Training the Leaf Model

- No latent space: We approximate the latent distribution of latent variable models such as HMMs with a deterministic function given by a decision tree.

Expressive: We use interactions that are non-linear and non-parametric.

Efficient: Dropping the latent space allows for a fast computation of log-likelihoods and filterings.

Dynamic Forests

Training the Leaf Model

- Each leaf node \( \ell \) contains a regression model \((A_{\ell}, \Sigma_{\ell})\) that is based on a subset \( D_{\ell} \) of the training data.

- Given a design matrix \( \Phi \) and a target matrix \( U \), we estimate \( A_{\ell} \) using ridge regression,
  \[
  A_{\ell} := U \Phi^{\top} (\Phi \Phi^{\top} + \gamma I)^{-1}.
  \]

- Regularization of high-dimensional covariance estimates is achieved by projection to an isotropic target,
  \[
  \Sigma_{\ell} := d^{-1} \text{tr} \left( \Sigma_{\ell} \right) I,
  \]

  where \( \Sigma_{\ell} \) is the sample covariance of the residual vectors \( \{r_{\ell}\}_{\ell \in D_{\ell}} \).

Training the Tree Structure

- Recursively split leaves by selecting the best among a set of hyperplane splits.

- A candidate split \( s \) at leaf node \( \ell \) introduces child nodes \( u \) and \( v \), each of which receives a subset of \( D_{\ell} \). Splits are scored by reduction in residual error,
  \[
  Z_{\ell} := E_{\ell} - \left( E_u + E_v \right),
  \]

  with \( E_{\ell} := \sum_{\ell \in D_{\ell}} \| r_{\ell} \|^{2} \).

Ensembles

- We train \( C \) dynamic tree models with bootstrap resampled training data and aggregate their predictions by averaging.

- Given features \( \phi (p_a(x_{t})) \) extracted from the previous \( K \) frames, the posterior of the Dynamic Forest is given by a multimodal Gaussian mixture,
  \[
  p (x_{t} \mid p_a(x_{t})) = \frac{1}{C} \sum_{c} N (x_{t} \mid A_{c}(x_{t}) \phi (p_a(x_{t})), \Sigma_{c}(x_{t})) ,
  \]

  where \( (c, t) \) denotes the leaf that is selected in tree \( c \) at time \( t \).

Latent Space View

- Leaf \( \ell \) corresponds to a latent distribution with a \( \delta \)-peak at state \( \ell \).

- Latent variables are conditionally independent given the observed sequence.

Motion Completion

- We take sequences from the CMU motion capture database and apply Dynamic Forests to a motion completion task.

- We remove 31 frames in the middle of each test sequence and fill in the missing frames using Dynamic Forests trained on the same gesture.

- The estimate for frame \( x_{t} \) given observations for its parents \( p_a(x_{t}) \) can be obtained by the conditional expectation under our motion model,
  \[
  \hat{x}_{t} = E [ x_{t} \mid p_a(x_{t}) ] .
  \]