Computing Orthogonal Projections Can Be Too Computational Expensive

Holger Boche, Volker Pohl and H. Vincent Poor

***Abstract*— Orthogonal projections are widely used to determine optimal controllers, filters and approximations. This paper shows that despite their simple theoretical foundation and their regular use in practical applications, the complexity for computing orthogonal projections might be very high. The paper proves that the causal projection on L^2 maps computable continuous functions onto not computable functions. Moreover, it is shown that the orthogonal projection associated with the polynomial approximation in L^2 shows complexity blowup in the sense that it maps polynomial-time computable continuous functions onto polynomials that are not polynomial-time computable. Finally, it is shown that the coefficients of the Wiener prediction filter for stationary stochastic processes might not be computable in polynomial time, even for smooth and polynomial-time computable spectral densities.**

I. Introduction

Orthogonal projections play a central role in many engineering problems such as the design of optimal controllers and filters [1], [2], signal recovery [3], estimation [4], [5], or for general signal processing methods [6]–[8]. Let $\mathcal{H}$ be an arbitrary Hilbert space and let $\mathcal{U} \subset \mathcal{H}$ be a closed convex subset of $\mathcal{H}$. Then many engineering problems can be formulated as the following minimization problem: Given $f \in \mathcal{H}$, find the best approximation of f by an element from $\mathcal{U}$, i.e. find $f_{\mathcal{U}} \in \mathcal{U}$ such that

$$\|f - f_{\mathcal{U}}\|_{\mathcal{H}} = \inf_{g\in\mathcal{U}} \|f - g\|_{\mathcal{H}} = d(f,\mathcal{U}) \ ,$$

where $d(f,\mathcal{U})$ is the minimum *distance* of f from $\mathcal{U}$. It is well known that this problem has a unique minimizer $f_{\mathcal{U}} \in \mathcal{U}$ and the linear mapping $\mathrm{P}_{\mathcal{U}} : f \mapsto f_{\mathcal{U}}$ is said to be the *orthogonal projection* from $\mathcal{H}$ onto $\mathcal{U}$ (cf. Fig. 1). It is clear from the definition that $\mathrm{P}_{\mathcal{U}}^2 = \mathrm{P}_{\mathcal{U}}$. If $\mathcal{U}$ is even a closed subspace of $\mathcal{H}$, then $\mathrm{I}_{\mathcal{H}} - \mathrm{P}_{\mathcal{U}}$ is the orthogonal projection onto the orthogonal complement $\mathcal{U}^{\perp}$ of $\mathcal{U}$ in $\mathcal{H}$, where $\mathrm{I}_{\mathcal{H}}$ denotes the identity operator on $\mathcal{H}$. Since $f = \mathrm{P}_{\mathcal{U}} f + (\mathrm{I}_{\mathcal{H}} - \mathrm{P}_{\mathcal{U}})f = f_{\mathcal{U}} + f_{\mathcal{U}^{\perp}}$ for every $f \in \mathcal{H}$, one has a decomposition of $\mathcal{H}$ into two orthogonal subspaces: $\mathcal{H} = \mathcal{U} \oplus \mathcal{U}^{\perp}$.

This paper investigates the computability and the computational complexity of $f_{\mathcal{U}}$ and $d(f,\mathcal{U})$. Since computations are usually done by digital computers, all objects f, $f_{\mathcal{U}}$, and $d(f,\mathcal{U})$ can not be computed exactly (apart from trivial cases) but they have to be approximated by objects that can be processed by digital computers (e.g. rational numbers, rational polynomials, etc.). If such an approximation is possible by *effectively controlling the approximation error*, the object is said to be *computable*. Assume $f \in \mathcal{H}$ is computable, then it is an interesting and important question whether also $f_{\mathcal{U}}$ and $d(f,\mathcal{U})$ are computable. If f is computable then there exists the four possible combinations shown in Table I.

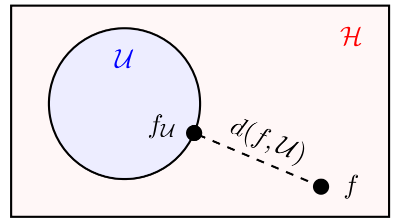

Fig. 1. Orthogonal projection of $f \in \mathcal{H}$ onto the subspace $\mathcal{U}$.

