

# Experiment results log.

---

## Planned order of scenarios

<b>Scenarios</b>	<b>Parameters</b>	<b>Week</b>
A1	N	March 23 - March 29
A2	B	March 30 - April 5
A3	Datasets	March 30 - April 5
B1	Chain Strength	April 6 - April 12
B2	Embedding	April 13 - April 19
B3	Shots	April 20 - April 26
B4	Annealing	April 27 - May 3

## Actual order of scenarios

<b>Scenarios</b>	<b>Parameters</b>	<b>Week</b>
<b>A1</b>	<b>N</b>	<b>March 23 - March 29</b>
<b>B1</b>	<b>Chain Strength</b>	<b>March 30 - April 5</b>
<b>A2</b>	<b>B</b>	<b>April 6 - April 12</b>
<b>B3</b>	<b>Shots</b>	<b>April 13 - April 19</b>
<b>A2B3</b>	<b>B and Shots</b>	<b>April 20 - April 26</b>
<b>B2</b>	<b>Embedding</b>	<b>April 20 - April 26</b>
<b>B4</b>	<b>Annealing</b>	<b>April 27 - May 3</b>
<b>A3</b>	<b>Datasets</b>	<b>May 4 - May 10</b>

## Sidenotes to research about

- Find what is the maximum **N** value that is supported by dwave

## Scenario A1 - N

We started by experimenting several values of **N**, in order to find the maximum possible value of **N** that could be solved in a reasonable time by the classical solver.

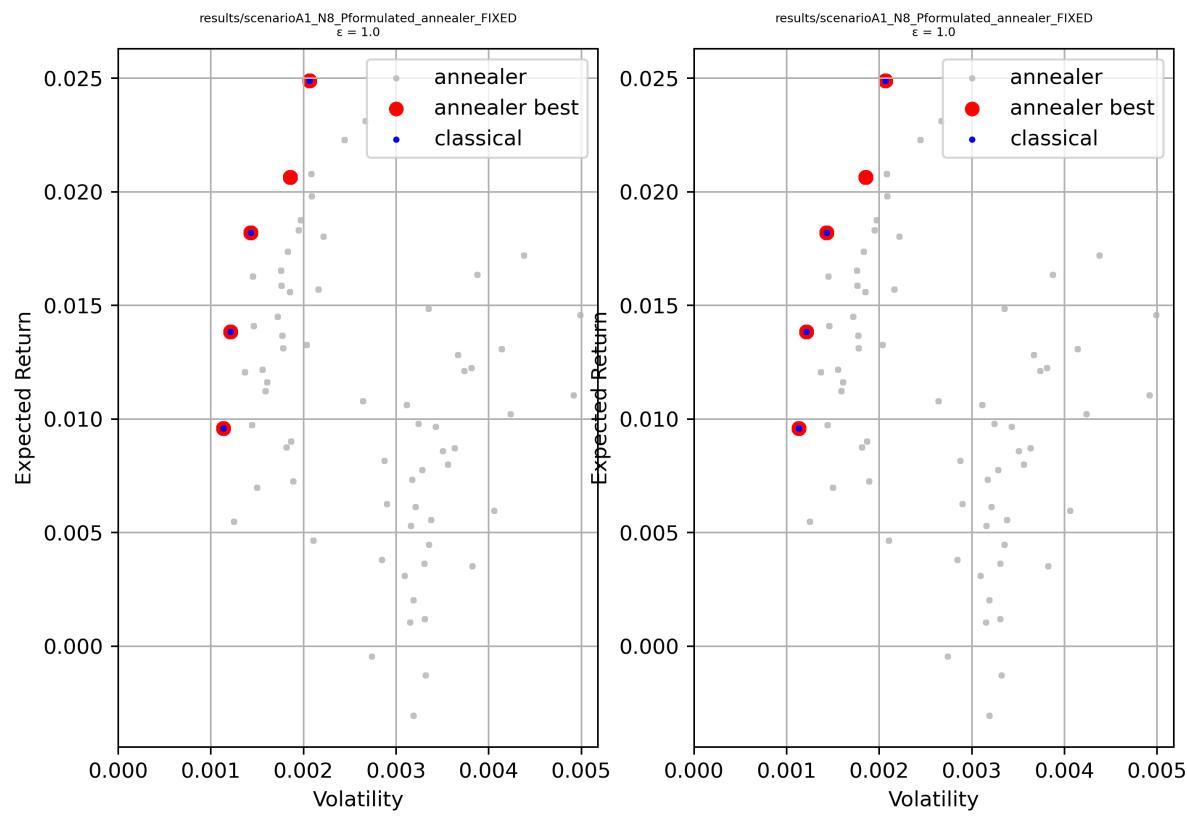
The **N** values are: 8, 16, 32, and 64. P was calculated as  $P = -q * \min\_sigma + \max\_mu$

For this scenario, we used the "diversified" dataset, 1000 shots per execution, and 5 tries. The **q\_values** are listed in the following table:

N	q values
8	0, 11, 20, 54
16	0, 2, 6, 100, 500
32	0, 0.4, 0.9, 2, 3, 9, 100
64	0, 0.2, 0.4, 0.6, 1.1, 1.3, 1.5, 2, 5, 6, 7, 8, 10, 100, 500

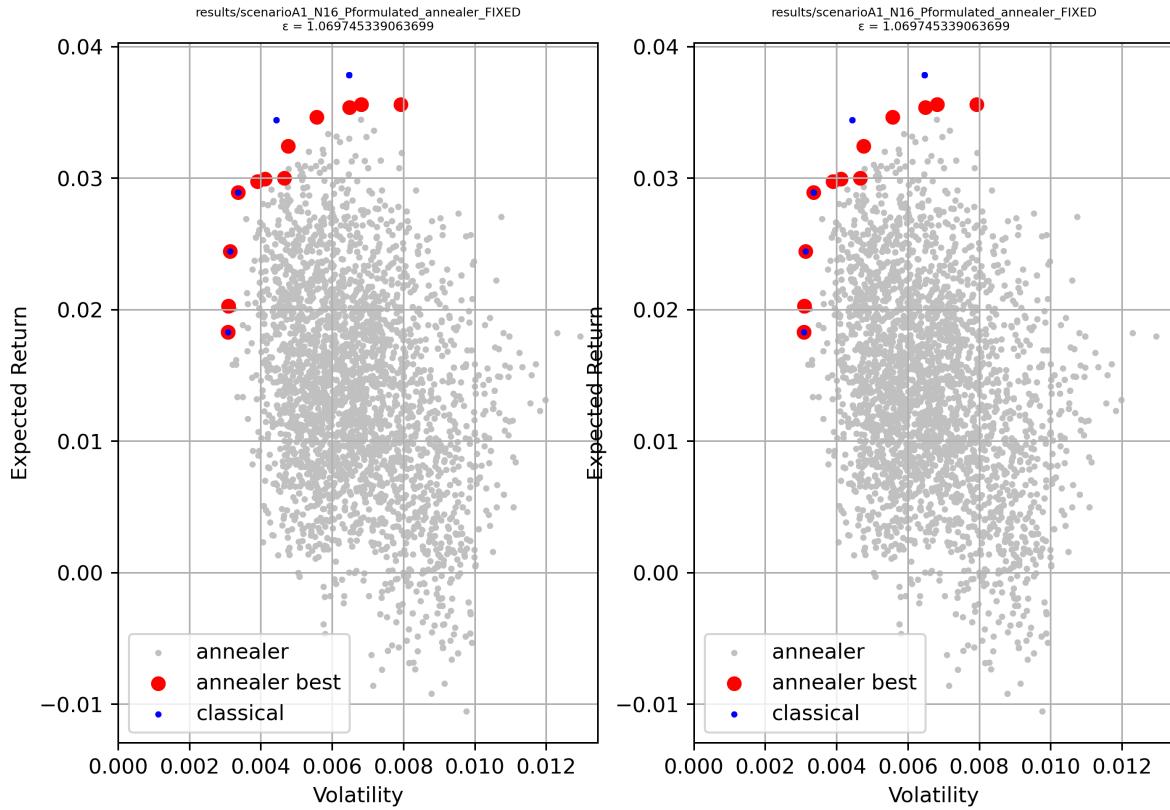
The following images show the first try for each N.

### Epsilon Indicator - scenario1Y2021M04D18h23m07s48



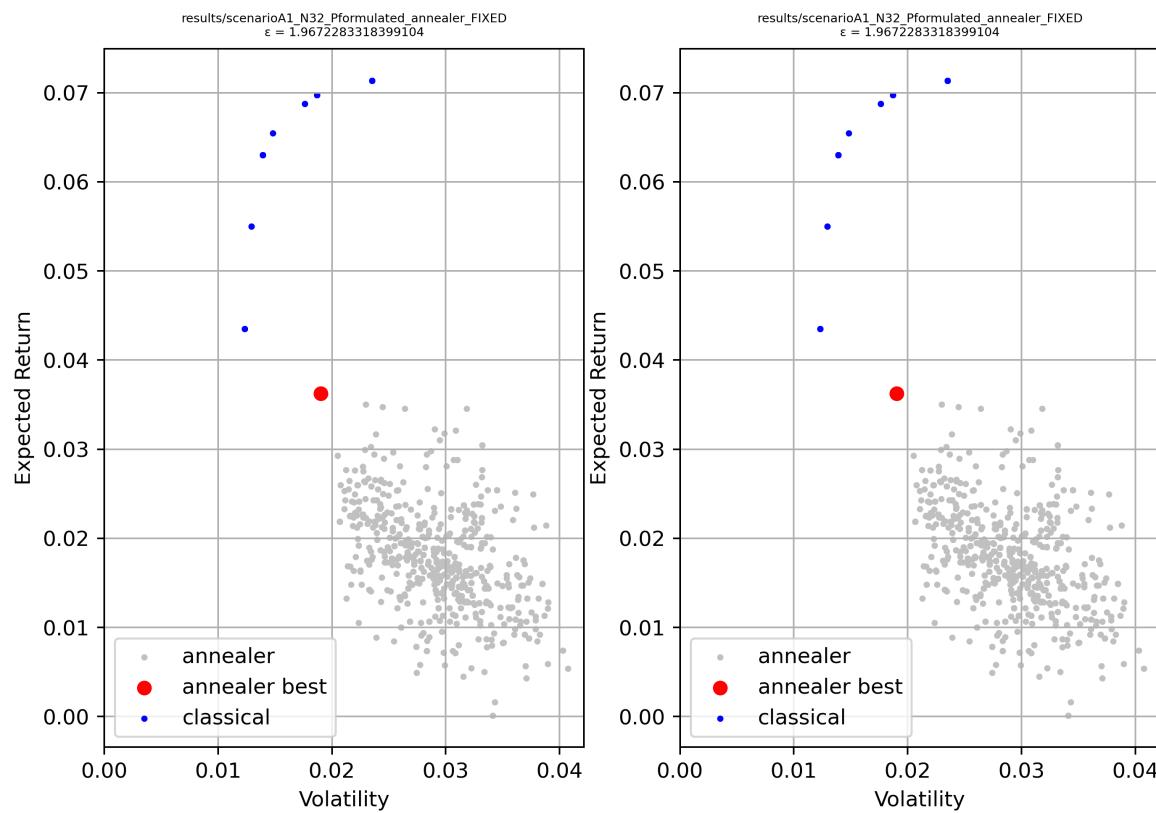
How to interpret: Blue markers are part of the efficient frontier. The epsilon indicator is the minimum factor by which the red set has to be multiplied in the objective so as to weakly dominate Hence, the closer to 1 is the epsilon indicator, the better the red set.

### Epsilon Indicator - scenario1Y2021M04D18h23m08s05

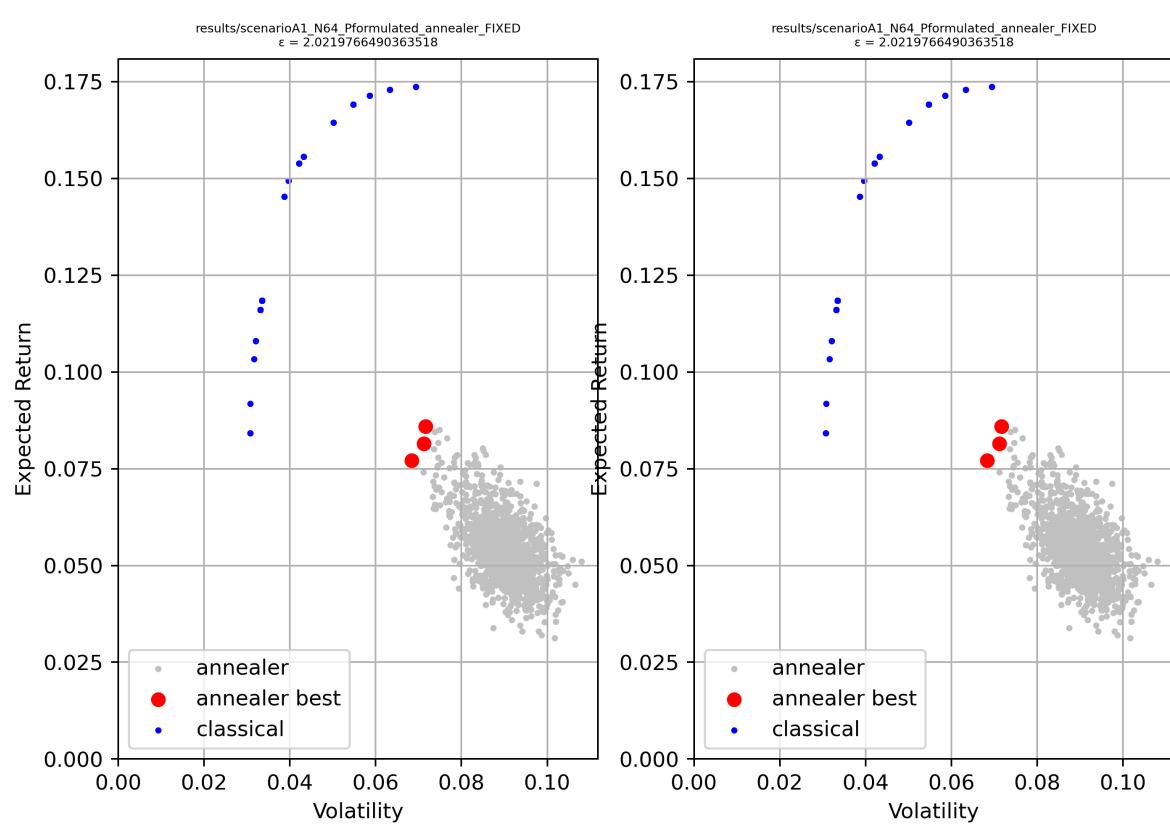


How to interpret: Blue markers are part of the efficient frontier. The epsilon indicator is the minimum factor by which the red set has to be multiplied in the objective so as to weakly dominate Hence, the closer to 1 is the epsilon indicator, the better the red set.

