Zeroth-Order Optimizer Benchmarking for 3D Performance Capture

A real-world use case analysis

Alexandros Doumanoglou
Petros Drakoulis*
Kyriaki Christaki*
aldoum@iti.gr
petros.drakoulis@iti.gr
kchristaki@iti.gr

Nikolaos Zioulis*
Vladimiros Sterzentsenko
Antonis Karakottas
nzioulis@iti.gr
vladster@iti.gr
ankarako@iti.gr

Dimitrios Zarpalas
Petros Daras
zarpalas@iti.gr
daras@iti.gr

Centre For Research and Technology
HELLAS
Thessaloniki, Greece

ABSTRACT

This document contains the 1st part of the supplementary material of the paper: https://doi.org/10.1145/3449639.3459354. Visit the project's page at https://vcl3d.github.io/nevergrad/. Part #2 of the supplementary material can be found at https://github.com/VCL3D/nevergrad/blob/pb-material/supplementary2.pdf, while part #3 at https://github.com/VCL3D/nevergrad/blob/pb-material/supplementary3.pdf.

ACM Reference Format:

Alexandros Doumanoglou, Petros Drakoulis, Kyriaki Christaki, Nikolaos Zioulis, Vladimiros Sterzentsenko, Antonis Karakottas, Dimitrios Zarpalas, and Petros Daras. 2021. Zeroth-Order Optimizer Benchmarking for 3D Performance Capture: A real-world use case analysis. In *Proceedings of the Genetic and Evolutionary Computation Conference 2021 (GECCO '21)*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 LOCAL FITTING

In Figs. 1 - 10 we present loss vs budget curves for the optimizers in the local fitting case per individual captured performance, while In Fig. 11 we give the averaged loss vs budget curve across all captured performances. For Figs 1 - 10 faint color curves indicate variation across repetitions, whilst faint color curves in Fig 11 indicate variation across captured performances. Finally, in Fig. 12 we give qualitative results, illustrating error heatmaps (warm colors indicate low error, while hot colors indicate high error) for the converged poses of each optimizer, in 3 different captured performances (target frames).

Figure 1: *Local fitting:* Loss vs Budget for sequence Pushups

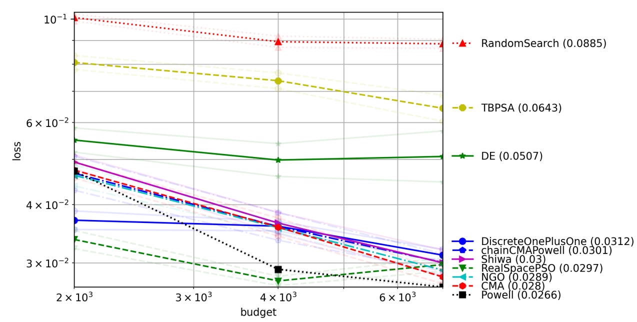
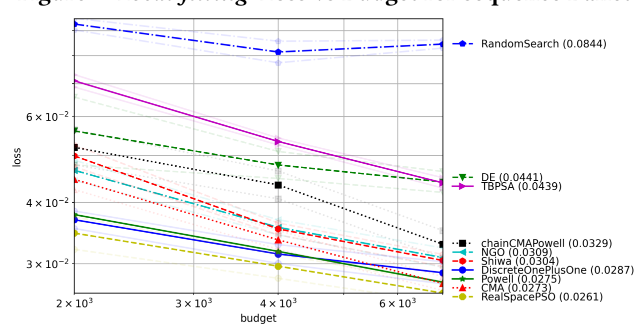


Figure 2: *Local fitting:* Loss vs Budget for sequence Ballet



*Denotes equal contribution

GECCO '21, July 10–14, 2021, Lille, France
2021. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Figure 3: *Local fitting:* Loss vs Budget for sequence Calisthenics

