Chapter 6

Geometry and Coarse-Grained Representations of Landscapes

Konstantin Klemm, Jing Qin, and Peter F. Stadler

Abstract. Basic geometric notions describing the structure of landscapes as well as the dynamics of local search on them include basins, saddles, reachability and funnels. We focus on discrete, combinatorial landscapes and emphasize the complications arising from local degeneracies. Local search in such landscapes is well described by adaptive walks, which we use to define reachability of a target from an initial configuration. Reachability introduces a topological structure on the configuration space. Combinatorial vector fields (CVFs) provide a more powerful mathematical framework in which the subtleties of local degeneracy can be conveniently formalized. Stochastic search dynamics has a direct representation as a probability space over the set of CVFs with the given landscape as a Lyapunov function. This ensemble of CVFs is amenable to the framework of standard statistical mechanics. The implications of landscape structure on search dynamics are elucidated further by the fact that the set of all CVFs on a landscape has a product structure, factorizing over extended plateaus (so called shelves) of the landscape. Finally, we discuss the coarse graining of landscapes from two perspectives. Traditionally, a partitioning (e.g. by gradient basins) of a given landscape is used to obtain a landscape with

Konstantin Klemm · Peter F. Stadler
Bioinformatics Group, Department of Computer Science, and Interdisciplinary Center for Bioinformatics, University of Leipzig, Härterstrasse 16-18, 04107 Leipzig, Germany
e-mail: klemm@bioinf.uni-leipzig.de

Jing Qin · Peter F. Stadler
Max Planck Institute for Mathematics in the Sciences,
Inselsetraße 22, 04103 Leipzig, Germany
e-mail: qin@bioinf.uni-leipzig.de

Peter F. Stadler
Fraunhofer Institut for Cell Therapy and Immunology,
Perlickstraße 1, 04103 Leipzig, Germany
Institute for Theoretical Chemistry, University of Vienna,
Währingerstrasse 17, A-1090 Vienna, Austria
Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM87501, U.S.A.
e-mail: studla@bioinf.uni-leipzig.de
fewer configurations called macrostates. A reverse, and less investigated, view on coarse graining considers finer landscapes, with a larger number of configurations than the original one and a non-injective mapping into the original configuration space. Such encodings of landscapes, when suitably defined, turn out advantageous for optimization by adaptive walks.

### 6.1 Introduction

Combinatorial landscape theory provides a framework for the description of the thermodynamics and kinetics of a large class of complex systems. It has proven to be a valuable concept in evolutionary biology, combinatorial optimization and the physics of disordered systems.

The notion of a “fitness landscape” originated in theoretical biology as a technique to visualize evolutionary adaption in 1932 \[39\]. The basic ingredients are a set of discrete genetic structures, a fitness function used to evaluate every possible structure and a “mutation” function measuring the feasibility of transitions between pairs of different structures. Due to the combined effects of mutation and selection, a population moves uphill/downhill on the landscape, which provides evolutionary information in the form of accessibility or reachability. The rationale behind this view of evolution on a landscape gives rise to the inception of evolutionary algorithms for global search or solving combinatorial optimization tasks such as the traveling salesman problem. The equivalent notion of “energy landscapes” arose in physics as a natural description of disordered systems. In spin glasses, for instance, each spin configuration is assigned an energy describing its Hamiltonian which specifies the model \[1\]. In theoretical chemistry, energy landscapes are viewed as discrete models to approximate the smooth potential energy surfaces \[24\]. In structural biology, energy landscapes are used to understand the folding of biopolymers such as RNAs and proteins into their three-dimensional structures \[7\].

In formal terms, a (combinatorial) landscape consists of a search space or configuration space $\mathcal{X} = (V, \mathcal{T})$ and a fitness or energy function $f : V \to \mathbb{R}$ that evaluates each configuration. In general, $\mathcal{T}$ denotes a (generalized) topological structure on $V$. In this contribution we will restrict ourselves to the simplest case, namely undirected finite graphs $G = (V, E)$ as search spaces. Similarly, we will assume that the values of $f$ are real numbers. We refer to \[9\] for some insights into landscapes over recombination spaces and to \[33\] for landscapes whose values are elements of a partially ordered set. For the sake of clarity we adopt the picture of physics and interpret $f$ as an energy function. Optimization thus seeks low energy configurations by dynamics that tends to minimize $f$.

In this contribution we focus on geometric and topological features of landscapes, i.e., on properties that arise from the interplay of the structure of $G$ with the function $f$. These are of particular interest for an understanding of processes on combinatorial landscapes that are governed by local transitions, including in particular a wide variety of heuristic optimization algorithms from simulated annealing to genetic algorithms. Although the relationship between dynamical processes on combinatorial...