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

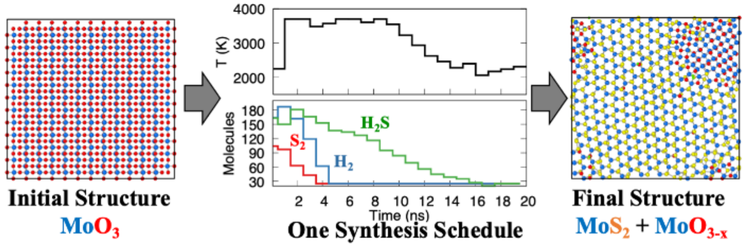
Abstract

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 learn 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 MoS2-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-sulfidized crystalline and phase-pure MoS2 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.

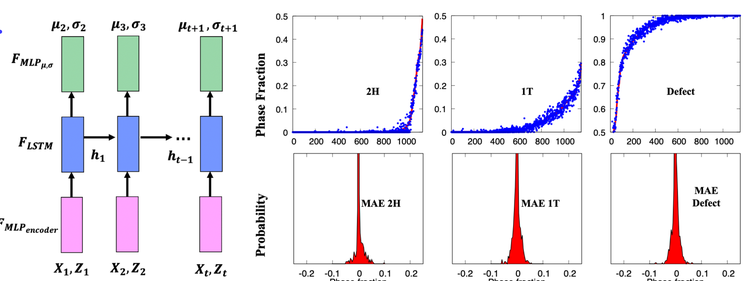
MOS_2 Synthesis by CVD using reactive molecular dynamics (RMD)



- Final MoS_2 + MoO_{3-x} is a complex structure consisting of *three phases 2H, 1T and defect.*
- 2H is the semiconducting phase, with tunable electronic properties to build new optoelectronic devices

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 $Z_t = \left(n_t^{2H}, n_t^{1T}, n_t^{Defect}\right)$, and it output μ_{t+1}, σ_{t+1} for next time step.
- Correlation and error analysis of test datasets shows mean absolute error (MAE) are ≤ 0.1% phase fraction for all three phases



$$P(Z_{2:T}|Z_1X) = P(Z_2|Z_1, X_1) \dots P(Z_{t+1}|Z_{1:t}, X_{1:n}) \dots P(Z_T|Z_{1:T-1}, X_{1:T-1})$$

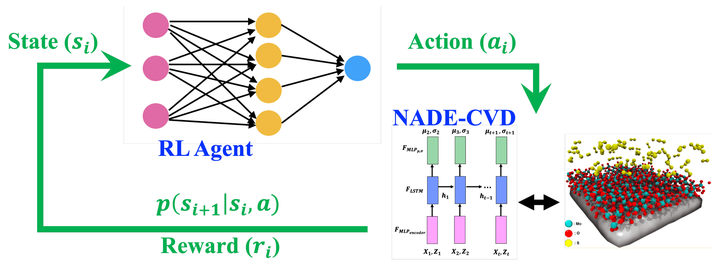
$$P(Z_{t+1}|Z_{1:t}, X_{1:t}) = \mathcal{N}(Z_{t+1}|\mu_{t+1}, \sigma_{t+1})$$

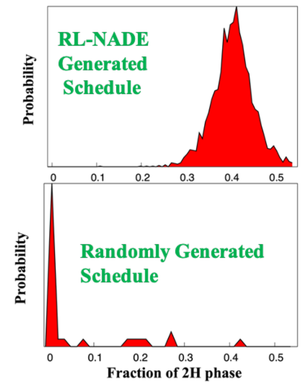
$$F(Z_{1:t}, X_{1:t}) = \mu_{t+1}, \sigma_{t+1} = \{\mu_{t+1}^{2H}, \mu_{t+1}^{1T}, \mu_{t+1}^{defect}, \sigma_{t+1}^{2H}, \sigma_{t+1}^{1T}, \sigma_{t+1}^{defect}\}$$

Reinforcement Learning for MoS_2 Synthesis

Objective: Learn a policy π_θ using policy gradient algorithm for optimal synthesis condition *in terms of temperature (T) and gas conc. (H_2, S_2, H_2S)* a to create MoS_2 structure with maximum 2H phase fraction in minimum time. By maximizing the following objective function

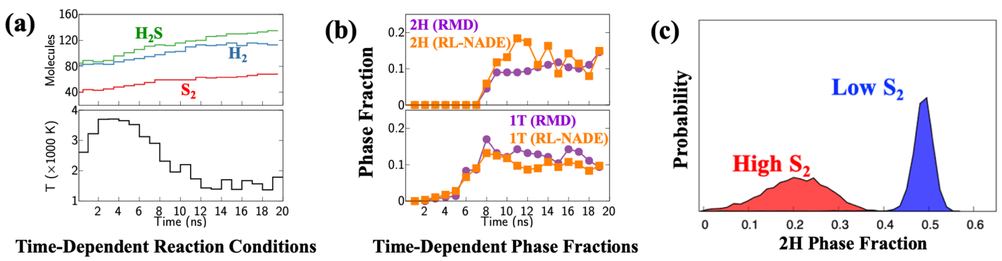
$$arg\,max_\theta\, \mathbb{E}_{\tau\sim\pi_\theta}\left[\sum_{t=1}^{t=20\,ns} r(s_t, a_t)\right]; r(s_t, a_t) = \begin{cases} 0.0 & if\ Z_{t+1}[n_{2H}] < 0.40 \\ 0.2Z_{t+1}[n_{2H}] & if\ Z_{t+1}[n_{2H}] \geq 0.40 \end{cases}$$





- State s_t is 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

- (a) shows a RL-agent generated synthesis condition and (b) corresponding phase fraction of 2H and 1T phases predicted by NADE-CVD and validated by RMD simulation.
- Captures the non-trivial mechanism of 1T nucleation before 2H, and initial high temperature needed to melt MoO_3 crystal.



- (c) shows effect of initial gas conc. on 2H phase fraction. RL policy promotes low initial S_2 conc. (0-3ns) when temperature is high, which promotes evolution of oxygen and self-reduction of MoO_3 surface.

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.