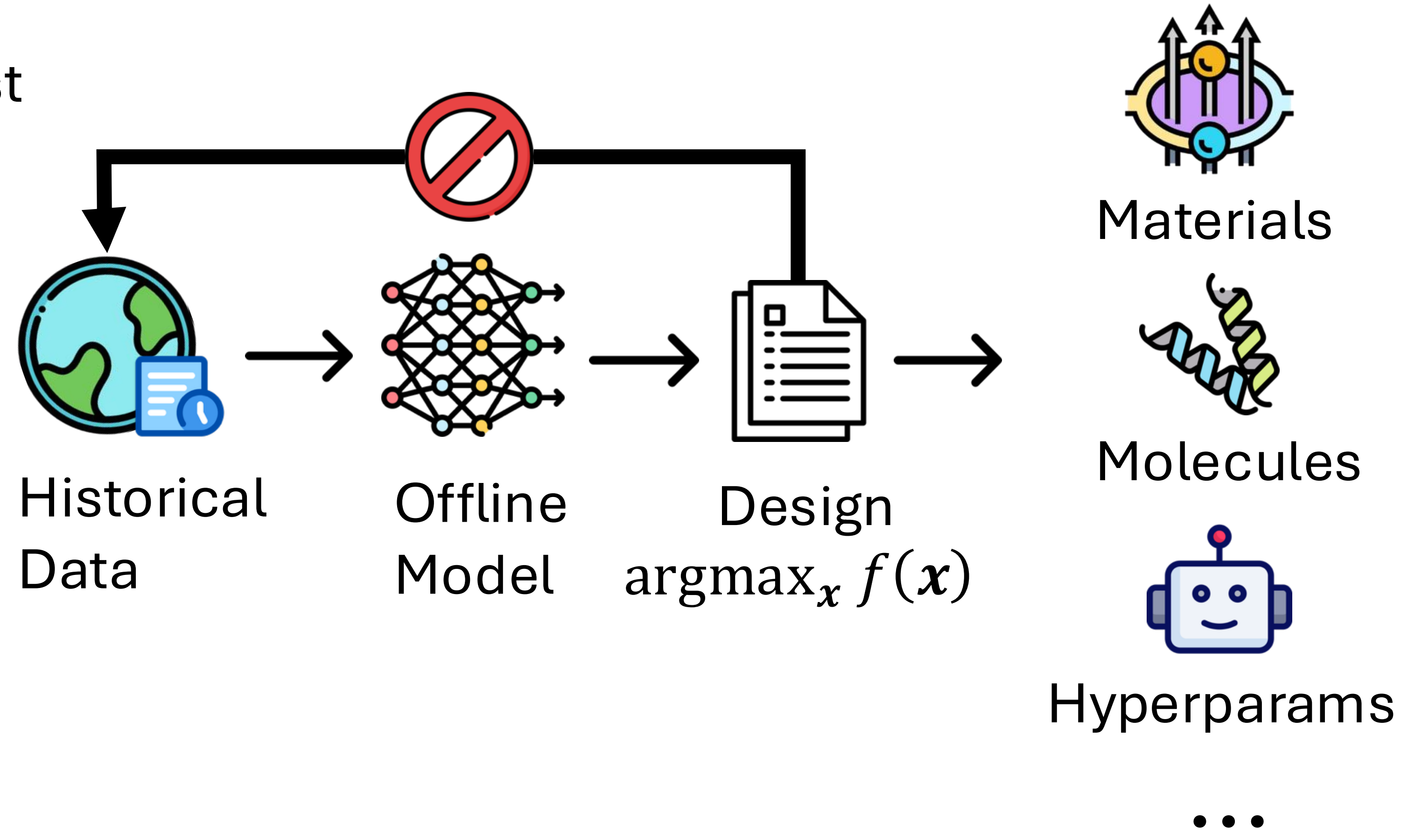
Support-Proximity Augmented Diffusion Estimation for Offline Black-Box Optimization

Yonghan Yang* [1], Ye Yuan* [1,2,3], Zipeng Sun [2,3], Linfeng Du [2,3],
Bowei He [1], Haolun Wu [2,3], Can Chen [3,4], Xue Liu [1,2,3]

[1] MBZUAI [2] McGill University [3] Mila - Quebec AI Institute [4] Amazon AGI
* equal contribution

