NM}\mathbf{K}_{MM}^{-1}\mathbf{u},\ \ \mathbf{K}_{NN} - \mathbf{K}_{NM}\mathbf{K}_{MM}^{-1}\mathbf{K}_{MN}\right)$$

New marginal likelihood:

$$
\begin{aligned}
p(\mathbf{y} \mid \mathbf{x}) &= \int_{\mathbf{f},\mathbf{u}} p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}) \\
&= \int_{\mathbf{f},\mathbf{u}} p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) p(\mathbf{u} \mid \mathbf{z})
\end{aligned}
$$

Variational inference:

1. Introduce approximate posterior: $q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$
2. Optimise evidence lower bound.

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$\log p(\mathbf{y} \mid \mathbf{x}) = \log \int_{\mathbf{f},\mathbf{u}} p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$\log \left(\mathbb{E}_{p(\mathbf{f},\mathbf{u}|\mathbf{x},\mathbf{z})}[p(\mathbf{y} \mid \mathbf{f})] \right)$$

$$\log p(\mathbf{y} \mid \mathbf{x}) = \log \int_{\mathbf{f},\mathbf{u}} p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$\log \left(\mathbb{E}_{p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}[p(\mathbf{y} \mid \mathbf{f})] \right)$$

$$\begin{aligned} \log p(\mathbf{y} \mid \mathbf{x}) &= \log \int_{\mathbf{f}, \mathbf{u}} p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}) \\ &= \log \int_{\mathbf{f}, \mathbf{u}} q(\mathbf{f}, \mathbf{u}) \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \end{aligned}$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$\log \left(\mathbb{E}_{p(\mathbf{f},\mathbf{u}|\mathbf{x},\mathbf{z})}[p(\mathbf{y} \mid \mathbf{f})] \right)$$

$$\log p(\mathbf{y} \mid \mathbf{x}) = \log \int_{\mathbf{f},\mathbf{u}} p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})$$

$$= \log \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}$$

$$\log \left(\mathbb{E}_{q(\mathbf{f},\mathbf{u})} \left[\frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \right] \right)$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$\log\left(\mathbb{E}_{p(\mathbf{f},\mathbf{u}|\mathbf{x},\mathbf{z})}[p(\mathbf{y} \mid \mathbf{f})]\right)$$

$$\log p(\mathbf{y} \mid \mathbf{x}) = \log \int_{\mathbf{f},\mathbf{u}} p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})$$

$$= \log \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \qquad \log\left(\mathbb{E}_{q(\mathbf{f},\mathbf{u})}\left[\frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}\right]\right)$$

$$\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \qquad \mathbb{E}_{q(\mathbf{f},\mathbf{u})}\left[\log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}\right]$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$\log\left(\mathbb{E}_{p(\mathbf{f},\mathbf{u}|\mathbf{x},\mathbf{z})}[p(\mathbf{y} \mid \mathbf{f})]\right)$$

$$\log p(\mathbf{y} \mid \mathbf{x}) = \log \int_{\mathbf{f},\mathbf{u}} p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})$$

$$= \log \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \qquad \log\left(\mathbb{E}_{q(\mathbf{f},\mathbf{u})}\left[\frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}\right]\right)$$

$$\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \qquad \mathbb{E}_{q(\mathbf{f},\mathbf{u})}\left[\log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}\right]$$

Equivalent to minimising $\mathbb{KL}[q(\mathbf{f}, \mathbf{u}) \parallel p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})]$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u})q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u})q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \\ &= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f})p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z})p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{f}, \mathbf{u})}\end{aligned}$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) p(\mathbf{u} \mid \mathbf{z})}{p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) q(\mathbf{u})}
\end{aligned}$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u})q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) \cancel{p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z})} p(\mathbf{u} \mid \mathbf{z})}{\cancel{p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z})} q(\mathbf{u})}
\end{aligned}$$

Deriving the evidence lower bound (ELBO)

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y}, \mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{f}, \mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) \cancel{p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z})} p(\mathbf{u} \mid \mathbf{z})}{\cancel{p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z})} q(\mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{u})}
\end{aligned}$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad\qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{u})}$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad\qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathbb{KL}[q(\mathbf{u}) \| p(\mathbf{u} \mid \mathbf{z})]
\end{aligned}$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathbb{KL}[q(\mathbf{u}) \| p(\mathbf{u} \mid \mathbf{z})] \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathbb{KL}[\mathcal{N}(\mathbf{u} \mid \mathbf{m}, \mathbf{S}) \| \mathcal{N}(\mathbf{u} \mid \mathbf{0}, \mathbf{K}_{MM}))
\end{aligned}$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad\qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log \frac{p(\mathbf{y} \mid \mathbf{f}) p(\mathbf{u} \mid \mathbf{z})}{q(\mathbf{u})} \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathbb{KL}[q(\mathbf{u}) \| p(\mathbf{u} \mid \mathbf{z})] \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathbb{KL}[\mathcal{N}(\mathbf{u} \mid \mathbf{m}, \mathbf{S}) \| \mathcal{N}(\mathbf{u} \mid \mathbf{0}, \mathbf{K}_{MM})) \\
&= \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathcal{L}_1
\end{aligned}$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned} \log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1 \\ &= \int_{\mathbf{f}} \left[\int_{\mathbf{u}} p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \mathcal{N}(\mathbf{u} \mid \mathbf{m}, \mathbf{S}) \right] \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1 \end{aligned}$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad\qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathcal{L}_1 \\
&= \int_{\mathbf{f}} \left[\int_{\mathbf{u}} p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \mathcal{N}(\mathbf{u} \mid \mathbf{m}, \mathbf{S}) \right] \log p(\mathbf{y} \mid \mathbf{f}) - \mathcal{L}_1 \\
&= \int_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathcal{L}_1
\end{aligned}$$

