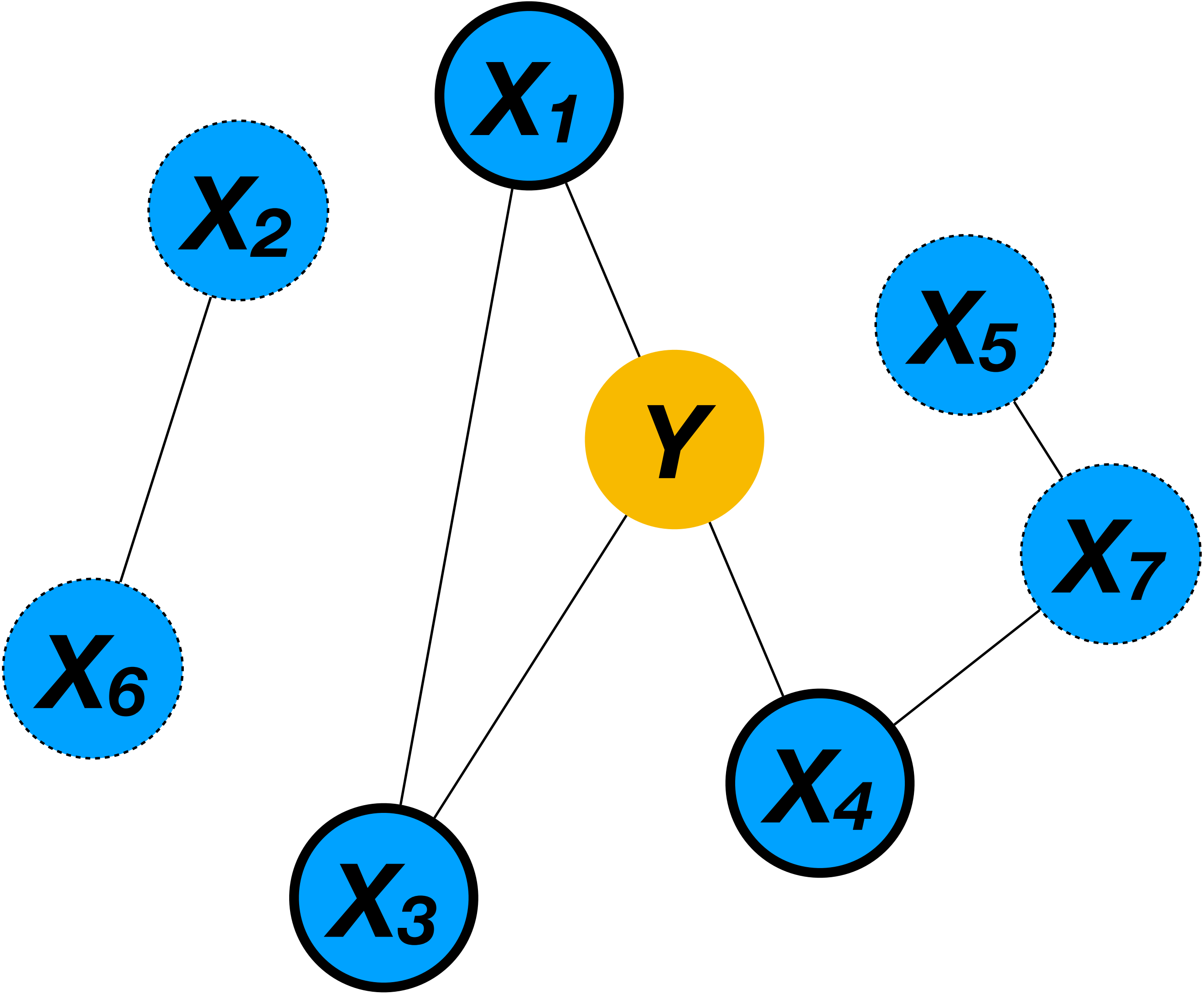


Variable selection



Let $(X, Y) \sim F_{X,Y}$ with $X \in \mathbb{R}^p$ being the features and Y being the response.

Want to identify the set of important features (conditionally dependent of Y given other features).