### Epsilon Indicator - scenario1Y2021M04D18h23m08s14

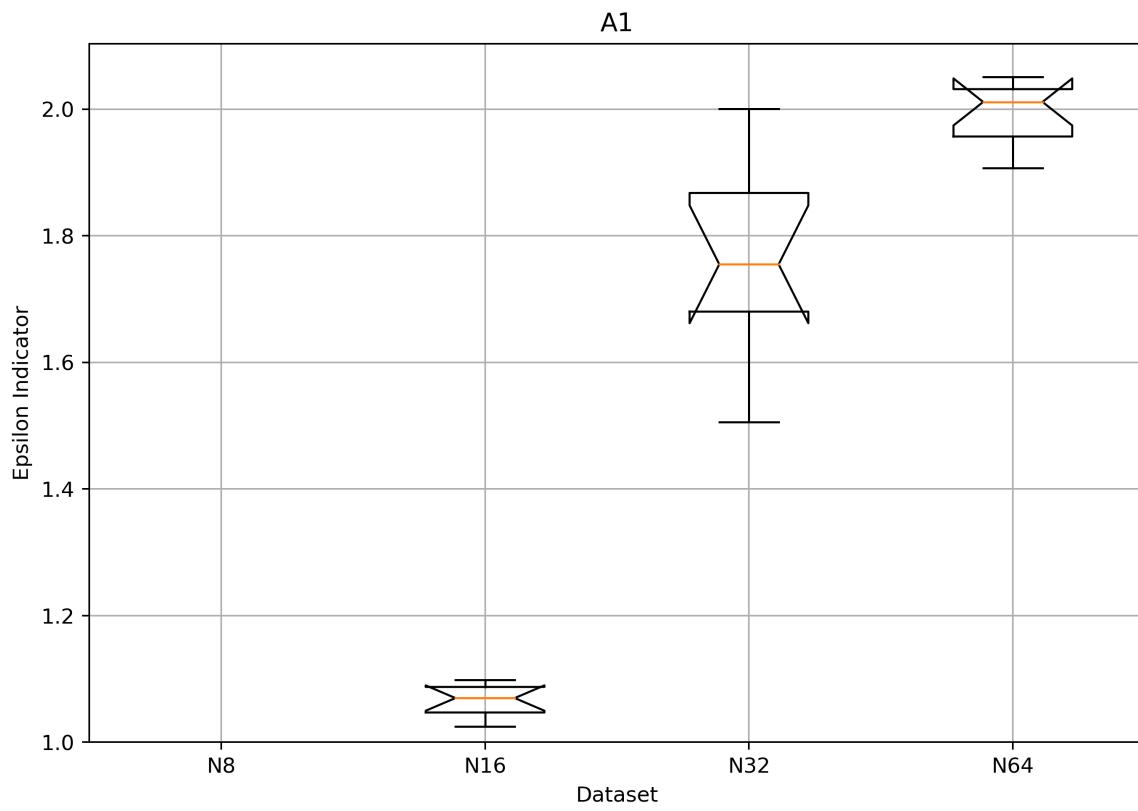


### Epsilon Indicator - scenario1Y2021M04D18h23m08s20



Comparing all the different tries, we have:

## Boxplots - scenarioA3Y2021M05D17h15m56s31



## Hypothesis Testing:

Kruskal-Wallis one-way ANOVA, since we cannot assume the normal distribution of the residuals.

The null hypothesis is that all samples have equal means.

Otherwise, at least two means are different.

We will use a confidence level of 0.05.

For 4 groups with 5 samples each, the critical value for H is 7.377.

Result: `KruskalResult(statistic=17.609472880061105, pvalue=0.0005294252543614154)`

The null hypothesis is rejected.

Post hoc analysis (Conover-Iman Test) needs to be done:

```
Results of Kruskal-Wallis One-way ANOVA:
H statistic: 36.19634076615208
p-value: 6.805708821785085e-08
Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's
test
      N8          N16          N32          N64
N8  1.000000e+00  2.601673e-07  1.219874e-15  1.767614e-20
N16  2.601673e-07  1.000000e+00  6.077688e-08  1.027386e-14
N32  1.219874e-15  6.077688e-08  1.000000e+00  5.001768e-06
N64  1.767614e-20  1.027386e-14  5.001768e-06  1.000000e+00
```

N8 and N16 ARE significantly different!  
 N8 and N32 ARE significantly different!  
 N8 and N64 ARE significantly different!  
 N16 and N32 ARE significantly different!  
 N16 and N64 ARE significantly different!  
 N32 and N64 ARE significantly different!

## Key Takeaways:

As expected, the epsilon indicator increases with the  $N$  value. However, during those executions, dwave's problem inspector warned that the chains were too weak, and that, in the case of  $N=64$ , all samples had broken chains. Based on this warning, we decided to immediately execute scenario B1, changing the original order of scenarios.

## Scenario B1 - Chain Strength

Looking at the fraction of chain breaks in Scenario A1, we know that on average each sample had almost a third ( $0.31$ ) of its chains broken when  $N=32$ . This fraction increases to over half ( $0.54$ ) when  $N=64$ ! Those values are very high and are another clue that the chain strength needs to be adjusted, especially for those values of  $N$ .

A good starting value for the chain strength is the maximum absolute value (`maxAbs`) of the QUBO matrix. However, this is not always the most optimal value. We need to test several values based on this initial value. By testing those values, we can find a value near the sweet spot between the probability that the chains are intact and the probability of finding optimal values. Refer to:

[https://www.dwavesys.com/sites/default/files/2\\_Wed\\_Am\\_PerfTips.pdf](https://www.dwavesys.com/sites/default/files/2_Wed_Am_PerfTips.pdf)

We have three tables, one for the epsilon indicator, one for the fractions of valid solutions, and one for the average fractions of chain breaks.

Starting with the average fractions of chain breaks (Lower is better):

Chain strength	N8	N16	N32	N64
default value	0.00081	0.01153	0.31350	0.54426
0.125 * maxAbs	0.00397	0.02741	0.31014	0.38301
0.250 * maxAbs	0.00034	0.00106	0.00170	0.00683
0.375 * maxAbs	0.00006	0.00032	0.00111	0.00453
0.500 * maxAbs	0.00006	0.00026	0.00149	0.00475
0.625 * maxAbs	0.00006	0.00031	0.00112	0.00453
0.750 * maxAbs	0.00006	0.00029	0.00130	0.00454
0.875 * maxAbs	0.00006	<b>0.00017</b>	0.00102	0.00461
1.000 * maxAbs	0.00003	0.00034	<b>0.00100</b>	0.00439
1.125 * maxAbs	<b>0.00000</b>	0.00030	0.00119	<b>0.00401</b>

<b>Chain strength</b>	<b>N8</b>	<b>N16</b>	<b>N32</b>	<b>N64</b>
1.250 * maxAbs	<b>0.00000</b>	0.00042	0.00125	0.00419
1.375 * maxAbs	0.00006	0.00028	0.00108	0.00424
1.500 * maxAbs	0.00009	0.00025	0.00201	0.00430

Next, we obtained the following fractions of valid solutions (Higher is better):

<b>Chain strength</b>	<b>N8</b>	<b>N16</b>	<b>N32</b>	<b>N64</b>
default value	0.877	<b>0.688</b>	0.121	0.094
0.125 * maxAbs	0.001	0.002	0.076	0.205
0.250 * maxAbs	<b>0.934</b>	0.622	<b>0.395</b>	<b>0.243</b>
0.375 * maxAbs	0.848	0.543	0.325	0.220
0.500 * maxAbs	0.781	0.485	0.299	0.186
0.625 * maxAbs	0.703	0.444	0.261	0.172
0.750 * maxAbs	0.665	0.388	0.252	0.170
0.875 * maxAbs	0.630	0.406	0.242	0.163
1.000 * maxAbs	0.598	0.366	0.235	0.151
1.125 * maxAbs	0.594	0.370	0.219	0.148
1.250 * maxAbs	0.556	0.342	0.223	0.129
1.375 * maxAbs	0.540	0.330	0.212	0.136
1.500 * maxAbs	0.512	0.310	0.198	0.138

Finally, we obtained the following epsilon indicators (Lower is better):

<b>Chain strength</b>	<b>N8</b>	<b>N16</b>	<b>N32</b>	<b>N64</b>
default value	1,000	1,070	1,967	2,022
0,125 * maxAbs	1,368	18,844	1,767	1,977
0,250 * maxAbs	1,000	1,075	1,178	1,474
0,375 * maxAbs	1,000	1,057	1,203	1,580
0,500 * maxAbs	1,000	1,099	1,331	1,500
0,625 * maxAbs	1,000	1,098	1,269	1,410
0,750 * maxAbs	1,000	1,120	1,429	1,523
0,875 * maxAbs	1,000	1,123	1,430	1,587
1,000 * maxAbs	1,000	1,119	1,250	1,526

<b>Chain strength</b>	<b>N8</b>	<b>N16</b>	<b>N32</b>	<b>N64</b>
1,125 * maxAbs	1,000	1,099	1,142	1,539
1,250 * maxAbs	1,000	1,092	1,355	1,610
1,375 * maxAbs	1,000	1,110	1,352	1,465
1,500 * maxAbs	1,000	1,101	1,345	1,423

To validate such results, this scenario has been repeated for **N=16**, **N=32**, and **N=64**.

<b>Chain strength</b>	<b>N16</b>	<b>N32</b>	<b>N64</b>
default value	1,092	1,739	1,981
0,125 * maxAbs	2,314	1,813	2,185
0,250 * maxAbs	1,098	1,256	1,550
0,375 * maxAbs	1,109	1,336	1,492
0,500 * maxAbs	1,114	1,257	1,502
0,625 * maxAbs	1,110	1,322	1,503
0,750 * maxAbs	1,077	1,299	1,516
0,875 * maxAbs	1,120	1,307	1,489
1,000 * maxAbs	1,141	1,350	1,485
1,125 * maxAbs	1,114	1,327	1,430
1,250 * maxAbs	1,101	1,266	1,549
1,375 * maxAbs	1,169	1,198	1,508
1,500 * maxAbs	1,126	1,325	1,597

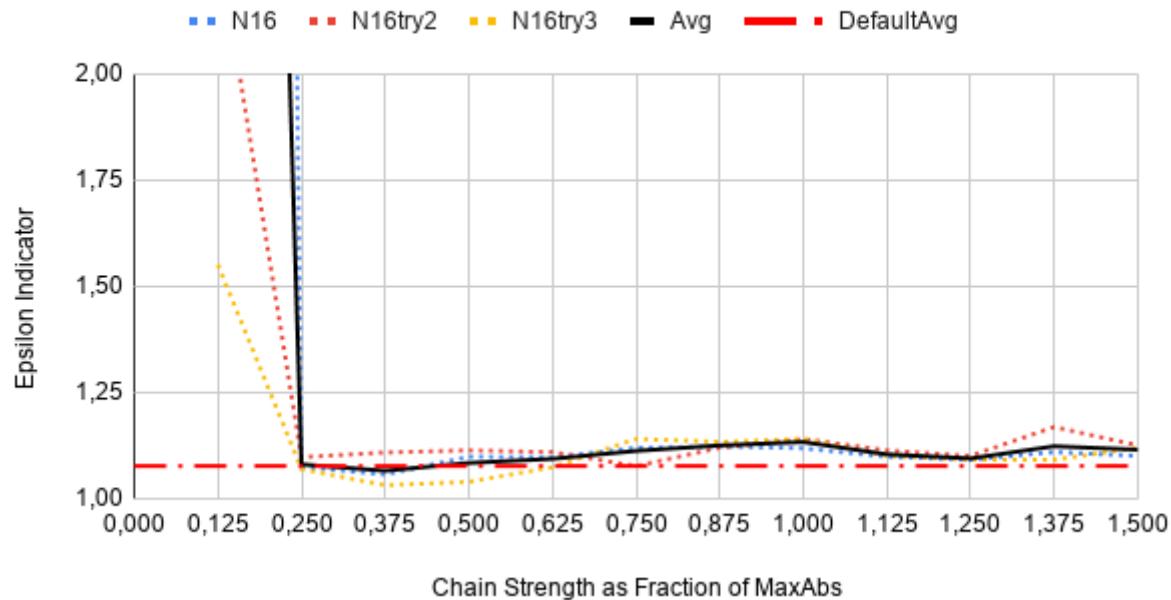
And one more time:

<b>Chain strength</b>	<b>N16</b>	<b>N32</b>	<b>N64</b>
default value	1,070	1,728	1,988
0,125 * maxAbs	1,551	1,760	1,906
0,250 * maxAbs	1,070	1,266	1,462
0,375 * maxAbs	1,032	1,235	1,583
0,500 * maxAbs	1,040	1,325	1,514
0,625 * maxAbs	1,074	1,332	1,551
0,750 * maxAbs	1,141	1,270	1,458
0,875 * maxAbs	1,134	1,229	1,515

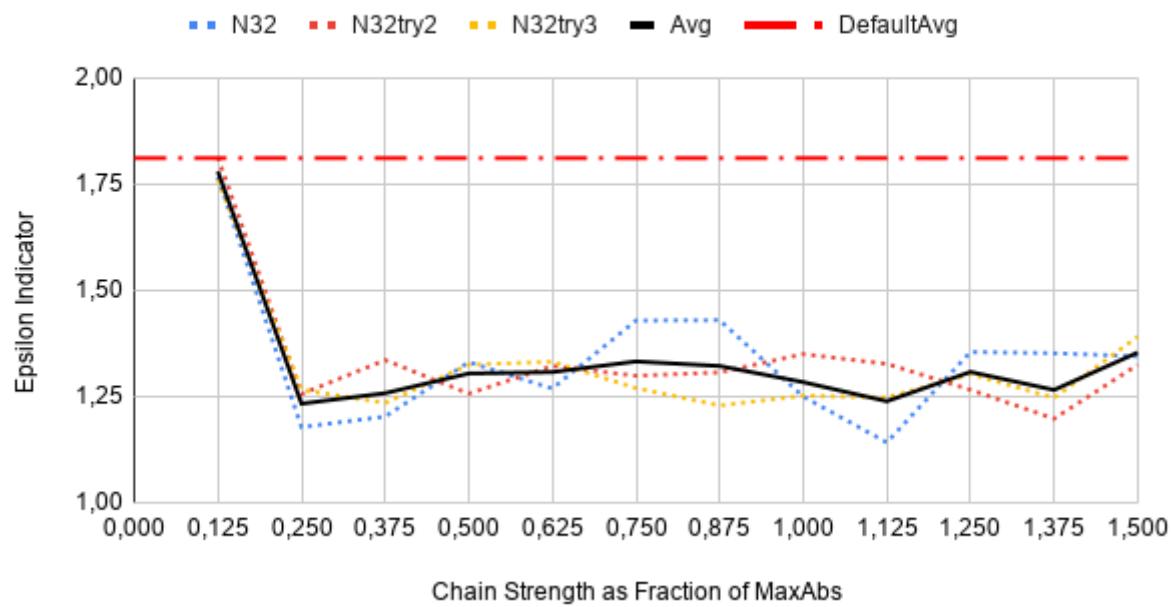
<b>Chain strength</b>	<b>N16</b>	<b>N32</b>	<b>N64</b>
1,000 * maxAbs	1,141	1,252	1,536
1,125 * maxAbs	1,101	1,248	1,547
1,250 * maxAbs	1,092	1,303	1,523
1,375 * maxAbs	1,092	1,247	1,560
1,500 * maxAbs	1,120	1,391	1,519
5,000 * maxAbs	1.177	1,297	1,627

The results are summarized in the following charts.