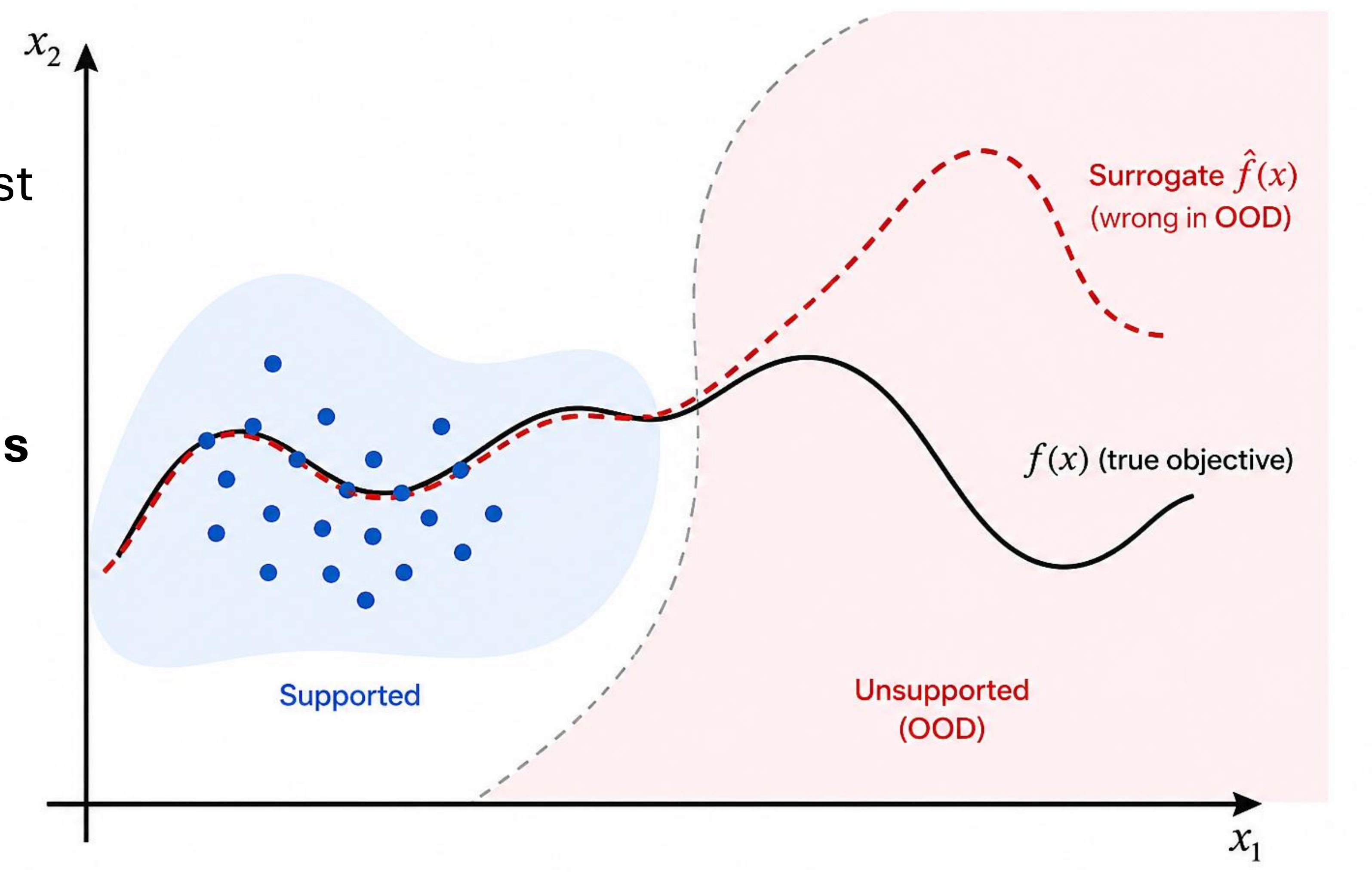
Background | **Method** | **Experiments**

- **Offline black-box optimization** searches best design with fixed dataset and no new evaluations.



