Classification of DNA nucleotides with transverse tunneling currents

It has been theoretically suggested and experimentally demonstrated that fast and low-cost sequencing of DNA, RNA, and peptide molecules might be achieved by passing such molecules between electrodes embedded in a nanochannel. The experimental realization of this scheme faces major challenges, however. In realistic liquid environments, typical currents in tunneling devices are of the order of picoamps. This corresponds to only six electrons per microsecond, and this number affects the integration time required to do current measurements in real experiments. This limits the speed of sequencing, though current fluctuations due to Brownian motion of the molecule average out during the required integration time. Moreover, data acquisition equipment introduces noise, and electronic filters create correlations in time-series data. We discuss how these effects must be included in the analysis of, e.g., the assignment of specific nucleobases to current signals. As the signals from different molecules overlap, unambiguous classification is impossible with a single measurement. We argue that the assignment of molecules to a signal is a standard pattern classification problem and calculation of the error rates is straightforward. The ideas presented here can be extended to other sequencing approaches of current interest.