Testing if CRBM kernels learn Ising symmetries

Emanuel Casiano-Diaz

July 6, 2022

In order to understand if the kernel of the CRBM learns the symmetries of the physical model (which in this case will be the $2D$ Ising Model) even when the kernel is not explicitly symmetrized, we performed a set of numerical tests.

1 Test: Learned kernels for original vs rotated training set

For this test, we train an RBM using a set of completely random visible layer configurations (i.e, they were not sampled from Monte Carlo nor they represent the full set of possible spin configurations). Additionally, the kernels were initialized to a shape that looked similar to what we have noticed that the CRBM learns, which is that all weights are zero, except three of them that form an "L"-shaped configuration. In this "L"-shaped configuration, the "elbow" consists of a weight that is larger in magnitude than the other two.

After the CRBM was trained using the original set of visible layer configurations, we proceeded to generate the exact same data set, but with each of the configurations rotated by 90° clockwise.

If the learned weights on both cases were the same, this might suggest that the CRBM kernel is indeed learning the symmetries of the physical model, even in the absence of an explicit kernel symmetrization subroutine in the training loop.

Fig. 1 shows the learned kernels. For this example, we chose two kernels of size 2 each. In the figure, the top two kernels were learned using the original data set while the bottom was learned using the same data set but rotated 90 degrees clockwise. Upon closer inspection, it seems like in both cases, the weights were indeed roughly for the original and rotated training set, although their positions on the kernels might have changed. This is evidenced by the positions of the top right and bottom left weights having swapped with each other.To see this more clearly, the numeric values of the learned kernels can be seen in Fig. 2

The results potentially suggest that the "L"-shaped patterns correspond to kernels that have learned the symmetries of the Ising model.
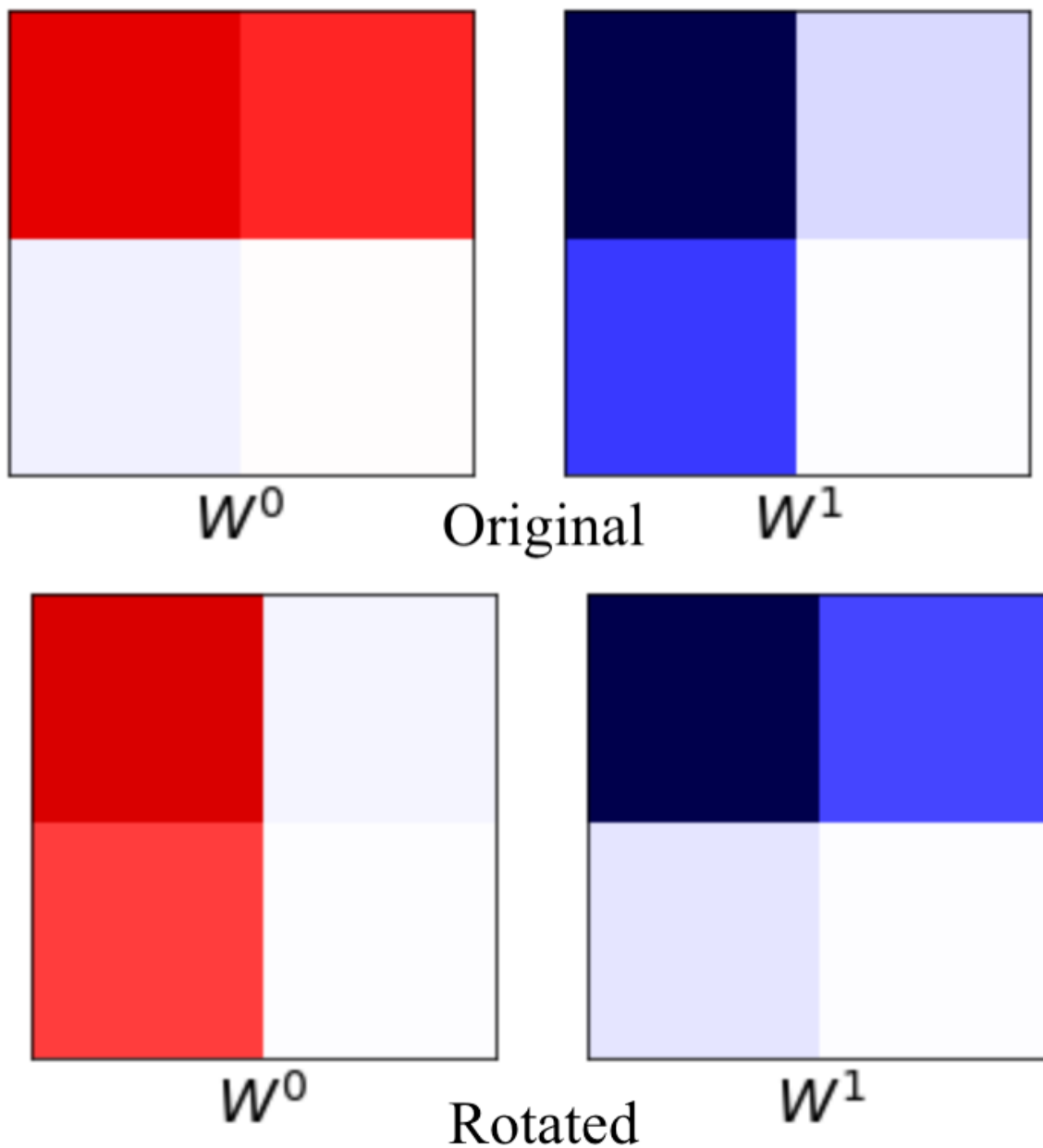
Figure 1: Heatmap of learned kernels. The top two kernels show the learned weights for the case in which the CRBM was trained using the original data set. The bottom kernel corresponds to the case of the rotated data set. The colors seem to suggest that the learned weights in both cases were roughly the same, with the only difference being that the top-right and bottom-left weights are swapped.

```
for i in range(kernel_dims[0]):
    print(W[:, :, 0, i],'\n')
```