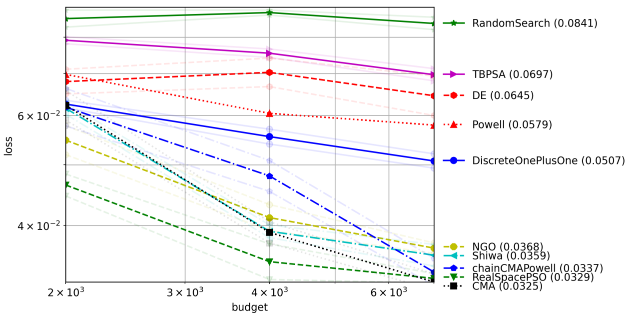


Figure 4: *Local fitting:* Loss vs Budget for sequence Boxing

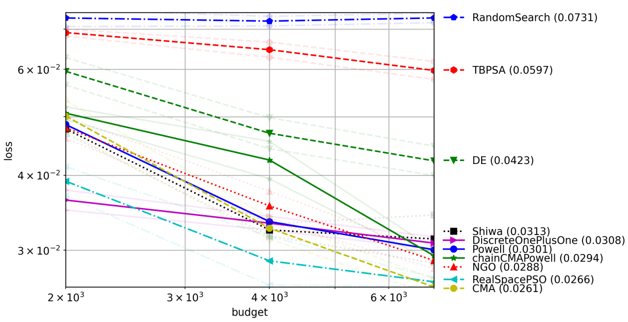


Figure 5: *Local fitting:* Loss vs Budget for sequence Yoga1

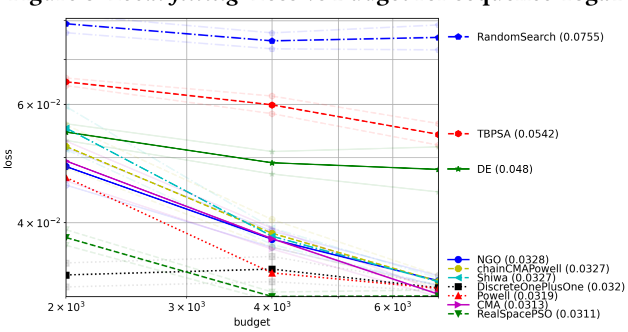


Figure 6: *Local fitting:* Loss vs Budget for sequence Yoga2

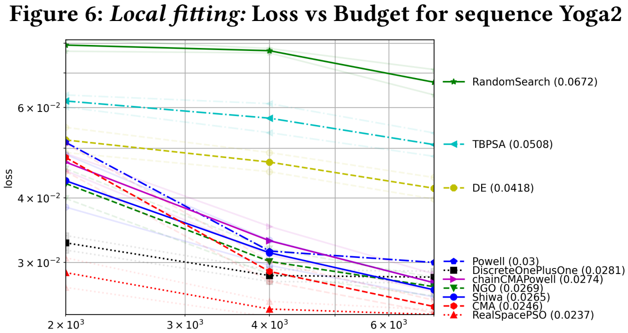


Figure 7: *Local fitting:* Loss vs Budget for sequence Jumping Jacks

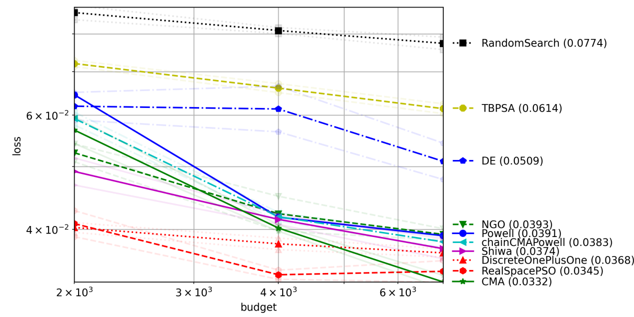


Figure 8: *Local fitting:* Loss vs Budget for sequence Football

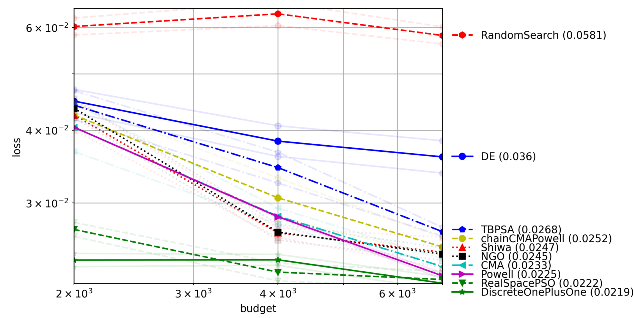


Figure 9: *Local fitting:* Loss vs Budget for sequence Dancing

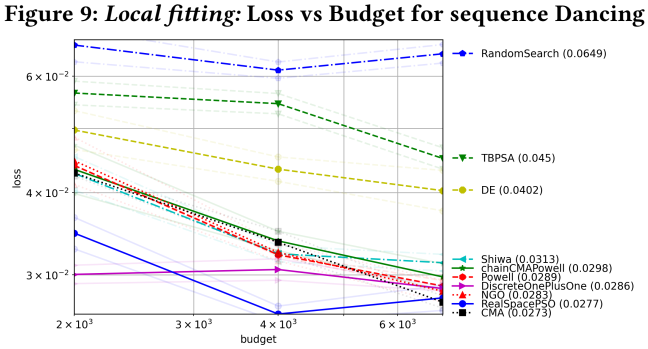


Figure 10: *Local fitting:* Loss vs Budget for sequence Stretching

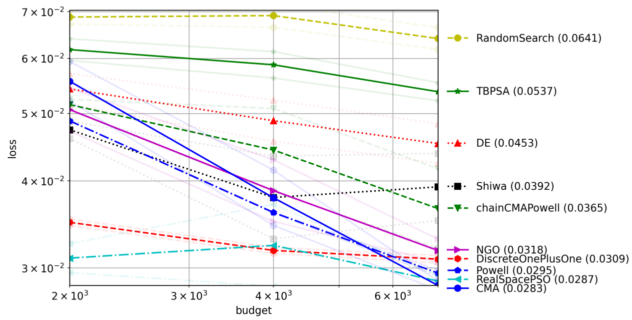
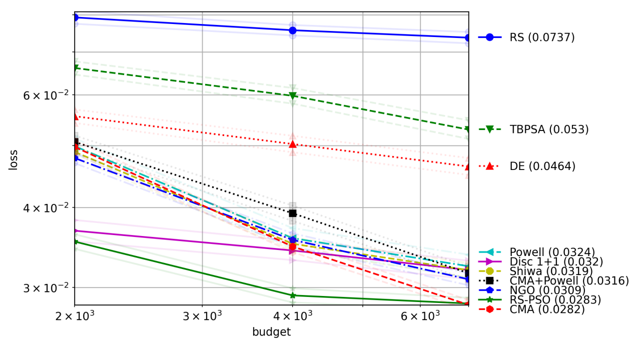


