Strong benchmark (20 runs, first removed); 95% conf. interval