Discovering cell types in flow cytometry data with random matrix theory YANG SHEN, Chemical Physics Graduate Program, University of Maryland, ROBERT NUSSENBLATT, Laboratory of Immunology, National Eye Institute, National Institutes of Health, WOLFGANG LOSERT, Chemical Physics Graduate Program, University of Maryland — Flow cytometry is a widely used experimental technique in immunology research. During the experiments, peripheral blood mononuclear cells (PBMC) from a single patient, labeled with multiple fluorescent stains that bind to different proteins, are illuminated by a laser. The intensity of each stain on a single cell is recorded and reflects the amount of protein expressed by that cell. The data analysis focuses on identifying specific cell types related to a disease. Different cell types can be identified by the type and amount of protein they express. To date, this has most often been done manually by labelling a protein as expressed or not while ignoring the amount of expression. Using a cross correlation matrix of stain intensities, which contains both information on the proteins expressed and their amount, has been largely ignored by researchers as it suffers from measurement noise. Here we present an algorithm to identify cell types in flow cytometry data which uses random matrix theory (RMT) to reduce noise in a cross correlation matrix. We demonstrate our method using a published flow cytometry data set. Compared with previous analysis techniques, we were able to rediscover relevant cell types in an automatic way.

1Department of Physics, University of Maryland, College Park, MD 20742