Data-driven decision making for autonomous materials synthesis

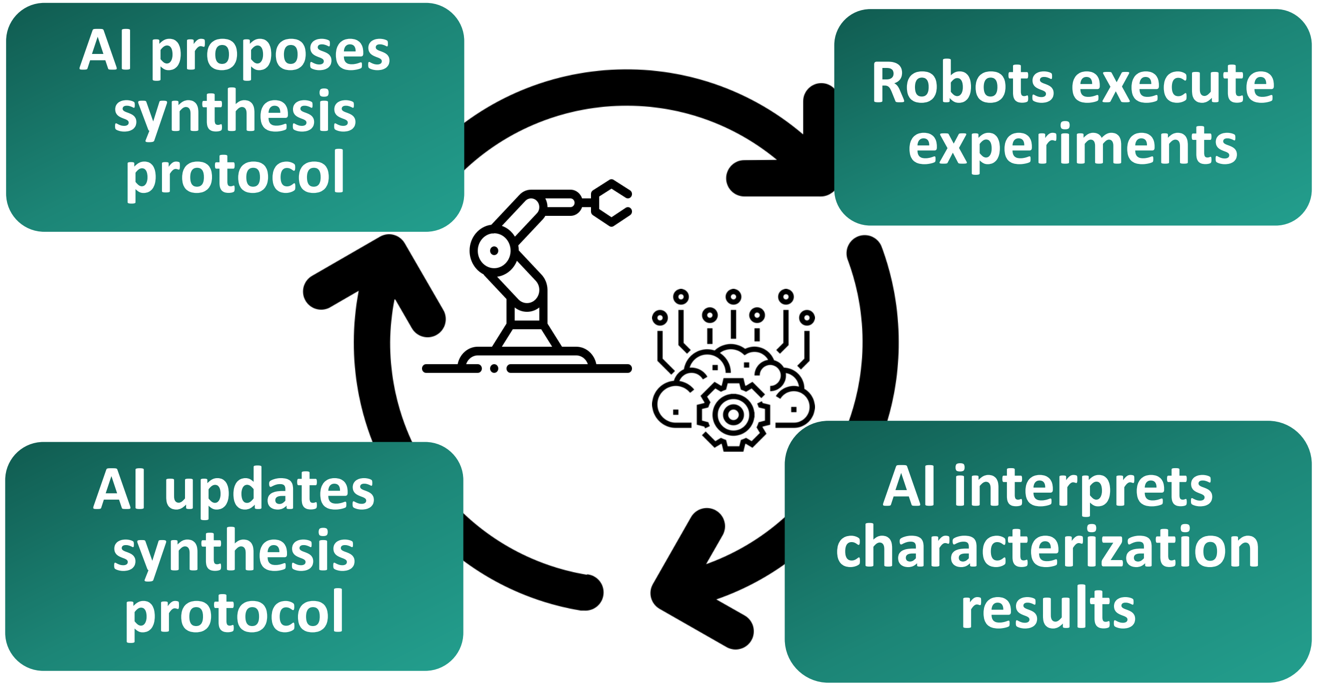
Nathan Szymanski
Ceder group at UC Berkeley
May 11, 2022

Spring MRS symposium DS01: Integrating ML and simulation for materials modeling, design, and manufacturing

Automated **solid-state synthesis** for inorganic materials

Given a target phase, how do we synthesize it?

AI-driven experiments



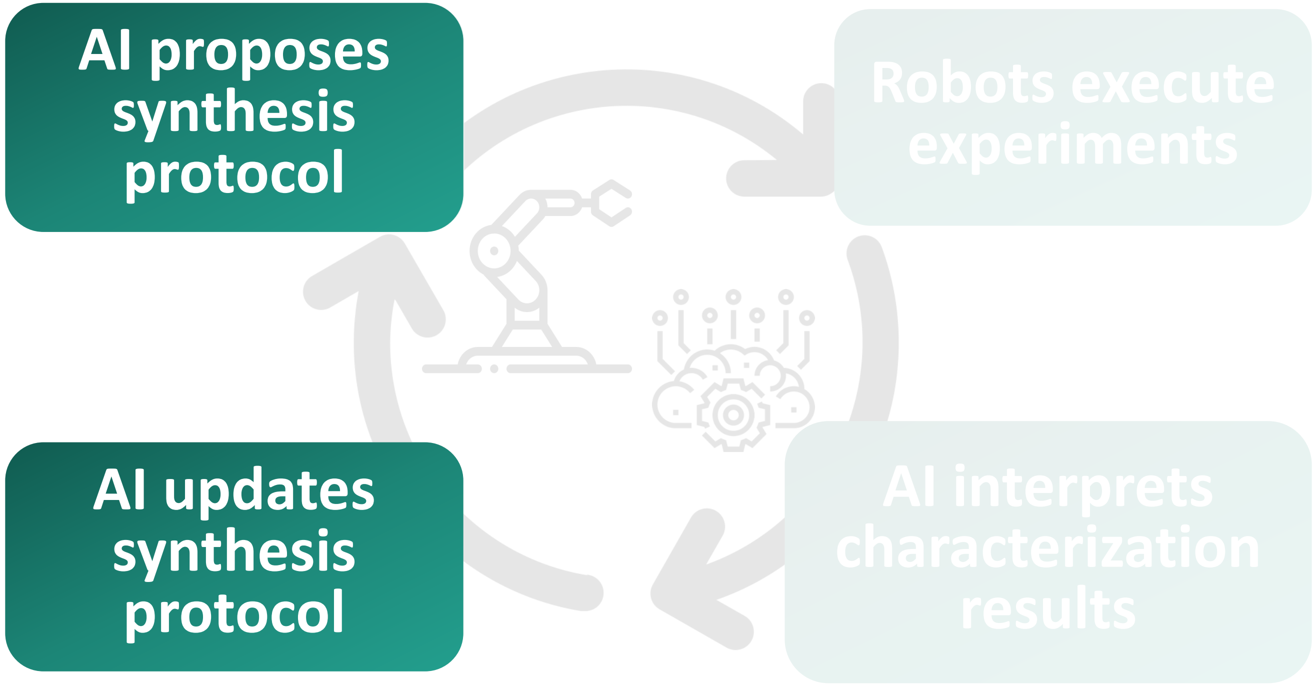
N. J. Szymanski, …, G. Ceder, Materials Horizons 2021.