Deriving the evidence lower bound (ELBO), cont.

$$q(\mathbf{f}, \mathbf{u}) \approx p(\mathbf{f}, \mathbf{u} \mid \mathbf{x}, \mathbf{z}, \mathbf{y})$$

$$q(\mathbf{f}, \mathbf{u}) = q(\mathbf{f} \mid \mathbf{u}) q(\mathbf{u})$$

$$q(\mathbf{f} \mid \mathbf{u}) \triangleq p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \qquad q(\mathbf{u}) \triangleq \mathcal{N}(\mathbf{m}, \mathbf{S})$$

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f},\mathbf{u}} q(\mathbf{f}, \mathbf{u}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathcal{L}_1 \\
&= \int_{\mathbf{f}} \left[\int_{\mathbf{u}} p(\mathbf{f} \mid \mathbf{u}, \mathbf{x}, \mathbf{z}) \mathcal{N}(\mathbf{u} \mid \mathbf{m}, \mathbf{S}) \right] \log p(\mathbf{y} \mid \mathbf{f}) - \mathcal{L}_1 \\
&= \int_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathcal{L}_1
\end{aligned}$$

$$p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S}) = \mathcal{N}(\mathbf{f} \mid \mu_f, \Sigma_f)$$

$$\mu_f = \mathbf{K}_{MM}^{-1} \mathbf{m}$$

$$\Sigma_f = \mathbf{K}_{NN} - \mathbf{K}_{NM} \mathbf{K}_{MM}^{-1} (\mathbf{K}_{MM} - \mathbf{S}) \mathbf{K}_{MM}^{-1} \mathbf{K}_{MN}$$

Deriving the evidence lower bound (ELBO), cont.

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1$$

$$\approx \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1, \qquad \tilde{\mathbf{f}} \sim p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S})$$

Deriving the evidence lower bound (ELBO), cont.

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1$$

$$\approx \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1, \qquad \tilde{\mathbf{f}} \sim p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S})$$

Since the likelihood factorises over data points:

$$\cdots = \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \log \prod_{i=1}^{N} p(y_i \mid f_i) - \mathscr{L}_1$$

Deriving the evidence lower bound (ELBO), cont.

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1$$

$$\approx \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1, \qquad \tilde{\mathbf{f}} \sim p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S})$$

Since the likelihood factorises over data points:

$$\cdots = \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \log \prod_{i=1}^{N} p(y_i \mid f_i) - \mathscr{L}_1$$

$$= \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \sum_{i=1}^{N} \log p(y_i \mid f_i) - \mathscr{L}_1$$

Deriving the evidence lower bound (ELBO), cont.

$$
\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f}} p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S}) \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1 \\
&\approx \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \log p(\mathbf{y} \mid \mathbf{f}) - \mathscr{L}_1, \qquad\qquad \tilde{\mathbf{f}} \sim p(\mathbf{f} \mid \mathbf{x}, \mathbf{z}, \mathbf{m}, \mathbf{S})
\end{aligned}
$$

Since the likelihood factorises over data points:

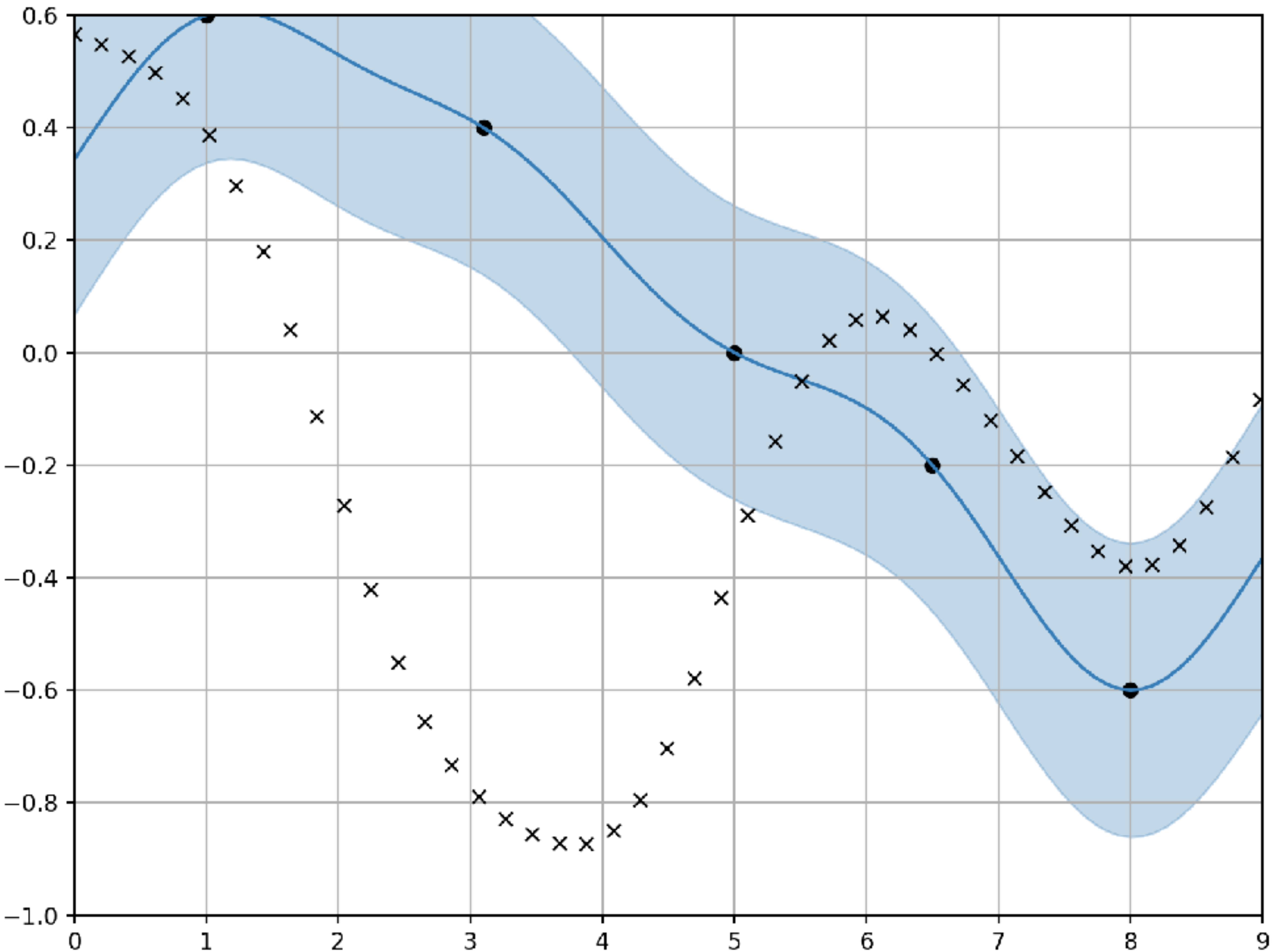
$$
\begin{aligned}
\cdots &= \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \log \prod_{i=1}^{N} p(y_i \mid f_i) - \mathscr{L}_1 \\
&= \frac{1}{K} \sum_{\tilde{\mathbf{f}}} \sum_{i=1}^{N} \log p(y_i \mid f_i) - \mathscr{L}_1 \\
&= \frac{1}{K} \sum_{i=1}^{N} \sum_{\tilde{f}_i} \log p(y_i \mid f_i) - \mathscr{L}_1, \qquad \tilde{f}_i \sim p(f_i \mid x_i, \mathbf{z}, \mathbf{m}, \mathbf{S})
\end{aligned}
$$

