University of Strathclyde

Department of Mathematics

Numerical Analysis Group

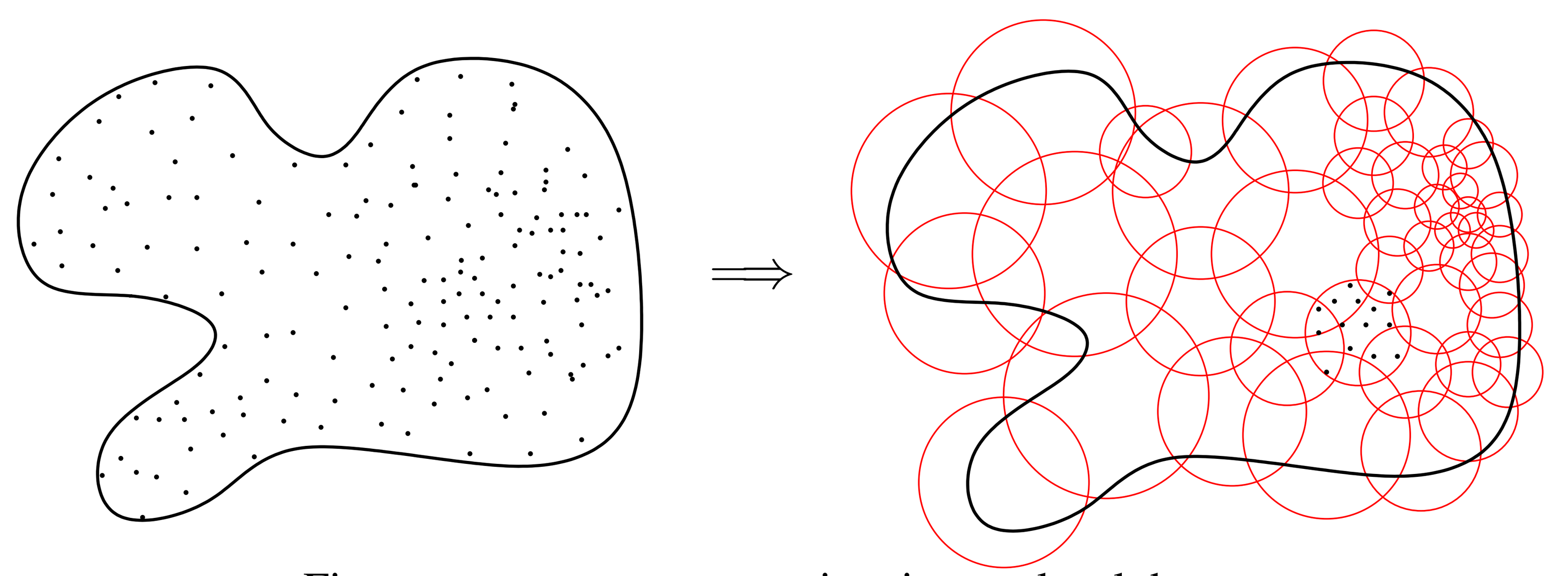
Efficient Fitting of Large Scattered Data

