Leveraging Discrete Function Decomposability for Scientific Design

James Bowden, Sergey Levine, Jennifer Listgarten

← blog post

full paper →

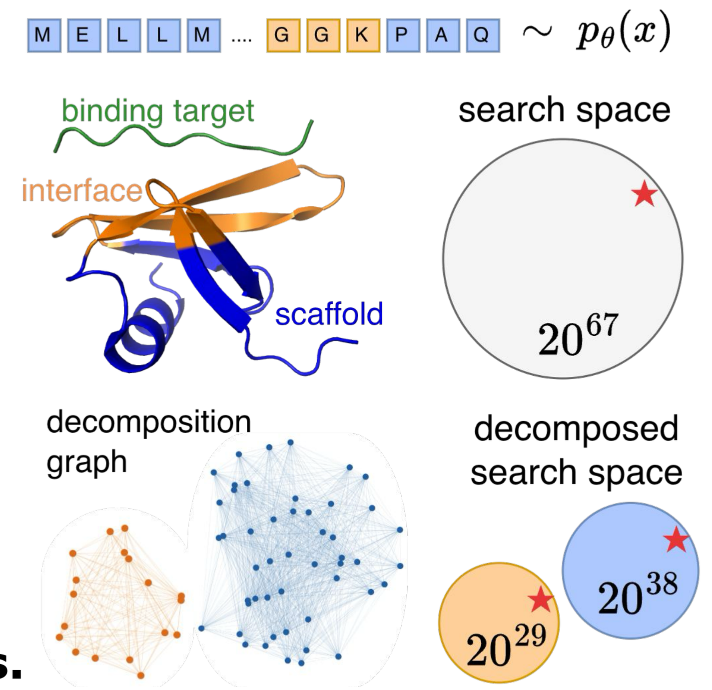
1. What does it mean to decompose a design problem?

GOAL: Train generative model to sample good designs, x, according to objective, $f(x)$: $\arg\max_\theta \mathbb{E}_{p_\theta(x)}[f(x)]$

INSIGHT: When designing proteins, scientists often assume that components can be designed separately. Current generative models don't account for this.

Operating in the much smaller decomposed search space means better designs can be found more quickly.

In example (right), leveraging the decomposition is easy b/c just separate optimizations. **We generalize for arbitrary topologies to enable leveraging realistic decompositions.**



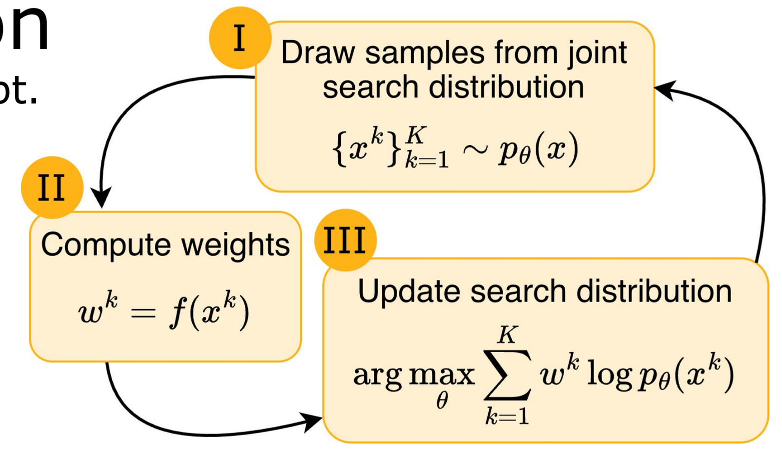
2. Primer: distributional optimization

a.k.a. estimation of distribution algorithm (EDA), RL policy opt.

GOAL: $\arg\max_\theta \mathbb{E}_{p_\theta(x)}[f(x)]$

INTUITION: Use generative model search distribution $p_\theta(x)$ to find high-$f(x)$ designs, via weighted ML

How can we take advantage of a decomposition?



