Published in final edited form as:

Stat Probab Lett. 2008 December 15; 78(18): 3307–3311. doi:10.1016/j.spl.2008.07.001.

Minimax estimation of the integral of a power of a density

Eric Tchetgen, Lingling Li, James Robins, and Aad van der Vaart*

Department of Biostatistics and Epidemiology Harvard University

*Department of Mathematics Vrije Universiteit

Abstract

We construct an estimator of $\int \mathcal{P}(z) dz$, based on a random sample of size n from a density f on the unit cube in $\mathbb{R}^d$. This estimator achieves the minimax rate for f known to belong to a multiple of the unit ball in a Hölder space of order α , where $\alpha \leq d/4$. We are mostly interested in the case that the power p is larger than 2 and/or the dimension d is large.

AMS 2000 subject classifications

Primary 62G05; 62G20

Keywords and phrases

Nonlinear functional; Nonparametric estimation; U -statistic

1. Introduction

Suppose that we observe an i.i.d. sample $Z_1, \dots, Z_n$ from a density f on the d -dimensional cube $[0, 1]^d$. We wish to estimate the functional $\psi(f; p) = \int \mathcal{P}(z) dz$, for a given, known integer $p \geq 2$, when it is known that f belongs to the Hölder space $C^\alpha[0, 1]^d$.

We concentrate on the case that the regularity α is low relative to the dimension:

$$\alpha \leq d/4. \quad (1.1)$$

In this case the square minimax rate of estimation over the unit ball of $C_\alpha[0, 1]^d$ is known (see [2]) to be not faster than

$$r_n := n^{-8\alpha/(d+4\alpha)}. \quad (1.2)$$

For values of α below the cut-off (1.1) this rate is slower than the standard parametric rate n^{-1} , and for large d it can be slow even for high regularity levels α . In this paper we show that the lower bound is sharp by constructing an estimator with mean squared error of order r_n .

This problem was previously considered by many authors, including [2], [1], [9], [4], as a canonical example of a nonlinear functional. A simple construction of a minimax estimator for the case $p = 2$ is given in [7], [8], and the case that $d = 1$ and $p = 3$ is covered by [6]. The authors of the latter paper also indicate that their construction extends to general smooth functionals ([6], Section 5).

Similar to the constructions by these authors, our estimator is based on approximations of f in a basis and an analysis of the bias of these approximations. The final estimator is a p th order U -statistic with a kernel determined by three approximation levels, but otherwise given by a simple and direct formula. Different from [6] our construction allows for general approximation schemes, not limited to the Haar basis. In fact, the latter basis cannot be used in the case $\alpha > 1$, as it gives suboptimal approximation in this case. The case $\alpha > 1$ can arise under (1.1) if $d > 4$. Not using the Haar basis leads to additional bias terms, which need to be estimated, whence our estimator is different from the one in [6], both for small and large d . However, actually the use of a general basis has lead us to simpler formulas.

The paper is organized as follows. In Section 2 we introduce the projection kernels used for the construction of our estimator. In Section 3 we very briefly mention the estimator in the quadratic case. In Section 4 we present our estimator for $p = 3$ and state the main result of the paper. Section 5 contains the proof of the main result.

Throughout the paper we use the following notation. Given a function $b: ([0, 1]^d)^m \rightarrow \mathbb{R}$ of m arguments, the U -statistic with kernel b_m is written

$$\mathbb{U}_n b = \frac{(n-m)!}{n!} \sum_{1 \leq i_1 \neq i_2 \neq \dots \neq i_m \leq n} b(Z_{i_1}, \dots, Z_{i_m})$$

The “kernel” b need not be permutation symmetric in this definition. As an alternative notation for this U -statistic we use $\mathbb{U}_n b_m(Z_{i_1}, \dots, Z_{i_m})$, where the unspecified indices $i_1, \dots, i_m$ serve as reminder of the arguments involved in the U -statistic. The notation $a \lesssim b$ means that $a \leq Cb$ for a constant C that is fixed within the context. Undelimited integrals are silently understood to be over the sample space $[0, 1]^d$.

2. Projections

Our estimator can be viewed as an unbiased estimator of an approximation to the functional $\psi(f, p)$, constructed from approximations $\bar{f}_k$ to f . In the case $p > 2$ we need to combine three such approximations, with different values of k , each taken as a projection onto a k -dimensional space.

