Novel Methods for Generating and Mutating Site-Disordered Materials in an Evolutionary Energy Minimization Framework

Ethan Meltzer[a], Mr. Victor T. Barone[b], Dr. Richard E. Irving[b], Dr. Sanjay V. Khare[b]

[a] Oberlin College Dept. of Computer Science

[b] University of Toledo Dept. Of Physics

Site-Disordered Cu_2AgBiI_6

- Lead-free semiconductor with high absorption coeff. ($1.0*10^5\ cm^{-1}$) in visible spectrum.
  - Compete with or ideally exceed efficiency of current gen. thin-film lead-perovskite and CdTe materials
- Original research done at University of Liverpool suggests material is highly site-disordered, ~10^{31} configurations possible for 3x3x1 cell
- Need to know energy-minimal configurations of the material for proper solar cell chemistry

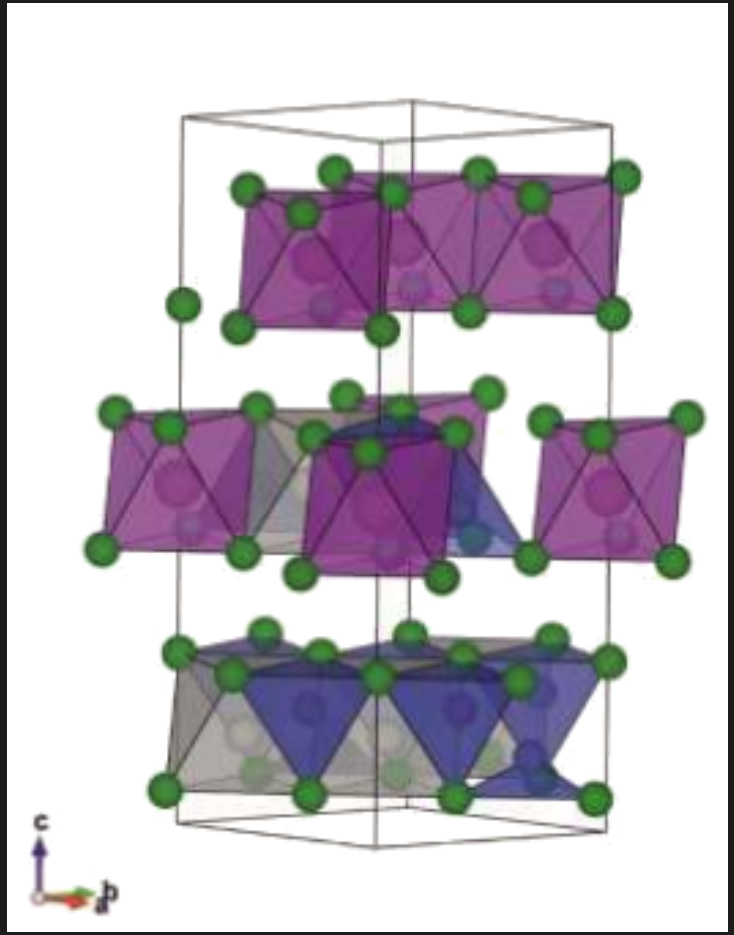
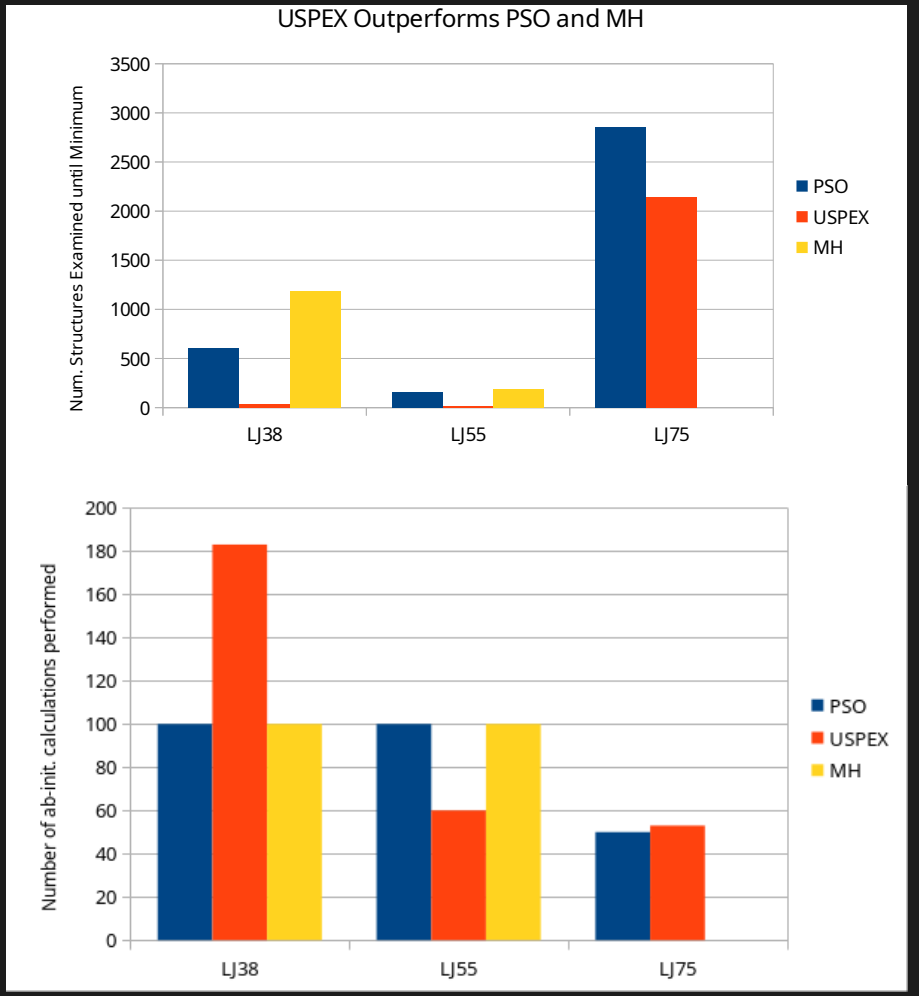