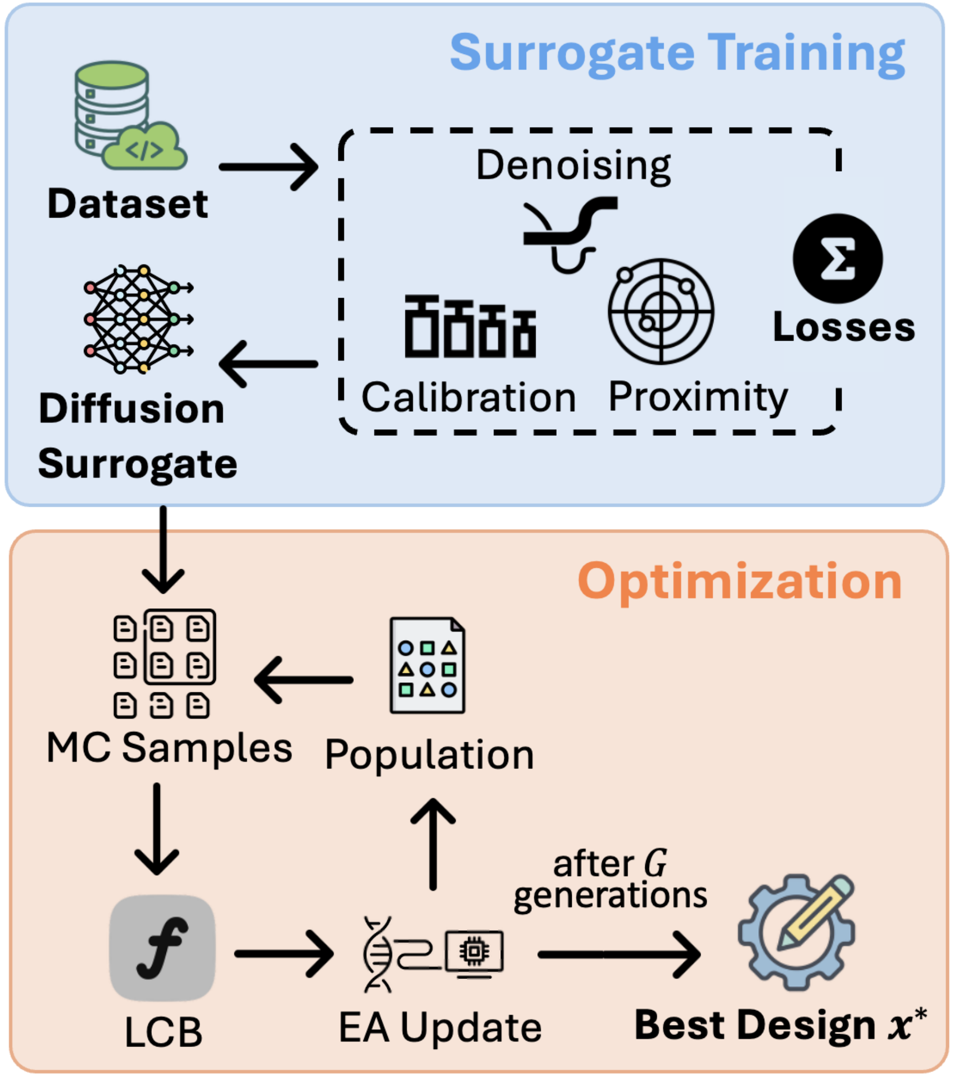
Background Method Experiments

- **Offline black-box optimization** searches best design with fixed dataset and no new evaluations.
- Optimizer **exploits surrogate overestimations** far from data manifold.



Background **Method** Experiments

- SPADE trains a **calibrated, support-aware diffusion surrogate** on offline design–score data.
- It then uses **MC sampling**, **LCB scoring**, and **evolutionary updates** to find the best design.



$$p(\boldsymbol{x} \mid \boldsymbol{y}) \propto p(\boldsymbol{y} \mid \boldsymbol{x})\, p(\boldsymbol{x})$$

SPADE is

- diffusion surrogate for likelihood $p(\boldsymbol{y} \mid \boldsymbol{x})$
- support prior $p(\boldsymbol{x})$ via kNN proximity