3. Primer: exact message-passing optimization

GOAL: $\arg\max_x f(x)$

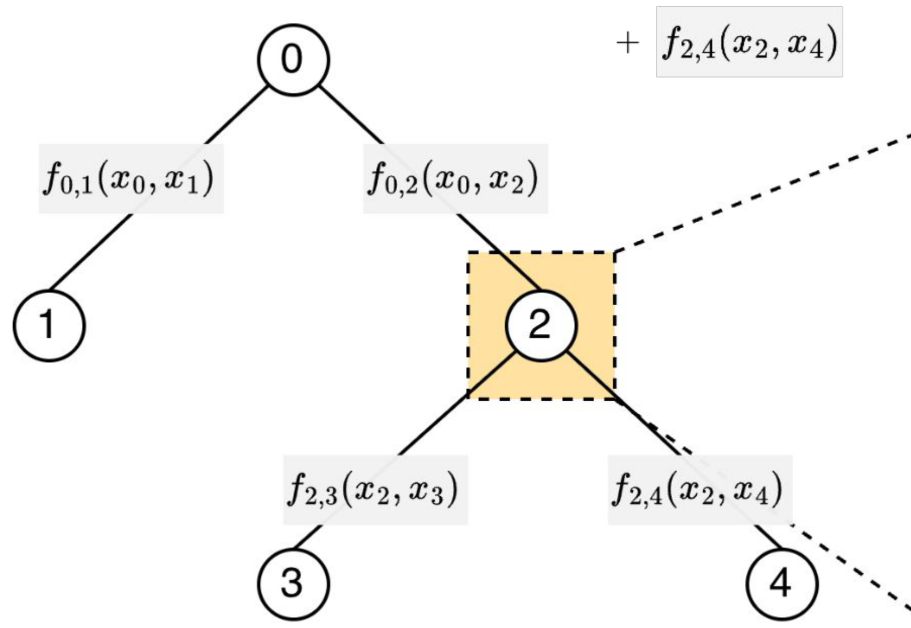
$f(x) = f_{0,1}(x_0,x_1) + f_{0,2}(x_0,x_2) + f_{2,3}(x_2,x_3) + f_{2,4}(x_2,x_4)$

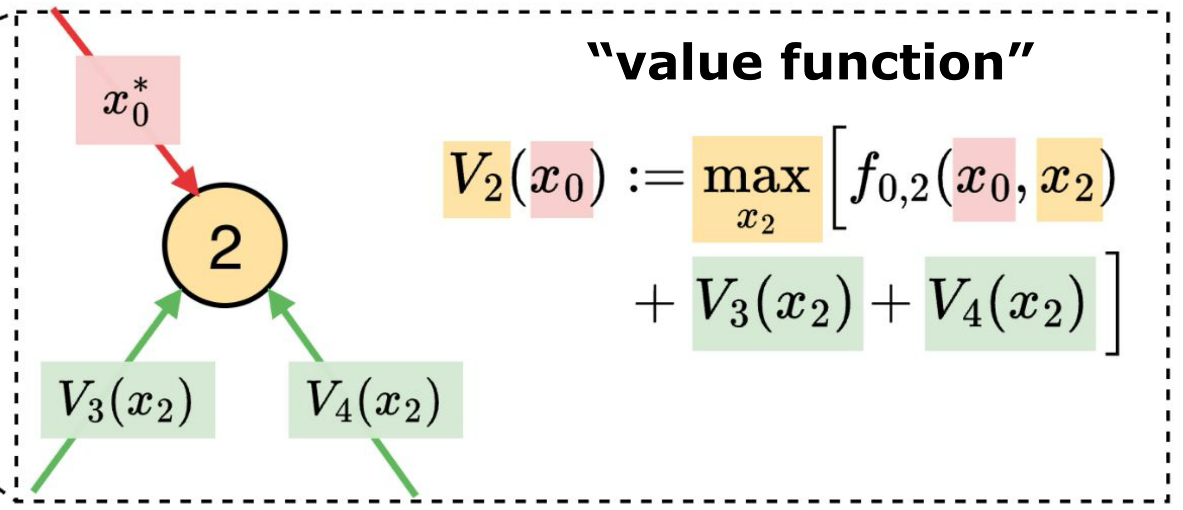
INTUITION:

Given a tree-decomposed function,

a) accumulate intermediate $f(x)$ at each node

b) solve smaller per-edge opt. problems.





