The General Problem

- We start with the physical system and denote the system value as \( y \in Y \). We often have observations on \( y \), denoted as \( z \), where \( z = H y + e \), where \( H \) is the incidence matrix, and \( e \) is the measurement error which is treated as independent of all other quantities.
- The simulator is a deterministic complex computer model for the physical system. We denote the simulator as \( f : X \rightarrow Y \), where \( x \in X \) are uncertain model parameters.
- We have \( n \) evaluations of the simulator at inputs \( X = \{x_1, \ldots, x_n\} \), and we denote the resulting evaluations as \( F = \{f(x_1), \ldots, f(x_n)\} \).
- We can also partition these quantities into historic values and future values to be predicted, i.e. \( F = (f_s, f_{\text{p}}) \).
- The emulator is then summarised by the mean function and variance function of \( f^* \), obtained from the mean function and variance function of the emulator for \( f \) (denoted \( \mu(x) \triangleq E[f(x)] \) and \( \kappa(x, x') \triangleq \text{Var}[f(x), f(x')] \)), which are the only features of the emulator that we are required to specify.
- Using these values we can then compute the unconditional mean and variance of \( f^* \) by first conditioning on \( x^* \) and then integrating out with respect to the prior distribution.
- Given \( E[f^*] \) and \( \text{Var}[f^*] \), it is then straightforward to compute the joint mean and variance of the collection \( (y, z) \) with no gaussian requirement on the error terms.
- We can now evaluate the adjusted mean and variance for \( y \) adjusted by \( z \) using the Bayes linear adjustment formulae.

Bayes or Bayes Linear?

- Full Bayesian Methods
  - Gives a joint distribution for \( (x^*, y) \);
  - Requires strong distributional choices for \( f \) and \( e \), as well as for \( x^* \);
  - In all but small problems, tractability requires a gaussian distribution for \( \{f, e\} \);
  - Even then, large problems can be prohibitively expensive;
  - Very intricate computations may be able to reduce this expense, if the application is important enough;
- Bayes Linear
  - Requires full specification for \( x^* \), but only mean and covariance specification for \( f(x^*) \) and \( y \);
  - Much more tractable for larger problems.

Implausibility and Refocusing

- Calibration is learning about \( x^* \) using the simulator evaluations and \( z \).
- Using the emulator we can obtain, for each set of inputs \( x \), \( E[f_s(x)] \) and \( \text{Var}[f_s(x)] \).
- We seek to rule out regions of \( x^* \in X \) which are unlikely to have given rise to \( z \).
- To achieve this, we calculate the implausibility: \( I_{\text{ly}}(x) = |E[f_s(x)] - z|/\sqrt{\text{Var}[f_s(x)] - z} \) for each component.
- This calculation can be performed univariately - there is no need for a full multivariate specification.
- The implausibilities are then combined, such as by using \( I_{\text{ly}}(x) = \max_i I_{\text{ly}}(x) \), and can then be used to identify regions of \( x \) with large \( I_{\text{ly}}(x) \) as implausible.
- With this information, we can then refocus our analysis on the non-implausible regions of the input space.

The Hat Run

- In calibrated prediction, we reintroduce calibration to the process whilst retaining tractability.
- We first find the Bayes Linear estimate of the best input \( x^* \) by using: \( \hat{x} = E[f^*] \).
- Next, we evaluate the simulator \( f \) at this estimate: \( \hat{x} = f(\hat{x}) \). This is called the hat run.
- By using the emulator, we then determine the mean and variance of \( f \) and the covariance between \( f \) and \( (z, y) \).
- Using all this information, we can now predict \( y \) by finding the adjusted mean and variance for \( y \) using both \( z \) and \( f \).
- The hat run introduces specially-identified local knowledge about \( f(x^*) \), as opposed to the global knowledge we get from general runs of the simulator.

Reification

- Reification concerns relating one or more simulators, \( \{f, f', \ldots\} \), to the underlying system \( y \).
- We link the existing simulators to the system by the reified simulator \( f^* \), which is an unknown simulator with higher accuracy than any of \( \{f, f', \ldots\} \) and which separates those simulators from the system.
- Advantages of reified modelling: \( x^* \) goes into \( f^* \) not \( f \); provides a coherent unification of several models (easy if they have markov relationship); does in two intuitive steps what we previously had to do in one.

Uncalibrated Forecasting

- The Bayes Linear approach is simpler in terms of belief specification and prediction.
- The key equations in the Bayes Linear approach are:
  \[
  E[y] = E[y] + \text{Cov}[y, z] \text{Var}[z]^{-1} (z - E[z]),
  \]
  \[
  \text{Var}[y] = \text{Var}[y] + \text{Cov}[y, z] \text{Var}[z]^{-1} \text{Cov}[z, y],
  \]
  where \( E[y] \) is the expectation for \( y \) adjusted by \( z \), and \( \text{Var}[y] \) is the variance of \( y \) adjusted by \( z \).

Figure 1: Independence graph representing the Best Input Approach

References