Universität Stuttgart

Fachbereich Mathematik

Strongly consistent detection for nonparametric hypotheses

László Györfi, Harro Walk

Strongly consistent detection for nonparametric hypotheses

László Györfi and Harro Walk

Consider two robust detection problems formulated by nonparametric hypotheses such that both pairs hypotheses are composite and indistinguishable. Strongly consistent testing rules are shown.

1 Composite hypotheses defined by half spaces of distributions

Let ν_0, ν_1 be fixed distributions on $\mathbb{R}^d$ which are the nominal distributions under two hypotheses. Let

$$V(\nu, \mu) = \sup_{A \subset \mathbb{R}^d} |\nu(A) - \mu(A)|$$

denote the total variation distance between two distributions ν and μ, where the supremum is taken over all Borel sets of $\mathbb{R}^d$.

Let $X, X_1, X_2, \ldots$ be i.i.d. random vectors according to a common distribution μ. We observe $X_1, \ldots, X_n$. Under the hypothesis H_j ($j = 0, 1$) the distribution μ is a distorted version of ν_j. Formally define the two hypotheses by

$$H_0 = \{\mu : V(\mu, \nu_0) < V(\mu, \nu_1)\}, \tag{1}$$

and

$$H_1 = \{\mu : V(\mu, \nu_1) < V(\mu, \nu_0)\}. \tag{2}$$

László Györfi

Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Stoczek u.2, 1521 Budapest, Hungary e-mail: gyorfi@cs.bme.hu. The research reported here was supported in part by the National Development Agency (NFÜ, Hungary) as part of the project Introduction of Cognitive Methods for UAV Collision Avoidance Using Millimeter Wave Radar, (grant no.: KMR-12-1-2012-0008)

Harro Walk

Department of Mathematics, University of Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany e-mail: walk@mathematik.uni-stuttgart.de.

Our aim is to construct a distribution-free strong consistent test, which makes error only finitely many times almost surely (a.s.). The concept of strong consistent test is quite unusual, it means that both on H_0 and on H_1 the test makes a.s. no error after a random sample size. In other words, denoting by $\mathbb{P}_0$ and $\mathbb{P}_1$ the probability under the hypotheses H_0 and H_1, we have

$$\mathbb{P}_0\{\text{rejecting } H_0 \text{ for only finitely many } n\} = 1$$

and

$$\mathbb{P}_1\{\text{rejecting } H_1 \text{ for only finitely many } n\} = 1.$$

In a real life problem, for example, when we get the data sequentially, one gets data just once, and should make good inference for these data. Strong consistency means that the single sequence of inference is a.s. perfect if the sample size is large enough. This concept is close to the definition of discernability introduced by Dembo and Peres [4]. For a discussion and references, we refer the reader to Biau and Györfi [5], Devroye and Lugosi [8], Gretton and Györfi [10] and Györfi and Walk [15].

Motivated by a related goodness of fit test statistic of Györfi and van der Meulen [14], put

$$L_{n,0} = \sum_{j=1}^{m_n} |\mu_n(A_{n,j}) - \nu_0(A_{n,j})|,$$

and

$$L_{n,1} = \sum_{j=1}^{m_n} |\mu_n(A_{n,j}) - \nu_1(A_{n,j})|,$$

where μ_n denotes the empirical measures associated with the sample $X_1, \ldots, X_n$, so that

$$\mu_n(A) = \frac{\#\{i : X_i \in A, i = 1, \ldots, n\}}{n}$$

for any Borel subset A, and $\mathscr{P}_n = \{A_{n,1}, \ldots, A_{n,m_n}\}$ is a finite partition of $\mathbb{R}^d$.

Introduce the test such that accept the hypothesis H_0 if

$$L_{n,0} < L_{n,1}, \tag{3}$$

and reject otherwise.

The sequence of partitions $\mathscr{P}_1, \mathscr{P}_2, \ldots$ is called asymptotically fine if for any sphere S centered at the origin

$$\lim_{n \to \infty} \max_{A \in \mathscr{P}_n, A \cap S \neq \emptyset} diam(A) = 0. \tag{4}$$

Theorem 1. *Assume that the sequence of partitions $\mathscr{P}_1, \mathscr{P}_2, \ldots$ is asymptotically fine and*

$$\lim_{n \to \infty} \frac{m_n}{n} = 0. \tag{5}$$

Then the test (3) is strongly consistent.

Proof. Assume H_0 without loss of generality. Then the error event means that

$$L_{n,0} \geq L_{n,1}.$$

Thus,

$$0 \leq \sum_{j=1}^{m_n} |\mu_n(A_{n,j}) - \nu_0(A_{n,j})| - \sum_{j=1}^{m_n} |\mu_n(A_{n,j}) - \nu_1(A_{n,j})|$$

$$\leq 2L_n + \sum_{j=1}^{m_n} |\mu(A_{n,j}) - \nu_0(A_{n,j})| - \sum_{j=1}^{m_n} |\mu(A_{n,j}) - \nu_1(A_{n,j})|,$$

where

$$L_n = \sum_{j=1}^{m_n} |\mu_n(A_{n,j}) - \mu(A_{n,j})|.$$

Introduce the notation

$$\varepsilon = -(V(\mu, \nu_0) - V(\mu, \nu_1)) > 0.$$

The sequence of partitions $\mathscr{P}_1, \mathscr{P}_2, \ldots$ is asymptotically fine, which implies that

$$\lim_{n \to \infty} \left(\sum_{j=1}^{m_n} |\mu(A_{n,j}) - \nu_0(A_{n,j})| - \sum_{j=1}^{m_n} |\mu(A_{n,j}) - \nu_1(A_{n,j})| \right)$$
$$= 2(V(\mu, \nu_0) - V(\mu, \nu_1))$$
$$= -2\varepsilon,$$