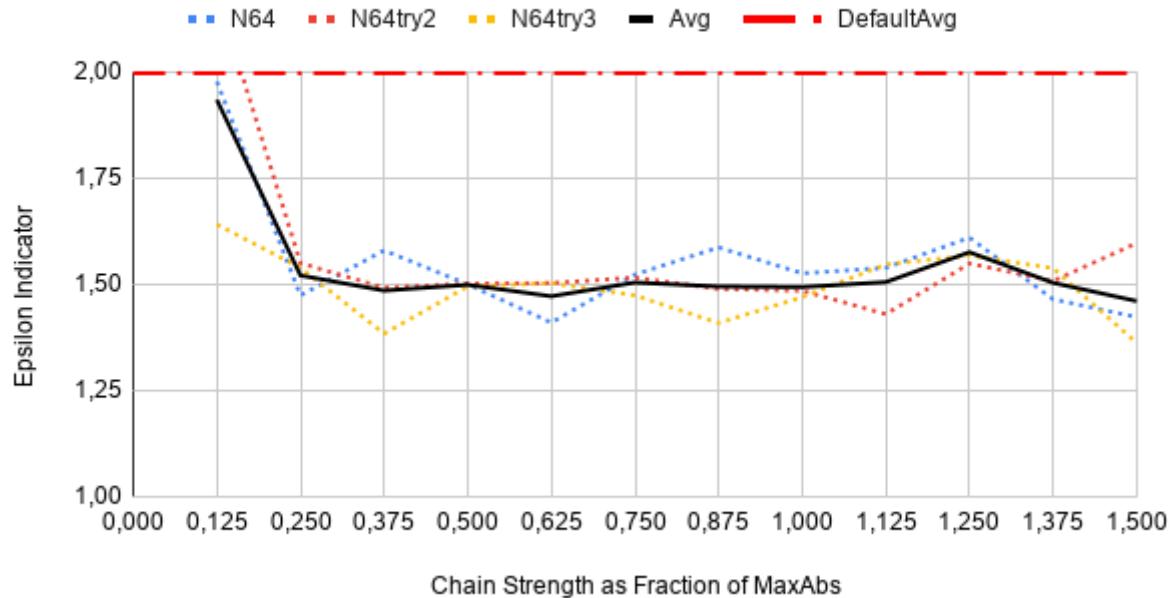
N = 16



N = 32



$N = 64$



### Key Takeaways:

Looking at the results, we notice that the impact of any change to the chain strength is higher for higher values of  $N$ .

It also becomes clear that, especially for higher values of  $N$ , the default chain strength is far from being the best value. It seems that for higher values of  $N$ , the farther is the default chain strength value from the best value.

Another thing that also becomes clear is that the fractions of chain breaks and valid solutions are not directly synonymous with the quality of the solutions.

For the case of  $N=8$ , every try gave a perfect score of  $1.000$ .

For the case  $N=16$ , the epsilon values are so similar that they fall under the margin of variation. Thus we cannot place conclusions based on these results. (Note: in this case, the default strength is always the best!)

There is an exception for both cases of  $N=8$  and  $N=16$ . When  $\text{chain\_strength} = 0.125 * \text{maxAbs}$  there is a high fraction of chain breaks and almost no samples are valid solutions. Thus, for this value of chain strength, the results are very bad.

This behavior is also noticeable for  $N=32$  and  $N=64$ , that present a relatively high epsilon indicator with this chain strength.

It seems that, after this very weak chain strength, the following values of chain strength rapidly attain the lowest epsilon indicators registered, with a very slow climb afterwards.

In the end, the results suggest that it is okay to choose any value that is part of the slow climb. However, from theory, we know that we should avoid any value over  $1.000 * \text{maxAbs}$ , since it scales down the problem.

Therefore, for all  $N$  values, a safe range seems to be between  $0.250 * \text{maxAbs}$  and  $1.000 * \text{maxAbs}$ .

Based on those findings, the case N=8 will not be tested in the remaining scenarios, since the annealer already achieved optimality.

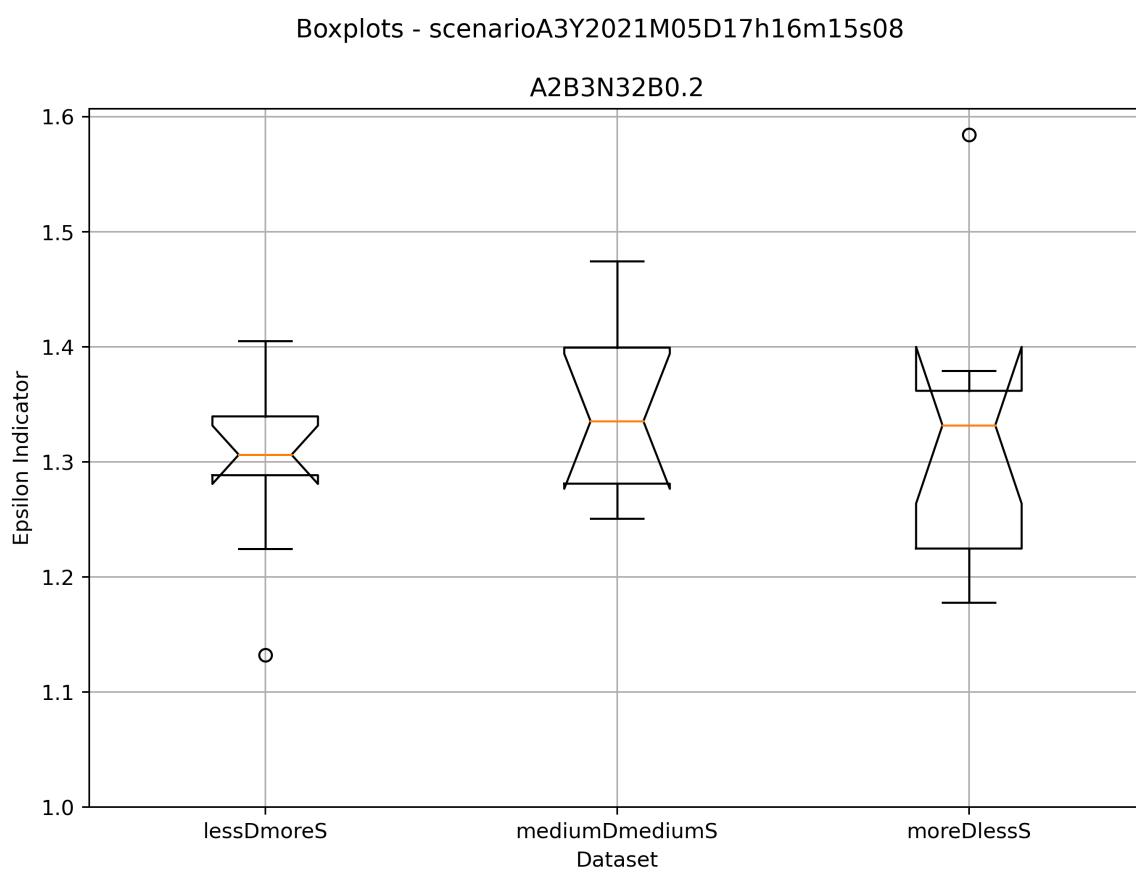
## Scenario A2 and B3 - B and Shots

2 factors: B and Shots

"B" factor has three levels: Small Budget, Medium Budget and Large Budget (fractions 0.2, 0.5, and 0.8, respectively).

"Shots" factor has three levels: Less directions and More shots per direction, Medium directions and Medium shots per direction, More directions and Less shots per direction (codenamed lessDmoreS, mediumDmediumS, and moreDlessS, respectively).

Starting with N=32:



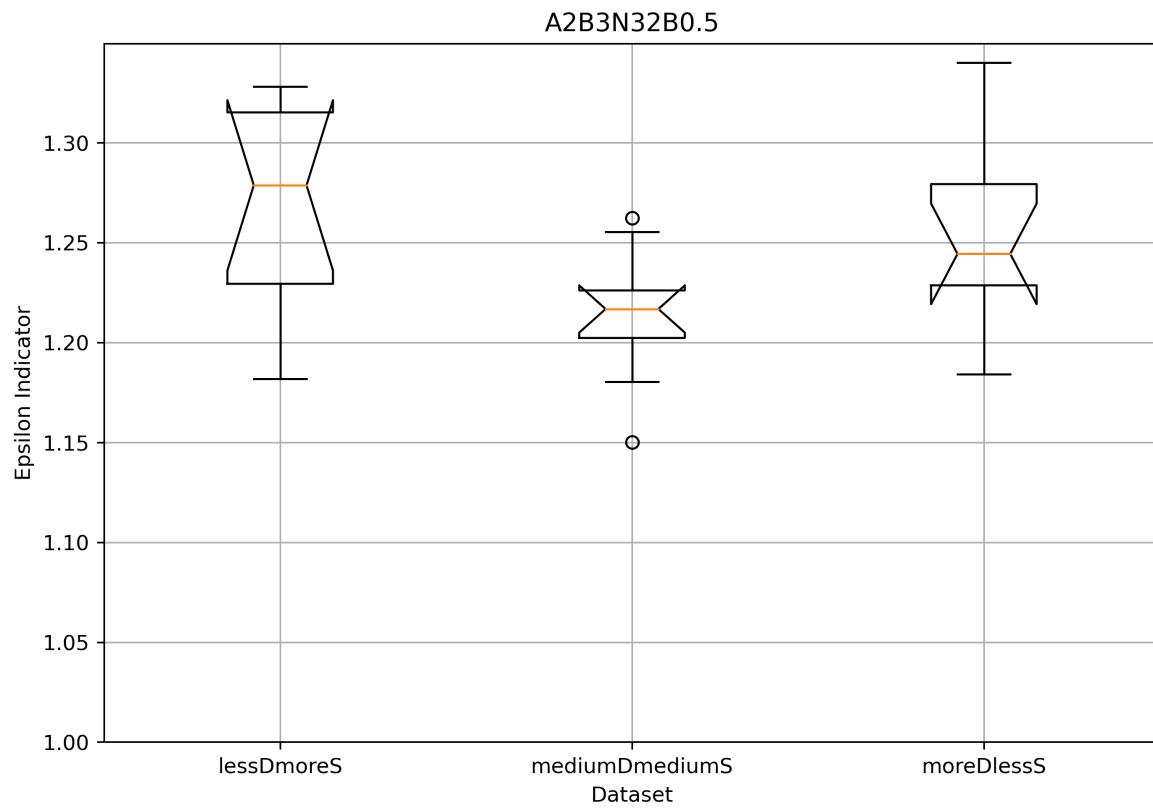
Results of Kruskal-Wallis One-way ANOVA:

H statistic: 1.02451612903225

p-value: 0.5991411508220548

The null hypothesis was not rejected!

## Boxplots - scenarioA3Y2021M05D17h16m17s58



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 6.080000000000125

p-value: 0.047834889494198084

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

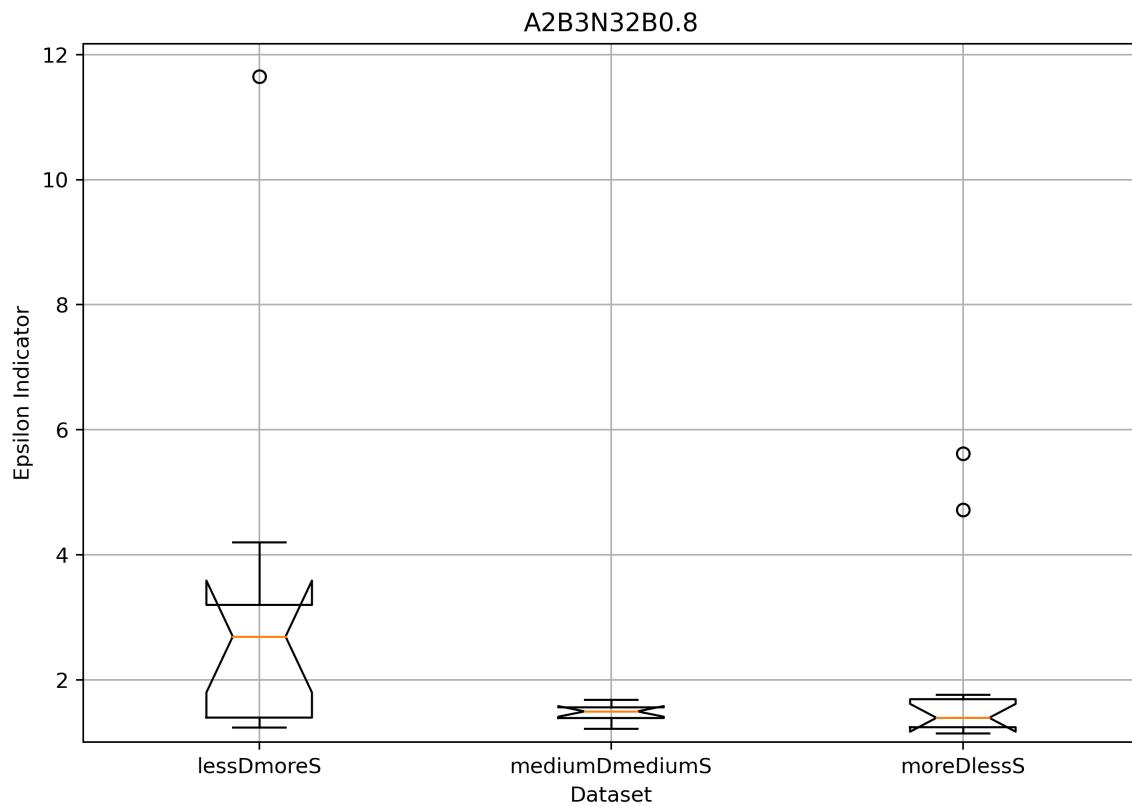
	lessDmoreS	mediumDmediumS	moreDlessS
lessDmoreS	1.000000	0.045701	1.000000
mediumDmediumS	0.045701	1.000000	0.215099
moreDlessS	1.000000	0.215099	1.000000

lessDmoreS and mediumDmediumS ARE significantly different!

lessDmoreS and moreDlessS are NOT!

mediumDmediumS and moreDlessS are NOT!