Automated **solid-state synthesis** for inorganic materials

AI-driven experiments

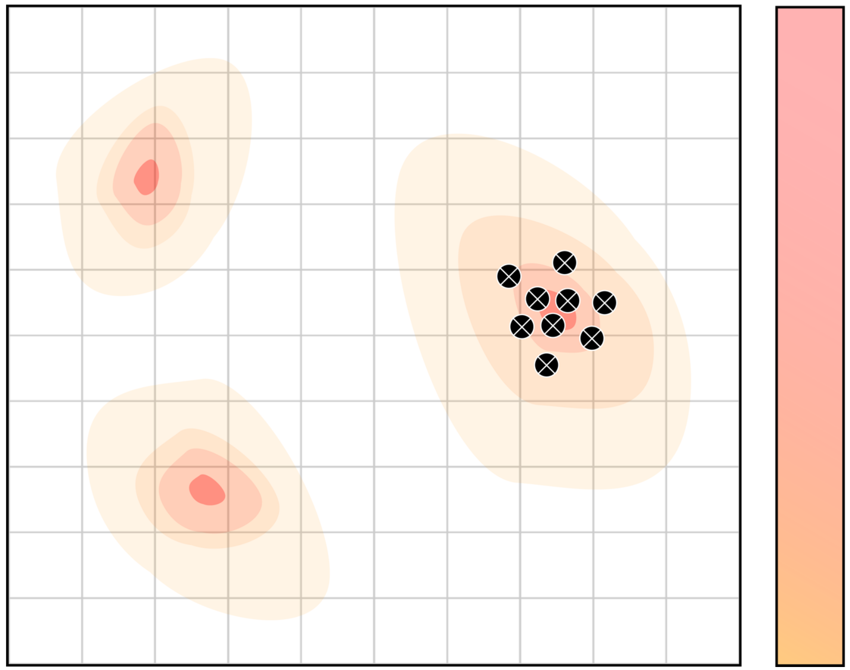
Decision making goal:

Identify precursors and conditions that lead to **maximal yield** of the target phase using a **minimal number of experiments**



N. J. Szymanski, …, G. Ceder, Materials Horizons 2021.

To develop and test an algorithm, we need more **data**



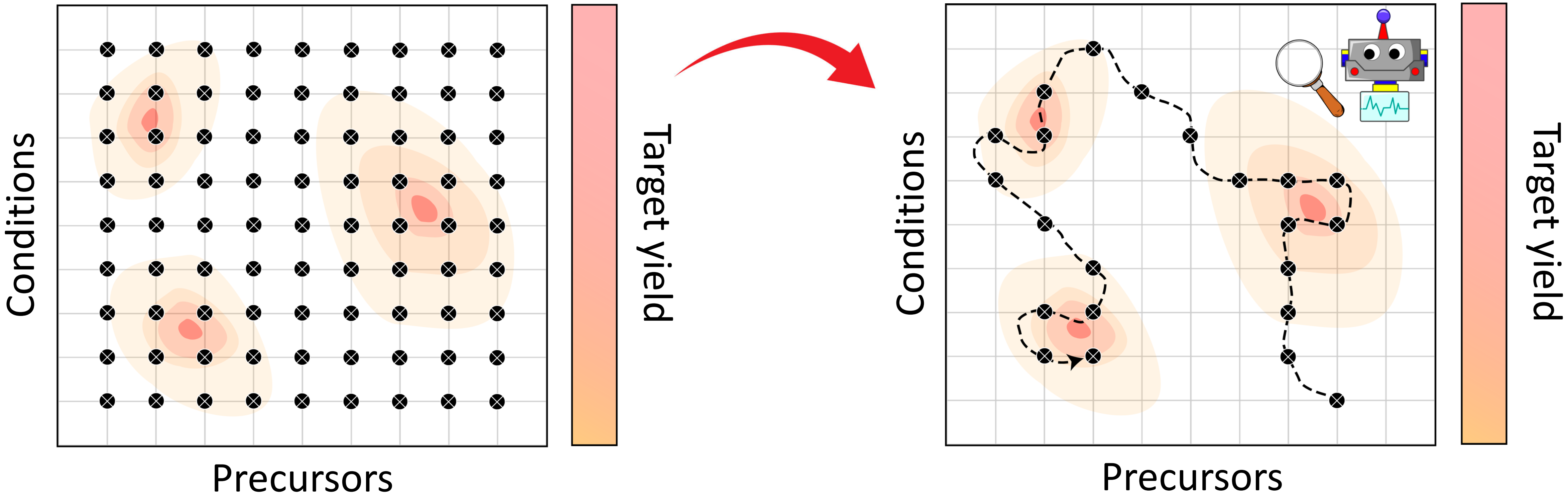
Published synthesis data:

- **Sparse** sampling of design space
- **Bias** toward positive results
- **Repetition** of similar procedures

Synthesis data from the literature

O. Kononova, …, G. Ceder, Scientific Data 2019.

To develop and test an algorithm, we need more **data**

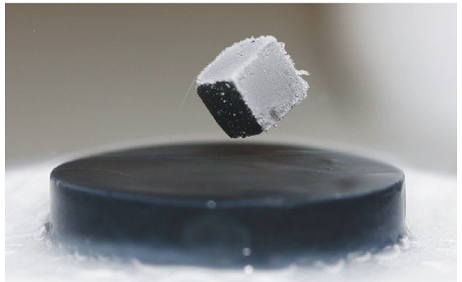


Completely populate the space by testing different precursors and conditions

Dataset acts as a **surrogate model** for the decision-making algorithm to query as it explores the space

A comprehensive experimental synthesis dataset

Target: YBCO ($YBa_2Cu_3O_{6+x}$) → $T_c \approx 91$ K superconductor



188 solid-state synthesis experiments

Temperatures (×4)

600 °C 700 °C

800 °C 900 °C

Hold time (×1)

4 hours

Precursors (×11)