(cf. Biau and Györfi [5]). Thus, for all n large enough,

$$P_{e,n} = \mathbb{P}\{\text{error}\} \leq \mathbb{P}\{0 \leq 2L_n - \varepsilon\}.$$

Beirlant, Devroye, Györfi and Vajda [2] and Biau and Györfi [5] proved that, for any $\delta > 0$,

$$\mathbb{P}\{L_n > \delta\} \leq 2^{m_n} e^{-n\delta^2/2}.$$

Therefore

$$P_{e,n} \leq 2^{m_n} e^{-n\varepsilon^2/8}.$$

Because of (5),

$$\sum_{n=1}^{\infty} P_{e,n} < \infty,$$

and so the Borel-Cantelli lemma implies that a.s.

$$L_{n,0} - L_{n,1} < 0$$

for all n large enough, i.e., the error

$$L_{n,0} - L_{n,1} \geq 0$$

occurs a.s. for only finitely many n. Thus, the strong consistency is proved. □

In a straightforward way, the proof of Theorem 1 can be extended to infinite partitions if we assume that for each sphere S centered at the origin

$$\lim_{n\to\infty} \frac{|\{j : A_{n,j} \cap S \neq \emptyset\}|}{n} = 0.$$

Next, a variant of the test (3) with much smaller computational complexity will be defined. The test statistic is based on recursive histogram. In this section assume that the partitions are infinite and all cells of the partitions have finite and positive Lebesgue measure λ. Let $A_n(x)$ denote the cell of $\mathscr{P}_n$ into which x belongs. The density estimate

$$f_n(x) := \frac{1}{n} \sum_{i=1}^{n} \frac{\mathbb{I}_{\{X_i \in A_i(x)\}}}{\lambda(A_i(x))}$$

is called recursive histogram.

For $A \in \mathscr{P}_n$, introduce the estimate

$$\mu_n^*(A) := \int_A f_n(x)dx.$$

Notice that $\mu_n^*(A)$ can be calculated in a recursive way, which is important in on-line applications. These definitions imply that

$$\mu_n^*(A) = \frac{1}{n} \sum_{i=1}^{n} \int_A \frac{\mathbb{I}_{\{X_i \in A_i(x)\}}}{\lambda(A_i(x))}dx = \frac{1}{n} \sum_{i=1}^{n} \int_A \frac{\mathbb{I}_{\{x \in A_i(X_i)\}}}{\lambda(A_i(X_i))}dx = \frac{1}{n} \sum_{i=1}^{n} \frac{\lambda(A \cap A_i(X_i))}{\lambda(A_i(X_i))}.$$

If the sequence of partitions $\mathscr{P}_1, \mathscr{P}_2, \ldots$ is nested, i.e., the sequence of σ-algebras $\sigma(\mathscr{P}_n)$ is non-decreasing, then for $A \in \mathscr{P}_n$ let the ancestor $B_A^{(i)} \in \mathscr{P}_i$ be such that $A \subset B_A^{(i)}$ ($i \leq n$). One can check that for nested partitions the estimate has the following form:

$$\mu_n^*(A) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{X_i \in B_A^{(i)}\}} \frac{\lambda(A)}{\lambda(B_A^{(i)})}.$$

Put

$$L_{n,j}^* := \sum_{A \in \mathscr{P}_n} |\mu_n^*(A) - \nu_j(A)|$$

($j = 0, 1$). Introduce the test such that accept the hypothesis H_0 if

$$L_{n,0}^* < L_{n,1}^*, \tag{6}$$

and reject otherwise.

Theorem 2. *Assume that the sequence of partitions $\mathscr{P}_1, \mathscr{P}_2, \ldots$ is asymptotically fine such that*

$$\sum_{n=1}^{\infty} \frac{1}{n^2 \inf_j \lambda(A_{n,j})} < \infty. \tag{7}$$

Further suppose that μ has a density. Then the test (6) is strongly consistent.

Proof. Assume H_0 without loss of generality. One notices

$$L_{n,0}^* - L_{n,1}^* \leq 2L_n^* + Q_n^*,$$

where

$$L_n^* = \sum_{A \in \mathscr{P}_n} |\mu_n^*(A) - \mu(A)|,$$

and

$$Q_n^* = \sum_{A \in \mathscr{P}_n} |\mu(A) - \nu_0(A)| - \sum_{A \in \mathscr{P}_n} |\mu(A) - \nu_1(A)|.$$

By Biau and Györfi [5],

$$Q_n^* \to 2(V(\mu, \nu_0) - V(\mu, \nu_1)) < 0,$$

the latter because of H_0. Next $L_n^* \to 0$ a.s. will be shown. Denote the density of μ by f. Thus

$$L_n^* = \sum_{A \in \mathscr{P}_n} \left| \int_A f_n(x)dx - \int_A f(x)dx \right| \leq \int |f_n(x) - f(x)|dx.$$

