J. Franke, S. Halim

Wild bootstrap tests for comparing signals and images

Vorwort

Das Tätigkeitsfeld des Fraunhofer-Instituts für Techno- und Wirtschaftsmathematik ITWM umfasst anwendungsnahe Grundlagenforschung, angewandte Forschung sowie Beratung und kundenspezifische Lösungen auf allen Gebieten, die für Techno- und Wirtschaftsmathematik bedeutsam sind.

In der Reihe »Berichte des Fraunhofer ITWM« soll die Arbeit des Instituts kontinuierlich einer interessierten Öffentlichkeit in Industrie, Wirtschaft und Wissenschaft vorgestellt werden. Durch die enge Verzahnung mit dem Fachbereich Mathematik der Universität Kaiserslautern sowie durch zahlreiche Kooperationen mit internationalen Institutionen und Hochschulen in den Bereichen Ausbildung und Forschung ist ein großes Potenzial für Forschungsberichte vorhanden. In die Berichtreihe sollen sowohl hervorragende Diplom- und Projektarbeiten und Dissertationen als auch Forschungsberichte der Institutsmitarbeiter und Institutsgäste zu aktuellen Fragen der Techno- und Wirtschaftsmathematik aufgenommen werden.

Darüber hinaus bietet die Reihe ein Forum für die Berichterstattung über die zahlreichen Kooperationsprojekte des Instituts mit Partnern aus Industrie und Wirtschaft.

Berichterstattung heißt hier Dokumentation des Transfers aktueller Ergebnisse aus mathematischer Forschungs- und Entwicklungsarbeit in industrielle Anwendungen und Softwareprodukte – und umgekehrt, denn Probleme der Praxis generieren neue interessante mathematische Fragestellungen.

Prof. Dr. Dieter Prätzel-Wolters
Institutsleiter

Kaiserslautern, im Juni 2001

Wild bootstrap tests for comparing signals and images

Jürgen Franke and Siana Halim
Department of Mathematics, University of Kaiserslautern,
and Models and Algorithms in Image Processing group, Fraunhofer ITWM
P.O.Box 3049, 67653 Kaiserslautern, Germany

In this expository article, we give an introduction into the basics of bootstrap tests in general. We discuss the residual-based and the wild bootstrap for regression models suitable for applications in signal and image analysis. As an illustration of the general idea, we consider a particular test for detecting differences between two noisy signals or images which also works for noise with variable variance. The test statistic is essentially the integrated squared difference between the signals after denoising them by local smoothing. Determining its quantile, which marks the boundary between accepting and rejecting the hypothesis of equal signals, is hardly possible by standard asymptotic methods whereas the bootstrap works well. Applied to the rows and columns of images, the resulting algorithm not only allows for the detection of defects but also for the characterization of their location and shape in surface inspection problems.

1 Introduction

Bootstrap techniques are general tools for approximating the distribution of statistics of interest. They replace and, for small or moderate sample sizes, frequently improve the classic asymptotic analysis. Typical applications are tests for comparing two random signals or two noisy images with each other. In the following, we present a new bootstrap test motivated by a problem from surface inspection in industrial quality control where detecting and also classifying defects plays a major role. Automatic inspection systems [17] are replacing human inspectors more and more frequently, but there are still enough situations where image analysis algorithms have difficulties to beat the scrutiny of the human eye.

We consider the problem of defect detection in woven textures. For that purpose, various methods have been proposed, e.g. MRF models [5], Independent Component Analysis [22] or exploiting regularity and local orientation [4]. We are, however, not only interested in detecting defects, but also want to get information on the location and shape of the defect areas. Therefore, we treat defect detection as a hypothesis testing problem with the hypothesis representing the absence of defects. As a test case, the following practical problem is considered: the quality of woven fabrics is monitored during production to maintain the quality standards required by customers. Up to fifty different kinds of defects are known to be present, and an automated system of several cameras and corresponding data-processing hard- and software has to detect and classify defects. The algorithms currently available are not completely satisfactory as, e.g., several tuning parameters have to be chosen in a heuristic fashion during implementation of the system [6]).

Woven textiles have a kind of semiregular and additionally noisy surface structure which is hard to reproduce by simple parametric models. Therefore, we adopt a nonparametric

view. In a first step, we smooth the surface under consideration locally, and, then, compare the resulting denoised image with another one derived from a similar specimen or from a different part of the same specimen which is already known to be free of defects. We construct an appropriate test where the significance bound is calculated approximately by the wild bootstrap. We prefer to use this resampling method to the more common residual-based bootstrap as it is simple to implement and works well in a *heteroscedastic* situation, i.e. the variance of the noise is not constant over the image which is a common feature of surface inspection problems, partly due to the effect of illumination [6].

2   Testing with the bootstrap

In the language of statistics, the signal detection or surface inspection problem is formulated as a decision problem. Given data combined to a N-dimensional random vector Y, we have to decide between a hypothesis H_0, e.g. "no defect present on a surface", and the corresponding alternative H_1. Typically, a test is performed by calculating some function $T(Y)$ of the data and comparing it with some bound c_α, chosen as the $(1 - \alpha)$-quantile of the distribution of $T(Y)$ under the hypothesis H_0. If $T(Y) \leq c_\alpha$, we decide for H_0, otherwise for H_1. α is the prescribed probability of an error of the first kind, i.e. under H_0 we have $\mathrm{pr}(T(Y) > c_\alpha) = \alpha$. Constructing the test boils down to the problem of determining c_α. However, for many appealing test statistics $T(Y)$, their distributions under H_0 are not known. The classical escape from this dilemma is provided by asymptotics, i.e. an approximation of the unknown distribution is derived which holds for sample size $N \to \infty$. In simple situations, where the data Y consist of independent random variables with a common distribution known up to a few parameters, this classical approach usually works well. In signal and image analysis, however the structure of the data is frequently too complicated. An asymptotic approximation can be still derived usually, but it frequently is not tractable, depending on unknown quantities itself, and, even more important, it provides reasonable approximations only for sample sizes N which are much larger than in the intended practical applications.

In such a situation, bootstrap tests may be applied. We move from the real world of our data Y to the bootstrap world of a *pseudo data vector* or *resample* Y^* which may be artificially generated from the original data and which has a similar random structure as Y itself. Then, we consider the test statistic $T(Y^*)$ calculated from the bootstrap data Y^*, and determine the $(1\text{-}\alpha)$-quantile c_α^* of its distribution: $\mathrm{pr}^*(T(Y^*) > c_\alpha^*) = \alpha$, where pr^* denotes the conditional probability given the data Y. Then, c_α^* is used as an approximation for the unknown bound c_α, and the bootstrap test decides for the hypothesis H_0 if $T(Y) \leq c_\alpha^*$, and for H_1 else.

In many cases, the bootstrap will work, i.e. the distribution of $T(Y^*)$ given the data Y provides a valid approximation for the distribution of $T(Y)$ for large enough N, and, as a consequence, $c_\alpha^* \approx c_\alpha$. However, this has to be proven, as there are several counter examples [8, 19] where intuitively appealing bootstrap methods do not work without some specific modifications. The applicability of the bootstrap depends on the way how the bootstrap data Y^* are generated as well as on the particular test statistic $T(Y)$ considered. On the other hand, it frequently can be proven by theoretical arguments and illustrated by simulations that the bootstrap provides a much better approximation to the quantities of interest, in our case c_α, than the usual asymptotic approximations.

Except for some rare simple cases, c_α^* cannot be calculated analytically. In the bootstrap world, however, we know in principle everything about the distribution of Y^* and, then, of $T(Y^*)$, as we can artificially generate independent realizations of those random variables.

Therefore, we can calculate quantities of interest numerically by Monte Carlo simulation. For the $(1 - \alpha)$-quantile c_α^*, the numerical algorithm is the following:

1. (a) Generate a realization $Y^*(b)$ of the bootstrap data

    (b) Calculate $T_b^* = T(Y^*(b))$.

    Repeat for $b = 1, \ldots, B$; keep $T_1^*, \ldots, T_B^*$ in storage.

2. Order $T_1^*, \ldots, T_B^*$ w.r.t. size to get $T_{(1)}^* \leq \ldots \leq T_{(B)}^*$.

3. Set $c_{\alpha,B}^* = T_{([1-\alpha)B])}^*$, where $[x]$ denotes the largest integer $\leq x$.

If B is chosen large enough, $c_{\alpha,B}^*$ will get arbitrarily close to c_α^*. Therefore, in the following we do no longer distinguish between the numerical approximation $c_{\alpha,B}^*$ and c_α^* itself, and we use the latter notation for both of them. For evaluating quantiles, a larger number B of resamples is needed than in other applications of the bootstrap, as they are essentially determined by only a small fraction of the largest or smallest values of $T_1^*, \ldots, T_B^*$. For $\alpha = 0.05$, e.g., $B = 1000$ usually suffices, for $\alpha = 0.01$, even $B = 5000$ or 10000 are needed.

