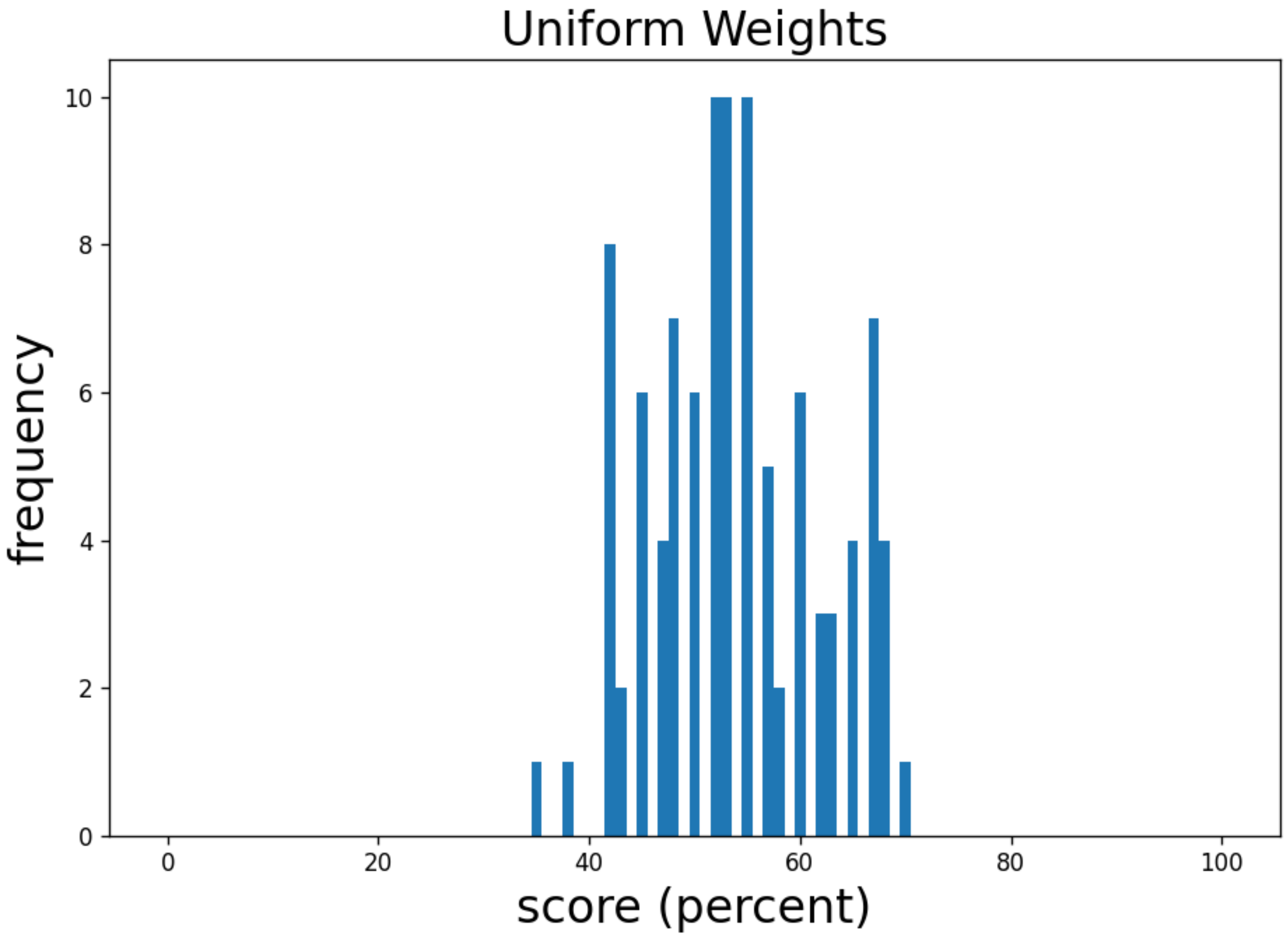
Constrained Optimisation

James Chok & Geoff Vasil - Dec 2022

Student	Q1	Q2	Q3	Q4	...	Q50
Scrooge	❌	✅	✅	✅		✅
Fezziwig	✅	✅	❌	❌		✅
Marley	❌	❌	✅	✅		❌
Emily	✅	✅	✅	✅		❌
...						

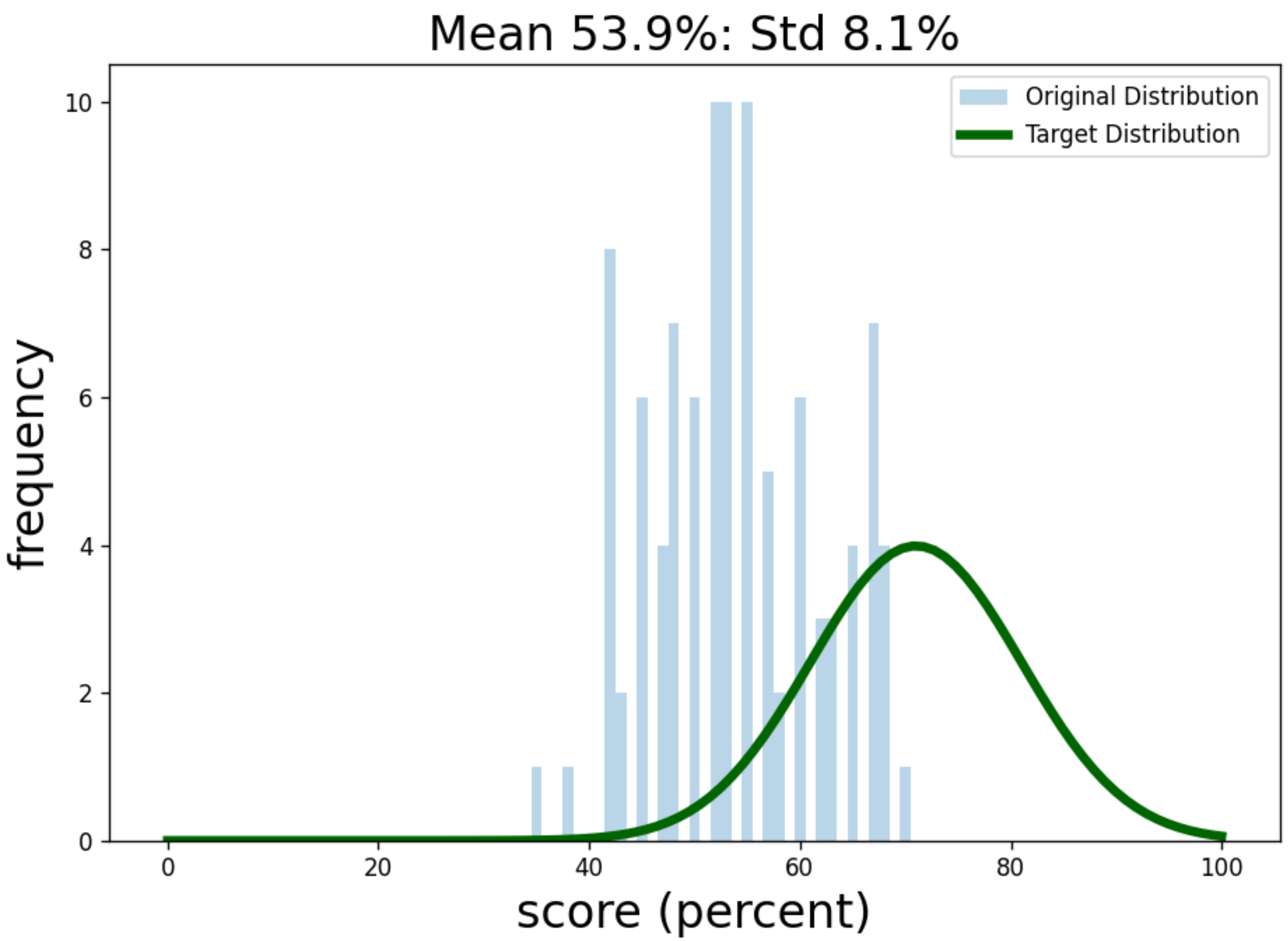
$$\text{Scrooge's Mark} = \sum_i w_i \cdot \text{(ith question is correct)}$$



