Referee report on the manuscript

"Behavioral Measures and their Correlation with IPM Iteration Counts on Semi-Definite Programming Problems"
by Robert M. Freund, Fernando Ordonez, and Kim Chuan Toh.

This manuscript is a follow up of [13], where the first two authors tested whether the numbers of iterations taken by an interior-point algorithm to solve a linear programming problem has a connection with certain condition measure of the problem instance. The new manuscript extends [13] in two ways: 1) It performs a similar experimental study for semidefinite programming problems. 2) It examines four different "behavioral" measures (the former paper only tested one).

The empirical results drawn reported in the paper provide an interesting contribution to the literature on condition measures for optimization. In particular, the high correlation between the number of iterations and the "aggregated geometry" measure, as well as the correlation between the number of iterations and Renegar's condition number $C(d)$ provide valuable "experimental evidence" of the relevance of these two measures.

Comments/suggestions:

1) It seems reasonable to use the popular SDPLIB suite as a test set for this experimental study. However, someone could argue that the conclusions drawn solely from this set of problems is likely to be biased. Indeed, the authors themselves acknowledge the likelihood of systematic patterns in the SDPLIB suite (see the last paragraph on page 23 as well as the very last paragraph on page 25). I was a bit surprised to see no attempt whatsoever to generate a pool of unbiased problem instances. The authors could have considered at least some naive approach such as a collection of problem instances whose entries are drawn from independent Gaussian random variables.

2) The choice of norm in (7) is crucial, as it yields tractable expressions to compute or estimate the aggregated geometry measure as well as $C(d)$. However, it is not at all obvious (at least not to this reader) that "$\|\cdot\|_{E1}$" is indeed a norm. The authors should provide a concrete relevant reference or proof.

3) Given that (12) is a measure for the dual problem, wouldn't it make sense to use the dual norm $\|z\|_*$ in the objective?

4) The first statement in the third paragraph of the introduction is misleading: the claim that the "aggregated geometry measure is a good predictor of IPM iterations" suggests that the aggregated geometry measure is some kind of *a priori* estimator. However, the inherent dependence of such measure on the solution set makes it rather *a posteriori* estimator.

5) By contrast to the estimate on the geometry measures, the estimate on the condition number $C(d)$ is fairly inaccurate (it involves a fairly rough approximation of $\|d\|$). Hence it is not that surprising to observe a lower correlation between the number of iterations and the estimated $C(d)$. The authors may want to discuss this more explicitly.

6) Unlike the geometry measures and Renegar's condition number, the theoretical properties of the measures of degeneracy and near absence of strict complementarity are much weaker. Strict complementarity and non-degeneracy are "local" properties of the optimal solutions and are only known to be related to the local convergence of interior-point methods. They have no connection with global properties such as overall complexity. Therefore, it is not at all surprising to observe little or no correlation between these measures and the iteration count of interior-point algorithms. Once again, the authors may want to discuss this.