Figure 11: *Local fitting:* Average Loss vs Budget across all sequences



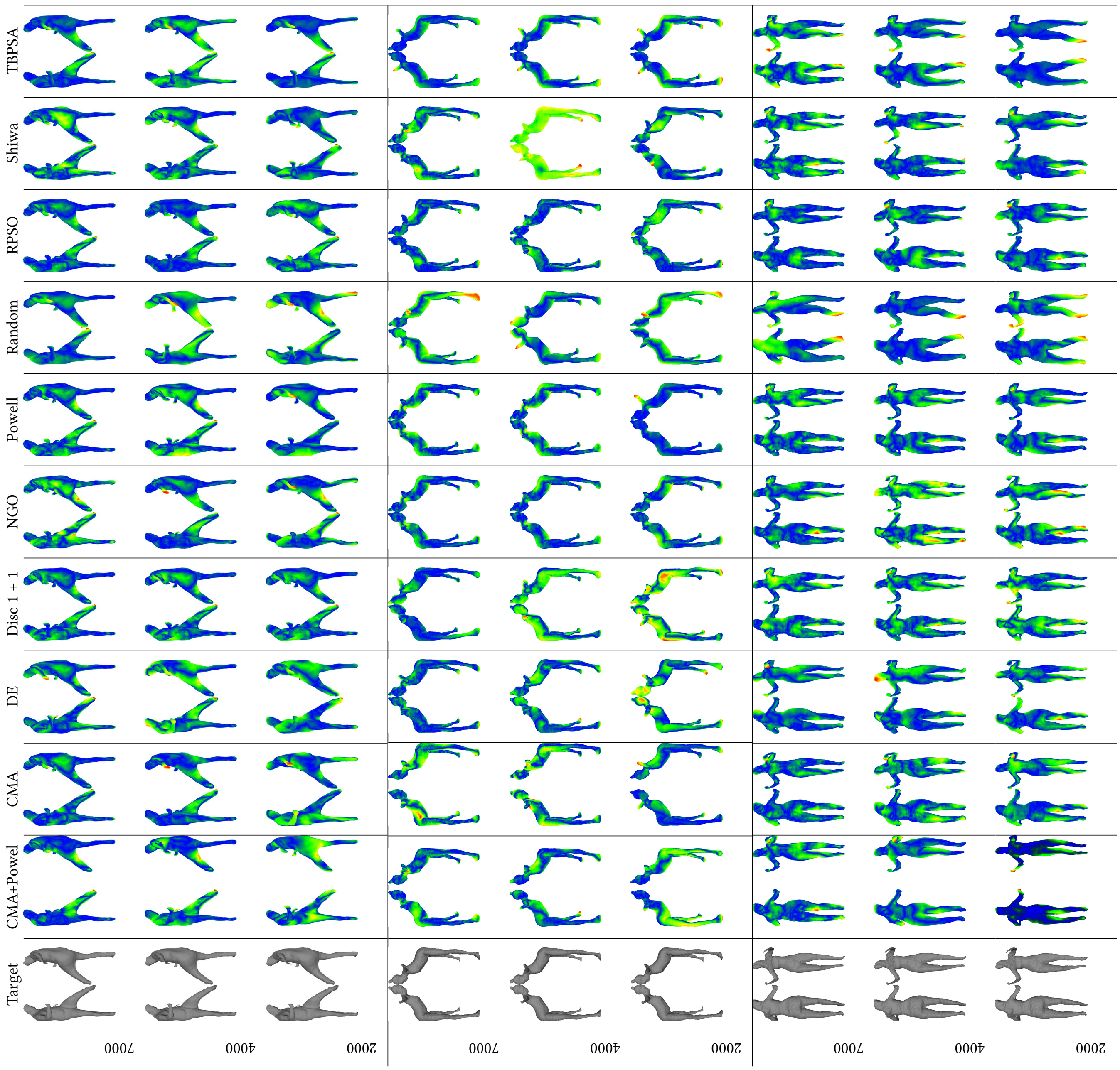

TBPSA
Shiwa
RPSO
Random
Powell
NGO
Disc 1 + 1
DE
CMA
CMA+Powel
Target
7000
4000
2000

Figure 12

2 GLOBAL FITTING

In Figs. 13 - 20 (22 - 29) we present loss vs budget curves for the optimizers in the global fitting case per individual captured performance for $\lambda_J = 0.1$ ($\lambda_J = 0.2$), while In Fig. 21 (30) we give the averaged loss vs budget curve across all captured performances. For Figs 13 - 20 (22 - 29) faint color curves indicate variation across repetitions, whilst faint color curves in Figs 21, 30 indicate variation across captured performances. Finally, in Fig. 31 we give qualitative results, illustrating error heatmaps (warm colors indicate low error, while hot colors indicate high error) for the converged poses of each optimizer, in 3 different captured performances (target frames).

Figure 13: *Global fitting:* Loss vs Budget for sequence Ballet, $\lambda_J = 0.1$

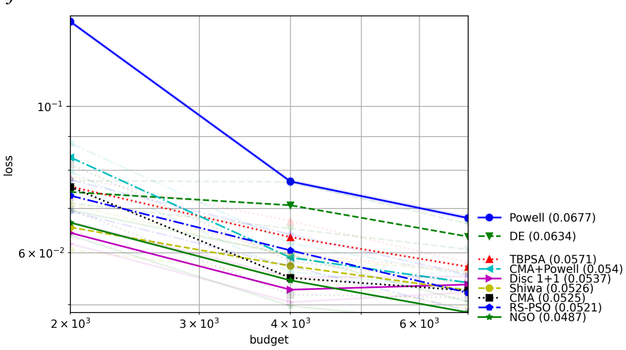


Figure 14: *Global fitting:* Loss vs Budget for sequence Boxing, $\lambda_J = 0.1$

