Nonparametric Denoising of Signals with Unknown Local Structure, I: Oracle Inequalities
Anatoli Juditsky, Arkadii S. Nemirovski

To cite this version:

HAL Id: hal-00318084
https://hal.archives-ouvertes.fr/hal-00318084
Submitted on 3 Sep 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Nonparametric Denoising of Signals with Unknown Local Structure, I: Oracle Inequalities

Anatoli Juditsky\textsuperscript{a,},\, Arkadi Nemirovski\textsuperscript{b}

\textsuperscript{a}LJK, B.P. 53, 38041 Grenoble Cedex 9, France
\textsuperscript{b}ISyE, Georgia Institute of Technology, 765 Ferst Drive, Atlanta GA 30332-0205 USA

Abstract

We consider the problem of pointwise estimation of multi-dimensional signals \( s \), from noisy observations \( (y_\tau) \) on the regular grid \( \mathbb{Z}^d \). Our focus is on the adaptive estimation in the case when the signal can be well recovered using a (hypothetical) linear filter, which can depend on the unknown signal itself.

The basic setting of the problem we address here can be summarized as follows: suppose that the signal \( s \) is “well-filtered”, i.e. there exists an adapted time-invariant linear filter \( q^*_T \) with the coefficients which vanish outside the “cube” \( \{0, \ldots, T\}^d \) which recovers \( s_0 \) from observations with small mean-squared error. We suppose that we do not know the filter \( q^* \), although, we do know that such a filter exists. We give partial answers to the following questions:

– is it possible to construct an adaptive estimator of the value \( s_0 \), which relies upon observations and recovers \( s_0 \) with basically the same estimation error as the unknown filter \( q^*_T \)?

– how rich is the family of well-filtered (in the above sense) signals?

We show that the answer to the first question is affirmative and provide a numerically efficient construction of a nonlinear adaptive filter. Further, we establish a simple calculus of “well-filtered” signals, and show that their family is quite large: it contains, for instance, sampled smooth signals, sampled modulated smooth signals and sampled harmonic functions.

Key words: Nonparametric denoising, oracle inequalities, adaptive filtering.

\textsuperscript{*} Corresponding author.

Email addresses: anatoli.juditsky@imag.fr (Anatoli Juditsky), nemirovs@isye.gatech.edu (Arkadi Nemirovski).
1 Introduction

In this paper, we focus on the problem of denoising of multi-dimensional signals. Let $F = (\Omega, \Sigma, P)$ be a probability space. We consider the problem of recovering unknown random field $(s_\tau = s_\tau(\xi))_{\tau \in \mathbb{Z}^d}$ over $\mathbb{Z}^d$ from noisy observations

$$y_\tau = s_\tau + e_\tau.$$  \hspace{1cm} (1)

It is convenient for us to assume that both the signal $(s_\tau)$ and the noises are complex-valued. Besides this, we assume that the field $(e_\tau)$ of observation noises is independent of $(s_\tau)$ and is of the form $e_\tau = \sigma \epsilon_\tau$, where $(\epsilon_\tau)$ are independent of each other standard Gaussian complex-valued variables; the adjective “standard” means that $\Re(\epsilon_\tau)$, $\Im(\epsilon_\tau)$ are independent of each other $N(0,1)$ random variables. Our focus here is at estimating the value $s_t$ of signal at a given location $t \in \mathbb{Z}^d$.

The above problem is “classical” in statistical estimation and signal processing, and as such, has received much attention. In particular, linear estimators (referred as linear filters in the signal processing community) are widely used in the statistical literature. To be more precise, suppose that our aim is to recover the value $s_0$ of the signal at zero given observations $(y_\tau)$ on the box $O_T = \{ \tau \in \mathbb{Z}^d : |\tau_j| \leq T, 1 \leq j \leq d \}$. We call the estimation $\hat{s}_t$ of $s_0$ linear if it is of the form

$$\hat{s}_t = \sum_{\tau \in O_T} q_\tau y_\tau$$

for some $q \in C(O_T)$, where $C(O_T)$ is the set of complex-valued fields $q = \{ q_\tau, \tau \in O_T \}$ over $O_T$.

The simplicity of linear estimators is responsible for their popularity in statistical signal processing. Another outstanding feature of such estimators is their minimax property. Suppose that the a priori information resumes to the fact that $(s_\tau)$ belongs to some convex compact set which is symmetric with respect to zero, let us call it $S$. One of the most renown results of estimation theory (see, for instance, [15,8,10]) states that the linear minimax estimator is, in a certain sense, an optimal estimator of $s_0$ in our problem. Indeed, consider the following linear minimax estimation strategy: let $q^{(T)}_t$ be the optimal solution \(^1\) to the problem

\(^1\) For evident reasons such a solution exists in the situation we are interested in.
\[
\min_{q \in C(\mathcal{O}_T)} \max_{s \in \mathcal{S}} E_s \left( s_0 - \sum_{\tau \in \mathcal{O}_T} q_{\tau} y_{\tau} \right)^2
\]

(here \( E_s \) stands for the expectation with respect to the distribution of \( (y_{\tau}) \) which corresponds to the underlying signal \( s \)). The linear minimax estimator \( \hat{s}_t^* \) of \( s_0 \) is defined by
\[
\hat{s}_t^* = \sum_{\tau \in \mathcal{O}_T} q_{\tau}^{(T)} y_{\tau}.
\]

Then
\[
\max_{s \in \mathcal{S}} E_s (s_0 - \hat{s}_t^*)^2 \leq C \inf_{\hat{s}} \max_{s \in \mathcal{S}} E_s (s_0 - \hat{s})^2,
\]

where the infimum in the right-hand side is taken over all possible estimators of \( s_0 \) from observations \( (y_{\tau}) \) and \( C \) is a moderate absolute constant (e.g., \( C \leq 1.25 \)). In other words, the linear estimator \( \hat{s}_t \) is a (almost) minimax estimator of \( s_0 \). We would like to stress the exceptional power of the above result – we only need \( \mathcal{S} \) to be convex and compact for the linear estimator to be minimax optimal. The evident downside of using linear minimax estimators is that the a priori information about the set \( \mathcal{S} \) of signals should be as precise as possible to achieve descent estimation accuracy. There was a significant research on adaptive estimation in the above setting (cf [6,7]). Those techniques allow to choose the “best” in a certain sense set which contains the signal from special finite families of convex sets. Another “classical” approach to adaptation for linear estimators has been developed in [20,21,22,23,28]. In the latter approach the “form” of the filter \( q_{\tau}^{(T)} \) is considered as given in advance (no information about sets of signals is used in this case), and the parameter \( T \) (the “window width”) is selected adaptively to achieve the best bias/variance tradeoff. Recently, more general adaptation techniques has been studied in [24,14], which allow to choose the best estimator from special finite families of available linear estimators.

The problem we are interested in here, when posed informally, is as follows: if we consider the form of the filter as a “free parameter”, is it possible to provide an estimation procedure which is adaptive with respect to this parameter? In other words, suppose that a “good” filter \( q_{\tau}^{(T)} \), with a small estimation error exists. Then, is it possible to construct a data-driven estimation method which has (almost) the same accuracy as the “oracle” – a hypothetic optimal estimation method which uses the “good” filter \( q_{\tau}^{(T)} \). It is natural, as it is common in adaptive nonparametric estimation, to measure the quality of an adaptive estimation routine with the factor by which the risk of the adaptive procedure is greater than that of the “oracle” estimator. What we look for is
the estimation method for which this factor is not too large. Let us consider, for instance, the following question:

(?) Suppose that know that the (deterministic or random) signal \((s_\tau(\xi))_{\tau \in \mathbb{Z}^d \equiv (s_{\tau,\xi})_{\tau \in \mathbb{Z}^d}, \xi \in \Omega}\) underlying observations (1) can be recovered from these observations “at a parametric rate” by “linear time-invariant filtering”: for a given \(T\), there exists (unknown in advance) filter \(q_s^{(T)}\) which recovers \(s_0\) via \(O(T^d)\) observations around zero such that

\[
E \left| s_0 - \sum_{\tau \in O_T} q_s^{(T)}(\tau) y_\tau \right|^2 \leq O(\sigma^2 T^{-d}).
\]

(2)

Can we mimic this filter?

We show that the answer to the question (?) is positive. Namely, whenever a discrete time signal (that is, a signal defined on a regular discrete grid) is well-filtered, i.e., can be recovered from its noisy observations at a parametric rate by a linear time-invariant filter, we can recover this signal at a “nearly parametric” rate without a priori knowledge of the associated filter.

Several points should be stressed in the above claim. First, we are able to mimic only ideal filters \(q_s^{(T)}\) of small \(l^2\)-norm. Indeed, the relation (2) implies that the stochastic term of the error \(E \left( \sum_{\tau \in O_T} q_s^{(T)}(\tau) e_\tau \right)^2\) is bounded with \(O(\sigma^2 T^{-d})\), which is conceivable only if \(|q_s^{(T)}|_{l^2} = O(T^{-d/2})\). This constraint is crucial, as the price for adaptation becomes prohibitive when the \(l^2\)-norm of the ideal filter is much larger than \(O(T^{-d/2})\). Though this assumption seems quite restrictive, the family of well-filtered signals is quite wide. As we shall see later, this family contains also “highly oscillating” sampled modulated smooth signals, sampled harmonic functions, etc.

In this paper we also treat the problem of adaptive prediction, when we are interested in recovering of a discrete time signal at a point \(t \in \mathbb{Z}^d\) via noisy observations taken at the points \(\{\tau \in \mathbb{Z}^d : t_j - T \leq \tau_j \leq t_j - \kappa\}\) “preceding” the point \(t\), with a given in advance “forecast horizon” \(\kappa \geq 0\).

The rest of our paper is organized as follows. In Section 2 we give a formal definition of a well-filtered (well-predicted) signal on a \(d\)-dimensional regular grid (the latter, w.l.o.g., is normalized to be \(\mathbb{Z}^d\)), and then show in Section 3 demonstrate that such a signal can be recovered at a nearly parametric rate without a priori knowledge of the corresponding “good filter” (Theorems 4 and 5). The underlying estimation routines (i.e., “Algorithm A” of Section 3.1 and “Algorithm B” of Section 3.2) constitute a substantial extension of the procedures proposed in [25] and [26]. In Section 4.1, we demonstrate that the
family of well-filtered signals is pretty wide – it contains a wide spectrum of “basic functions” (for example, exponential polynomials) and is closed with respect to a number of basic operations, including modulation, taking linear combinations and tensor products.

To make the exposition more readable, all proofs are collected in the appendix.

The denoising procedures, described in this paper constitute the basic bricks of the construction of adaptive estimators of locally well-filtered signals, which we describe in the companion paper[18]. The results of [18] extend to the wide classes of modulated signals the results of [27,12,16,13] on spatial adaptive estimates of signals with inhomogeneous smoothness.

2 Problem statement

In order to proceed we need some notations.

Fields over \( \mathbb{Z}^d \). Let \( C(\mathbb{Z}^d) \) be the linear space of complex-valued fields \( r = \{r_\tau : \tau \in \mathbb{Z}^d \} \) over \( \mathbb{Z}^d \).

- Given nonnegative integer \( T \) and \( p \in [1, \infty] \), we define semi-norms \(| \cdot |_{T,p}
\] on \( C(\mathbb{Z}^d) \) by \( |r|_{T,p} = \left( \sum_{|\tau| \leq T} |r_\tau|^p \right)^{1/p} \), \(|\tau| = \max\{|\tau_1|,\ldots,|\tau_d|\}\), with the standard interpretation of the right hand side when \( p = \infty \), and we set \(|r|_p = \lim_{T \to \infty} |r|_{T,p} \in \mathbb{R} \cup \{+\infty\}\). A field \( r \in C(\mathbb{Z}^d) \) with finitely many nonzero entries \( r_\tau \) is called a filter, and the smallest \( T \) such that \( r_\tau = 0 \) whenever \(|\tau| > T \), is called the order \( \text{ord}(r) \) of a filter \( r \); we write \( C_T(\mathbb{Z}^d) = \{r \in C(\mathbb{Z}^d) : \text{ord}(r) \leq T \} \). We identify a filter \( r \) with the multivariate Laurent sum \( r(z_1, \ldots, z_d) = \sum_{\tau} r_\tau z_1^{\tau_1} \cdots z_d^{\tau_d} \).

- We call a filter \( r \) polynomial, if the corresponding Laurent sum is a polynomial (i.e., if the entries \( r_\tau \) vanish when any of \( \tau_j < 0, j = 1, \ldots, d \)). The set of all polynomials is denoted \( P(\mathbb{Z}^d) \). For integers \( k, T, 0 \leq k \leq T \), we denote by \( P^k_T(\mathbb{Z}^d) \) the subspace of \( P(\mathbb{Z}^d) \) formed by polynomials \( r \) for which the entries \( r_\tau \) vanish outside the set \( k \leq \tau_j \leq T, j = 1, \ldots, d \).

- We denote by \( \Delta_j, j = 1, \ldots, d \), the “basic shift operators” on \( C(\mathbb{Z}^d) \):

\[
(\Delta_j r)_{\tau_1, \ldots, \tau_d} = r_{\tau_1, \ldots, \tau_j-1, \tau_j+1, \ldots, \tau_d}.
\]

Further, we use the notation \( \Delta_j^{-1} \) for the inverse of \( \Delta_j \):

\[
(\Delta_j^{-1} r)_{\tau_1, \ldots, \tau_d} = r_{\tau_1, \ldots, \tau_j-1, \tau_j-1, \tau_j+1, \ldots, \tau_d}.
\]

- Finally, we define the output of a filter \( r \), the input to the filter being a field \( x \in C(\mathbb{Z}^d) \), as the field \( r(\Delta)x \equiv r(\Delta_1, \Delta_2, \ldots, \Delta_d)x \), so that \( (r(\Delta)x)_t = \sum_{\tau} r_\tau x_{t-\tau} \).
Fourier transform. Let $T$ be a nonnegative integer, let $\Gamma_T$ be the set of roots of $1$ of the degree $2T + 1$, and let $C(\Gamma_T^d)$ be the space of complex-valued functions on $\Gamma_T^d \equiv (\Gamma_T)^d$.

- We define the Fourier transform $F_T : C(\mathbb{Z}^d) \to C(\Gamma_T^d)$ as $(F_T r)(\mu) = \left( \sum_{|\tau| \leq T} r_{\tau} \mu_{\tau_1} \cdots \mu_{\tau_d} \right)^{1/p}$, where $r \in C_T(\mathbb{Z}^d)$, and $\mu \in \Gamma_T$. Note that $r_{\tau} = \left( \frac{1}{(2T+1)^d/2} \sum_{\mu \in \Gamma_T^d} (F_T r)(\mu) \mu_{\tau_1} \cdots \mu_{\tau_d} \right)$. The Fourier transform allows to equip $C(\mathbb{Z}^d)$ with semi-norms coming from the standard $p$-norms on $C(\Gamma_T^d)$:

$$|r|_{T,p} = |F_T r|_p \equiv \left( \sum_{\mu \in \Gamma_T^d} |(F_T r)(\mu)|^p \right)^{1/p},$$

with the standard interpretation of the right hand side for $p = \infty$.