Sampling based approach

1. Draw a mini-batch $\tilde{\mathbf{f}} = \{\tilde{f}_j \quad | \quad \tilde{f}_j \sim p(f_j \mid x_j, \mathbf{z}, \mathbf{m}, \mathbf{S})\}$

2. Evaluate $\displaystyle\sum_{\tilde{f}_j \in \tilde{\mathbf{f}}} \log p(y_j \mid f_j) - \frac{1}{|\tilde{\mathbf{f}}|}\mathscr{L}_1$

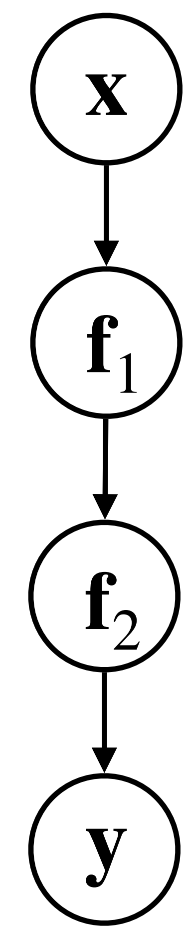
3. Make stochastic update to **z**, **m** and **S** (and potentially hyper-parameters)



0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
0
1
2
3
4
5
6
7
8
9

Deep Gaussian processes, revisited

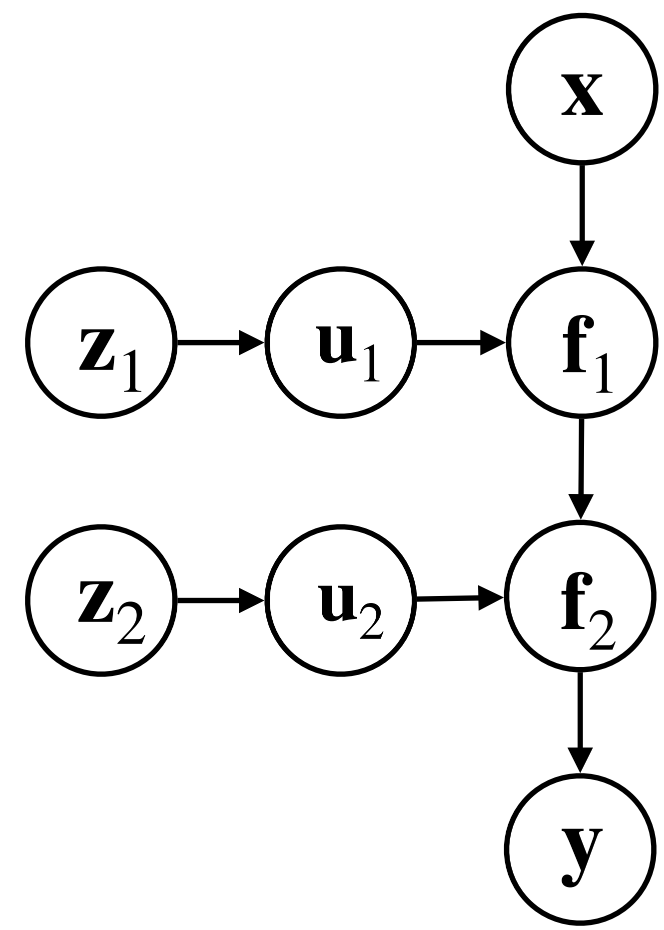
$$p(\mathbf{y} \mid \mathbf{x}) = \int_{\mathbf{f}_1,\mathbf{f}_2} p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{f}_2 \mid \mathbf{f}_1)p(\mathbf{f}_1 \mid \mathbf{x})$$