TABLE I

Possible combinations for computability/non-computability.

f	$f_{\mathcal{U}}$	$d(f,\mathcal{U})$
computable	computable	computable
computable	computable	non-computable
computable	non-computable	computable
computable	non-computable	non-computable

Our first question in this paper is whether all of these possible combinations can actually occur. For example, is it possible that $f_{\mathcal{U}}$ is not-computable but that $d(f,\mathcal{U})$ is computable, or vice versa? To date there seems to exist no publications addressing this fundamental question, even though orthogonal projections play such an important role in control, signal processing, and many other fields.

We will show that there exist simply and very common applications that illustrate the behavior of Rows 1, 3, and 4. In particular, Section III will show that a behavior according to Row 1 occurs for $\mathcal{H} = \ell^2(\mathbb{Z})$ and for $\mathcal{U} = \ell^2(\mathbb{N})$, the closed subspace of all causal discrete signals. Then Section IV shows that for $\mathcal{H} = L^2(\mathbb{T})$, the space of square integrable functions on the unit circle $\mathbb{T} \subset \mathbb{C}$, and $\mathcal{U} = L^2_+(\mathbb{T})$, the closed subspace of all functions $f \in L^2(\mathbb{T})$ for which all Fourier coefficients $c_n(f)$ with $n < 0$ vanish, one gets a behavior as in Row 3. Finally, Section V provides an example for Row 4. This example is closely related to a central question in prediction theory for stationary stochastic processes. Here $\mathcal{H}$ is a weighted L^2-space with a smooth

This work was partly supported by the German Ministry of Education and Research (BMBF) within the program "Souverän. Digital. Vernetzt" in the joint project 6G-life under Grant 16KISK002 and within the joint project Post Shannon Communication (NewCom) under Grant 16KIS1003K and also by the German Research Foundation (DFG) under Grant BO 1734/37–1. The work of H. V. Poor was partly supported by the U.S National Science Foundation under Grants CNS-2128448 and ECCS-2335876.

H. Boche and V. Pohl are with the School of Computation, Information and Technology, Technical University of Munich, 80333 München, Germany, {boche, volker.pohl}@tum.de. H. Boche is also with the Munich Center for Quantum Science and Technology (MCQST), 80799 München, Germany. H. V. Poor is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544, USA poor@princeton.edu.

weight function and $\mathcal{U}$ corresponds to the subspace of FIR prediction filters.

As a second class of problems, we consider situations where $\mathcal{H}$ and $\mathcal{U}$ are such that a computable f always implies that $f_{\mathcal{U}}$ or $d(f,\mathcal{U})$ or both are computable (Rows $1-3$ of Table I). Then we ask for the complexity of computing $f_{\mathcal{U}}$ or $d(f,\mathcal{U})$. This question is of particular interest if $\mathcal{U}$ is a finite dimensional subspace of $\mathcal{H}$, because then $f_{\mathcal{U}}$ and $d(f,\mathcal{U})$ are always computable and one may ask for the complexity of computing $f_{\mathcal{U}}$ and $d(f,\mathcal{U})$. Here we consider this question for the computation of $f_{\mathcal{U}}$ in the application of optimal prediction of stochastic processes (Section V). Finally, Section VI investigates this question for the computation of $f_{\mathcal{U}}$ for the projection related to polynomial approximations of continuous functions. In both examples, the orthogonal projection shows *complexity blowup*, i.e. there exists an $f \in \mathcal{H}$ that is polynomial-time computable but for which $\mathrm{P}_{\mathcal{U}} f$ is not polynomial-time computable. So even if $f_{\mathcal{U}}$ or $d(f,\mathcal{U})$ are computable, the actual computation might still be very complex in the sense of complexity theory.