We use a fixed orthogonal projection $K_k: L_2[0, 1]^d \rightarrow L_2[0, 1]^d$ given by a kernel operator, with kernel denoted by the same symbol as the operator: $K_k f(u) = \int K_k(u, v) f(v) dv$. The projection property $K_k^2 = K_k$ of the operator can be expressed in the kernel as

$$\int K_k(u, w) K_k(w, v) dw = K_k(u, v), \quad \text{a.e. } (u, v). \quad (2.1)$$

For simplicity of notation we assume that the kernel $(u, v) \rightarrow K_k(u, v)$ is such that this equation is valid for every $u, v \in [0, 1]^d$. Then we can formulate the first requirement on the kernel as: for every $m, k_1, \dots, k_m \in \mathbb{N}$,

$$\int \prod_{s=1}^m K_{k_s}(z, z) dz \lesssim \prod_{s=1}^m k_s. \quad (2.2)$$

We derive this property for a variety of kernels below. The second requirement is that the projection operator is suitable for approximation in our model space, the Hölder space $C^\alpha[0, 1]^d$: for every k , and $\|\cdot\|_\alpha$ a norm for $C^\alpha[0, 1]^d$,

$$\sup_{\|f\|_\alpha \leq 1} \|f - K_k f\|_\infty \lesssim \left(\frac{1}{k}\right)^{\alpha/d}. \quad (2.3)$$

This is a typical target for the construction of finite-dimensional approximations in approximation theory, and is known to be valid for Fourier, polynomial, spline and suitable wavelet approximations.

Condition (2.3) is the only condition where the Hölder class as prior model for the minimax result plays a role. If the kernels provide this rate of approximation for a different class of functions, then our estimator will attain the same mean square error over this class.

The kernel of an orthogonal projection on the linear span of an orthonormal set of functions $e_1, \dots, e_k$ in $L_2[0, 1]^d$ is given by $K_k(u, v) = \sum_{i=1}^k e_i(u)e_i(v)$. Thus condition (2.2) involves the functions $u \mapsto \sum_{i=1}^k e_i^2(u)$. If the basis functions are uniformly bounded, then (2.2) is clearly satisfied. Another case of interest is that the basis functions are spatially localized. We shall discuss this in detail for wavelets.

An orthonormal wavelet basis of $L_2(\mathbb{R}^d)$ is given in terms of functions $\psi_{i,j}^v$ indexed by a “resolution” (or scale) parameter $i \in \mathbb{N}$, a “location” parameter $j \in \mathbb{Z}^d$, and a “dimension index” $v \in \{0, 1\}^d$. Each function $\psi_{i,j}^v$ is a scaled and translated version of a fixed base function $\psi_{0,0}^v$ through

$$\psi_{i,j}^v(z) = 2^{id/2} \psi_{0,0}^v(2^i z - j), \quad i \in \mathbb{N}, j \in \mathbb{Z}^d, v \in \{0, 1\}^d.$$

The multiscale property of wavelets entails that for each resolution level I the functions $\psi_{i,j}^v$ with $i \geq I, j \in \mathbb{Z}^d$ and $v \in \{0, 1\}^d$, with $v = 0$ if $i > I$, span $L_2(\mathbb{R}^d)$. A suitable projection is obtained by leaving out the contributions of the base functions at resolution levels $i > I$, and retaining only the projections on the functions $\psi_{i,j}^v$ with $j \in \mathbb{Z}^d$ and $v \in \{0, 1\}^d$.

For base functions $\psi_{0,0}^v$ with compact support, for fixed I only of the order 2^{Id} functions $\psi_{I,j}^v$ are nonzero on $[0, 1]^d$. We denote the corresponding set of indices j by J_I . Even if the wavelets are orthonormal in $L_2(\mathbb{R}^d)$, the (restrictions of the) resulting set of functions $\{\psi_{I,j}^v : j \in J_I, v \in \{0, 1\}^d\}$ may not be orthonormal in $L_2[0, 1]^d$, as the support of some basis functions may not be fully included in $[0, 1]^d$. Following [3] this can be remedied by substituting appropriate “edge wavelets” for any such basis function. Thus we end up with an orthonormal basis $\{\psi_{I,j}, j \in J_I, v \in \{0, 1\}^d\}$ of a $k_I = O(2^{Id})$ -dimensional subspace of $L_2[0, 1]^d$.