3   Generating bootstrap data in regression models

The critical point in applying the bootstrap is the generation of the pseudo data Y^*. If $Y = (Y_1, \ldots, Y_N)$ is just a sequence of independent, identically distributed (i.i.d.) real-valued observation $Y_i, i = 1, \ldots, N$, then we may draw $Y_1^*, \ldots, Y_N^*$ independently from the sample distribution of $Y_1, \ldots, Y_N$, i.e. we set $Y_i^* = Y_{L(i)}, i = 1, \ldots, N$, where $L(1), \ldots, L(N)$ are independent random numbers with $\mathrm{pr}(L(i) = k) = \frac{1}{N}, i, k = 1, \ldots, N$.

In signal or image analysis, we usually have to deal with data which are not identically distributed and/or not independent. Here, we discuss only the first case and make some remarks concerning dependent data at the end. Let us consider a deterministic signal with additive white noise: $Y_i = m_i + \varepsilon_i, i = 1, \ldots, N$, where $\varepsilon_1, \ldots, \varepsilon_N$ are i.i.d. with mean 0 and finite variance σ^2. We assume that m_i changes more slowly with time i than the random errors ε_i such that we can rewrite the data generating model as

$$Y_i = m(x_i) + \varepsilon_i , \quad i = 1, \ldots, N, \tag{1}$$

where $m(x)$ is a reasonably smooth function on the unit interval $[0, 1]$ and $x_i = \frac{i}{N}$ denotes rescaled time.

Here, it makes no sense to consider the sample distribution of $Y_1, \ldots, Y_N$, as these data do not have a common distribution; their means $m_1, \ldots, m_N$ differ. For regression models like (1), we resample from the residuals $\varepsilon_1, \ldots, \varepsilon_N$ instead. As they are not observable directly, we first need an initial approximation $\hat{m}_i^0 = \hat{m}^0(x_i), i = 1, \ldots, N$, of the signal which provides a first approximation $\hat{\varepsilon}_i = Y_i - \hat{m}_i^0, i = 1, \ldots, N$, for the noise variables. Before generating the bootstrap data by random sampling, we have to make one further modification. The true noise distribution has mean 0, whereas the sample distribution of $\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_N$ usual does not. This may lead to a severe distortion in applying the bootstrap [1], and, therefore, we center the empirical residuals around 0 and set $\hat{\varepsilon}_i^0 = \hat{\varepsilon}_i - \frac{1}{N} \sum_{j=1}^N \hat{\varepsilon}_j$. Then, we generate the noise $\varepsilon_1^*, \ldots, \varepsilon_N^*$ in the bootstrap world by randomly drawing from $\hat{\varepsilon}_1^0, \ldots, \hat{\varepsilon}_N^0$. Finally, we get the bootstrap data Y_i^* by adding the noise to the initial signal approximation: $Y_i^* = \tilde{m}_i + \varepsilon_i^*, i = 1, \ldots, N$, where $\tilde{m}_i = \tilde{m}(x_i), i = 1, \ldots, N$, is another preliminary approximation of the signal. We could choose $\hat{m}_i^0 = \tilde{m}_i$, but it is sometimes convenient to
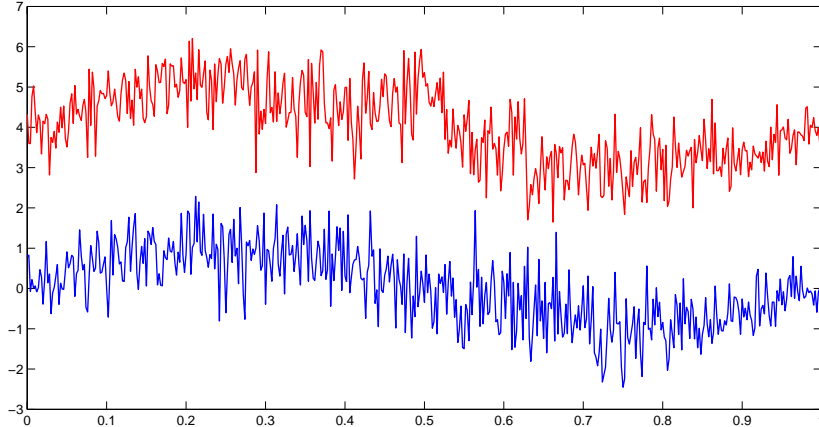
Figure 1: The noisy reference signal (blue) and the noisy disturbed signal (red), the latter shifted by $+4$ to avoid overlap

have a bit more freedom. For the performance of the bootstrap, the choice of $\tilde{m}_i$ is critical as can be seen from the short glimpse at theory in the appendix.

Sometimes, the assumption that the noise variables are identically distributed is not justified, e.g. the average absolute size of errors may be larger if the unknown signal is large. Such an effect may be due to inherent physical reasons, as in the example from image analysis below, or it may be due to an unknown dependence of the observations Y_i not only on the signal m_i but also on some hidden variables. In this case, drawing bootstrap residuals from the sample distribution of $\varepsilon_1, \ldots, \varepsilon_N$ or of its approximations $\hat{\varepsilon}_1^0, \ldots, \hat{\varepsilon}_N^0$ again makes no sense. Instead, for each $i = 1, \ldots, N$, we generate the bootstrap residual ε_i^* based on the information on only one approximate observation, i.e. of $\hat{\varepsilon}_i$. This seems to be a wild idea but it works in theory and in practice. The details of this so-called *wild bootstrap* [18, 19, 20, 23] are the following. We again start from the noise approximations $\hat{\varepsilon}_i = Y_i - \hat{m}_i^0$, $i = 1, \ldots, N$. Then, for each i, we artificially generate a random variable ε_i^* satisfying

$$E^* \varepsilon_i^* = 0, \quad E^* (\varepsilon_i^*)^2 = (\hat{\varepsilon}_i)^2, \quad E^* (\varepsilon_i^*)^3 = (\hat{\varepsilon}_i)^3 \tag{2}$$

which mimicks the random behaviour of ε_i as far as we can get it from only one approximate observation $\hat{\varepsilon}_i$. E^* denotes the expectation conditional on the original data and, therefore, $\hat{\varepsilon}_i$, to be given. A popular choice for ε_i^* is a random variable which assumes only two different values a_i, b_i with probabilities γ and $1 - \gamma$ each. a_i, b_i, γ are uniquely determined by the constraints (2), and we get

$$a_i = \frac{1 - \sqrt{5}}{2} \, \hat{\varepsilon}_i \;, \quad b_i = \frac{1 + \sqrt{5}}{2} \, \hat{\varepsilon}_i \;, \quad \gamma = \frac{5 + \sqrt{5}}{10}.$$

Alternatively, we could choose $\varepsilon_i^* = \{\frac{1}{\sqrt{2}} V_i + \frac{1}{2}(V_i^2 - 1)\} \, \hat{\varepsilon}_i$, $i = 1, \ldots, N$, where $V_1, \ldots, V_N$ are i.i.d. standard normal variables. Other constructions are possible [18]. Once we have $\varepsilon_1^*, \ldots, \varepsilon_N^*$, we proceed as above and get the bootstrap data as $Y_i^* = \tilde{m}_i + \varepsilon_i^*$, $i = 1, \ldots, N$.
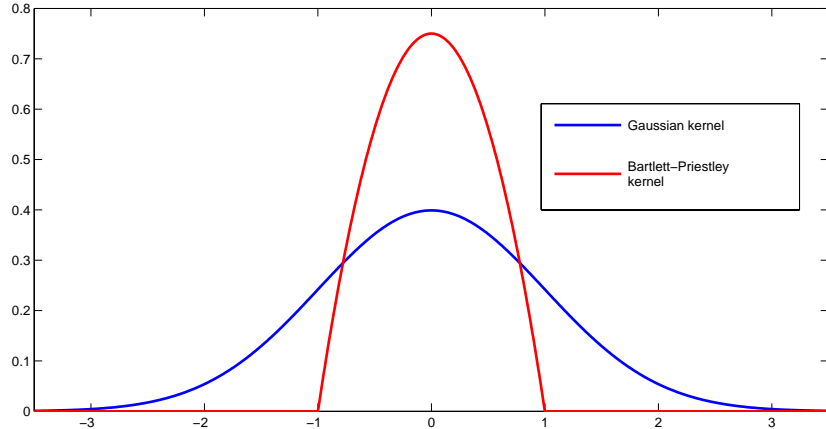
4

Figure 2: Examples of common kernels $K(u) : \frac{3}{4}(1-u^2)_+$ (Bartlett-Priestley or Epanechnikov) and $\frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}u^2)$ (Gaussian)

4 Comparison of two noisy signals

As a first step towards the surface inspection algorithm, we consider the one-dimensional counterpart where we compare two noisy signals of the form (1)

$$Y_i = m_I(x_i) + \varepsilon_i, \qquad \tilde{Y}_i = m_{II}(x_i) + \tilde{\varepsilon}_i, \qquad x_i = \frac{i}{n}, \quad i = 0, ..., n. \tag{3}$$

The noise variables $\varepsilon_i, \tilde{\varepsilon}_i, i = 0, \ldots, n$, are independent with mean zero and finite variances $\mathrm{var}(\varepsilon_i) = \mathrm{var}(\tilde{\varepsilon}_i) = \sigma^2(x_i)$. m_I is a reference signal, and we want to know if the second signal coincides with the first one up to the contribution of noise, i.e. we want to test the hypothesis $H_0 : m_I = m_{II}$.