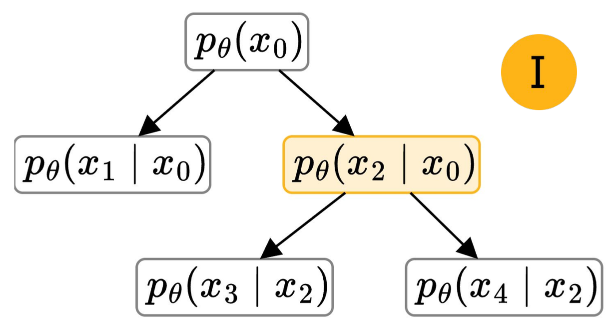
4. Our method: Decomposition-Aware Distributional Optimization (**DADO**)

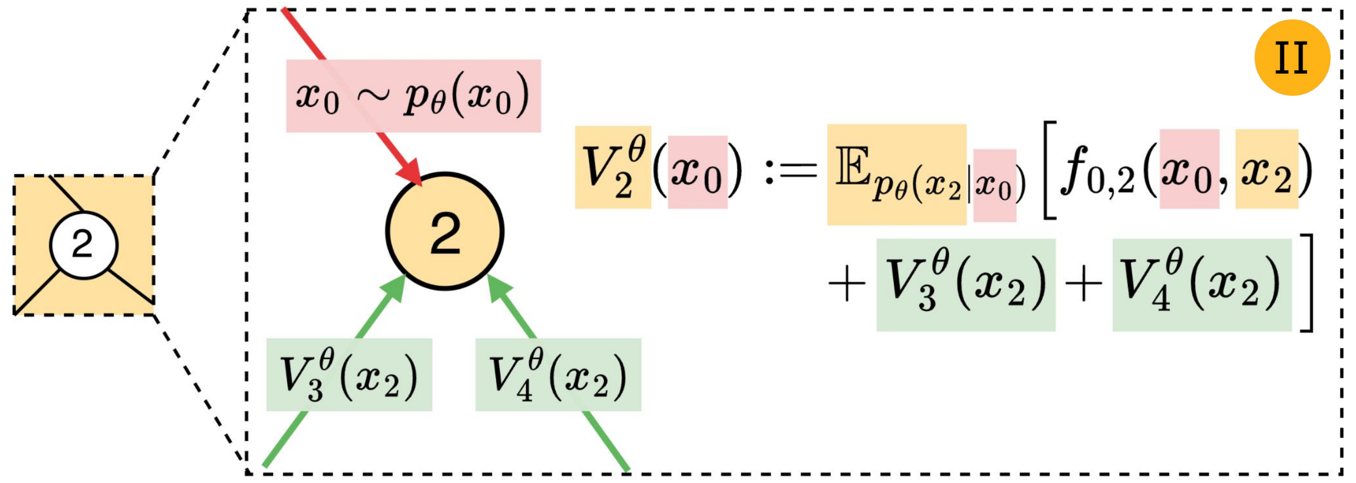
What inspiration can we draw from exact message-passing?