Therefore we have to prove the strong L_1-consistency of the recursive histogram. Consider the bias part. Introduce the ordinary histogram:

$$\tilde{f}_n(x) := \frac{1}{n} \sum_{i=1}^{n} \frac{\mathbb{I}_{\{X_i \in A_n(x)\}}}{\lambda(A_n(x))},$$

and put

$$\bar{f}_n(x) := \mathbb{E}\{\tilde{f}_n(x)\} = \frac{\mu(A_n(x))}{\lambda(A_n(x))}.$$

According to the Abou-Jaoude theorem, if the sequence of partitions $\mathscr{P}_1, \mathscr{P}_2, \ldots$ is asymptotically fine, then

$$\int |\bar{f}_n - f| \to 0$$

(cf. Devroye and Györfi [6]). Thus, for the bias term of the recursive histogram, we get

$$\int |\mathbb{E}\{f_n\} - f| = \int \left| \frac{1}{n} \sum_{i=1}^{n} \bar{f}_i - f \right| \leq \frac{1}{n} \sum_{i=1}^{n} \int |\bar{f}_i - f| \to 0. \tag{8}$$

For the variation term of the recursive histogram, we apply the generalized theorem of Kolmogorov: Let U_n, $n = 1, 2, \ldots$ be an L_2-valued sequence of independent, zero mean random variables such that

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\{\|U_n\|_2^2\}}{n^2} < \infty$$

where $\|\cdot\|_2$ denotes the L_2 norm. Then

$$\lim_{n \to \infty} \left\| \frac{1}{n} \sum_{i=1}^{n} U_i \right\|_2 = 0$$

a.s. (cf. Győrfi et al. [11]). For

$$U_n := \frac{\mathbb{I}_{\{X_n \in A_n(\cdot)\}}}{\lambda(A_n(\cdot))} - \mathbb{E}\left\{ \frac{\mathbb{I}_{\{X_n \in A_n(\cdot)\}}}{\lambda(A_n(\cdot))} \right\},$$

one has to verify the condition of the generalized Kolmogorov theorem:

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\left\{ \left\| \frac{\mathbb{I}_{\{X_n \in A_n(\cdot)\}}}{\lambda(A_n(\cdot))} - \mathbb{E}\left\{ \frac{\mathbb{I}_{\{X_n \in A_n(\cdot)\}}}{\lambda(A_n(\cdot))} \right\} \right\|_2^2 \right\}}{n^2} \leq \sum_{n=1}^{\infty} \frac{\mathbb{E}\left\{ \left\| \frac{\mathbb{I}_{\{X_n \in A_n(\cdot)\}}}{\lambda(A_n(\cdot))} \right\|_2^2 \right\}}{n^2}$$

$$= \sum_{n=1}^{\infty} \frac{\mathbb{E}\left\{ \int \frac{\mathbb{I}_{\{X_n \in A_n(x)\}}}{\lambda(A_n(x))^2} dx \right\}}{n^2}$$

$$= \sum_{n=1}^{\infty} \frac{\mathbb{E}\left\{ \int \frac{\mathbb{I}_{\{x \in A_n(X_n)\}}}{\lambda(A_n(X_n))^2} dx \right\}}{n^2}$$

$$= \sum_{n=1}^{\infty} \frac{\mathbb{E}\left\{ \frac{1}{\lambda(A_n(X_n))} \right\}}{n^2}$$

$$\leq \sum_{n=1}^{\infty} \frac{1}{n^2 \inf_j \lambda(A_{n,j})} < \infty,$$

by the condition of the theorem, and so

$$\int |f_n - \mathbb{E}\{f_n\}|^2 \to 0. \tag{9}$$

a.s. From Lemma 3.1 in Győrfi and Masry [13] we get that the limit relations (8) and (9) imply

$$\int |f_n - f| \to 0$$

a.s. Therefore a.s.

$$L_{n,0}^* - L_{n,1}^* < 0$$

for all n large enough, and so the the strong consistency is proved. $\square$

2 Composite hypotheses defined by half spheres of distributions

Again, under the hypothesis H'_j ($j = 0, 1$) the distribution μ is a distorted version of ν_j. In this section we assume that the true distribution lies within a certain total variation distance of the underlying nominal distribution.

We formally define the two hypotheses by

$$H'_j = \left\{ \mu : V(\mu, \nu_j) < \Delta \right\}, \quad j = 0, 1 , \tag{10}$$

where

$$\Delta := (1/2)V(\nu_0, \nu_1).$$

Because of

$$H'_j \subset H_j, \quad j = 0, 1 ,$$

the test (3) in the previous section is strongly consistent. In this section we introduce a simpler test. For the notations

$$f = \frac{d\nu}{d(\nu + \mu)} \quad \text{and} \quad g = \frac{d\mu}{d(\nu + \mu)},$$

the general version of Scheffé's theorem implies that

$$V(\nu, \mu) = \nu(A^*) - \mu(A^*),$$

where

$$A^* = \{ x : f(x) > g(x) \} .$$

Introduce the notation

$$A_{0,1} = \{ x : f_0(x) > f_1(x) \} = \{ x : f_0(x) > 1/2 \} ,$$

where

$$f_0 = \frac{d\nu_0}{d(\nu_0 + \nu_1)} \quad \text{and} \quad f_1 = \frac{d\nu_1}{d(\nu_0 + \nu_1)}.$$

The proposed test is the following: accept hypothesis H'_0 if

$$\mu_n(A_{0,1}) \geq \frac{\nu_0(A_{0,1}) + \nu_1(A_{0,1})}{2}, \tag{11}$$

and reject otherwise.