Uniform Weights
frequency
score (percent)
0
2
4
6
8
10
0
20
40
60
80
100

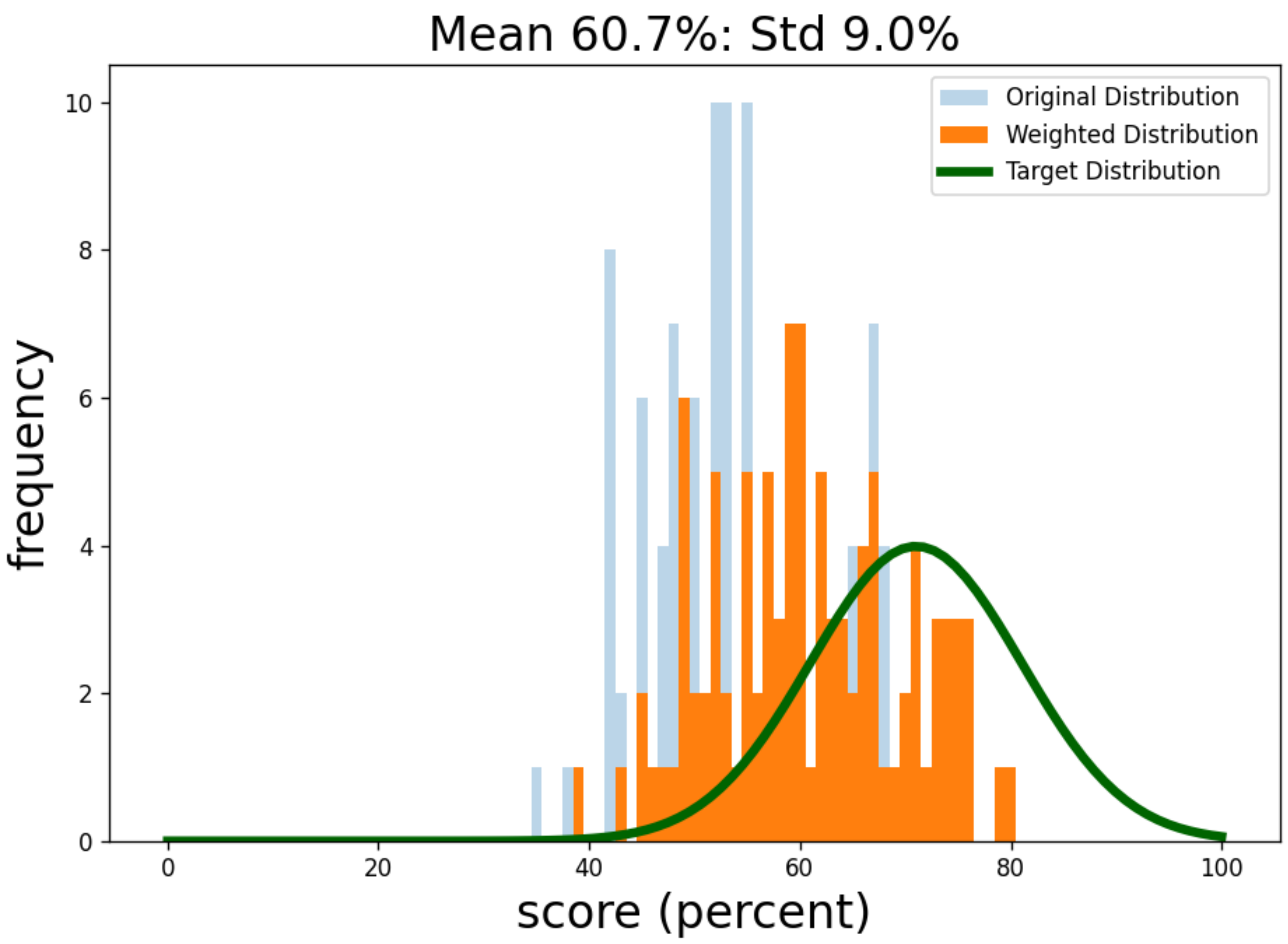

Uniform Weights
Random Weights
frequency
score (percent)
0
20
40
60
80
100
0
2
4
6
8
10
12
14
16
Uniform Weights
Random Weights
score (percent)
0
20
40
60
80
100
0
2
4
6
8
10
12