Y_2O_3 $Y_2C_3O_9$ BaO

$BaCO_3$ $BaCuO_2$ BaO_2

$Ba_2Cu_3O_6$ $Y_2Cu_2O_5$

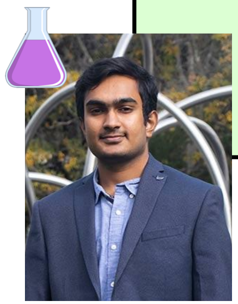
$CuCO_3$ Cu_2O CuO

×47 Precursor Sets

10 experiments (out of 188) led to phase pure YBCO

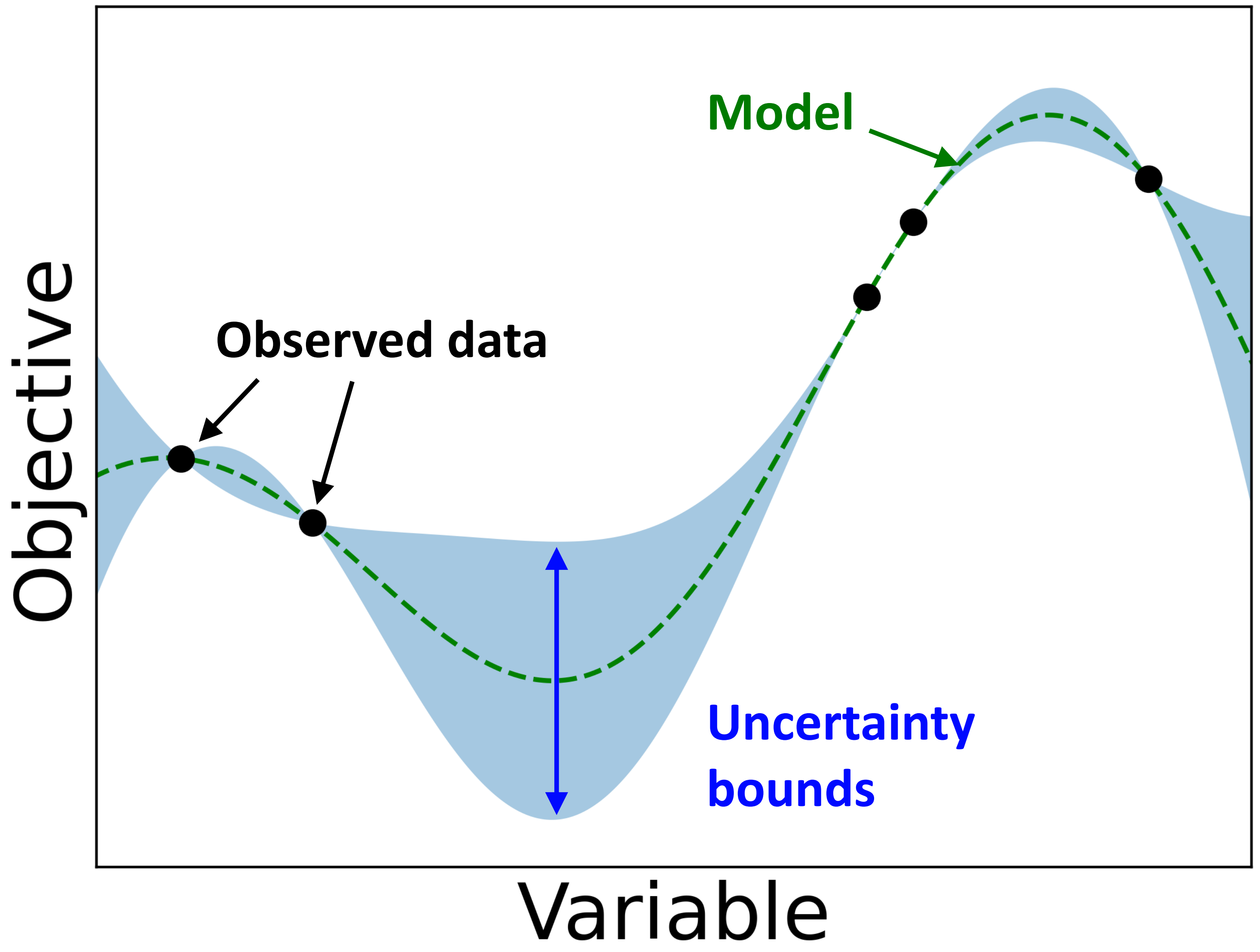
~5% of the space is optimal

How to efficiently identify?



Pragnay Nevatia

Black box approach: Bayesian optimization (BO)



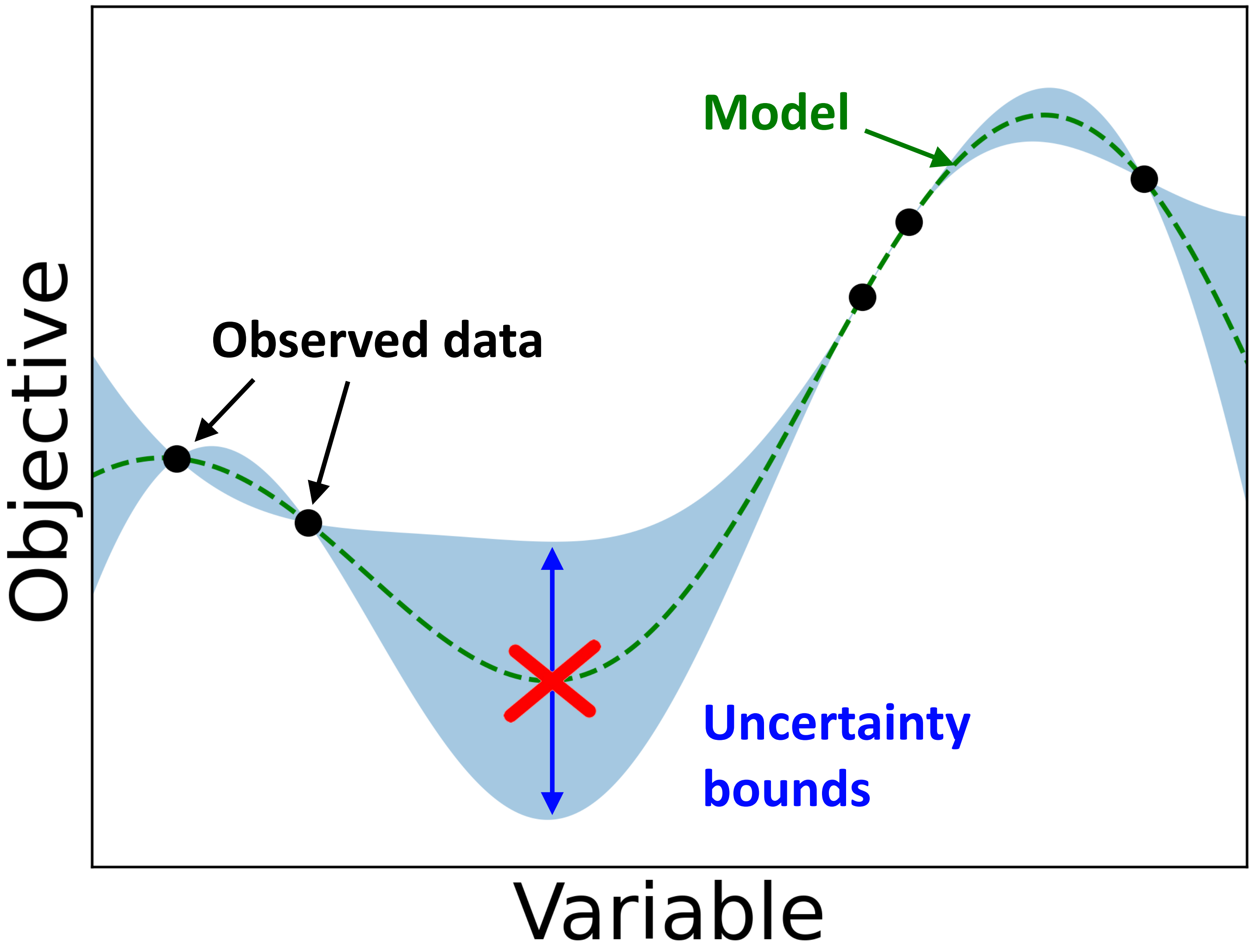
Exploration

Sample points to
minimize uncertainty

Exploitation

Sample points to
maximize the objective

Black box approach: Bayesian optimization (BO)