Figure: Sansom et al. *Highly Absorbing Lead-Free Semiconductor Cu_2AgBiI_6 for Photovoltaic Applications from the Quaternary CuI-AgI-BiI_3 Phase Space.*
Journal of the American Chemical Society 2021 143 (10), 3983-3992
DOI: 10.1021/jacs.1c00495

Evolutionary Algorithim and USPEX

- Can survey population and find minimum quickly
  - Takes advantage of the parallel nature of simulating a generation
- Initialize, Select, Breed (combine, crossover, mutate), Repeat (until criteria are met), Terminate
- Universal Structure Predictor: Evolutionary Xtallography (USPEX) software implements efficiently, interfaces with VASP
- Not optimized for site-disordered materials, will compare with our novel approach focused on first-gen selection



Figures adapted from data available on USPEX website: https://uspex-team.org/en/uspex/overview

First-Generation Binning

- Bias first-generation towards high-symmetry configurations in hopes of finding lower energy configurations faster.

$$B_i = \left\lfloor \begin{cases} a\frac{\mathrm{X}(\frac{zi}{n})}{\mathrm{X}(z)}, z > 0 \\ a\frac{i}{n}, z = 0 \\ a\left(1 - \frac{\mathrm{X}(\frac{zi}{n})}{\mathrm{X}(z)}\right), z < 0 \end{cases} \right\rfloor$$

$$\mathrm{X}(x) = \int_0^x \exp\left(-\frac{t^2}{2}\right) dt$$

B_i is the i-th bin, a is the number of space groups, n is the number of bins, and z is the bias factor (higher = more aggressive bias in the positive direction)