Iteration 0



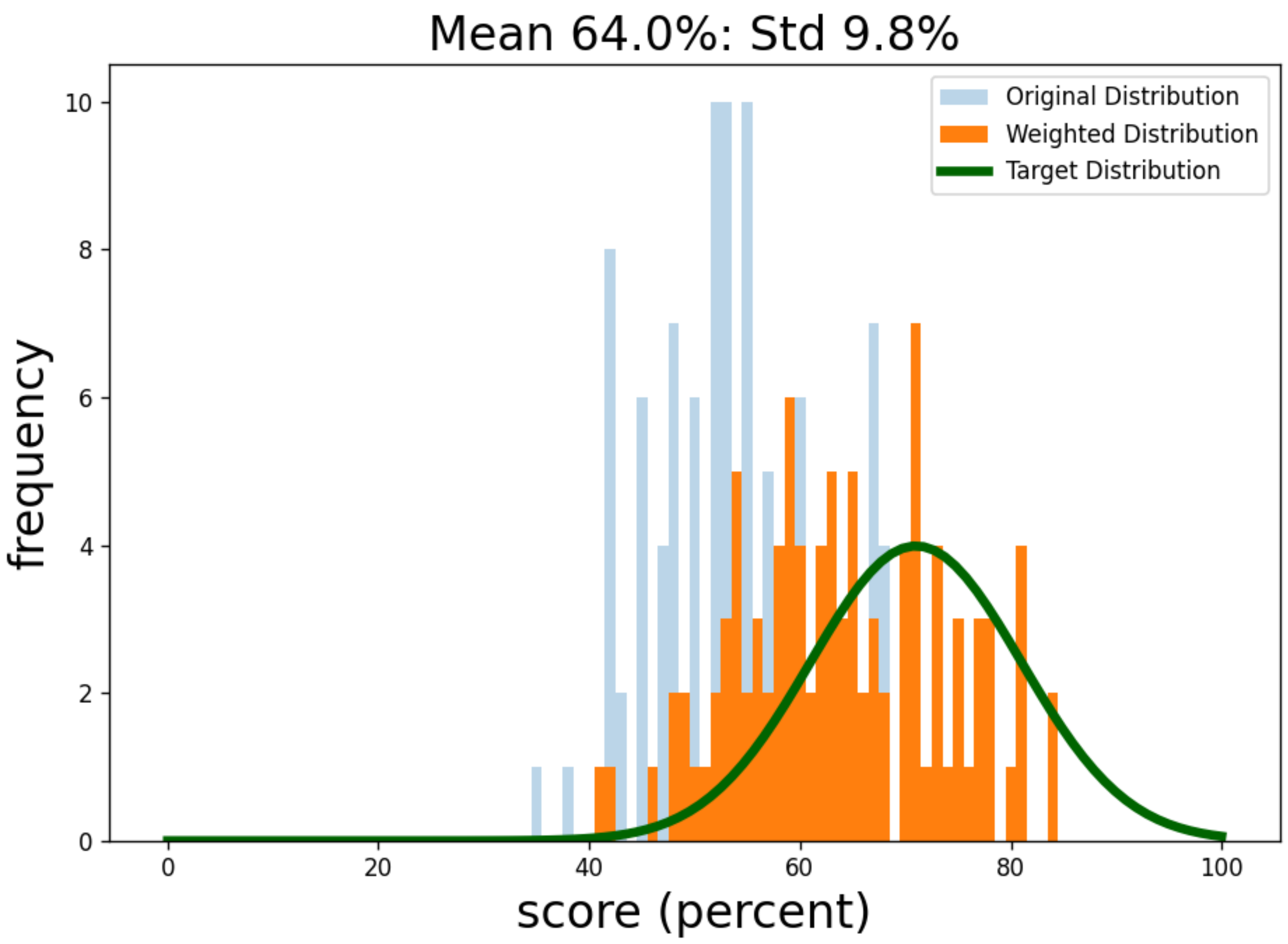
Mean 53.9%: Std 8.1%
Original Distribution
Target Distribution
frequency
score (percent)
0
2
4
6
8
10
0
20
40
60
80
100