Dr. Oleg Davydov, Department of Mathematics, University of Strathclyde

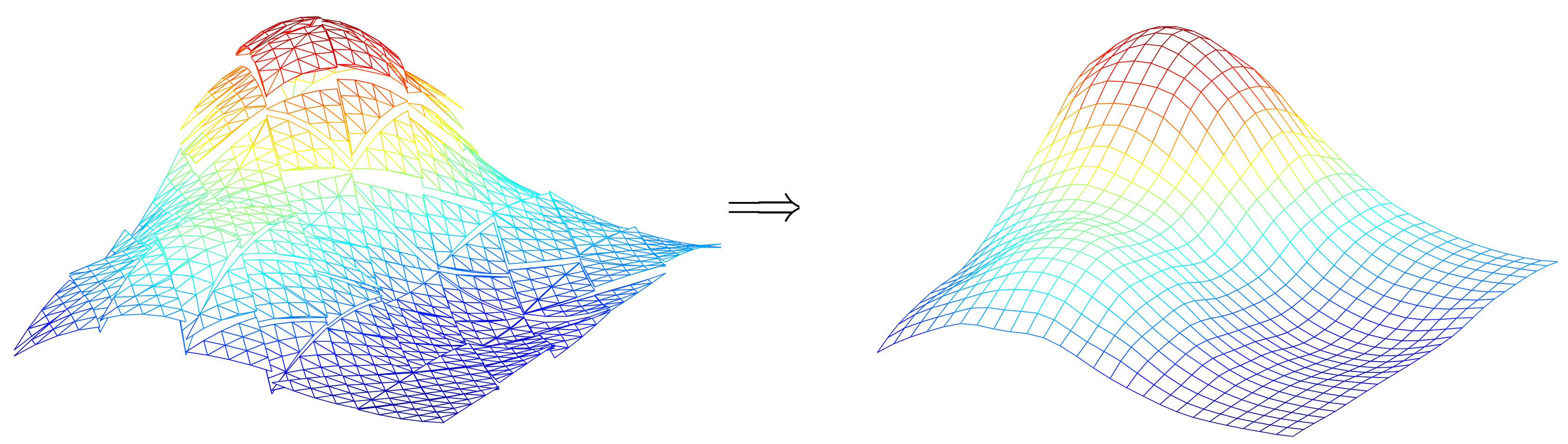
Scattered Data Fitting. Let $\Xi = \{\xi_i\}_{i=1}^N \subset \Omega$ be a finite set of arbitrarily distributed points in a domain $\Omega \subset \mathbb{R}^d$, with a real number z_i assigned to each ξ_i, $i = 1, \dots, N$. The scattered data fitting problem is to find a (smooth) function s defined on Ω that approximates these data, i.e. $s(\xi_i) \approx z_i$, $i = 1, \dots, N$. The quality of the approximation s is not only measured in terms of the errors $|s(\xi_i) - z_i|$ at the data points. It is also expected that s does not exhibit artificial oscillations, and, if the data come from a function f defined on Ω, i.e. $z_i = f(\xi_i)$, $i = 1, \dots, N$, then s approximates f well everywhere on Ω, as far as the information contained in the data allows this.

The scattered data fitting problem arises in numerous applied fields, for instance, in Geosciences, Medical Imaging, Computer Aided Geometric Design, Reverse Engineering, and Data Mining. Typical real world data sets, e.g. those obtained by remote sensing, contain millions of points and therefore require fast and efficient fitting algorithms, ideally of linear complexity $\mathcal{O}(N)$ and of high approximation power.

Two-Stage Methods. The most effective methods for scattered data fitting, such as those based on the idea of energy minimisation, require solving full linear systems or global optimisation problems. Their computational cost becomes prohibitively high for large data. To overcome this difficulty, a two-stage method can be used, where reliable approximations to the local patches of the data are computed in the first stage, and then 'glued together' into a global smooth surface in the second stage, with linear total computational cost.



