Improvement in Convergence Speed of Fully-parallel Annealing Algorithm with Spin-update Restriction

June 24, 2021
Poster D in AQC2021 Virtual Conference

Kazushi Kawamura
AI Computing Research Unit (ArtIC), Tokyo Institute of Technology, Japan
Our Previous Work: SCA

- **SCA**: Stochastic Cellular Automata annealing
  - is one of the extension algorithms of SA (Simulated Annealing)
  - has capability to update spins of a fully connected Ising model in a completely parallel manner

- The breakthrough of this technique is the conversion from Ising model to cellular automata

![Ising spin system](image1)

![SCA spin system](image2)
Comparison of SA and SCA

**Spin Connection**

**Hamiltonian**

\[ H(\sigma) = -\frac{1}{2} \sum_{i,j \in E} J_{ij} \sigma_i \sigma_j - \sum_{i \in V} h_i \sigma_i \]

**Spin Update**

\[ H(\sigma, \tau) = -\frac{1}{2} \sum_{i,j \in E} J_{ij} \sigma_i \tau_j - \frac{1}{2} \sum_{i \in V} h_i (\sigma_i + \tau_i) \]

\[ + \frac{1}{2} \sum_{i \in V} q (1 - \sigma_i \tau_i) \]

**Pinning Parameter**
Our Previous Work: STATICA

• STATICA: Hardware Architecture for SCA

Prototype LSI chip (512 spins)

For detail, see the following paper:
Behavior of SCA for different $q$’s

- $q$ has a role to ensure the proper convergence in SCA

$$H(\sigma, \tau) = -\frac{1}{2} \sum_{i,j \in V} J_{ij} \sigma_i \tau_j - \frac{1}{2} \sum_{i \in V} h_i (\sigma_i + \tau_i) + \frac{1}{2} \sum_{i \in V} q (1 - \sigma_i \tau_i) \quad (q > 0)$$

Pinning Parameter

$$\begin{cases} 
\sigma_i = \tau_i \iff q (1 - \sigma_i \tau_i) = 0 \\
\sigma_i \neq \tau_i \iff q (1 - \sigma_i \tau_i) = 2q 
\end{cases}$$

- However, using large $q$ is not practical :(

Simulation w/ 2000-spin Ising model (G22)

- large $q$: spin flips are strongly restricted
  - Slow convergence
  - Easy to be trapped in local minima

$q = 10.5$

- 500

$q = 5.3$

- 700
Behavior of SCA for different $q$’s

• **Next question**: Should $q$ be smaller? 😕

• **Answer**: It is not always so
  - because the proper convergence of SCA might not be ensured

Simulation w/ 2000-spin Ising model (G22)

- **small** $q$ : $\min H(\sigma, \tau) \neq \min H(\sigma)$
  - Unable to achieve $\min H(\sigma)$
  - Oscillation: All spins are flipped every step

\[
\begin{align*}
H(\sigma, \tau) = & -6,511 \\
H(\sigma) = & -6
\end{align*}
\]

\[
\begin{align*}
H(\sigma, \tau) = & 19,990 \\
H(\sigma) = & 19
\end{align*}
\]
SCA is in a dilemma

• $q$ must be large for ensuring the proper convergence

• $q$ should be small for improving the convergence

• Simulation results show that only making $q$ as small as possible is not enough for solving combinatorial optimization problems (see slide 11)
Proposed Algorithm: $\varepsilon$SCA

- **Observation**: Using small $q$ in SCA caused spin flip oscillation

- **Question**: Can we stop (or restrain) the oscillation? If yes, how can we do so?

- **Strategy**: That may be realized by restricting the number of spins that are updated simultaneously
Proposed Algorithm: εSCA

- **εSCA**: SCA w/ spin-update restriction

In εSCA, the i-th spin is flipped when satisfying both $P_i > \text{rand}_1$ and $\varepsilon > \text{rand}_2$

- $P_i$: flip probability of i-th spin
- $\varepsilon$: parameter
- $\text{rand}_1$, $\text{rand}_2$ = [0, 1)

Flip trials are done for the spins surrounded by the epsilon symbol ($\varepsilon = 0.6$)
Proposed Algorithm: $\varepsilon$SCA

- $\varepsilon$SCA : SCA w/ spin-update restriction

Simulation w/ 2000-spin Ising model (G22)

**SCA**

$q = 5.3$

$q = 3.5$

**$\varepsilon$SCA**

$q = 3.5$

$\varepsilon = 1.0 \rightarrow 0.5$

$q = 3.5$

$\varepsilon = 0.95 \rightarrow 0.2$

$q = 1.5$

$\varepsilon = -6,511$

$\varepsilon = -6,554$

$\varepsilon = -6,582$
Benchmarking of SCA & $\varepsilon$SCA for N-Queen Problem

**N-queen Problem**

Place N chess queens on an N x N chessboard so that no two queens attack each other

**Setup**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>#MC steps</td>
<td>${1000 \times 2}$</td>
</tr>
<tr>
<td>$T_{\text{init}}$</td>
<td>${30, 10, 5}$</td>
</tr>
<tr>
<td>$T_{\text{fin}}$</td>
<td>${0.8}$</td>
</tr>
<tr>
<td>$q_{\text{init}}$</td>
<td>${10, 9, 8, 7, 6}$</td>
</tr>
<tr>
<td>$q_{\text{fin}}$</td>
<td>$= q_{\text{init}}$</td>
</tr>
<tr>
<td>$\varepsilon_{\text{init}}$</td>
<td>$-$</td>
</tr>
<tr>
<td>$\varepsilon_{\text{fin}}$</td>
<td>$= \varepsilon_{\text{init}}$</td>
</tr>
</tbody>
</table>

**Results**

<table>
<thead>
<tr>
<th>SCA</th>
<th>$\varepsilon$SCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>N = 22, 1000 trials</td>
<td></td>
</tr>
</tbody>
</table>

Success Rate vs $q$ for SCA and $\varepsilon$SCA with different $T$ values.
Our Future Direction

• Examine the control method for $\varepsilon$ that can bring out the best performance of $\varepsilon_{\text{SCA}}$

• Mathematically prove that $\varepsilon_{\text{SCA}}$ can reach the same ground state as SA

• Integrate the effect of $\varepsilon$ into the hardware architecture