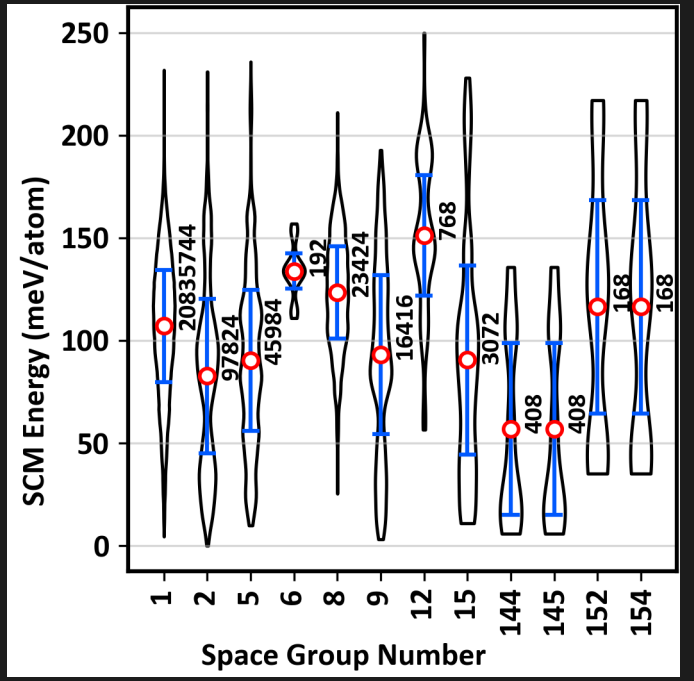


Figure: in-progress research from Khare group

Verifying Validity of Space Groups

- Random generation of configurations *failed* to produce any high-symmetry configurations (space group > 1)
  - Generated ~50k configurations, expected frequency of high-order configurations is 0.5%, expect on average 250 high-symmetry configurations for a sample of that size.
- Vast majority of space groups are impossible to obtain for this stoichiometry, restricting focus to trigonal groups 143 – 161
- “Closed” symmetry group sets demand a num. of sites that is a whole number multiple of num. of operations, “open” sets require that the number of sites shares a factor with num. of operations.

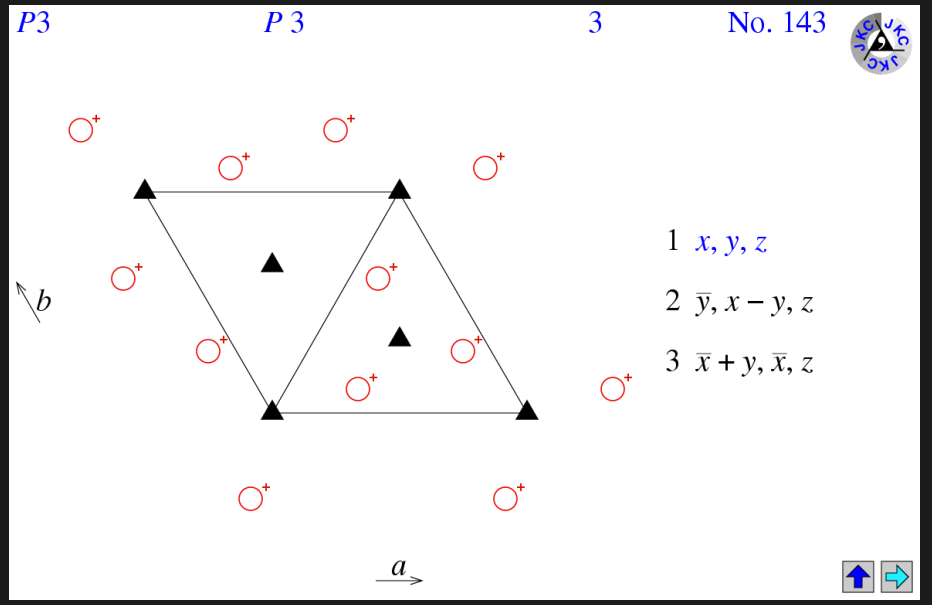


Diagram for space group 143 with listed symmetry operations, from University of London http://img.chem.ucl.ac.uk/sgp/large/143az1.htm