Now it is time to give a precise meaning to the basic question (?) of Introduction. In order to do this, we should specify our a priori knowledge of the constant factor hidden in $O(\cdot)$ and on the ranges on values of $T$ and $\tau$ where (2) holds true.

2.1 Nice signals

Since the observation noises are independent of $(s_\tau)$, we have

$$E \left\{ |s_\tau - (q(\Delta)y)_\tau|^2 \right\} = 2\sigma^2 q^2_2 + E_\xi \left\{ |s_\tau - (q(\Delta)s(\xi))_\tau|^2 \right\}; \quad (3)$$

therefore in order to ensure (2), both terms in the right hand side of the latter inequality should be of order of $T^{-d}$. This observation motivates the following

**Definition 1** Let $\theta \geq 0$, $\rho \geq 1$ be reals, let $L$ be a nonnegative integer or $+\infty$, and let $t \in \mathbb{Z}^d$. Finally, let $(s_\tau)_{\tau \in \mathbb{Z}^d} \equiv (s_\tau(\xi))_{\tau \in \mathbb{Z}^d}$ be a random field on $\mathbb{Z}^d$.

1. **[T-well-filtered signals]** Let $T$ be a nonnegative integer. We say that $(s_\tau)$ is $T$-well-filtered, with the parameters $\theta$, $\rho$, $L$, at the point $t$ (notation: $(s_\tau) \in S^d_L(\theta, \rho, T)$), if there exists a filter $q = q(T) \in C_T(\mathbb{Z}^d)$, $|q|_2 \leq \frac{\rho}{(2T+1)^d/2}$, which reproduces $(s_\tau)$ in the box $\{ \tau : |\tau - t| \leq L \}$ with the mean square error not exceeding $\theta(2T + 1)^{-d/2}$:

$$\max_{\tau: |\tau - t| \leq L} \left[ E \left\{ |s_\tau - (q(\Delta) s)_\tau|^2 \right\} \right]^{1/2} \leq \theta(2T + 1)^{-d/2}. \quad (4)$$

2. **[well-filtered signals]** We say that $(s_\tau)$ is well-filtered, with the parameters
\( \theta, \rho, L, \) at the point \( t \) (we use the notation: \( (s_\tau) \in F_L^t(\theta, \rho) \)), if, for every integer \( T, 0 \leq T \leq L, \) \( (s_\tau) \) is \( T \)-well-filtered, with the parameters \( \theta, \rho, L, \) at \( t. \)

In the above definition we were focusing on the case of de-noising – recovering a well-filtered signal \( (s) \) at a point \( t \in \mathbb{Z}^d \) via a given number observations “around” this point.\(^2\) Another interesting problem is that of prediction, where the goal is to recover \( s_t \) via observations \( y_{\tau} \) “preceding by a given horizon \( \kappa \in \mathbb{Z}_+ \)” the point \( t \), i.e., observations with \( \tau_j \leq t_j - \kappa, j = 1, \ldots, d. \)

**Definition 2** Let \( \theta \geq 0, \rho \geq 1 \) be reals, let \( \Lambda_0 \geq \kappa \) be nonnegative integers, \( L \) be a nonnegative integer or \( +\infty \), and let \( t \in \mathbb{Z}^d \). Finally, let \( (s_\tau)_{\tau \in \mathbb{Z}^d} \equiv (s_\tau(\xi))_{\tau \in \mathbb{Z}^d} \) be a random field on \( \mathbb{Z}^d \).

(1) **[T-well-predicted signals]** Let \( T \) be a nonnegative integer. We say that \( (s_\tau) \) is \( T \)-well predicted with the parameters \( \theta, \rho, \kappa, L, \) at the point \( t \) (notation: \( (s_\tau) \in Q_{\Lambda_0, T}^t(\theta, \rho, T) \)), if there exists a filter \( q = q(T) \in P_T^\kappa(\mathbb{Z}^d), \) \( |q|_2 \leq \frac{\rho}{(2T+1)^{d/2}}, \) which reproduces \( (s_\tau) \) in the box \( \{\tau : |\tau - t| \leq L \} \) with the mean square error not exceeding \( \theta(2T + 1)^{-d/2}: \)

\[
\max_{\tau : |\tau - t| \leq L} \left[ E \left\{ |s_\tau - q(\Delta)s_\tau|^2 \right\} \right]^{1/2} \leq \theta(2T + 1)^{-d/2}. \tag{5}
\]

(2) **[well-predicted signals]** We say that \( (s_\tau) \) is well-predicted, with the parameters \( \theta, \rho, \kappa, \Lambda_0, L, \) at the point \( t \) (notation: \( (s_\tau) \in P_{\Lambda_0, \Lambda_0, \theta, \rho}^t(\theta, \rho) \)), if, for every integer \( T, \Lambda_0 \leq T \leq L, \) \( (s_\tau) \) is \( T \)-well-predicted, with the parameters \( \theta, \rho, \kappa, L, \) at \( t. \)

**Remark 3** Note that the quantitative description of a well-predicted field, when compared with the description of a well-filtered field, involves an extra parameter \( \Lambda_0 \) – the smallest “window width” starting with which a possibility to predict \( s_t \) is postulated. In the case of well-filtered fields, this width is just 0, in full accordance with the fact that in the de-noising problem every signal is 0-well-filtered, at every point \( t, \) with parameters \( \theta = 0, \rho = 1, L = \infty \) due to the existence of the trivial “single-point” filter \( q(z) \equiv 1. \)

In the sequel, we qualify as nice a signal which fulfils the requirements of Definition 2 or 1 above. The filters \( q(T) \) associated, in the sense of the above definitions, with a nice signal \( (s_\tau) \) as to filters certifying the “niceness” (“well-filterability” of “well-predictability”) of the signal.

We are about to demonstrate that in the framework, suggested by the above definitions, the answer to the question (7) is affirmative, i.e., a signal which is nice (\( T \)-well-filtered or \( T \)-well-predicted, with parameters \( \theta, \rho, L = 3T \)) at a point \( t \) can be recovered at this point “at a nearly parametric rate” with no

\(^2\) To be more precise, in the filtering literature this case is referred to as interpolation.
a priori knowledge of the corresponding “good filter”; all we should know in advance are the parameters $\rho$ and $T$.

3 Main result

We start the recovering routine for the adaptive filtering problem.

3.1 Adaptive filtering

The estimator we intend to use is as follows:

**Algorithm A:** Given a setup $(\rho \geq 1, T)$ and a point $t \in \mathbb{Z}^d$, we build an estimation $\hat{s}_t[T, y]$ of $s_t$ via observations $(y_\tau)$, $|\tau - t| \leq 4T$, as follows:

1. When $T = 0$, we merely set $\hat{s}_t[0, y] = y_t$
2. When $T > 0$, we set $\hat{s}_t[T, y] = (\hat{\phi}^t(\Delta)y)_t$, where $\hat{\phi}^t \in C_{2T}(\mathbb{Z}^d)$ is an optimal solution to the following optimization problem:

$$\min_{\phi \in C_{2T}(\mathbb{Z}^d)} \left\{ \frac{|\Delta^{1-t_1} \ldots \Delta^{d-t_d}(1 - \phi(\Delta))y|_2}{2T, \infty} : |\phi|_{2T, 1} \leq 2^{d/2} \rho^2 (2T + 1)^{-d/2} \right\}$$

where $y_L^t = \{y_\tau : |t - \tau| \leq L\}$.

Note that the objective in (6) is affected only by observations $y_{4T}^t$, so that our algorithm recovers $s_t$ via $(8T + 1)^d$ observations “around” the point $t$.

**Theorem 4** Assume that the signal $(s_\tau)$ underlying observations (1) is $T$-well-filtered, with parameters $\theta$, $\rho$, $L \geq 3T$: $(s_\tau) \in S_L^d(\theta, \rho, T)$ with $L \geq 3T$. Then the mean square error of the estimate $\hat{s}_t[T, \cdot]$ of $s_t$ yielded by Algorithm A with setup $(\rho, T)$ can be bounded from above as follows:

$$\left( \mathbb{E}\left\{ |\hat{s}_t[T, y] - s_t|^2 \right\} \right)^{1/2} \leq c(d) \rho^3 \frac{\theta + \sigma^2}{(2T + 1)^{d/2}}$$

where $c(d) = 3(2^d + 2^{3d-1})$.

In particular, if $(s_\tau)$ is well-filtered, with the parameters $\theta$, $\rho$, $L$, at a point $t$, then for every integer $T$, $0 \leq T \leq \lfloor L/3 \rfloor$, the accuracy of the estimate $\hat{s}_t[T, y]$ of $s_t$ yielded by Algorithm A can be bounded by (7). Finally, in the case of
**deterministic** \((s)\), we have

\[
|s_t - \tilde{s}_t[T, y]| \leq c(d)\rho^2 [\theta + \sigma\rho\Theta_T] (2T + 1)^{-d/2},
\]

\[
\Theta_T' = \sigma^{-1} \max_{|\tau| \leq 2T} |\Delta_1^{t_1} \ldots \Delta_d^{t_d} e|^{*}_{2T, \infty}.
\]

Comments: note that Theorem 4 gives an affirmative answer to the question (\(?\)). Indeed, let a signal \((s_\tau)\) admit, for some \(T\), a filter-type estimate \(\tilde{s}_\tau = (q^*(\Delta)y)_\tau\) with “window width” \(T\) (i.e., with \(q^* \in C_T(\mathbb{Z}^d)\)) and with the mean square error which, in an \(O(T)\)-neighborhood of a point \(t\), is of the “parametric” order \(O(\sigma(2T + 1)^{-d/2})\):

\[
\max_{|\tau - t| \leq 3T} E\{|s_\tau - \tilde{s}_\tau|^2\} \leq \kappa^2 \equiv \frac{\sigma^2 \mu^2}{(2T + 1)^{d/2}}
\]

with some known \(\mu \geq 1\). We do not know what is this estimate, although do know that it exists (i.e., know the associated \(T, \mu\)), and we want to recover \(s_t\) from observations \(y_{4T}^t\) nearly as well as if we were using our hypothetic estimate \(\tilde{s}_t\). Theorem 4 says that Algorithm A basically achieves this goal. Indeed, from (3), (9) it follows that \(|q^*|_2 \leq \frac{\mu}{(2T + 1)^{d/2}}\) and \((s_\tau) \in S_{3T}(\sigma, \mu, T)\). Applying Theorem 4 with \(\rho = \mu, \theta = \sigma\mu, L = 3T\), we conclude that with the estimate yielded by Algorithm A, the mean square error of recovering \(s_t\) does not exceed \(O(1)\mu^3 \left[1 + \sqrt{\ln(2T + 1)}\right] \kappa\). We see that as far as the dependence on “observation time” \(T^d\) is concerned, the estimate yielded by Algorithm A is just by a logarithmic in \(T\) factor worse than the estimate \(\tilde{s}_t\) we wish to mimic.

In the literature on nonparametric estimation the bounds as in Theorem 4 are often referred to as oracle inequalities. Since the pioneering work [1] a number of oracle inequalities have been established for a wide variety of estimation problems (cf. the papers [19], [2], [3], [4], [9], [11], [5] among many others). In that context one refer to the filter \(q\), which certifies the niceness of the signal, as the oracle, and the bound (7) describes the ability of a particular adaptive method (Algorithm A above) to reproduce the oracle.

Note that the “upper bound” of Theorem 4 may be compared to the lower bound of Theorem 2 of [17] for the 1-dimensional situation. The latter result states that one can exhibit a family of signals which 1) each member of the family can be recovered with the rate \(O(\sigma\rho\sqrt{T})\) using the corresponding certifying filter; 2) the rate of estimation of signals from the family using the observation (1) is at best \(O(\sigma\rho^2 \sqrt{T})\). In other words, it states that the factor \(\rho\sqrt{T}\) is an unavoidable “price” for adaptation. When comparing the result of Theorem 4 to that lower bound, we observe an extra factor \(\rho^2 \geq 1\) in the corresponding upper bound (7). By now we do not know if this extra factor can be completely eliminated. Nevertheless, in light of these results, we
can claim that recovering of signals with certifying filter of large $l_2$-norm is a rather desperate task – the price for adaptation is then proportional to $\rho \gg 1$ in this case.

### 3.2 Adaptive prediction

We now turn to the problem of adaptive prediction. The predictor we intend to use is as follows:

**Algorithm B:** Given a setup $(\rho \geq 1, \kappa, T)$ and a point $t \in \mathbb{Z}^d$, we build a prediction $\hat{s}_t[\kappa, T, y]$ of $s_t$ via observations $(y_\tau)$, $\kappa \leq t - \tau_j \leq 4T$, $j = 1, \ldots, d$, as $\hat{s}_t[\kappa, T, y] = (\hat{\psi}^t(\Delta)y)_t$, where $\hat{\psi}^t \in P_{2T}^\kappa(\mathbb{Z}^2)$ is an optimal solution to the following optimization problem:

$$\min_{\psi \in C_{2T}^\kappa(\mathbb{Z}^d)} \left\{ \frac{1}{\| \Delta_1 \ldots \Delta_d (1 - \psi(\Delta))y \|_{2T,\infty}} : \| \psi \|_{2T,1} \leq \frac{2^{d/2} \rho^2}{(2T + 1)^{d/2}} \right\};$$

where $y_{\kappa,L} = \{y_\tau : \kappa \leq t - \tau_j \leq L, j = 1, \ldots, d\}$.

Note that the objective in (10) is affected only by observations $y_{\kappa,4T}$, so that our algorithm recovers $s_t$ via $(4T - \kappa + 1)^d$ observations “around” the point $t$.

**Theorem 5** Assume that the signal $(s_\tau)$ underlying observations (1) is $T$-well-predicted, with parameters $\theta$, $\rho$, $\kappa$, $L \geq 3T$: $(s_\tau) \in Q^{\kappa,L}_{\kappa,4T}(\theta, \rho, T)$ with $L \geq 3T$. Then the mean square error of the estimate $\hat{s}_t[\kappa, T, \cdot]$ of $s_t$, provided by Algorithm B with setup $(\rho, \kappa, T)$, can be bounded from above as follows:

$$\left( E \left\{ |\hat{s}_t[\kappa, T, y] - s_t|^2 \right\} \right)^{1/2} \leq c(d)\rho^{3/2} \frac{\theta + \sigma \rho \sqrt{\ln(2T + 1) + 1}}{(2T + 1)^{d/2}},$$

$$c(d) = 3(2^d + 2^{3d-1}).$$

In particular, if $(s_\tau)$ is well-predicted, with the parameters $\theta$, $\rho$, $\kappa$, $T_0$, $L$, at a point $t$, then for every integer $T$, $T_0 \leq T \leq |L/3|$, the accuracy of the estimate $\hat{s}_t[\kappa, T, \cdot]$ of $s_t$ yielded by Algorithm B can be bounded by (11).