Deep Gaussian processes, revisited

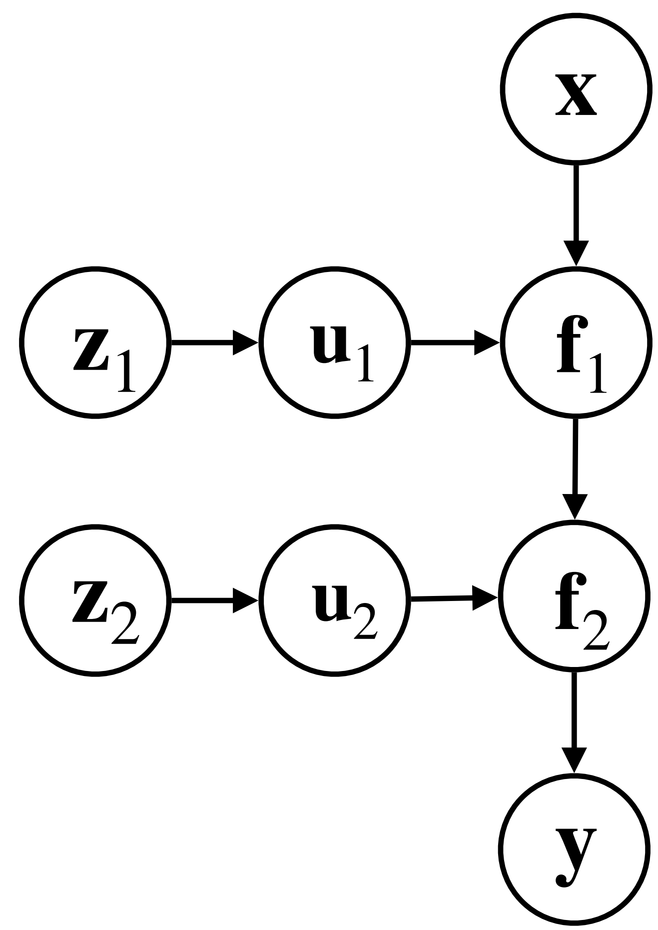
$$p(\mathbf{y} \mid \mathbf{x}) = \int_{\mathbf{f}_1,\mathbf{f}_2} p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{f}_2 \mid \mathbf{f}_1)p(\mathbf{f}_1 \mid \mathbf{x})$$

$$= \int_{\mathbf{f}_1,\mathbf{u}_1,\mathbf{f}_2,\mathbf{u}_2} p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})$$



New evidence lower bound

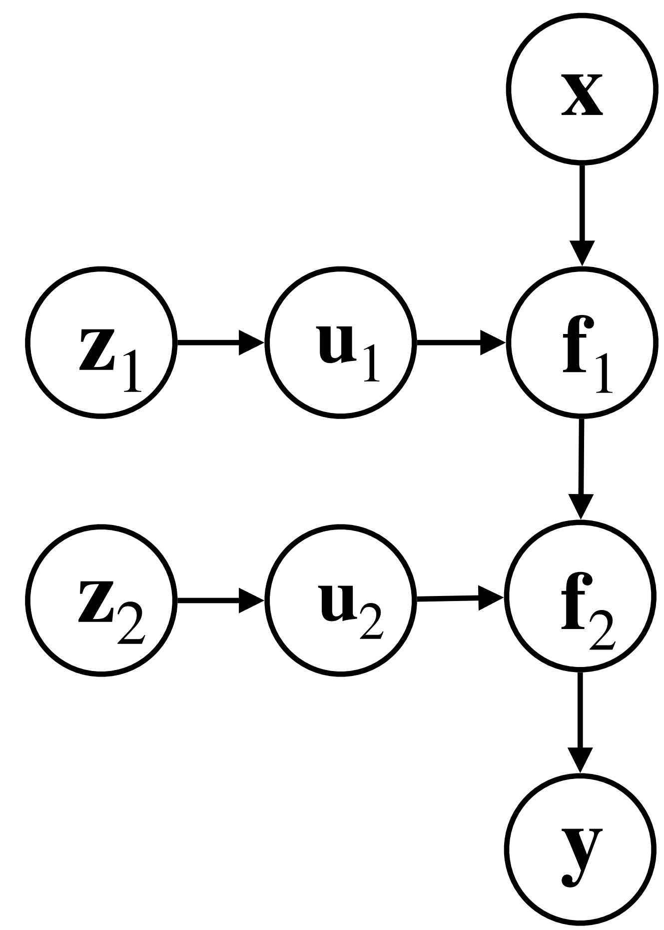
$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)}$$