The kernel of the orthogonal projection onto this subspace is given by

$$K_{k_I}(z_1, z_2) = \sum_{j \in J_I} \sum_{v \in \{0,1\}^d} \psi_{I,j}^v(z_1) \psi_{I,j}^v(z_2).$$

If the functions $\psi_{0,0}^v$ are uniformly bounded, then the uniform norm of each function $\psi_{I,j}^v$ is bounded up to a constant by $2^{Id/2}$. In view of the construction by scaling and translating, at a given point (z, z) only a fixed number of base functions $\psi_{I,j}^v$ are nonzero, for every fixed I . It follows that trivially $K_{k_I}(z, z) \lesssim 2^{Id} \sim k_I$, and hence (2.2) is satisfied.

To ensure the approximation property (2.3) the base functions must be appropriately chosen. For $\alpha = 1$ the Haar basis qualifies. In general one may construct a higher-dimensional wavelet basis as tensor products $\psi_{0,0}^v = \phi^{v_1} \times \cdots \times \phi^{v_d}$ (with $v = (v_1, \dots, v_d) \in \{0, 1\}^d$), given a pair of a one-dimensional father wavelet ϕ^0 and a mother wavelet ϕ^1 . Then the approximation property (2.3) is ensured if the mother wavelet has no less than $\lceil \alpha \rceil$ vanishing moments (see [5]).

3. Quadratic functional

To estimate the quadratic function $\psi(f; 2)$ we estimate the approximation $\psi(K_k f; 2)$ unbiasedly by the second-order U -statistic $\mathbb{U}_n K_k$. For appropriate projections this is precisely the estimator considered by [7], [8], and [6], and used in [10] to construct adaptive confidence sets. Thanks to (2.3) the bias is of the order

$$\psi(K_k f; 2) - \psi(f; 2) = \int (K_k f - f)^2(z) dz \lesssim \left(\frac{1}{k}\right)^{2\alpha/d}.$$

By standard computations on U -statistics (e.g. [11], Chapter 12) and (2.2) with $m = 1$ the variance can be seen to be of the order

$$\max\left(\frac{1}{n}, \frac{k}{n^2}\right).$$

The best trade-off $k \sim n^{2d/(d+4\alpha)}$ achieves mean square error r_n .

4. Main result

The estimation of $\psi(f; p)$ for $p \geq 3$ necessitates a more elaborate approximation scheme. We use three levels k_1, k_2, k_3 of truncation and define a p th order kernel by

$$\begin{aligned} b_{k_1, k_2, k_3}(Z_1, \dots, Z_p) = & \\ = \sum_{0 \leq t \leq 2} \binom{p}{t} \int \prod_{s=1}^{p-t} K_{k_1}(z, Z_s) \prod_{r=p-t+1}^p (K_{k_3}(z, Z_r) - K_{k_1}(z, Z_r)) dz & \\ + \binom{p}{3} \int \prod_{s=1}^{p-3} K_{k_1}(z, Z_s) \prod_{r=p-2}^p (K_{k_2}(z, Z_r) - K_{k_1}(z, Z_r)) dz & \\ + 3 \binom{p}{3} \int \prod_{s=1}^{p-3} K_{k_1}(z, Z_s) \prod_{r=p-2}^{p-1} (K_{k_2}(z, Z_r) - K_{k_1}(z, Z_r)) \times & \\ \times (K_{k_3}(z, Z_p) - K_{k_2}(z, Z_p)) dz. & \end{aligned}$$

Theorem 4.1. Let K_k be a projection kernel satisfying (2.2) and (2.3). Fix $p \geq 3$. For sequences $k_1(n) \sim n$, $k_3(n) \sim n^{2(1+4a/d)}$ and $k_2(n)$ such that $n^{(3/2-2a/d)(1+4a/d)} \lesssim k_2(n) \lesssim n^{(3/2+2a/d)(1+4a/d)}$, the sequence of estimators $\mathbb{U}_n b_{k_1(n), k_2(n), k_3(n)}$ has mean square error for estimating $\psi(f; p)$ of order r_n given in (1.2) uniformly over the class of densities $f \in \mathcal{C}^\alpha[0, 1]$ with $\|f\|_\alpha$ bounded by a fixed constant.