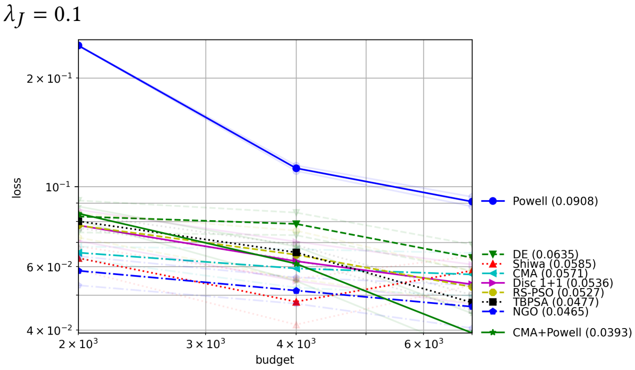


Figure 15: *Global fitting:* Loss vs Budget for sequence Calisthenics, $\lambda_J = 0.1$

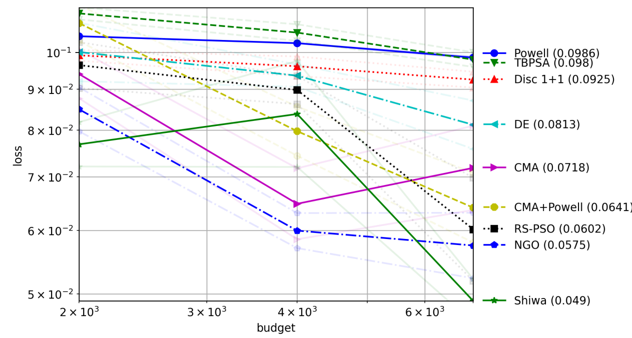


Figure 16: *Global fitting:* Loss vs Budget for sequence Dancing1, $\lambda_J = 0.1$

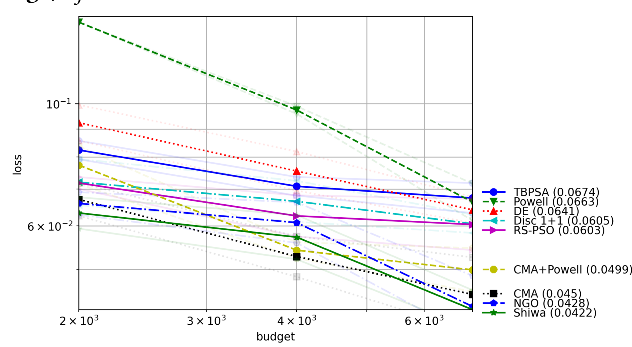


Figure 17: *Global fitting:* Loss vs Budget for sequence Dancing2, $\lambda_J = 0.1$

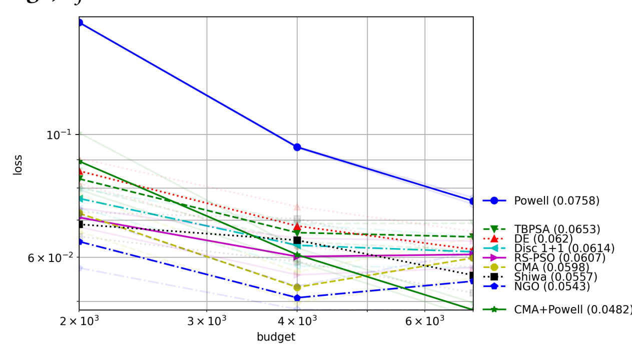


Figure 18: *Global fitting:* Loss vs Budget for sequence Football, $\lambda_J = 0.1$

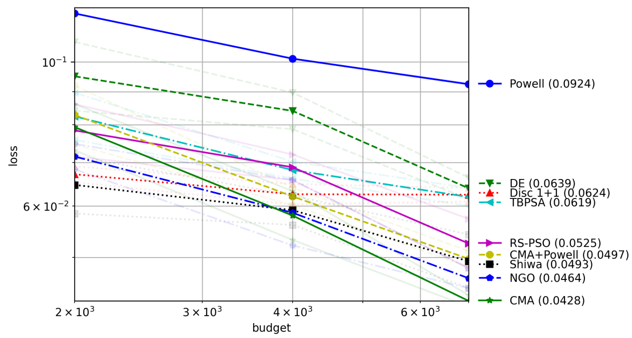


Figure 19: *Global fitting:* Loss vs Budget for sequence Karate, $\lambda_J = 0.1$

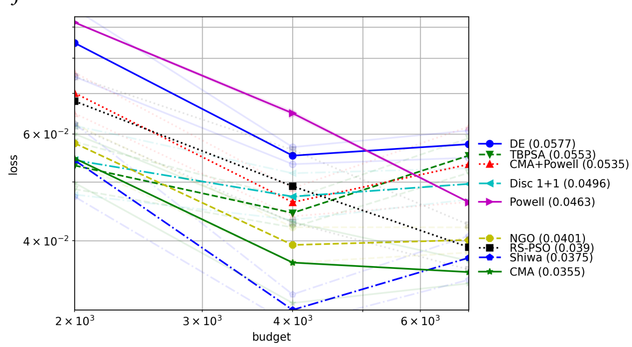


Figure 20: *Global fitting:* Loss vs Budget for sequence Stretching, $\lambda_J = 0.1$

