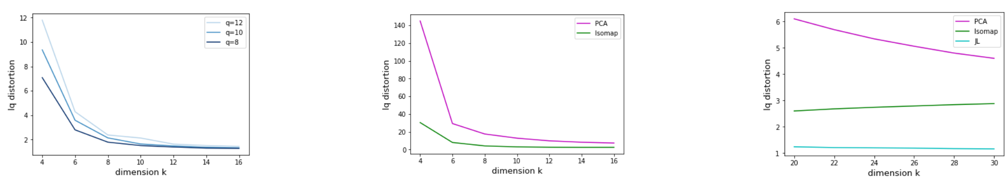
We first address all referees' request to include experimental validation of the theoretical results in the paper.

Empirical Experiments. In all the experiments, we have followed VL [NIPS'18] using similar distributions for sampling random Euclidean input spaces, tests were made for a large range of parameters, averaging over at least 10 independent tests. The results are consistent for all settings, and measures, and will be provided in full in the final paper.

Tightness of the bounds, phase transition phenomenon, and superiority of JL. In our paper we proved theoretical bounds for the distortion measures of the JL transform into $k \geq 1$ dimensions. In particular, we showed that for $q < k$ the ℓ_q-distortion is bounded by $1 + O(1/\sqrt{k}) + O(q/k)$, and all the rest measures are bounded by $O(\sqrt{q/k})$. Particularly, the bounds are independent of n - the size and dimension d of the input data set. In addition, we proved that for the ℓ_q-distortion and REM_q measures a phase transition must occur at $q \sim k$ for any dim. reduction method, where the bounds dramatically increase from being bounded by some constant to grow with n, in particular as $\mathrm{poly}(n)$ for $q > k$.

The graphs in Fig. 1 and Fig.2a describe the following setting: A random X of a fixed size and dimension $n = 800$ was embedded into $k \in [4, 30]$ dimensions, by the JL/PCA/Isomap methods; the value of $q = 10$. We stress that we run many more experiments a wide range of parameter values of $n \in [100, 3000]$, $k \in [2, 100]$, $q \in [1, 10]$, and obtained essentially identical qualitative behavior. In Fig. 1a, the ℓ_q-distortion as a function of k of the JL embedding is shown for $q = 8, 10, 12$. The phase transitions are seen at around $k \sim q$ as predicted by our theorems. In Fig. 1b the bounds and the phase transitions of the PCA and Isomap methods are presented for the same setting ($d = 800, q = 10$), as predicted by our lower bounds. In Fig. 1c, ℓ_q-distortion bounds are shown for increasing values of $k > q$. Note that the ℓ_q-distortion of the JL is a *small constant close to* 1, as predicted, compared to values significantly > 2 for the compared heuristics. Overall, Fig. 1 clearly shows that JL dramatically outperforms the other methods for all the range of values of k. Below is **Fig. 1: Validating ℓ_q-distortion behavior**. The same conclusions as above hold for



(a) Phase transition of JL (b) Phase transition of PCA/Isomap (c) Comparing ℓ_q-dists for $k > q$

σ-distortion as well, as shown in Fig. 2a, on the same sample data set. In the last experiment shown in Fig. 2b, we tested the behavior of the σ-distortion as a function of d-the dimension of the input data set, similarly to that of VL[18](Fig. 2), and tests are shown for embedding dimension $k = 20$ and $q = 2$. According to our theorems, the σ-dist of the JL transform is $O(\sqrt{q/k})$, which is bounded by constant for $q < k$. It is seen that the σ-dist is growing as d increases for both PCA/ISOMAP, whereas it is a *constant* for JL, as predicted. Moreover, JL obtains a *significantly smaller value* of σ-distotion. Below is **Fig. 2: Validating σ-distortion behavior**. In the final paper, we will include further experiments

(a) sigma-distortion (b) sigma-distortion as a function of dimension d

on the JL-based approximation algorithms, which are expected to show similar to more dramatic qualitative behavior.

Discussion/Conclusion Section. Two of the referees #2 and #5 rightfully requested the inclusion of such a section, discussing consequences of the work for practical considerations. We note that a shortened version implicitly appears in the last 3 paragraphs prior to section 1.1 in the supp. material. The discussion section will greatly expand on these.

Further improvement suggestions. We thank referee #2 for his detailed comments that we'll happily incorporate. We shall adopt referee #4's suggestion to use numeric citations (we didn't realize it was possible). Referee #5 asks to improve clarity and writing, in contrast to the others who seem impressed by it. He mentioned "sections that do not exist (on page 3)" - can be found in supp. material. We realize that NIPS has a wide range of audience and we will make an effort to rewrite in a way that will be clear for all. The referee also criticizes the theoretical methodology of the paper, yet the paper contains very detailed proofs for all theorems. The *only* exception mentioned by the referee: proofs of the properties in page 5 will be included in the full paper. In particular, the "translation invariance" property mentioned by the referee, trivially holds for any distance based measure in any metric space *by definition*. We note that we do not "propose a new distortion measure" but new dim. reduction methods, based on JL, which we have addressed above.