Then, we get that

Theorem 3. *The test (11) is strongly consistent.*

Proof. Assume H_0 without loss of generality. Put

$$\varepsilon = \Delta - V(\mu, \nu_0) > 0.$$

Observe that by the Scheffé theorem [22],

$$
\begin{aligned}
\nu_0(A_{0,1}) - \mu(A_{0,1}) &\leq V(\nu_0, \mu) \\
&= \Delta - \varepsilon \\
&= \frac{1}{2} V(\nu_0, \nu_1) - \varepsilon \\
&= \frac{1}{2} \left(\nu_0(A_{0,1}) - \nu_1(A_{0,1}) \right) - \varepsilon.
\end{aligned}
$$

Rearranging the obtained inequality, we get that

$$\mu(A_{0,1}) \geq \frac{\nu_0(A_{0,1}) + \nu_1(A_{0,1})}{2} + \varepsilon. \tag{12}$$

Therefore, (12) and Hoeffding's inequality [16] imply that

$$
\begin{aligned}
\mathbb{P}\{\text{error}\} = \mathbb{P}\left\{ \mu_n(A_{0,1}) < \frac{\nu_0(A_{0,1}) + \nu_1(A_{0,1})}{2} \right\} \\
\leq \mathbb{P}\left\{ \mu(A_{0,1}) - \mu_n(A_{0,1}) > \varepsilon \right\} \\
\leq e^{-2n\varepsilon^2}.
\end{aligned}
$$

Therefore the Borel-Cantelli lemma implies the strong consistency. □

3 Discussion

INDISTINGUISHABILITY.
For the hypotheses H_0 and H_1 there is no positive margin, because the gap between H_0 and H_1 is just the hyperplane

$$\{\mu : V(\mu, \nu_0) = V(\mu, \nu_1)\}.$$

Moreover, the margin is zero:

$$\inf_{\mu \in H_0, \nu \in H_1} V(\mu, \nu) = 0.$$

Without any positive margin condition it is impossible the derive uniform bound on the error probabilities. The pair (H_0, H_1) of hypotheses is called distinguishable if there is a sequence of uniformly consistent tests, which means that the errors of the first and second kind tend to zero uniformly. For a test T_n with sample size n, let $\alpha_{n,\mu}(T_n)$ and $\beta_{n,\mu}(T_n)$ denote the errors of the first and second kind, resp. Put

$$\alpha_n(T_n, H_0) = \sup_{\mu \in H_0} \alpha_{n,\mu}(T_n) \qquad \beta_n(T_n, H_1) = \sup_{\mu \in H_1} \beta_{n,\mu}(T_n).$$

A sequence of tests T_n, $n = 1, 2, \ldots$ is called uniformly consistent if

$$\lim_{n \to \infty} (\alpha_n(T_n, H_0) + \beta_n(T_n, H_1)) = 0.$$

It is known that a necessary condition of the distinguishable property is that for any distribution μ

$$\max \left\{ \inf_{\nu \in H_0} V(\mu, \nu), \inf_{\nu \in H_1} V(\mu, \nu) \right\} > 0.$$

(See Barron [1], Ermakov [9], Hoeffding and Wolfowitz [17], Kraft [18], LeCam [19], LeCam and Schwartz [20], Schwartz [23].) Obviously, this necessary condition is not satisfied when $\mu^* = (\nu_1 + \nu_2)/2$. Because of

$$\max \left\{ \inf_{\nu \in H_0'} V(\mu^*, \nu), \inf_{\nu \in H_1'} V(\mu^*, \nu) \right\} = 0,$$

the pair (H_0', H_1') of hypotheses is indistinguishable, too.

COMPUTATION.
The hypothesis testing method (11) proposed above is computationally quite simple. The set $A_{0,1}$ and the nominal probabilities $\nu_0(A_{0,1})$ and $\nu_1(A_{0,1})$ may be computed and stored before seeing the data. Then one merely needs to calculate $\mu_n(A_{0,1})$.

HYPOTHESES FORMULATED BY DENSITIES.
Devroye, Györfi, Lugosi [7] formulated a special case of hypotheses (H_0', H_1'), when μ, ν_0 and ν_1 have densities f, f_0 and f_1. Under some mild margin condition they proved uniform exponential bounds for the probability of failure for $k \geq 2$ hypotheses. Moreover, they illustrated the robustness for additive noise scheme, and showed examples, where the test (11) is consistent and the maximum likelihood test does not work. Formally, the maximum likelihood test T_n is defined by

$$T_n = \begin{cases} 0 & \text{if } \sum_{i=1}^n (\log f_0(X_i) - \log f_1(X_i)) > 0 \\ 1 & \text{otherwise.} \end{cases}$$

For $f \in H_0'$, the strong law of large numbers implies the strong consistency of the maximum likelihood test if both integrals $\int f \log f_0$ and $\int f \log f_1$ are well defined, and

$$\int f \log f_0 > \int f \log f_1.$$

ROBUSTNESS.
Note that Theorem 3 does not require any assumption for the nominal distributions. The test is robust in a very strong sense: we obtain consistency under the sole assumption that the distorted distribution remains within a certain total variation distance of the nominal distribution. For example, if μ is either $(1 - \delta)\nu_0 + \delta \tau$, or