Finally, in the case of deterministic $(s)$, we have

$$|s_t - \hat{s}_t[T, y]| \leq c(d)\rho^{3/2} \frac{\theta + \sigma \rho \Theta_T}{(2T + 1)^{d/2}},$$

$$\Theta_T = \sigma^{-1} \max_{|\tau| \leq 2T} |\Delta_1^{t_1 - t_1} \ldots \Delta_d^{t_d - t_d} e|_{2T,\infty}.$$  \hspace{1cm} (12)

The proof of Theorem 5 is identical to that of Theorem 4.
4 Families of nice signals

When applying Algorithms A, B and Theorems 4, 5, the crucial question is how to recognize niceness. We are about to give a partial answer to this question.

4.1 Calculus of nice signals

Our current goal is to understand how wide are the families of nice signals, and our plan is as follows: (a) we list a number of operations which preserve the property in question, and (b) we present a list of examples of signals possessing the property. Applying to “raw materials” from (b) operations from (a), one can produce a wide variety of nice signals. Here is a sample of operations preserving niceness of signals.

I. “Scale” of nice signals. We start with the following evident observation:

$$\rho' \geq \rho, \theta' \geq \theta, L' \leq L \Rightarrow F^t_{L'}(\theta, \rho) \subset F^t_L(\theta, \rho')$$

and

$$\rho' \geq \rho, \theta' \geq \theta, \kappa' \leq \kappa, T_0' \geq T_0, L' \leq L \Rightarrow P^t_{\kappa,T_0,L'}(\theta, \rho) \subset P^t_{\kappa',T_0,L}(\theta', \rho').$$

II. Taking linear combinations. Our next observation is that a linear combination of “good” signals is again good, with properly updated parameters:

Proposition 6 (i) Let $$(s^j_t) \in F^t_L(\theta_j, \rho_j),$$ and let $\lambda_j \in \mathbb{C}$ be random variables independent of $$(s^j_t)$$ and such that $E\{|\lambda_j|^2\} < \infty$, $j = 1, \ldots, m$. Then

$$(s_t \equiv \sum_{j=1}^m \lambda_j s^j_t) \in F^t_{L^+}(\theta^+, \rho^+),$$

$$\theta^+ = (2m - 1)^{d/2}2^{m-1}\rho_1\ldots\rho_m \sum_{j=1}^m \frac{\theta_j E\{|\lambda_j|^2\}^{1/2}}{\rho_j},$$

$$\rho^+ = (2m - 1)^{d/2}2^{m-1}\rho_1\ldots\rho_m, L^+ = \lfloor L/2 \rfloor.$$ (13)

In the case of $m = 1$, one can set $\rho^+ = \rho_1$, $\theta^+ = |\lambda_1|\theta_1$, $L^+ = L$. The filters certifying the well-filterability of $$(s_t)$$ can be chosen to be independent of the coefficients $\lambda_j$.

(ii) Let $$(s^j_t) \in P^t_{\kappa_j,T_0^j,L}(\theta_j, \rho_j),$$ $j = 1, \ldots, m$, and let $\lambda_j \in \mathbb{C}$ be random variable
independent of $(s^j)$ and such that $E\{|\lambda_j|^2\} < \infty$, $j = 1, \ldots, m$. Then

$$
(s_\tau \equiv \sum_{j=1}^m \lambda_j s^j) \in \mathbf{P}^\tau_{\kappa^+,L^+}(\theta^+, \rho^+),
$$

$$
\theta^+ = (2m - 1)d/2^{m-1} \rho_1 \ldots \rho_m \sum_{j=1}^m 2|E\{|\lambda_j|^2\}|^{1/2},
$$

$$
\rho^+ = (2m - 1)d/2^{m-1} \rho_1 \ldots \rho_m, \quad \kappa^+ = \min_{1 \leq j \leq m} \kappa_j, \quad T_0^+ = \max_{1 \leq j \leq m} T_0^j,
$$

$$
L^+ = [L/2].
$$

In the case of $m = 1$, one can set $\rho^+ = \rho_1$, $\theta^+ = |\lambda_1|\theta_1$, $\kappa^+ = \kappa$, $T_0^+ = T_0$, $L^+ = L$. The filters certifying the well-predictability of $(s_\tau)$ can be chosen to be independent of the coefficients $\lambda_j$.

III. Modulation and conjugation. Next we notice that the families of nice signals are closed w.r.t. “modulation” and conjugation:

**Proposition 7** (i) Let $(s_\tau) \in \mathbf{F}_L^t(\theta, \rho)$, and let $\omega \in \mathbb{R}^d$, $\phi \in \mathbb{R}$ be deterministic. Then the signal $(\bar{s}_\tau = \exp\{i[\omega^T \tau + \phi]\} s_\tau)_{\tau \in \mathbb{Z}^d}$ belongs to $\mathbf{F}_L^t(\theta, \rho)$ along with $(s_\tau)$, and the signal $(\bar{s}_\tau = \bar{\pi}_\tau)_{\tau}$ (where $\bar{\pi}$ is the complex conjugate of a $\in \mathbb{C}$) belongs to $\mathbf{F}_L^t(\theta, \rho)$.

(ii) Let $(s_\tau) \in \mathbf{P}_{\kappa, T_0, L}^{\tau+}(\theta, \rho)$, and let $\omega \in \mathbb{R}^d$, $\phi \in \mathbb{R}$ be deterministic. Then the signal $(\bar{s}_\tau = \exp\{i[\omega^T \tau + \phi]\} s_\tau)_{\tau \in \mathbb{Z}^d}$ also belongs to $\mathbf{P}_{\kappa, T_0, L}^{\tau+}(\theta, \rho)$, and the signal $(\bar{s}_\tau = \bar{\pi}_\tau)_{\tau}$ belongs to $\mathbf{P}_{\kappa, T_0, L}^{\tau+}(\theta, \rho)$.

IV. Lifting. We are about to show that a nice signal in a dimension $d \leq d^+$ can be viewed as a nice signal, with properly updated parameters, in a dimension $d^+ > d$:

**Proposition 8** (i) Let $1 \leq d \leq d^+$, and let $(s_\tau)_{\tau \in \mathbb{Z}^d}$ be a signal which is well-filtered, with parameters $\theta, \rho, L$, at a point $t \in \mathbb{Z}^d$. Then the signal $(s_\tau^+ = s_{\tau_1, \ldots, \tau_d})$ is well-filtered, with the parameters $\theta^+ = (2L + 1)(d^+ - d)/2 \theta$, $\rho^+ = \rho$, $L^+ = L$ at every point $t^+ \in \mathbb{Z}^{d^+}$ such that $(t_1^+, \ldots, t_d^+) = t$.

(ii) Let $1 \leq d \leq d^+$, and let $(s_\tau)_{\tau \in \mathbb{Z}^d}$ be a signal which is well-predictable, with parameters $\theta, \rho, \kappa, T_0, L$, at a point $t \in \mathbb{Z}^d$. Then the signal $(s_\tau^+ = s_{\tau_1, \ldots, \tau_d})$ is well-predictable, with the parameters $\theta^+ = (2L + 1)(d^+ - d)/2 \theta$, $\rho^+ = (2\kappa + 1)(d^+ - d)/2 \rho$, $\kappa^+ = \kappa$, $T_0^+ = T_0$, $L^+ = L$, at every point $t^+ \in \mathbb{Z}^{d^+}$ such that $(t_1^+, \ldots, t_d^+) = t$.

V. “Tensor product”. Let $d = d^* + d^\prime$ with positive integers $d^*$, $d^\prime$, so that $\mathbb{Z}^d = \mathbb{Z}^{d^*} \times \mathbb{Z}^{d^\prime}$. Given random fields $(s_\tau^*(\xi))_{\tau \in \mathbb{Z}^{d^*}}$, $(s_\tau^\prime(\xi))_{\tau \in \mathbb{Z}^{d^\prime}}$, we define their tensor product as the field $(s_\tau(\xi) = s_\tau^*(\xi)s_\tau^\prime(\xi))_{\tau \in \mathbb{Z}^{d^*} \times \mathbb{Z}^{d^\prime}}$.

**Proposition 9** (i) Let $(s_\tau^*(\xi))_{\tau \in \mathbb{Z}^{d^*}} \in \mathbf{F}_L^{d^*}(0, \rho^*)$, $(s_\tau^\prime(\xi))_{\tau \in \mathbb{Z}^{d^\prime}} \in \mathbf{F}_L^{d^\prime}(0, \rho^\prime)$. 

12
Then \((s_\tau) \in \mathbf{F}_L^{(t',t'')} (0, \rho \rho'').\)

(ii) Let \((s_{\tau'}(\xi))_{\tau' \in \mathbf{E}} \in \mathbf{P}_{\kappa,T_0,L} (0, \rho'),\) \((s_{\tau''}(\xi))_{\tau'' \in \mathbf{E}} \in \mathbf{P}_{\kappa,T_0,L} (0, \rho'').\) Then \((s_\tau) \in \mathbf{P}_{\kappa,T_0,L} (0, \rho' \rho'').\)

4.2 Examples of nice signals

I. Exponential and algebraic polynomials. Let us define an exponential polynomial \((s_\tau)\) on \(\mathbb{Z}^d\) as a finite sum of exponential monomials

\[
\exp \{ \omega^T \tau \} \equiv \sum_{\ell=1}^{M} c_\ell \tau^\alpha(\ell) \exp \{ \omega^T(\ell) \tau \},
\]

where \(\omega(\ell)\) and \(\alpha(\ell)\) are deterministic, and \(c_\ell\) may be random. Given an exponential polynomial \((s_\tau)\) on \(\mathbb{Z}^d\), we define its partial sizes \(N_j, j = 1, ..., d,\) as follows: let \(m_j\) be the maximum of the degrees \(\alpha_j(\ell), \ell = 1, ..., M,\) of the variable \(\tau_j\) in the monomials of the sum (15), and \(M_j\) be the number of distinct from each other complex numbers among the “partial frequencies” \(\omega_j(\ell): M_j = \text{Card} O_j, O_j = \{ \omega_j(\ell) : 1 \leq \ell \leq M \}.\) The \(j\)-th partial size \(N_j(s)\) of exponential polynomial (15) is, by definition, the integer \((m_j + 1)M_j.\) For example, with all frequencies equal to 0, an exponential polynomial becomes an algebraic polynomial, and its \(j\)-th size is by 1 larger than the degree of the polynomial w.r.t. \(j\)-th variable \(\tau_j.\)

**Proposition 10** Let \((s_\tau)\) be an exponential polynomial on \(\mathbb{Z}^d\) of partial sizes \(N_1, ..., N_d.\) Then for all \(t \in \mathbb{Z}^d\) one has

\[
(s_\tau) \in \mathbf{F}_L^t (0, \rho_d(N_1, ..., N_d)), \rho_d(N_1, ..., N_d) = \prod_{j=1}^{d} [(2N_j - 1)^{1/2} 2^{3N_j/2}],\]

and the filters \(q^T\) certifying this inclusion can be chosen to be dependent solely on \(T\) and on the collection of \(d\) sets \(O_j = \{ \omega_j(\ell) : 1 \leq \ell \leq M \}\) of partial frequencies.

**Remark 11** A major shortcoming of (16) is a dramatic growth of \(\rho_d(N, N, ..., N)\) with \(N\) and \(d.\) In several important cases, better bounds for \(\rho\) can be found. For example, an algebraic polynomial of degree \(m\) in every variable

\[
p_\tau = \sum_{\alpha \geq 0, |\alpha| \leq m} c_\alpha \tau^\alpha\]

13
belongs to $F^t_{\infty}(0,(16m)^d)$ for every $t$, and the filters $q^{(T)}$ certifying this inclusion can be chosen to depend solely on $T,d,m$.

II. Solutions to homogeneous difference equations and harmonic functions.

Consider a difference operator $D$:

$$(Df)_\tau = \sum_{\ell=1}^{k} w_\ell f_{\tau-\alpha(\ell)}; \quad (18)$$

here $\alpha(1),...,\alpha(k) \in \mathbb{Z}^d$ and $w_1,...,w_k \in \mathbb{C}$. For a positive integer $N$ and $t \in \mathbb{Z}^d$, let

$$B^t_N = \{\tau \in \mathbb{Z}^d \mid |\tau-t| \leq N\}, \quad B^t_N(D) = \{\tau \in B^t_N \mid \tau + \alpha(\ell) \in B^t_N, \ell = 1,...,k\},$$

$$H^t_N(D) = \{(s) \in C(\mathbb{Z}^d) \mid s_\tau = (Ds)_\tau \quad \forall \tau \in B^t_N(D)\}.$$  

For example, with

$$(Df)_\tau = \frac{1}{2d} \sum_{i=1}^{d} f_{\tau_1,...,\tau_{i-1},\tau_i+\epsilon,\tau_{i+1},...,\tau_d}; \quad (19)$$

the linear space $H^t_N(D)$ is the space of fields which are “discrete harmonic” on $B^t_N$, that is, $s_\tau = \frac{1}{2d} \sum_{i=1}^{d} s_{\tau_1,...,\tau_{i-1},\tau_i+e,\tau_{i+1},...,\tau_d}$ for all $\tau$ with $|\tau-t| \leq N-1$. Let us call a difference operator $D$ regular, if it possesses the following properties:

R.1 The vectors $\{\alpha(\ell)\}_{1 \leq \ell \leq k}$ span the entire $\mathbb{R}^d$;

R.2 The coefficients $w_\ell = \rho_\ell \exp\{i\phi_\ell\}$ ($\rho_\ell \geq 0$, $\phi_\ell \in \mathbb{R}$) are nonzero, and

$$(a) \quad \sum_{\ell=1}^{k} \rho_\ell \leq 1; \quad (b) \quad \sum_{\ell=1}^{k} \rho_\ell \alpha(\ell) = 0. \quad (20)$$

For example, the averaging operator (19) and its degrees are regular. It turns out that the solutions of homogeneous difference equations with regular difference operators are well-filtered:

**Proposition 12** Let $D$ be a regular difference operator. Then there exists a constant $c = c(D) > 0$ such that

$$\forall N > 0 : \quad H^t_N(D) \subset F^t_{\lfloor cN \rfloor}(0,c^{-1}). \quad (21)$$

As a nontrivial application example for Proposition 12, consider the families of random fields defined as follows. Let $d \leq 4$, $M$ be a positive integer, and $R$ be a positive real. Consider the family $H^+(M)$ of all deterministic continuous
functions $f$ on $\mathbb{R}^d$ which are harmonic in the interior of the box $D_{2M}^0 = \{ x \in \mathbb{R}^d : |x_j| \leq 2M, j \leq d \}$: \( \left( \sum_{j=1}^d \frac{\partial^2 f}{\partial x_j^2} \right) f(x) = 0, \ x \in \text{int}D_{2M}^0. \) Now let $H^+(M,R)$ be the family of random functions $f$ such that all realizations of a function belong to $H^+(M)$ and, besides this, \( E\{\| f \|_{\infty,2M}^2 \} \leq R^2, \) where $\| f \|_{\infty,2M}$ is the uniform norm on $D_{2M}^0$. Restricting functions $f$ from $H^+(M,R)$ on $\mathbb{Z}^d$, we get a family of random fields $H(M,R)$ on $\mathbb{Z}^d$.