New evidence lower bound

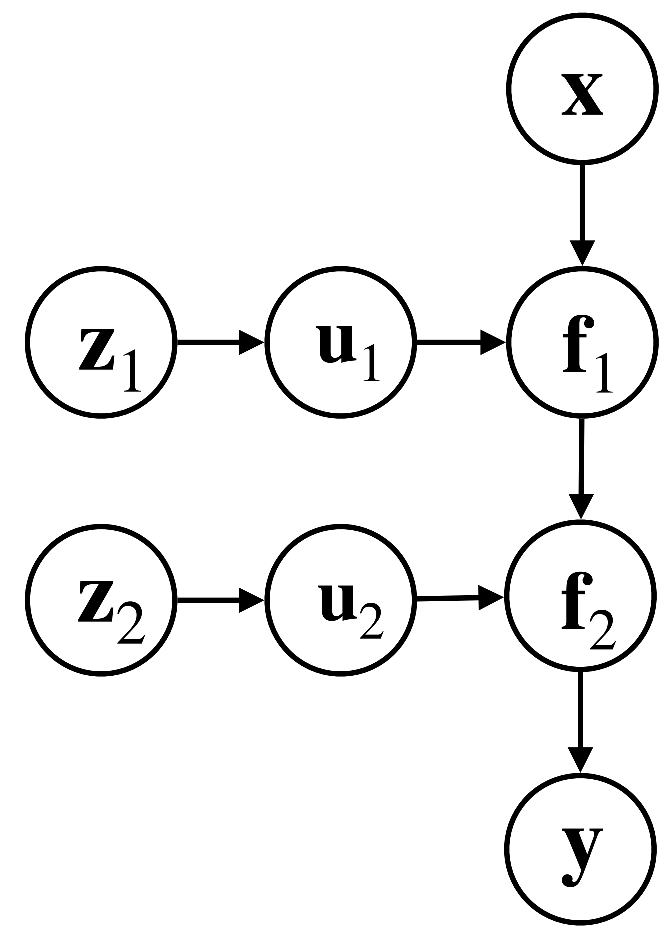
$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)}$$

$$= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) p(\mathbf{f}_2 \mid \mathbf{u}_2, \mathbf{z}_2, \mathbf{f}_1) p(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2 \mid \mathbf{u}_2) q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)}$$



New evidence lower bound

$$
\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) \cancel{p(\mathbf{f}_2 \mid \mathbf{u}_2, \mathbf{z}_2, \mathbf{f}_1)} p(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{\cancel{q(\mathbf{f}_2 \mid \mathbf{u}_2)} q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)}
\end{aligned}
$$

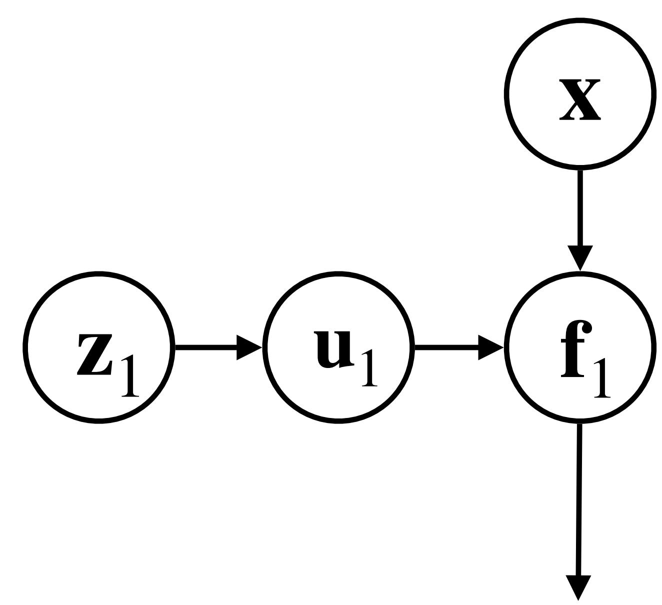


New evidence lower bound

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1) \log \frac{p(\mathbf{y},\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1 \mid \mathbf{z}_2,\mathbf{z}_1,\mathbf{x})}{q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1)}$$

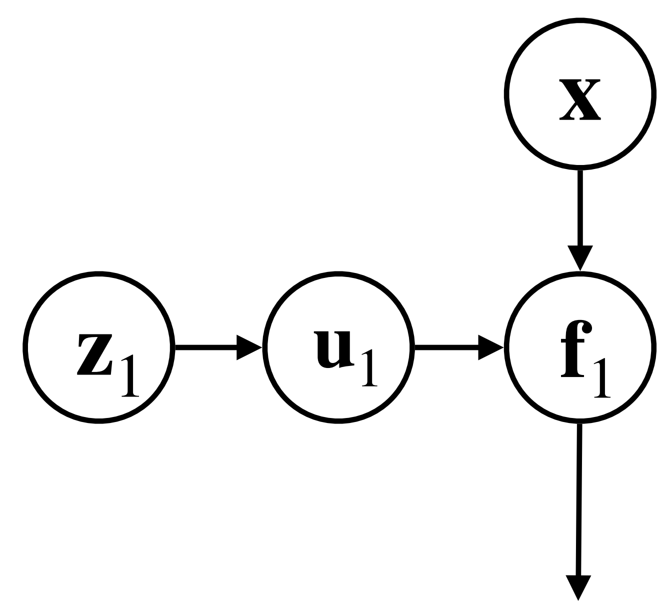
$$= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)\cancel{p(\mathbf{f}_2 \mid \mathbf{u}_2,\mathbf{z}_2,\mathbf{f}_1)}p(\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1 \mid \mathbf{z}_2,\mathbf{z}_1,\mathbf{x})}{\cancel{q(\mathbf{f}_2 \mid \mathbf{u}_2)}q(\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1)}$$

$$= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)p(\mathbf{f}_1,\mathbf{u}_1 \mid \mathbf{z}_1,\mathbf{x})}{q(\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1)}$$



New evidence lower bound

$$\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1) \log \frac{p(\mathbf{y},\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1 \mid \mathbf{z}_2,\mathbf{z}_1,\mathbf{x})}{q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)\cancel{p(\mathbf{f}_2 \mid \mathbf{u}_2,\mathbf{z}_2,\mathbf{f}_1)}p(\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1 \mid \mathbf{z}_2,\mathbf{z}_1,\mathbf{x})}{\cancel{q(\mathbf{f}_2 \mid \mathbf{u}_2)}q(\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)p(\mathbf{f}_1,\mathbf{u}_1 \mid \mathbf{z}_1,\mathbf{x})}{q(\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)p(\mathbf{f}_1 \mid \mathbf{u}_1,\mathbf{z}_1,\mathbf{x})p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2)q(\mathbf{f}_1 \mid \mathbf{u}_1)q(\mathbf{u}_1)}
\end{aligned}$$



