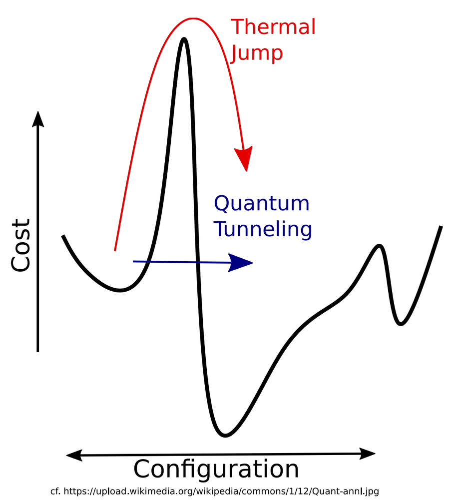
QuAnt: Quantum Annealing with Learnt Couplings

Marcel Seelbach Benkner[1] Maximilian Krahn[2,3] Edith Tretschk[2] Zorah Lähner[1]
Michael Möller[1] Vladislav Golyanik[2]

[1]University of Siegen [2]Max Planck Institute for Informatics, SIC [3]Aalto University

Quantum Annealing

- Quantum annealing as implemented in the D-Wave machines is a metaheuristic to solve quadratic unconstrained binary optimization problems:
$$\underset{x\in\{0,1\}^m}{\operatorname{argmin}}\ x^T A x\ .$$
- Advantageous to simulated annealing for energy landscapes with high, narrow spikes [1]:
- Major progress in experimental realization: D-Wave 2000Q has 2048 solid state qubits and D-Wave Advantage has over 5000 qubits.
- But connectivity restricted to hardware graph: m ≤ 65, 150



Motivation

- Previously the A matrix for the quantum annealer was crafted by hand.
- It is necessary to develop one algorithm per problem setting.
- Deep learning covers multiple problem instances with one setting.

Loss Function

Total loss function: $L(A) = L_{\text{acc}}(A) + \lambda_1 L_{\text{uniqueness}}(A) + \lambda_2 L_{\text{mlp}}$

Accuracy loss:
$$L_{\text{acc}}(A) = x_{\text{gt}}^T A x_{\text{gt}} - x_{\text{min}}^T A x_{\text{min}},$$
where $x_{\text{min}} = \underset{x\in\{0,1\}^m}{\operatorname{argmin}}\, x^T A x$ is the vector with the lowest energy.

Uniqueness loss:
$$L_{\text{uniqueness}}(A) = -|x_{\text{gt}}^T A x_{\text{gt}} - x_{\text{snd}}^T A x_{\text{snd}}|,$$
where x_{snd} is the vector with the second lowest energy.

MLP loss: $L_{\text{mlp}} = \sum_{f\in F} \frac{1}{|f|} ||f||_1$,
where F is the set of all network layer outputs.

Derivative of $L_{\text{acc}}(A)$ with regards to the entries of A:
$$\frac{\partial L_{\text{acc}}(A)}{\partial A_{i,j}} = 2(x_{\text{gt}})_i (x_{\text{gt}})_j - 2(x_{\text{min}})_i (x_{\text{min}})_j - 2\frac{\partial x_{\text{min}}(A)}{\partial A_{i,j}} A x_{\text{min}}$$

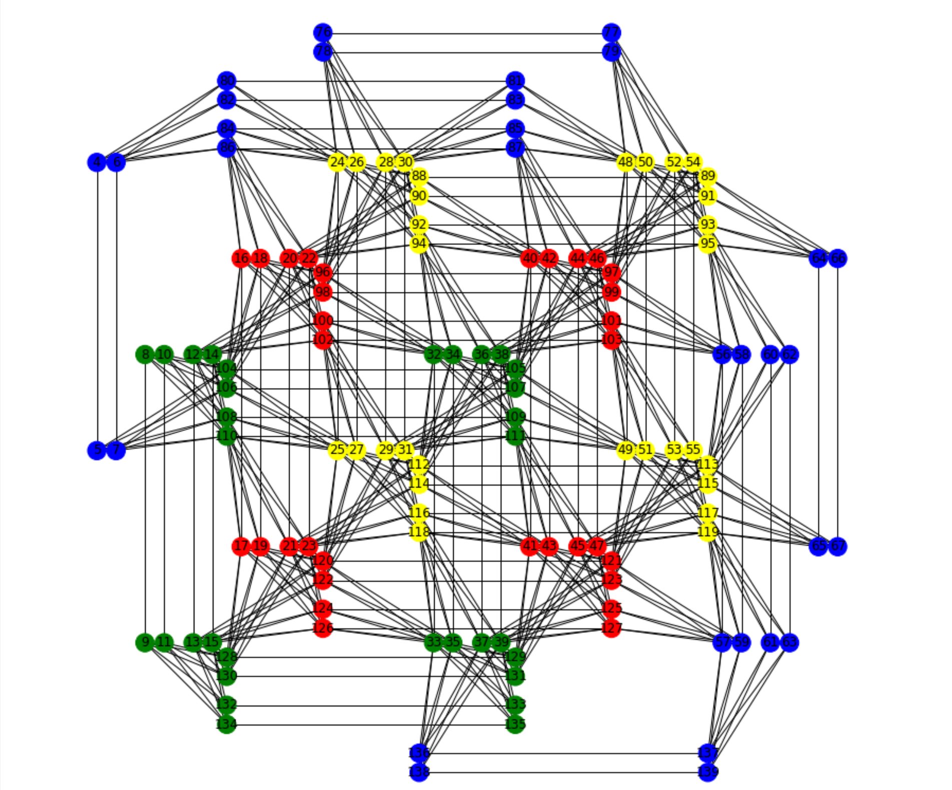
Backpropagation through Quantum annealer is hardly feasible, but the blue term is mostly 0.
Hence we approximate the derivative of $L_{acc}(A)$ with:
$$\frac{\partial L_{\text{acc}}(A)}{\partial A_{i,j}} \approx 2(x_{\text{gt}})_i (x_{\text{gt}})_j - 2(x_{\text{min}})_i (x_{\text{min}})_j$$