## Boxplots - scenarioA3Y2021M05D17h16m24s52



Results of Kruskal-Wallis One-way ANOVA:

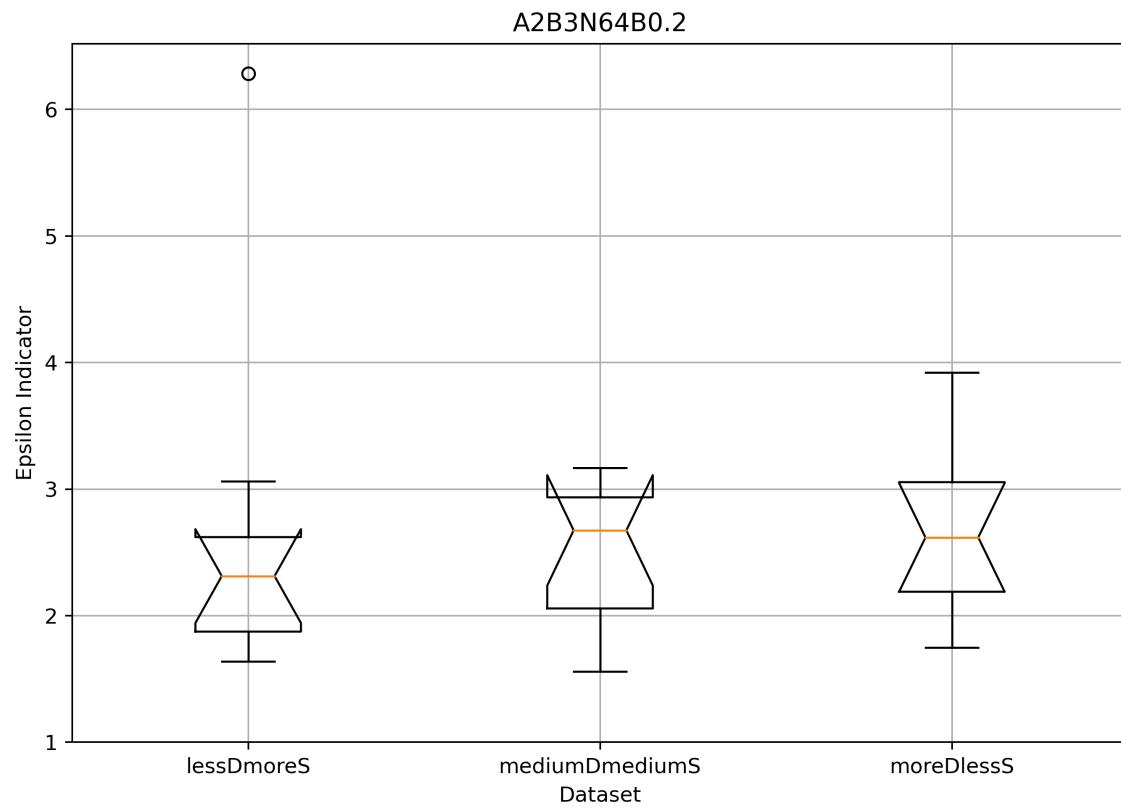
H statistic: 2.6348387096774104

p-value: 0.2678255736803744

The null hypothesis was not rejected!

And for N=64:

## Boxplots - scenarioA3Y2021M05D17h16m49s22



How to interpret: The epsilon indicator is the minimum factor by which the annealer set has to be multiplied in the objective so as to weakly dominate the classical set.  
Hence, the closer to 1 is the epsilon indicator, the better the annealer set.

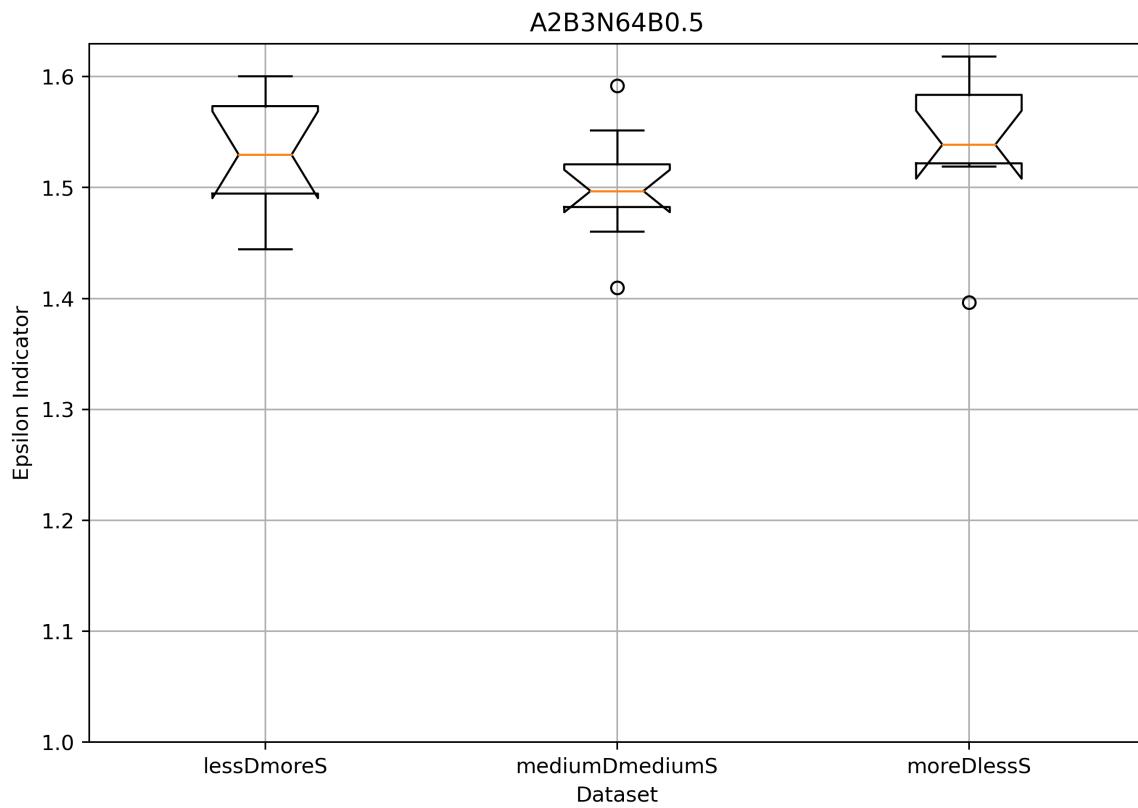
## Results of Kruskal-Wallis One-way ANOVA:

H statistic: 1.2619354838709569

p-value: 0.5320766388987637

The null hypothesis was not rejected!

## Boxplots - scenarioA3Y2021M05D17h16m55s41



How to interpret: The epsilon indicator is the minimum factor by which the annealer set has to be multiplied in the objective so as to weakly dominate the classical set.  
Hence, the closer to 1 is the epsilon indicator, the better the annealer set.

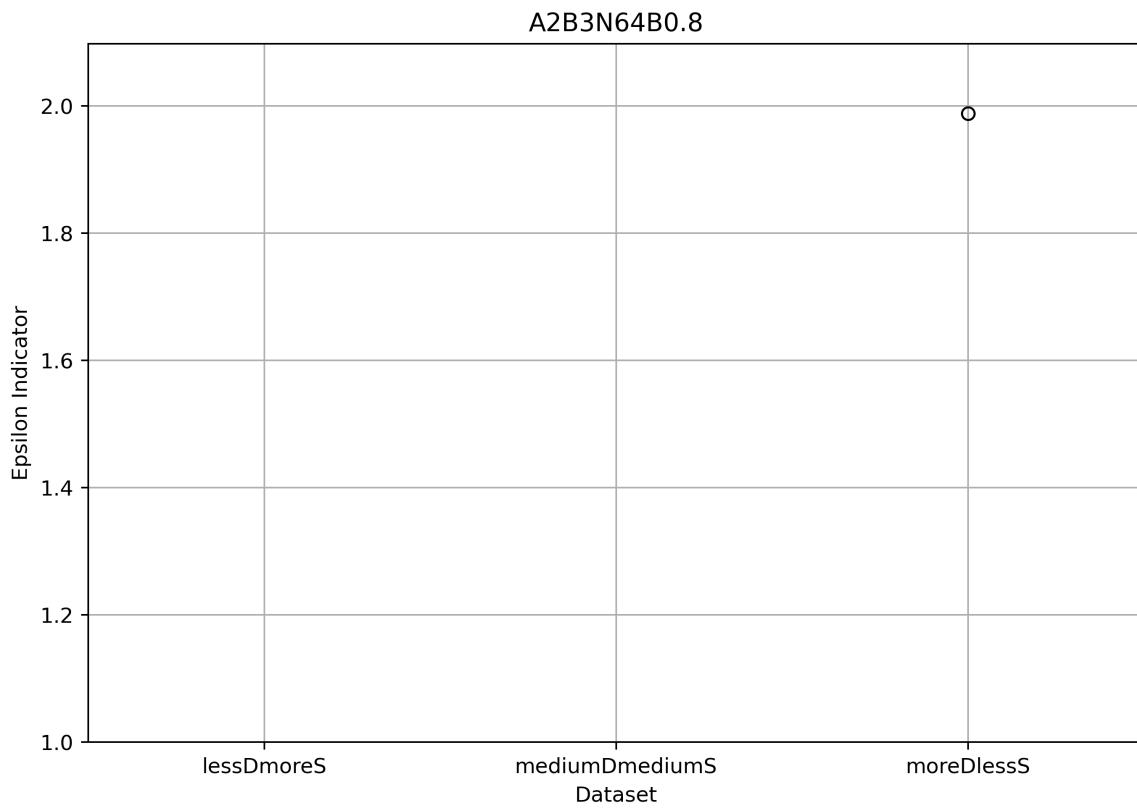
## Results of Kruskal-Wallis One-way ANOVA:

H statistic: 3.8735483870967755

p-value: 0.14416825938986216

The null hypothesis was not rejected!

## Boxplots - scenarioA3Y2021M05D17h16m52s43



How to interpret: The epsilon indicator is the minimum factor by which the annealer set has to be multiplied in the objective so as to weakly dominate the classical set.  
Hence, the closer to 1 is the epsilon indicator, the better the annealer set.

## Results of Kruskal-Wallis One-way ANOVA:

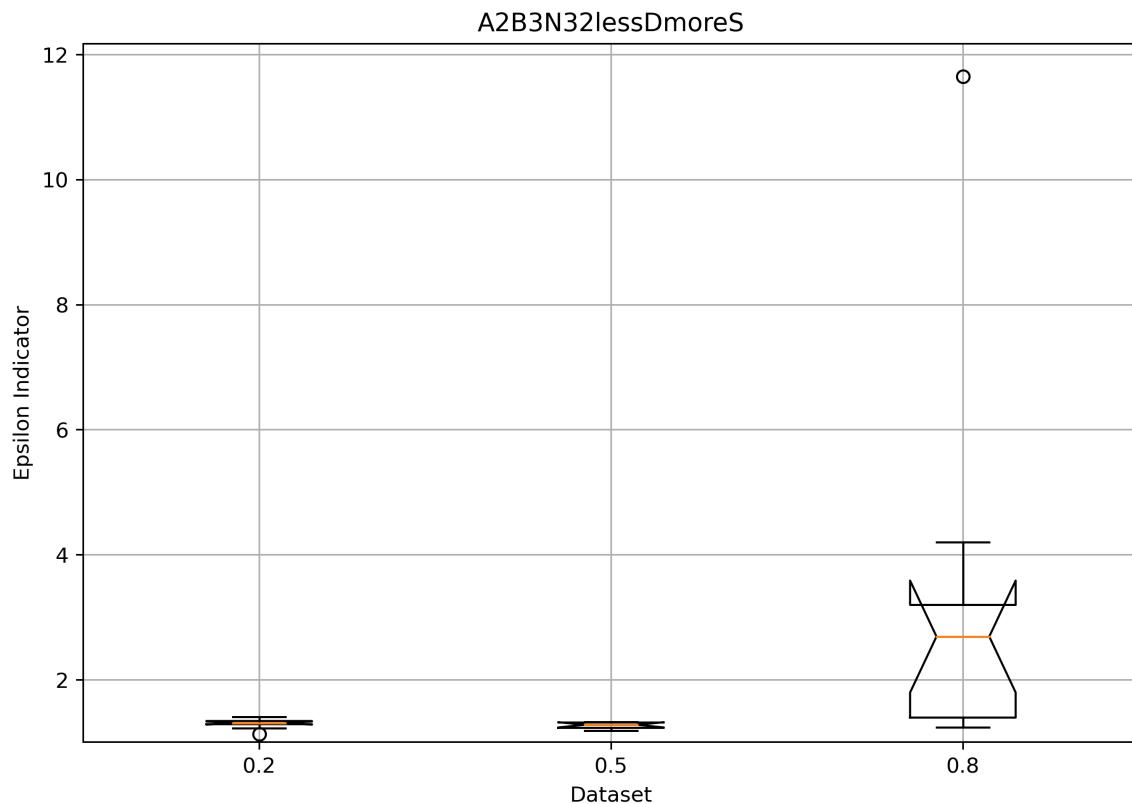
H statistic: 1.999999999999427

p-value: 0.367879441171453

The null hypothesis was not rejected!