II. Notation and Preliminaries

For $1 \le p < \infty$, $L^p(\mathbb{T})$ denotes the usual spaces of integrable functions on the *unit circle* $\mathbb{T} = \{z \in \mathbb{C} \ : \ |z| = 1\}$. In particular, $L^2(\mathbb{T})$ is a Hilbert space with inner product $\langle f, g\rangle_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\overline{g(\mathrm{e}^{\mathrm{i}\theta})}\,\mathrm{d}\theta$. The Banach space of continuous functions on $\mathbb{T}$ with norm $\|f\|_\infty = \max_{\zeta\in\mathbb{T}} |f(\zeta)|$ is denoted by $\mathcal{C}(\mathbb{T})$, and $\mathcal{P}_N \subset \mathcal{C}(\mathbb{T})$ stands for the set of trigonometric polynomials of degree not larger than $N \in \mathbb{N}$, i.e. of functions p of the form

$$p(\mathrm{e}^{\mathrm{i}\theta}) = \tfrac{a_0(p)}{2} + \textstyle\sum_{n=1}^{N} a_n(p)\cos(n\theta) + b_n(p)\sin(n\theta)$$

with the *Fourier coefficients*

$$\begin{aligned} a_n(p) &= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} p(\mathrm{e}^{\mathrm{i}\theta})\,\cos(nt)\,\mathrm{d}t \quad \text{and} \\ b_n(p) &= \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} p(\mathrm{e}^{\mathrm{i}\theta})\,\sin(nt)\,\mathrm{d}t\ . \end{aligned} \tag{1}$$

We write $H(\mathbb{D})$ for the set of functions that are holomorphic (i.e. analytic) in the *unit disk* $\mathbb{D} = \{z \in \mathbb{C} \ : \ |z| < 1\}$ and $H^\infty(\mathbb{D})$ is the Banach space of all $f \in H(\mathbb{D})$ with norm $\|f\|_\infty = \sup_{z\in\mathbb{D}} |f(z)| < \infty$.

This paper paper investigates *causal projections* on several Hilbert spaces ($\ell^2(\mathbb{Z})$, $L^2(\mathbb{T})$, $L^2(\varphi)$). To simplify notation, all these different projections will be denoted by P_+. So the operator P_+ has always be viewed in the actual context.

We heavily rely on concepts and notation from computational analysis and complexity theory. Because of space constraints, we refer to standard literature (e.g., [9]–[12]) for corresponding introductions. As far as notation is concerned, we will basically follow [13]. In particular, $\mathbb{R}_{\mathrm{c}}$ and $\mathbb{C}_{\mathrm{c}}$ stand for the set of *computable real and complex numbers*, respectively, and $\mathcal{C}_{\mathrm{c}}(\mathbb{T})$ denotes the set of all *computable continuous functions* on $\mathbb{T}$. Some of our results are based on the assumption that the complexity class FP_1 is strictly smaller than the complexity class $\#P_1$. Even thought there exists no formal proof of this conjecture, it is widely assumed that it is true. Moreover, this assumption is true if the better known conjecture $P \neq NP$ is true.

III. Causal Projection on ℓ^2

This section presents an example for the most favorable behavior of Table I, namely for Row 1. This example will also be used in Section IV to provide an example for Row 3. We consider one of the most simple Hilbert spaces, namely the space $\mathcal{H} = \ell^2(\mathbb{Z})$ of all sequences $\boldsymbol{a} = \{a_n\}_{n\in\mathbb{Z}} \subset \mathbb{C}$ with finite ℓ^2-norm $\|\boldsymbol{a}\|_2^2 = \big(\sum_{n\in\mathbb{Z}} |a_n|^2\big)^{1/2}$. Then we investigate the causal and anti-causal projection

$$(\mathrm{P}_+\boldsymbol{a})(n) = \begin{cases} a_n : n \ge 0 \\ 0 : n < 0 \end{cases} \text{ and } (\mathrm{P}_-\boldsymbol{a})(n) = \begin{cases} 0 : n \ge 0 \\ a_n : n < 0 \end{cases}$$

respectively. These projections play a fundamental role in control, signal processing and communications.

Remark III.1: Note that both operators have a simple closed form representation as $(\mathrm{P}_+\boldsymbol{a})(n) = s(n)\,a_n$ and $(\mathrm{P}_+\boldsymbol{a})(n) = [1 - s(n)]\,a_n$ with the step response given by $s(n) = 1$ for $n \ge 1$ and $s(n) = 0$ for $n < 0$. Based on this representation, the computability of P_+ and P_- according to Table I could immediately be verified.