Website (Code is available):

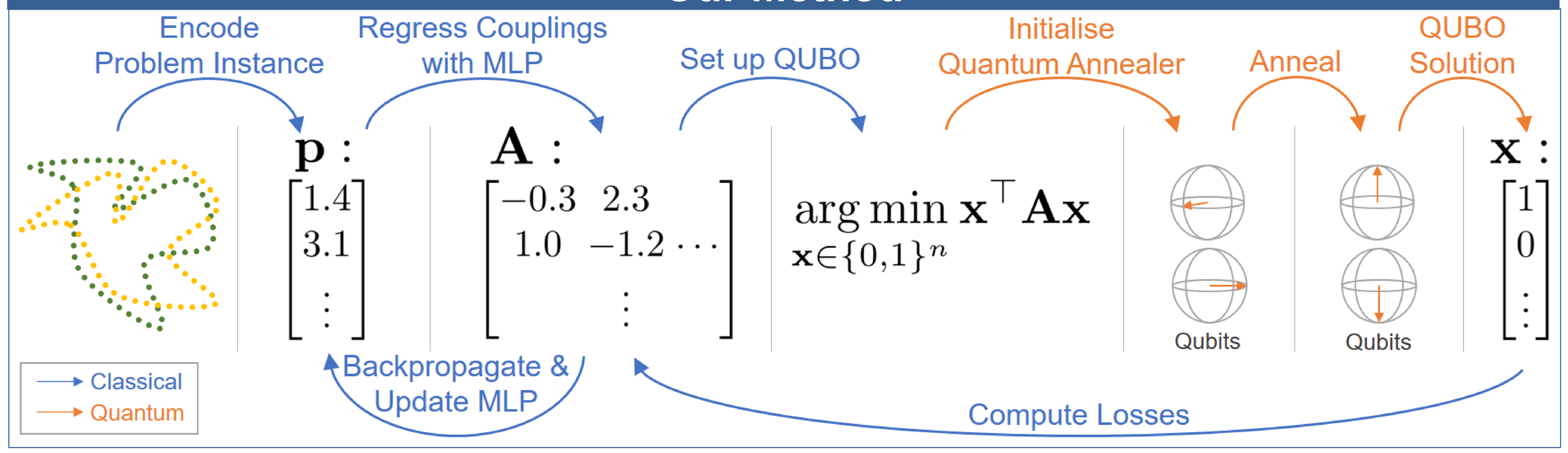
https://4dqv.mpi-inf.mpg.de/QuAnt/

4D and Quantum Vision Group

Hardware Graph



Our Method



Architecture & Problem Encoding

- Networks are either 3 or 5 layer MLPs with hidden dimensions of 32 or 78.
- Hidden layer activation is ReLU and the final activation is a sine function.
- 5 layer setup contains a skip connection from the first to the third layer and uses concatenation for the combination of feature streams.
- Encoding of solutions:
  - **Graph matching:** Binary representation of permutation table
  - **Point set registration**: Binning the angle intervals into equal sized bins
  - **3D rotation estimation:** Concatenating 3 binned angles into one vector
- Baselines:
  - **Diag**: Our method but we set all off-diagonal elements to zero
  - **Pure**: Same network but direct prediction without any QUBO solver
  - **Direct**: Brute Force solution of the input quadratic assignment problem

Results

- **Graph matching** on the Willow Dataset [2] (% of correct permutations):