$$\tilde{A}(x) \approx A(\mu, \sigma) + \kappa \log \hat{p}_{knn}(x)$$

Acquisition ≈ predicted utility + data-support prior

Background | Method | **Experiments**

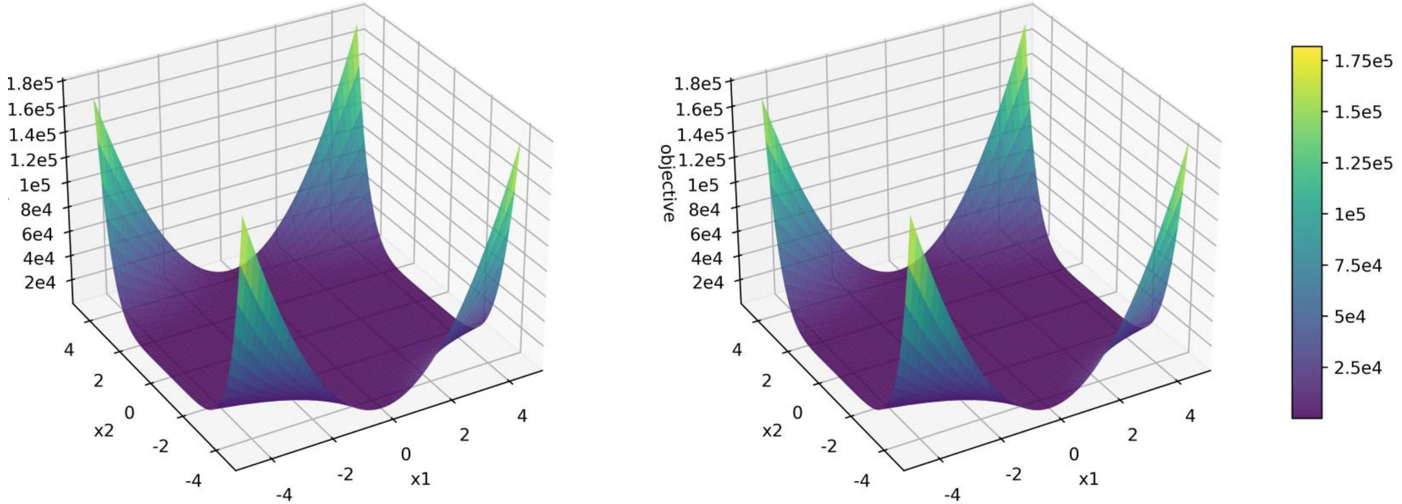
- **SOTA on benchmarks**: mean & median ranks are 2.8 and 1.5 out of all 24 methods

