presented here. These computational results are consistent with the theoretically known fact that standard PCA is optimal for one single Gaussian distribution. Interestingly, however, the results in Table 1 indicate that ℓ^1 MCDA performs quite well—albeit sub-optimally—for one single Gaussian distribution. Thus, there is no major disadvantage in using ℓ^1 MCDA as a default PCA (instead of standard PCA), since it performs well both for the cases when standard PCA is known to be optimal and, as we have seen, for the many other cases when standard PCA and robust PCAs perform less well, poorly or not at all.

5. Conclusions and Future Work

The assumptions underlying ℓ^1 MCDA are less restrictive and more practical than those underlying standard PCA and currently available robust PCAs. The 2D ℓ^1 MCDA that we have developed differs from standard PCA and all previously proposed robust PCAs in that it (1) allows use of a wide variety of distance functions in the data space (while noting the advantages of using the ℓ^1 norm to define the distance function); (2) replaces all (not just some) of the ℓ^2-based procedures and concepts in standard PCA with ℓ^1-based procedures and concepts; (3) has a theoretical foundation in heavy-tailed statistics but works well for data from both heavy-tailed and light-tailed distributions; (4) is applicable for data that need not have (but can have) mutually orthogonal main directions, can have multiple spokes and can contain patterned artificial outliers (clutter) and (5) does not require assumption of sparsity of the principal components or of the error. Most of the robust PCAs (with the exception of Ke and Kanade's) that have previously been proposed in the literature involve use of the ℓ^1 norm not at all or only to a limited extent and continue to rely on ℓ^2-based items including singular values, inner products, orthogonal projection, averaging and second moments (variances, covariances). The ℓ^1 MCDA that we propose comes exclusively from a unified theoretical framework based on the ℓ^1 norm. The computational results presented in Section 4 show that the ℓ^1 MCDA proposed here significantly outperforms not only the standard PCA but also two robust PCAs in terms of accuracy.

This ℓ^1 MCDA is a foundation for a new, robust procedure that can be used for identification of dimensionality, identification of structure (including nonconventional spoke structure) and data compression in $\mathbb{R}^n$, $n \geq 3$, a topic on which the authors of this paper are currently working. In designing ℓ^1 MCDA for higher dimensions, the guiding principles will continue to be direct connection with heavy-tailed statistics and exclusive reliance on ℓ^1 operations. The higher-dimensional versions of Steps 1 and 2 of ℓ^1 MCDA are feasible as long as appropriate higher-dimensional angular coordinates can be defined. The reader may question whether the "ℓ^1 polar coordinates" that are used in 2D can be extended to n dimensions. Indeed they can in the following manner. In direct analogy to the definition of the polar coordinate θ_m in (8) in two dimensions, one defines angular coordinates one defines $n-1$ angular coordinates of an n-dimensional data point $(x_m^{(1)}, x_m^{(2)}, \ldots, x_m^{(n)})$ to be the quotients with the components of the point as numerators and the ℓ^1 radius $\sum_{j=1}^{n} |x_m^{(j)}|$ of the point as denominators. In passing we note that these higher-dimensional ℓ^1 angular coordinates are computationally cheaper than standard ℓ^2 hyperspherical angular coordinates, which require calculation of square roots of sums of squares.

For samples of M vectors in $\mathbb{R}^n$, $M > n$, the cost of classical PCA is $O(Mn^2)$. The costs of Croux and Ruiz-Gazen's method and of Ke and Kanade's method are not specified in the literature but