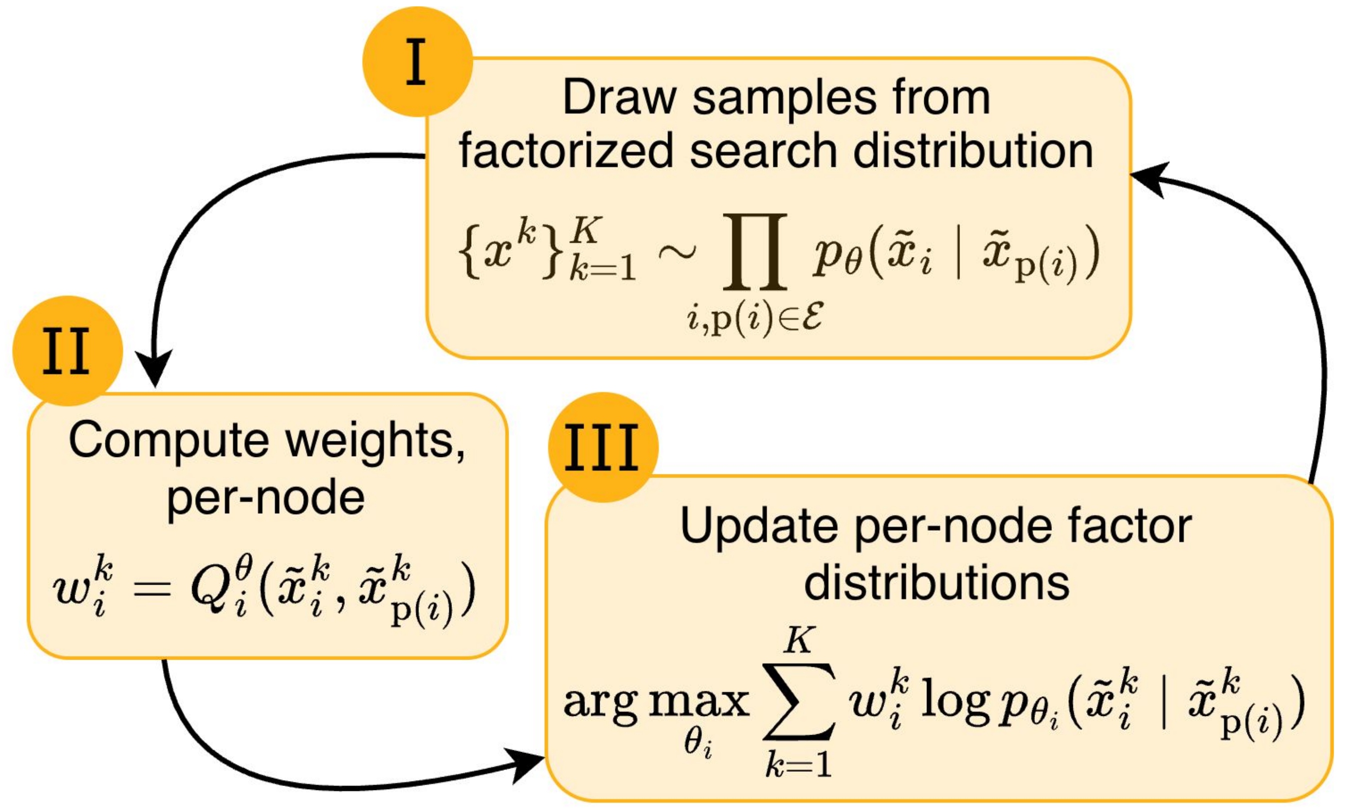
- Given a decomposition, can compute a **value function** for each node.
- Then, choosing each node's state using **only** its parent's chosen state and its value function results in a **coordinated decentralized** optimization.