[[3.3362691e+00 2.3674216e+00]
 [-1.6785726e-01 1.5885308e-07]]

[[-5.6069160e+00 -3.9531565e-01]
 [-2.1835535e+00 -7.7671564e-07]]

```
for i in range(kernel_dims[0]):
    print(W[:, :, 0, i],'\n')
```

[[3.7490754e+00 -1.3243775e-01]
 [2.1889460e+00 -2.4946399e-05]]

[[-5.7980390e+00 -2.1234174e+00]
 [-2.9494333e-01 -1.2887945e-05]]

Figure 2: Numeric values of learned kernels for original (top) and rotated (bottom) data sets. Color coded ovals are used to match the weights that seemingly swapped positions after training with the rotated data set.

2 Test: Free-energy for non-symmetrized vs symmetrized kernels

Recall that we want the free-energy of the visible layer of the CRBM to match the physical distribution of the model:

$$P(v) = P(\sigma)$$
$$\frac{1}{\mathcal{Z}_{\text{CRBM}}}e^{-F(v)} = \frac{1}{\mathcal{Z}_{\text{phys}}}e^{-E(\sigma)}, \tag{1}$$

which is equivalent to saying that the free-energy $F(v)$ and the physical energy $E(\sigma)$ should be equal, up to an additive constant:

$$E(\sigma) = F(v) + \log\left(\frac{\mathcal{Z}_{\text{CRBM}}}{\mathcal{Z}_{\text{phys}}}\right) \equiv F(v) + C \tag{2}$$

Since the free and physical energies are the same, up to a constant, and the physical energy is invariant under reflection and rotation symmetries, the free-energy should also be invariant under these symmetries.

Still need to perform this test. Go to next section :) –ECD

3 Test: What free-energy do non-symmetrized vs symmetrized kernels learn?

We were thinking that, for reasons that are not clear yet, both the "L"-shaped and the explicitly symmetrized kernels are encoding the same information about the symmetries.

To test this, we trained two CRBMs on the same data set. For one of them, the rotation and reflection symmetrizations were performed explicitly on the kernel at regular intervals in the training loop, while for the other CRBM, the kernel was not explicitly symmetrized. The learned kernels can be seen in Fig. 3. The "L"-shaped kernel is for the case of no explicit symmetrization and the cross-shaped kernel has been explicitly symmetrized for reflection and rotation. The non-zero weights of the symmetrized kernel vs the non-symmetrized one seem to differ by a factor of 2. Interestingly, for the heavy weight, the symmetrized one is roughly doubled. But, for the lighter weights, the symmetrized ones are instead halved.