We require that the input sequences of these projections are computable sequences of computable numbers, i.e. every element of the sequences can effectively be approximated by a rational number. Since our sequences are in $\ell^2(\mathbb{Z})$, it is natural to require that also the norm is a computable number.

Definition III.1 (Computable ℓ^2-sequences): *A computable sequence of computable numbers $\boldsymbol{a} \in \ell^2(\mathbb{Z})$ is said to be an ℓ^2-*computable sequence *if there exists a (recursive) function $\phi : \mathbb{N} \to \mathbb{N}$ such that for all $M \in \mathbb{N}$*

$$N_1 \ge N_0 = \phi(M) \textit{ implies } \textstyle\sum_{N_0 \le |n| \le N_1} |a_n|^2 < 2^{-M}. \tag{2}$$

We write $\ell^2_{\mathrm{c}}(\mathbb{Z})$ for the set of all ℓ^2-computable sequences.

Remark III.2: Condition (2) holds if and only if $\|\boldsymbol{a}\|_2 \in \mathbb{R}_{\mathrm{c}}$.

Now we ask whether for any $\boldsymbol{a} \in \ell^2_{\mathrm{c}}(\mathbb{Z})$ also the projections $\mathrm{P}_+\boldsymbol{a}$ and $\mathrm{P}_-\boldsymbol{a}$ are again ℓ^2-computable sequences.

Theorem III.1: *For every $\boldsymbol{a} \in \ell^2_{\mathrm{c}}(\mathbb{Z})$, we have $\mathrm{P}_+\boldsymbol{a} \in \ell^2_{\mathrm{c}}(\mathbb{Z})$ and $\mathrm{P}_-\boldsymbol{a} \in \ell^2_{\mathrm{c}}(\mathbb{Z})$.*

In view of Table I, Theorem III.1 shows that the projections P_+ and P_- on $\mathcal{H} = \ell^2(\mathbb{Z})$ show the behavior of Row 1.

Proof: It is clear that $\mathrm{P}_+\boldsymbol{a}$ and $\mathrm{P}_-\boldsymbol{a}$ are computable sequences of computable numbers. So we only have to verify that $\|\mathrm{P}_+\boldsymbol{a}\|_2$ and $\|\mathrm{P}_-\boldsymbol{a}\|_2$ are computable numbers. We only show the proof for $\mathrm{P}_+\boldsymbol{a}$. First, we define by

$$\alpha_N = \big(\textstyle\sum_{n=0}^{N} |a_n|^2\big)^{1/2}\ , \qquad N \in \mathbb{N}\ ,$$

a monotonically increasing computable sequence $\{\alpha_N\}_{N\in\mathbb{N}}$ of computable numbers with $\lim_{N\to\infty}\alpha_N = \|\mathrm{P}_+\boldsymbol{a}\|_2$. On the other hand, since $\|\mathrm{P}_+\boldsymbol{a}\|_2^2 \le \|\boldsymbol{a}\|_2^2 - \sum_{n=-N}^{-1} |a_n|^2$ for all $N \in \mathbb{N}$, we define by

$$\beta_N = \big(\|\boldsymbol{a}\|_2^2 - \textstyle\sum_{n=-N}^{-1} |a_n|^2\big)^{1/2}\ , \qquad N \in \mathbb{N}\ ,$$

a monotonically decreasing computable sequence $\{\beta_N\}_{N\in\mathbb{N}}$ of computable numbers with $\lim_{N\to\infty}\beta_N = \|\mathrm{P}_+\boldsymbol{a}\|_2$. Thus, $\|\mathrm{P}_+\boldsymbol{a}\|_2$ is the limit of a computable sequence that converges from below and a computable sequence that converges from above to $\|\mathrm{P}_+\boldsymbol{a}\|_2$. Therefore $\|\mathrm{P}_+\boldsymbol{a}\|_2 \in \mathbb{R}_{\mathrm{c}}$. ∎

IV. Causal Projection on $L^2(\mathbb{T})$

With every sequence $\boldsymbol{a} \in \ell^2(\mathbb{Z})$, we may associate its Fourier series. This gives rise to the following description of the projection operator P_+ on $\mathcal{H} = L^2(\mathbb{T})$.