Later on, the Y_i's and $\tilde{Y}_i$'s will be corresponding rows or columns of two images representing woven textiles. Those surfaces show particular semiregular patterns which are hard to model in a parametric way. Therefore, we do not assume a specific parametric form for m_I, m_{II}, but they may be arbitrary functions required only to have a certain degree of smoothness. Then, (3) is a setup which has been studied extensively in nonparametric regression analysis [7, 12]. As the x_i's are equidistant, we may estimate m_I, m_{II} by Priestley-Chao kernel estimates $\widehat{m}_I, \widehat{m}_{II}$ [21] which are local weighted averages of the data and closely related to kernel spectrum estimates:

$$\widehat{m}_I(x, h) = \frac{1}{n+1} \sum_{i=0}^{n} K_h(x - x_i) Y_i, \qquad \widehat{m}_{II}(x, h) = \frac{1}{n+1} \sum_{i=0}^{n} K_h(x - x_i) \tilde{Y}_i,$$

where $K_h(.) = h^{-1} K(./h)$ denotes a rescaled kernel function. K typically is a probability density symmetric around 0; Figure 2 shows two popular choices. The scaling parameter, the *bandwidth* $h > 0$, controls the smoothness of the function estimates. For sake of simplicity, we consider here only the case where the order h is given by $cn^{-1/5}$ for some constant $c > 0$ which is the optimal rate for minimizing the integrated mean-squared error $\int E(\widehat{m}_I(x, h) - m_I(x))^2 dx$ for $n \to \infty$ [12].

Several tests for the equality of functions based on kernel estimates have been studied, in particular for testing if a certain parametric model adequately describes the data. There, the

nonparametric kernel fit is compared with a parametric function estimate based on the same sample [11, 14]. We adapt this approach to comparing two nonparametric kernel estimates based on two different samples. To perform the test, we look at some distance between the function estimates $\widehat{m}_I(x, h)$ and $\widehat{m}_{II}(x, h)$, and we reject the hypothesis $m_I = m_{II}$ if this distance is too large. We use the standardized integrated squared difference between the two estimates

$$T_n = n\sqrt{h} \int \left(\widehat{m}_{II}(x, h) - \widehat{m}_I(x, h)\right)^2 dx \approx \sqrt{h} \sum_{i=0}^{n} \left(\widehat{m}_{II}(x_i, h) - \widehat{m}_I(x_i, h)\right)^2. \qquad (4)$$

The difference between the representation of T_n as an integral or as a sum is of order $1/n$ and, therefore, negligible [10]. The test rejects the hypothesis $m_I = m_{II}$ on a given level α if $T_n > c_\alpha$. We now use the wild bootstrap to get an approximation c_α^* of c_α.

As a first step of the bootstrap, we estimate the residuals by

$$\hat{\varepsilon}_i = Y_i - \widehat{m}_I(x_i, h), \qquad \hat{\tilde{\varepsilon}}_i = \tilde{Y}_i - \widehat{m}_{II}(x_i, h), \quad i = 0, \dots, n.$$

Then, as described in the previous section, we use these two independent sets of sample residuals to generate two sets of wild bootstrap residuals $\varepsilon_i^*, \tilde{\varepsilon}_i^*, i = 0, \dots, n$. To calculate c_α^* numerically, we have to determine the distribution of our test statistic in the bootstrap world assuming that the hypothesis is satisfied. For that purpose, we consider two resamples for the case where the two signals coincide, i.e. we set

$$Y_i^* = \widehat{m}_I(x_i, g) + \varepsilon_i^*, \qquad \tilde{Y}_i^* = \widehat{m}_I(x_i, g) + \tilde{\varepsilon}_i^*, \quad i = 0, \dots, n. \qquad (5)$$

As an initial approximation of the common signal, we use a Priestley-Chao kernel estimate based on the reference data, but allowing for a different bandwidth g. Using the resamples, we calculate kernel estimates in the bootstrap world

$$\widehat{m}_I^*(x, h) = \frac{1}{n+1} \sum_{i=0}^{n} K_h(x - x_i) Y_i^*, \qquad \widehat{m}_{II}^*(x, h) = \frac{1}{n+1} \sum_{i=0}^{n} K_h(x - x_i) \tilde{Y}_i^*.$$

Mark that, e.g., $\widehat{m}_I^*(x, h)$ estimates $\widehat{m}_I(x, g)$ as, in the bootstrap world, (5) is the counterpart of our original model (3) for the case where H_0 ist satisfied. For the bootstrap to work, we need that the estimation error $\widehat{m}_I(x, h) - m_I(x)$ in the world of real data and $\widehat{m}_I^*(x, h) - \widehat{m}_I(x, g)$ in the bootstrap world of artificially generated data show practically the same stochastic behaviour, at least for large sample sizes n. Therefore, we have necessarily to choose $g \gg h$ or, more precisely, $h, g \to 0, \frac{h}{g} \to 0$ for $n \to \infty$, i.e. $\widehat{m}_I(x, g)$ is an oversmooth estimate of $m_I(x)$. Otherwise, the bias $E \, \widehat{m}_I(x, h) - m_I(x)$ would not be correctly reproduced in the bootstrap world [8].

Finally, the bootstrap test statistic can be constructed as follows. Analogously to (4) we define

$$T_n^* = nh^{1/2} \int \left(\widehat{m}_I^*(x, h) - \widehat{m}_{II}^*(x, h)\right)^2 dx. \qquad (6)$$

We can repeatedly generate realizations of T_n^* to get a Monte Carlo approximation $c_{\alpha, B}^*$ of the quantile c_α^* as described above.

We apply the bootstrap test to compare the two artificially generated noisy signals of Figure 1. Here, $n = 500$, and the two noiseless signals resp. the standard deviation of zero-mean Gaussian noise are given by

$$m_I(x) = \sin(2\pi x), \quad m_{II}(x) = m_I(x) + e^{-800(x-0.5)^2}, \quad \sigma(x) = 0.7 - 1.4(x - 0.5)^2.$$
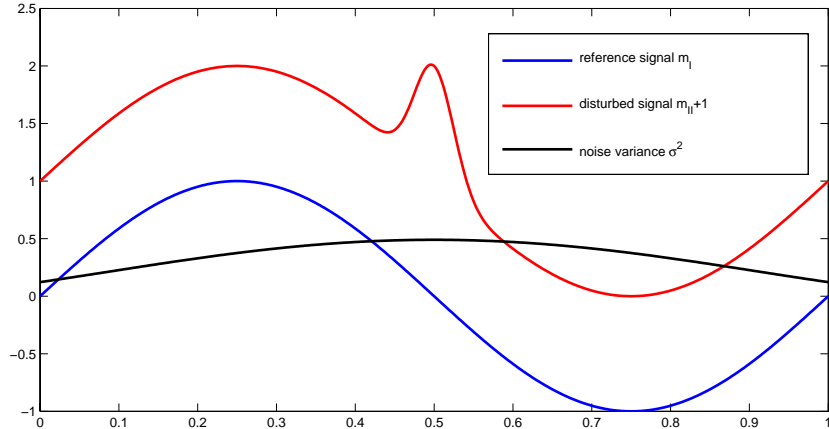
Figure 3: The reference signal $m_I(x)$ and the second signal $m_{II}(x)$ with disturbance, the latter shifted by $+1$, and the noise variance $\sigma^2(x)$

m_{II} differs from m_I essentially by a local disturbance around $x = 0.5$ as can be seen from Figure 3 showing those functions.

We have done a Monte Carlo simulation with $M = 10,000$ independently generated pairs of samples satisfying the hypothesis, resulting in M independent copies of our test statistic under H_0. This allows to estimate the true unknown probability density of the test statistic T_n under H_0 by a kernel density estimate [12] as well as calculating the true quantile $c_\alpha = 1.155$ for $\alpha = 0.95$. For real data, where the data generating process is unknown, we only can get approximations by the bootstrap or by asymptotic approximations, here based on asymptotic normality discussed in the appendix. Figure 4 shows the true density $p(t)$ of T_n for bandwidth $h = 0.02$, its wild bootstrap approximation based on the data of Figure 1, using $g = 0.03$ and $B = 1000$ resamples, and its asymptotic normal approximation, in each case under the hypothesis. For $n = 500$ the bootstrap gets quite close to the truth, whereas the asymptotic approximation does not represent the asymmetry of $p(t)$ well and shows a considerable bias. The approximations of c_α are $c_\alpha^* = 1.243$ by the bootstrap and $c_\alpha^{asy} = 2.104$ by asymptotics. Calculating the test statistic from the data of Figure 1, we get $T_n = 2.5517 \gg c_\alpha^*$, such that the bootstrap test rejects the hypothesis well. Using the asymptotic quantile c_α^{asy}, we also would reject H_0 but not so clearly.

For comparison, Figure 5 shows the the same kind of picture as Figure 4, but for data with a constant noise standard deviation $\sigma(x) = 0.7$ and for $n = 1000$. Doubling the sample size leads to a much better performance of the asymptotic normal approximation, but the bootstrap looks still considerably better.