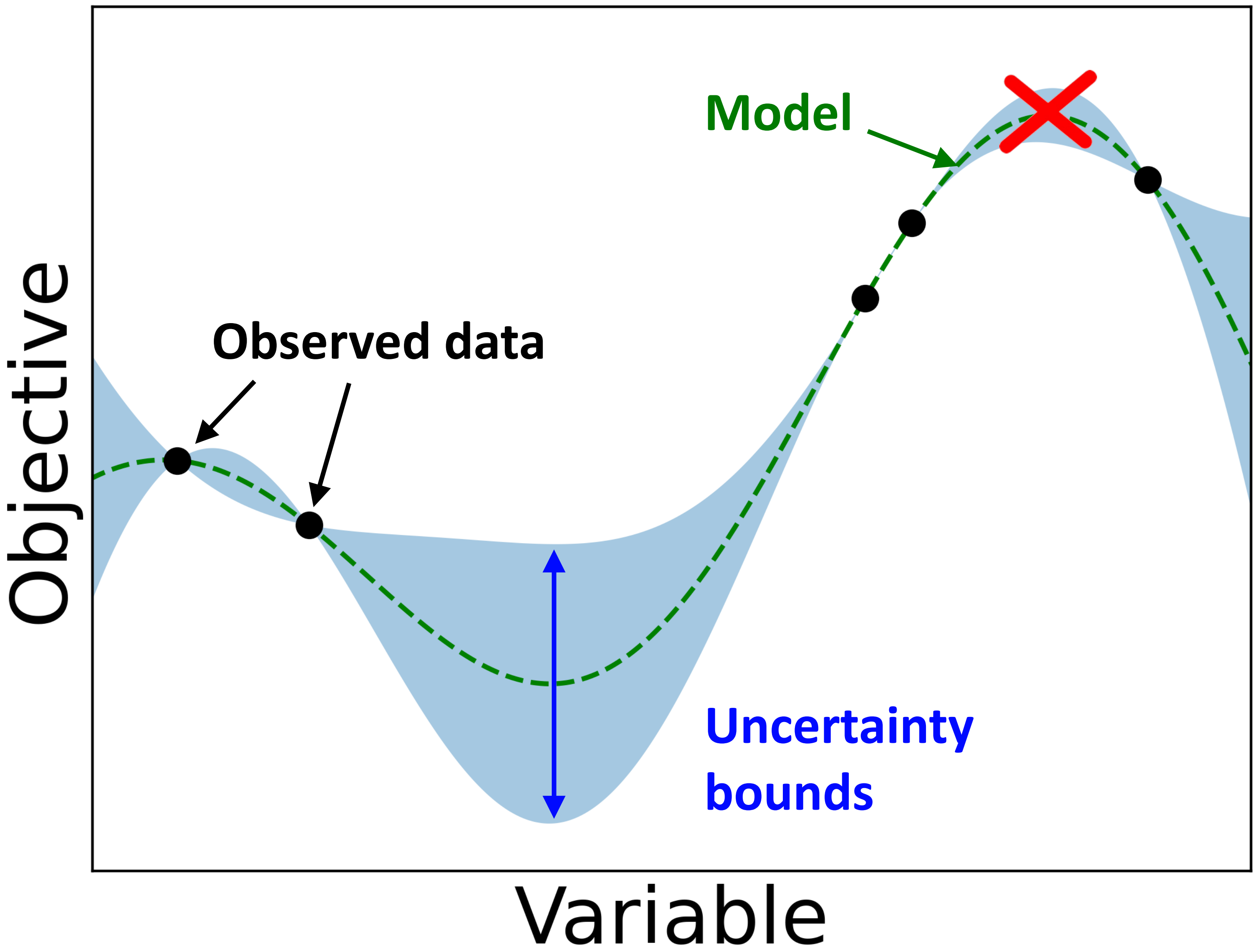
Exploration

Sample points to
minimize uncertainty

Exploitation

Sample points to
maximize the objective

Black box approach: Bayesian optimization (BO)