Since $\mathcal{C}_\mathrm{c}(\mathbb{T}) \subset L^2(\mathbb{T})$, every $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ can be written as a Fourier series

$$f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n\in\mathbb{Z}} c_n(f)\,\mathrm{e}^{\mathrm{i}n\theta}\ , \quad \theta \in [-\pi, \pi)\ ,$$

with the *Fourier coefficients* $c_n(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}n\theta}\mathrm{d}\theta$, $n \in \mathbb{Z}$, and where the sum converges in $L^2(\mathbb{T})$. Then the causal orthogonal projection is a mapping from $L^2(\mathbb{T})$ onto $L^2_+(\mathbb{T}) = \left\{ f \in L^2(\mathbb{T}) : c_n(f) = 0 \text{ for } n < 0 \right\}$, defined by

$$\mathrm{P}_+ : \sum_{n\in\mathbb{Z}} c_n(f)\,\mathrm{e}^{\mathrm{i}n\theta} \mapsto \sum_{n=0}^{\infty} c_n(f)\,\mathrm{e}^{\mathrm{i}n\theta}$$

or equivalently as the solution of the minimization problem

$$\inf_{p\in\mathcal{P}_+} \|f - p\|^2_{L^2(\mathbb{T})} = \|f - \mathrm{P}_+ f\|^2_{L^2(\mathbb{T})} = d(f, L^2_+(\mathbb{T})) \quad (3)$$

where $\mathcal{P}_+$ is the set of all trigonometry polynomials of the form $p(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=0}^{N} a_n\,\mathrm{e}^{\mathrm{i}n\theta}$ for some degree $N \in \mathbb{N}$. Similarly, we may consider the anti-causal projection, defined by $\mathrm{P}_- : \sum_{n\in\mathbb{Z}} c_n(f)\,\mathrm{e}^{\mathrm{i}n\theta} \mapsto \sum_{n=-\infty}^{-1} c_n(f)\,\mathrm{e}^{\mathrm{i}n\theta}$.

First, we use the ℓ^2-results from Section III to show that $d(f, L^2_+(\mathbb{T}))$ is a computable real number for every computable continuous function $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$.

Theorem IV.1: *For all $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$, $d(f, L^2_+(\mathbb{T}))$ is in $\mathbb{R}_\mathrm{c}$.*

Proof: Since $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ there exists an algorithm [9] with inputs f and $n \in \mathbb{Z}$ that computes the Fourier coefficient $c_n(f)$. Therefore $\boldsymbol{c} = \{c_n(f)\}_{n\in\mathbb{Z}} \in \ell^2(\mathbb{Z})$ is a computable sequence of computable numbers. The mapping $x \mapsto x^2$ is a computable continuous function on $\mathbb{R}$. Consequently, also $\left\{ |c_n(f)|^2 \right\}_{n\in\mathbb{Z}}$ is a computable sequence of computable numbers, and since $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ also $|f|^2 \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ and so $\|f\|^2_{L^2(\mathbb{T})} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|f(\mathrm{e}^{\mathrm{i}\theta})\right|^2 \mathrm{d}\theta$ is a computable real number. For $N \in \mathbb{N}$, let $d_N(f) = \sum_{n=-N}^{N} |c_n(f)|^2$, then $\{d_N\}_{N\in\mathbb{N}}$ is a monotonically increasing computable sequence of computable numbers. Therefore d_N effectively converges [9] to $\sum_{n\in\mathbb{Z}} |c_n(f)|^2 = \|f\|^2_{L^2(\mathbb{T})} \in \mathbb{R}_\mathrm{c}$ as $N \to \infty$, showing that $\boldsymbol{c} \in \ell^2_\mathrm{c}$. Using Parseval's theorem, (3) yields

$$d(f, L^2_+(\mathbb{T})) = \sum_{n=-\infty}^{-1} |c_n(f)|^2 = \|\mathrm{P}_- \boldsymbol{c}\|^2_{\ell^2(\mathbb{Z})}\ .$$

So Theorem III.1 implies $d(f, L^2_+(\mathbb{T})) \in \mathbb{R}_\mathrm{c}$. ∎

Next, we ask whether for any $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ the projections $\mathrm{P}_+ f$ and $\mathrm{P}_- f$ are again computable continuous functions. The following theorem shows that this is not the case, even if f is additional smooth.

