

# SEED-SET

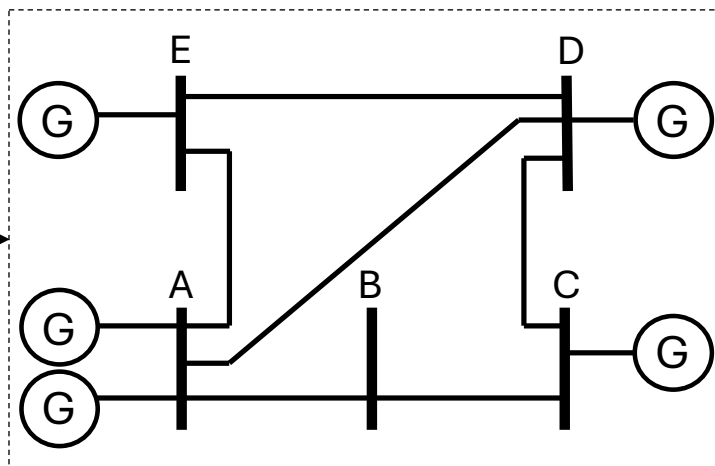
Prompt  
Design



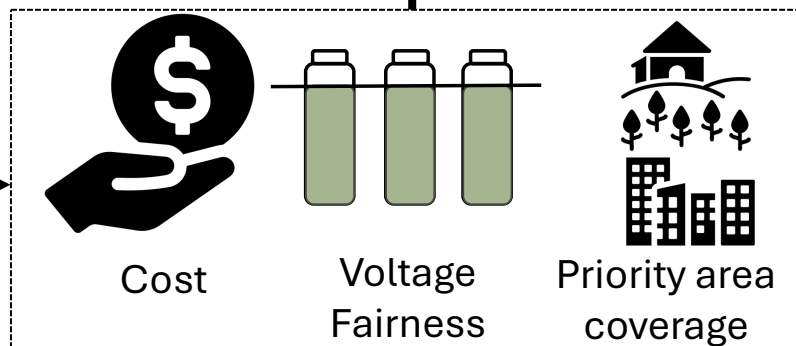
Criteria



According to the provided criteria [redacted], and the observed results: [redacted], provide your preference over the new design w.r.t. the previous one.



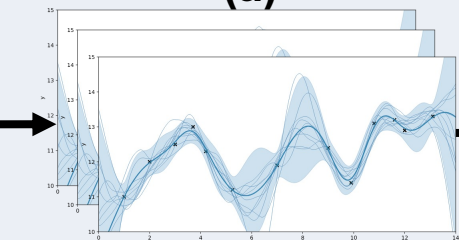
(e) System-Level Simulations



Objective evaluation  
metrics ( $y$ )

Normalization

(a)

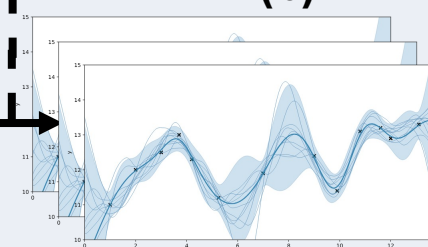


Objective GPs ( $f: x \rightarrow y$ )



Objective  
Judgement

(b)



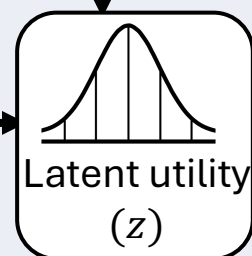
Subjective GPs ( $g: y \rightarrow z$ )



(c)

LLM Evaluations

$y_1 > y_2$

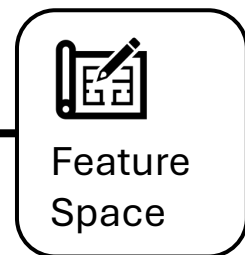


Latent utility  
( $z$ )

(d) New Design Acquisition



Bayesian Experimental Design



Feature  
Space

Normalization

( $x$ )