First stage: compute approximations to local data

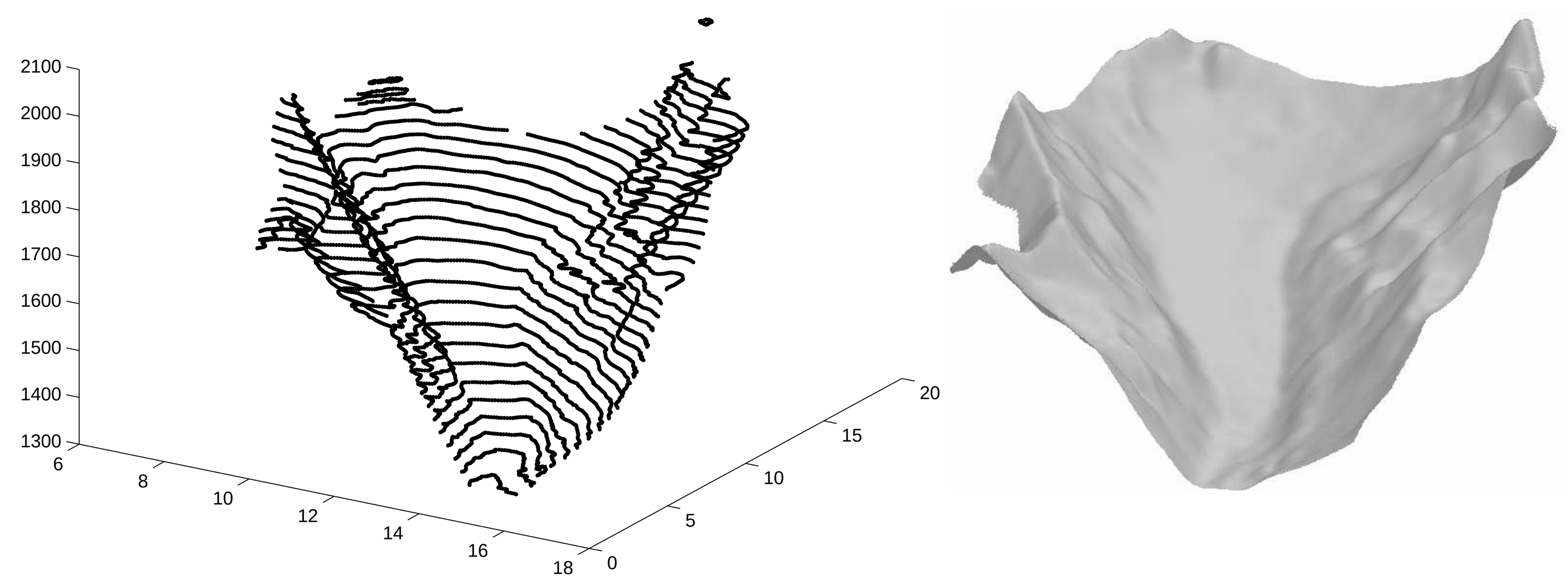


Second stage: gluing together

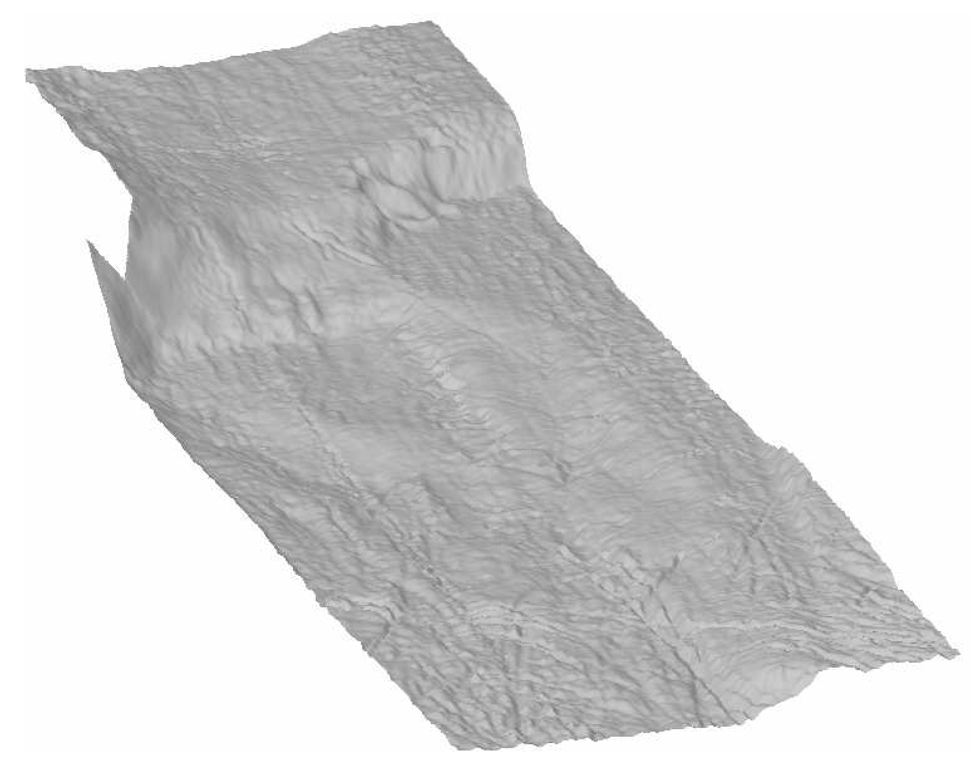
Local Approximations. In the first stage we employ reliable discrete least squares approximations (with polynomials and/or radial basis functions) computed adaptively taking into account recent new approximation error bounds.