The estimator given in Theorem 4.1 is easiest understood in the cubic case $p=3$. In this case, we can expand the functional as, with $\bar{f}_k = K_k f$,

$$\begin{aligned} \psi(f; 3) = & \int \bar{f}_{k_1}^3 + 3 \int (\bar{f}_{k_3} - \bar{f}_{k_1})^2 \bar{f}_{k_1} + 3 \int (\bar{f}_{k_3} - \bar{f}_{k_1}) \bar{f}_{k_1}^2 \\ & + \int (\bar{f}_{k_2} - \bar{f}_{k_1})^3 + 3 \int (\bar{f}_{k_2} - \bar{f}_{k_1})^2 (\bar{f}_{k_3} - \bar{f}_{k_2}) \\ & + \int (\bar{f}_{k_3} - \bar{f}_{k_2})^3 + 3 \int (\bar{f}_{k_3} - \bar{f}_{k_2})^2 (\bar{f}_{k_2} - \bar{f}_{k_1}) + \int (f^3 - \int \bar{f}_{k_3}^3). \end{aligned} \quad (4.1)$$

In view of (2.3) the three terms in the last line of (4.1) can be shown to be of order $r_n^{1/2}$. The estimator $\mathbb{U}_n b_{k_1, k_2, k_3}$ is constructed to be unbiased for the sum of the first five terms in (4.1).

The expansion (4.1) is similar to the expansion used in [6]. An advantage of the Haar basis used in [6] is that the third, fifth and seventh term are then zero, and hence need not be estimated. However, to handle regularity $\alpha > 1$ we must use higher order kernels, for which the simplification no longer holds. It can be shown that for $\alpha/d \geq 1/20$ an estimator that ignores the third, fifth and seventh terms yields too large an approximation bias to give a mean square error of the desired order r_n .

5. Proof

In this section we present two lemmas followed by the proof of the main theorem. Throughout the section we assume the conditions of Theorem 4.1. The upper bounds are uniform in f ranging over a fixed multiple of the unit ball in $\mathcal{C}^\alpha[0, 1]^d$.

Lemma 5.1

For any positive integers $m, k_1, \dots, k_m$ and bounded function g ,

$$\mathbb{E} \left[\int g(z) \prod_{s=1}^m K_{k_s}(z, Z_s) dz \right]^2 \lesssim \|f\|_\infty^m \|g\|_\infty^2 \prod_{s=1}^{m-1} k_s.$$

Proof—The expected value is an integral relative to the density of $(Z_1, \dots, Z_m)$. By bounding this density by $\|f\|_\infty^m$, the integral is turned into an integral relative to the uniform measure. Thus we can assume without loss of generality that $Z_1, \dots, Z_m$ are uniformly distributed.

By writing the square of the integral as a repeated integral and using Fubini's theorem, we can rewrite the left side of the lemma as

$$\int \int \mathbb{E} \prod_{s=1}^m K_{k_s}(u, Z_s) K_{k_s}(v, Z_s) g(u) g(v) du dv.$$

The independence of $Z_1, \dots, Z_m$ and the projection property (2.1) permit to rewrite this in the form (note that the kernel is symmetric by orthogonality)

$$\begin{aligned} & \int \int \prod_{s=1}^m K_{k_s}(u, v) g(u) g(v) du dv \\ & \leq \int \int \prod_{s=1}^{m-1} \sqrt{K_{k_s}(u, u)} \sqrt{K_{k_s}(v, v)} g(u) g(v) K_{k_m}(u, v) du dv, \end{aligned}$$

by the Cauchy-Schwarz inequality applied to (2.1) and the identity $\int K_k^2(u, w) dw = K_k(u, u)$.

