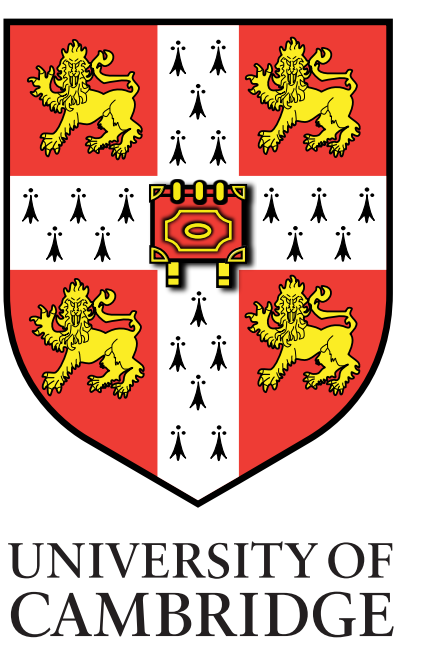




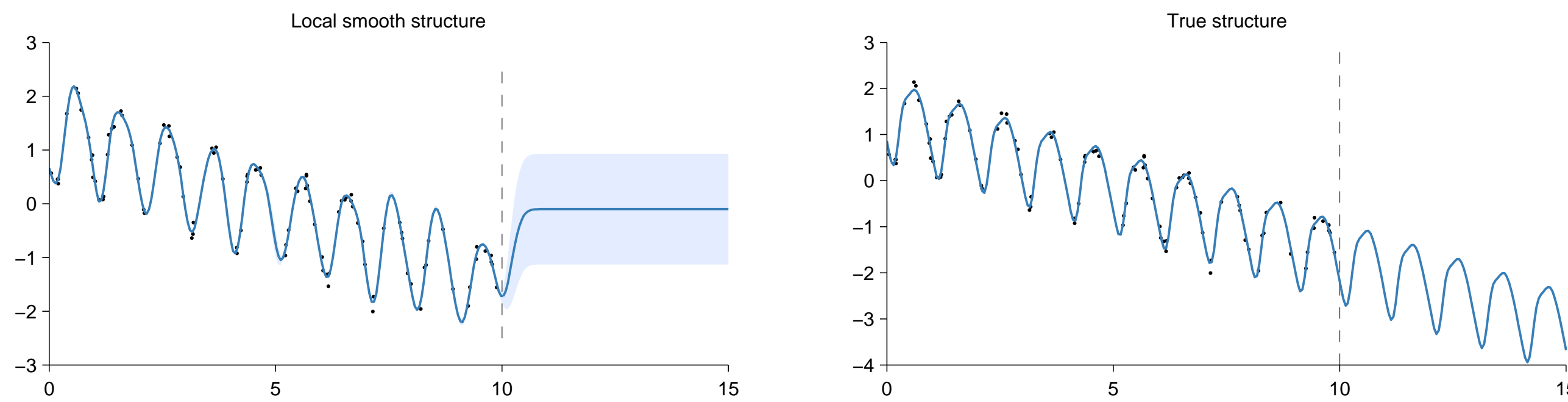
Massachusetts
Institute of
Technology

Structure Discovery in Nonparametric Regression through Compositional Kernel Search

David Duvenaud, James Robert Lloyd, Roger Grosse,
Joshua B. Tenenbaum, Zoubin Ghahramani