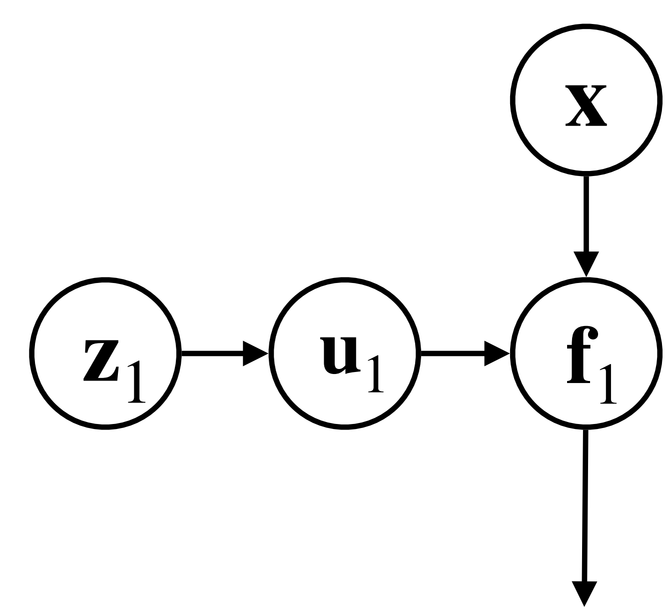
New evidence lower bound

$$\log p(\mathbf{y} \mid \mathbf{x}) \geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)}$$

$$= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) \cancel{p(\mathbf{f}_2 \mid \mathbf{u}_2, \mathbf{z}_2, \mathbf{f}_1)} p(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{\cancel{q(\mathbf{f}_2 \mid \mathbf{u}_2)} q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)}$$

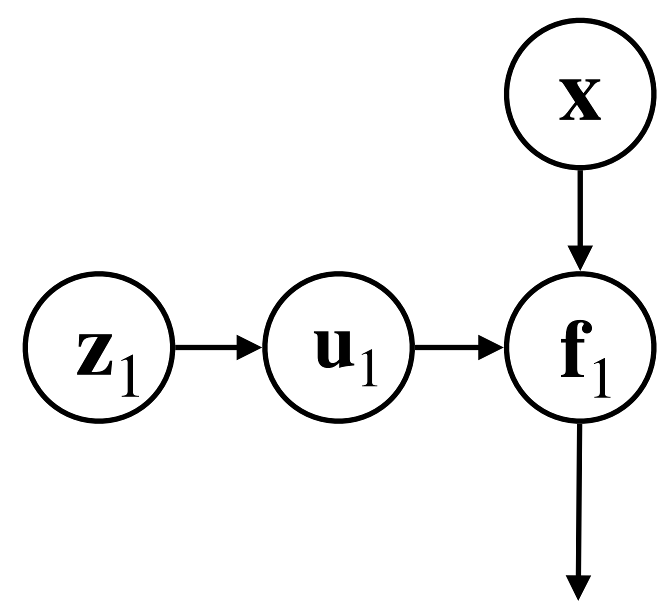
$$= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) p(\mathbf{u}_2 \mid \mathbf{z}_2) p(\mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_1, \mathbf{x})}{q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)}$$

$$= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) p(\mathbf{u}_2 \mid \mathbf{z}_2) \cancel{p(\mathbf{f}_1 \mid \mathbf{u}_1, \mathbf{z}_1, \mathbf{x})} p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2) \cancel{q(\mathbf{f}_1 \mid \mathbf{u}_1)} q(\mathbf{u}_1)}$$



New evidence lower bound

$$
\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) \cancel{p(\mathbf{f}_2 \mid \mathbf{u}_2, \mathbf{z}_2, \mathbf{f}_1)} p(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{\cancel{q(\mathbf{f}_2 \mid \mathbf{u}_2)} q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) p(\mathbf{u}_2 \mid \mathbf{z}_2) p(\mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_1, \mathbf{x})}{q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) p(\mathbf{u}_2 \mid \mathbf{z}_2) \cancel{p(\mathbf{f}_1 \mid \mathbf{u}_1, \mathbf{z}_1, \mathbf{x})} p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2) \cancel{q(\mathbf{f}_1 \mid \mathbf{u}_1)} q(\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2) p(\mathbf{u}_2 \mid \mathbf{z}_2) p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2) q(\mathbf{u}_1)}
\end{aligned}
$$