Proposition 13 Let $d \leq 4$, $M$ be a positive integer and $R > 0$ be a real. For an appropriately chosen absolute constant $c > 0$, for all deterministic fields $(s_\tau) \in H(M,R)$ one has
\[
|t| \leq cM, L \leq cM \Rightarrow (s_\tau) \in F_L^t(c^{-1}R,c^{-1}),
\]
and the filters $q^{(T)}$ certifying the above inclusion can be chosen depending solely on $d$, $T$.

4.3 Basic example of well-predicted signal: quasi-stable exponential polynomial

Let us define a quasi-stable exponential polynomial $(s_\tau)$ on $\mathbb{Z}^d$ as an exponential polynomial
\[
s_\tau = \sum_{\ell=1}^M c_\ell T^{a(\ell)} \exp\{\omega^T(\ell)\tau\}
\]
where all partial frequencies $\omega_j(\ell)$ satisfy the restriction $\Re(\omega_j(\ell)) \leq 0$. For example, an algebraic polynomial (partial frequencies are zero) and a trigonometric polynomial (partial frequencies are imaginary) are quasi-stable.

Proposition 14 Let $(s_\tau)$ be a quasi-stable exponential polynomial on $\mathbb{Z}^d$ of partial sizes $N_1,...,N_d$. Then for every integer $\kappa \geq 0$ and all $t \in \mathbb{Z}^d$ one has
\[
(s_\tau) \in P^t_{\kappa,T_0,\infty}(0,\rho_{\kappa,d}(N_1,...,N_d)),
\]
\[
\rho_{\kappa,d}(N_1,...,N_d) = \prod_{j=1}^d \left( (2N_j - 1)^{1/2} 2^{N_j} (\text{max}[2,2\kappa + 1])^{N_j/2} \right),
\]
\[
T_0 = \kappa \max_{1 \leq j \leq d} N_j
\]
and the filters $q^{(T)}$ certifying this inclusion can be chosen to be depending solely on $T, \kappa$ and on the collection of $d$ sets $O_j = \{ \omega_j(\ell) : 1 \leq \ell \leq M \}$ of partial frequencies.
5 Appendix

5.1 Preliminaries

Norm relations. Let us list several evident relations between the introduced semi-norms on $C(\mathbb{Z}^d)$.

- [Parseval equality]:
  \[
  (r, s)_T = \sum_{t:|t| \leq T} r_t \overline{s_t} = \sum_{\mu \in \Gamma_T^d} (F_Tr)(\mu)(F Ts)(\mu) \equiv \langle F_Tr, F Ts \rangle_T, \tag{25}
  \]
  where $\overline{a}$ is the complex conjugate of $a \in \mathbb{C}$; in particular,
  \[
  |r|_{T,2} = |r|^*_T; \tag{26}
  \]

A useful corollary of Parseval’s equality combined with the fact that $|q|^*_T,p = |\overline{q}|^*_T,p$ is the relation
  \[
  \left| \sum_{|t| \leq T} a_t b_t \right| \leq |a|^*_T,1 |b|^*_T,\infty. \tag{27}
  \]

- [Norms of convolutions of filters]
  \[
  r, s \in C(\mathbb{Z}^d) \Rightarrow |r(z_1, ..., z_d)s(z_1, ..., z_d)|_p \leq |r|_1 |s|_p; \tag{28}
  \]

- [Relations between $|\cdot|$ and $|\cdot|^*$]: for $p, q \in [1, \infty]$ one has
  \[
  |r|^*_T,p \leq (2T + 1)^d((1/p-1/2)_+ + (1/2-1/q)_+) |r|_{T,q}, \quad a_+ = \max[a, 0]; \tag{29}
  \]
  \[
  \text{ord}(r) + \text{ord}(s) \leq T \Rightarrow |r(z_1, ..., z_d)s(z_1, ..., z_d)|^*_T,p \leq |r|_1 |s|^*_T,p. \tag{30}
  \]

Useful fact. In the sequel, we need the following simple and well-known fact:

**Lemma 15** Let $f_j = \xi_j + i\eta_j$, $0 \leq j < N$, be a sequence of $N$ standard Gaussian complex-valued random variables, not necessarily independent of each other. Then
  \[
  \left[ E\{ \max_{0 \leq j < N} |f_j|^2 \} \right]^{1/2} \leq \sqrt{2 \ln N + 2};
  \]
  \[
  P\{ \max_{0 \leq j < N} |f_j| > u + \sqrt{2 \ln N} \} \leq \exp\{-u^2/2\} \quad \forall u \geq 0. \tag{31}
  \]
Proof. We have

\[ \psi(r) \equiv P\{ \max_{0 \leq j < N} |f_j| > r \} \leq \min[1, N \exp\{-r^2/2\}] \Rightarrow \\
\exp\{-(u + \sqrt{2 \ln N})^2/2\} \leq \exp\{-u^2/2\}; \]

\[ E\{ \max_{0 \leq j < N} |f_j|^2 \} = -\int_0^\infty r^2 d\psi(r) = 2 \int_0^\infty r\psi(r) dr \leq 2 \sqrt{2 \ln N} \int_0^\infty r dr = 2 \ln N + 2. \quad \Box \]

5.2 Proof of Theorem 4

W.l.o.g., we may assume that \( t = 0 \). We denote by \( q^* \) the filter associated with \((s_\tau)\) via the description of the inclusion \((s_\tau) \in S_{2T}(\theta, \rho, T)\). Let us set

\[ |q^*|_2^2 = \hat{\rho}^2(2T + 1)^{-d/2}; \quad \kappa = \theta(2T + 1)^{-d/2} \quad [\hat{\rho} \leq \rho], \tag{32} \]

so that

\[ \bar{s} = q^*(\Delta)s \Rightarrow \max_{r:|r| \leq 2T} E\{|s_r - \bar{s}_r|^2\} \leq \kappa^2. \tag{33} \]

Finally, let

\[ \Theta_T = \max_{r:|r| \leq 2T} |\Delta^T_1 \ldots \Delta^T_d e|_{2T, \infty}, \tag{34} \]

and let \( \tilde{\phi} \) be the optimal solution, used in Algorithm A, of the optimization problem (6).

1°. We start with simple technical lemma:

**Lemma 16** Let \( r(z_1, \ldots, z_d) = (q^*(z_1, \ldots, z_d))^2 \). Then \( r \in C_{2T}(\mathbb{Z}^d) \) possesses the following properties:

\[ |r|_2 \leq |r|_{2T, 1} \leq 2^{d/2} \hat{\rho}^2(2T + 1)^{-d/2}; \tag{35} \]

\[ |r|_1 \leq \rho^2; \tag{36} \]

\[ \left[ E\{(1 - r(\Delta))s|_{2T, 2}^2\} \right]^{1/2} \leq \kappa(\hat{\rho} + 1)(4T + 1)^{d/2}; \tag{37} \]

\[ |(1 - r(\Delta)) y|_{2T, \infty} \leq |(1 - r(\Delta)) s|_{2T, 2} + (1 + \hat{\rho}^2)\Theta_T \tag{38} \]

\[ \left[ E\{((1 - r(\Delta)) y|_{2T, \infty}^2) \} \right]^{1/2} \leq \sigma(1 + \hat{\rho}^2)\sqrt{4 \ln(4T + 1) + 2} + \kappa(\hat{\rho} + 1)(4T + 1)^{d/2}. \tag{39} \]
Proof. (35): We have

\[ |r|_{2T,1}^2 = \sum_{\mu \in T} |r|_{\mu}^2 = \sum_{\mu \in T} |q^*(\mu)|^2 = (4T + 1)^{d/2} \sum_{\mu \in T} |q^*(\mu)|^2 \]

\[ = (4T + 1)^{d/2} (|q^*|_{2T,2})^2 = (4T + 1)^{d/2} (|q^*|_{2T,2})^2 \leq 2^{d/2} \hat{\rho}^2 (2T + 1)^{-d/2}. \]

Since \(|r|_2 = |r|_{2T,2} = |r|_{2T,1}^2 \leq |r|_{2T,1}^2 \), (35) follows.

(36): We clearly have \(|q^*|^2 \leq ((2T + 1)^{d/2} |q^*|^2 \leq \hat{\rho}^2 \).

(37): Let \( h = (1 - q^*(\Delta))s \), so that by virtue of \((s) \in S^0(\theta, \rho, T)\) and in view of the origin of \( q^* \) we have

\[ \max_{\tau, |r| \leq 3T} E\{|h_r|^2\} \leq \kappa^2. \] (41)

Setting \( g = (1 - r(\Delta))s \), we have

\[ g_r = ((1 + q^*(\Delta))(1 - q^*(\Delta))s)_r = (1 + q^*(\Delta))h_r = h_r + (q^*(\Delta)h)_r \]

\[ \Rightarrow |g_r| \leq |h_r| + |q^*|_{[\Delta_1^{-\tau_1} ... \Delta_d^{-\tau_d} h]}_{T,2} \]

\[ \Rightarrow (E\{|g_r|^2]\})^{1/2} \leq (E\{|h_r|^2\})^{1/2} + |q^*|_{2} \left( \sum_{\tau, |r| \leq 3T} E\{|h_{r-r}|^2\} \right)^{1/2} ; \]

applying (41) and taking into account that \(|q^*|^2 \leq \hat{\rho}^2 (2T + 1)^{-d/2} \), we come to

\[ \max_{\tau, |r| \leq 3T} E\{|((1 - r(\Delta))s)_r|^2\} \leq [\kappa (\hat{\rho} + 1)]^2 , \] (42)

and (37) follows.

(38), (39): We have

\[ |(1 - r(\Delta))y|_{2T,\infty} \leq |(1 - r(\Delta))s|_{2T,\infty} + |(1 - r(\Delta))e|_{2T,\infty} \]

\[ \leq |(1 - r(\Delta))s|_{2T,2} + |e|_{2T,\infty} \leq |(1 - r(\Delta))s|_{2T,2} + |(1 - r(\Delta))e|_{2T,\infty} \]

\[ \leq |(1 - r(\Delta))s|_{2T,2} + |e|_{2T,\infty} + \sum_{\tau, |r| \leq 2T} |r||\Delta_1^{\tau_1} ... \Delta_d^{\tau_d} e|_{2T,\infty} \]

\[ \leq |(1 - r(\Delta))s|_{2T,2} + (1 + |r|_1) \max_{\tau, |r| \leq 2T} |\Delta_1^{\tau_1} ... \Delta_d^{\tau_d} e|_{2T,\infty}. \]
The resulting inequality combines with (36) to yield (38). Further, from the resulting inequality and (37) it follows that

\[
E \left\{ \left( (1 - r(\Delta)) y_{|\tau| \leq 2T}^* \right)^2 \right\}^{1/2} 
\leq \kappa (\hat{\nu} + 1) (4T + 1)^{d/2} + (1 + |r|_1) \left( E \left\{ \max_{\tau: |\tau| \leq 2T} |\Delta_{1\tau^1} \ldots \Delta_{d\tau^d} e|_{2T,\infty}^* \right\} \right)^{1/2} 
\leq \kappa (\hat{\nu} + 1) (4T + 1)^{d/2} + (1 + \hat{\nu}^2) \left( E \{ \Theta_T^2 \} \right)^{1/2}
\]

(we have used (36)). To derive (39) from the resulting inequality, it remains to note that

\[
\left( E \{ \Theta_T^2 \} \right)^{1/2} \leq \sigma \sqrt{4d \ln(4T + 1) + 2}.
\] (43)

Indeed, the coordinates of the Fourier transform of \( \Delta_{1\tau^1} \ldots \Delta_{d\tau^d} e \) are, up to factor \( \sigma \), standard complex-valued Gaussian random variables, so that \( \sigma^{-2}\Theta_T^2 \) is the maximum of squared modulæ of \( (4T + 1)^{2d} \) of these variables; therefore \( E \{ \Theta_T^2 \} \leq \sigma^2 (4d \ln(4T + 1) + 2) \) by Lemma 15. \( \square \)

2. We now study the properties of the solution \( \hat{\phi} \) of problem (6).

**Lemma 17** One has

\[
\begin{align*}
|\hat{\phi}|_{2T, 2} &\leq 2^{d/2} \rho^2 (2T + 1)^{-d/2}; \\
|(1 - \hat{\phi}(\Delta))e|_{2T, \infty}^* &\leq (1 + 2^d \rho^2) \Theta_T; \\
E \left\{ \left( |(1 - \hat{\phi}(\Delta))e|_{2T, \infty}^* \right)^2 \right\}^{1/2} &\leq \sigma (1 + 2^d \rho^2) \sqrt{4d \ln(4T + 1) + 2}; \\
|\hat{\phi}(\Delta)|s_{2T, \infty}^* &\leq |(1 - r(\Delta))s|_{2T, 2} + 2(1 + 2^d \rho^2) \Theta_T; \\
E \left\{ \left( |(1 - \hat{\phi}(\Delta))s|_{2T, \infty}^* \right)^2 \right\}^{1/2} &\leq 2\sigma (1 + 2^d \rho^2) \sqrt{4d \ln(4T + 1) + 2} + \kappa (\hat{\nu} + 1) (4T + 1)^{d/2}.
\end{align*}
\] (44), (45), (46)

**Proof.** (44): \( |\hat{\phi}|_{2T, 2} = |\hat{\phi}|_{2T, 2}^* \leq |\hat{\phi}|_{2T, 1} \leq 2^{d/2} \rho^2 (2T + 1)^{-d/2} \), (the concluding inequality comes from the fact that \( \hat{\phi} \) is feasible for (6)).