Defining a function H by $H(z) = \prod_{s=1}^{m-1} (K_{k_s}(z, z))^{1/2} g(z)$, we can write this as

$$\int (K_{k_m} H)(z) H(z) dz \leq \int H^2(z) dz.$$

Finally we bound g by $\|g\|_\infty$ and apply assumption (2.2). ■

Corollary 5.1

For any positive integers m and $k_1 \cdots k_m$ and bounded function g , there exists a constant C_m with

$$\mathbb{E} \left[\mathbb{U}_n \int g(z) \prod_{s=1}^m (K_{k_s}(z, Z_{i_s}) - K_{k_s} f(z)) dz \right]^2 \leq C_m \left(\frac{1}{n} \right)^m \|f\|_\infty^{3m} \|g\|_\infty^2 \prod_{s=1}^{m-1} k_s.$$

Proof—Because the U -statistic in the left side is degenerate, its variance as given in the left side is bounded by n^{-m} times the second moment of its kernel. By expanding the product of

the differences, we can bound the latter by a multiple of the sum over $v = (v_1, \dots, v_m) \in \{0, 1\}^m$ of expectations of the form

$$\mathbb{E} \left[\int \left(g(z) \prod_{s=1}^m K_{k_s} f(z)^{1-v_s} \right) \prod_{s=1}^m K_{k_s}(z, Z_{i_s})^{v_s} dz \right]^2.$$

To each of these expectations we can apply Lemma 5.1, with the function g in the lemma taken equal to the present functions $g \prod_{s=1}^m (K_{k_s} f)^{1-v_s}$. The term with $v = (1, \dots, 1)$ then yields a bound of the order $\prod_{s=1}^{m-1} k_s$, whereas the other terms yield upper bounds of smaller order. ■

Proof of Theorem 4.1—We first show that the variance of the estimator is of the order r_n . The p th order contribution to the Hoeffding decomposition of the estimator can be written as a linear combination of degenerate U -statistics of the form

$\mathbb{V}_p(t) = \mathbb{U}_n \left[\int \prod_{s=1}^p (K_{k_{t_s}}(z, Z_{i_s}) - \bar{f}_{k_{t_s}}(z)) dz \right]$, where $k_i = k_i(n)$ for $i = 1, 2, 3$ and $t = (t_1, \dots, t_p) \in \{0, 1, 2\}^p$ is one of the vectors

$$1_p, 1_{p-1}2_1, 1_{p-1}3_1, 1_{p-2}2_2, 1_{p-2}2_13_1, 1_{p-2}3_2, 1_{p-3}2_3, 1_{p-3}2_23_1.$$

(Here $1_k 2_l 3_m$ denotes the vector with k times coordinate 1, l times 2 and m times 3.) By

Corollary 5.1 the variance of $\mathbb{V}_p(t)$ is of the order $n^{-p} \prod_{s=1}^{p-1} k_{t_s}$. It follows that the variance of the p th order contribution to the Hoeffding decomposition is of the order

$n^{-p} \times \max(k_1^{p-1}, k_1^{p-2} k_2, k_1^{p-2} k_3, k_1^{p-3} k_2^2)$, which is of the order r_n . The lower order contributions to the Hoeffding decomposition of the estimator can be written as finite linear combinations of degenerate U -statistics of the form

$\mathbb{U}_n \left[\int \prod_{s=1}^p (K_{k_{t_s}}(z, Z_{i_s}) - \bar{f}_{k_{t_s}}(z))^{v_s} \bar{f}_{k_{t_s}}^{(1-v_s)}(z) dz \right]$, for $(v_1, \dots, v_p) \in \{0, 1\}^p$ and t ranging

over the same set as before. By Corollary 5.1 the variance is of the order $n^{-m} \prod_{s=1}^{p-1} k_{t_s}^{v_s}$, for

$m = \sum_{s=1}^p v_s$. As $k_i(n) \rightarrow n$ for every $i = 1, 2, 3$ these variances are of lower order than the variance of the p th order degenerate part, which we previously found to be of the desired order.

To study the bias of the estimator we expand the functional $\psi(f; p)$ as, with $a_t = \binom{p}{t}$,

$$\begin{aligned}
& \sum_{0 \leq t \leq 2} a_t \int \bar{f}_{k_1}^{p-t} (\bar{f}_{k_3} - \bar{f}_{k_1})^t + a_3 \int \bar{f}_{k_1}^{p-3} (\bar{f}_{k_2} - \bar{f}_{k_1})^3 \\
& \quad + 3a_3 \int \bar{f}_{k_1}^{p-3} (\bar{f}_{k_2} - \bar{f}_{k_1})^2 (\bar{f}_{k_3} - \bar{f}_{k_2}) \\
& + 3a_3 \int \bar{f}_{k_1}^{p-3} (\bar{f}_{k_2} - \bar{f}_{k_1}) (\bar{f}_{k_3} - \bar{f}_{k_2})^2 + a_3 \int \bar{f}_{k_1}^{p-3} (\bar{f}_{k_3} - \bar{f}_{k_2})^3 \\
& + \sum_{4 \leq t \leq p} a_t \int \bar{f}_{k_1}^{p-t} (\bar{f}_{k_3} - \bar{f}_{k_1})^t + \sum_{1 \leq t \leq p} a_t \int \bar{f}_{k_3}^{p-t} (f - \bar{f}_{k_3})^t.
\end{aligned}$$