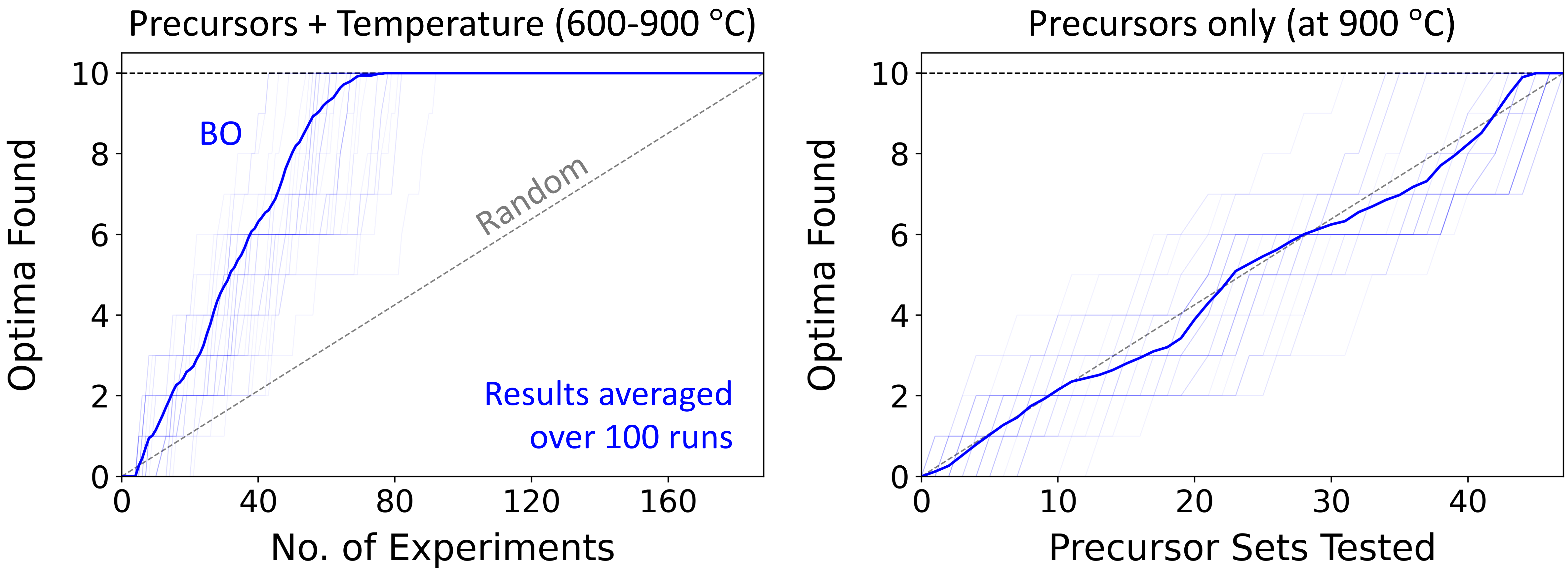
Exploration

Sample points to
minimize uncertainty

Exploitation

Sample points to
maximize the objective

BO optimizes temperature, but not precursors

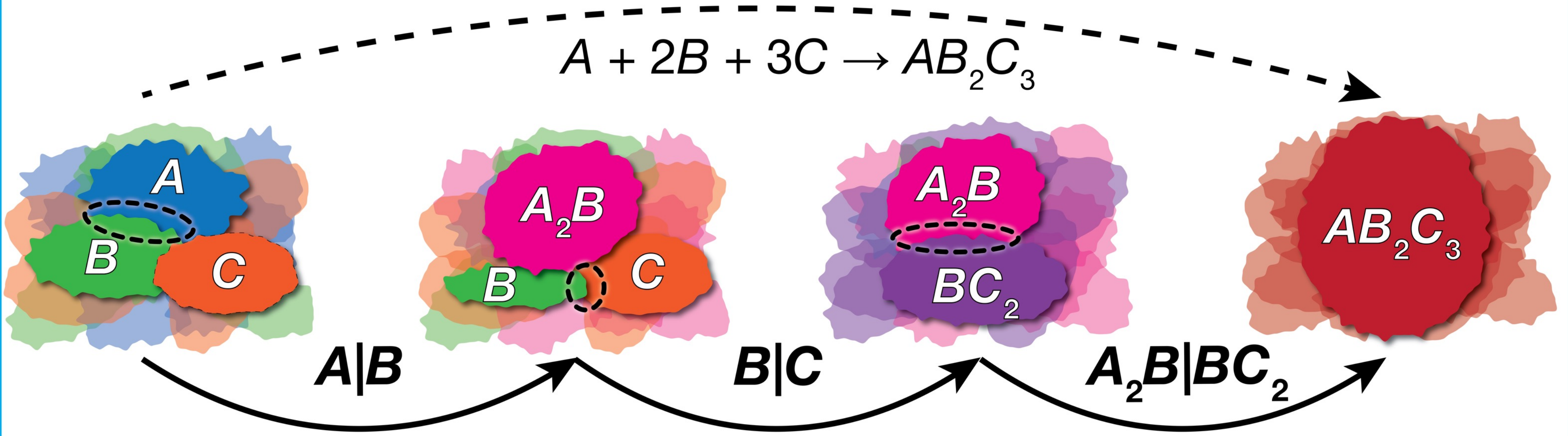


BO learns to focus on high T (900 °C) to maximize YBCO yield ☺

But…precursor selection is basically random ☹

Synthesis concepts can guide precursor selection

Solid-state reactions occur sequentially between *pairs* of phases



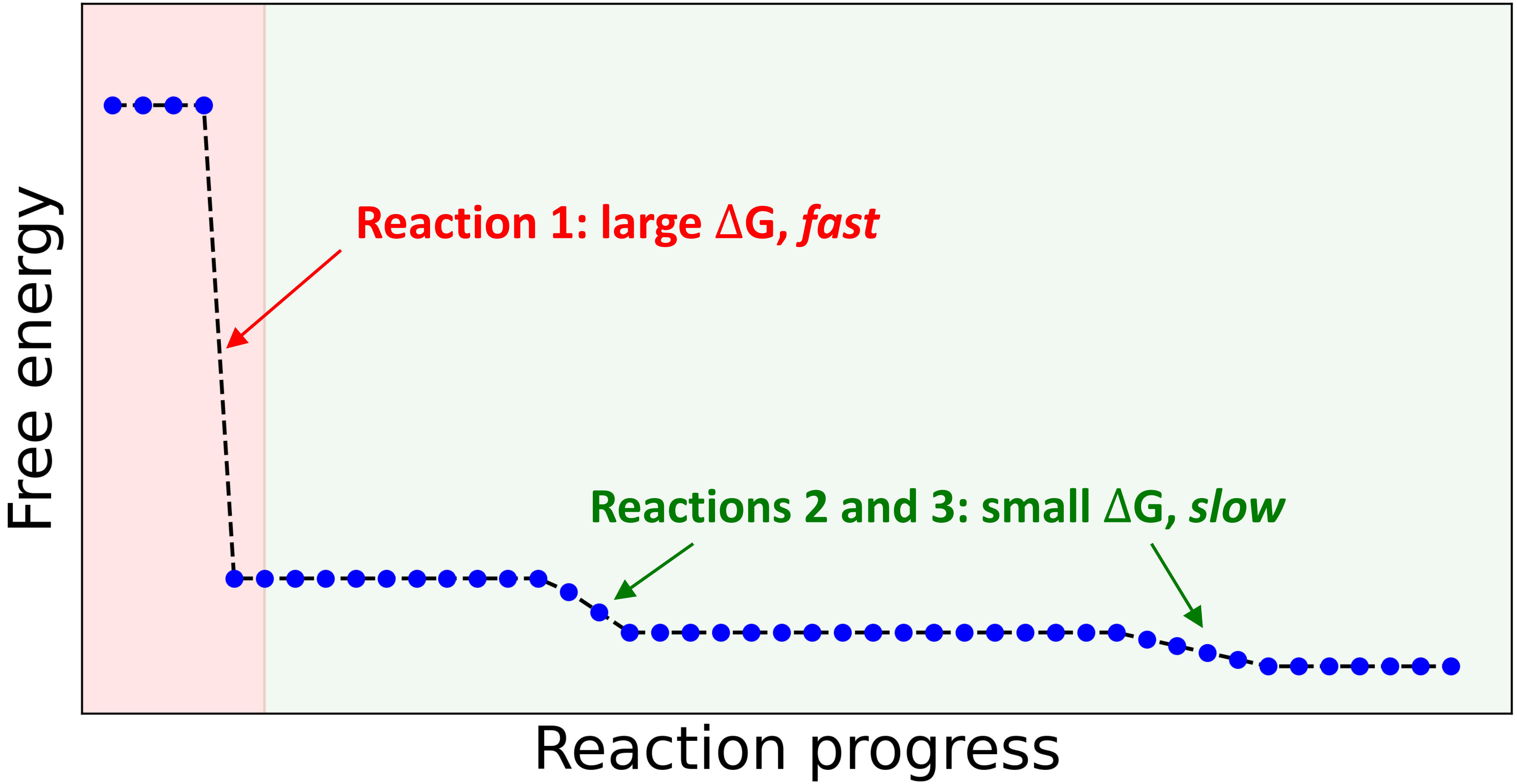
Pairwise reactions

Step-by-step sequence of "2×2" reactions until equilibrium is reached

A. Muira, C. Bartel, …, G. Ceder, W. Sun, Adv. Mater. 2021.

Synthesis concepts can guide precursor selection