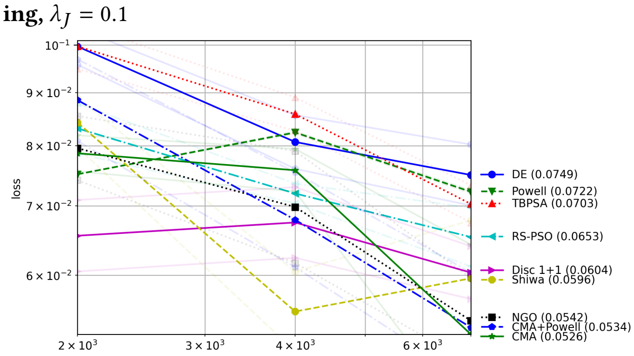


Figure 21: *Global fitting:* Average Loss vs Budget across sequences, $\lambda_J = 0.1$

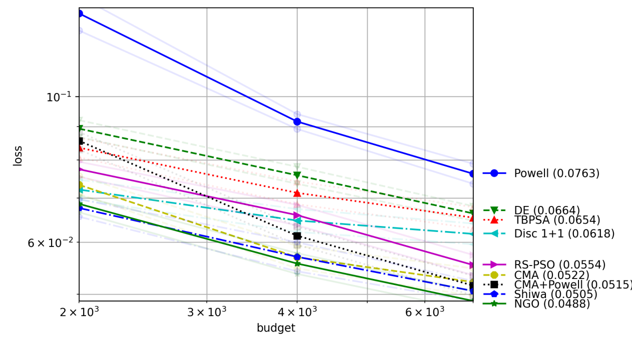


Figure 22: *Global fitting:* Loss vs Budget for sequence Ballet, $\lambda_J = 0.2$

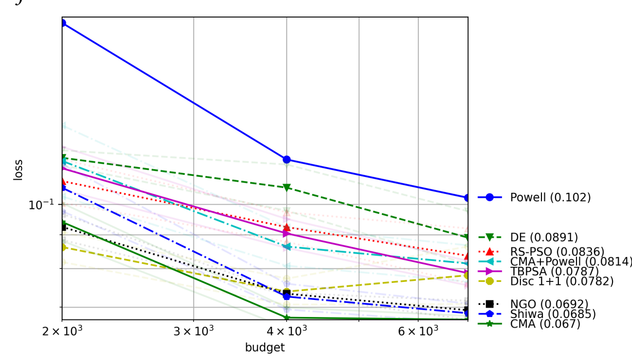


Figure 23: *Global fitting:* Loss vs Budget for sequence Boxing, $\lambda_J = 0.2$

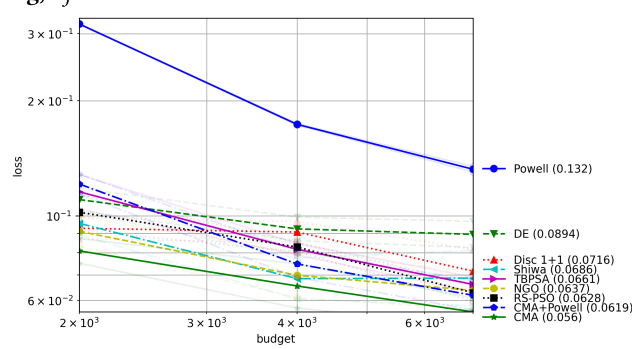


Figure 24: *Global fitting:* Loss vs Budget for sequence Calisthenics, $\lambda_J = 0.2$

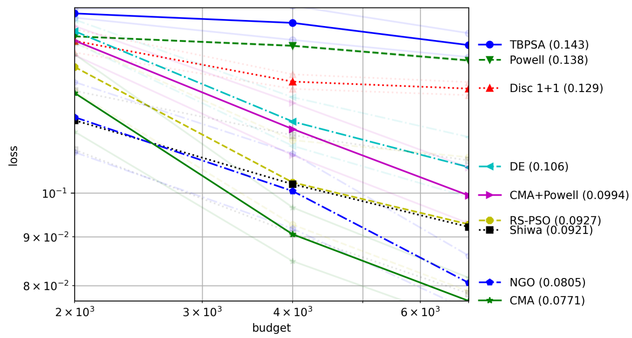


Figure 25: *Global fitting:* Loss vs Budget for sequence Dancing1, $\lambda_J = 0.2$

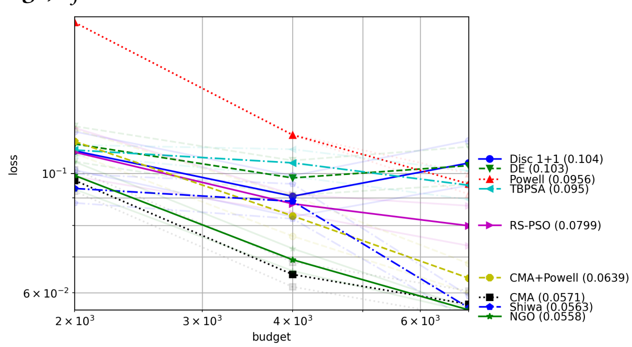


Figure 26: *Global fitting:* Loss vs Budget for sequence Dancing2, $\lambda_J = 0.2$

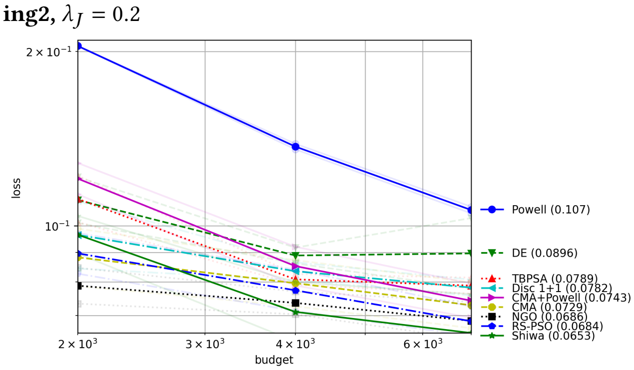


Figure 27: *Global fitting:* Loss vs Budget for sequence Football, $\lambda_J = 0.2$