Ours	Diag	Pure	Direct
69	53	90	**97**

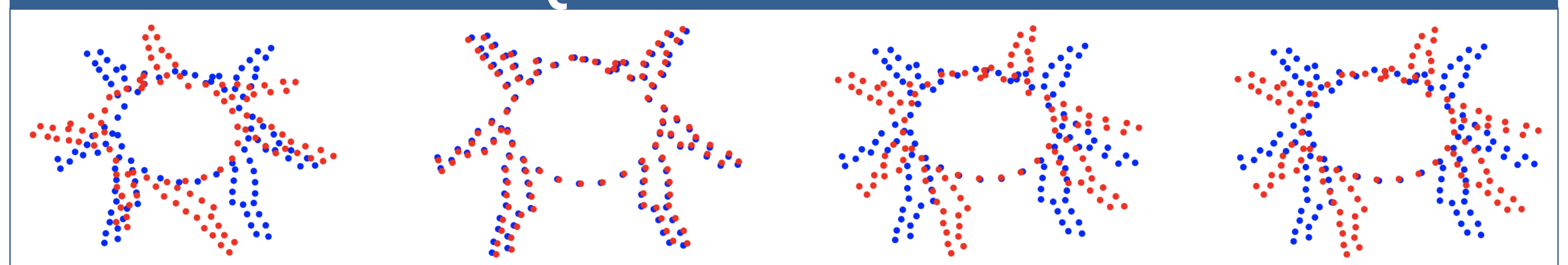
- **Point set registration** on 2D shape dataset [3] (angle difference to ground truth):

	Ours	Diag	Pure
L=3, H=32	8.4 ± 0.8	11.1 ± 1.3	**8.2 ± 1.2**
L=3, H=78	**7.2 ± 1.1**	8.3 ± 0.7	9.3 ± 1.9
L=5, H=32	**8.6 ± 0.5**	10.9 ± 1.2	9.3 ± 1.9
L=5, H=78	**6.8 ± 0.3**	7.7 ± 0.5	11.3 ± 4.5

- **3D rotation estimation** on ModelNet10 [4] (angle difference to ground truth):

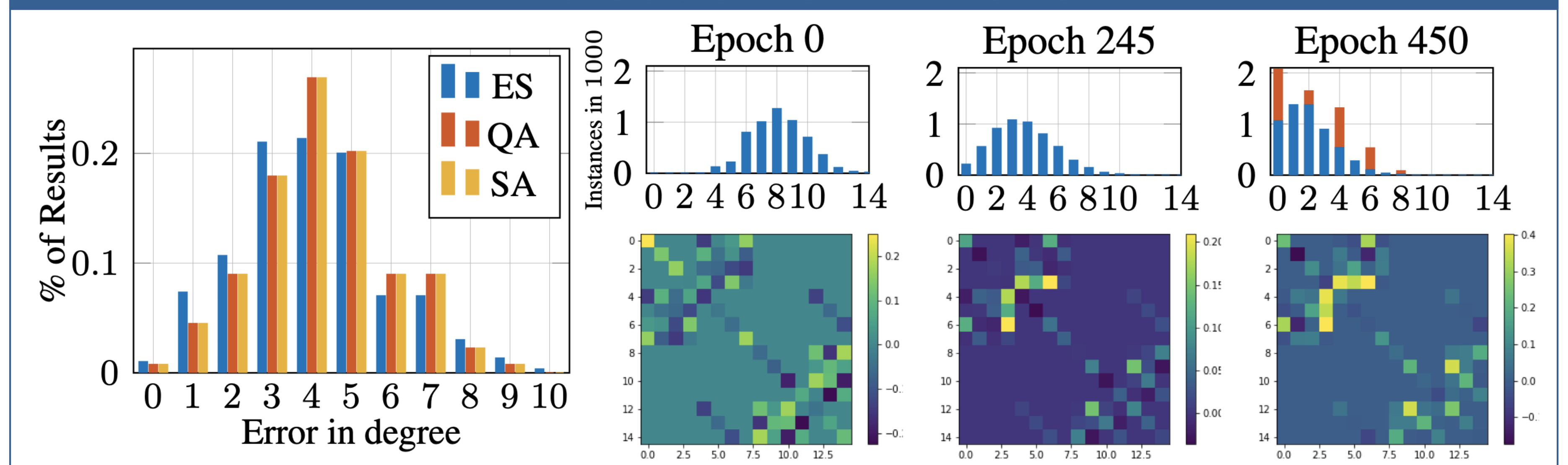
	Ours	Diag	Pure
L=3, H=32	5.9 ± 3.0	**5.4 ± 1.0**	7.9 ± 0.5
L=3, H=78	**4.1 ± 0.5**	5.0 ± 0.3	7.1 ± 0.1
L=5, H=32	**3.7 ± 0.8**	5.0 ± 0.4	16.2 ± 7.1
L=5, H=78	**3.4 ± 0.4**	4.7 ± 0.2	10.1 ± 1.8

Qualitative Results



(a) Initial problem instance (b) Ours with the full loss (c) Ours without L_{MLP} (d) Ours without L_{unique}

Evaluation on D-Wave



(a) Errors on rotation estimation. (b) Evolution of result (top) and coupling matrix (bottom).

- Evaluation with exhaustive search (ES), quantum annealing (QA) and simulated annealing (SA) returns similar results for test data in rotation estimation.
- Right hand side is the evolution during training on a quantum annealer.
- Red bars in right-most histogram indicate projection to valid permutation.

References

- [1] A. Das, et al., Quantum annealing in a kinetically constrained system, 2005
- [2] Cho et al., Learning on Graphs ICCV, 2013
- [3] Carlier at al., The 2d shape structure dataset: A user annotated open access database, Computer & Graphics, 2016
- [4] Wu et al., 3d shapenets: A deep representation for volumetric shapes, CVPR, 2015