Now, let's look at the budget, starting with N32:

## Boxplots - scenarioA3Y2021M05D17h16m58s16



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 13.223225806451609

p-value: 0.0013446615904149253

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

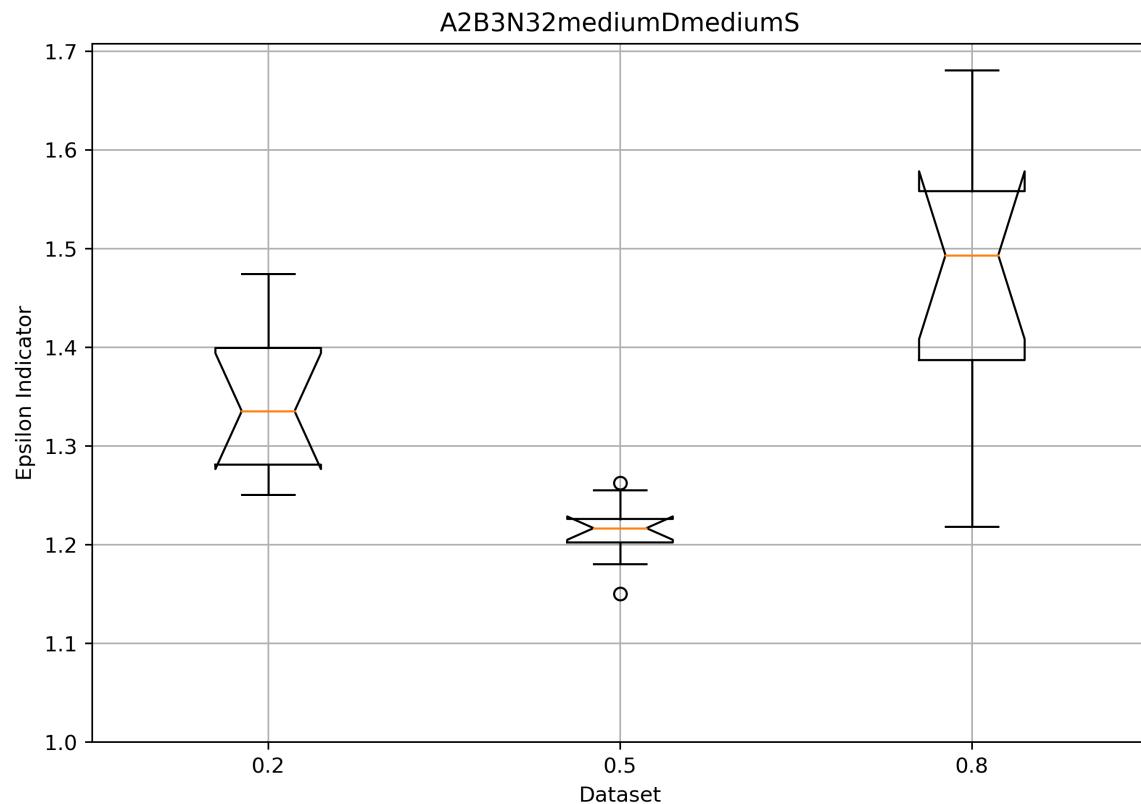
	0.2	0.5	0.8
0.2	1.000000	0.726036	0.006507
0.5	0.726036	1.000000	0.000278
0.8	0.006507	0.000278	1.000000

0.2 and 0.5 are NOT!

0.2 and 0.8 ARE significantly different!

0.5 and 0.8 ARE significantly different!

## Boxplots - scenarioA3Y2021M05D17h17m01s09



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 18.691612903225817

p-value: 8.733087853773252e-05

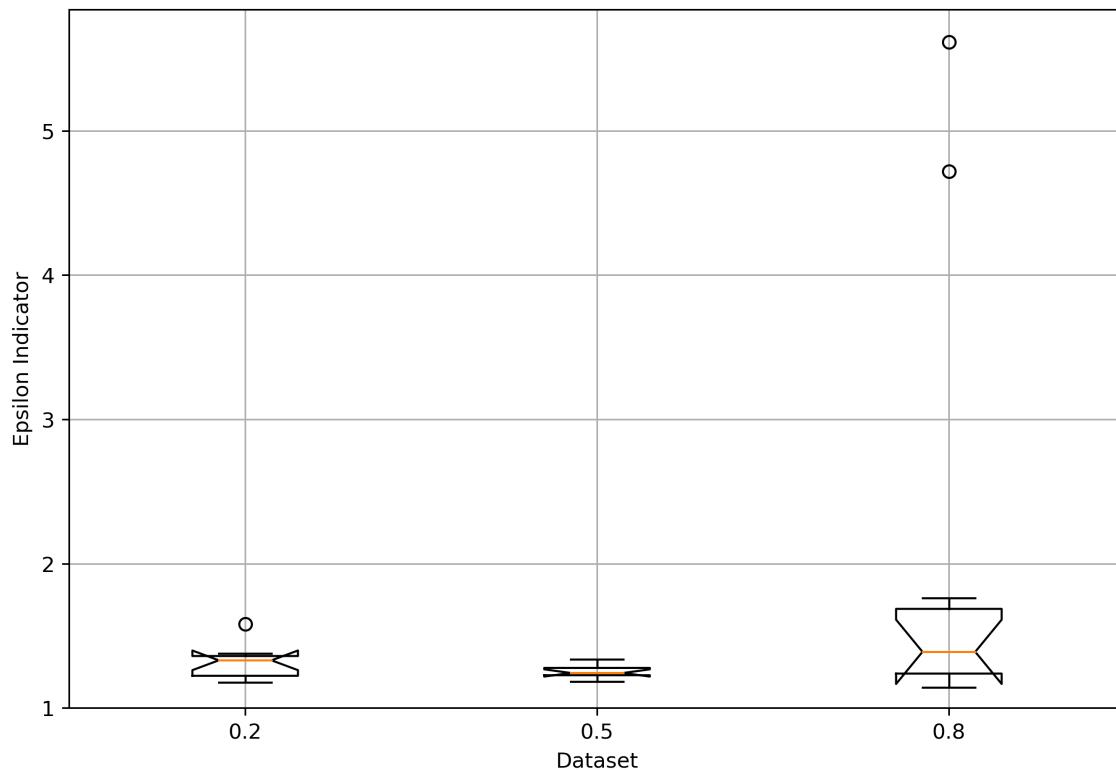
Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	0.2	0.5	0.8
0.2	1.000000	2.648012e-04	9.609807e-02
0.5	0.000265	1.000000e+00	6.732613e-07
0.8	0.096098	6.732613e-07	1.000000e+00

0.2 and 0.5 ARE significantly different!  
0.2 and 0.8 are NOT!  
0.5 and 0.8 ARE significantly different!

## Boxplots - scenarioA3Y2021M05D17h17m03s36

A2B3N32moreDlessS



How to interpret: The epsilon indicator is the minimum factor by which the annealer set has to be multiplied in the objective so as to weakly dominate the classical set.  
Hence, the closer to 1 is the epsilon indicator, the better the annealer set.

Results of Kruskal-Wallis One-way ANOVA:

H statistic: 2.8980645161290397

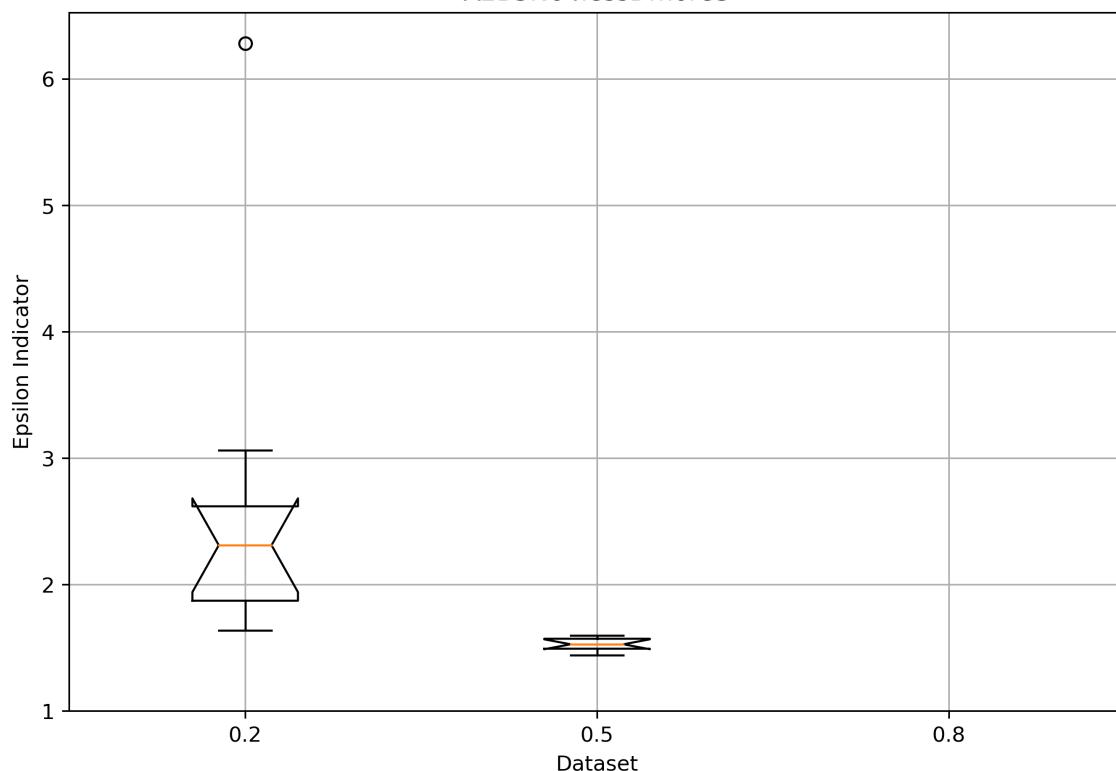
p-value: 0.2347974014742501

The null hypothesis was not rejected!

And for N=64:

## Boxplots - scenarioA3Y2021M05D17h17m08s19

A2B3N64lessDmoreS



How to interpret: The epsilon indicator is the minimum factor by which the annealer set has to be multiplied in the objective so as to weakly dominate the classical set.  
Hence, the closer to 1 is the epsilon indicator, the better the annealer set.

## Results of Kruskal-Wallis One-way ANOVA:

H statistic: 26.789838337182452

p-value: 1.5228618937077617e-06

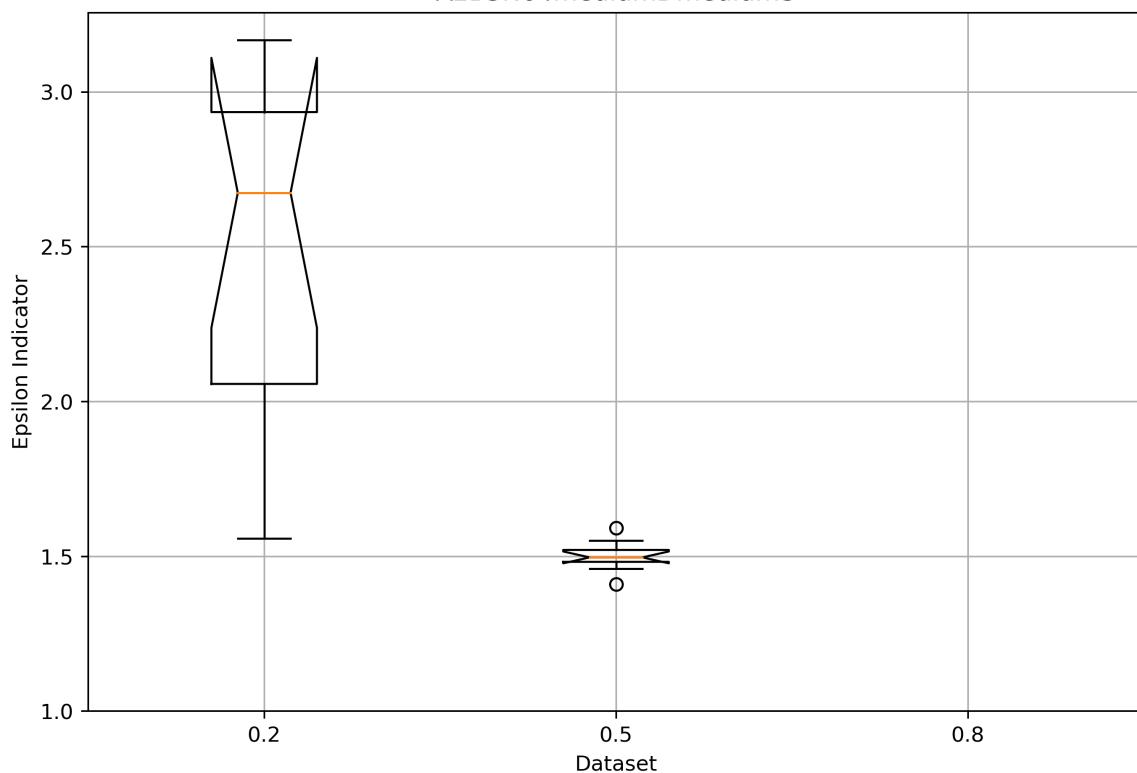
Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	0.2	0.5	0.8
0.2	1.000000e+00	3.502739e-09	3.502739e-09
0.5	3.502739e-09	1.000000e+00	3.825579e-16
0.8	3.502739e-09	3.825579e-16	1.000000e+00

0.2 and 0.5 ARE significantly different!  
0.2 and 0.8 ARE significantly different!  
0.5 and 0.8 ARE significantly different!

