
Summary

In this thesis I consider the Langevin equation with an alternative mobility matrix. One often considers the mobility matrix in the Langevin equation as a constant matrix, whereas here I consider the mobility matrix as a space-dependent matrix which needs to be updated at every time step. The Langevin equation is a stochastic differential equation which describes the time evolution of a set of changing macroscopic variables. This set of variables changes relatively slow compared to other (microscopic) variables. One of the earliest descriptions with the Langevin equation is the description of Brownian dynamics: the motion of (macroscopic) pollen grains in water due to the rapidly moving water molecules.

In a molecular system, the positions (coordinates) of the larger-sized particles (molecules/atoms) can be considered as a set of macroscopic variables and the positions of the smaller-sized particles as microscopic variables. This relates to the relatively fast movements of the smaller particles with respect to the movements of the larger particles. The subdivision of the variables enables the use of the Langevin equation for describing these coarse-grained models.

In a simulation with conventional molecular dynamics (MD), the atoms/molecules are represented as particles and the movements of these particles are calculated from Newton's equation of motion. However, simulations of a complex system with many different length and time scales based on the fundamental equations of motion take a very long simulation time before capturing the functional and relevant motions. This problem is called critical slowing down. To avoid this problem multi-scale simulation methods are applied, which permits the use of different size time steps and thus enables acceleration of the relevant (slow) movements.

The aim of this thesis is to develop a stochastic quasi Newton (S-QN) method, such that by incorporating multi-scaling the relevant motions are effectively taken with a

larger time step. Due to the integration of the slow motions with a larger time step, critical slowing down can be avoided. At every time step the slowest modes (corresponding to the relevant movements) can be detected and the acceleration of the movements towards these modes is a possibility. However, detecting the different modes of a system at each time step is computationally very expensive. Therefore, a method that automatically changes the mobility matrix in the Langevin equation is developed, such that in the resulting movements the slow modes are automatically integrated with a larger time step.

In Chapter 2, the new mobility matrix is presented: the inverse Hessian of the energy potential. Due to this choice for the mobility matrix, the deterministic term of the Langevin equation is similar to the displacement in the Newton minimization method. We expect, analogous to the performance difference between the Newton method and the steepest descent method, a better convergence towards the minimum of the energy potential compared to a constant mobility matrix. The stochastic term (noise term) in the Langevin equation, constructed using the factorized mobility, is needed for having correct thermodynamic sampling. Simple one and two dimensional systems show that using the alternative mobility matrix is thermodynamically consistent: the potential energy has a Boltzmann distribution. The choice for the inverse Hessian contributes to a faster barrier crossing. This has been determined quantitatively using the mean first passage time, which measures the average time needed to cross a barrier (getting to the top of the energy landscape starting from the bottom). Using the alternative mobility matrix, the mean first passage time is one order of magnitude smaller than for the constant mobility matrix. Since the calculation of the inverse Hessian is computationally expensive, an approximation for the inverse Hessian has been used. This approximation is derived from an existing update scheme in the quasi Newton method: the DFP-update scheme.

In Chapter 3, a factorized secant update (FSU) is constructed. This update is a factorized equivalent of the DFP-update, such that no explicit factorisation of the mobility is needed, which reduces the complexity from $O(n^3)$ to $O(n^2)$. The FSU is constructed such that it forms a basis for the construction of a limited memory version (LFSU), where no matrices are stored and no matrix multiplications are performed. Instead of storing n^2 matrix elements, $3m$ vectors of length n are stored, where m is the truncation parameter or the history depth. Applying the new update scheme to a harmonic chain of particles shows that the update scheme converges to the inverse Hessian. The FSU and LFSU contribute automatically to the acceleration of the different length and time scales, which results in a better sampling performance

compared to the conventional Langevin dynamics (constant mobility matrix). In Chapter 4, the Langevin dynamics, with the mobility matrix constructed by FSU, is applied to a model protein. For the stability of the simulation, the FSU-scheme has been adapted such that regularization occurs for ill-conditioned systems; i.e. systems with large condition numbers (largest eigenvalue of the mobility matrix divided by the smallest eigenvalue). The model protein is based on 3 types of coarse grained particles which either represent a hydrophobic or hydrophilic or a neutral peptide. This protein model enables the determination of sequence-dependent properties of the protein. Using the S-QN method the slow modes are accelerated, which enables faster crossing of energy barriers and samples a larger area of the energy landscape. As a consequence, different inherent states of the protein are found, including the native state. Detailed analysis shows that automatic multi-scaling first minimizes the bond length, bond angle and torsion potentials to its equilibrium value. Afterwards, a sudden drop in the total potential occurs caused by the non-bonded interaction which starts to play a role. The sudden drop in the total potential coincides with a drop in the overlap function. The overlap function is a measure of the amount of (conformation-wise) overlap with a reference structure, for which the native structure is often taken, such that the overlap function is zero for total overlap and one for no overlap at all. The evolution of the overlap function of the partial structures (sub-domains) indicates that these domains are already established before the drop in the total overlap occurs and that the partial structures move collectively. In the conventional Langevin dynamics the collective movements are absent and less inherent states are found due to critical slowing down.

The proposed FSU scheme enables automatic multi-scaling in the Langevin dynamics and contributes to efficient calculation of the noise term. The construction of the FSU also enables the construction of a limited memory version for the mobility. However, the Langevin equation with a space dependent mobility has a spurious drift term, so that calculation of the divergence of the mobility is needed. This term is absent in the conventional Langevin equation (divergence of a constant mobility has no contribution). The computationally expensive calculation of the divergence of the mobility is avoided by using a prediction and correction scheme, which requires the calculation of the inverse mobility. In Chapter 5, the computation of the Langevin equation with space dependent mobility is considered. For the inverse of the mobility a limited memory version is constructed, so that the whole Langevin equation can be written in a limited memory variant. Standard computational methods are of $O(n^3)$, FSU is of $O(n^2)$ and LFSU is of $O(mn)$ and requires less storage. Together with the

multi-scaling feature of (L)FSU, a powerful method for molecular simulations has been provided.