5   An algorithm for defect detection and localization

We return to the surface inspection problem. We assume that we have a pair of noisy grey-scale images $\mathbf{Y}, \tilde{\mathbf{Y}}$ of same size, the first one known to be without defect, the other one to be tested for the presence of a defect. In practice, both will be corresponding parts of the same textile, where the first one already has been found to be free of defects, and the second part now is to be checked. The images are digitized on a regular grid $\mathbf{x}_{ij} = (\frac{i}{m}, \frac{j}{n}), i =$

Figure 4: Kernel density estimate of the true density of $T_n, n = 500$, under the hypothesis, based on 10,000 Monte Carlo runs, the wild bootstrap approximation of the density based on the data of Figure 1, and the asymptotic normal density with mean B_h^0 and variance σ_T^2
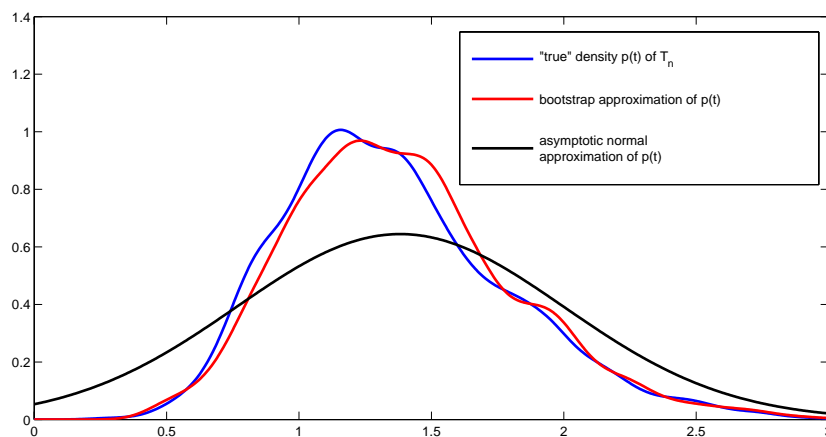

Figure 5: Kernel density estimate of the true density of $T_n, n = 1000$, based on Monte Carlo simulations with constant noise variance, its wild bootstrap approximation and the corresponding asymptotic normal density as in Figure 4
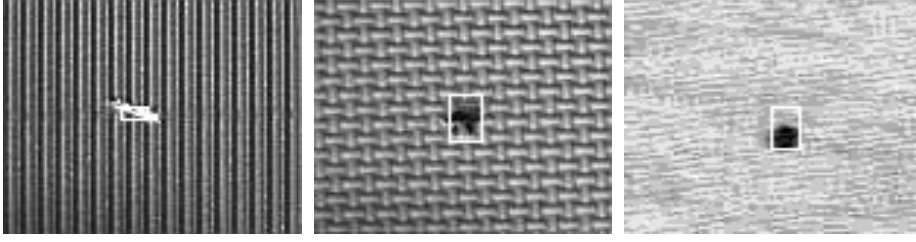
Figure 6: Some examples of defect detection

$0, \ldots, n, j = 0, \ldots, n$, in the unit square

$$\mathbf{Y}_{ij} = \mathbf{m}_I(\mathbf{x}_{ij}) + \varepsilon_{ij}, \qquad \tilde{\mathbf{Y}}_{ij} = \mathbf{m}_{II}(\mathbf{x}_{ij}) + \tilde{\varepsilon}_{ij}, \quad i = 0, \ldots, m, j = 0, \ldots, n. \qquad (7)$$

The random residuals $\varepsilon_{ij}, \tilde{\varepsilon}_{ij}$ are independent with mean zero and finite variances $\mathrm{var}(\varepsilon_{ij}) = \mathrm{var}(\tilde{\varepsilon}_{ij}) = \sigma^2(\mathbf{x}_{ij})$. They do not only represent the observational noise, but also the random fine-scale structure in the woven textile.

To detect and to localize defects, we compare corresponding rows resp. columns of the two images, thus reducing the two-dimensional image testing problem to $(m + 1)(n + 1)$ one-dimensional signal testing problems as in the previous section. The results are summarized as binary vectors $\mathbf{c} \in \{0, 1\}^{n+1}, \mathbf{r} \in \{0, 1\}^{m+1}$ defined as follows.

For any $k = 0, \ldots, n$, we model the two k^{th} columns $Y_i = \mathbf{Y}_{ik}, \tilde{Y}_i = \tilde{\mathbf{Y}}_{ik}, i = 0, \ldots, n$, of $\mathbf{Y}$ resp. $\tilde{\mathbf{Y}}$ like in (3). Then we test if $m_I = m_{II}$ holds or not, using the wild bootstrap test of the previous section. If the hypothesis is rejected, we set $\mathbf{c}_k = 1$, and $\mathbf{c}_k = 0$ else.

Analogously, we compare the two rows $Y_i = \mathbf{Y}_{ki}, \tilde{Y}_i = \tilde{\mathbf{Y}}_{ki}, i = 0, \ldots, n$, and we set $\mathbf{r}_k = 1$ if the hypothesis $m_I = m_{II}$ is rejected here, and $\mathbf{r}_k = 0$ else. If any defect is detected at all, we set

$$i_{min} = \min\{i; \mathbf{r}_i = 1\}, i_{max} = \max\{i; \mathbf{r}_i = 1\}, j_{min} = \min\{j; \mathbf{c}_j = 1\}, j_{max} = \max\{j; \mathbf{c}_j = 1\}.$$

Then the rectangle

$$\big[(i_{min}, j_{min}), (i_{max}, j_{max})\big] = \{(i, j); \ i_{min} \le i \le i_{max}, j_{min} \le j \le j_{max}\}$$

is detected as the defect area of the image.

Figure 6 shows some examples of areas of defect, marked as white rectangles, detected by the algorithm in real textures. In all cases, the image size was 102×96 pixels. We have used as a bandwidth $h \approx 0.4$, and as a reference bandwidth for constructing the bootstrap data $g = 2h$. Using the wild bootstrap, we approximated the c_α^* for level $\alpha = 0.05$ by Monte Carlo simulation based on $B = 500$ artificially generated samples of bootstrap data. Before applying the test, we did some preprocessing of the data, removing linear trends in the column resp. row series data due to illumination [15].

6 Conclusion and extensions

In this contribution, we have derived tests based on the wild bootstrap which allow to check for differences between two irregular signals observed with additive noise where the latter

consists of independent, but not necessarily identically distributed random variables. The signals are first denoised by kernel estimates and, then, compared by looking at the integrated squared difference. The bound between accepting and rejecting the hypothesis of equal signals are determined by the wild bootstrap and numerically calculated by Monte Carlo simulation. The test is applied to all pairwise rows and columns of two images which results in an algorithm which allows to detect defects and to get additionally information on their location and shape.

The idea and theory of the test may be straightforwardly extended to the direct comparison of two images [10]. This is computationally less expensive than doing all the row- and columnwise tests, but it provides less information.

The defect areas which may detected are rectangles parallel to the coordinate axes. The algorithm may be modified by applying the one-dimensional bootstrap test not only to rows and columns, but also to diagonal cuts through the images. This allows for more flexibility at the price of higher computational costs.

Whereas the choice of the kernel K has only minor influence on the performance, the choice of bandwidth h is an important practical problem. We have chosen it based on intuition and experience, but there are many adaptive methods like crossvalidation [12] or plug-in schemes [2] to get h from the data. This presumably has no influence on the validity of the bootstrap test provided that the data-adaptive and therefore random bandwidth $\hat{h}$ stabilizes for increasing sample size in the sense that the relative difference $(\hat{h} - h)/h$ to a deterministic sequence satisfying the necessary assumptions vanishes. Corresponding theoretical results are known for related problems [19].

We have discussed only the case of deterministic signals with additive independent noise which can be formulated as a regression model. The basic ideas of the residual-based or wild bootstrap may be applied equally well to dependent data, e.g. to kernel spectrum estimates [8] or kernel estimates of the autoregressive function in nonlinear autoregression $X_t = m(X_{t-1}, \ldots, X_{t-p}) + \varepsilon_t$ [9].

7   Appendix: Some theoretical considerations

The bootstrap provides a straightforward method for deriving formal tests from intuitive ideas which are quite good even for smaller sample sizes. However, it does not work automatically and, for each specific application, has to be justified by theoretical arguments or at least by extensive simulation studies.

