Quantum state estimation [1] is central to many tasks that process quantum information such as the characterization of a quantum source, the verification of the properties of a quantum channel, the monitoring of a transmission line used for quantum key distribution. In state estimation, several copies of the state to be estimated are measured, and then the collected data are used to determine a point estimator—a point (state) in the state space that represents our best guess of the actual state. A popular choice of point estimator is the maximum-likelihood estimator (MLE) [2], the state for which the data are more likely than for any other state.

Since the data have statistical noise, one needs to supplement a point estimator with error bars of some sort—error regions in the state space, more generally, for higher-dimensional problems. Ad-hoc recipes have been proposed for attaching a vicinity of states to a given point estimator, often relying on having a lot of data [3], involve data resampling [4], or consider all data that one might have been observed [5, 6]. By contrast, we present a simple, yet systematic, procedure for determining optimal error regions from only the data that we did observe.

For this purpose, we consider two types of error regions: maximum-likelihood regions (MLRs) and smallest credible regions (SCRs). The MLR is that region of pre-chosen size, for which the given data are more likely than any other region of the same size. In agreement with Ref. [7], we measure the size of a region by its prior probability—the probability that we assign to the region before any data are at hand. The SCR is the smallest (in size) region with the pre-chosen value of the credibility. Here, the credibility of a region is its posterior probability, that is: the probability that the actual state lies in the region, conditioned on the data [8].

It turns out that the problems of finding the MLR and the SCR are duals of each other. Each SCR is also a MLR, and each MLR is a SCR. In the case of MLR, we maximize the joint probability—of finding the state in the region and obtaining the data—for a given size. However, in the case of SCR, we minimize the size for a given joint probability. Both optimizations lead to the same conclusion: The MLR or SCR contains all states for which the likelihood of the data exceeds a certain threshold value $\lambda$. Here, we specify the threshold value $\lambda$ as a fraction of the maximum value of the likelihood. In particular, the MLR (SCR) is a small vicinity of the MLE in the limit of very small size (credibility). In addition, the MLR and the SCR can be used to construct the confidence regions for quantum state estimation presented in Ref. [5].

Quite remarkably and somewhat surprisingly, the shape of the optimal error region is fully determined by the likelihood and the threshold value; however, it is independent of the prior. The prior enters only when the size and credibility are calculated. Much guidance on choosing priors can be found in standard statistics literature; we provide a summary in Ref. [9], focusing on points relevant in quantum contexts.

Since the MLR or SCR is characterized by the threshold value $\lambda$, the error regions can be conveniently and concisely communicated even in higher-dimensional problems: One simply needs to report the size $s_\lambda$ and the credibility $c_\lambda$ as functions of $\lambda$ for the observed data with his or her prior choice. The end users interested in the MLR with the size of his liking or the SCR of her wanted credibility can thus determine the corresponding values of $\lambda$. It is then an easy matter...
For the given data and chosen size or credibility, the MLR or the SCR is a neighborhood of the MLE. In this sense, one can regard them as systematically constructed error regions for the MLE. While there are efficient methods for computing the MLE [2] [11], we are currently lacking equally efficient algorithms for finding the MLR and the SCR. Progress on this front is needed before one can approach high-dimensional problems. We have just begun to enter this unexplored territory and will report progress in due course.

For the given data and chosen size or credibility, the MLR or the SCR is a neighborhood of the MLE. In this sense, one can regard them as systematically constructed error regions for the MLE. While there are efficient methods for computing the MLE [2] [11], we are currently lacking equally efficient algorithms for finding the MLR and the SCR. Progress on this front is needed before one can approach high-dimensional problems. We have just begun to enter this unexplored territory and will report progress in due course.