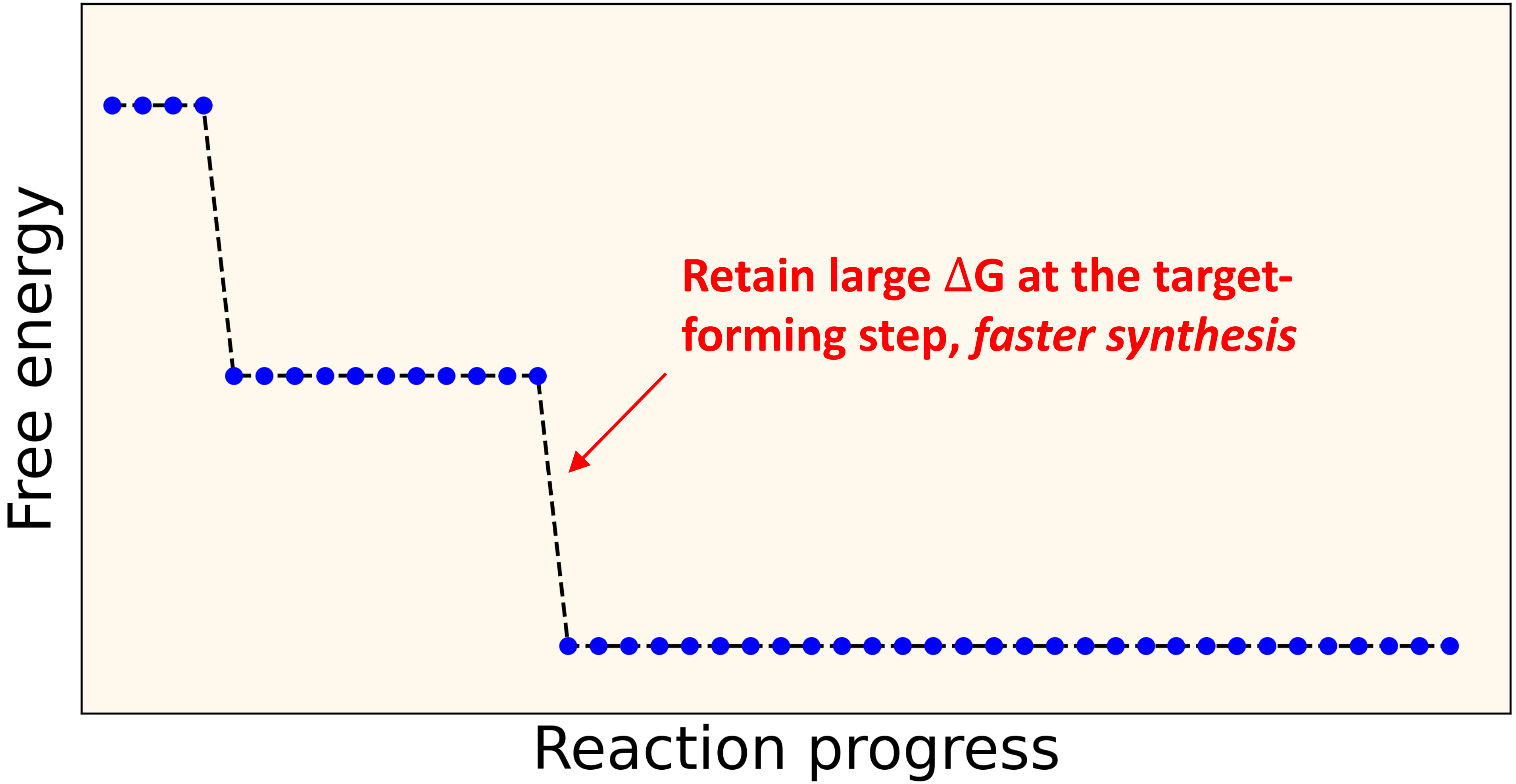
Precursors → intermediate phases → final products



M. Bianchini, …, G. Ceder, Nature Materials, 2020.

Synthesis concepts can guide precursor selection

Re-designed reaction pathway for accelerated synthesis

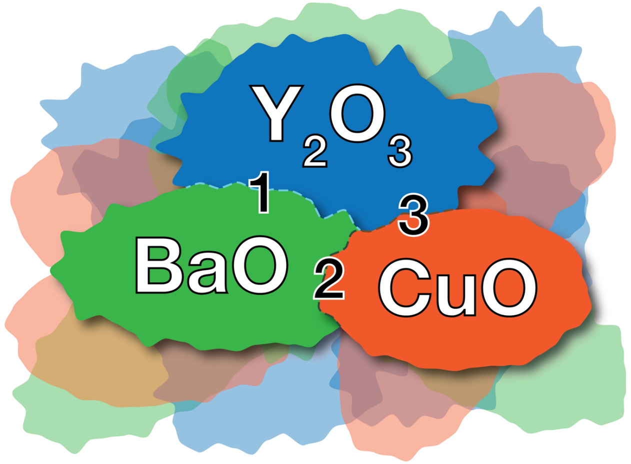


M. Bianchini, …, G. Ceder, Nature Materials, 2020.

Physics informed optimization

Exploration

- Prioritize precursor sets with most **<u>new</u> pairs**
- Record pairwise reactions observed experimentally



Three possible pairwise reactions

Exploitation

- Prioritize precursors with **large ΔG** to form the target phase
- Update ΔG based on observed intermediate reactions