Iteration 2

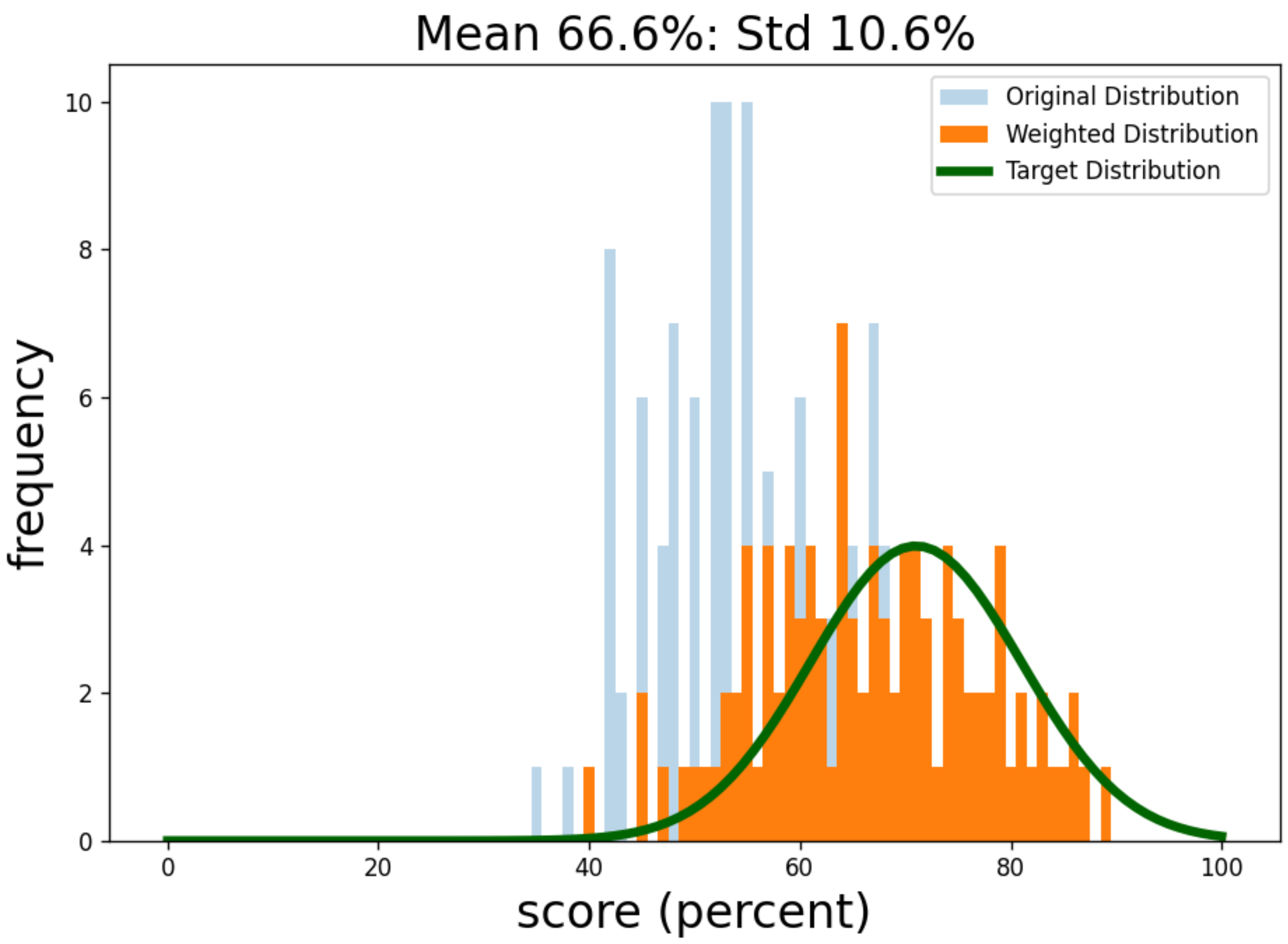


Mean 60.7%: Std 9.0%
Original Distribution
Weighted Distribution
Target Distribution
frequency
score (percent)
0
2
4
6
8
10
0
20
40
60
80
100