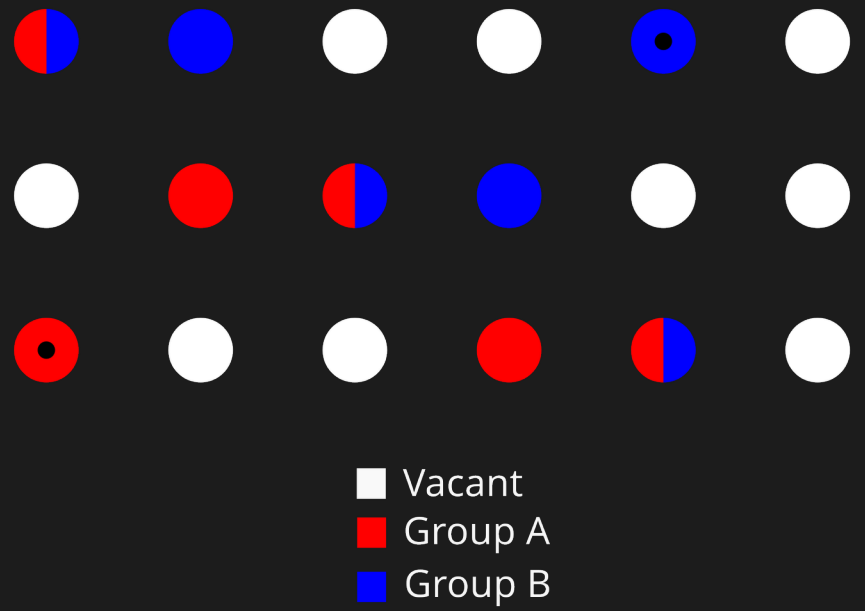
The *c* basis vector is normal to the plane of the screen.

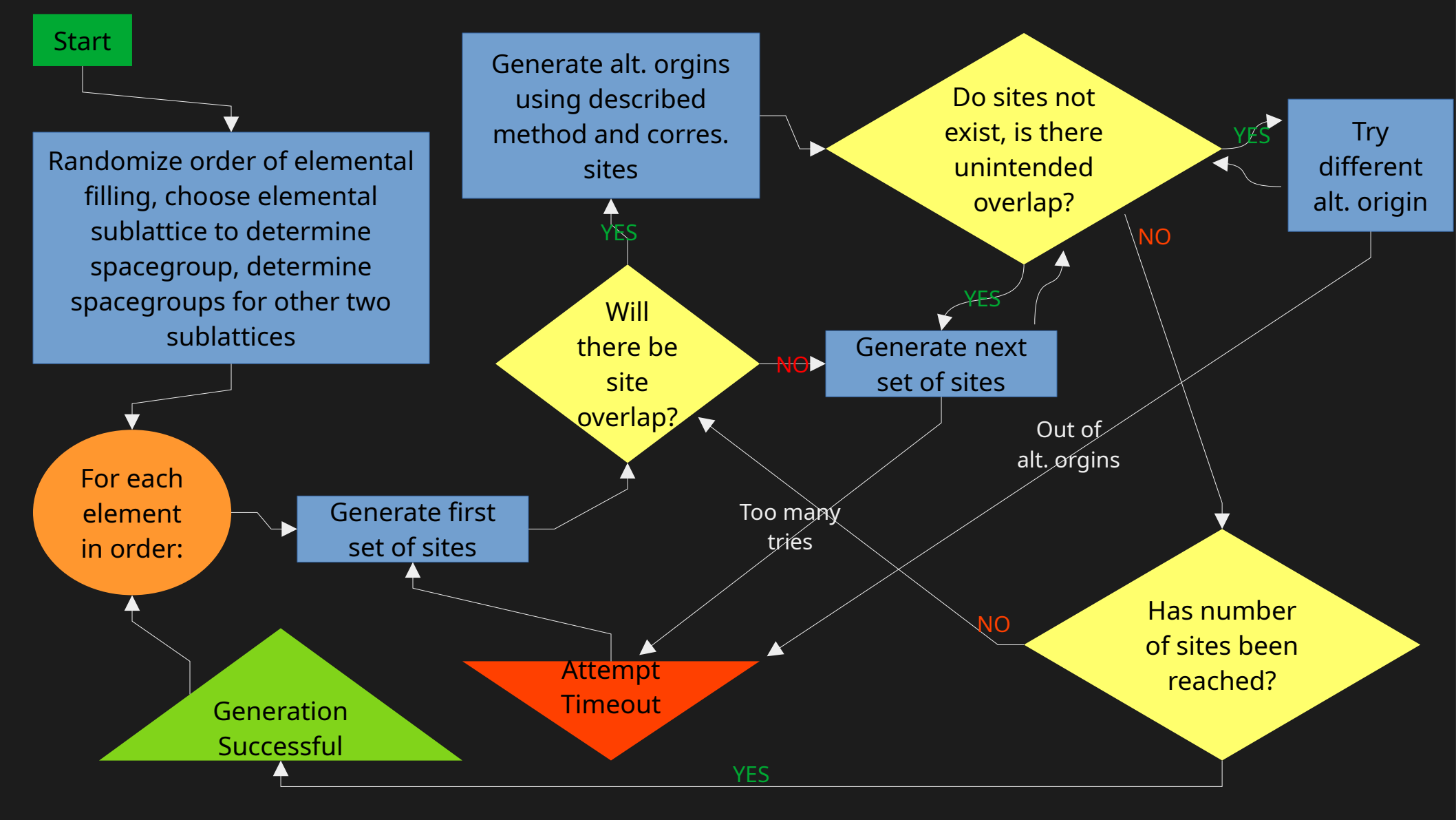
Forced Generation of High-Symmetry Configurations