## Boxplots - scenarioA3Y2021M05D17h17m09s56

A2B3N64mediumDmediumS



Results of Kruskal-Wallis One-way ANOVA:

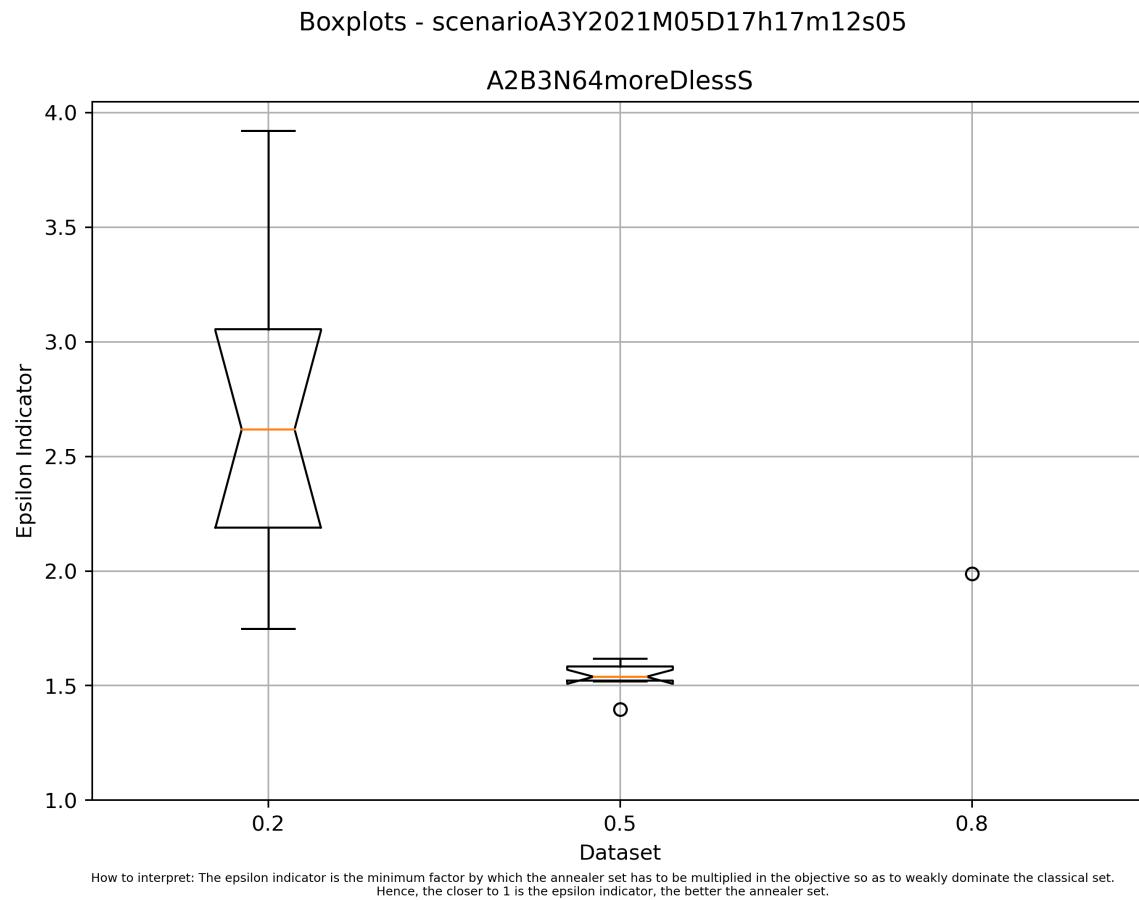
H statistic: 26.52461893764435

p-value: 1.7388102050638365e-06

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	0.2	0.5	0.8
0.2	1.000000e+00	1.647582e-08	9.039993e-09
0.5	1.647582e-08	1.000000e+00	1.776432e-15
0.8	9.039993e-09	1.776432e-15	1.000000e+00

0.2 and 0.5 ARE significantly different!  
0.2 and 0.8 ARE significantly different!  
0.5 and 0.8 ARE significantly different!



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 24.56283428571427

p-value: 4.637119681113234e-06

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	0.2	0.5	0.8
0.2	1.000000e+00	6.834875e-07	3.720831e-05
0.5	6.834875e-07	1.000000e+00	5.155262e-12
0.8	3.720831e-05	5.155262e-12	1.000000e+00

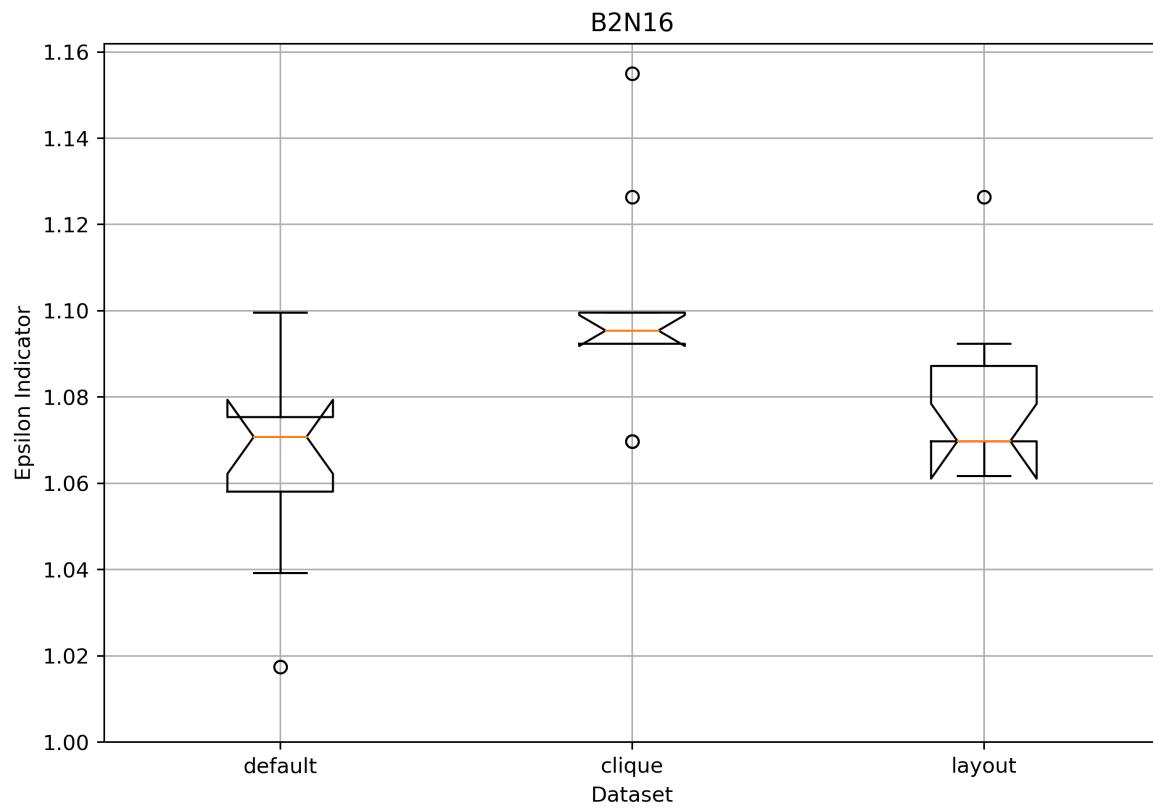
0.2 and 0.5 ARE significantly different!  
0.2 and 0.8 ARE significantly different!  
0.5 and 0.8 ARE significantly different!

Key Takeaways:

## Scenario B2 - Embedding

So far, we used the **general** embedding. D-Wave offers another two embedding options, **clique** and **layout** embeddings. The three options are going to be compared.

## Boxplots - scenarioA3Y2021M05D17h17m27s34



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 6.207853285328522

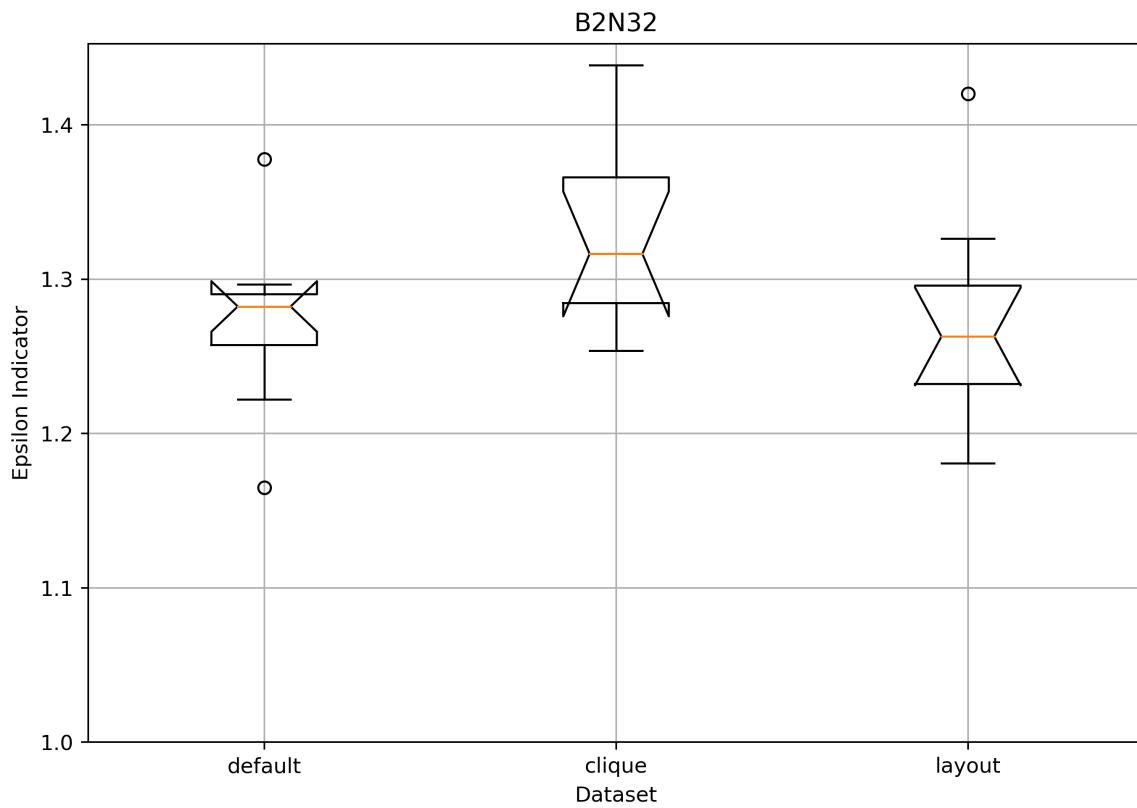
p-value: 0.04487265711539793

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	default	clique	layout
default	1.000000	0.050978	1.000000
clique	0.050978	1.000000	0.139886
layout	1.000000	0.139886	1.000000

default and clique are NOT!  
default and layout are NOT!  
clique and layout are NOT!

## Boxplots - scenarioA3Y2021M05D17h17m25s23



How to interpret: The epsilon indicator is the minimum factor by which the annealer set has to be multiplied in the objective so as to weakly dominate the classical set.  
Hence, the closer to 1 is the epsilon indicator, the better the annealer set.

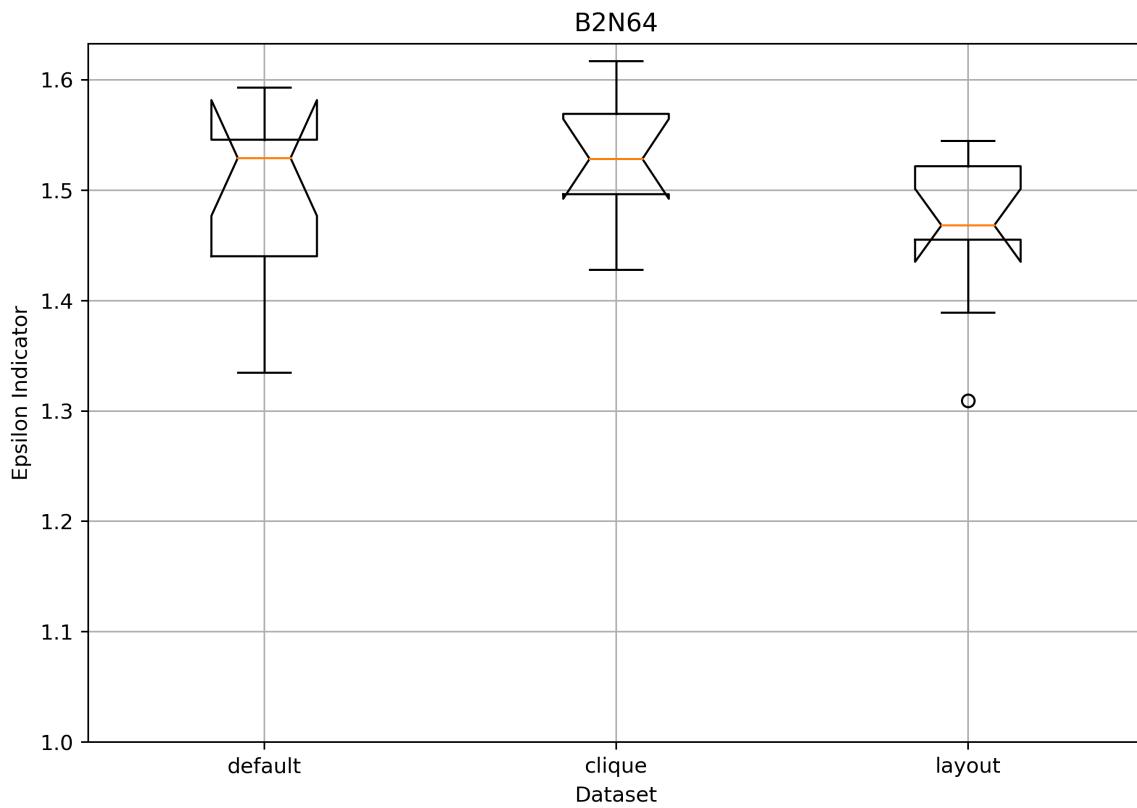
## Results of Kruskal-Wallis One-way ANOVA:

H statistic: 5.46838709677418

p-value: 0.06494636208860985

The null hypothesis was not rejected!