As an example from [19], where a seemingly straightforward bootstrap approach fails, let us consider a regression model similar to the signal detection problem of section 4. For i.i.d. data (X_i, Y_i), $i = 0, \ldots, n$, we have $Y_i = m(X_i) + \varepsilon_i$ where, given X_i, the residuals ε_i have mean 0 and finite variance which may depend on X_i. We have some idea about the function $m(x)$, e.g. that it is a polynomial of degree 3: $m(x) = m_b(x) = b_1 + b_2 x + b_3 x^2 + b_4 x^3$ for some $b = (b_1, \ldots, b_4)$, but we are not sure and want to test this particular modeling assumption. An intuitive idea is estimating b by, e.g., a least-squares estimate $\hat{b}$ and comparing the resulting parametric function estimate $m_{\hat{b}}(x)$ with a nonparametric function estimate $\hat{m}(x, h)$ like in section 4 which does not assume a specific form of $m(x)$. As, here, the X_i are random and, in particular, not equidistant, we have to use the Nadaraya-Watson estimate [12] instead of

the Priestley-Chao estimate

$$\hat{m}(x,h) = \frac{\sum_{i=0}^{n} K_h(x - X_i) Y_i}{\sum_{i=0}^{n} K_h(x - X_i)} \ .$$

Intuitively, if our model assumption that $m(x)$ is of the parametric form $m_b(x)$ holds, we should have $m_{\hat{b}}(x) \approx \hat{m}(x,h)$. Looking a bit closer at both estimates, we see that $\hat{m}(x,h)$ has the typical bias of local smoothers whereas $m_{\hat{b}}(x)$ is unbiased if the parametric model is not misspecified. Therefore, we should not compare $\hat{m}(x,h)$ with $m_{\hat{b}}(x)$, but with a correspondingly smoothed version $m_{\hat{b}}(x,h)$ which has a bias of the same order as $\hat{m}(x,h)$:

$$m_{\hat{b}}(x,h) = \frac{\sum_{i=0}^{n} K_h(x - X_i) m_{\hat{b}}(X_i)}{\sum_{i=0}^{n} K_h(x - X_i)} \ .$$

For comparing $m_{\hat{b}}(x,h)$ with $\hat{m}(x,h)$ we consider the same type of distance as in (4), i.e.

$$T_n = n\sqrt{h} \int (\hat{m}(x,h) - m_{\hat{b}}(x,h))^2 dx.$$

As the data are i.i.d., we may generate bootstrap data (X_i^*, Y_i^*), $i = 0, \ldots, n$, by drawing them independently from the sample distribution of $(X_0, Y_0), \ldots, (X_n, Y_n)$. However, the resulting test statistic T_n^* in the bootstrap world **does not** provide an adequate approximation of the distribution of T_n. E.g., asymptotically for $n \to \infty$, the variance of T_n^* is 3 times as large as the variance of T_n [19], and consequently bootstrap quantiles c_α^* are far from the true quantiles c_α which we need for the test. So, this naive bootstrap fails here. The wild bootstrap, however, leads to a valid bootstrap test in this situation too [19].

We have seen that resampling techniques like the bootstrap are not guranteed to provide valid approximations of distributional parameters like the mean-square error of estimators, confidence intervals or significance bounds c_α of test statistics. So, some theory is useful to convince ourselves that the bootstrap does not lead to false conclusions. We illustrate the type of arguments by studying the distributions of the test statistic T_n of (4) and its bootstrap counterpart T_n^* of (6). For the validity of the bootstrap approximation of c_α, those two distributions have to be close to each other only under the hypothesis $m_I = m_{II}$. However, it is useful to know that the bootstrap test has not only approximately the right level, but that it also has a similar power. Specifically, we consider local alternatives of the form

$$m_I(x) \neq m_{II}(x) = m_I(x) + (n\sqrt{h})^{-1/2} \Delta_n(x) \tag{8}$$

for some sequence of functions $\Delta_n(x)$ which are uniformly bounded in n and x. For $\Delta_n(.) \equiv 0$, this incluues the hypothesis. So, m_{II} gets closer and closer to m_I for $n \to \infty$, i.e. the signal detection problem becomes harder and harder. That type of asymptotics provides more insight into the performance of tests, as, for a fixed distance between m_I, m_{II}, even mediocre tests would detect the violation of the hypothesis with probability going to 1.

Under suitable technical assumptions on the signals m_I, m_{II}, on the kernel and on the noise, we may show [10] as an auxiliary result that the distribution of T_n is approximately normal with mean B_h and variance σ_T^2 given by

$$B_h = B_h^0 + B_h^1,$$

$$B_h^0 = \frac{2}{\sqrt{h}} \int \sigma^2(x) dx \int K^2(u) du,$$

$$B_h^1 = \int (K_h \star \Delta_n(x))^2 dx,$$

$$\sigma_T^2 = 8 \int \sigma^4(x) dx \, K^{(\star 4)}(0).$$

In particular, $B_h^1 \geq 0$ and, under the hypothesis $m_I = m_{II}$, $B_h^1 = 0$. $\star$ denotes the convolution of two functions, and $K^{(\star 4)}(x)$ denotes the 4-fold convolution of K with itself.

Mark that the variance of T_n converges to the finite σ_T^2 for $n \to \infty$ whereas the mean diverges as $h \to 0$. This does not matter for the test, as the difference B_h^1 between the mean under the hypothesis $m_I = m_{II}$ and under the alternative (8) is of order h^{-1} and, therefore, of larger order than B_h^0, as, e.g., for $\Delta_n(x) \equiv \delta$ we have $B_h^1 = \delta^2 \int K_h^2(x)\, dx = h^{-1}\delta^2 \int K^2(u)\, du$.

We could use the asymptotic normality of T_n instead of the bootstrap to get an approximation of the signifance bound c_α and, then, a valid test. However, that is not directly possible as, even under the hypothesis, $B_h = B_h^0$ and σ_T^2 depend on the unknown function $\sigma^2(x)$ characterizing the variable noise variance. This function could be estimated by local smoothing too and plugged into the above equations, but the accumulating approximation errors have the effect that the normal approximation of the distribution of T_n is good enough only for extremely large sample sizes n - compare Figures 4 and 5, where even the normal approximation (black curve) with known mean B_h^0 and variance σ_T^2 is not performing well. So, the bootstrap really helps if it works.

The validity of the wild bootstrap approximation of the test statistics T_n is given by the following result. Again, under appropriate assumptions, we can show that the so-called Mallows distance between the distribution of T_n and the distribution of T_n^*, the latter conditional on the original data, converges to 0 for $n \to \infty$ [10]. This is based on the fact that the estimation errors $\widehat{m}_I(x,h) - m_I(x)$ in the real world and $\widehat{m}_I^*(x,h) - \widehat{m}_I(x,g)$ in the bootstrap world both converge to 0 with rate $(nh)^{-1/2}$, whereas the distance between their distributions converges to 0 faster [3, 13]. Under the hypothesis, where $m_I(x) = m_{II}(x)$, this extends correspondingly to $\widehat{m}_I(x,h) - \widehat{m}_{II}(x,h)$ and, then, to T_n.

References

[1] P. J. Bickel and D. A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Statist.* 9: 1196–1217, 1981.

[2] M. Brockmann, T. Gasser and E. Hermann. Locally adaptive bandwidth choice for kernel regression estimators. *J. Amer. Statist. Assoc.* 88: 1302–1309, 1993.

[3] R. Cao-Abad. Rate of convergence for the wild bootstrap in nonparametric regression. *Ann. Statist.* 19: 2226–2231, 1991.

[4] D. Chetverikov and A. Hanbury. Finding defects in texture using regularity and local orientation. *Pattern Recognition* 35: 2165–2180, 2002.

[5] F. S. Cohen, Z. Fan and S. Attali Automated inspection of textile fabrics using textural models. *IEEE Trans. Pattern Analysis and Machine Intelligence* 13: 803–808, 1991.

[6] C. Daul, R. Rösch, B. Claus, J. Grotepaß, U. Knaak and R. Föhr A fast image processing algorithm for quality control of woven textiles. In: *Mustererkennung, DAGM 1998*, ed. P. Levi et al., Springer, Heidelberg, 1998.

[7] J. Fan and I. Gijbels. *Local Polynomial Modelling and its Applications.* Chapman and Hall, London, 1988.

[8] J. Franke and W. Härdle. On bootstrapping kernel spectrum estimates. *Ann. Statist.* 20: 121–145, 1992.

[9] J. Franke, J. P. Kreiss and E. Mammen. Bootstrap of kernel smoothing in nonlinear time series. *Bernoulli* 8: 1–37, 2002.

[10] J. Franke and S. Halim. A bootstrap test for comparing images in surface inspection. Preprint 150, DFG-Priority Program 1114

[11] P. Hall and J. D. Hart. Bootstrap test for difference between means in nonparametric regression. *J. Amer. Statist. Assoc.* 85: 1039–1049, 1990.

[12] W. Härdle. *Applied Nonparametric Regression.* Cambridge University Press, Cambridge, 1990.

[13] W. Härdle and J. S. Marron. Bootstrap simultaneous error bars for nonparametric regression. *Ann. Statist.* 19: 778–796, 1991.

[14] W. Härdle and E.Mammen. Comparing nonparametric versus parametric regression fits. *Ann. Statist.* 21: 1926–1947, 1993.

[15] S. Halim. *Spatially Adaptive Detection of Local Disturbances in Time Series and Stochastic Processes on the Integer Lattice Z^2.* PhD Thesis, University of Kaiserslautern, 2005.

[16] P. Hall. *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, Berlin-Heidelberg-New York, 1992.

[17] K. Kohrt. Automatische Qualitätskontrolle - Bildverarbeitung in der Industrie. *Keramische Zeitschrift* 2: 2005.

[18] R. Liu. Bootstrap procedures under some non i.i.d models. *Ann. Statist.* 16: 1696–1708, 1988.