Quasi-interpolation. In the second stage quasi-interpolation with multivariate splines is used, thus producing piecewise polynomial surfaces most welcome in computer aided design applicaitons.

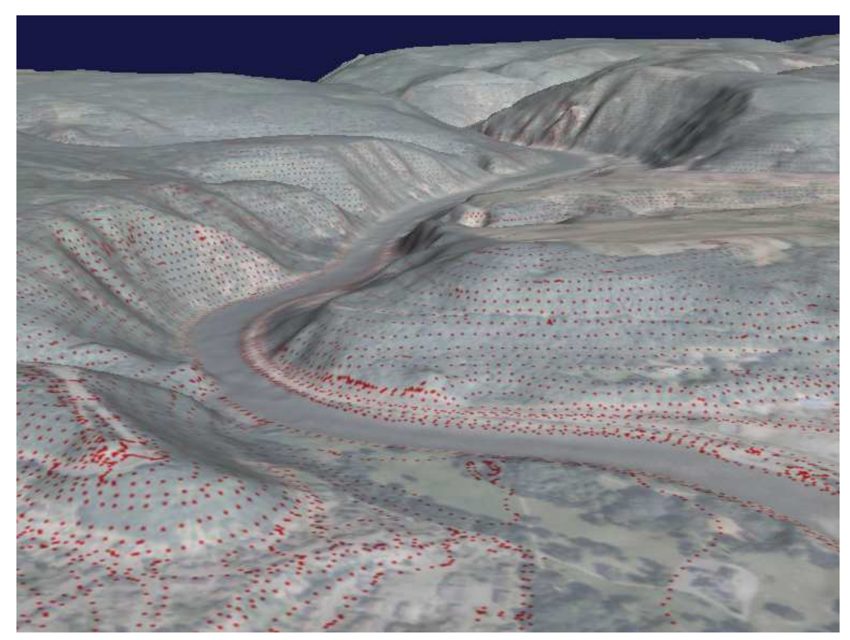
Numerical Examples.



Reconstruction of the surface of a glacier from contours



Reconstruction of Rotterdam's harbour floor from multibeam echosounder data



Reconstruction of a terrain from topography data

Current and Future Work. Interesting research topics include for example (a) development of adaptive algorithms that automatically adjust the parameters, such as the degree of the approximation or the spline mesh, to the data features; (b) development of anisotropic methods specifically adjusted to data possessing strong local directional features; (c) compression of the resulting spline functions using wavelet type approaches; (d) fitting data on manifolds, such as global topography or aircraft surface pressure; (e) surface denoising.

Software: TSFIT, available under GNU General Public License from `http://www.maths.strath.ac.uk/~aas04108/tsfit/index.html`

Collaborations: Max Planck Institute for Informatics (Saarbrücken), University of Florence, University of Mannheim, Vanderbilt University.
Data provided by Landesamt für Kataster-, Vermessungs- und Kartenwesen des Saarlandes and Quality Positioning Services (Zeist).