$(1-\delta)\nu_1 + \delta\tau$ with arbitrary "strange" distribution τ such that $\delta < \Delta$, then we have (10):

$$\begin{aligned} V(\mu, \nu_0) &= V((1-\delta)\nu_0 + \delta\tau, \nu_0) \\ &= V(\delta\tau, \delta\nu_0) \\ &\leq \delta \\ &< \Delta. \end{aligned}$$

τ can be the outliers' distribution, it is really arbitrary, for example, not having expectations, or can be even a discrete distribution. The probability of outlier δ can be at most equal to Δ. The outliers can be formulated such that we are given three, independent i.i.d. sequences $\{U_i\}$, $\{V_i\}$, $\{I_i\}$, where $\{U_i\}$, $\{V_i\}$ are $\mathbb{R}^d$-valued, and $\{I_i\}$ are binary. Put

$$X_n = (1-I_n)U_n + I_n V_n.$$

If U_n is ν_0 distributed, V_n is τ distributed, $\mathbb{P}\{I_n = 1\} = \delta$, then we get the previous scheme. Other application can be the case of censored observations, when V_n is a distortion of U_n such that some components of the vector U_n are censored. In this scheme δ is the probability of censoring. Notice that in order to estimate the distribution from censored observations one needs samples $\{(X_i, I_i)\}_{i=1}^n$ (cf. Györfi et al. [12]), while for detection it is enough to have $\{X_i\}_{i=1}^n$.

OPEN PROBLEMS.
1. Characterize the distributions $\mu \in H_0 \setminus H_0'$, where the simple test (11) is strongly consistent. As in the proof of Theorem 3 the strong consistency can be verified if

$$\mu(A_{0,1}) > \frac{\nu_0(A_{0,1}) + \nu_1(A_{0,1})}{2}.$$

We are interested in non-consistent examples, too.
2. Maybe one can improve the test (3), since in the construction of the partitions we don't take into account the properties of ν_0 and ν_1. For example, we can include somehow the set $A_{0,1}$.

SEQUENTIAL TESTS.
We dealt with sequences of nonparametric tests with increasing sample size n, where almost surely type I and II errors occur only for finitely many n. One has to distinguish them from nonparametric sequential tests with power one (cf. Darling and Robbins [3], Section 6 in Robbins [21], Section 9.2 in Sen [24]). Such tests almost surely terminate at a random sample size with rejection of a null hypothesis H_0 after finitely many observations, if the alternative hypothesis is valid, and with positive probability do not terminate if H_0 is valid (open-ended procedures). In the latter case an upper bound of the complementary probabilities is an upper bound for the type I error probability.

References

1. Barron A.R.: Uniformly powerful goodness of fit tests. *Annals of Statistics*, **17**, 107–124 (1989)
2. Beirlant, J., Devroye, L., Györfi, L. and Vajda, I.: Large deviations of divergence measures on partitions, *J. Statistical Planning and Inference*, **93**, 1–16 (2001)
3. Darling, D. A. and Robbins, H.: Some nonparametric sequential tests with power one, *Proc. National Academy of Sciences U.S.A.*, **61**, 804–809 (1968)
4. Dembo, A. and Peres, Y.: A topological criterion for hypothesis testing, *Annals of Statistics*, **22**, 106–117 (1994)
5. Biau, G. and Györfi, L.: On the asymptotic properties of a nonparametric l_1-test statistic of homogeneity. *IEEE Transactions on Information Theory*, **51**, 3965–3973 (2005)
6. Devroye, L. and Györfi, L.: *Nonparametric Density Estimation: the L_1 View*. Wiley, New York (1985)
7. Devroye, L., Györfi, L. and Lugosi, G.: A note on robust hypothesis testing, *IEEE Trans. Information Theory*, **48**, 2111–2114 (2002)
8. Devroye, L. and Lugosi, G.: Almost sure classification of densities, *J. Nonparametric Statistics*, **14**, 675–698 (2002)
9. Ermakov, M.: On distinguishability of hypotheses. http://arxiv.org/pdf/1308.4295.pdf, (2013)
10. Gretton, A. and Györfi, L.: Consistent nonparametric tests of independence, *Journal of Machine Learning Research*, **11**, 1391–1423 (2010)
11. Györfi, L., Györfi, Z. and Vajda, I.: A strong law of large numbers and some applications. *Studia Scientiarum Mathematicarum Hungarica*, **12**, 233–244 (1977)
12. Györfi, L., Kohler, M., Krzyżak, A. and Walk, H.: *A Distribution-Free Theory of Nonparametric Regression*. Springer, New York (2002)
13. Györfi, L. and Masry, E.: The L_1 and L_2 strong consistency of recursive kernel density estimation from dependent samples. *IEEE Trans. Information Theory*, **36**, 531–539 (1990)
14. Györfi, L. and van der Meulen, E. C.: A consistent goodness of fit test based on the total variation distance, in *Nonparametric Functional Estimation and Related Topics*, (G. Roussas, Ed.), Kluwer, Dordrecht, pp. 631–645 (1990)
15. Györfi, L. and Walk, H.: Strongly consistent nonparametric tests of conditional independence, *Statistics and Probability Letters*, **82**, 1145–1150 (2012)
16. Hoeffding, W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, **58**, 13–30 (1963)
17. Hoeffding, W. and Wolfowitz, J.: Distinguishability of sets of distributions. *The Annals of Mathematical Statistics*, **29**, 700–718 (1958)
18. Kraft, C.: Some conditions for consistency and uniform consistency. *University of California Publications in Statistics*, **2**, 125–142 (1955)
19. Le Cam, L.: Convergence of estimates under dimensionality restrictions. *Annals of Statistics*, **1**, 38–53 (1973).
20. Le Cam, L. and Schwartz, L.: A necessary and sufficient condition for the existence of consistent estimates. *Annals of Mathematical Statistics*, **31**, 40–150 (1960)
21. Robbins, H.: Statistical methods related to the law of the iterated logarithm. *The Annals of Mathematical Statistics*, **41**, 1397–1409 (1970)
22. Scheffé, H.: A useful convergence theorem for probability distributions. *The Annals of Mathematical Statistics*, **18**, 434–438 (1947)
23. Schwartz, L.: On Bayes procedures. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **4**, 10–26 (1965)
24. Sen, P. K.: *Sequential Nonparametrics: Invariance Principles and Statistical Inference*. Wiley, New York (1981)

