**The Bayesian Model Design:**

The derivation of the analytic expression of the Bayesian estimate density function requires a series of modelling choices so that we may satisfy the following goals:

1. full integration of a precomputed density function prior into the expression,
2. the result estimate density function is non-negative,
3. allow design flexibility of assigning our uncertainty of different measurements, and
4. for prior density function with only using the measurement data, m, kernel, K, measurement noise, and the prior.

The discussion of the creation of this model is expands on work done by Paul Teal [].

We seek to simplify Bayes’ theorem in (eq) so that we may have one function of the mean.

**Gaussian Assumption for the Prior**

To allow for tractable manipulation of the probability density functions such as probability marginalisation and Bayes’ theorem, we assume each of the discretised points that make the T2 density function have a Gaussian distribution. As in,

f = \mathbb{N}(\mu\_f | C\_f)

There are two major implications with this assumption:

1. A Gaussian function is not bounded -- $N \in (-\inf, +\inf) $) -- meaning there is a possibility of a negative density function, violating the non-negativity constraint.
2. It may not describe the true density the prior density points if we do not have a large number of density functions making up the prior.

The second implication means that the prior must have a sufficient sum of density functions to form it such that the Central Limit Theorem so that we may assume normality. Therefore, this constrains usage to a use case where we already know the type of density function we are trying to find. This is valid for BFF estimation, as we typically know the type of porous media’s density function but not the bound fluid fraction of our specific measurement.

**Bayes Theorem for Multivariate Gaussians**

If we have two random multivariate variables x and y, then

X \sim N (\mu\_X, C\_X) \text{, and }

Y|X \sim N(MX + \mu\_Y, C\_Y) {, where} Y = MX + \epsilon \text{.}

Then according to Bayes’ theorem:

X |Y \sim N (R)

**Probability of a Measurement Given a Density Function p(m|f)**

The choice of how we model the measured data given the density function we are trying to predict adversely affects our Bayesian density function estimate.

For a direct expression of measurement data given a density function we have:

m|f \sim N(Kf + \mu\_f, C\_m)

The prior estimate must be compatible with the rest of the Bayesian framework to be usable. For example, 100 T2 relaxation bins may describe an experimental T2 density function. However, it requires conversion into 30 relaxation bins to be compatible with the estimator framework. Interpolation of the prior to the actual framework`s dimensionality is used to bridge between these two domains.

Any extrapolation is set to zero as we assume that the measurement tool is only sensitive for the range of T2 relaxation values we provide.

\begin{figure} [h]

\centering

\includegraphics[width=0.8\textwidth]{design/interpolation\_choice.eps}

\caption{Test of different interpolation techniques for interpolation from a lower dimensionality ($N\_y = 10$) to a higher dimensionality ($N\_y = 100$)}

\label{fig:interpolation\_comparison}

\end{figure}

Figure \ref{fig:interpolation\_comparison} demonstrates different interpolation schemes conversion from 10 T2 relaxation bins to 100 T2 relaxation bins. The nearest point interpolation and next point interpolation maintain the same discretisation so they are insufficient to recreate a density function. The linear interpolation does not have a smoothed result -- something unrealistic for T2 density functions. Examining all of the candidate interpolation techniques, we can see that:

\begin{enumerate}

\item Nearest point and next point interpolation maintains the discretised coarseness, making them poor at maintaining a representative T2 density function.

\item The linear interpolation is less coarse than the nearest and next point interpolates but it does not provide a smooth function. Not being smooth makes it insufficient for providing a function with the expected smoothness of a density function.

\item The spline interpolation and Makima (Modified Akima cubic Hermite) interpolation directly violate the non-negativity constraint on the density function.

\item This all leaves the shape-preserving piece-wise cubic interpolation (PCHIP) \cite{fritsch1980monotone} as it returns a valid non-negative function that smooth. These two crucial aspects make it representative of a T2 density function attained through NMR relaxation.

\end{enumerate}

The resulting design decision made is to use PCHIP interpolation. It is fit-for-purpose in adapting experimental data to the Bayesian framework while preserving properties that describe T2 relaxation.

The first implication complicates the correctness of the model, as there is a danger of violating the physical non-negativity constraint. In order to provide an analytically tractable result, we make a trade off and relax this constraint. The prior mean is still non-negative as it corresponds directly to physical measurements, but an instantiation of a prior density function will not have this constraint.

This estimates the covariance directly from the high quality density functions \cite{DiscreteRandomSignalsBookCovarianceEst}. This modelled the dependence between different T2 relaxation bins in the density function. Hence, this covariance estimate fully describes the uncertainty and dependence of different points of the prior. However, it comes at a danger of overfitting directly to the prior data we are using.

The error of the estimate of the bound fluid volume will take the form of absolute error. A comparative error is used instead of displaying the actual estimate as we are utilising cross validation over thirty different samples of high quality experimental data. We want to establish a generalised measure of performance.