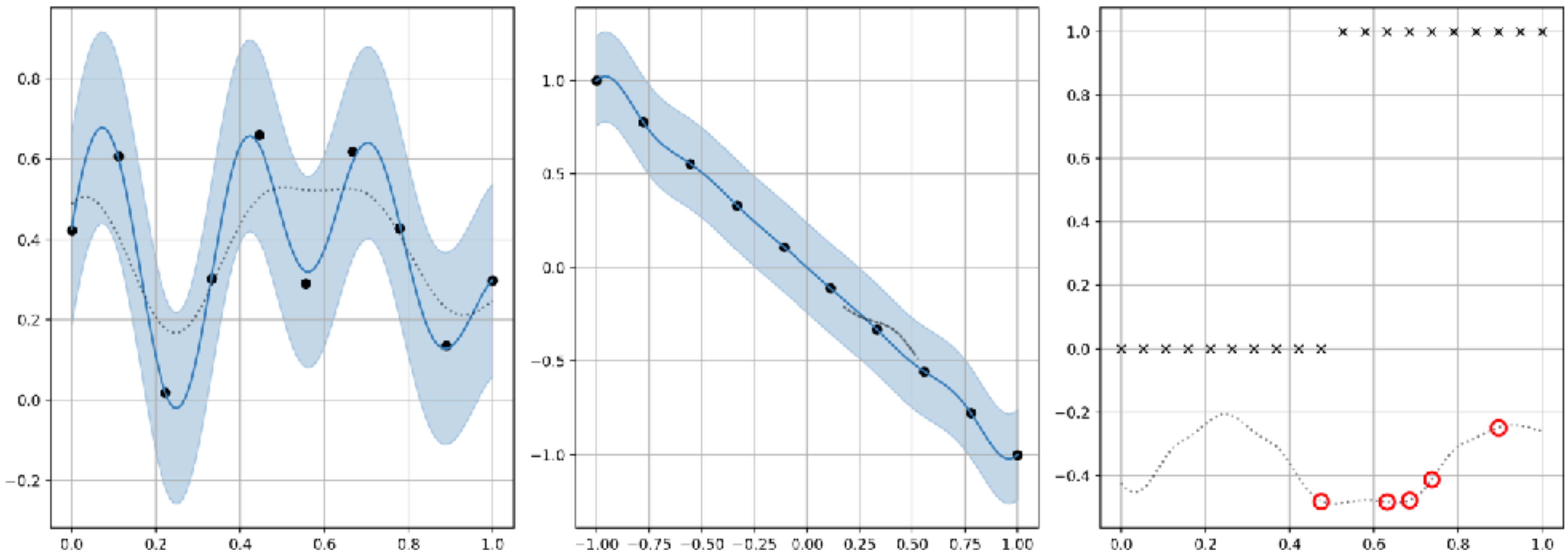
New evidence lower bound

$$
\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)\cancel{p(\mathbf{f}_2 \mid \mathbf{u}_2, \mathbf{z}_2, \mathbf{f}_1)}p(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{\cancel{q(\mathbf{f}_2 \mid \mathbf{u}_2)}q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)p(\mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_1, \mathbf{x})}{q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)\cancel{p(\mathbf{f}_1 \mid \mathbf{u}_1, \mathbf{z}_1, \mathbf{x})}p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2)\cancel{q(\mathbf{f}_1 \mid \mathbf{u}_1)}q(\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2)q(\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log p(\mathbf{y} \mid \mathbf{f}_2) - \mathbb{KL}[q(\mathbf{u}_2) \| p(\mathbf{u}_2 \mid \mathbf{z}_2)] - \mathbb{KL}[q(\mathbf{u}_1) \| p(\mathbf{u}_1 \mid \mathbf{z}_1)]
\end{aligned}
$$

