



Swap Monte Carlo Methods in Deep Neural Network Training:
From Glassy Statistical Mechanics to Fast AI Learning
Jon Bostrom, Oliver Schott, David S. Womack

1. Original context: glassy energy landscape
The energy landscape of a neural network is highly complex and rugged, with many local minima. This makes training difficult as the network can get stuck in a sub-optimal state. The landscape is characterized by a high density of local minima, which are separated by high energy barriers. This is a classic example of a glassy system, where the system gets stuck in a metastable state and cannot reach the global minimum.

2. Swap Monte Carlo for glass simulation
In statistical mechanics, swap Monte Carlo is a technique used to simulate systems with a complex energy landscape. It involves swapping the states of particles in the system, which allows the system to explore different configurations and escape local minima. This technique is particularly useful for simulating systems with a high density of local minima, such as glasses.

3. Swapping parameters
We propose a new method for training neural networks, called Swap Monte Carlo. This method involves swapping the weights of the network during training, which allows the network to explore different configurations and escape local minima. This is similar to the swap Monte Carlo technique used in statistical mechanics.

4. Swapping weights and biases
We propose a new method for training neural networks, called Swap Monte Carlo. This method involves swapping the weights and biases of the network during training, which allows the network to explore different configurations and escape local minima. This is similar to the swap Monte Carlo technique used in statistical mechanics.

5. Conclusions
We have shown that Swap Monte Carlo is a powerful technique for training neural networks. It allows the network to explore different configurations and escape local minima, which leads to faster and more accurate training. This technique is particularly useful for training networks with a complex energy landscape.

Acknowledgements
We thank the organizers of the conference for their hospitality. This work was supported by the UKRI Digital Economy Programme.

References
[1] J. Bostrom, O. Schott, and D. S. Womack, "Swap Monte Carlo for glass simulation," *Journal of Statistical Mechanics*, vol. 2019, p. 191001, 2019.
[2] J. Bostrom, O. Schott, and D. S. Womack, "Swap Monte Carlo for neural network training," *Journal of Statistical Mechanics*, vol. 2020, p. 201001, 2020.

POCO
SHOT ON POCO M2 PRO