Fig. 4 shows the free-energies computed for a fixed set of random configurations of the visible layer. The top result is with no explicit symmetrization of the kernel, while the bottom one is with explicit symmetrization. Notice that up to constant factor, the free-energies are the same. Thus, the free-energy is invariant under symmetrization of the kernel. Somehow, the cross and "L"-shaped kernels are encoding the same information.

We are not sure yet if this is a consequence of the simplicity of the Ising model or if both representations are somehow equivalent to each other.
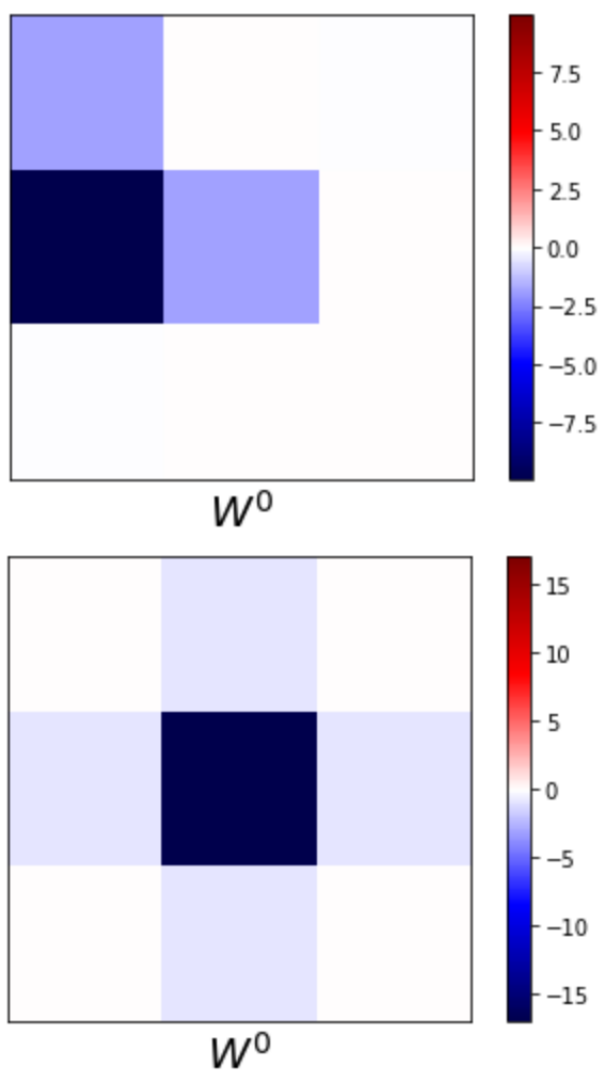
Figure 3: Learned kernels with no explicit symmetrization (top) and with explicit rotation and reflection symmetrization (bottom).

```
crbm.n_log_like(v) - np.min(crbm.n_log_like(v))

<tf.Tensor: shape=(10,), dtype=float32, numpy=
array([0.0000000e+00, 1.7770844e+00, 3.5557137e+00, 5.3354797e+00,
       1.7763557e+00, 1.0795593e-03, 1.7764435e+00, 1.7778015e+00,
       1.7778015e+00, 1.7792435e+00], dtype=float32)>

crbm.n_log_like(v) - np.min(crbm.n_log_like(v))

<tf.Tensor: shape=(10,), dtype=float32, numpy=
array([0.        , 1.7771301, 3.5555954, 5.3320847, 1.7764816, 0.        ,
       1.7764816, 1.7771301, 1.7771225, 1.7768936], dtype=float32)>
```

Figure 4: Learned kernels with no explicit symmetrization (top) and with explicit rotation and reflection symmetrization (bottom).

4 Moving forward

It does not seem like implementing the explicit symmetrization will give much benefit in the case of the zero-field, nearest-neighbor Ising model. Some potential avenues of study could instead be to implement the sampler that Sakib suggested, or to study a more complicated model in which explicitly implementing the symmetries might provide more benefit.