Boxplots - scenarioA3Y2021M05D17h17m23s54



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 3.2309677419354728

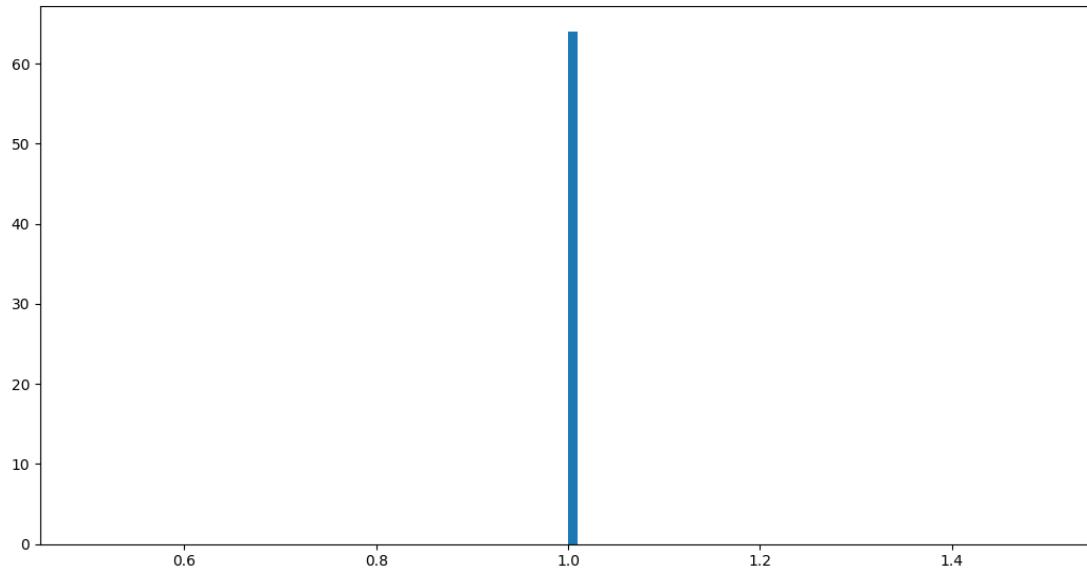
p-value: 0.19879445629913187

The null hypothesis was not rejected!

## Key Takeaways:

In conversations with Jose Pinilla, a Ph.D. student that authored an implementation of a layout-aware embedding, `layout` embedding is much more suited for *sparse* graphs, which is not the case of the POP. In fact, POP usually generates fully connected graphs. However, Jose Pinilla said "if there are clusters of high connectivity, you'll immediately be rewarded with faster results, or a higher chance of at least finding an embedding". I noticed that, in fact, `layout` embedding was much faster than the other two options.

It is interesting that those faster results were also accompanied by better performance. Again, in conversations with Jose Pinilla, he provided me with some code to plot a histogram that let us confirm that the graph is in fact fully connected.



The graph is fully connected, which means that there are no clusters of high connectivity and no speed boost should be expected. **So, why did it have better performance? This is an interesting question that I pose for further research.**

For the remaining scenarios, we will use the `layout` embedding.

## Scenario B4 - Annealing

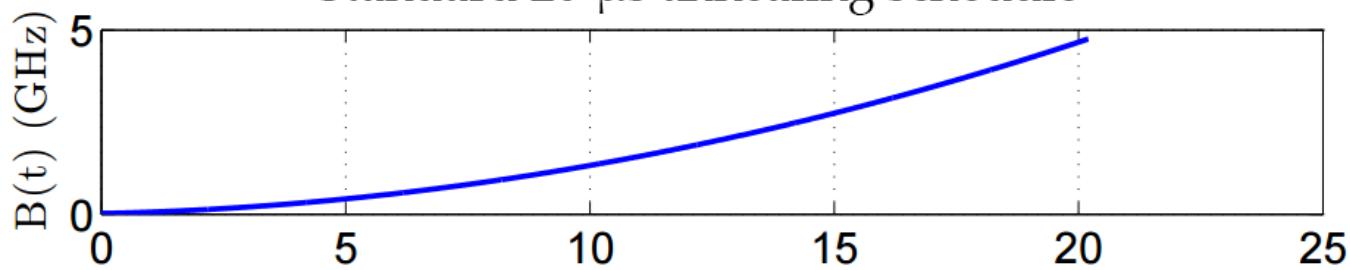
It is time to study the impact of Annealing, if it has any!

So far, we used the `default` annealing strategy. We will study another three common strategies that may provide significant improvements to the annealer performance.

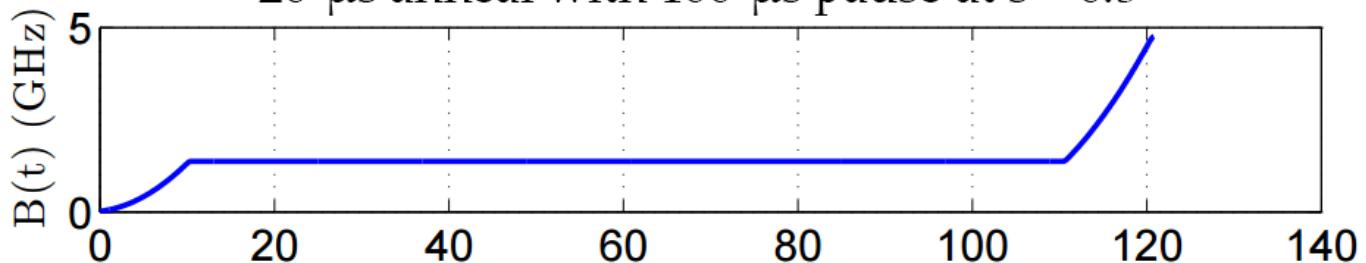
- `default` Standard 20 $\mu$ s annealing schedule
- `long` Standard 100 $\mu$ s annealing schedule
- `pause` 20 $\mu$ s anneal with 100 $\mu$ s pause at s=0.5
- `quench` 20 $\mu$ s anneal with 2 $\mu$ s quench at s=0.5

`default`, `pause`, and `quench` are illustrated in the following image:

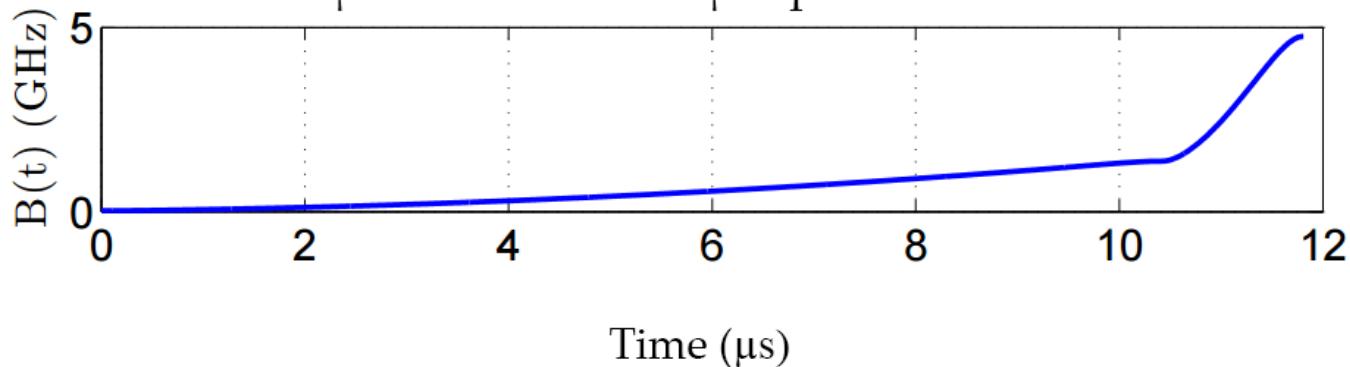
Standard 20- $\mu$ s annealing schedule



20- $\mu$ s anneal with 100- $\mu$ s pause at  $s = 0.5$

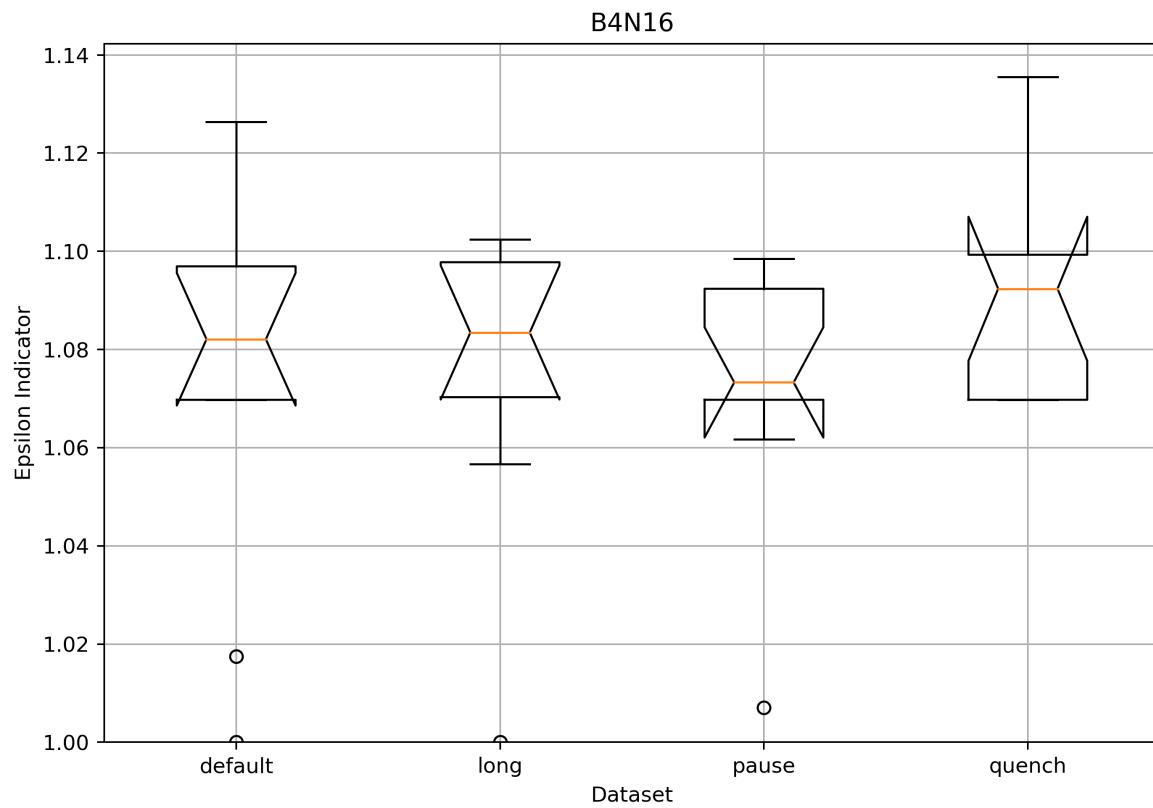


20- $\mu$ s anneal with 2- $\mu$ s quench at  $s = 0.5$



The experiments were run five times:

## Boxplots - scenarioA3Y2021M05D17h17m31s50



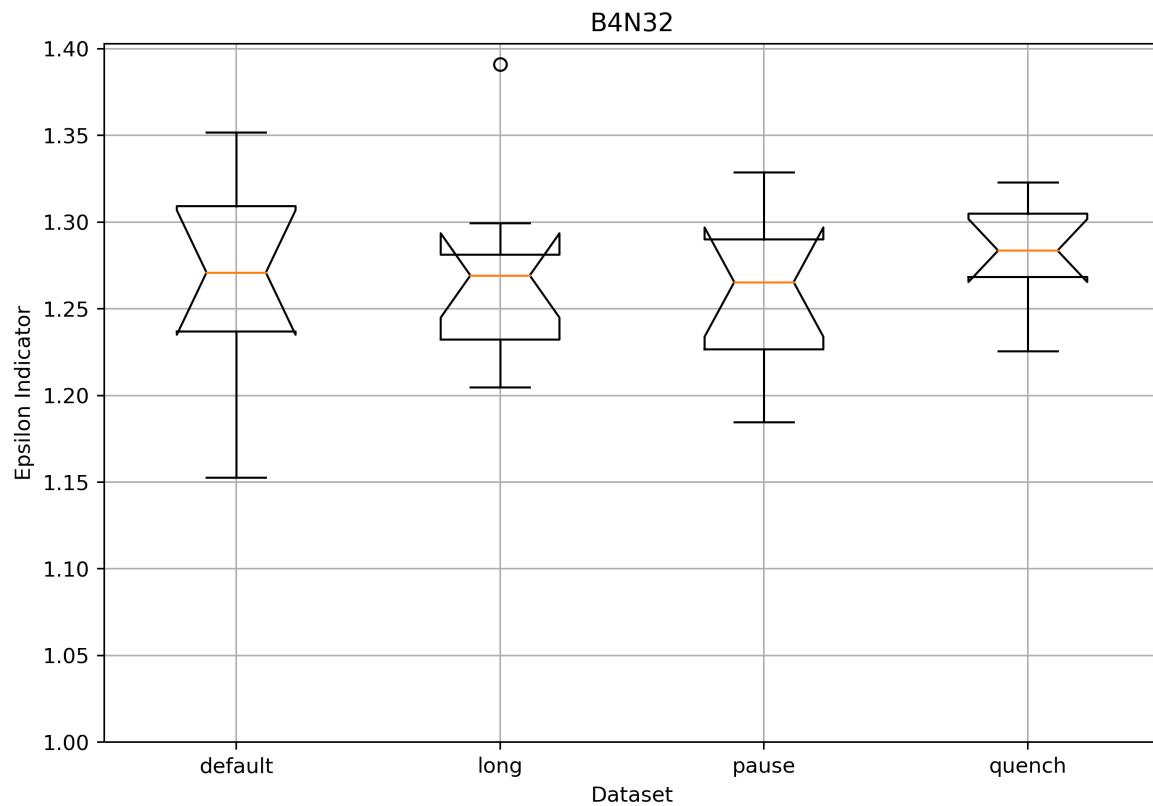
Results of Kruskal-Wallis One-way ANOVA:

H statistic: 0.9966749002470076

p-value: 0.8020565329549938

The null hypothesis was not rejected!