László Györfi
Department of Computer Science and Information Theory
Budapest University of Technology and Economics
Stoczek u.2
1521 Budapest
Hungary
E-Mail: `gyorfi@cs.bme.hu`

Harro Walk
Universität Stuttgart
Fachbereich Mathematik
Pfaffenwaldring 57
70569 Stuttgart
Germany
E-Mail: `walk@mathematik.uni-stuttgart.de`

Erschienene Preprints ab Nummer 2007/2007-001

2014-014 *Kohler, M.; Krzyżak, A.; Walk, H.:* Nonparametric recursive quantile estimation

2014-013 *Kohler, M.; Krzyżak, A.; Tent, R.; Walk, H.:* Nonparametric quantile estimation using importance sampling

2014-012 *Györfi, L.; Ottucsák, G.; Walk, H.:* The growth optimal investment strategy is secure, too.

2014-011 *Györfi, L.; Walk, H.:* Strongly consistent detection for nonparametric hypotheses

2014-010 *Köster, I.:* Finite Groups with Sylow numbers $\{q^x, a, b\}$

2014-009 *Kahnert, D.:* Hausdorff Dimension of Rings

2014-008 *Steinwart, I.:* Measuring the Capacity of Sets of Functions in the Analysis of ERM

2014-007 *Steinwart, I.:* Convergence Types and Rates in Generic Karhunen-Loève Expansions with Applications to Sample Path Properties

2014-006 *Steinwart, I.; Pasin, C.; Williamson, R.; Zhang, S.:* Elicitation and Identification of Properties

2014-005 *Schmid, J.; Griesemer, M.:* Integration of Non-Autonomous Linear Evolution Equations

2014-004 *Markhasin, L.:* L_2- and $S_{p,q}^r B$-discrepancy of (order 2) digital nets

2014-003 *Markhasin, L.:* Discrepancy and integration in function spaces with dominating mixed smoothness

2014-002 *Eberts, M.; Steinwart, I.:* Optimal Learning Rates for Localized SVMs

2014-001 *Giesselmann, J.:* A relative entropy approach to convergence of a low order approximation to a nonlinear elasticity model with viscosity and capillarity

2013-016 *Steinwart, I.:* Fully Adaptive Density-Based Clustering

2013-015 *Steinwart, I.:* Some Remarks on the Statistical Analysis of SVMs and Related Methods

2013-014 *Rohde, C.; Zeiler, C.:* A Relaxation Riemann Solver for Compressible Two-Phase Flow with Phase Transition and Surface Tension

2013-013 *Moroianu, A.; Semmelmann, U.:* Generalized Killling spinors on Einstein manifolds

2013-012 *Moroianu, A.; Semmelmann, U.:* Generalized Killing Spinors on Spheres

2013-011 *Kohls, K; Rösch, A.; Siebert, K.G.:* Convergence of Adaptive Finite Elements for Control Constrained Optimal Control Problems

2013-010 *Corli, A.; Rohde, C.; Schleper, V.:* Parabolic Approximations of Diffusive-Dispersive Equations

2013-009 *Nava-Yazdani, E.; Polthier, K.:* De Casteljau's Algorithm on Manifolds

2013-008 *Bächle, A.; Margolis, L.:* Rational conjugacy of torsion units in integral group rings of non-solvable groups

2013-007 *Knarr, N.; Stroppel, M.J.:* Heisenberg groups over composition algebras

2013-006 *Knarr, N.; Stroppel, M.J.:* Heisenberg groups, semifields, and translation planes

2013-005 *Eck, C.; Kutter, M.; Sändig, A.-M.; Rohde, C.:* A Two Scale Model for Liquid Phase Epitaxy with Elasticity: An Iterative Procedure

2013-004 *Griesemer, M.; Wellig, D.:* The Strong-Coupling Polaron in Electromagnetic Fields

2013-003 *Kabil, B.; Rohde, C.:* The Influence of Surface Tension and Configurational Forces on the Stability of Liquid-Vapor Interfaces