The expectation of the estimator is equal to the sum of the first, second and third terms, whence the bias is the sum of the remaining terms. In view of (2.3) the norms $\|\bar{f}_{k_i}\|_\infty$ are uniformly bounded and, for every $i, j \in \{1, 2, 3\}$,

$$\|\bar{f}_{k_i} - \bar{f}_{k_j}\|_\infty \lesssim \left(\frac{1}{k_i \wedge k_j} \right)^{\alpha/d}.$$

The bias is therefore easily seen to be bounded by a multiple of

$$\left(\frac{1}{k_1} \right)^{\alpha/d} \left(\frac{1}{k_2} \right)^{2\alpha/d} + \left(\frac{1}{k_2} \right)^{3\alpha/d} + \left(\frac{1}{k_1} \right)^{4\alpha/d} + \left(\frac{1}{k_3} \right)^{2\alpha/d} + \left| \int \bar{f}_{k_3}^{p-1} (f - \bar{f}_{k_3}) \right|.$$

The first four terms can be seen to be of order $r_n^{1/2}$. The last term we rewrite as

$$\left| \int (f^{p-1} - K_{k_3}(f^{p-1}))(f - \bar{f}_{k_3}) + \int (\bar{f}_{k_3}^{p-1} - f^{p-1})(f - \bar{f}_{k_3}) \right| \lesssim \left(\frac{1}{k_3} \right)^{2\alpha/d}.$$

This concludes the proof of the theorem. ■

References

1. Bickel PJ, Ritov Y. Estimating integrated squared density derivatives: sharp best order of convergence estimates. *Sankhyā Ser A*. 1988; 50(3):381–393. MRMR1065550 (91e:62079).
2. Birgé L, Massart P. Estimation of integral functionals of a density. *Ann Statist*. 1995; 23(1):11–29. MRMR1331653 (96c:62065).
3. Cohen A, Daubechies I, Vial P. Wavelets on the interval and fast wavelet transforms. *Appl Comput Harmon Anal*. 1993; 1(1):54–81. MRMR1256527 (94m:42074).
4. Emery, M., Nemirovski, A., Voiculescu, D. Lectures on probability theory and statistics. Vol. 1738. Springer-Verlag; Berlin: 2000. Lecture Notes in Mathematics Lectures from the 28th Summer School on Probability Theory held in Saint-Flour, August 17–September 3, 1998, Edited by Pierre Bernard. MRMR1775638 (2001c:60003)
5. Härdle, W., Kerkycharian, G., Picard, D., Tsybakov, A. Lecture Notes in Statistics. Vol. 129. Springer-Verlag; New York: 1998. Wavelets, approximation, and statistical applications. MRMR1618204 (99f:42065)
6. Kerkycharian G, Picard D. Estimating nonquadratic functionals of a density using Haar wavelets. *Ann Statist*. 1996; 24(2):485–507. MRMR1394973 (97e:62062).
7. Laurent B. Efficient estimation of integral functionals of a density. *Ann Statist*. 1996; 24(2):659–681. MRMR1394981 (97d:62079).

8. Laurent B. Estimation of integral functionals of a density and its derivatives. *Bernoulli*. 1997; 3(2): 181–211. MRMR1466306 (99c:62144).
9. Laurent B, Massart P. Adaptive estimation of a quadratic functional by model selection. *Ann Statist*. 2000; 28(5):1302–1338. MRMR1805785 (2002c:62052).
10. Robins J, van der Vaart A. Adaptive nonparametric confidence sets. *Ann Statist*. 2006; 34(1):229–253. MRMR2275241.
11. van der Vaart, AW. Asymptotic statistics. Vol. 3. Cambridge University Press; Cambridge: 1998. Cambridge Series in Statistical and Probabilistic MathematicsMRMR1652247 (2000c:62003)