## Boxplots - scenarioA3Y2021M05D17h17m34s27



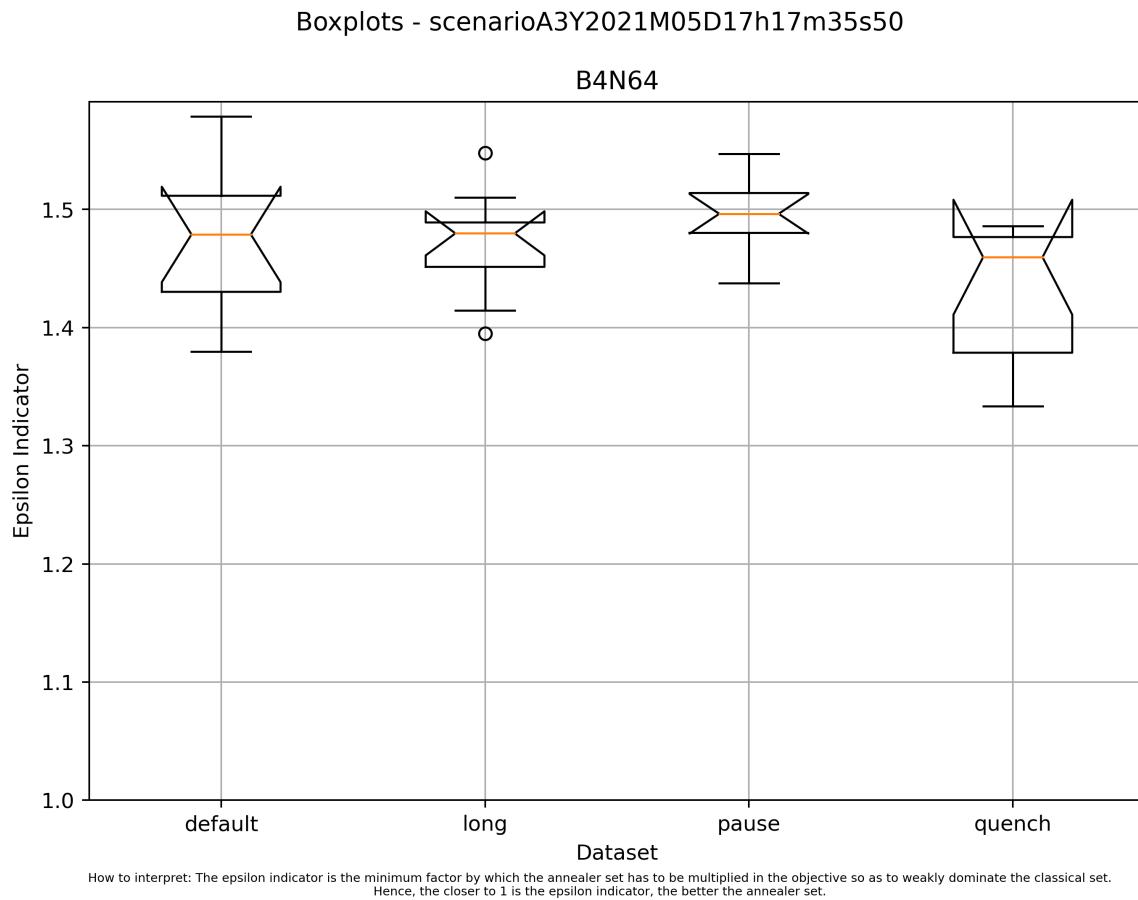
How to interpret: The epsilon indicator is the minimum factor by which the annealer set has to be multiplied in the objective so as to weakly dominate the classical set.  
Hence, the closer to 1 is the epsilon indicator, the better the annealer set.

## Results of Kruskal-Wallis One-way ANOVA:

H statistic: 1.4648780487805055

p-value: 0.690399853798688

The null hypothesis was not rejected!



Results of Kruskal-Wallis One-way ANOVA:  
 H statistic: 6.924878048780471  
 p-value: 0.0743311017252379  
 The null hypothesis was not rejected!

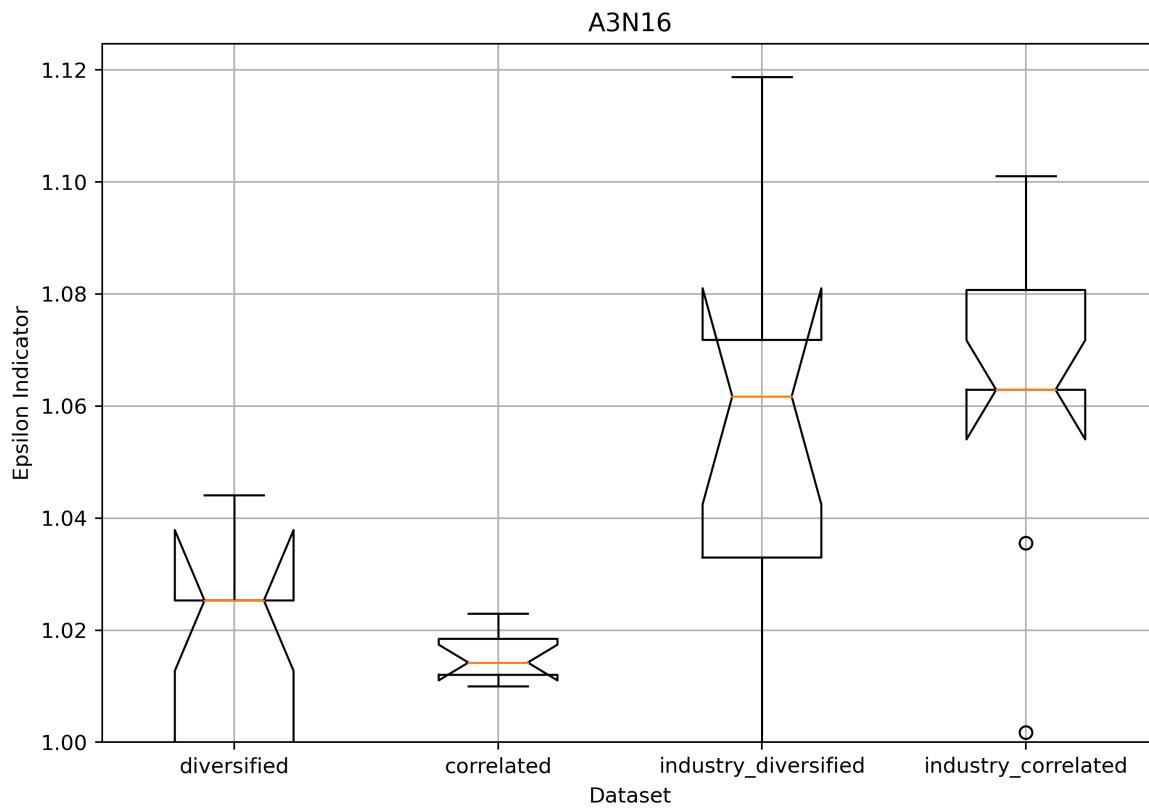
Key Takeaways:

## Scenario A3 - Datasets

For this scenario, we will study the influence from the dataset. Previous scenarios used a `industry_diversified` dataset, with assets from different industries. Therefore, we are going to introduce another dataset, called `industry_correlated`, from the same source, however, with assets from the same industry. Moreover, we will also introduce two datasets that are statistically correlated and diversified, `diversified` and `correlated`.

The results are executed for sizes `N=16`, `N=32`, and `N=64`, with parameters `chain_strength = 1.000 * maxAbs`, `B=0.5`, `mediumDmediumS` q values, `layout` embedding, and `default` schedule.

## Boxplots - scenarioA3Y2021M05D17h18m09s22



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 15.414601018675732

p-value: 0.0014945310210257902

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	diversified	correlated	industry_diversified
industry_correlated			
diversified	1.000000	1.000000	0.099833
0.004966			
correlated	1.000000	1.000000	0.030382
0.001261			
industry_diversified	0.099833	0.030382	1.000000
1.000000			
industry_correlated	0.004966	0.001261	1.000000
1.000000			

diversified and correlated are NOT!

diversified and industry\_diversified are NOT!

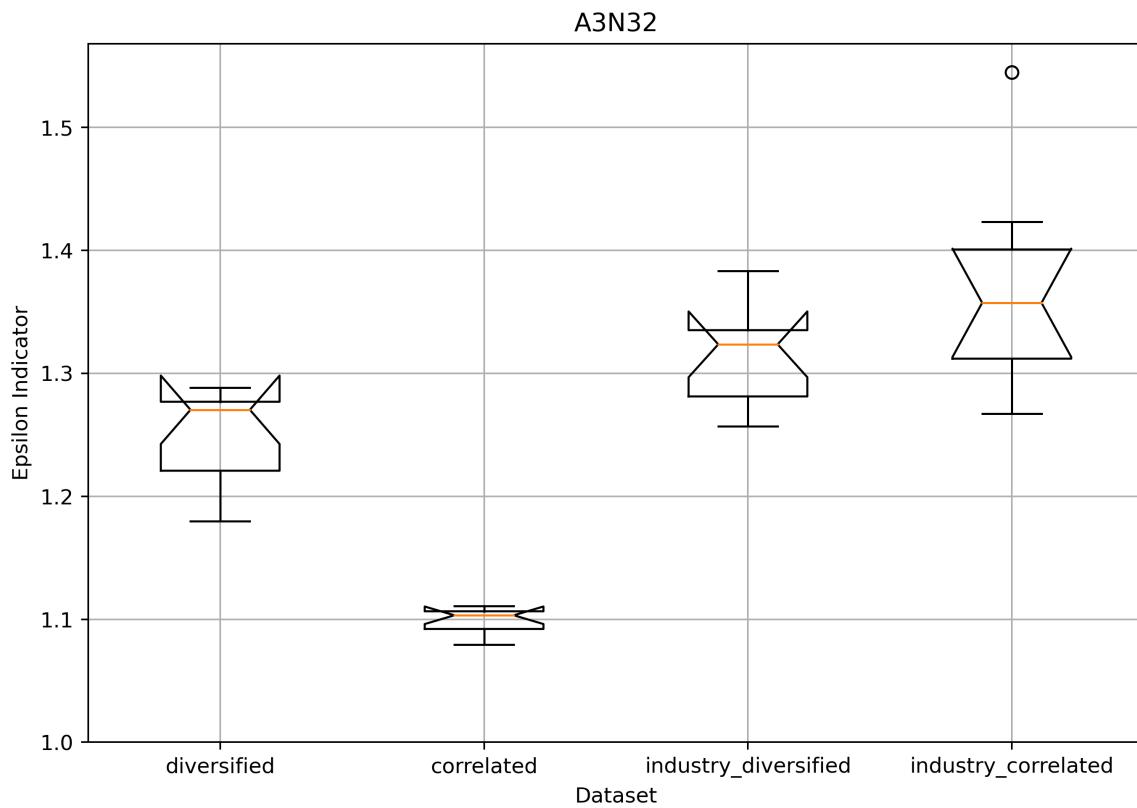
diversified and industry\_correlated ARE significantly different!

correlated and industry\_diversified ARE significantly different!

correlated and industry\_correlated ARE significantly different!

industry\_diversified and industry\_correlated are NOT!

## Boxplots - scenarioA3Y2021M05D17h18m10s33



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 28.54097560975609

p-value: 2.7961951388847623e-06

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	diversified	correlated	industry_diversified
industry_correlated	1.000000	4.024953e-04	2.324991e-02
diversified	2.356543e-04		
correlated	3.378219e-10	0.000402	1.000000e+00
industry_diversified	7.142090e-01	0.023250	3.326977e-08
industry_correlated	1.000000e+00	0.000236	3.378219e-10

diversified and correlated ARE significantly different!

diversified and industry\_diversified ARE significantly different!

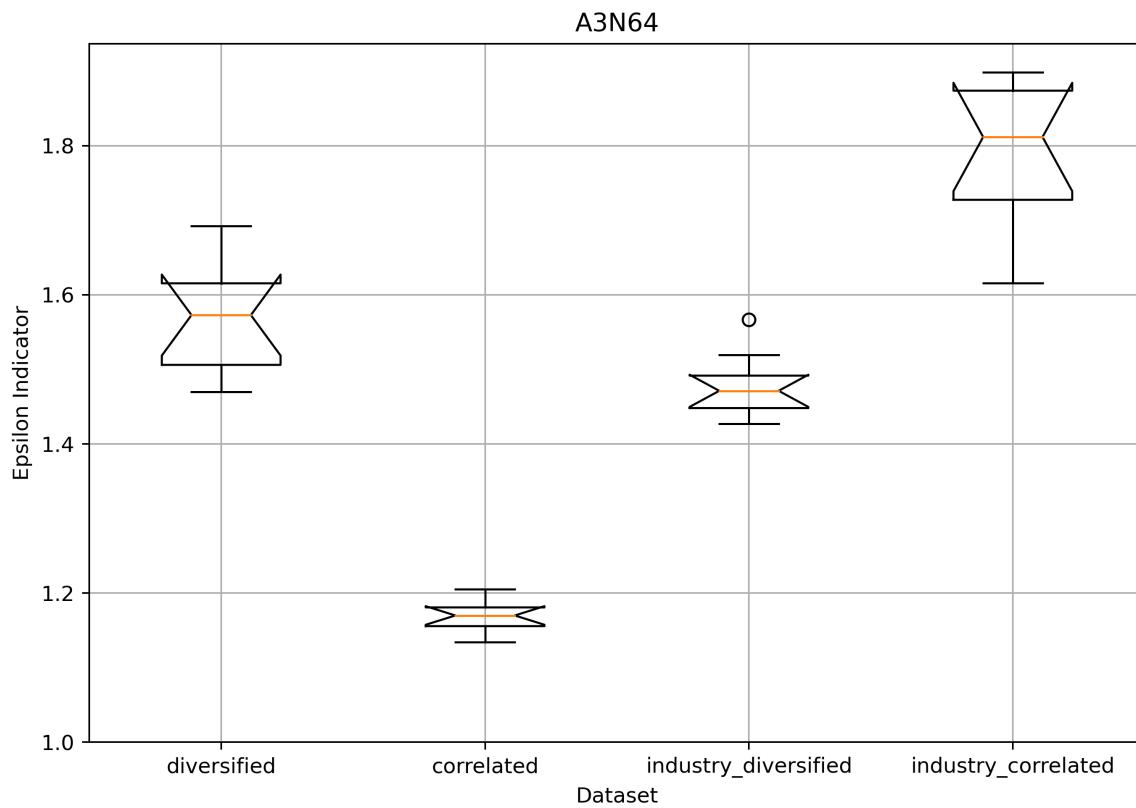
diversified and industry\_correlated ARE significantly different!

correlated and industry\_diversified ARE significantly different!

correlated and industry\_correlated ARE significantly different!

industry\_diversified and industry\_correlated are NOT!

## Boxplots - scenarioA3Y2021M05D17h18m13s15



Results of Kruskal-Wallis One-way ANOVA:

H statistic: 34.17951219512196

p-value: 1.8156272199120692e-07

Null hypothesis rejected! Now performing pairwise comparison with Conover-Iman's test

	diversified	correlated	industry_diversified
industry_correlated	1.000000e+00	4.477898e-11	1.959452e-03
diversified	1.714440e-05		
correlated	1.052151e-16	1.000000e+00	4.731989e-06
industry_diversified	1.959452e-03	4.731989e-06	1.000000e+00
industry_correlated	1.714440e-05	1.052151e-16	1.389476e-10
diversified	1.000000e+00		

diversified and correlated ARE significantly different!

diversified and industry\_diversified ARE significantly different!

diversified and industry\_correlated ARE significantly different!

correlated and industry\_diversified ARE significantly different!

correlated and industry\_correlated ARE significantly different!

industry\_diversified and industry\_correlated ARE significantly different!

Key Takeaways: