Quantum Material Synthesis by Reinforcement Learning

Pankaj Rajak$^1$, Aravind Krishnamoorthy$^2$, Aiichiro Nakano$^2$, Rajiv K. Kalia$^2$, Priya Vashishta$^2$

$^1$Argonne National Laboratory, $^2$University of Southern California

Abstract

Designing conditions for experimental synthesis is the primary bottleneck for the realization of new functional quantum materials. Current strategies to synthesize new promising materials with desired properties are based upon a trial-and-error approach, which is a time-consuming process and does not generalize to different materials. Here, we use deep reinforcement learning to design synthesis schedules, which are time-dependent synthesis conditions of temperatures and reactant concentrations for a prototypical quantum material, monolayer MoS$_2$ via chemical vapor deposition (CVD). The reinforcement learning (RL) agent is coupled to a deep generative model that captures the probability density function of MoS$_2$-CVD dynamics and is trained on 10,000 computational synthesis simulations. After training, the RL agent successfully learns the optimal policy in terms of threshold temperatures and chemical potentials for the onset of chemical reactions and provides mechanistic insight to predict new synthesis schedules that produce well-sulfurized crystalline and phase-pure MoS$_2$ in minimum time, which is validated by reactive molecular dynamics.

Background and Synthesis of Quantum Material

- Quantum Materials have exotic physical properties, arising from quantum mechanical nature of their electrons.
- Example: 2D semiconducting materials like MoS$_2$, where quantum confinement controls electronic properties.
- Existing scalable techniques like chemical vapor deposition (CVD) have numerous synthesis parameters that need to be tuned/optimized to produce quantum materials with desired properties.

Reinforcement Learning for MoS$_2$ Synthesis

Objective: Learn a policy $\pi_\theta$ using policy gradient algorithm for optimal synthesis condition in terms of temperature ($T$) and gas conc. ($H_2$, $S_2$, $H_2S$) to a MoS$_2$ structure with maximum 2H phase fraction in minimum time. By maximizing the following objective function

$$\arg\max_{\theta} \mathbb{E}_t \left[ \sum_{t=0}^{T-1} r(s_t, a_t) \right]$$

- State $s_t$ encodes synthesis history $(Z_{1:t}, X_{1:t})$
- Output action $a_t$ is a gaussian distribution over the change in simulation condition $(\Delta T, \Delta S_2, \Delta H_2, \Delta H_2S)$
- $r(s_t, a_t)$ reward received by RL agent

(c) shows effect of initial gas conc. on 2H phase fraction. RL policy promotes low initial $S_2$ concentration, which promotes evolution of oxygen and self-diffusion of MoO$_2$ crystal.

Conclusions

- Developed a RL scheme for predictive synthesis of MoS$_2$ quantum material using CVD.
- RL agent successfully proposed several new reaction schedules, i.e., time-dependent reaction conditions, to synthesize highly crystalline MoS$_2$ with maximum 2H phase fraction.
- RL agent provides insight into the mechanism of material synthesis and understanding of the role of synthesis conditions (temperature, chemical environment) on the quality of the synthesized crystal.
- RL scheme provides the first viable high-throughput approach to screening material synthesis conditions to tackle the as-yet unsolved problem of predictive synthesis of novel nanomaterials.

Probability Density Function of CVD Dynamics

- Each RMD simulation is expensive (~3 days).
- We modeled the probability density function of CVD dynamics with an autoregressive density function (NADE-CVD), where each conditional probability is modeled with Gaussian distribution.
- Each RMD simulation is discretized into 1 ns interval, where the input to NADE-CVD at timestep $t$ is $X_t = (T_t, H_2, S_2, H_2S)$ and $Z_t = \left( n_{2H}^{\text{defect}}, n_{1T}^{\text{defect}}, \sigma_{1T}^{\text{defect}} \right)$, and it output $n_{2H}, \sigma_{1T}$ for next time step.
- Correlation and error analysis of test datasets shows mean absolute error (MAE) are $\leq 0.1\%$ phase fraction for all three phases.

Probability Distributions of CVD Processes

- $P\left(Z_{1:2} | Z_{1} \right) = P\left(Z_{2} | Z_{1}, X_{1} \right) = P\left(n_{1T}^{\text{defect}}, n_{1T}^{\text{defect}}, n_{2H}^{\text{defect}, \text{defect}} \right)$
- $P\left(Z_{2:h+1} | X_{1:h+1} \right) = \mathcal{N}\left(\mathbf{z}_{h+1}, \sigma_{h+1}\right)

<table>
<thead>
<tr>
<th>Probability Distribution</th>
<th>Time-Dependent Reaction Conditions</th>
<th>Time-Dependent Phase Fractions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $S_2$</td>
<td><img src="image" alt="Low $S_2$" /></td>
<td><img src="image" alt="Low $S_2$ Phase Fraction" /></td>
</tr>
</tbody>
</table>