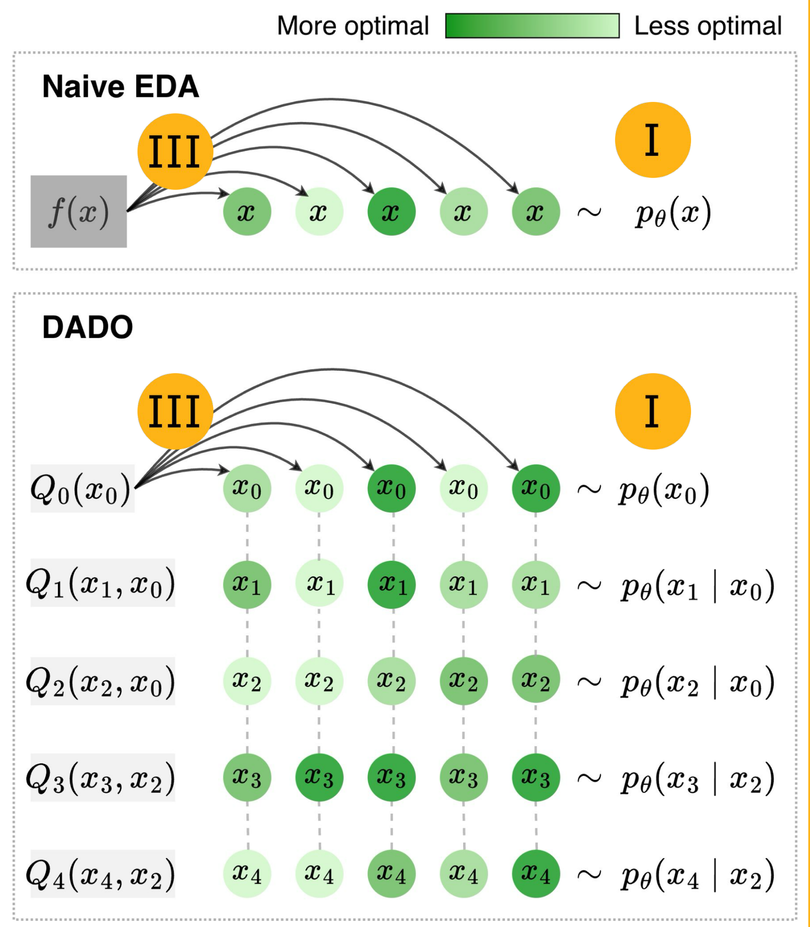
To infuse decomposition-awareness into distributional optimization:

- **Factorize the generative model** according to decomposition, yielding per-node factors. These only "see" the smaller decomposed space.
- Compute a **distributional value function** for each node in expectation under its factor.







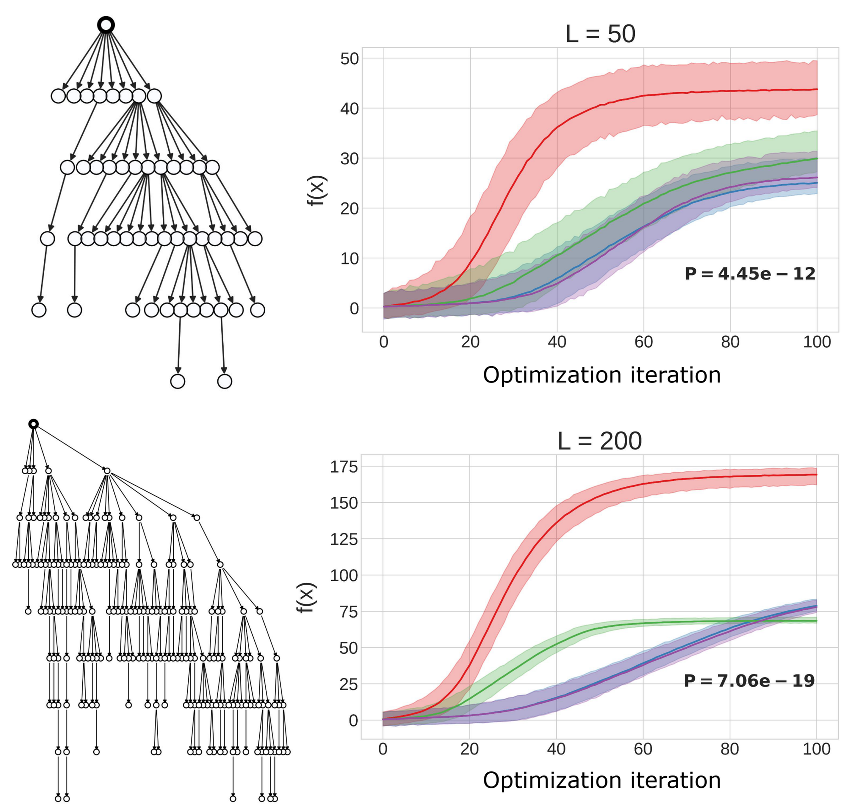


Using per-node value functions as weights allows separate, per-factor updates which are **more statistically efficient.**