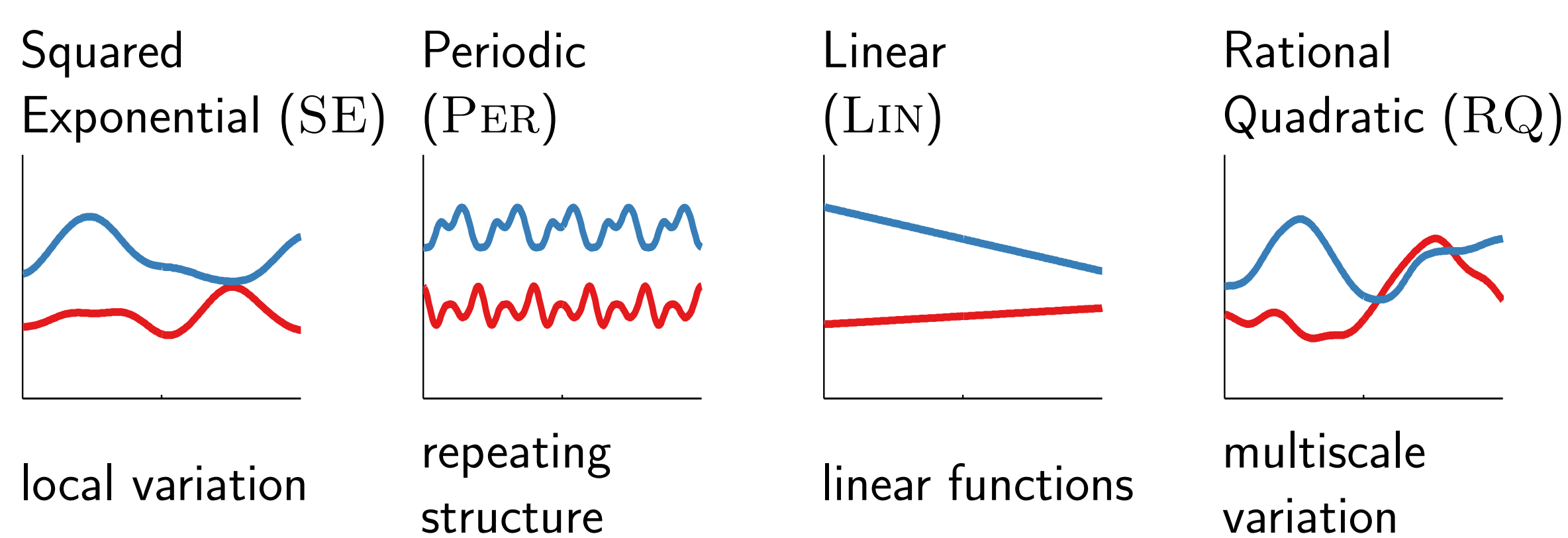
Identifying structure is crucial for extrapolation



- Traditionally, a statistician proposes an appropriate model for the type of structures present
- Automatic model selection techniques already exist, typically choosing between a finite or restricted set of models
- Instead, we automate statistical model *construction*