- To fill first generation, manually generate high-symmetry configurations
  - Up-front computational cost of doing so should be less than the additional cost needed to simulate large highly-random early generations
- Compiled a database of the symmetry operations in each space group
- Verified their closedness v. openness – all groups 143 to 161 are open.
- Design considerations for generation script:
  - Should be able to sample the *entire* population of high-symmetry configurations – no unintended bias
  - Should be able to input a space group number, index file containing all possible sites in the lattice, and number of configurations desired and output that many configurations in that spacegroup
  - In order to qualify for a certain space group, each elemental sublattice must be target space group or greater.

Resolve Overlapping Symmetry Sets

- Ag and Bi both need to occupy 9 sites, groups 147-161 have 6 symmetry ops
  - Overlap 2 symmetry sets so they share 3 sites
  - Generate one set of 6 and all its size-3 subsets
  - Generate all size-3 subsets of the inverse symmetry operations *except* the one that corresponds to the points in the original set
  - Convolute these two lists to find possible alternate orgins
  - Choose at random one of the possible orgins if there are multiple, generate the rest of points in that set, and make sure that they are within the elemental sublattice




Start
Randomize order of elemental filling, choose elemental sublattice to determine spacegroup, determine spacegroups for other two sublattices
For each element in order:
Generate first set of sites
Will there be site overlap?
YES
Generate alt. orgins using described method and corres. sites
NO
Generate next set of sites
YES
Do sites not exist, is there unintended overlap?
YES
Try different alt. origin
NO
Out of alt. orgins
Too many tries
Has number of sites been reached?
NO
YES
Attempt Timeout
Generation Successful

Next Steps, Continued Research

- Complete implementation of config. generation algorithm
- Generate and simulate first generation using modified binning strategy, use EA to converge on minima
  - Tweak and rerun multiple times to refine parameters of evolution, first-generation filling.
- Implement and simulate crystal using USPEX software
- Compare resulting minima found by both methods, also analyze runtime data to determine computational cost/efficiency
- Report findings, hopefully publish results

Thank you!

Ethan Meltzer, Oberlin College Dept. of Computer Science
Mr. Victor T. Barone, Dr. Richard E. Irving, Dr. Sanjay V. Khare
University of Toledo Dept. Of Physics

Questions?
emeltzer@oberlin.edu