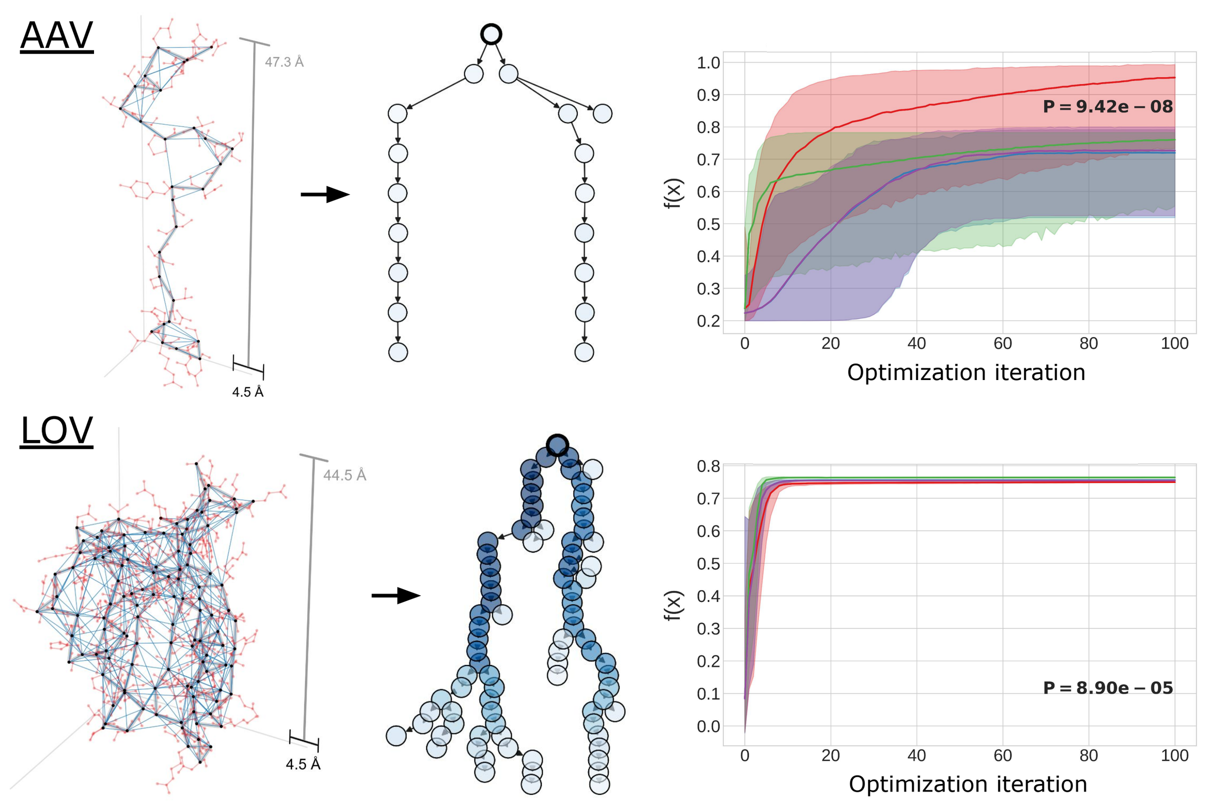
5. Experimental results

Legend: DADO EDA PPO FDA

SYNTHETIC OBJECTIVES with random tree decompositions (D=20):



REALISTIC PROTEIN DESIGN OBJECTIVES with decompositions from AlphaFold contact graphs:



Potential efficiency gain will in large part depend on how sparse the decomposition graph is.

6. Summary

- If your objective function has any decomposability, using it can unlock much more efficient design!
- Settings that may be amenable: proteins, multi-layered insulators, telescopes, circuits, mechanical systems, crystal lattices, …
- DADO optimizes functions characterized by **any** decomposition graph topology by synthesizing message-passing + distributional optimization

7. Future directions

- Not straightforward to find a useful decomposition for each design problem
  - Open area of research; may benefit from domain-specific approaches
  - Existing approaches infer from labeled data or auxiliary information
- Extensions of DADO for generative models w/o access to likelihoods
- Using a pre-trained model for initialization or regularization with DADO
- DADO for optimizing objectives on continuous design spaces
- Lower-variance value function estimates