[19] E. Mammen. *When Does Bootstrap Work?.* Lecture Notes in Statistics, Springer, Berlin-Heidelberg-New York, 1992.

[20] E. Mammen. Bootstrap and wild bootstrap for high-dimensional linear models. *Ann. Statist.* 21: 255–285, 1993.

[21] M. B. Priestley and M. T. Chao. Nonparametric function fitting. *J. Royal Stat. Soc., Ser. B,* 34: 385–392, 1972.

[22] O. G. Sezer, A. Ertüzün and A. Erçil. Using perceptual relation of regularity and anisotropy in the texture with independent component model for defect detection. *Pattern Recognition* 40: 121–133, 2007.

[23] C. F. J. Wu. Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Statist.* 14: 1261–1295, 1986.

Published reports of the Fraunhofer ITWM

The PDF-files of the following reports are available under:
www.itwm.fraunhofer.de/de/zentral__berichte/berichte

1. D. Hietel, K. Steiner, J. Struckmeier

A Finite - Volume Particle Method for Compressible Flows

(19 pages, 1998)

2. M. Feldmann, S. Seibold

Damage Diagnosis of Rotors: Application of Hilbert Transform and Multi-Hypothesis Testing

Keywords: Hilbert transform, damage diagnosis, Kalman filtering, non-linear dynamics
(23 pages, 1998)

3. Y. Ben-Haim, S. Seibold

Robust Reliability of Diagnostic Multi-Hypothesis Algorithms: Application to Rotating Machinery

Keywords: Robust reliability, convex models, Kalman filtering, multi-hypothesis diagnosis, rotating machinery, crack diagnosis
(24 pages, 1998)

4. F.-Th. Lentes, N. Siedow

Three-dimensional Radiative Heat Transfer in Glass Cooling Processes

(23 pages, 1998)

5. A. Klar, R. Wegener

**A hierarchy of models for multilane vehicular traffic
Part I: Modeling**

(23 pages, 1998)

Part II: Numerical and stochastic investigations

(17 pages, 1998)

6. A. Klar, N. Siedow

Boundary Layers and Domain Decomposition for Radiative Heat Transfer and Diffusion Equations: Applications to Glass Manufacturing Processes

(24 pages, 1998)

7. I. Choquet

**Heterogeneous catalysis modelling and numerical simulation in rarified gas flows
Part I: Coverage locally at equilibrium**

(24 pages, 1998)

8. J. Ohser, B. Steinbach, C. Lang

Efficient Texture Analysis of Binary Images

(17 pages, 1998)

9. J. Orlik

Homogenization for viscoelasticity of the integral type with aging and shrinkage

(20 pages, 1998)

10. J. Mohring

Helmholtz Resonators with Large Aperture

(21 pages, 1998)

11. H. W. Hamacher, A. Schöbel

On Center Cycles in Grid Graphs

(15 pages, 1998)

12. H. W. Hamacher, K.-H. Küfer

Inverse radiation therapy planning - a multiple objective optimisation approach

(14 pages, 1999)

13. C. Lang, J. Ohser, R. Hilfer

On the Analysis of Spatial Binary Images

(20 pages, 1999)

14. M. Junk

On the Construction of Discrete Equilibrium Distributions for Kinetic Schemes

(24 pages, 1999)

15. M. Junk, S. V. Raghurame Rao

A new discrete velocity method for Navier-Stokes equations

(20 pages, 1999)

16. H. Neunzert

Mathematics as a Key to Key Technologies

(39 pages (4 PDF-Files), 1999)

17. J. Ohser, K. Sandau

Considerations about the Estimation of the Size Distribution in Wicksell's Corpuscle Problem

(18 pages, 1999)

18. E. Carrizosa, H. W. Hamacher, R. Klein, S. Nickel

Solving nonconvex planar location problems by finite dominating sets

Keywords: Continuous Location, Polyhedral Gauges, Finite Dominating Sets, Approximation, Sandwich Algorithm, Greedy Algorithm
(19 pages, 2000)

19. A. Becker

A Review on Image Distortion Measures

Keywords: Distortion measure, human visual system
(26 pages, 2000)

20. H. W. Hamacher, M. Labbé, S. Nickel, T. Sonneborn

Polyhedral Properties of the Uncapacitated Multiple Allocation Hub Location Problem

Keywords: integer programming, hub location, facility location, valid inequalities, facets, branch and cut
(21 pages, 2000)

21. H. W. Hamacher, A. Schöbel

Design of Zone Tariff Systems in Public Transportation

(30 pages, 2001)

22. D. Hietel, M. Junk, R. Keck, D. Teleaga

The Finite-Volume-Particle Method for Conservation Laws

(16 pages, 2001)

23. T. Bender, H. Hennes, J. Kalcsics, M. T. Melo, S. Nickel

Location Software and Interface with GIS and Supply Chain Management

Keywords: facility location, software development, geographical information systems, supply chain management
(48 pages, 2001)

24. H. W. Hamacher, S. A. Tjandra

Mathematical Modelling of Evacuation Problems: A State of Art

(44 pages, 2001)

25. J. Kuhnert, S. Tiwari

Grid free method for solving the Poisson equation

Keywords: Poisson equation, Least squares method, Grid free method
(19 pages, 2001)

26. T. Götz, H. Rave, D. Reinel-Bitzer, K. Steiner, H. Tiemeier

Simulation of the fiber spinning process

Keywords: Melt spinning, fiber model, Lattice Boltzmann, CFD
(19 pages, 2001)

27. A. Zemitis

On interaction of a liquid film with an obstacle

Keywords: impinging jets, liquid film, models, numerical solution, shape
(22 pages, 2001)

28. I. Ginzburg, K. Steiner

Free surface lattice-Boltzmann method to model the filling of expanding cavities by Bingham Fluids

Keywords: Generalized LBE, free-surface phenomena, interface boundary conditions, filling processes, Bingham viscoplastic model, regularized models
(22 pages, 2001)

29. H. Neunzert

**»Denn nichts ist für den Menschen als Menschen etwas wert, was er nicht mit Leidenschaft tun kann«
Vortrag anlässlich der Verleihung des Akademiepreises des Landes Rheinland-Pfalz am 21.11.2001**

Keywords: Lehre, Forschung, angewandte Mathematik, Mehrskalenanalyse, Strömungsmechanik
(18 pages, 2001)

30. J. Kuhnert, S. Tiwari

Finite pointset method based on the projection method for simulations of the incompressible Navier-Stokes equations

*Keywords: Incompressible Navier-Stokes equations, Meshfree method, Projection method, Particle scheme, Least squares approximation
AMS subject classification: 76D05, 76M28*
(25 pages, 2001)

31. R. Korn, M. Krekel

Optimal Portfolios with Fixed Consumption or Income Streams

Keywords: Portfolio optimisation, stochastic control, HJB equation, discretisation of control problems.
(23 pages, 2002)

32. M. Krekel

Optimal portfolios with a loan dependent credit spread

Keywords: Portfolio optimisation, stochastic control, HJB equation, credit spread, log utility, power utility, non-linear wealth dynamics
(25 pages, 2002)

33. J. Ohser, W. Nagel, K. Schladitz

The Euler number of discretized sets – on the choice of adjacency in homogeneous lattices

Keywords: image analysis, Euler number, neighborhod relationships, cuboidal lattice
(32 pages, 2002)

34. I. Ginzburg, K. Steiner

Lattice Boltzmann Model for Free-Surface flow and Its Application to Filling Process in Casting

Keywords: Lattice Boltzmann models; free-surface phenomena; interface boundary conditions; filling processes; injection molding; volume of fluid method; interface boundary conditions; advection-schemes; up-wind-schemes
(54 pages, 2002)

35. M. Günther, A. Klar, T. Materne, R. Wegener

Multivalued fundamental diagrams and stop and go waves for continuum traffic equations

Keywords: traffic flow, macroscopic equations, kinetic derivation, multivalued fundamental diagram, stop and go waves, phase transitions
(25 pages, 2002)

36. S. Feldmann, P. Lang, D. Prätzel-Wolters

Parameter influence on the zeros of network determinants

Keywords: Networks, Equicofactor matrix polynomials, Realization theory, Matrix perturbation theory
(30 pages, 2002)

37. K. Koch, J. Ohser, K. Schladitz

Spectral theory for random closed sets and estimating the covariance via frequency space

Keywords: Random set, Bartlett spectrum, fast Fourier transform, power spectrum
(28 pages, 2002)

38. D. d'Humières, I. Ginzburg

Multi-reflection boundary conditions for lattice Boltzmann models

Keywords: lattice Boltzmann equation, boudary condistions, bounce-back rule, Navier-Stokes equation
(72 pages, 2002)

39. R. Korn

Elementare Finanzmathematik

Keywords: Finanzmathematik, Aktien, Optionen, Portfolio-Optimierung, Börse, Lehrerweiterbildung, Mathematikunterricht
(98 pages, 2002)

40. J. Kallrath, M. C. Müller, S. Nickel

Batch Presorting Problems: Models and Complexity Results

Keywords: Complexity theory, Integer programming, Assigment, Logistics
(19 pages, 2002)