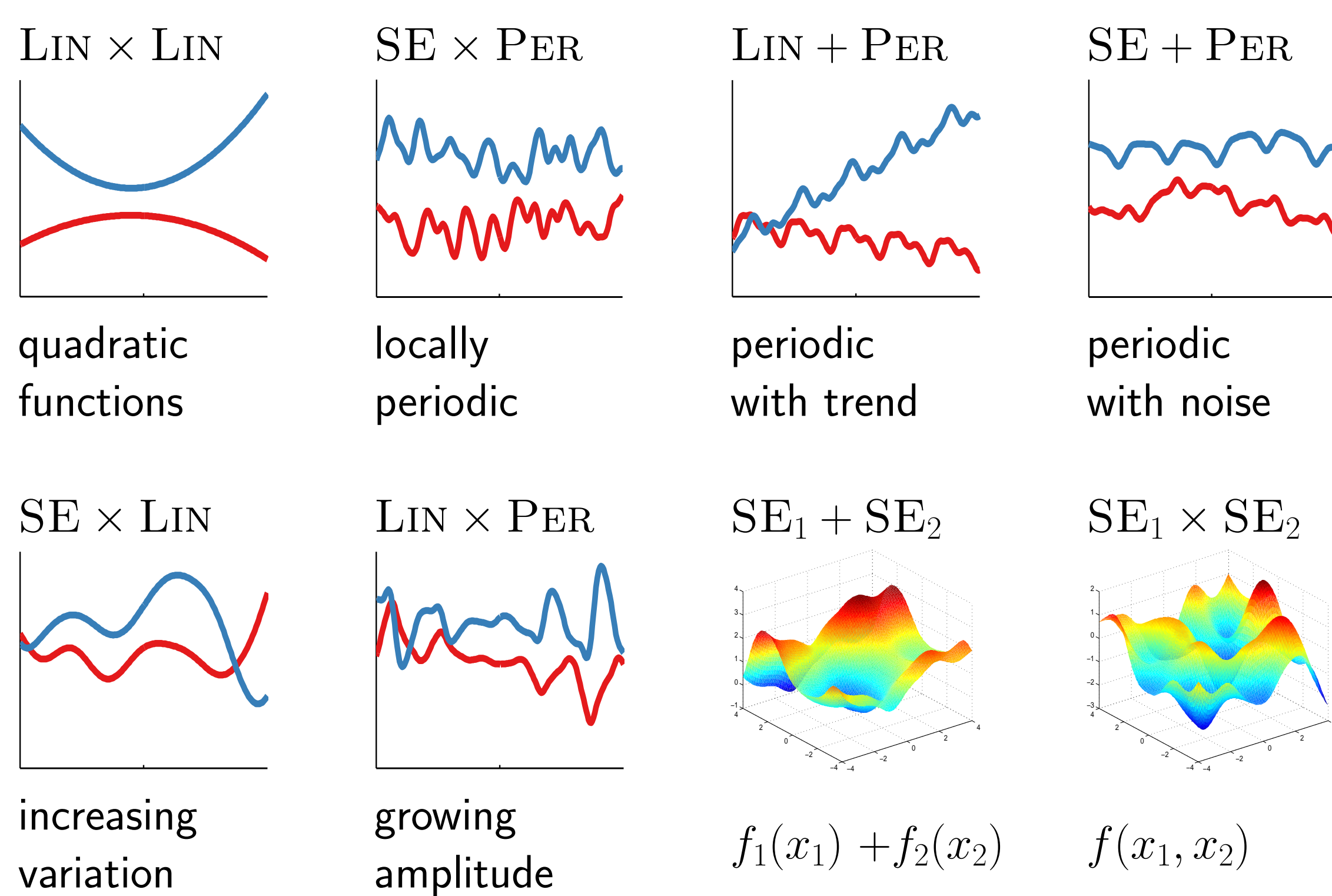
Gaussian processes model structure through kernels

- The kernel specifies which structures are likely under the GP prior, which in turn determines the generalization properties of the model.
- Below we list standard base kernels, and draws from the corresponding GP priors:



Kernels can be composed. . .

- Constructing appropriate composite kernels has previously been done by experts



. . . defining a rich, open-ended set of models. . .

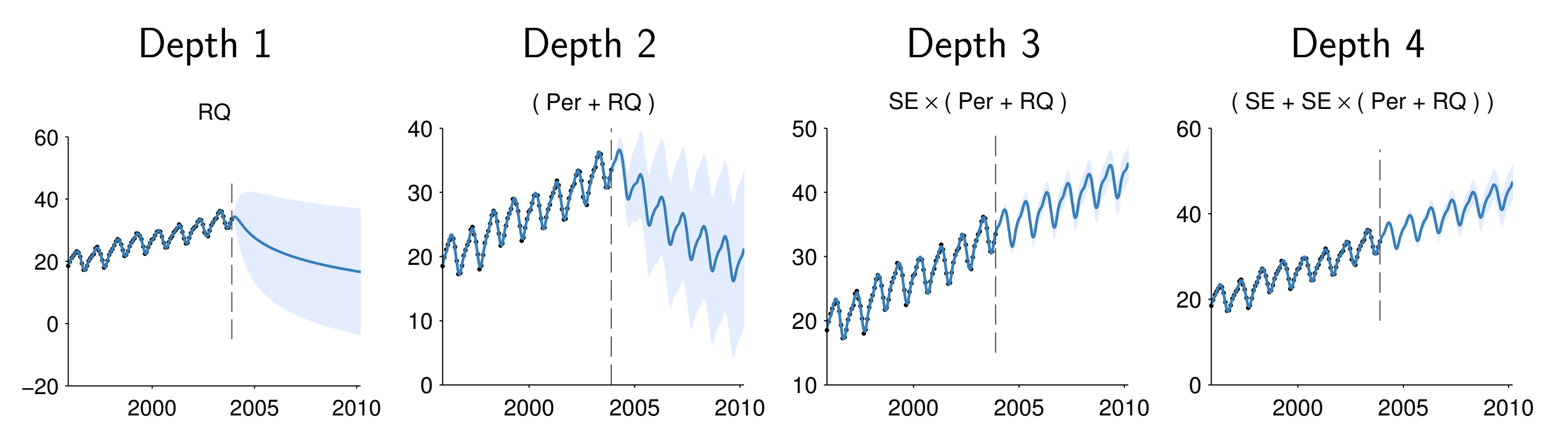
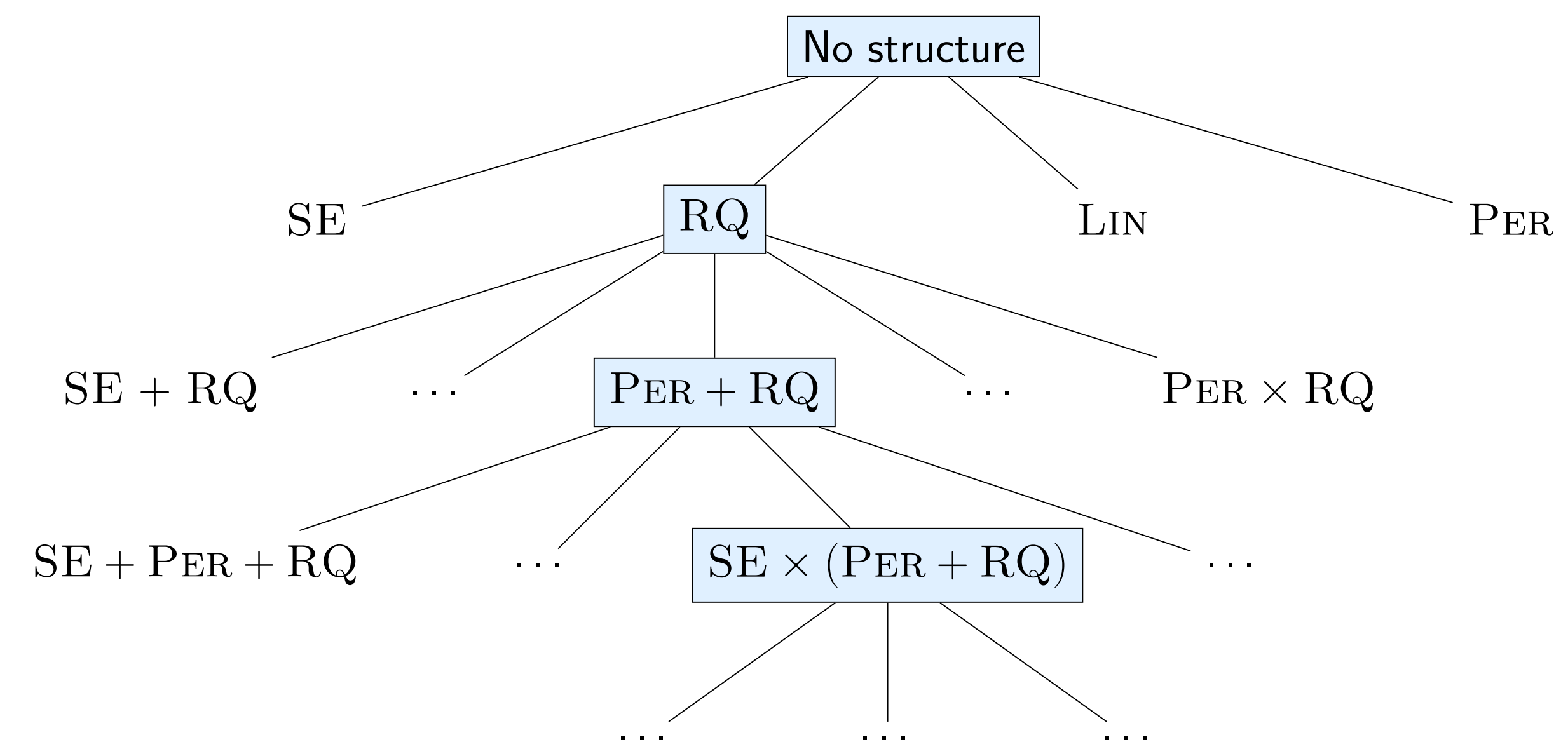
- We consider all algebraic expressions composed a small number of base kernels and the operations '+' and 'x'