2013-002 *Devroye, L.; Ferrario, P.G.; Györfi, L.; Walk, H.:* Strong universal consistent estimate of the minimum mean squared error

2013-001 *Kohls, K.; Rösch, A.; Siebert, K.G.:* A Posteriori Error Analysis of Optimal Control Problems with Control Constraints

2012-018 *Kimmerle, W.; Konovalov, A.:* On the Prime Graph of the Unit Group of Integral Group Rings of Finite Groups II

2012-017 *Stroppel, B.; Stroppel, M.:* Desargues, Doily, Dualities, and Exceptional Isomorphisms

2012-016 *Moroianu, A.; Pilca, M.; Semmelmann, U.:* Homogeneous almost quaternion-Hermitian manifolds

2012-015 *Steinke, G.F.; Stroppel, M.J.:* Simple groups acting two-transitively on the set of generators of a finite elation Laguerre plane

2012-014 *Steinke, G.F.; Stroppel, M.J.:* Finite elation Laguerre planes admitting a two-transitive group on their set of generators

2012-013 *Diaz Ramos, J.C.; Dominguez Vázquez, M.; Kollross, A.:* Polar actions on complex hyperbolic spaces

2012-012 *Moroianu; A.; Semmelmann, U.:* Weakly complex homogeneous spaces

2012-011 *Moroianu; A.; Semmelmann, U.:* Invariant four-forms and symmetric pairs

2012-010 *Hamilton, M.J.D.:* The closure of the symplectic cone of elliptic surfaces

2012-009 *Hamilton, M.J.D.:* Iterated fibre sums of algebraic Lefschetz fibrations

2012-008 *Hamilton, M.J.D.:* The minimal genus problem for elliptic surfaces

2012-007 *Ferrario, P.:* Partitioning estimation of local variance based on nearest neighbors under censoring

2012-006 *Stroppel, M.:* Buttons, Holes and Loops of String: Lacing the Doily

2012-005 *Hantsch, F.:* Existence of Minimizers in Restricted Hartree-Fock Theory

2012-004 *Grundhöfer, T.; Stroppel, M.; Van Maldeghem, H.:* Unitals admitting all translations

2012-003 *Hamilton, M.J.D.:* Representing homology classes by symplectic surfaces

2012-002 *Hamilton, M.J.D.:* On certain exotic 4-manifolds of Akhmedov and Park

2012-001 *Jentsch, T.:* Parallel submanifolds of the real 2-Grassmannian

2011-028 *Spreer, J.:* Combinatorial 3-manifolds with cyclic automorphism group

2011-027 *Griesemer, M.; Hantsch, F.; Wellig, D.:* On the Magnetic Pekar Functional and the Existence of Bipolarons

2011-026 *Müller, S.:* Bootstrapping for Bandwidth Selection in Functional Data Regression

2011-025 *Felber, T.; Jones, D.; Kohler, M.; Walk, H.:* Weakly universally consistent static forecasting of stationary and ergodic time series via local averaging and least squares estimates

2011-024 *Jones, D.; Kohler, M.; Walk, H.:* Weakly universally consistent forecasting of stationary and ergodic time series

2011-023 *Györfi, L.; Walk, H.:* Strongly consistent nonparametric tests of conditional independence

2011-022 *Ferrario, P.G.; Walk, H.:* Nonparametric partitioning estimation of residual and local variance based on first and second nearest neighbors

2011-021 *Eberts, M.; Steinwart, I.:* Optimal regression rates for SVMs using Gaussian kernels

2011-020 *Frank, R.L.; Geisinger, L.:* Refined Semiclassical Asymptotics for Fractional Powers of the Laplace Operator

2011-019 *Frank, R.L.; Geisinger, L.:* Two-term spectral asymptotics for the Dirichlet Laplacian on a bounded domain

2011-018 *Hänel, A.; Schulz, C.; Wirth, J.:* Embedded eigenvalues for the elastic strip with cracks

2011-017 *Wirth, J.:* Thermo-elasticity for anisotropic media in higher dimensions

2011-016 *Höllig, K.; Hörner, J.:* Programming Multigrid Methods with B-Splines

2011-015 *Ferrario, P.:* Nonparametric Local Averaging Estimation of the Local Variance Function

2011-014 *Müller, S.; Dippon, J.:* k-NN Kernel Estimate for Nonparametric Functional Regression in Time Series Analysis

2011-013 *Knarr, N.; Stroppel, M.:* Unitals over composition algebras

2011-012 *Knarr, N.; Stroppel, M.:* Baer involutions and polarities in Moufang planes of characteristic two

2011-011 *Knarr, N.; Stroppel, M.:* Polarities and planar collineations of Moufang planes

2011-010 *Jentsch, T.; Moroianu, A.; Semmelmann, U.:* Extrinsic hyperspheres in manifolds with special holonomy

2011-009 *Wirth, J.:* Asymptotic Behaviour of Solutions to Hyperbolic Partial Differential Equations

2011-008 *Stroppel, M.:* Orthogonal polar spaces and unitals

2011-007 *Nagl, M.:* Charakterisierung der Symmetrischen Gruppen durch ihre komplexe Gruppenalgebra

2011-006 *Solanes, G.; Teufel, E.:* Horo-tightness and total (absolute) curvatures in hyperbolic spaces