(45), (46): We have

\[
\begin{align*}
|\hat{\phi}(\Delta)|e_{2T, \infty}^* &\leq (1 + \hat{\phi}|_{2T, 1}) \max_{\tau: |\tau| \leq 2T} |\Delta_{1\tau^1} \ldots \Delta_{d\tau^d} e|_{2T, \infty}^* \\
&\leq \left( 1 + (4T + 1)^{d/2} |\hat{\phi}|_{2T, 2} \right) \max_{\tau: |\tau| \leq 2T} |\Delta_{1\tau^1} \ldots \Delta_{d\tau^d} e|_{2T, \infty}^* \\
&\leq (1 + 2^d \rho^2) \max_{\tau: |\tau| \leq 2T} |\Delta_{1\tau^1} \ldots \Delta_{d\tau^d} e|_{2T, \infty}^* 
\end{align*}
\]
(we have used (44)). The resulting inequality implies that

\[
\left[ E \left\{ (\left| (1 - \hat{\phi}(\Delta))e \right|_{2T,\infty}^2) \right\} \right]^{1/2} \leq (1 + 2d\rho^2) \left[ E \left\{ \max_{|r| \leq 2T} \left( \left| \Delta_1^T \cdots \Delta_d^T e \right|_{2T,\infty}^2 \right) \right\} \right]^{1/2} \\
\leq (1 + 2d\rho^2)\sigma \sqrt{4d \ln(4T + 1) + 2}
\]

(we have used (43)).

(48), (48): Note that the polynomial \( r \) defined in Lemma 16 is a feasible solution of the optimization problem (6) by the first relation in (35), so that the optimal value in the problem does not exceed \( J(r, y_{4T}^0) \). It follows that

\[
(a) \ J(\hat{\phi}, y_{4T}^0) \leq J(r, y_{4T}^0) \\
\Rightarrow (b) \ |(1 - \hat{\phi}(\Delta))y|_{2T,\infty} \leq |(1 - r(\Delta))y|_{2T,\infty} \\
\Rightarrow (c) \ |(1 - \hat{\phi}(\Delta))s|_{2T,\infty} \leq |(1 - \hat{\phi}(\Delta))e|_{2T,\infty} + |(1 - r(\Delta))y|_{2T,\infty} \\
\Rightarrow (d) \ E \left\{ \left( |(1 - \hat{\phi}(\Delta))s|_{2T,\infty}^2 \right)^{1/2} \right\} \leq E \left\{ \left( |(1 - \hat{\phi}(\Delta))e|_{2T,\infty}^2 \right)^{1/2} \right\} \\
+ E \left\{ \left( |(1 - r(\Delta))y|_{2T,\infty}^2 \right)^{1/2} \right\}
\]

Relation (47) follows from (c) combined with (38) and (45) (recall that \( \hat{\rho} \leq \rho \)). Relation (48) follows from (d) combined with (46) and (39). \( \square \)

Our next step is to prove

**Lemma 18** One has

\[
\left| \left( (1 - r(\Delta))(1 - \hat{\phi}(\Delta))s \right)_0 \right| \leq \left| (1 - r(\Delta))s \right|_0 \\
+ 2^{d/2}\rho^2(2T + 1)^{-d/2} \left| (1 - r(\Delta))s \right|_{2T,2}; \quad (49)
\]

\[
E \left\{ \left( \left| (1 - r(\Delta))(1 - \hat{\phi}(\Delta))s \right|_0 \right)^2 \right\}^{1/2} \leq \kappa(\hat{\rho} + 1)(2^d\rho^2 + 1). \quad (50)
\]

**Proof.** We have

\[
\left| \left( (1 - r(\Delta))(1 - \hat{\phi}(\Delta))s \right)_0 \right| \leq \left| (1 - r(\Delta))s \right|_0 + \left| \hat{\phi}(\Delta)(1 - r(\Delta))s \right|_0 \\
\leq \left| (1 - r(\Delta))s \right|_0 + |\hat{\phi}|_{2T,2} \left| (1 - r(\Delta))s \right|_{2T,2} \\
\leq \left| (1 - r(\Delta))s \right|_0 + 2^{d/2}\rho^2(2T + 1)^{-d/2} \left| (1 - r(\Delta))s \right|_{2T,2} \quad \text{[see (44)]}
\]

\[20\]
as required in (49). From the resulting inequality it follows that
\[
E \left\{ \left| ((1 - r(\Delta))(1 - \hat{\phi}(\Delta))s) \right|^2 \right\}^{1/2} \leq E \left\{ \left| ((1 - r(\Delta))s) \right|^2 \right\}^{1/2} + 2d/\rho^2(2T + 1)^{-d/2} \left[ E \left\{ \left| (1 - r(\Delta))s \right|^2 \right\} \right]^{1/2}
\]
\[
\leq \kappa(\hat{\rho} + 1) + 2d/\rho^2(2T + 1)^{-d/2} \left[ E \left\{ \left| (1 - r(\Delta))s \right|^2 \right\} \right]^{1/2}
\]
\[
\leq \kappa(\hat{\rho} + 1) + 2d/\rho^2(2T + 1)^{-d/2}\kappa(\hat{\rho} + 1)(4T + 1)^{d/2}
\]
and (50) follows. \(\square\)

4°. Now we are able to complete the proof of Theorem 4. The error of the estimate \(\hat{s}\) at the point \(t = 0\) is
\[
s_0 - \hat{s}_0 = s_0 - (\hat{\phi}(\Delta)y)_0 = \left( (1 - \hat{\phi}(\Delta))s \right)_0 - (\hat{\phi}(\Delta)e)_0 \equiv \epsilon_0^{(1)} + \epsilon_0^{(2)},
\]
\[
\epsilon_0^{(1)} = \left( (1 - \hat{\phi}(\Delta))s \right)_0, \quad \epsilon_0^{(2)} = (\hat{\phi}(\Delta)e)_0.
\]
Setting \(f_\tau = e_\tau\), we have
\[
|\epsilon_0^{(2)}| = \left| \sum_{|\tau|+|\hat{\phi}_\tau| \leq 2T} \hat{\phi}_\tau e_\tau \right| \leq |\hat{\phi}|^*_{2T,1} |f|_2^{*}_{T,\infty}
\]
[see (27)]
\[
\leq 2d/\rho^2(2T + 1)^{-d/2} |f|_2^{*}_{T,\infty},
\]
since \(\hat{\phi}\) is feasible for (6)]
whence, by definition of \(\Theta_T\),
\[
|\epsilon_0^{(2)}| \leq 2d/\rho^2(2T + 1)^{-d/2}\Theta_T.
\]
(52)

Applying (43), we derive from the latter inequality that
\[
E \left\{ |\epsilon_0^{(2)}|^2 \right\}^{1/2} \leq 2d/\sigma \rho^2(2T + 1)^{-d/2}\sqrt{2d\ln(4T + 1) + 2}.
\]
(53)

We further have
\[
|\epsilon_0^{(1)}| = \left| \left( (1 - \hat{\phi}(\Delta))s \right)_0 \right|
\]
\[
\leq \left| (r(\Delta)(1 - \hat{\phi}(\Delta))s) \right| + \left| (1 - r(\Delta))(1 - \hat{\phi}(\Delta))s \right|
\]
\[
\leq \left| r^*_{2T,1} (1 - \hat{\phi}(\Delta))s \right| + \left| (1 - r(\Delta))(1 - \hat{\phi}(\Delta))s \right|
\]
\[
\leq 2d/\rho^2(2T + 1)^{-d/2} |(1 - \hat{\phi}(\Delta))s|^{*}_{2T,\infty}
\]
\[
+ \left| (1 - r(\Delta))(1 - \hat{\phi}(\Delta))s \right|
\]
(54)
(the inequality \( a \) is given by (27), and \( b \) follows from the feasibility of \( \hat{\phi} \) for (6)), whence

\[
\left[ E \left\{ |c_0(1)|^2 \right\} \right]^{1/2} \leq 2^{d/2} \rho^2 (2T + 1)^{-d/2} \left[ E \left\{ \left( |(1 - \hat{\phi}(\Delta)) s|_{2T,\infty}^* \right)^2 \right\} \right]^{1/2} \\
+ \left[ E \left\{ \left( |(1 - r(\Delta))(1 - \hat{\phi}(\Delta)) s\right)_0^2 \right\} \right]^{1/2} \leq 2^{d/2} \rho^2 (2T + 1)^{-d/2} \left[ 2^d \sigma(1 + 2^d \rho^2) \sqrt{4d \ln(4T + 1) + 2} + \kappa(\hat{\rho} + 1)(4T + 1)^{d/2} \right] + \kappa(\hat{\rho} + 1)(2^d \rho^2 + 1)
\]

(55)

(see (48), (50)). Combining (51), (53), (55), we finally get

\[
\left[ E \left\{ |s_0 - \hat{s}_0|^2 \right\} \right]^{1/2} \leq 2^{d/2} \sigma^2 (2T + 1)^{-d/2} \sqrt{2d \ln(4T + 1) + 2} + 2^{d/2} \rho^2 (2T + 1)^{-d/2} \left[ 2^d \sigma(1 + 2^d \rho^2) \sqrt{4d \ln(4T + 1) + 2} + \kappa(\hat{\rho} + 1)(4T + 1)^{d/2} \right] + \kappa(\hat{\rho} + 1)(2^d \rho^2 + 1)
\]

(56)

Recalling that \( \hat{\rho} \leq \rho \), \( \kappa = \theta(2T + 1)^{-d/2} \) and that \( \rho \geq 1 \), (7) follows.

Now assume that \((s)\) is deterministic. In this case, from (54) combined with (47) and (49) implies that

\[
|\epsilon_0(1)| \leq 2^{1+d/2} \rho^2 (2T + 1)^{-d/2} |(1 - r(\Delta)) s|_{2T,2} + 2^{1+d/2} \rho^2 (1 + 2^d \rho^2)(2T + 1)^{-d/2} \Theta_T + \left\{ |(1 - r(\Delta)) s\right)_0^1 \right|,
\]

(57)

while from (37), (42) it follows that

\[
|(1 - r(\Delta)) s|_{2T,2} \leq \kappa(\hat{\rho} + 1)(4T + 1)^{d/2} \leq 2^{d/2} \theta(1 + \rho), \]
\[
|((1 - r(\Delta)) s)_0| \leq \kappa(1 + \hat{\rho}) \leq \theta(1 + \rho)(2T + 1)^{-d/2}.
\]

(58)

Therefore (57) implies that

\[
|\epsilon_0(1)| \leq 3^{3+d} \rho^3 [\theta + \rho \Theta_T] (2T + 1)^{-d/2}.
\]

(59)

Combining this relation with (52) and (51), we arrive at (8). \( \square \)
5.2.1 Proof of Proposition 6

In the proofs to follow, we focus on the case of well-filtered signals; the reasoning in the case of well-predicted signals is completely similar.

The case of $m = 1$ is evident. Now let $m \geq 2$, let $T^+$ be an integer, $0 \leq T^+ \leq L^+$, and let $T = [m^{-1}T^+]$. Since $s^j \in \mathbf{F}_L^j(\theta, \rho)$ and clearly $T \leq L$, there exist filters $q^j$ such that

\[ (a) : \text{ord}(q^j) \leq T; \quad (b) : |q^j|^2 \leq \rho_j(2T + 1)^{-d/2}; \]
\[ (c) : |q^j|_1 = |q^j|_{T,1} \leq (2T + 1)^{d/2}|q^j|^2 \leq \rho_j; \]
\[ (d) : \left[E \{ |s^j_\tau - (q^j(\Delta)s^j)_\tau|^2 \right]^{1/2} \leq \theta_j(2T + 1)^{-d/2} \quad \forall (\tau : |\tau - t| \leq L). \]

Now let filter $q$ be defined by

\[ 1 - q(z) = \prod_{j=1}^{m} (1 - q^j(z)), \quad z = (z_1, ..., z_d). \]

Observe that

\[ \text{ord}(q) \leq mT \leq T^+. \quad (61) \]

Note that

\[ |q|_2 \leq 2^m \rho_1...\rho_m(2T + 1)^{-d/2} \leq (2m - 1)^{d/2}2^m \rho_1...\rho_m(2T^+ + 1)^{-d/2}. \quad (62) \]

Indeed, we clearly have

\[
|q(z)|_2 = \left| \sum_{\ell=1}^{m} (-1)^{\ell+1} \sum_{1 \leq j_1 < j_2 < ... < j_{\ell} \leq m} q^{j_1}(z)q^{j_2}(z)...q^{j_{\ell}}(z) \right|_2
\leq \sum_{\ell=1}^{m} \sum_{1 \leq j_1 < j_2 < ... < j_{\ell} \leq m} \left| q^{j_1}(z)q^{j_2}(z)...q^{j_{\ell}}(z) \right|_2
\leq \sum_{\ell=1}^{m} \sum_{1 \leq j_1 < j_2 < ... < j_{\ell} \leq m} \frac{\rho_{j_1}\rho_{j_2}...\rho_{j_{\ell}}}{(2T + 1)^{\ell d/2}}
\leq [(1 + \rho_1)...(1 + \rho_m) - 1](2T + 1)^{-d/2}
\leq 2^m \rho_1...\rho_m(2T + 1)^{-d/2}
\]

(a is by (60.b - c) since $|u(z)v(z)| \leq |u|_1|v|_2$, $|u(z)v(z)| \leq |u|_1|v|_1$, $b$ is due to $\rho_j \geq 1$, as required in (62). Further, by (60.c), for the filters

\[ Q^j(z) = \left( \prod_{\ell=1}^{j-1} (1 - q^\ell(z)) \right) \left( \prod_{\ell=j+1}^{m} (1 - q^\ell(z)) \right) \]
one has
\[ |Q^j|_1 \leq (1 + \rho_1)...(1 + \rho_{j-1})(1 + \rho_{j+1})...(1 + \rho_m) \leq \frac{2^{m-1}\rho_1...\rho_m}{\rho_j}. \] (63)

Now let \( \tau \in \mathbb{Z}^d \) be such that \( |\tau - t| \leq L^+ \). We have

\[
\begin{align*}
&E \left\{ |(1-q(\Delta))s|_{\tau}^2 \right\}^{1/2} = \left[ E \left\{ \sum_{j=1}^{m} \lambda_j \left( 1 - q(\Delta) \right) s_j \right\}_\tau^2 \right]^{1/2} \\
&\leq \sum_{j=1}^{m} \left[ E \left\{ |\lambda_j (1-q(\Delta))s_j|_{\tau}^2 \right\} \right]^{1/2} \\
&= \sum_{j=1}^{m} \left[ E \left\{ |\lambda_j| |1-q(\Delta)| |s_j|_{\tau}^2 \right\} \right]^{1/2} \\
&\leq \max_{\tau'} \sum_{j=1}^{m} \left[ E \left\{ |\lambda_j| |1-q^j(\Delta)| |s_j|_{\tau'}^2 \right\} \right]^{1/2} \\
&\leq 2^{m-1}\rho_1...\rho_m(2T + 1)^{-d/2} \sum_{j=1}^{m} \frac{\theta_j [E\{\lambda_j\}^2]}{\rho_j} \\
&\leq \left( \frac{2m-1}d \cdot 2^{m-1}\rho_1...\rho_m \sum_{j=1}^{m} \frac{\theta_j [E\{\lambda_j\}^2]}{\rho_j} \right) (2T^+ + 1)^{-d/2}
\end{align*}
\]

where \( a \) is due to independence of \( \lambda_j \) and \( (s^j) \) and \( b \) follows from (63), (60.d), and since
\[ |\tau'-\tau| \leq (m-1)T, |\tau-t| \leq L^+ \Rightarrow |\tau'-t| \leq L^+ + T^+ \leq L. \]