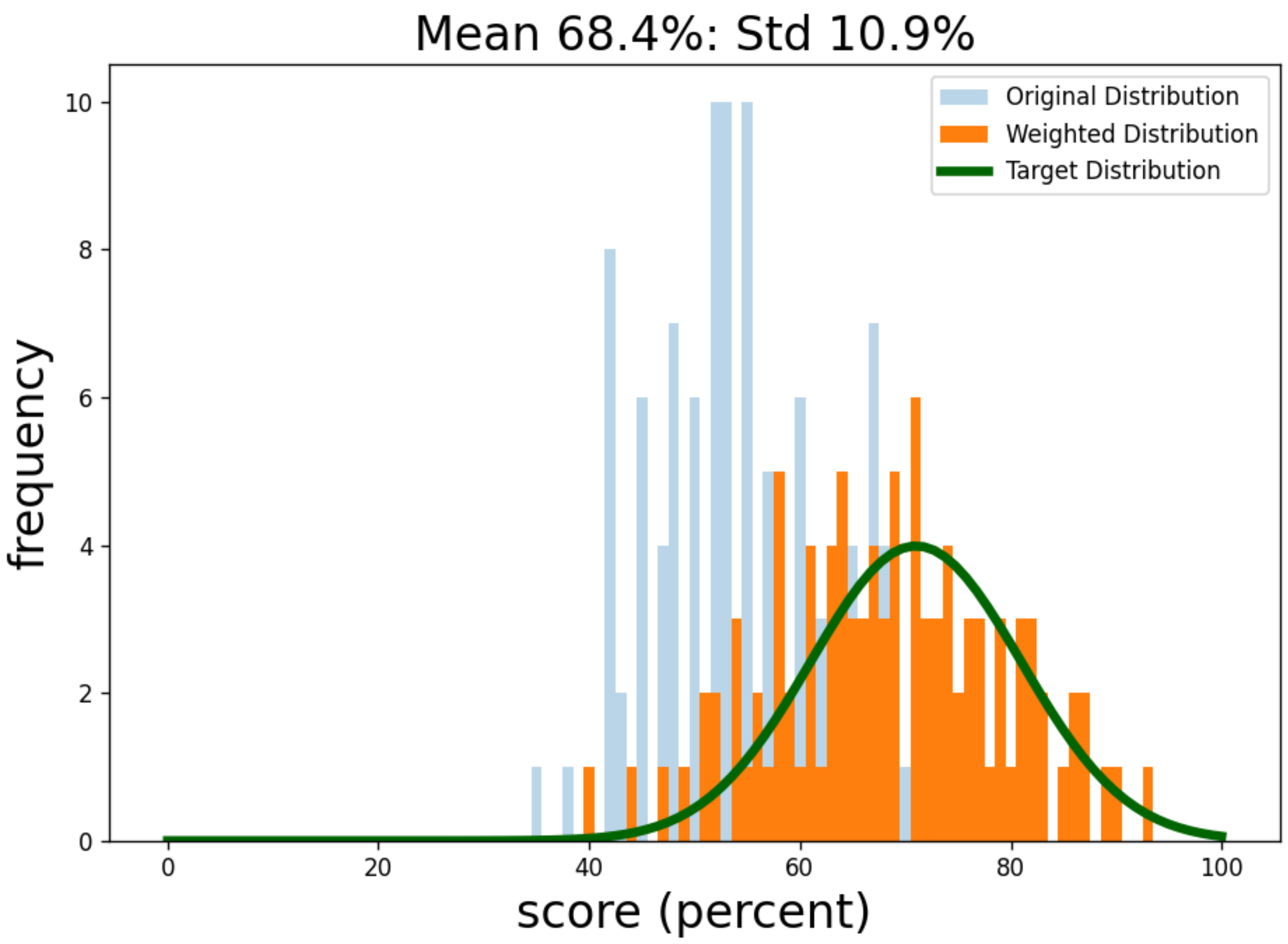
Iteration 10



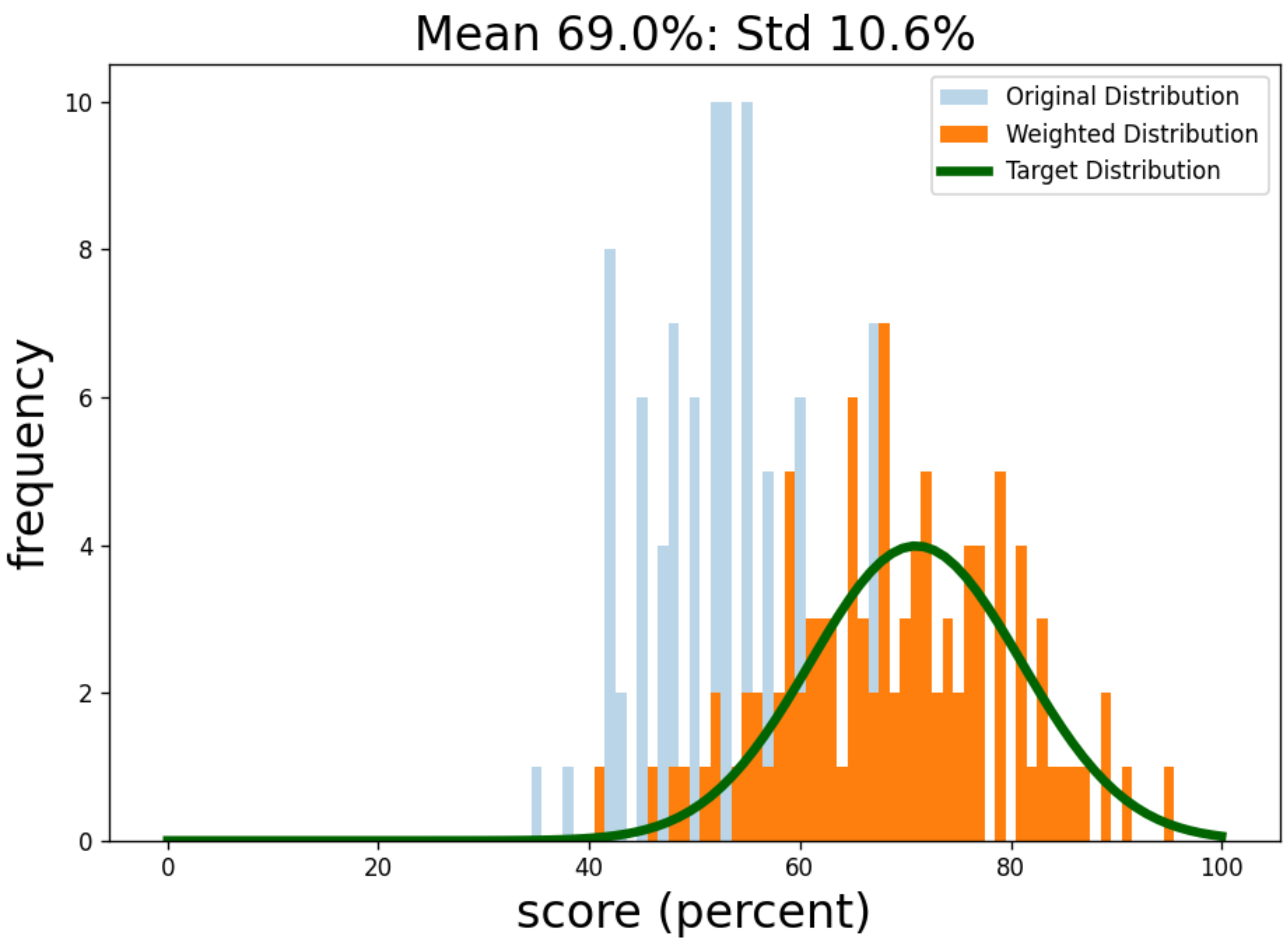
Iteration 40



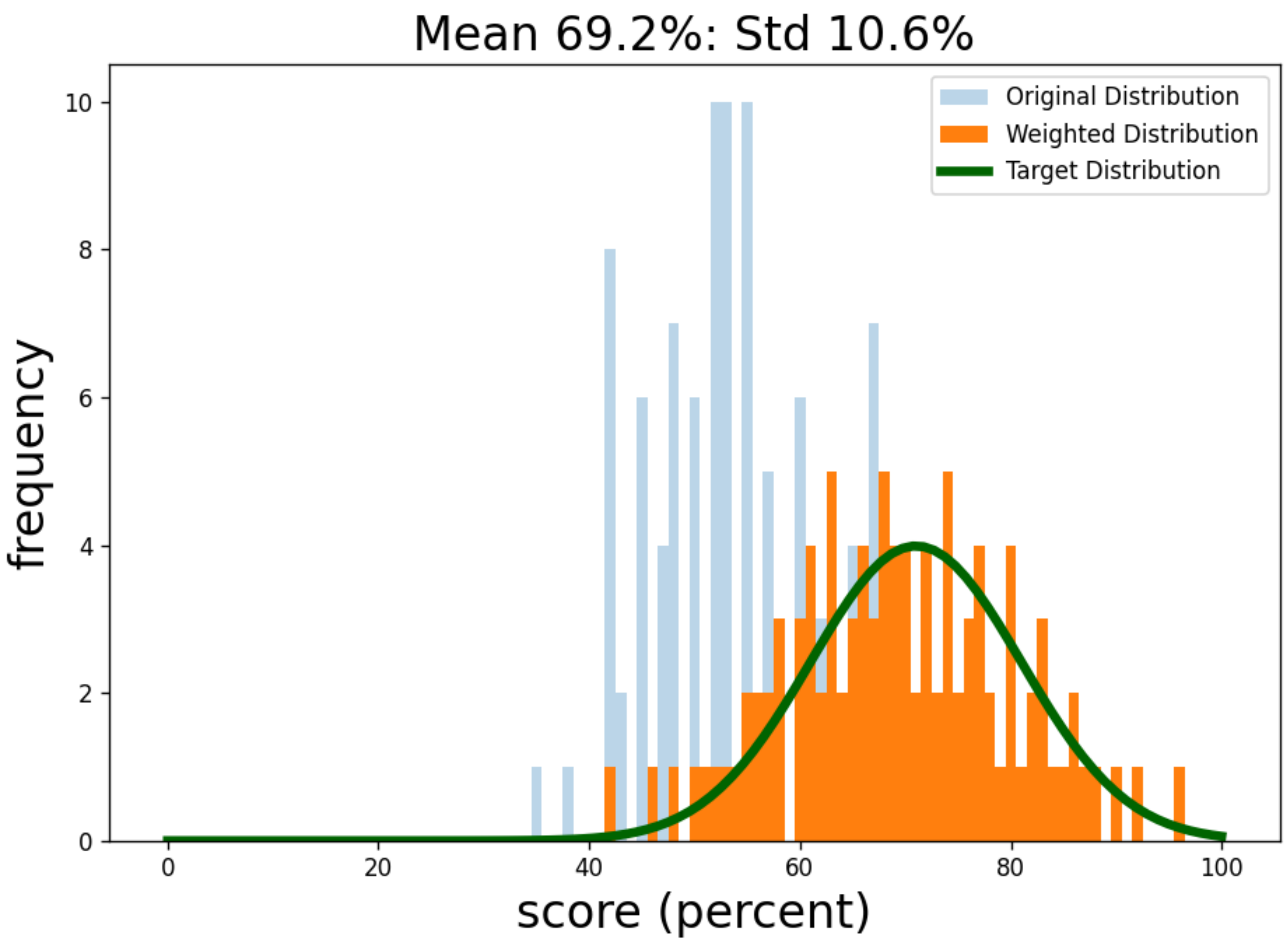
Iteration 170



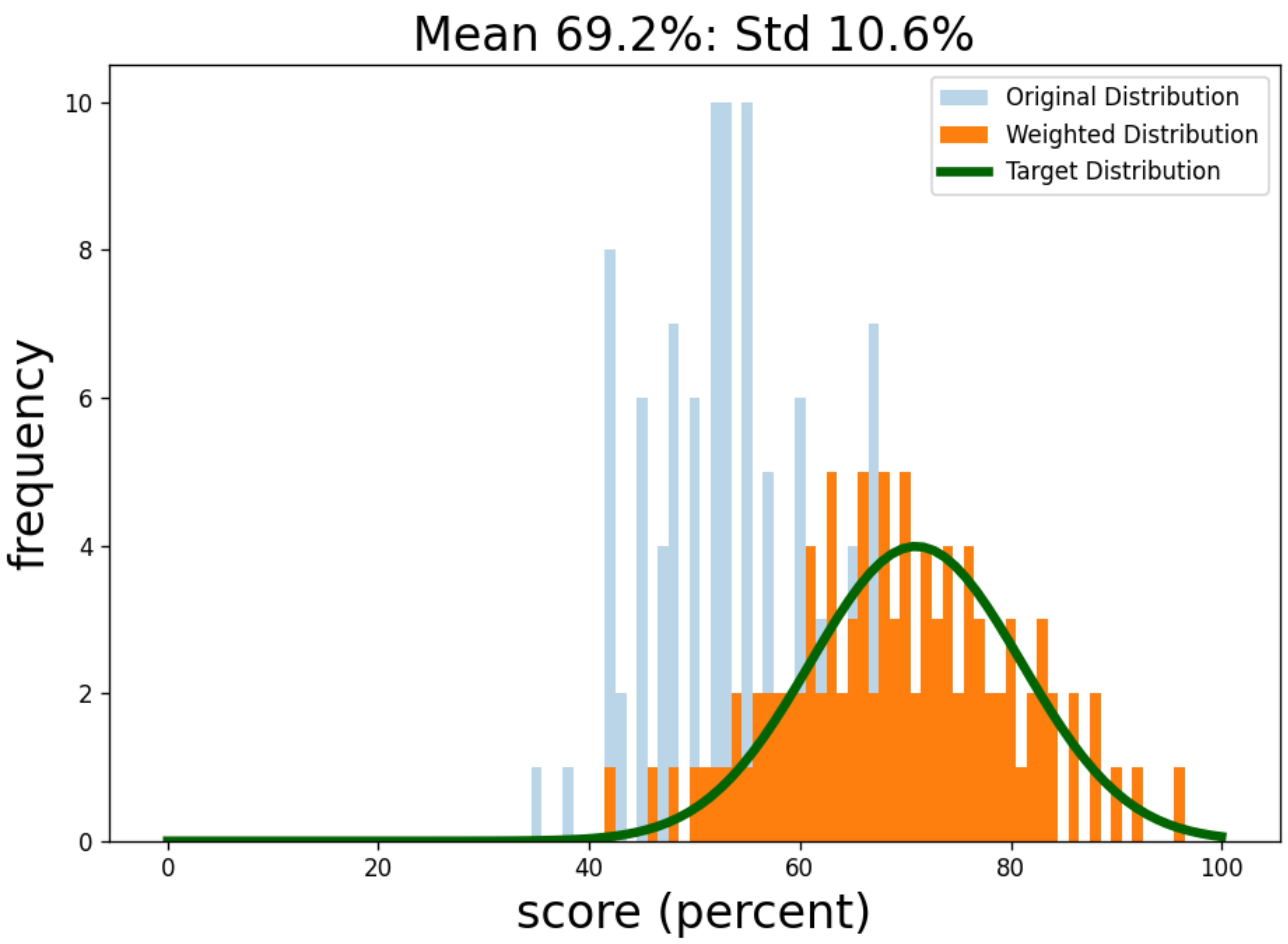
Iteration 700



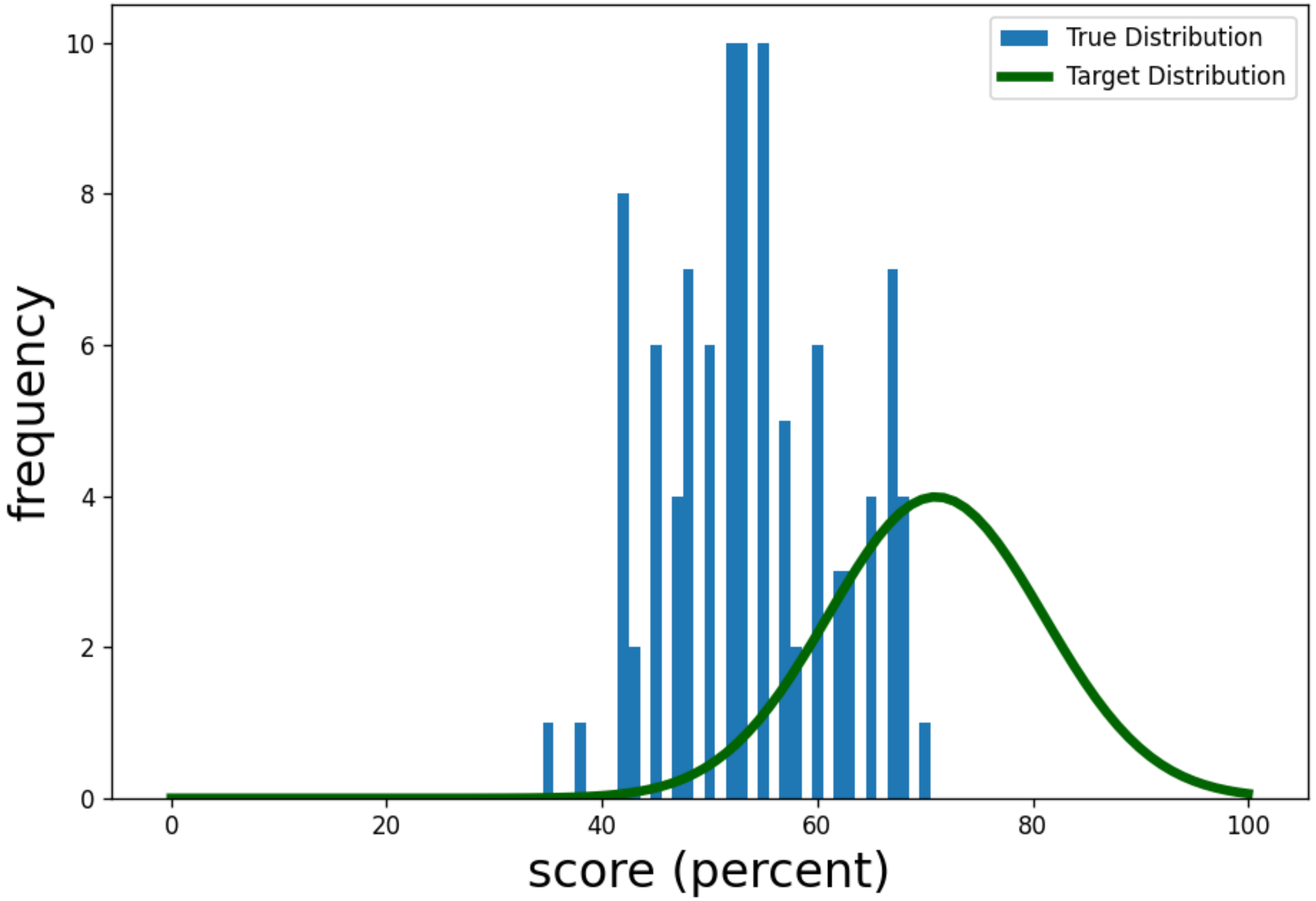
Iteration 2000



Iteration 11,000

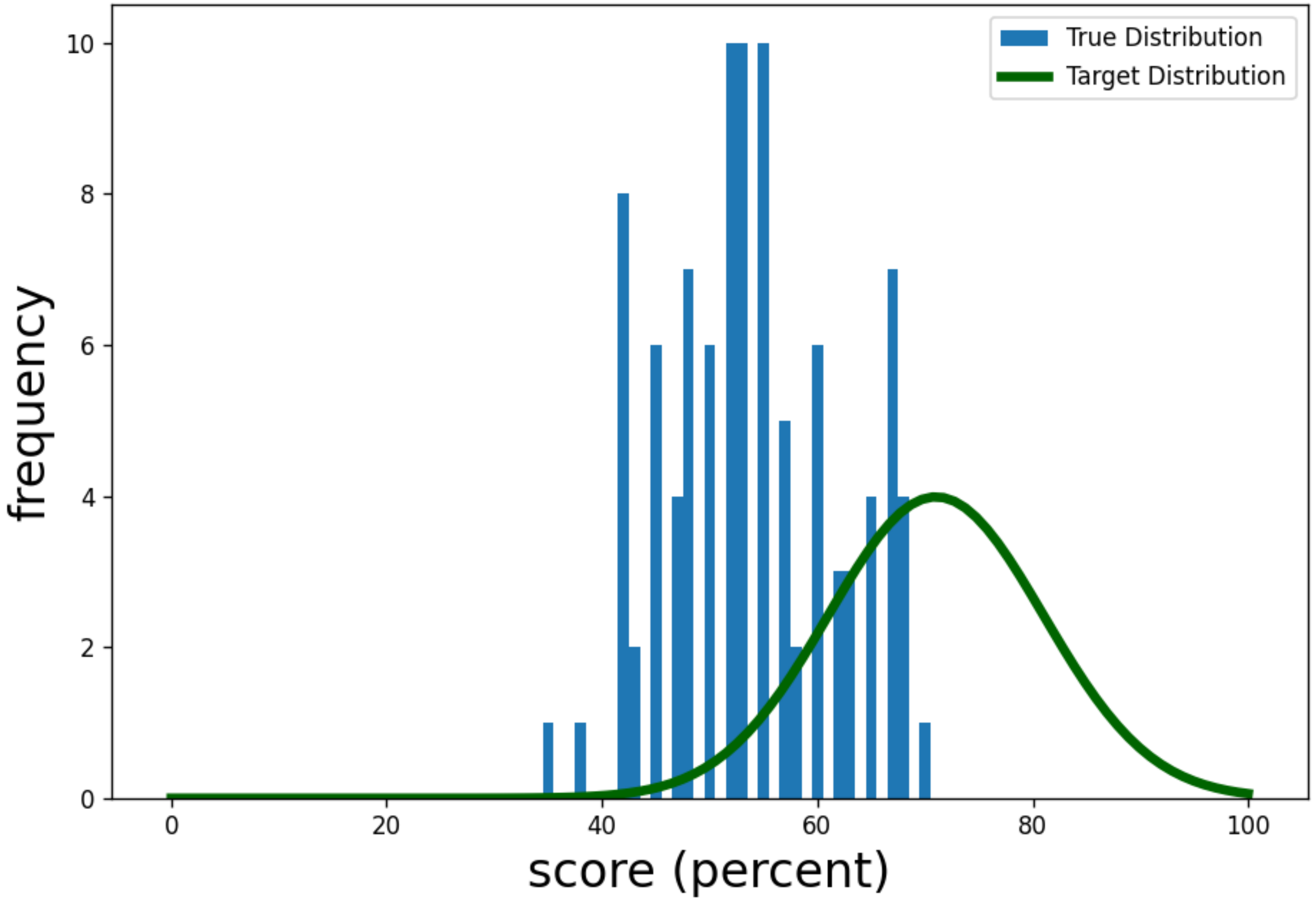


$$\min_{w} \; \|\text{Dist. of Weighted Marks} \; - \; \text{Target Dist.}\|^2$$



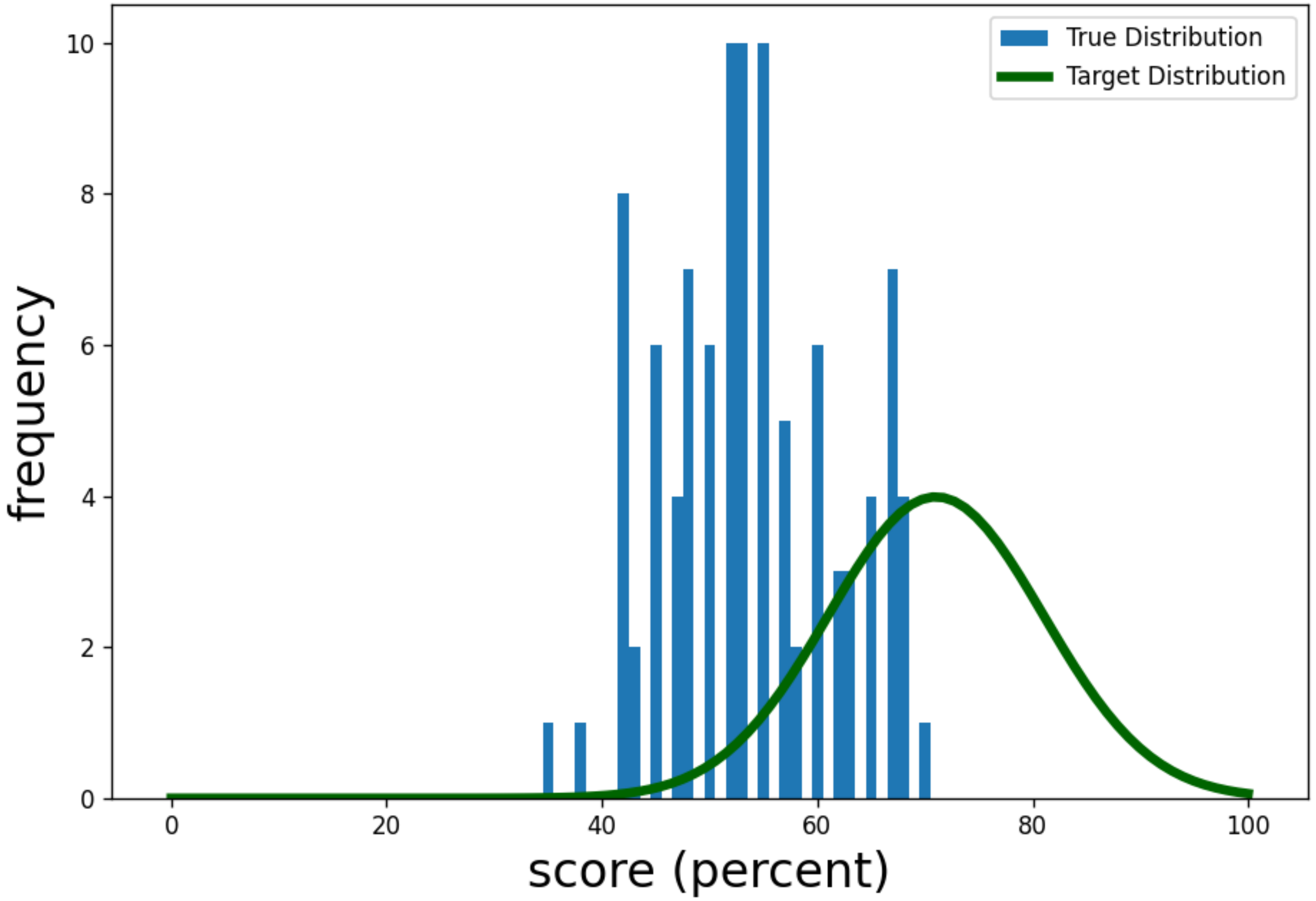
