Data Management Plan

The research results will be shared with the academic community and general public through conference presentations, journal articles, the QMCPy.org blog and posts on the investigators web site. Software generated by the project will be freely available for educational, research and non-profit purposes on Github. When appropriate, preprints will be posted on https://arxiv.org/. Slide decks for talks will be posted on SpeakerDeck or similar platform. Publications, both articles and software, will be uploaded to NSF Public Access Repository. The emphasis on transparency and accountability of data management will be maintained throughout the project. We will maintain records, so that all mathematical calculations and simulations are reproducible.

Products of Research: The main contributions of this research project are the development and analysis of new simulation algorithms, including both theory and software. Details of the main research products will first appear in the investigators' notebooks and preprints. The significant research results will be published in peer-reviewed journal articles, book chapters, or conference proceedings. The primary journals for the published work will include those in computational mathematics, statistics, and the application areas where our algorithms will be used.

Data Format, Content, and Backup: All researchers will maintain research folders with the current state of the project. The electronic files will be backed up on a daily basis. Team members will use a shared Google/Dropbox folder and/or Github repository to provide access to computer files to all team members. In this way, all team members will have access to the ongoing files of the project. These files will contain computer code, text, plots, and images. Google/Dropbox will be used to share data and files among the project researchers and the files will be reasonably organized and coherent. The papers will be written in LaTeX or MS Word, the images will be in a standard format, such as jpg, tiff, or eps. The data files will be in format that is easily readable by a wide community of users, e.g. Excel or ascii text.

Data Access and Sharing: All participants in the project will publish the results of their work. The model data and other supporting materials created or gathered in the course of the work will be shared with other researchers upon reasonable request and within a reasonable time of the request, if the investigators have the authority to share the data. To ensure that data generated with this project is widely available and archived, the estimates for model parameters and references will be included in the research papers, and auxiliary material provided with their publication.

Software developed—in particular QMCPy algorithms and demos—will be available on a public repository. The software will be open to pull requests from those who wish to contribute to its development. This is consistent with our intention to make QMCPy a community-owned library.

Reuse and Redistribution: Public access to research products will be regulated in order to protect privacy and confidentiality concerns, as well to respect any proprietary or intellectual property rights. Legal offices will be consulted on a case-by-case basis to address any concerns, if necessary. Terms of use will include proper attribution to the PIs and authors along with disclaimers of liability in connection with any use or distribution of the research data.

Archiving and Preservation of Access: Research products will be made available immediately after publication. Journal publications will be available online from respective journal websites and

linked to by the PIs' university websites. Again, publications, both articles and software, will be uploaded to NSF Public Access Repository. All computer data and files generated as a result of this project will backed up daily to protect from loss of data from hardware failures, fire, theft, etc.