Combining the resulting inequality, (61), (62) and taking into account that \( T^+ \in \{0, 1, ..., L^+ \} \) is arbitrary, we conclude that \( s \in F^t_{L^+}(\theta^+, \rho^+) \). Note that by construction, the filters certifying the latter inclusion are independent of \( \lambda_j \). \( \square \)

5.2.2 Proof of Proposition 7

(i): Let \( T \leq L \), and let \( q \) be such that
\[
\begin{align*}
\text{ord}(q) &\leq T, |q|_2 \leq \frac{\rho}{(2T+1)^{d/2}}, \\
\max_{\tau:|\tau-t| \leq L} \left[ E \left\{ |(1-q(\Delta))s|_{\tau}^2 \right\} \right]^{1/2} &\leq \frac{\rho}{(2T+1)^{d/2}}.
\end{align*}
\] (64)
Let us set \( \bar{q}_\tau = \exp\{i\omega^T\tau\} q_\tau \), \( \tau \in \mathbb{Z}^d \). Then \( \text{ord}(\bar{q}) \leq T \), \( |\bar{q}|_2 = |q|_2 \) and

\[
((1 - \bar{q}(\Delta))\bar{s})_\tau = \exp\{i[\omega^T\tau + \phi]\} s_\tau
- \sum_{\tau'}(\exp\{i\omega^T\tau\prime\}\{q_\tau\prime\}(\exp\{i[\omega^T(\tau - \tau') + \phi]\} s_{\tau - \tau'})
= \exp\{i[\omega^T\tau + \phi]\} ((1 - q(\Delta))s)_\tau ,
\]
so that (64) remains valid when \( q, (s) \) are replaced with \( \bar{q}, (\bar{s}) \). Thus, \( (\bar{s}) \in \mathbf{F}_L^*(\theta, \rho) \). (i) is proved; (ii) is evident. \( \square \)

5.2.3 Proof of Proposition 8

Let \( T \leq L \), and let \( q = (q_\tau)_{\tau \in \mathbb{Z}^d} \) be such that \( \text{ord}(q) \leq T \), \(|q|_2 \leq \rho(2T + 1)^{-d/2} \),

\[
\left[ E \left\{ |((1 - q(\Delta))s)_\tau|^2 \right\} \right]^{1/2} \leq \theta(2T + 1)^{-d/2} \quad \forall (\tau \in \mathbb{Z}^d : |\tau - t| \leq L).
\]

Setting \( q_{\tau_1,...,\tau_d}^+ = (2T + 1)^{-d/2} q_{\tau_1,...,\tau_d} \), we clearly have \( \text{ord}(q^+) \leq T \), \(|q^+|_2 \leq \rho(2T + 1)^{-d/2} \) and

\[
\left[ E \left\{ |((1 - q^+(\Delta))s^+_\tau)|^2 \right\} \right]^{1/2} \leq \theta(2T + 1)^{-d/2} \quad \forall (\tau \in \mathbb{Z}^d^+ : |\tau - t^+| \leq L).
\]

It remains to note that \( \theta(2T + 1)^{-d/2} \leq \theta^+(2T + 1)^{-d/2} \) for \( 0 \leq T \leq L \). \( \square \)

5.2.4 Proof of Proposition 9

Let \( T \leq L \), and let \( q' \in C_T(\mathbb{Z}^{d'}) \), \( q'' \in C_T(\mathbb{Z}^{d''}) \) be such that

\[
(a) : |q'|_2 \leq \rho'(2T + 1)^{-d'/2} , |q''|_2 \leq \rho''(2T + 1)^{-d''/2} , \quad s'_{\tau - \tau'}(\xi) = \sum_{\tau'} s'_{\tau - \tau'} q'_{\tau'} , |\tau' - t'| \leq L, \quad \text{(65)}
\]
\[
s''_{\tau - \tau}''(\xi) = \sum_{\tau'} s''_{\tau - \tau'} q''_{\tau'} , |\tau'' - t''| \leq L
\]

Let \( q(z_1,...,z_d) = q'(z_1,...,z_{d'}) q''(z_{d'+1},...,z_d) \), so that

\[
q \in C_T(\mathbb{Z}^d), \quad |q|_2 = |q'|_2 |q''|_2 \leq \rho' \rho''(2T + 1)^{-d/2} \quad \text{(66)}
\]

(see (65.a)). Now let \( \tau = (\tau', \tau'') \) be such that \(|\tau - (t', t'')| \leq L \). We have

\[
(q(\Delta)s(\xi))_\tau = \sum_{(\nu', \nu'') \in \mathbb{Z}^{d'} \times \mathbb{Z}^{d''}} s'_{\tau' - \nu'}(\xi) s''_{\tau'' - \nu''}(\xi) q'_{\nu'} q''_{\nu''} = \sum_{\nu' \in \mathbb{Z}^{d'}} s'_{\tau' - \nu'} q'_{\nu'}, s''_{\tau''}(\xi)
\]
\[
= s'_{\tau'}(\xi) s''_{\tau''}(\xi) = s_{(\tau', \tau'')},
\]

25
which combines with (66) to yield that \((s_\tau) \in F_L^{(\ell_t, c_n)}(0, \rho \rho'')\). □

5.2.5 Proof of Proposition 10

We start with the following two evident facts:

Lemma 19 Let \((s^j) \subset C(Z^d)\) be deterministic fields belonging to \(F^j_L(\theta, \rho)\), \(j = 1, 2, \ldots\) such that \(s^j_t \to s_\tau\), \(j \to \infty\), for every \(\tau \in Z^d\). Then \((s) \in F^j_L(\theta, \rho)\).

Indeed, for every \(T\), \(0 \leq T \leq L\), the filters \(q^jT \subset C_T(Z^d)\) which certify the inclusions \((s^j) \subset F^j_L(\theta, \rho)\) satisfy \(|q^jT|_2 \leq \rho(2T + 1)^{-d/2}\) and therefore have a limiting point \(q^T \subset C_T(Z^d)\) with \(|q^j|_2 \leq \rho(2T + 1)^{-d/2}\). The filters \(\{q^T\}_{0 \leq T \leq L}\) clearly certify the inclusion \((s) \in F^j_L(\theta, \rho)\). □

Lemma 20 For every \(t \in Z\), the univariate exponential field \((s_\tau = \exp\{\omega \tau\})\), \(\omega \in C\), belongs to \(F^\infty_\omega(0, \sqrt{2})\).

Indeed, assuming \(\Re(\omega) \geq 0\) and given \(T \geq 0\), let us set \(q(z) = \frac{1}{T+1}[1 + \exp\{-\omega\}z^{-1} + \exp\{-2\omega\}z^{-2} + \ldots + \exp\{-T\omega\}z^{-T}]\). Then \(q \subset C_T(Z)\), \(|q|_2 = (T+1)^{-1/2} \leq 2^{1/2}(2T+1)^{-1/2}\), while clearly \(q(\Delta)s \equiv s\). In the case of \(\Re(\omega) < 0\), the same reasoning holds true for \(q(z) = \frac{1}{T+1}[1 + \exp\{\omega\}z + \exp\{2\omega\}z^2 + \ldots + \exp\{T\omega\}z^T]\). □

To complete the proof, we need the following fact:

Lemma 21 Let \((s_\tau)\) be a “simple” exponential polynomial – a deterministic exponential polynomial of the form \((s_\tau) = \sum_{\ell=1}^M c_{\ell} \exp\{\omega T(\ell)\}\). Then

\[
\forall t \in Z^d : (s_\tau) \subset F^t_\infty(0, \rho_\omega(N_1, \ldots, N_d)),
\]

where \(\rho_\omega(\cdot, \ldots, \cdot)\) is given by (16) and \(N_1, \ldots, N_d\) are the partial sizes of the polynomial. Besides this, the filters \(q^{(\ell)}\) certifying the above inclusion can be chosen to depend solely on \(T\) and on the collection of the \(d\) sets \(O_j = \{\omega_j(\ell) : \ell = 1, \ldots, M\}\).

Lemma 21 ⇒ Proposition 10: Assume first that the coefficients \(c_{\ell}\) in (15) are deterministic. Since every one of the univariate functions \(f(t) = tk, 0 \leq k \leq m\), is, uniformly on compact sets, the limit, as \(\epsilon \to +0\), of appropriate linear combinations of the \(m+1\) exponents \(\exp\{-k\epsilon\}\), the exponential polynomial (15) is the pointwise, on \(Z^d\), limit, as \(i \to \infty\), of simple exponential polynomials \((s^j_\tau)\) with extended sets of “frequencies” \(\{\omega_j(\ell)\}_{j, \ell}\); in the approximating polynomials, every one of these frequencies is replaced by \((m_j + 1)\) frequencies \(\omega_j(\ell) - k\epsilon_i\), \(0 \leq k \leq m_j\). Note that by the definition of partial sizes of exponential polynomials, the approximating polynomials have exactly the
same partial sizes as the original polynomial \((s_\tau)\). Combining Lemmas 21 and (19), we immediately conclude that the exponential polynomial (15) belongs to \(F^d_m(0, \rho_d(N_1, ..., N_d))\). Since the filters \(q^{(T)}\) certifying well-filterability of the approximating polynomials \((s_\tau')\) can be chosen to depend solely on \(T\) and the sets of partial frequencies of these approximating polynomials, from the proof of Lemma 19 it follows that the filters \(q^{(T)}\) certifying the inclusion \((s_\tau) \in F^d_m(0, \rho_d(N_1, ..., N_d))\) can be chosen to depend solely on \(T\) and the sets of partial frequencies of \((s_\tau)\), as required in Proposition 10. We have proved Proposition 10 for the case of a deterministic exponential polynomial; since the filters certifying well-filterability of such a polynomial are independent of the coefficients \(c_\ell\), the result is valid for random polynomials as well. □

**Proof of Lemma 21.** Proof is by induction in \(d\).

**Base** \(d = 1\) is readily given by Lemma 20 combined with Proposition 6.

**Step** \(1 \leq d \Rightarrow d + 1\): Let \(s_\tau = \sum c_\ell \exp\{\omega^T(\ell)\tau\}\) be a simple exponential polynomial on \(\mathbb{Z}^{d+1}\) with partial sizes \(N_j\) and the sets of partial frequencies \(O_j\), \(j = 1, ..., N\). Let \(T \geq 0\), and let \(t \in \mathbb{Z}^{d+1}\). By the inductive hypothesis, there exist filters \(g^{(T)} \in C_T(\mathbb{Z}^{d})\), \(h^{(T)} \in C_T(\mathbb{Z})\) (depending solely on \(T\) and on \(O_1, ..., O_{d+1}\)) such that

\[
\begin{align*}
(a) : |g^{(T)}|_2 &\leq \rho_d(N_1, ..., N_d)(2T + 1)^{-d/2}, \\
(a') : |h^{(T)}|_2 &\leq \rho_d(N_{d+1})(2T + 1)^{-1/2}, \\
(b) : r_\tau = \sum_{\nu \in \mathbb{Z}^d} r_{\tau - \nu}g^{(T)}_\nu \ \forall \tau \in \mathbb{Z}^d \ \forall (r_\tau) \in \mathcal{E}(O_1, ..., O_d), \\
(b') : p_\tau = \sum_{\nu \in \mathbb{Z}} p_{\tau - \nu}h^{(T)}_\nu \ \forall \tau \in \mathbb{Z} \ \forall (p_\tau) \in \mathcal{E}(O_{d+1}),
\end{align*}
\]

(68)

where \(\mathcal{E}(O^1, ..., O^m)\) is the space of all simple exponential polynomials on \(\mathbb{Z}^m\) with the sets of partial frequencies \(O^1, ..., O^m\). Setting \(q^{(T)}_\tau = g^{(T)}_\tau h^{(T)}_\tau\), \(\tau \in \mathbb{Z}^{d+1}\), we clearly have

\[
\begin{align*}
q^{(T)} \in C_T(\mathbb{Z}^{d+1}), \quad |q^{(T)}|_2 &\leq |g^{(T)}|_2|h^{(T)}|_2 \leq \rho_d(N_1, ..., N_d)\rho_d(N_{d+1}) \\
= \rho_{d+1}(N_1, ..., N_{d+1})
\end{align*}
\]

(69)

(see (68.a, a')). Further, for every \((s_\tau) \in \mathcal{E}(O_1, ..., O_{d+1})\) we have, setting \(\tau = (\tau', \tau'')\) with \(\tau' \in \mathbb{Z}^d\), \(\tau'' \in \mathbb{Z}\):

\[
\sum_{\nu \in \mathbb{Z}^{d+1}} q^{(T)}_\nu s_{\tau - \nu} = \sum_{\nu \in \mathbb{Z}^d} g^{(T)}_\nu \left( \sum_{\nu' \in \mathbb{Z}} h^{(T)}_{\nu'} s_{\tau' - \nu', \tau'' - \nu''} \right) = \sum_{\nu \in \mathbb{Z}^{d+1}} g^{(T)}_\nu s_{\tau' - \nu', \tau''}
\]

(see (68.a, a')). Further, for every \((s_\tau) \in \mathcal{E}(O_1, ..., O_{d+1})\) we have, setting \(\tau = (\tau', \tau'')\) with \(\tau' \in \mathbb{Z}^d\), \(\tau'' \in \mathbb{Z}\):
(a is by (68.1) since \((s_{\tau} - \nu, \mu)_{\mu \in \mathbb{Z}} \in \mathbf{E}(\mathbf{O}_{d+1})\), b is by (68.2) since \((s_{\mu, \tau, \nu})_{\mu \in \mathbb{Z}^d} \in \mathbf{E}(\mathbf{O}_1, ..., \mathbf{O}_d)\), which combines with (69) to imply that 
\[
(s_{\tau}) \in \mathbf{S}_\infty^t (0, \rho_{d+1}(N_1, ..., N_{d+1}), T).
\]
Thus, the filters \(q^{(T)}\) (which depend solely on \(T\) and \(\mathbf{O}_1, ..., \mathbf{O}_{d+1}\)) certify the inclusion \((s_{\tau}) \in \mathbf{L}_\infty^t (0, \rho_{d+1}(N_1, ..., N_{d+1})).\) The inductive step is completed. □

5.2.6 Proof of statement in Remark 11

It suffices to prove that for every nonnegative integer \(T\) and every \(m, d\) there exists a filter \(q^{(T)}\), \(\text{ord}(q^{(T)}) \leq T\), depending solely on \(T, m, d\), such that

\[
\begin{align*}
(a) &\quad q^{(T)}(\Delta)p = p \text{ for every polynomial (17)}, \\
(b) &\quad |q^{(T)}|_2 \leq \left(\frac{16m}{\sqrt{2T+1}}\right)^d \equiv \Theta^d.
\end{align*}
\]

This well-known fact can be proved by induction in \(d\) completely similar to the one used to prove Lemma 21; the only difference is in the Base, which now should be replaced with the following statement:

**Lemma 22** Let \(p(\tau) = \sum_{\ell=0}^{m} p_{\ell} \tau^\ell\) be a deterministic univariate algebraic polynomial of degree \(m\). Then for every \(T \geq 0\) there exists a filter \(q \in C_T(\mathbb{Z})\), depending solely on \(T, m\), with \(|q|_2 \leq 16m(2T+1)^{-1/2}\) such that \(p(t) = \sum_{\nu} q^{(T)}(p(t) - \nu)\) for all \(t \in \mathbb{Z}\).