41. J. Linn

On the frame-invariant description of the phase space of the Folgar-Tucker equation

Key words: fiber orientation, Folgar-Tucker equation, injection molding
(5 pages, 2003)

42. T. Hanne, S. Nickel

A Multi-Objective Evolutionary Algorithm for Scheduling and Inspection Planning in Software Development Projects

Key words: multiple objective programming, project management and scheduling, software development, evolutionary algorithms, efficient set
(29 pages, 2003)

43. T. Bortfeld , K.-H. Küfer, M. Monz, A. Scherrer, C. Thieke, H. Trinkaus

Intensity-Modulated Radiotherapy - A Large Scale Multi-Criteria Programming Problem

Keywords: multiple criteria optimization, representative systems of Pareto solutions, adaptive triangulation, clustering and disaggregation techniques, visualization of Pareto solutions, medical physics, external beam radiotherapy planning, intensity modulated radiotherapy
(31 pages, 2003)

44. T. Halfmann, T. Wichmann

Overview of Symbolic Methods in Industrial Analog Circuit Design

Keywords: CAD, automated analog circuit design, symbolic analysis, computer algebra, behavioral modeling, system simulation, circuit sizing, macro modeling, differential-algebraic equations, index
(17 pages, 2003)

45. S. E. Mikhailov, J. Orlik

Asymptotic Homogenisation in Strength and Fatigue Durability Analysis of Composites

Keywords: multiscale structures, asymptotic homogenization, strength, fatigue, singularity, non-local conditions
(14 pages, 2003)

46. P. Domínguez-Marín, P. Hansen, N. Mladenovi´c , S. Nickel

Heuristic Procedures for Solving the Discrete Ordered Median Problem

Keywords: genetic algorithms, variable neighborhood search, discrete facility location
(31 pages, 2003)

47. N. Boland, P. Domínguez-Marín, S. Nickel, J. Puerto

Exact Procedures for Solving the Discrete Ordered Median Problem

Keywords: discrete location, Integer programming
(41 pages, 2003)

48. S. Feldmann, P. Lang

Padé-like reduction of stable discrete linear systems preserving their stability

Keywords: Discrete linear systems, model reduction, stability, Hankel matrix, Stein equation
(16 pages, 2003)

49. J. Kallrath, S. Nickel

A Polynomial Case of the Batch Presorting Problem

Keywords: batch presorting problem, online optimization, competetive analysis, polynomial algorithms, logistics
(17 pages, 2003)

50. T. Hanne, H. L. Trinkaus

knowCube for MCDM – Visual and Interactive Support for Multicriteria Decision Making

Key words: Multicriteria decision making, knowledge management, decision support systems, visual interfaces, interactive navigation, real-life applications.
(26 pages, 2003)

51. O. Iliev, V. Laptev

On Numerical Simulation of Flow Through Oil Filters

Keywords: oil filters, coupled flow in plain and porous media, Navier-Stokes, Brinkman, numerical simulation
(8 pages, 2003)

52. W. Dörfler, O. Iliev, D. Stoyanov, D. Vassileva

On a Multigrid Adaptive Refinement Solver for Saturated Non-Newtonian Flow in Porous Media

Keywords: Nonlinear multigrid, adaptive refinement, non-Newtonian flow in porous media
(17 pages, 2003)

53. S. Kruse

On the Pricing of Forward Starting Options under Stochastic Volatility

Keywords: Option pricing, forward starting options, Heston model, stochastic volatility, cliquet options
(11 pages, 2003)

54. O. Iliev, D. Stoyanov

Multigrid – adaptive local refinement solver for incompressible flows

Keywords: Navier-Stokes equations, incompressible flow, projection-type splitting, SIMPLE, multigrid methods, adaptive local refinement, lid-driven flow in a cavity
(37 pages, 2003)

55. V. Starikovicius

The multiphase flow and heat transfer in porous media

Keywords: Two-phase flow in porous media, various formulations, global pressure, multiphase mixture model, numerical simulation
(30 pages, 2003)

56. P. Lang, A. Sarishvili, A. Wirsen

Blocked neural networks for knowledge extraction in the software development process

Keywords: Blocked Neural Networks, Nonlinear Regression, Knowledge Extraction, Code Inspection
(21 pages, 2003)

57. H. Knaf, P. Lang, S. Zeiser

Diagnosis aiding in Regulation Thermography using Fuzzy Logic

Keywords: fuzzy logic,knowledge representation, expert system
(22 pages, 2003)

58. M. T. Melo, S. Nickel, F. Saldanha da Gama

Largescale models for dynamic multicommodity capacitated facility location

Keywords: supply chain management, strategic planning, dynamic location, modeling
(40 pages, 2003)

59. J. Orlik

Homogenization for contact problems with periodically rough surfaces

Keywords: asymptotic homogenization, contact problems
(28 pages, 2004)

60. A. Scherrer, K.-H. Küfer, M. Monz, F. Alonso, T. Bortfeld

IMRT planning on adaptive volume structures – a significant advance of computational complexity

Keywords: Intensity-modulated radiation therapy (IMRT), inverse treatment planning, adaptive volume structures, hierarchical clustering, local refinement, adaptive clustering, convex programming, mesh generation, multi-grid methods
(24 pages, 2004)

61. D. Kehrwald

Parallel lattice Boltzmann simulation of complex flows

Keywords: Lattice Boltzmann methods, parallel computing, microstructure simulation, virtual material design, pseudo-plastic fluids, liquid composite moulding
(12 pages, 2004)

62. O. Iliev, J. Linn, M. Moog, D. Niedziela, V. Starikovicius

On the Performance of Certain Iterative Solvers for Coupled Systems Arising in Discretization of Non-Newtonian Flow Equations

Keywords: Performance of iterative solvers, Preconditioners, Non-Newtonian flow
(17 pages, 2004)

63. R. Ciegis, O. Iliev, S. Rief, K. Steiner

On Modelling and Simulation of Different Regimes for Liquid Polymer Moulding

Keywords: Liquid Polymer Moulding, Modelling, Simulation, Infiltration, Front Propagation, non-Newtonian flow in porous media
(43 pages, 2004)

64. T. Hanne, H. Neu

Simulating Human Resources in Software Development Processes

Keywords: Human resource modeling, software process, productivity, human factors, learning curve
(14 pages, 2004)

65. O. Iliev, A. Mikelic, P. Popov

Fluid structure interaction problems in deformable porous media: Toward permeability of deformable porous media

Keywords: fluid-structure interaction, deformable porous media, upscaling, linear elasticity, stokes, finite elements
(28 pages, 2004)

66. F. Gaspar, O. Iliev, F. Lisbona, A. Naumovich, P. Vabishchevich

On numerical solution of 1-D poroelasticity equations in a multilayered domain

Keywords: poroelasticity, multilayered material, finite volume discretization, MAC type grid
(41 pages, 2004)

67. J. Ohser, K. Schladitz, K. Koch, M. Nöthe

Diffraction by image processing and its application in materials science

Keywords: porous microstructure, image analysis, random set, fast Fourier transform, power spectrum, Bartlett spectrum
(13 pages, 2004)

68. H. Neunzert

Mathematics as a Technology: Challenges for the next 10 Years

Keywords: applied mathematics, technology, modelling, simulation, visualization, optimization, glass processing, spinning processes, fiber-fluid interaction, trubulence effects, topological optimization, multicriteria optimization, Uncertainty and Risk, financial mathematics, Malliavin calculus, Monte-Carlo methods, virtual material design, filtration, bio-informatics, system biology
(29 pages, 2004)

69. R. Ewing, O. Iliev, R. Lazarov, A. Naumovich

On convergence of certain finite difference discretizations for 1D poroelasticity interface problems

Keywords: poroelasticity, multilayered material, finite volume discretizations, MAC type grid, error estimates
(26 pages,2004)

70. W. Dörfler, O. Iliev, D. Stoyanov, D. Vassileva

On Efficient Simulation of Non-Newtonian Flow in Saturated Porous Media with a Multigrid Adaptive Refinement Solver

Keywords: Nonlinear multigrid, adaptive renement, non-Newtonian in porous media
(25 pages, 2004)

71. J. Kalcsics, S. Nickel, M. Schröder

Towards a Unified Territory Design Approach – Applications, Algorithms and GIS Integration

Keywords: territory desgin, political districting, sales territory alignment, optimization algorithms, Geographical Information Systems
(40 pages, 2005)

72. K. Schladitz, S. Peters, D. Reinel-Bitzer, A. Wiegmann, J. Ohser

Design of acoustic trim based on geometric modeling and flow simulation for non-woven

Keywords: random system of fibers, Poisson line process, flow resistivity, acoustic absorption, Lattice-Boltzmann method, non-woven
(21 pages, 2005)

73. V. Rutka, A. Wiegmann

Explicit Jump Immersed Interface Method for virtual material design of the effective elastic moduli of composite materials

Keywords: virtual material design, explicit jump immersed interface method, effective elastic moduli, composite materials
(22 pages, 2005)

74. T. Hanne

Eine Übersicht zum Scheduling von Baustellen

Keywords: Projektplanung, Scheduling, Bauplanung, Bauindustrie
(32 pages, 2005)

