

For the experiment, we chose a termination point of 200 which is twice the length of the previous simulation. We keep the mutation rate at 0.01 for each mutation type. The remaining procedure follows the process described in ??.

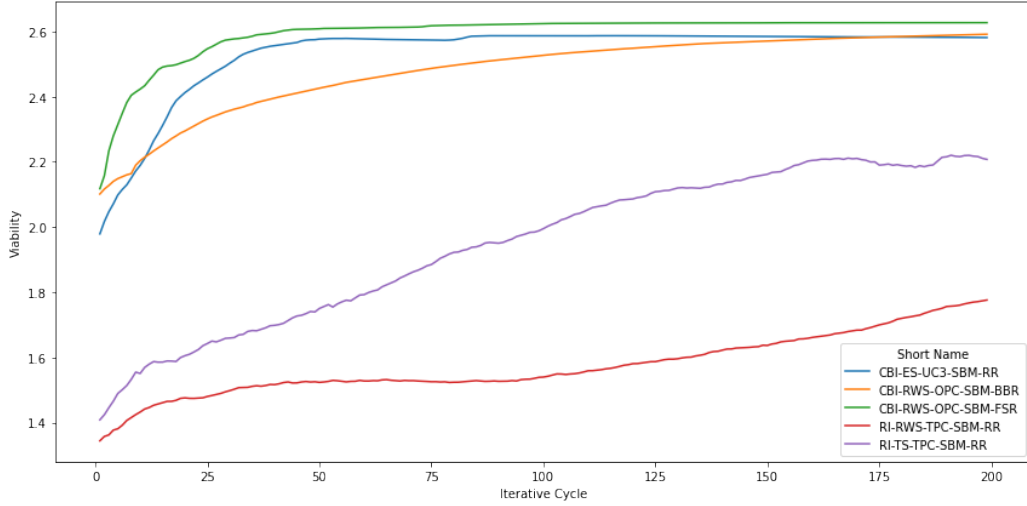


Figure 1: This figure shows the viability across the iteration cycles.

In Figure 1, we see a general increase in viability for each termination point. It shows that increasing the termination point also yields better results at the end of the generation process. We see that *CBI-ES-UC3-SBM-RR* returns the best results in the shortest time span. The model converges after roughly 50 iterative cycles. *CBI-RWS-OPC-SBM-BBR* appears to have not reached convergence. The randomly initiated models have not reached convergence as well. However, they remain far below models that use a more sophisticated method to initialise their population.

Figure 2 shows a decomposed view on how the viability measure evolves. Furthermore, we show the average amount of events within a generated counterfactual. In terms of similarity and sparsity, all models behave similarly. This is no surprise as both measures are inherently interlinked. We see that the randomly initiated models (RI-x) decrease the number of events they generate. Case-based initiated models appear to gain more viability slightly. Although, *CBI-RWS-OPC-SBM-BBR* appears that reaches its saturation point significantly later (100th cycle). Interestingly, the *CBI-RWS-OPC-SBM-BBR* model struggles to maintain feasibility and collapses to near 0 after the 100th iterative cycle. Another surprise is the steep ascension of the only model that uses tournament selection (*RI-TS-TPC-SBM-RR*) towards the end of the generation process. The model even overtakes the model

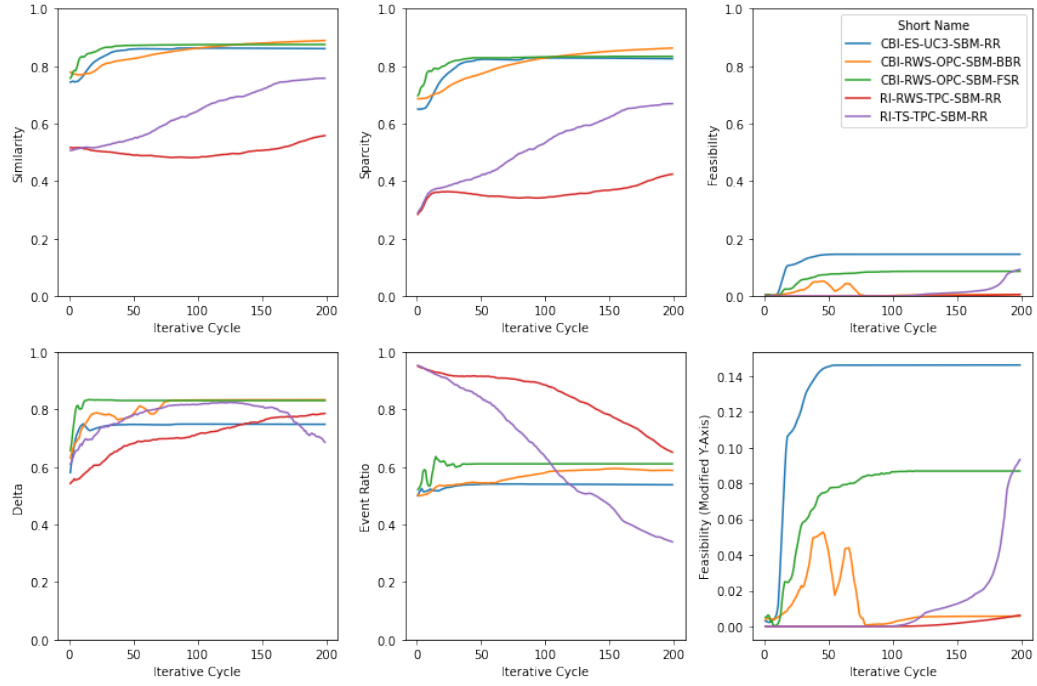


Figure 2: This figure shows the remaining measure components. Additionally, we show the ratio of events within the population. We also show a magnified version of the feasibility measure.

that leads the model configurations in terms of *viability*. Furthermore, we see that *CBI-ES-UC-SBM-RR* has the highest feasibility among all models. However, it also quickly converges after 50 iterative cycles.