**Proof.** By evident reasons, it suffices to prove that for a given \(T \geq 0\) there exists a collection of weights \(q_t, -T \leq t \leq T\), such that

\[
\sum_{t=-T}^{T} q_t = 1, \quad \sum_{t=-T}^{T} q_t t^i = 0, \quad i = 1, ..., m, \quad \sum_{t=-T}^{T} q_t^2 \leq \Theta^2 \equiv \frac{256m^2}{2T+1}.
\]

By the standard separation arguments, this is the same as to prove that for every real algebraic polynomial \(r(t)\) of degree \(\leq m\) such that \(r(0) = 1\) one has \(\sum_{t=-T}^{T} r^2(t) \geq \frac{2T+1}{256m^2}\), or, which is the same, that for the real trigonometric polynomial \(\rho(\phi) = r(T \sin(\phi))\) one has

\[
\sum_{t=-T}^{T} \rho^2(\phi_t) \geq \frac{2T+1}{256m^2}, \quad \phi_t = \text{asin}(t/T).
\]

(71)
Note that the degree of the trigonometric polynomial \( \rho(\cdot) \) is \( \leq m \) and that \( \rho(0) = 1 \). Besides this, \( \rho(\phi) = \rho(\pi - \phi) \); due to the latter fact,

\[
M \equiv \max_{\phi} |\rho(\phi)| = \max_{|\phi| \leq \frac{\pi}{2}} |\rho(\phi)| \geq |\phi(0)| = 1.
\]

By Bernstein’s Theorem on trigonometric polynomials, we have \( |\rho'(\phi)| \leq m M \).

Now let \( \phi \in (-\pi/2, \pi/2) \) be a point such that \( |\rho(\phi_\star)| = M \), let \( \Delta \) be the segment of the length \( \frac{1}{m} \) centered at \( \phi_\star \), and \( \Delta \) be the part of this segment in \( (-\pi/2, \pi/2) \). Note that the length of \( \Delta \) is at least \( \frac{1}{2m} \) and that for \( \phi \in \Delta \) one has \( |\rho(\phi)| \geq |\rho(\phi_\star)| - \frac{1}{2m} (mM) \geq M/2 \). Let \( n \) be the minimum number of points \( \phi_i \) belonging to a segment \( \delta \subset (-\pi/2, \pi/2) \) of the length \( 1/(2m) \), the minimum being taken over all positions of \( \delta \) in \( (-\pi/2, \pi/2) \). It is immediately seen that \( n \geq (1 - \sin(\pi/2 - 1/(2m)))T - 2 \geq \frac{T}{16m^2} - 2 \), whence

\[
\sum_{t=-T}^{T} \rho^2(\phi_t) \geq \sum_{t: \phi_t \in \Delta} \rho^2(\phi_t) \geq \frac{M^2}{4} n \geq \frac{M^2}{4} \left[ \frac{T}{16m^2} - 2 \right] \geq \frac{1}{4} \left[ \frac{T}{16m^2} - 2 \right].
\]

When \( T \geq 64m^2 \), the latter quantity is \( \geq \frac{2T+1}{256m^2} \), and in any case \( \sum_{t=-T}^{T} \rho^2(\phi_t) \geq \rho^2(\phi_0) = 1 \). Thus, we always have \( \sum_{t=-T}^{T} \rho^2(\phi_t) \geq \frac{2T+1}{256m^2} \), as required in (71). \( \square \)

5.2.7 Proof of Proposition 12

In the proof to follow, \( c_i \) stand for positive constants depending solely on \( D \).

\({}^0\) We start with the following evident observation:

**Lemma 23** There exists \( c_1 \) such that for every polynomial \( p(t) \) of one variable satisfying the relation \( p(1) = 1 \) one has

\[
M \leq c_1 N, \quad \deg(p) \leq c_1 N, \quad (s) \in H^1_N(D) \Rightarrow s_{\tau} = (p(D)s)_{\tau}, \quad \forall (\tau : |\tau - t| \leq M).
\]

\( ^0 \) Let us fix a positive integer \( N \), and let

\[
\delta(\omega) = \sum_{l=1}^{k} w_l \exp\{i\omega^T \alpha(\ell)\} : [-\pi, \pi]^d \to \mathbb{C},
\]

\[
\Omega_N^d = \left\{ \omega \in \mathbb{R}^d \mid \omega \in \left\{ \frac{2\pi j}{2N+1} \right\}_{|\omega| \leq N} : j = 1, ..., d \right\},
\]

and let \( \nu \) be the normalized counting measure on \( \Omega_N^d : \nu(\{\omega\}) = (2N+1)^{-d}, \omega \in \Omega_N^d \). Observe that in view of R.2 the function \( \delta(\cdot) \) maps \( \Omega_N^d \) into the unit disk \( D = \{ \zeta \in \mathbb{C} \mid |\zeta| \leq 1 \} \). Let \( \mu \) be the distribution of values of \( \delta(\Omega_N^d) \), so that \( \mu \) is

29
the measure supported by the finite set \( M = \{ \zeta \mid \exists \omega \in \Omega_N^d : \zeta = \delta(\omega) \} \), and
\[ \mu(\{\zeta\}) = \sum_{\omega \in \Omega_N^d : \delta(\omega) = \zeta} \nu(\{\omega\}) \]. Let also \( F(\alpha) = \mu(\{\zeta \mid \Re(\zeta) \geq 1 - \alpha\}) \), \( \alpha \geq 0 \).

**Lemma 24** There exists \( c_2 \in (0, 1) \) such that

\[
M \subset \tilde{M} = \left\{ \zeta \mid |\zeta| \leq 1, |\Im(\zeta)| \leq c_2^{-1}(1 - \Re(\zeta))^{3/2} \right\},
\]
\[ F(\alpha) \leq c_2^{-1}[\alpha^{d/2} + N^{-d}], \quad 0 \leq \alpha \leq 2. \tag{75} \]

**Proof.** (74), (75) are evident when \( \sum_{\ell=1}^{k} \rho_\ell < 1 \), since then \( |\delta(\omega)| \leq 1 - c_2 \) for properly chosen \( c_2 \) and all \( \omega \). Thus, in the sequel we focus on the case of \( \sum_{\ell=1}^{k} \rho_\ell = 1 \) (recall that \( \sum_{\ell=1}^{k} \rho_\ell \leq 1 \) by R.2).

2.1 Let \( K = \{ \omega \in [-\pi, \pi]^d : \delta(\omega) = 1 \} \). Since \( \rho_\ell > 0 \), \( \sum_{\ell=1}^{k} \rho_\ell = 1 \) and \( \delta(\omega) = \sum_{\ell=1}^{k} \rho_\ell \exp\{i\phi_\ell + \omega^T\alpha(\ell)\} \), a point \( \omega \in K \) must satisfy the equations

\[
\exp\{i[\phi_\ell + \omega^T\alpha(\ell)]\} = 1 \quad \forall (1 \leq \ell \leq k), \tag{76}
\]

whence \( \phi_\ell + \omega^T\alpha(\ell) \in 2\pi\mathbb{Z} \quad \forall (1 \leq \ell \leq k) \). Since \( \text{Rank}\{\alpha(\ell) : 1 \leq \ell \leq k\} = d \), the latter system of equations implies that \( K \) belongs to a set of the form \( r + A\mathbb{Z}^d \) with certain \( d \times d \) nonsingular matrix \( A \) (depending solely on \( D \)). The cardinality of the intersection of latter set with the cube \([-\pi, \pi]^d \) does not exceed certain \( c_3 \). Thus, \( \text{Card}K \leq c_3 \).

2.2 Let \( \omega \in K \), and let \( d\omega \in \mathbb{R}^n \) be such that \( |d\omega| \leq 1 \). Then

\[
\delta(\omega + d\omega) = \sum_{\ell=1}^{k} \rho_\ell \exp\{i[\phi_\ell + \omega^T\alpha(\ell)]\} \exp\{i(d\omega)^T\alpha(\ell)\}
\]
\[
\leq \sum_{\ell=1}^{k} \rho_\ell \sum_{\ell=1}^{k} \rho_\ell \exp\{i(d\omega)^T\alpha(\ell)\}
\]
\[
\leq \sum_{\ell=1}^{k} \rho_\ell \left(1 + \frac{1}{2} \left((d\omega)^T\alpha(\ell)\right)^2\right) \leq c_4|d\omega|^3
\]
\[
= \sum_{\ell=1}^{k} \rho_\ell \left(1 + \frac{1}{2} \left((d\omega)^T\alpha(\ell)\right)^2\right) \leq 1 - c_5|d\omega|^2 + c_4|d\omega|^3
\]

(for \( a \), see (76), \( b \) is by (20.b), \( c \) is due to \( \text{Rank}\{\alpha(\ell)\}_{\ell} = d \)). It follows that with properly chosen \( c_6 \) one has

\[
\forall (\omega \in [-\pi, \pi]^d, |\delta(\omega) - 1| \leq \alpha) \quad \exists \bar{\omega} \in K : |\omega - \bar{\omega}| \leq c_6^{-1}\sqrt{\alpha}. \tag{77}
\]

30
Since \( \text{Card}(\mathbb{K}) \leq c_3 \) by (20), and \(|\delta(\omega)| \leq 1\) for all \(\omega\), we conclude that

\[
\nu \left( \{\omega \in \Omega^d_\alpha : |\delta(\omega) - 1| \leq \alpha\} \right) \leq c_7 [\alpha^{d/2} + N^{-d}] \quad \forall \alpha \leq 2. \tag{78}
\]

(20.3) Now we can complete the proof of (74), (75). Let \(\bar{\omega} \in \mathbb{K}, d\omega \in \mathbb{R}^d, |d\omega| \leq 1\). We have

\[
\delta(\bar{\omega} + d\omega) = \sum_{l=1}^{k} \rho_l \exp\{i[\phi_l + \bar{\omega} T(\alpha)]\} \exp\{i(d\omega) T(\alpha)\}
\]

\[
= \sum_{a} \sum_{l=1}^{k} \rho_l \exp\{i(d\omega) T(\alpha)\} \left( 1 + i(d\omega) T(\alpha) + \frac{1}{2} \left( (d\omega) T(\alpha) \right)^2 - \frac{1}{6} \left( (d\omega) T(\alpha) \right)^3 + r_l(\omega, d\omega) \right), \tag{79}
\]

(for \(a\), see (76), for \(b\), see (20)). Taking into account that \(\sum \rho_l = 1\) and \(c_{11}|d\omega|^2 \leq \sum \rho_l \left( (d\omega) T(\alpha) \right)^2 \leq c_{12}|d\omega|^2\), we conclude from (77) combined with (79) that for properly chosen \(c_{13}\) one has

\[
\omega \in [-\pi, \pi]^d \Rightarrow |\Im(\delta(\omega))| \leq c_{13}(1 - \Re(\delta(\omega)))^{3/2},
\]

and (74) follows. By (74) one has \(|1 - \delta(\omega)| \leq c_{14}(1 - \Re(\delta(\omega)))\), so that (75) follows from (78). \(\square\)

31. Let \(n\) be a positive integer, and let \(T_n(\zeta)\) be the Tschebyshev polynomial of degree \(n\). Recall that this polynomial is defined as follows:

\[
T_n(\zeta) = \frac{w^n + w^{-n}}{2}, \quad \text{where} \ w = \zeta + i\sqrt{1 - \zeta^2}. \tag{80}
\]

In (80), the choice of the branch of \(\sqrt{\cdot}\) affects the value of \(w\), but does not affect the value of \(w^n + w^{-n}\); since we intend to work with \(\zeta\) from the unit disk, so that \(\Re(1 - \zeta^2) > 0\), in the calculations to follow we deal with the main branch of \(\sqrt{\cdot}\) in the closed right half-plane. On the segment \([-1, 1]\) of the real axis one has \(T_n(\zeta) = \cos(n \acos(\zeta))\), whence \(T_n(1) = 1, T'_n(1) = n^2\). From these relations it follows that the function \(P_n(\zeta) = \frac{1 - T_n(\zeta)}{n^2(1 - \zeta)}\) is a polynomial of degree \(n - 1\), and \(P_n(1) = 1\).
Lemma 25 One has

\[ p_n(\alpha) \equiv \max\{|P_n(\zeta)| : \zeta \in \hat{\mathcal{M}}, \Re(\zeta) = 1 - \alpha\} \leq q_n(\alpha) = \begin{cases} c_{15}, & 0 \leq \alpha \leq \frac{1}{n^2}, \\ \frac{c_{15}(1 + c_{15})^{n}}{n^2}, & \frac{1}{n^2} < \alpha \leq 2. \end{cases} \tag{81} \]

**Proof.** Let \( \zeta = 1 - \alpha + i\beta \in \hat{\mathcal{M}} \), so that

\[ |\beta| \leq c_{16}\alpha^{3/2}. \tag{82} \]

We have

\[
\begin{align*}
w &:= \zeta + i\sqrt{1 - \zeta^2} = 1 - \alpha + i\beta + i\sqrt{2\alpha - \alpha^2 - 2i(1 - \alpha)\beta + \beta^2} \\
&= 1 - \alpha + i\beta + i\sqrt{2\alpha - 1 - 0.5\alpha} + |0.5\beta - i(1 - \alpha)|(\beta/\alpha) \leq 1 + i\sqrt{2\alpha} + r_1(\zeta), \\
|r_1(\zeta)| &\leq c_{17}\alpha
\end{align*}
\]

(since \(|\beta/\alpha| \leq c_{16}\alpha\sqrt{\alpha}\) by (83)). Note that completely similar considerations demonstrate that

\[
w^{-1} = \zeta - i\sqrt{1 - \zeta^2} = 1 - i\sqrt{2\alpha} + r_2(\zeta), \quad |r_2(\zeta)| \leq c_{17}\alpha
\]