Reaction	ΔG
$Y_2O_3 + BaO_2 + CuCO_3 \rightarrow YBCO + O_2/CO_2$	-684 meV
$Y_2Cu_2O_5 + BaO_2 + CuO \rightarrow YBCO + O_2$	-591 meV
…	…

Physics informed optimization

Exploration

- Prioritize precursor sets with most **new pairs**
- Record pairwise reactions observed experimentally

Three possible pairwise reactions

Exploitation

- Prioritize precursors with **large ΔG** to form the target phase
- Update ΔG based on observed intermediate reactions

Reaction	**ΔG**
$Y_2O_3 + BaO_2 + CuCO_3 \rightarrow$ YBCO + O_2/CO_2	-684 meV
$Y_2Cu_2O_5 + BaO_2 + CuO \rightarrow$ YBCO + O_2	-591 meV
…	…

Physics informed optimization

$$Cu_2O + BaO_2 + Y_2O_3 \xrightarrow[\textit{Predict}]{Cu_2O|BaO_2} Ba_2Cu_3O_6 + Y_2O_3$$

$\Delta G = -349\ \mathrm{meV}$ → YBCO

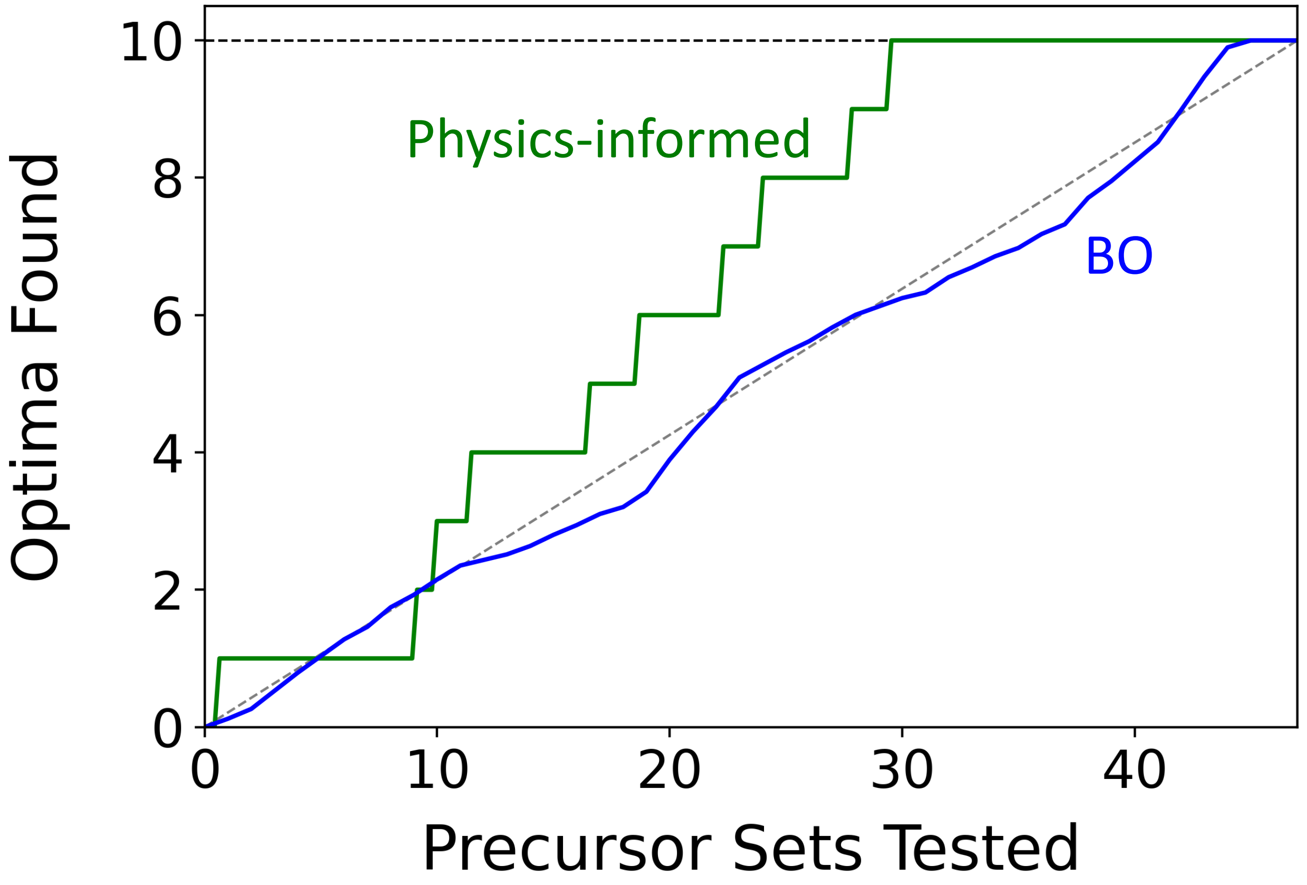
$\Delta G' = -77\ \mathrm{meV}$ → YBCO

Exploitation

- Prioritize precursors with **large ΔG** to form the target phase
- Update ΔG based on observed intermediate reactions