Theorem IV.2: *There exists a computable continuous function $f_1 \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ with the properties*

1) *f_1 is absolute continuous*
2) *f_1' is a computable $L^1(\mathbb{T})$ function*
3) *$\mathrm{P}_+ f_1$ is a continuous function*

but such that $(\mathrm{P}_+ f_1)(1) \notin \mathbb{C}_\mathrm{c}$, i.e. such that $\mathrm{P}_+ f_1$ is not a computable continuous function on $\mathbb{T}$.

Remark IV.1: Since $f(1) = (\mathrm{P}_+ f)(1) + (\mathrm{P}_- f)(1)$, it follows that $(\mathrm{P}_- f)(1) \notin \mathbb{C}_\mathrm{c}$, i.e. also $\mathrm{P}_- f$ is not a computable continuous function on $\mathbb{T}$.

Remark IV.2: There exist infinitely many projections with the same property. For example, for any $K \in \mathbb{N}$, we may define $(\mathrm{P}_{+,K} f)(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n=-K}^{\infty} c_n(f)\,\mathrm{e}^{\mathrm{i}n\theta}$. Then Theorem IV.2 also holds for $\mathrm{P}_{+,K}$ with the same function f_1.

Remark IV.3: The operator P_+ on $L^2(\mathbb{T})$ can only be represented by a singular integral via the Hilbert transform and so there exists no closed form representation to determine the continuous function $\mathrm{P}_+ f_1$ for $f_1 \in \mathcal{C}_\mathrm{c}(\mathbb{T})$. If such a simple formula for the determination of $\mathrm{P}_+ f$ would exist (e.g. as for P_+ on $\ell^2(\mathbb{Z})$ as in Remark III.1) then $\mathrm{P}_+ f$ would be a computable function for every $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$.

Sketch of Proof: The proof is based on a proof in [14, Theorem III.1]. There a computable continuous function $f_1 \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ was constructed that is absolute continuous and that has an absolute continuous conjugate function $\widetilde{f}_1 \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ with $\widetilde{f}_1(1) \notin \mathbb{R}_\mathrm{c}$. Using the properties of conjugate functions, the causal projection can be written as

$$(\mathrm{P}_+ f_1)(\zeta) = \tfrac{1}{2}\left[c_0(f_1) + f_1(\zeta) + \mathrm{i}\widetilde{f}_1(\zeta)\right]\ , \quad \zeta \in \mathbb{T}\ .$$

Recalling that $c_0(f_1)$ is a computable number (cf. proof of Theorem IV.1), it immediately follows that $\mathrm{P}_+ f_1$ is a continuous function on $\mathbb{T}$ and that $(\mathrm{P}_+ f_1)(1) \notin \mathbb{C}_\mathrm{c}$. That f_1 satisfies also Property 2) follows from the construction in [14] but is not shown here, because of space constraints. ∎

Theorems IV.1 and IV.2 show that the orthogonal projection $\mathrm{P}_+ : L^2(\mathbb{T}) \to L^2_+(\mathbb{T})$ shows a behavior as in Row 3 of Table I, i.e. a computable f is mapped onto a non-computable $f_\mathcal{U}$ whereas the distance $d(f, \mathcal{U})$ is computable for every computable f. We also observe that even thought the projections discussed in Sections III and IV are closely related, their computability behavior is different.

In view of Theorem IV.2, it is an interesting question for future research to derive sufficient conditions on $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ such that $\mathrm{P}_+ f$ is again a computable continuous function.

V. Prediction Theory of Stochastic Processes

This section provides an example for the behavior in Row 4 of Table I. It is shown that the projection operator from Wiener's prediction theory of stationary stochastic processes (see, e.g., [15]–[19]) shows exactly such a behavior.