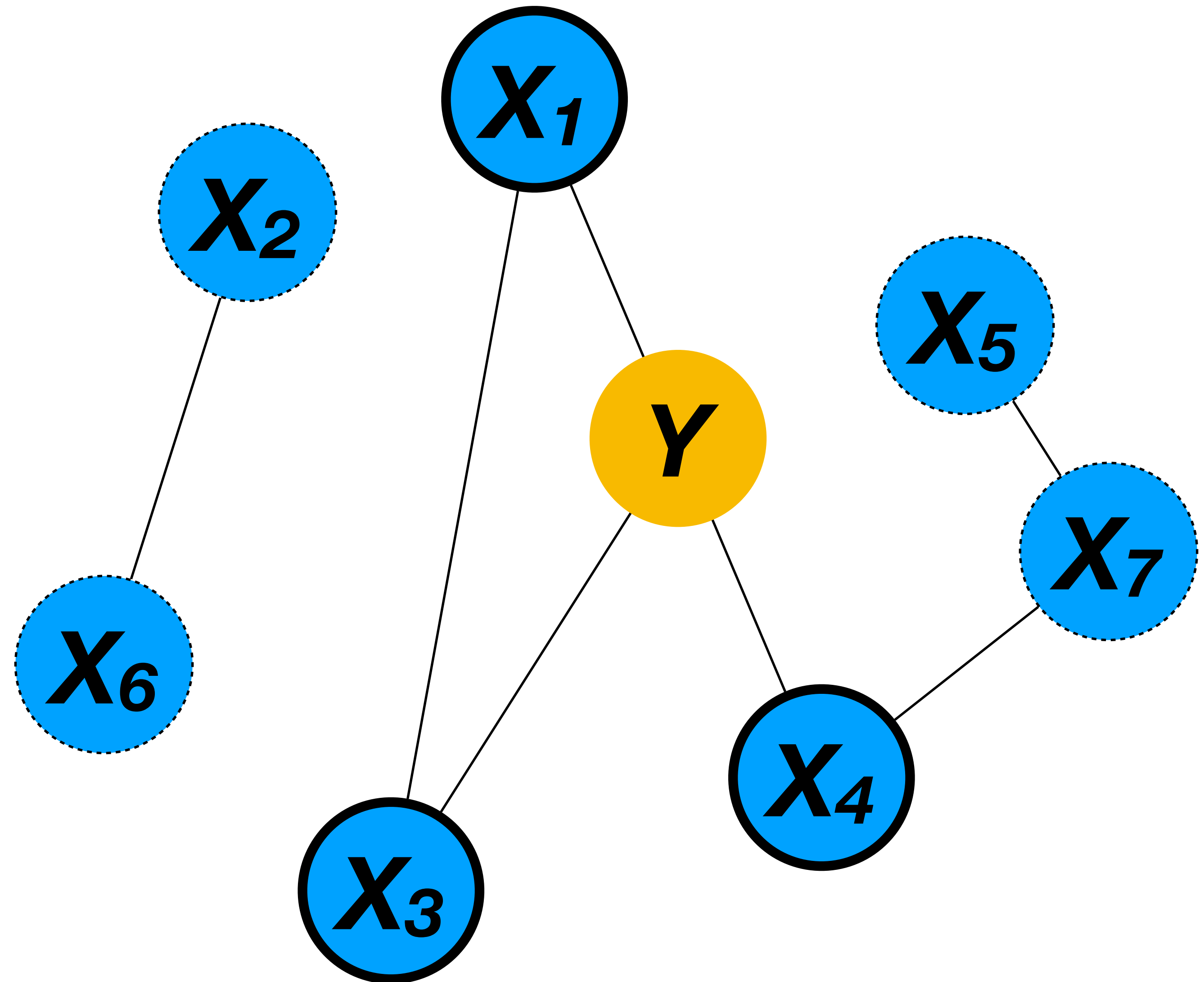
Select variables while controlling an error rate (e.g., expected fraction of false discoveries).

Variable selection

Let $(X, Y) \sim F_{X,Y}$ with $X \in \mathbb{R}^p$ being the features and Y being the response.

Want to identify the set of important features (conditionally dependent of Y given other features).

Select variables while controlling an error rate (e.g., expected fraction of false discoveries).



Model-X knockoffs