Method	SuperC	Ant	D'Kitty	LLM-DM	TF8	TF10	Mean Rank	Median Rank
$\mathcal{D}$(best)	0.399	0.565	0.884	1.000	0.439	0.511	–	–
Standard Optimization Baselines								
CMA-ES	0.465 ± 0.024	$\mathbf{1.561 \pm 0.896}$	0.724 ± 0.001	0.975 ± 0.016	0.939 ± 0.039	0.692 ± 0.013	13.3/24	13.5/24
REINFORCE	0.481 ± 0.013	0.263 ± 0.026	0.573 ± 0.204	0.305 ± 0.027	0.961 ± 0.034	0.632 ± 0.012	19.8/24	23.0/24
BO-qEI	0.402 ± 0.034	0.812 ± 0.000	0.896 ± 0.000	0.953 ± 0.022	0.825 ± 0.091	0.663 ± 0.011	18.7/24	19.5/24
Forward Surrogate Methods								
Standard GA	0.505 ± 0.013	0.293 ± 0.029	0.860 ± 0.021	0.998 ± 0.000	0.923 ± 0.011	0.732 ± 0.041	14.0/24	14.5/24
GA on GP	0.499 ± 0.019	0.948 ± 0.013	0.946 ± 0.001	0.846 ± 0.029	0.770 ± 0.087	0.599 ± 0.004	16.0/24	17.0/24
MC-Dropout	0.535 ± 0.064	0.805 ± 0.021	0.934 ± 0.004	0.781 ± 0.251	0.911 ± 0.042	0.817 ± 0.031	13.0/24	13.0/24
COMs	0.481 ± 0.028	0.878 ± 0.031	0.929 ± 0.016	0.815 ± 0.008	0.937 ± 0.025	0.755 ± 0.017	15.3/24	15.5/24
RoMA	0.509 ± 0.015	0.592 ± 0.059	0.825 ± 0.016	0.968 ± 0.011	0.662 ± 0.000	0.801 ± 0.000	16.2/24	16.0/24
ICT	0.503 ± 0.017	0.911 ± 0.030	0.945 ± 0.011	0.839 ± 0.021	0.888 ± 0.047	0.814 ± 0.027	14.0/24	13.5/24
Tri-mentoring	0.514 ± 0.018	0.944 ± 0.033	0.950 ± 0.015	0.750 ± 0.002	0.899 ± 0.045	0.811 ± 0.039	12.2/24	9.5/24
BDI	0.513 ± 0.000	0.964 ± 0.000	0.941 ± 0.000	0.988 ± 0.025	0.973 ± 0.000	$\underline{0.882 \pm 0.000}$	5.7/24	4.0/24
LTR	0.514 ± 0.022	0.904 ± 0.036	0.958 ± 0.012	0.982 ± 0.041	0.973 ± 0.010	0.849 ± 0.023	6.2/24	5.0/24
MATCH-OPT	0.504 ± 0.021	0.931 ± 0.011	0.957 ± 0.014	0.850 ± 0.033	0.977 ± 0.004	0.824 ± 0.008	9.2/24	8.5/24
PGS	$\mathbf{0.563 \pm 0.058}$	0.949 ± 0.017	0.966 ± 0.013	0.595 ± 0.027	$\underline{0.981 \pm 0.015}$	0.793 ± 0.021	8.0/24	4.0/24
Inverse Generative Methods								
CbAS	0.503 ± 0.069	0.846 ± 0.033	0.895 ± 0.016	0.919 ± 0.043	0.903 ± 0.028	0.652 ± 0.032	16.2/24	16.0/24
MINs	0.499 ± 0.017	0.894 ± 0.022	0.939 ± 0.004	0.982 ± 0.007	0.908 ± 0.063	0.647 ± 0.012	14.0/24	14.0/24
DDOM	0.481 ± 0.015	0.926 ± 0.025	0.923 ± 0.009	0.983 ± 0.028	0.884 ± 0.042	0.708 ± 0.015	14.7/24	17.0/24
GABO	0.508 ± 0.007	0.224 ± 0.051	0.719 ± 0.001	0.975 ± 0.019	0.939 ± 0.038	0.739 ± 0.009	15.0/24	13.5/24
GTG	0.480 ± 0.055	0.865 ± 0.040	0.935 ± 0.010	0.910 ± 0.030	0.901 ± 0.039	0.801 ± 0.004	15.0/24	14.0/24
RGD	0.515 ± 0.011	0.922 ± 0.020	0.883 ± 0.014	$\underline{1.004 \pm 0.008}$	0.889 ± 0.068	0.825 ± 0.039	10.7/24	9.5/24
BONET	0.434 ± 0.021	0.948 ± 0.025	0.955 ± 0.010	0.874 ± 0.039	0.894 ± 0.086	0.796 ± 0.018	13.3/24	14.5/24
DEMO	0.520 ± 0.012	0.948 ± 0.013	0.954 ± 0.013	0.907 ± 0.005	0.808 ± 0.044	0.836 ± 0.042	9.7/24	6.5/24
ROOT	0.525 ± 0.012	0.965 ± 0.014	$\underline{0.972 \pm 0.005}$	0.905 ± 0.006	$\mathbf{0.986 \pm 0.007}$	0.833 ± 0.046	$\underline{5.0/24}$	$\underline{3.5/24}$
SPADE (Ours)	$\underline{0.546 \pm 0.013}$	$\underline{0.978 \pm 0.006}$	$\mathbf{0.981 \pm 0.003}$	$\mathbf{1.019 \pm 0.064}$	0.923 ± 0.015	$\mathbf{0.915 \pm 0.010}$	**2.8/24**	**1.5/24**

Background | Method | **Experiments**

- **SOTA on benchmarks**: mean & median ranks are 2.8 and 1.5 out of all 24 methods
- **Ablation validates regularization design:** both components are necessary
- **Observation on synthetic functions**: SPADE recovers global structure.

Task	Base	w/o Prox	w/o Calib	Full
SuperC	0.519 ± 0.012	0.538 ± 0.011	0.542 ± 0.010	$\mathbf{0.546 \pm 0.013}$
Ant	0.932 ± 0.011	0.952 ± 0.008	0.963 ± 0.007	$\mathbf{0.978 \pm 0.006}$
D'Kitty	0.962 ± 0.006	0.972 ± 0.005	0.975 ± 0.004	$\mathbf{0.981 \pm 0.003}$
LLM-DM	0.957 ± 0.070	0.979 ± 0.060	0.998 ± 0.058	$\mathbf{1.019 \pm 0.064}$
TF8	0.890 ± 0.018	0.912 ± 0.016	0.897 ± 0.017	$\mathbf{0.923 \pm 0.015}$
TF10	0.870 ± 0.014	0.895 ± 0.012	0.882 ± 0.012	$\mathbf{0.915 \pm 0.010}$



Thanks!

arXiv:
2605.11246

Paper

Code