2011-005 *Ginoux, N.; Semmelmann, U.:* Imaginary Kählerian Killing spinors I

2011-004 *Scherer, C.W.; Köse, I.E.:* Control Synthesis using Dynamic D-Scales: Part II — Gain-Scheduled Control

2011-003 *Scherer, C.W.; Köse, I.E.:* Control Synthesis using Dynamic D-Scales: Part I — Robust Control

2011-002 *Alexandrov, B.; Semmelmann, U.:* Deformations of nearly parallel G_2-structures

2011-001 *Geisinger, L.; Weidl, T.:* Sharp spectral estimates in domains of infinite volume

2010-018 *Kimmerle, W.; Konovalov, A.:* On integral-like units of modular group rings

2010-017 *Gauduchon, P.; Moroianu, A.; Semmelmann, U.:* Almost complex structures on quaternion-Kähler manifolds and inner symmetric spaces

2010-016 *Moroianu, A.; Semmelmann,U.:* Clifford structures on Riemannian manifolds

2010-015 *Grafarend, E.W.; Kühnel, W.:* A minimal atlas for the rotation group $SO(3)$

2010-014 *Weidl, T.:* Semiclassical Spectral Bounds and Beyond

2010-013 *Stroppel, M.:* Early explicit examples of non-desarguesian plane geometries

2010-012 *Effenberger, F.:* Stacked polytopes and tight triangulations of manifolds

2010-011 *Györfi, L.; Walk, H.:* Empirical portfolio selection strategies with proportional transaction costs

2010-010 *Kohler, M.; Krzyżak, A.; Walk, H.:* Estimation of the essential supremum of a regression function

2010-009 *Geisinger, L.; Laptev, A.; Weidl, T.:* Geometrical Versions of improved Berezin-Li-Yau Inequalities

2010-008 *Poppitz, S.; Stroppel, M.:*   Polarities of Schellhammer Planes

2010-007 *Grundhöfer, T.; Krinn, B.; Stroppel, M.:*   Non-existence of isomorphisms between certain unitals

2010-006 *Höllig, K.; Hörner, J.; Hoffacker, A.:*   Finite Element Analysis with B-Splines: Weighted and Isogeometric Methods

2010-005 *Kaltenbacher, B.; Walk, H.:*   On convergence of local averaging regression function estimates for the regularization of inverse problems

2010-004 *Kühnel, W.; Solanes, G.:*   Tight surfaces with boundary

2010-003 *Kohler, M; Walk, H.:*   On optimal exercising of American options in discrete time for stationary and ergodic data

2010-002 *Gulde, M.; Stroppel, M.:*   Stabilizers of Subspaces under Similitudes of the Klein Quadric, and Automorphisms of Heisenberg Algebras

2010-001 *Leitner, F.:*   Examples of almost Einstein structures on products and in cohomogeneity one

2009-008 *Griesemer, M.; Zenk, H.:*   On the atomic photoeffect in non-relativistic QED

2009-007 *Griesemer, M.; Moeller, J.S.:*   Bounds on the minimal energy of translation invariant n-polaron systems

2009-006 *Demirel, S.; Harrell II, E.M.:*   On semiclassical and universal inequalities for eigenvalues of quantum graphs

2009-005 *Bächle, A, Kimmerle, W.:*   Torsion subgroups in integral group rings of finite groups

2009-004 *Geisinger, L.; Weidl, T.:*   Universal bounds for traces of the Dirichlet Laplace operator

2009-003 *Walk, H.:*   Strong laws of large numbers and nonparametric estimation

2009-002 *Leitner, F.:*   The collapsing sphere product of Poincaré-Einstein spaces

2009-001 *Brehm, U.; Kühnel, W.:*   Lattice triangulations of E^3 and of the 3-torus

2008-006 *Kohler, M.; Krzyżak, A.; Walk, H.:*   Upper bounds for Bermudan options on Markovian data using nonparametric regression and a reduced number of nested Monte Carlo steps

2008-005 *Kaltenbacher, B.; Schöpfer, F.; Schuster, T.:*   Iterative methods for nonlinear ill-posed problems in Banach spaces: convergence and applications to parameter identification problems

2008-004 *Leitner, F.:*   Conformally closed Poincaré-Einstein metrics with intersecting scale singularities

2008-003 *Effenberger, F.; Kühnel, W.:*   Hamiltonian submanifolds of regular polytope

2008-002 *Hertweck, M.; Höfert, C.R.; Kimmerle, W.:*   Finite groups of units and their composition factors in the integral group rings of the groups $PSL(2,q)$

2008-001 *Kovarik, H.; Vugalter, S.; Weidl, T.:*   Two dimensional Berezin-Li-Yau inequalities with a correction term

2007-006 *Weidl, T.:*   Improved Berezin-Li-Yau inequalities with a remainder term

2007-005 *Frank, R.L.; Loss, M.; Weidl, T.:*   Polya's conjecture in the presence of a constant magnetic field

2007-004 *Ekholm, T.; Frank, R.L.; Kovarik, H.:*   Eigenvalue estimates for Schrödinger operators on metric trees

2007-003 *Lesky, P.H.; Racke, R.:*   Elastic and electro-magnetic waves in infinite waveguides

2007-002 *Teufel, E.:*   Spherical transforms and Radon transforms in Moebius geometry

2007-001 *Meister, A.:*   Deconvolution from Fourier-oscillating error densities under decay and smoothness restrictions