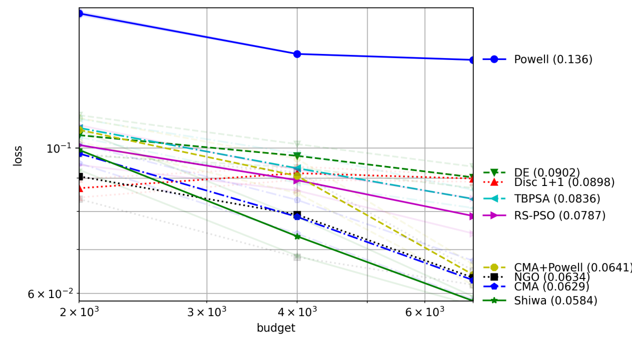


Figure 28: *Global fitting:* Loss vs Budget for sequence Karate, $\lambda_J = 0.2$

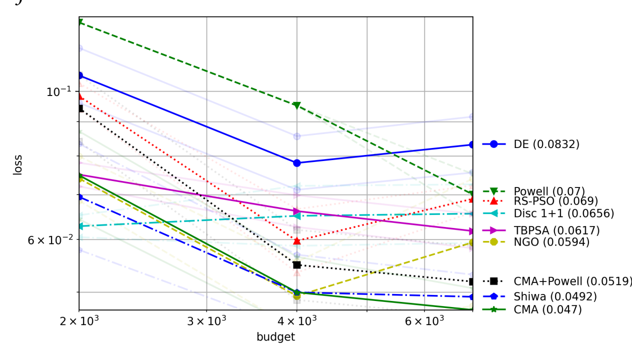


Figure 29: *Global fitting:* Loss vs Budget for sequence Stretching, $\lambda_J = 0.2$

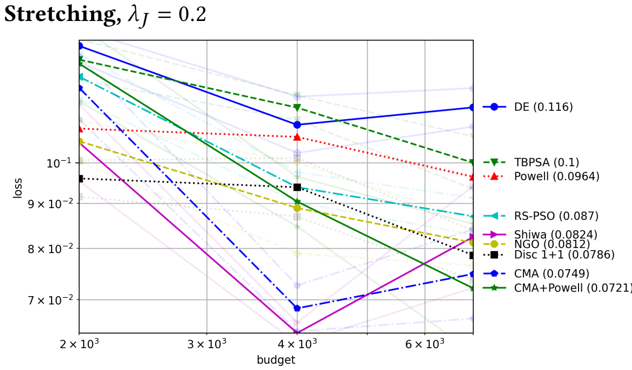
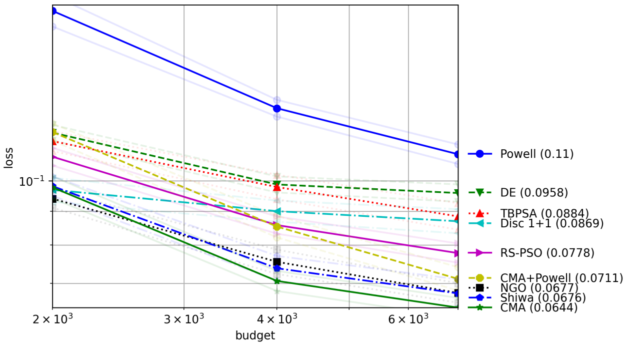


Figure 30: *Global fitting:* Average Loss vs Budget across sequences, $\lambda_J = 0.2$



2.1 Results and Analysis

Table 1: *Global fitting*: Average Loss for experiments with the same budget and optimizer, aggregating across experiment repetitions and performances. Optimizers are sorted with respect to the average achieved loss. The relative performance (RP) column indicates incremental relative performance with respect to the best performing optimizer for each budget. Results in this table are for $\lambda_J = 0.2$

B=2000			B=4000			B=7000		
Optimizer	RP	Loss	Optimizer	RP	Loss	Optimizer	RP	Loss
NGO	+0.00%	0.09397	CMA	+0.00%	0.07061	CMA	+0.00%	0.06436
Disc 1+1	+3.13%	0.09692	Shiwa	+4.51%	0.07379	Shiwa	+5.05%	0.06762
CMA	+4.01%	0.09774	NGO	+6.87%	0.07546	NGO	+5.22%	0.06772
Shiwa	+4.43%	0.09814	CMA+Powell	+20.79%	0.08529	CMA+Powell	+10.48%	0.07111
RS-PSO	+15.68%	0.10871	RS-PSO	+21.40%	0.08572	RS-PSO	+20.85%	0.07778
TBPSA	+22.12%	0.11476	Disc 1+1	+27.43%	0.08998	Disc 1+1	+34.99%	0.08688
DE	+25.77%	0.11819	TBPSA	+38.58%	0.09786	TBPSA	+37.28%	0.08836
CMA+Powell	+26.02%	0.11843	DE	+39.82%	0.09873	DE	+48.89%	0.09583
Powell	+92.00%	0.18043	Powell	+82.20%	0.12866	Powell	+70.36%	0.10965

Table 2: *Global fitting:* Average Hausdorff-RMS for experiments with the same budget and optimizer, aggregated across experiment repetitions and performances. Optimizers are sorted with respect to the average Hausdorff-RMS achieved. The relative performance (RP) column indicates incremental relative performance with respect to the best performing optimizer for each budget. Results in this table are for $\lambda_J = 0.2$.