$$\min_{w} \ \|\text{Dist. of Weighted Marks} \ - \ \text{Target Dist.}\|^2$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$



$$\min_{w} \; f(w)$$

$$w_i \;\geq\; 0, \qquad \sum_{i} w_i = 1$$



Gradient Descent

Discrete: $$w^{t+1} = w^t - \eta_t \nabla_w f$$

Continuous: $$\frac{dw}{dt} = -\nabla_w f$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$

Exponentiated Gradient Descent

Discrete: $$w^{t+1} = w^t \exp(-\eta_t \nabla_w f)$$

Continuous: $$\frac{d}{dt} \log(w) = -\nabla_w f$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$

Exponentiated Gradient Descent

Discrete: $$w^{t+1} = \frac{w^t \exp(-\eta_t \nabla_w f)}{\sum_i w_i^t \exp(-\eta_t \nabla_{w_i} f)}$$

Continuous: $$\frac{d}{dt} \log(w) = -\nabla_w f$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$

Projected Gradient

Discrete: $$w^{t+1} = w^t - \eta_t \left(\nabla_w f - \frac{1}{n} \sum_i \nabla_{w_i} f \right)$$

Continuous: $$\frac{dw}{dt} = - \left(\nabla_w f - \frac{1}{n} \sum_i \nabla_{w_i} f \right)$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$

Our Proposal

Continuous: $$\frac{dw}{dt} = -w(\nabla_w f - w \cdot \nabla_w f)$$

Our Proposal

Continuous: $$\frac{dw}{dt} = -w(\nabla_w f - w \cdot \nabla_w f)$$

Discrete: $$w^{t+1} = w^t - \eta_t w^t (\nabla_w f - w^t \cdot \nabla_w f)$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$

Our Proposal

$$\frac{dw}{dt} = -w\,(\nabla_w f - w \cdot \nabla_w f)$$

Continuous: $$\frac{d}{dt}\log(w) = -(\nabla_w f - w \cdot \nabla_w f)$$

Discrete: $$w^{t+1} = w^t \exp(-\eta_t (\nabla_w f - w^t \cdot \nabla_w f)),$$

$$w_i \geq 0, \qquad \sum_i w_i = 1$$