Special cases of our model

Bayesian linear regression	LIN
Bayesian polynomial regression	LIN x LIN x ...
Generalized Fourier decomposition	PER + PER + ...
Generalized additive models	$\sum_{d=1}^D SE_d$
Automatic relevance determination	$\prod_{d=1}^D SE_d$
Linear trend with local deviations	LIN + SE
Linearly growing amplitude	LIN x SE

. . . which we search greedily

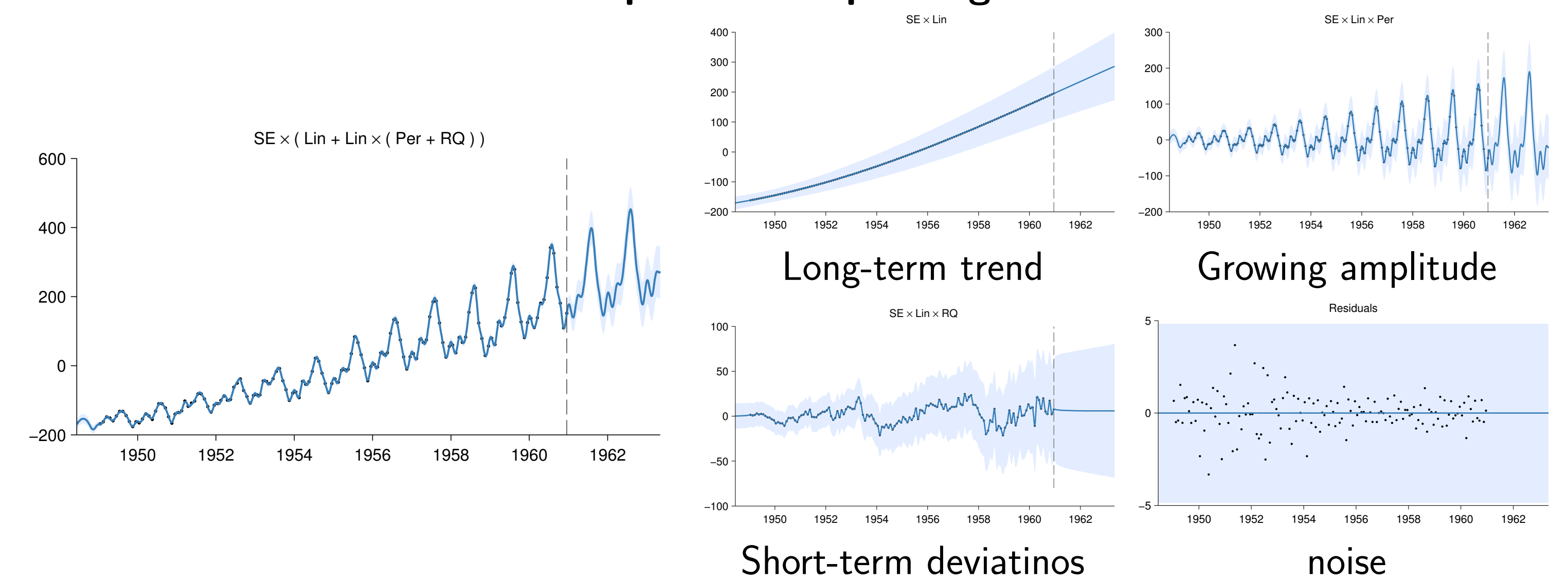
- We try all base kernels, selecting the one with the highest (approximate) marginal likelihood which balances data fit and model complexity
- The search continues by adding an extra term to the current best kernel, stopping when marginal likelihood no longer improves



Compound kernels are interpretable

- Compound kernels decompose functions into additive components

Example: Airline passengers



Example: Radio critical frequency