Optimizer	RP	Hauss	Optimizer	RP	Hauss	Optimizer	RP	Hauss
NGO	+0.00%	0.02550	CMA	+0.00%	0.01961	NGO	+0.00%	0.019195
Shiwa	+4.42%	0.02663	Shiwa	+6.55%	0.02089	CMA	+0.21%	0.019234
CMA	+5.49%	0.02690	NGO	+8.66%	0.02131	Shiwa	+7.66%	0.020665
RS-PSO	+15.58%	0.02948	RS-PSO	+18.54%	0.02325	CMA+Powell	+11.02%	0.021309
Disc 1+1	+16.18%	0.02963	CMA+Powell	+22.14%	0.02395	RS-PSO	+15.01%	0.022075
TBPSA	+19.20%	0.03040	DE	+28.98%	0.02529	TBPSA	+20.19%	0.023070
CMA+Powell	+27.92%	0.03262	Disc 1+1	+29.53%	0.02540	Disc 1+1	+22.05%	0.023426
DE	+28.85%	0.03286	TBPSA	+31.74%	0.02583	DE	+33.83%	0.025689
Powell	+128.82%	0.05836	Powell	+122.97%	0.04372	Powell	+90.01%	0.036472

In this section we give additional results and analysis for the *Global* fitting case when $\lambda_J = 0.2$.

In Table 1, we present average loss experiments, with respect to budget and optimizer, aggregated over experiment repetitions and performances, while Table 3 illustrates wilcoxon test results for the same cases. First, in all case budgets, Powel performs worst. For $B = 2000$, the top performing optimizers are NGO, Disc 1+1, CMA and Shiwa. Following them, RS-PSO and TBPSA are ranked 5th and 6th whilst DE and CMA+Powell perform third to last and second to last, respectively. For $B = 4000$ and $B = 7000$, the rankings are exactly the same, with CMA, Shiwa and NGO performing best, CMAPowell being ranked 4th, Disc 1+1 and RS-PSO ranked 5th and 6th, and TBPSA and DE consistently being ranked third to last and second to last, respectively.

In Hausdorff-RMS terms (Table 2), For $B = 2000$, the rankings are similar to the loss terms case, with the exception of Disc 1+1, which while in loss terms is ranked 2nd, in Hausforff-RMS terms is ranked 5th. For $B = 4000$, and $B = 7000$ the rankings are close to the loss terms case. Thus, in general, loss performance in on-par with the objective quality metric Hausdorff-RMS.

Table 3: Wilcoxon test for the *Global fitting case* when $\lambda_J = 0.2$, at a confidence level of 0.05. When the optimizer at cell's row performs better than the optimizer at cell's column, the cell is marked with '+'. Similarly, when the optimizer at cell's row performs worse than the optimizer at cell's column, the cell is marked with '-'. When the performance of the two optimizers corresponding to cell's row and column do not exhibit significant statistical difference, the cell is marked with '≈'

	B=2000								
	CMA+Powell	CMA	DE	Disc 1+1	NGO	Powell	RS-PSO	Shiwa	TBPSA
CMA+Powell	N/A	-	≈	-	-	+	-	-	≈
CMA	+	N/A	+	≈	≈	+	+	≈	+
DE	≈	-	N/A	-	-	≈	-	-	≈
Disc 1+1	+	≈	+	N/A	≈	+	≈	≈	+
NGO	+	≈	+	≈	N/A	+	+	≈	+
Powell	-	-	≈	-	-	N/A	-	-	-
RS-PSO	+	-	+	≈	-	+	N/A	≈	≈
Shiwa	+	≈	+	≈	≈	+	≈	N/A	+
TBPSA	≈	-	≈	-	-	+	≈	-	N/A
	B=4000								
	CMA+Powell	CMA	DE	Disc 1+1	NGO	Powell	RS-PSO	Shiwa	TBPSA
CMA+Powell	N/A	-	+	≈	-	+	≈	-	+
CMA	+	N/A	+	+	≈	+	+	≈	+
DE	-	-	N/A	≈	-	+	-	-	≈
Disc 1+1	≈	-	≈	N/A	-	+	≈	-	≈
NGO	+	≈	+	+	N/A	+	+	≈	+
Powell	-	-	-	-	-	N/A	-	-	-
RS-PSO	≈	-	+	≈	-	+	N/A	-	+
Shiwa	+	≈	+	+	≈	+	+	N/A	+
TBPSA	-	-	≈	≈	-	+	-	-	N/A
	B=7000								
	CMA+Powell	CMA	DE	Disc 1+1	NGO	Powell	RS-PSO	Shiwa	TBPSA
CMA+Powell	N/A	-	+	+	≈	+	≈	≈	+
CMA	+	N/A	+	+	≈	+	+	≈	+
DE	-	-	N/A	≈	-	≈	-	-	≈
Disc 1+1	-	-	≈	N/A	-	+	≈	-	≈
NGO	≈	≈	+	+	N/A	+	+	≈	+
Powell	-	-	≈	-	-	N/A	-	-	≈
RS-PSO	≈	-	+	≈	-	+	N/A	≈	≈
Shiwa	≈	≈	+	+	≈	+	≈	N/A	+
TBPSA	-	-	≈	≈	-	≈	≈	-	N/A

Table 4: *Global* fitting: Pearson correlation coefficient between RMS Hausdorff and Loss, for various λ_J

	λ_J			
	0	**0.05**	**0.1**	**0.2**
Correlation	0.817	0.860	0.863	0.895

2.2 Loss - RMS Hausdorff correlation analysis

In this section, we do a short study on a correlation analysis between loss and RMS-Hausdorff with respect to the silhouette weight coefficient λ_J. We collected data from the converged poses of the optimizers for all budgets, performances and repetitions and calculated the pearson correlation coefficient for four values of $\lambda_J = \{0, 0.05, 0.1, 0.2\}$. The results are shown in Table 4. We observe that higher correlation is achieved as the parameters λ_J increases, with $\lambda_J = 0.2$ resulting into the highest correlation of 0.895.

