Foreword (by Christos H. Papadimitriou)

To project is to throw forward, what a sower does to seeds—an image of deliberate randomness. The Latin root of the word survives in Caesar's famous proclamation on the randomness inherent in life and history, and the decisiveness and skill with which it must be met: "Alea jacta est." Projection is, of course, a plunge onto a lower-dimensional domain (canvas, retina, screen, projective plane). But it need not be an uncomfortable dimensional compromise; a projection can reveal, enhance, evince, illuminate—as long as it is the right projection.

This book is about the radical idea that even a *random* projection is often useful. Two decades ago, analysts stumbled upon a surprising fact of this sort, the Johnson-Lindenstrauss Lemma, as a crucial tool in their project of extending functions in continuous ways. This result (several variants of which are proved in this book, the first in page 2) says that, if you project n points in some high-dimensional space down to a random $O(\log n)$-dimensional plane, the chances are overwhelming that all distances will be preserved—within a small relative error. So, if distance is all you care about, there is no reason to stay in high dimensions!

Over the past ten years, random projection (and its extensions and variants) has been used increasingly often in Computer Science. For example (to mention one application of the technique, due to Luca Trevisan, that is not considered here), the traveling salesman problem in $\log n$ dimensions is as hard to approximate as the same problem in any higher dimensional space—that is to say, very hard.

Random projection and high-dimensional geometry lie at the heart of several important approximation algorithms for problems of a nongeometric nature. For example, the best known algorithm for approximating the maximum cut of a graph works by first embedding the graph in a high-dimensional space (the "best" such embedding can be found by semidefinite programming). A random hyperplane through the origin will then cut a near-maximum number of edges (cutting by a random hyperplane is, of course, tantamount to projecting on a random line, in a sense the ultimate random projection). There are several other important graph theoretic problems, including multicommodity flow, that are best approximated through a geometric consideration (though not always by random projection), and these constitute the subject of the first part of the book.

The second part of the book is about learning—and its geometry. How is it, asks Vempala, that we can learn complex concepts by focusing on the right aspects of the world, ignoring thousands of irrelevant attributes? He speculates that the key to the mystery lies in geometry. What is shown here is that, if concepts are geometrically well separated, or if the natural ways to separate them use criteria (hyperplanes) that are highly interdependent algebraically, then certain natural algorithms converge quickly to the right answer. (It would be very intriguing to verify that the genres of concepts in the learning of which we excel have indeed such

positive properties.) These arguments involve, among other interesting geometric and algebraic techniques, a simpler variant of random projection that is expressly plausible in the context of learning and networks of neurons.

The last part deals with what the author calls "information retrieval": The nearest neighbor problem, for which random projection works so well, and his own geometric take on clustering and principal components analysis (using yet another variant of random projection that is best adapted to the current application).

This is an elegant monograph, dense in ideas and techniques, diverse in its applications, and above all vibrant with the author's enthusiasm for the area. I believe that randomized projection and related geometric techniques will continue to have more and more applications in Computer Science—and that this book, by making the area accessible to graduate students and researchers (sowing the seeds, so to speak), will contribute importantly in this regard.