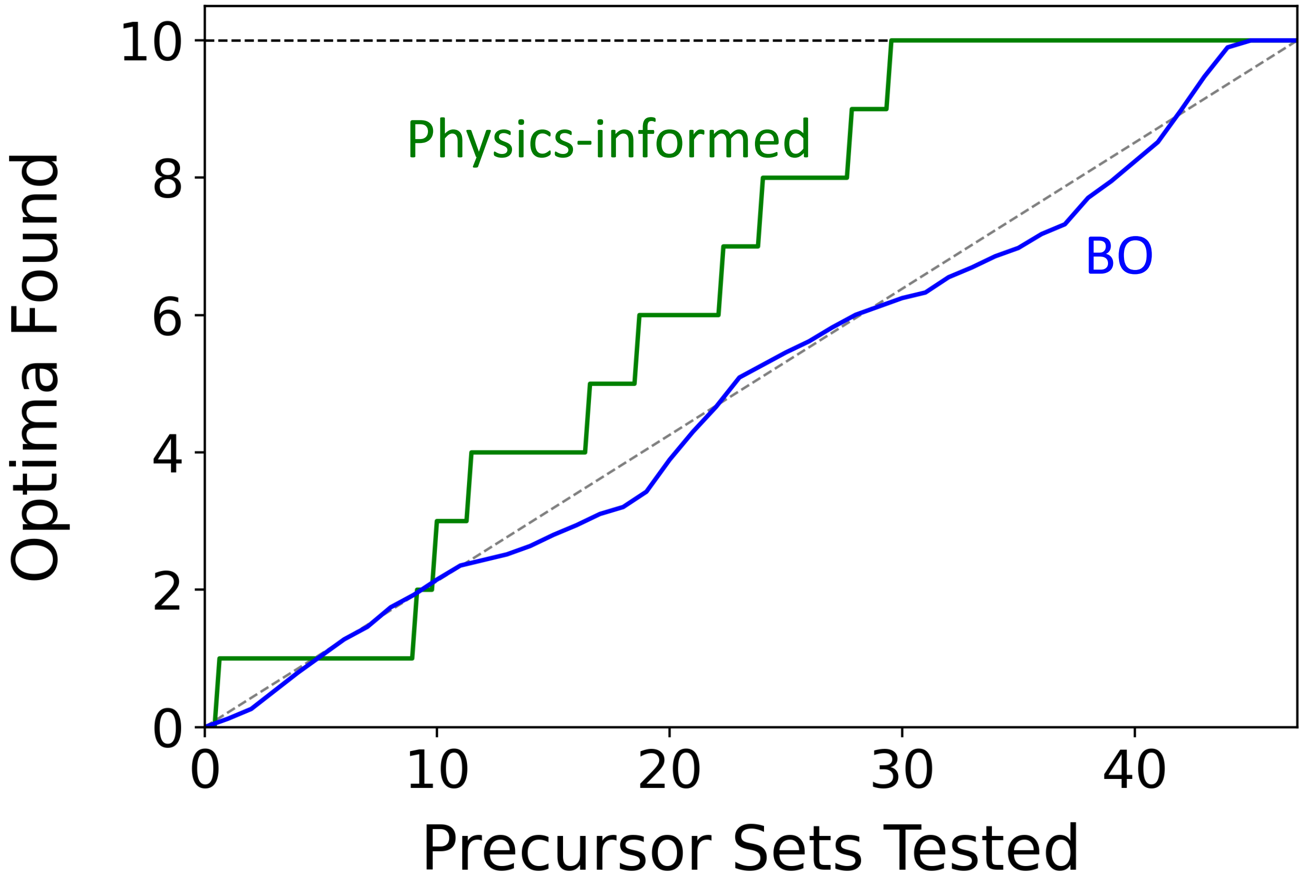
Reaction	ΔG
$Y_2O_3 + BaO_2 + CuCO_3 \rightarrow$ YBCO $+ O_2/CO_2$	-684 meV
$Y_2Cu_2O_5 + BaO_2 + CuO \rightarrow$ YBCO $+ O_2$	-591 meV
…	…

Pairwise analysis leads to better precursor selection



- Physics-informed approach outperforms Bayesian optimization
- **29 precursor sets** tested to identify *all* optimal synthesis routes

Pairwise analysis leads to better precursor selection

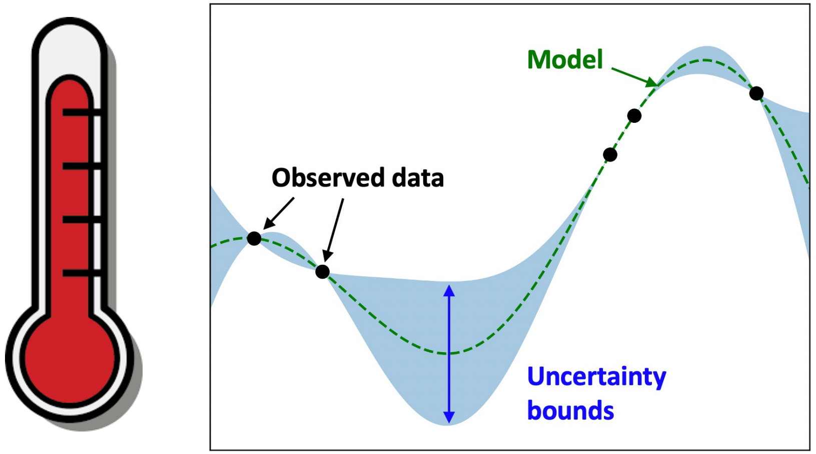


- Physics-informed approach outperforms Bayesian optimization
- **29 precursor sets** tested to identify *all* optimal synthesis routes
- **51 pairwise reactions** discovered

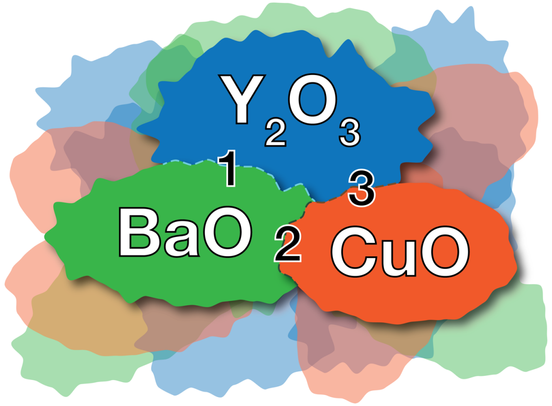
Phases	Temp.	Products
Y_2O_3|$BaCuO_2$	800-900 °C	Y_2BaCuO_5
$BaCuO_2$|$Y_2Cu_2O_5$	700-800 °C	YBCO
…	…	

Conclusions

Black-box techniques can optimize synthesis temperature



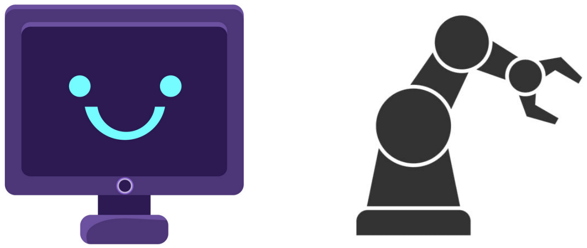
Domain knowledge is needed to guide precursor selection



ΔG

Grand vision: integration of decision making with automated experiments

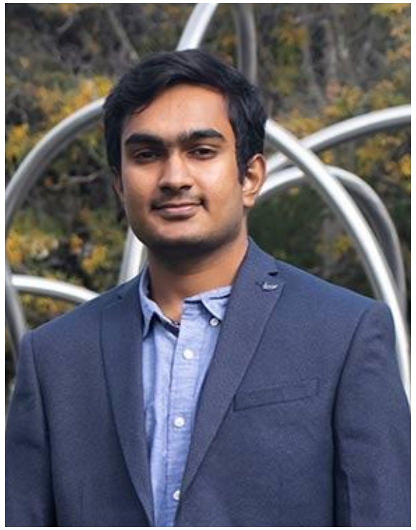
Target phase



Optimized synthesis recipe

Acknowledgements

Professor Gerbrand Ceder



Pragnay Nevatia

Dr. Yan Zeng

Dr. Chris Bartel

Nathan_Szymanski@berkeley.edu

@NJSzymanski