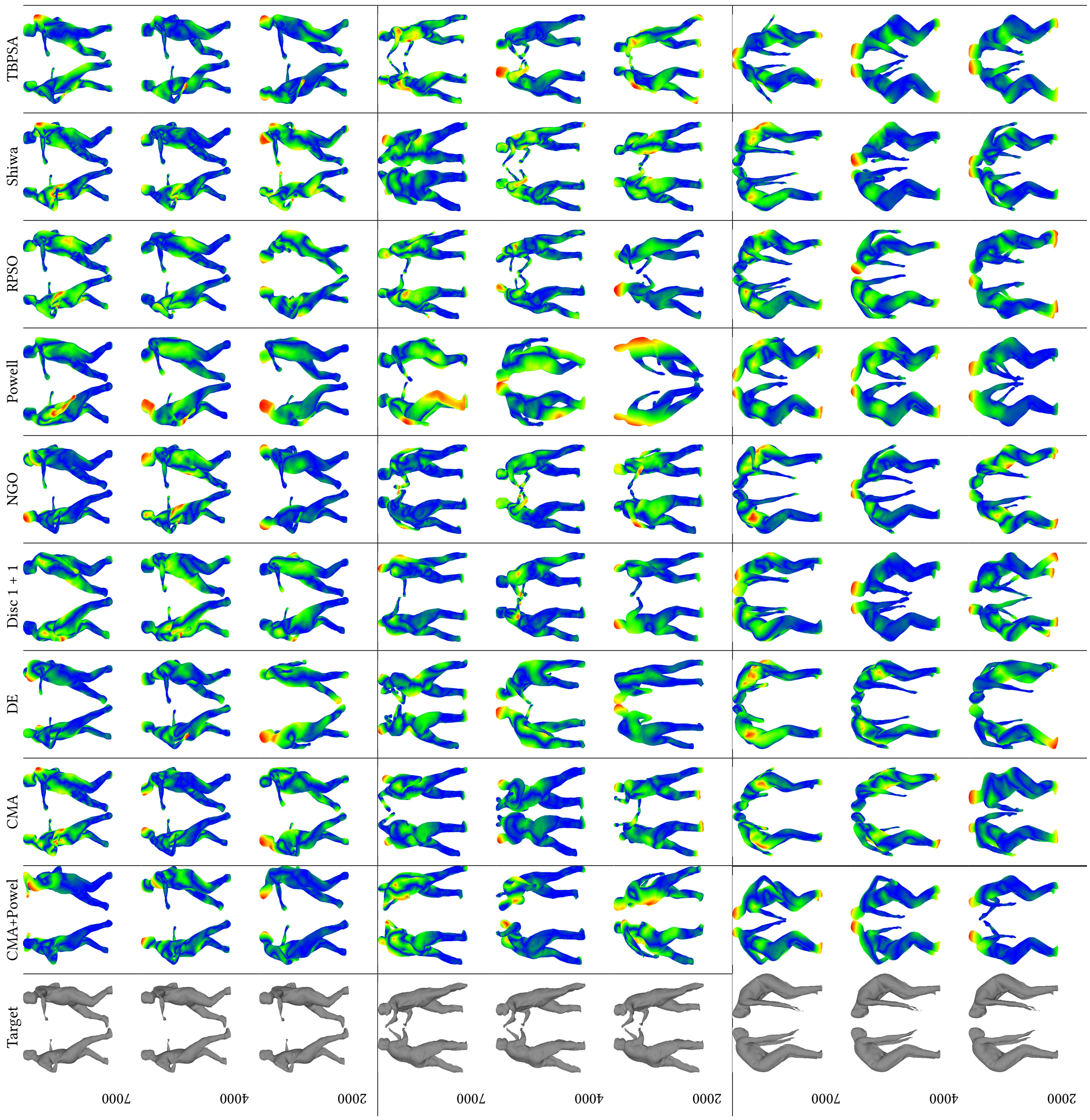

TBPSA
Shiwa
RPSO
Powell
NGO
Disc 1 + 1
DE
CMA
CMA+Powel
Target
7000
4000
2000
7000
4000
2000
7000
4000
2000

Figure 31

2.3 Optimizer exploration strategy visualization

In [1] and [2], we give animated renderings of the optimization process for each optimizer illustrating its search and convergence behaviour across iterations in some global experiments. The (truncated) Chamfer distance is visualized on top of the animated template, while the target mesh is Phong shaded with green color.

REFERENCES

[1] 2021. Supplementary Material 1. https://drive.google.com/file/d/1h4_cV8QyA-rITR_sMdJJGXNjjLAGgh9C/view?usp=sharing.
[2] 2021. Supplementary Material 2. https://drive.google.com/file/d/1LUrGaVQvhik5i4a8X_3kr8IltHmvk43d/view?usp=sharing.