75. J. Linn

The Folgar-Tucker Model as a Differetial Algebraic System for Fiber Orientation Calculation

Keywords: fiber orientation, Folgar–Tucker model, invariants, algebraic constraints, phase space, trace stability
(15 pages, 2005)

76. M. Speckert, K. Dreßler, H. Mauch, A. Lion, G. J. Wierda

Simulation eines neuartigen Prüfsystems für Achserprobungen durch MKS-Modellierung einschließlich Regelung

Keywords: virtual test rig, suspension testing, multibody simulation, modeling hexapod test rig, optimization of test rig configuration
(20 pages, 2005)

77. K.-H. Küfer, M. Monz, A. Scherrer, P. Süss, F. Alonso, A. S. A. Sultan, Th. Bortfeld, D. Craft, Chr. Thieke

Multicriteria optimization in intensity modulated radiotherapy planning

Keywords: multicriteria optimization, extreme solutions, real-time decision making, adaptive approximation schemes, clustering methods, IMRT planning, reverse engineering
(51 pages, 2005)

78. S. Amstutz, H. Andrä

A new algorithm for topology optimization using a level-set method

Keywords: shape optimization, topology optimization, topological sensitivity, level-set
(22 pages, 2005)

79. N. Ettrich

Generation of surface elevation models for urban drainage simulation

Keywords: Flooding, simulation, urban elevation models, laser scanning
(22 pages, 2005)

80. H. Andrä, J. Linn, I. Matei, I. Shklyar, K. Steiner, E. Teichmann

OPTCAST – Entwicklung adäquater Strukturoptimierungsverfahren für Gießereien Technischer Bericht (KURZFASSUNG)

Keywords: Topologieoptimierung, Level-Set-Methode, Gießprozesssimulation, Gießtechnische Restriktionen, CAE-Kette zur Strukturoptimierung
(77 pages, 2005)

81. N. Marheineke, R. Wegener

Fiber Dynamics in Turbulent Flows
Part I: General Modeling Framework

Keywords: fiber-fluid interaction; Cosserat rod; turbulence modeling; Kolmogorov's energy spectrum; double-velocity correlations; differentiable Gaussian fields
(20 pages, 2005)
Part II: Specific Taylor Drag

Keywords: flexible fibers; k-ε turbulence model; fiber-turbulence interaction scales; air drag; random Gaussian aerodynamic force; white noise; stochastic differential equations; ARMA process
(18 pages, 2005)

82. C. H. Lampert, O. Wirjadi

An Optimal Non-Orthogonal Separation of the Anisotropic Gaussian Convolution Filter

Keywords: Anisotropic Gaussian filter, linear filtering, orientation space, nD image processing, separable filters
(25 pages, 2005)

83. H. Andrä, D. Stoyanov

Error indicators in the parallel finite element solver for linear elasticity DDFEM

Keywords: linear elasticity, finite element method, hierarchical shape functions, domain decom-position, parallel implementation, a posteriori error estimates
(21 pages, 2006)

84. M. Schröder, I. Solchenbach

Optimization of Transfer Quality in Regional Public Transit

Keywords: public transit, transfer quality, quadratic assignment problem
(16 pages, 2006)

85. A. Naumovich, F. J. Gaspar

On a multigrid solver for the three-dimensional Biot poroelasticity system in multilayered domains

Keywords: poroelasticity, interface problem, multigrid, operator-dependent prolongation
(11 pages, 2006)

86. S. Panda, R. Wegener, N. Marheineke

Slender Body Theory for the Dynamics of Curved Viscous Fibers

Keywords: curved viscous fibers; fluid dynamics; Navier-Stokes equations; free boundary value problem; asymptotic expansions; slender body theory
(14 pages, 2006)

87. E. Ivanov, H. Andrä, A. Kudryavtsev

Domain Decomposition Approach for Automatic Parallel Generation of Tetrahedral Grids

Key words: Grid Generation, Unstructured Grid, Delaunay Triangulation, Parallel Programming, Domain Decomposition, Load Balancing
(18 pages, 2006)

88. S. Tiwari, S. Antonov, D. Hietel, J. Kuhnert, R. Wegener

A Meshfree Method for Simulations of Inter-actions between Fluids and Flexible Structures

Key words: Meshfree Method, FPM, Fluid Structure Interaction, Sheet of Paper, Dynamical Coupling
(16 pages, 2006)

89. R. Ciegis, O. Iliev, V. Starikovicius, K. Steiner

Numerical Algorithms for Solving Problems of Multiphase Flows in Porous Media

Keywords: nonlinear algorithms, finite-volume method, software tools, porous media, flows
(16 pages, 2006)

90. D. Niedziela, O. Iliev, A. Latz

On 3D Numerical Simulations of Viscoelastic Fluids

Keywords: non-Newtonian fluids, anisotropic viscosity, integral constitutive equation
(18 pages, 2006)

91. A. Winterfeld

Application of general semi-infinite Pro-gramming to Lapidary Cutting Problems

Keywords: large scale optimization, nonlinear programming, general semi-infinite optimization, design centering, clustering
(26 pages, 2006)

92. J. Orlik, A. Ostrovska

Space-Time Finite Element Approximation and Numerical Solution of Hereditary Lin-ear Viscoelasticity Problems

Keywords: hereditary viscoelasticity; kern approximation by interpolation; space-time finite element approximation, stability and a priori estimate
(24 pages, 2006)

93. V. Rutka, A. Wiegmann, H. Andrä

EJIIM for Calculation of effective Elastic Moduli in 3D Linear Elasticity

Keywords: Elliptic PDE, linear elasticity, irregular domain, finite differences, fast solvers, effective elastic moduli
(24 pages, 2006)

94. A. Wiegmann, A. Zemitis

EJ-HEAT: A Fast Explicit Jump Harmonic Averaging Solver for the Effective Heat Conductivity of Composite Materials

Keywords: Stationary heat equation, effective thermal conductivity, explicit jump, discontinuous coefficients, virtual material design, microstructure simulation, EJ-HEAT
(21 pages, 2006)

95. A. Naumovich

On a finite volume discretization of the three-dimensional Biot poroelasticity sys-tem in multilayered domains

Keywords: Biot poroelasticity system, interface problems, finite volume discretization, finite difference method.
(21 pages, 2006)

96. M. Krekel, J. Wenzel

A unified approach to Credit Default Swaption and Constant Maturity Credit De-fault Swap valuation

Keywords: LIBOR market model, credit risk, Credit Default Swaption, Constant Maturity Credit Default Swapmethod.
(43 pages, 2006)

97. A. Dreyer

Interval Methods for Analog Circiuts

Keywords: interval arithmetic, analog circuits, tolerance analysis, parametric linear systems, frequency response, symbolic analysis, CAD, computer algebra
(36 pages, 2006)

98. N. Weigel, S. Weihe, G. Bitsch, K. Dreßler

Usage of Simulation for Design and Optimi-zation of Testing

Keywords: Vehicle test rigs, MBS, control, hydraulics, testing philosophy
(14 pages, 2006)

99. H. Lang, G. Bitsch, K. Dreßler, M. Speckert

Comparison of the solutions of the elastic and elastoplastic boundary value problems

Keywords: Elastic BVP, elastoplastic BVP, variational inequalities, rate-independency, hysteresis, linear kine-matic hardening, stop- and play-operator
(21 pages, 2006)

100. M. Speckert, K. Dreßler, H. Mauch

MBS Simulation of a hexapod based sus-pension test rig

Keywords: Test rig, MBS simulation, suspension, hydraulics, controlling, design optimization
(12 pages, 2006)

101. S. Azizi Sultan, K.-H. Küfer

A dynamic algorithm for beam orientations in multicriteria IMRT planning

Keywords: radiotherapy planning, beam orientation optimization, dynamic approach, evolutionary algo-rithm, global optimization
(14 pages, 2006)

102. T. Götz, A. Klar, N. Marheineke, R. Wegener

A Stochastic Model for the Fiber Lay-down Process in the Nonwoven Production

Keywords: fiber dynamics, stochastic Hamiltonian sys-tem, stochastic averaging
(17 pages, 2006)

103. Ph. Süss, K.-H. Küfer

Balancing control and simplicity: a variable aggregation method in intensity modulated radiation therapy planning

Keywords: IMRT planning, variable aggregation, clus-tering methods
(22 pages, 2006)

104. A. Beaudry, G. Laporte, T. Melo, S. Nickel

Dynamic transportation of patients in hos-pitals

Keywords: in-house hospital transportation, dial-a-ride, dynamic mode, tabu search
(37 pages, 2006)

105. Th. Hanne

Applying multiobjective evolutionary algo-rithms in industrial projects

Keywords: multiobjective evolutionary algorithms, dis-crete optimization, continuous optimization, electronic circuit design, semi-infinite programming, scheduling
(18 pages, 2006)

106. J. Franke, S. Halim

Wild bootstrap tests for comparing signals and images

Keywords: wild bootstrap test, texture classification, textile quality control, defect detection, kernel estimate, nonparametric regression
(13 pages, 2007)

Status quo: January 2007