**3.1** Assume, first, that \( 0 \leq \alpha \leq \frac{1}{n^2} \). In this case from (83) it follows that

\[ |1 - w| \leq \sqrt{2n^{-1}}, \text{ whence, taking into account (83)}, \]

\[
\begin{align*}
|w^n - (1 + n(w - 1) + \frac{n(n-1)}{2}(w - 1)^2)| &\leq c_{17}(n|w - 1|)^3 \leq c_{18}n^3\alpha^{3/2}, \\
|w^{-n} - (1 - n(w - 1) + \frac{n(n+1)}{2}(w - 1)^2)| &\leq c_{17}(n|w - 1|)^3 \leq c_{18}n^3\alpha^{3/2} \\
\Rightarrow \frac{|w^n + w^{-n}|}{2} - 1 &\leq \frac{n^2}{2}|w - 1|^2 + c_{18}n^3\alpha^{3/2} \leq c_{19}(n^2\alpha + n^3\alpha^{3/2}) \leq c_{20}n^2\alpha.
\end{align*}
\]

Thus, one has \(|P_n(\zeta)| = \frac{|w^n + w^{-n}|}{2n^2|\alpha - i\beta|} - 1 \leq c_{15}, \text{ as required in (81)}\) for the case of

\( 0 \leq \alpha \leq \frac{1}{n^2}. \)

**3.2** Now consider the case of \( \frac{1}{n^2} \leq \alpha \leq 2 \). From (83), (84) it follows that

\[ |w| \leq 1 + c_{21}\alpha, \quad |w^{-1}| \leq 1 + c_{21}\alpha, \text{ whence } |P_n(\zeta)| = \frac{|w^n + w^{-n}|}{2n^2|\alpha - i\beta|} \leq c_{22}(1 + c_{21}\alpha)^n, \]

as required in (81). \( \square \)

**4.** Let \( Q(\zeta) = \frac{1 + \zeta}{2} \). It is immediately seen that

\[ \zeta = 1 - \alpha + i\beta \in \hat{\mathcal{M}} \Rightarrow |Q(\zeta)| \leq 1 - c_{23}\alpha \quad [c_{23} < \frac{1}{2}]. \tag{85} \]

32
Now let $c_{24}$ be a positive integer which is $\geq \frac{\alpha}{c_{23}}$ (see (81)). Consider the polynomial $S_n(\zeta) = P_n(\zeta)Q^{c_{24}n}(\zeta)$.

**Lemma 26** For every positive integer $n$, the polynomial $S_n(\zeta)$ possesses the following properties:

$$(a) \text{ : } \deg(S_n) \leq c_{25}n; \quad (b) \text{ : } S_n(1) = 1; \quad (c) \text{ : } \max_{\zeta} \{|S_n(\zeta)| : \zeta \in \overline{\mathbb{M}}, \Re(\zeta) = 1 - \alpha\} \leq c_{15} \min \left[ \frac{1}{n^\alpha}, 1 \right].$$  \hspace{1cm} (86)

**Proof.** Relations $(86.a - b)$ are evident (take into account that $P_n(1) = 1$ and $\deg(P_n) \leq n$). To verify $(86.c)$, note that if $\zeta = 1 - \alpha + i\beta \in \overline{\mathbb{M}}$, then in view of $(81)$ one has

$$0 \leq \alpha \leq \frac{1}{n^\alpha} \Rightarrow |S_n(\zeta)| \leq |P_n(\zeta)||Q(\zeta)|^{c_{24}n} \leq |P_n(\zeta)| \leq c_{15};$$

$$\frac{1}{n^\alpha} \leq \alpha \leq 2 \Rightarrow |S_n(\zeta)| \leq |P_n(\zeta)||Q(\zeta)|^{c_{24}n} \leq c_{15} \frac{(1+c_{23}n)^\alpha}{n^\alpha} (1 - c_{23}\alpha)^{c_{24}n} \leq c_{15} \frac{\exp(c_{15}n\alpha)}{n^\alpha} \exp \left\{ -c_{23}c_{24}n\alpha \right\} \leq \frac{c_{24}}{n^\alpha}$$

(for $a$, see (85), $b$ is due to $c_{23}c_{24} \geq c_{15}$).

Now we are ready to complete the proof of Proposition 12. Given a positive integer $n$, let us set $R_n(\zeta) = S_n(\zeta)$. In view of $(86)$ one has

$$(a) \text{ : } \deg(R_n) \leq c_{26}n; \quad (b) \text{ : } R_n(1) = 1; \quad (c) \text{ : } \max_{\zeta} \{|R_n(\zeta)| : \zeta \in \overline{\mathbb{M}}, \Re(\zeta) = 1 - \alpha\} \leq r_n(\alpha) \text{ (87)}$$

Consider the filters $q^{(n)}(z)$ given by $q^{(n)}(\Delta) = R_n(D)$, $n = 0, 1, ...$ By $(87.b)$ and Lemma 23 we have

$$\begin{align*}
T & \leq c_{27}N \\
1 \leq n(T) & \equiv \lfloor c_{27}T \rfloor \\
(s) & \in H_N^\alpha(D) \Rightarrow \begin{cases} \text{ord}(q^{(n(T))}) \leq T, \\ s_\tau = (q^{(n(T))}(\Delta)s)_\tau \forall (\tau : |\tau - t| \leq c_{27}N). \end{cases} \text{ (88)}
\end{align*}$$

By Parseval’s equality, we have also (in what follows, $n = n(T)$)

$$|q^{(n)}|_2^2 = \int_{\mathbb{R}^d_N} |R_n(\delta(\omega))|^2 \nu(d\omega) = \int_{\mathbb{M}} |R_n(\zeta)|^2 \mu(d\zeta) \leq \int_0^2 \int_{\mu_n(\alpha)} r_n(\alpha) dF(\alpha). \text{ (89)}$$
with \( a \) given by (87.\( c \)), (74) and the definition of \( F(\cdot) \). Let \( \gamma \) be the measure on \([0, 2]\) defined by \( G(\alpha) \equiv \gamma([0, \alpha]) = c_2^{-1}(\alpha^{d/2} + N^{-d}) \), so that

\[
F(\alpha) \leq G(\alpha) \equiv \gamma([0, \alpha]) \quad \forall \alpha \in [0, 2]
\]  
(90)

(see (75)). We have

\[
\begin{align*}
\frac{2}{\rho_n}(\alpha)dF(\alpha) &= \rho_n(2) - \frac{2}{\rho_n}(\alpha)F(\alpha)d\alpha \\
&\leq \rho_n(2) - \int \rho_n(\alpha)G(\alpha)d\alpha = \rho_n(2) - \rho_n(2)G(2) + \int \rho_n(\alpha)\gamma(d\alpha) \\
&\leq \int \rho_n(\alpha)\gamma(d\alpha) \equiv c_2^{-1} \left[ c_{28} \int_0^2 \min^2 \left[ n^{-2d} \alpha^{-d}, 1 \right] \alpha^{d-1}d\alpha \\
&+ \rho_n(0)N^{-d} \right] \leq c_{30} \left[ N^{-d} + n^{-d} \right] \leq c_{31}(2T + 1)^{-d}
\end{align*}
\]  
(91)

(a holds since \( \rho_n(\cdot) \) is nonincreasing, see (87.\( c \)), and by (90), b holds since \( c_2 \in (0, 1) \), see Lemma 24, c is by (87.\( c \)) and (89), \( d \) is due to \( n = n(T) = \lfloor c_{27}T \rfloor \). Combining (89) and (91), we conclude that

\[
|q^{(n(T))}|_2 \leq c_{32}(2T + 1)^{-d/2}.
\]  
(92)

From (88) and (92) we conclude that if \( L = \lfloor c_{33}N \rfloor \) and \( T \leq L \) is such that \( n(T) \equiv \lfloor c_{27}T \rfloor \geq 1 \), then

\[
\exists q^{(T)} \in C_T(\mathbb{Z}^d) : \begin{cases} 
|q^{(T)}|_2 \leq c_{32}(2T + 1)^{-d/2}, \\
\tau = (q^{(T)}(\Delta)s), \forall (\tau, |\tau - t| \leq L, (s) \in H_N^d(D))
\end{cases}
\]  
(93)

(Indeed, one can choose, as a required \( q^{(T)} \), the filter \( q^{(n(T))} \)). Setting \( q^{(T)}(z) \equiv 1 \) for \( T < \frac{1}{c_{27}} \), we enforce the validity of (93) for all \( T, 0 \leq T \leq L \). Thus, \( H_N^d(D) \subset F^d_{c_{29}L}(0, c_{34}) \). □

5.2.8 Proof of Proposition 13

**Lemma 27** Let \( f \in H^+(M) \) be a deterministic function, let \( N \leq M/2 \), and let \( t \in \mathbb{Z}^d, |t| \leq N \). Consider the “discrete box” \( B^t_N = \{ \tau \in \mathbb{Z}^d : |\tau - t| \leq N \} \), and let \( \phi \) be a deterministic function on \( B^t_N \) which coincides with \( f \) on the “discrete boundary” \( \partial B^t_N \equiv \{ \tau \in \mathbb{Z}^d : |\tau - t| = N \} \) of \( B^t_N \) and is “discrete
harmonic": \( \tau \in \mathbb{Z}^d, |\tau - t| < N \Rightarrow \phi_\tau = \frac{1}{2d} \sum_{e=(e_1, \ldots, e_d)} \phi_{\tau + e} \). Then

\[
\tau \in B_N^t \Rightarrow |f(\tau) - \phi_\tau| \leq c_1 \|f\|_{\infty, 2M} N^{-2}
\] (94)

(from now on, \( c_i \) are positive absolute constants).

Proof. First, we should prove that the “discrete harmonic” function \( \phi \) on \( B_N^t \) which coincides with \( f \) on \( \partial B_N^t \) does exist. This fact is well known; we present here its proof just for the sake of completeness. Let \( \psi \) be a function on \( \partial B_N^t \). Consider the following random walk on \( B_N^t \): arriving for the first time at a point \( \tau \) from \( \partial B_N^t \), we pay penalty \( \psi(\tau) \) and terminate; from an “interior point” \( \tau \in \text{int} B_N^t \equiv B_N^t \setminus \partial B_N^t \), we make a random step of length 1 along one of the coordinate axes, choosing every one of 2\( d \) possible steps with probability \( 1/(2d) \). It is immediately seen that the expected penalty payed at the termination, treated as a function of the initial state, is a discrete harmonic function with the boundary values \( \psi \).

Now, since \( |t| \leq N \) and \( 2N \leq M \), the function \( f \) is harmonic in the “continuous box” \( D_N^t = \{ \tau \in \mathbb{R}^d : |\tau - t| \leq 2N \} \), and the uniform norm of \( f \) in this square does not exceed \( \|f\|_{\infty, 2M} \). From the standard results on harmonic functions it follows that

\[
\forall (\tau \in D_N^t) : \left| \frac{\partial^\kappa}{\partial x_j^\kappa} f(\tau) \right| \leq c_2 \|f\|_{\infty, 2M} N^{-\kappa}, \ k = 1, 2, 3, 4, \ j = 1, \ldots, d.
\] (95)

Consequently, for the basic orths \( e_j, j = 1, \ldots, d \) we have

\[
\tau \in D_N^t, \ |s| \leq 1 \Rightarrow \left| f(\tau + s e_j) - \sum_{\kappa=0}^{3} \frac{1}{\kappa!} \frac{\partial^\kappa}{\partial x_j^\kappa} f(\tau) s^\kappa \right| \leq c_3 |s|^4 \|f\|_{\infty, 2M} N^{-4}.
\]

Since \( f \) is harmonic, we conclude that

\[
|\langle D f \rangle_\tau| \leq c_4 \|f\|_{\infty, 2M} N^{-4}, \ \tau \in B_N^t.
\] (96)

Now let \( h = f|_{z^d} - \phi \in C(B_N^t) \) and let \( h^\pm_\tau = h_\tau \mp \frac{2c_4 \|f\|_{\infty, 2M}}{N^{4}} \sum_{j=1}^{d} (\tau_j - t_j)^2 \).

Taking into account (96) and the fact that \( \phi \) is discrete harmonic, we have for \( \tau \in \text{int} B_N^t \):

\[
\langle Dh^+ \rangle_\tau = \langle Dh \rangle_\tau + \frac{2c_4 \|f\|_{\infty, 2M}}{N^{4}} \geq 0, \ \langle Dh^- \rangle_\tau = \langle Dh \rangle_\tau - \frac{2c_4 \|f\|_{\infty, 2M}}{N^{4}} < 0,
\]

whence both the maximum of \( h^+ \) and the minimum of \( h^- \) over \( B_N^t \) are attained at \( \partial B_N^t \). Since at the discrete boundary of \( B_N^t \) we have \( f = \phi \) and therefore
$h^+ \leq 4c_4\|f\|_{\infty,2M}N^{-2}$, we conclude that $\tau \in B_N^t \Rightarrow h_\tau \leq h^+_\tau \leq \max_{\tau \in \partial B_N^t} h^+_{\tau}$, $2dc_4\|f\|_{\infty,2M}N^{-2}$. By similar reasons, $\tau \in B_N^t \Rightarrow h_\tau \geq h^-_\tau \geq \min_{\tau \in \partial B_N^t} h^-_{\tau}$, $-2dc_4\|f\|_{\infty,2M}N^{-2}$. □

Now let $|t| \leq M/8$ and $L \leq M/8$. Given $T$, $0 \leq T \leq L$, and applying Proposition 12, we can build filter $q^{(T)} \in C_T(\mathbb{Z}^d)$ such that

$$|q^{(T)}|_2 \leq c_5(2T + 1)^{-1}, \quad \phi_\tau = \sum_{|\nu| \leq T} \phi_{\tau-\nu}q^{(T)}_\nu \forall (\tau : |\tau - t| \leq L)$$

(97)

for every $\phi$ which is discrete harmonic in the discrete box $B_{2L}^t$. Now let $f \in H(M,R)$. Applying Lemma 27, we can find function $\phi$ which is discrete harmonic in the box $B_{2L}^t$ and such that $|\phi_\tau - f_\tau|^2 \leq c_6^2\|f\|^2_{\infty,2M}L^{-4}$ for $\tau \in B_{2L}^t$. From (97) it now follows that

$$\forall (\tau : |\tau - t| \leq L) : \left[ E \left( \left| f_\tau - \sum_{|\nu| \leq T} f_{\tau-\nu}q^{(T)}_\nu \right|^2 \right) \right]^{1/2} \leq c_6 \left[ E \left[ \|f\|^2_{\infty,2M} \right] \right]^{1/2} L^{-2} (1 + |q^{(T)}_t|_1) \leq c_8 RL^{-2} (1 + |q^{(T)}_t|_2 (2T + 1)^{d/2}) \leq c_9 RL^{-2} \leq c_9 R(2T + 1)^{-d/2}$$

(recall that $d \leq 4$). □

The proof of Proposition 14 is completely similar to that of Proposition 10.

References