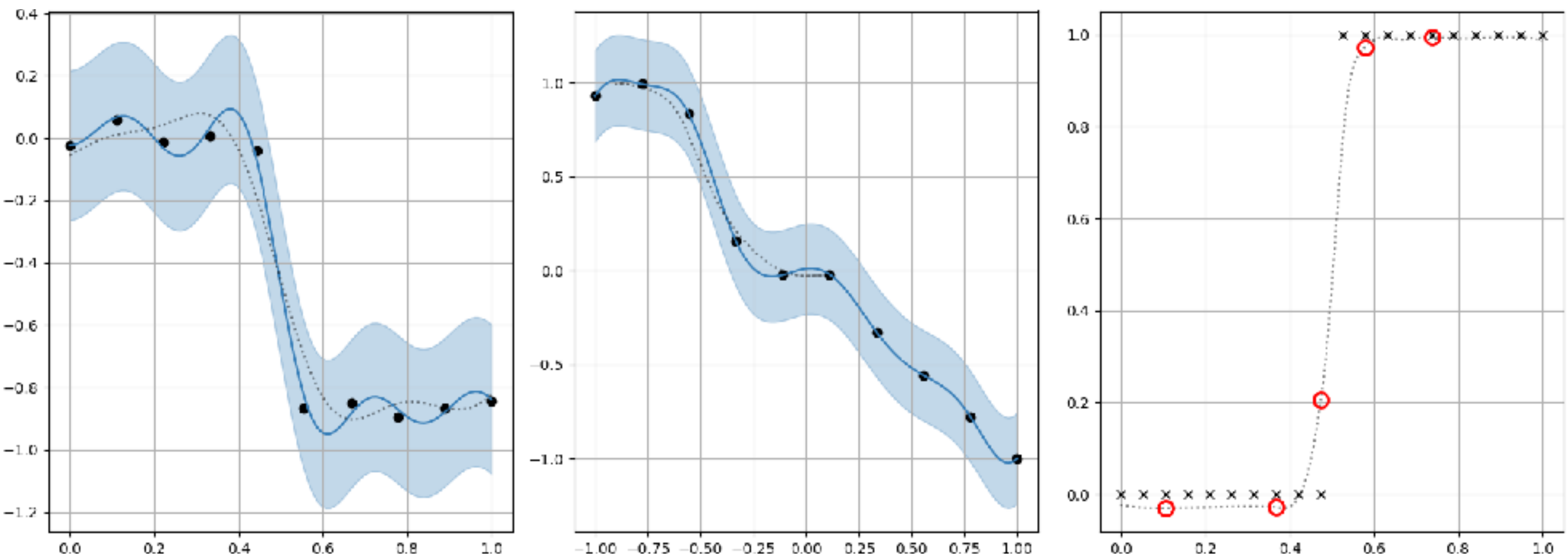
New evidence lower bound

$$
\begin{aligned}
\log p(\mathbf{y} \mid \mathbf{x}) &\geq \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y}, \mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)\cancel{p(\mathbf{f}_2 \mid \mathbf{u}_2, \mathbf{z}_2, \mathbf{f}_1)}p(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_2, \mathbf{z}_1, \mathbf{x})}{\cancel{q(\mathbf{f}_2 \mid \mathbf{u}_2)}q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)p(\mathbf{f}_1, \mathbf{u}_1 \mid \mathbf{z}_1, \mathbf{x})}{q(\mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)\cancel{p(\mathbf{f}_1 \mid \mathbf{u}_1, \mathbf{z}_1, \mathbf{x})}p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2)\cancel{q(\mathbf{f}_1 \mid \mathbf{u}_1)}q(\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log \frac{p(\mathbf{y} \mid \mathbf{f}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)p(\mathbf{u}_2 \mid \mathbf{z}_2)}{q(\mathbf{u}_2)q(\mathbf{u}_1)} \\
&= \int_{\mathbf{f}_2,\mathbf{u}_2,\mathbf{f}_1,\mathbf{u}_1} q(\mathbf{f}_2, \mathbf{u}_2, \mathbf{f}_1, \mathbf{u}_1) \log p(\mathbf{y} \mid \mathbf{f}_2) - \mathbb{KL}[q(\mathbf{u}_2) \| p(\mathbf{u}_2 \mid \mathbf{z}_2)] - \mathbb{KL}[q(\mathbf{u}_1) \| p(\mathbf{u}_1 \mid \mathbf{z}_1)] \\
&\approx \frac{1}{K} \sum_{i=1}^{N} \sum_{\tilde{f}_{2,i}} \log p(y_i \mid \tilde{f}_{2,i}) - \mathscr{L}_2 - \mathscr{L}_1, \qquad \tilde{f}_{2,i} \sim p(f_{2,i} \mid \tilde{f}_{1,i}, \mathbf{z}_2, \mathbf{m}_2, \mathbf{S}_2)
\end{aligned}
$$

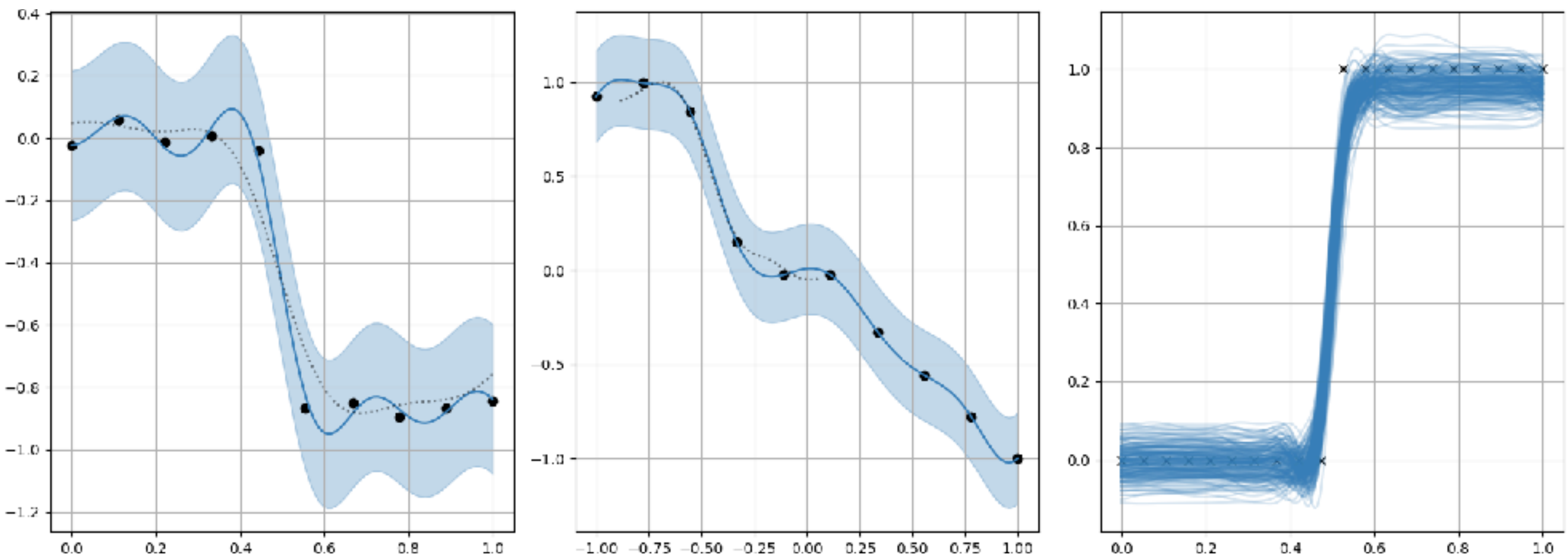
$$\tilde{f}_{1,i} \sim p(f_{1,i} \mid x_i, \mathbf{z}_1, \mathbf{m}_1, \mathbf{S}_1)$$



0.8
0.6
0.4
0.2
0.0
−0.2
0.0 0.2 0.4 0.6 0.8 1.0
1.0
0.5
0.0
−0.5
−1.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
0.0 0.2 0.4 0.6 0.8 1.0



0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
−1.2
0.0
0.2
0.4
0.6
0.8
1.0
1.0
0.5
0.0
−0.5
−1.0
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
−1.2
0.0
0.2
0.4
0.6
0.8
1.0
1.0
0.5
0.0
−0.5
−1.0
−1.00
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

Deep Gaussian processes research topics

- Multi-fidelity Bayesian optimisation
- Deep convolutional GP’s
- Recurrent processes
- Latent variable model for representation learning

- Optimisation and inference methods

Software

- DeepGPy (Damianou, Lawrence, et. al)
- GPFlow
- MXFusion
- …