A. Preliminaries: Prediction of stochastic processes

For a probability space $(\Omega, \mathscr{F}, \nu)$ let $\mathcal{R} = \mathcal{R}(\Omega, \mathscr{F}, \nu)$ be the Hilbert space of all (complex) random variables (rvs) x with zero mean and finite second moment and with inner product $\langle x, y\rangle_\mathcal{R} = \mathrm{E}[x\overline{y}] = \int_\Omega x(\omega)\,\overline{y(\omega)}\,\mathrm{d}\nu(\omega)$. Let $\boldsymbol{x} = \{x_n\}_{n\in\mathbb{Z}} \subset \mathcal{R}$ be a *wide-sense stationary (wss)* with *auto-covariance function* $\gamma_{\boldsymbol{x}}(n) = \langle x_n, x_0\rangle_\mathcal{R}$, $n \in \mathbb{Z}$ and assume that $\boldsymbol{x}$ is completely non-deterministic. Then $\gamma_{\boldsymbol{x}}$ has the spectral representation $\gamma_{\boldsymbol{x}}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi_{\boldsymbol{x}}(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}n\theta}\mathrm{d}\theta$, $n \in \mathbb{Z}$, with the *spectral density* $\varphi_{\boldsymbol{x}} \in L^1(\mathbb{T})$ that satisfies the *Szegö's condition*

$$\int_{-\pi}^{\pi} \log \varphi_{\boldsymbol{x}}(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta > -\infty\ . \quad (4)$$

Consider the problem of finding the best (in the norm of $\mathcal{R}$) linear predictor $\widehat{x}_0$ of x_0 from past observations $\{x_n : n = -1, -2, -3, ...\}$ of the sequence $\boldsymbol{x}$. This predictor has the form of a linear filter H

$$\widehat{x}_0 = \mathrm{H}(\boldsymbol{x}) = \sum_{n=1}^{\infty} h[n]\, x_{-n}$$

where $\boldsymbol{h} = \{h[n]\}_{n\in\mathbb{N}}$ is the *impulse response* of H, and

$$h(z) = \sum_{n=1}^{\infty} h[n]\, z^n \;, \quad z \in \mathbb{D} \;, \tag{5}$$

its *transfer function*[1]. Thus, we search for the transfer function h of a filter H that minimizes the *mean square error*

$$\sigma_h^2 = \|x_0 - \sum_{n=1}^{\infty} h[n]\, x_{-n}\|_{\mathcal{R}}^2 \;.$$

It is well known that the optimal h is given as the solution of the minimization problem

$$\begin{aligned}
&\inf_{h\in L_+^2(\varphi_{\boldsymbol{x}})} \|x_0 - \sum_{n=1}^{\infty} h[n]\, x_{-n}\|_{\mathcal{R}}^2 \\
&\quad= \inf_{h\in L_+^2(\varphi_{\boldsymbol{x}})} \tfrac{1}{2\pi}\int_{-\pi}^{\pi} \left|1 - h(\mathrm{e}^{\mathrm{i}\theta})\right|^2 \varphi_{\boldsymbol{x}}(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{d}\theta \\
&\quad= \inf_{h\in L_+^2(\varphi_{\boldsymbol{x}})} \|1 - h\|_{L^2(\varphi_{\boldsymbol{x}})}^2 = d(1, L_+^2(\varphi_{\boldsymbol{x}})) \;,
\end{aligned} \tag{6}$$

wherein $L^2(\varphi_{\boldsymbol{x}})$ is the space of functions on $\mathbb{T}$ with inner product $\langle f, g\rangle_{L^2(\varphi_{\boldsymbol{x}})} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\overline{g(\mathrm{e}^{\mathrm{i}\theta})}\, \varphi_{\boldsymbol{x}}(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{d}\theta$, and where $L_+^2(\varphi_{\boldsymbol{x}}) = \overline{\mathrm{span}}\left\{\mathrm{e}^{\mathrm{i}n(\cdot)} : n = 1, 2, 3, \dots\right\}$. So (6) shows that the optimal transfer function is given by

$$h_{\mathrm{opt}} = \mathrm{P}_+(1) \;,$$

with the orthogonal projection P_+ from $\mathcal{H} = L^2(\varphi_{\boldsymbol{x}})$ onto $\mathcal{U} = L_+^2(\varphi_{\boldsymbol{x}})$, and that h_{opt} satisfies

$$d(1, L_+^2(\varphi_{\boldsymbol{x}})) = \|1 - h_{\mathrm{opt}}\|_{L^2(\varphi_{\boldsymbol{x}})}^2 \;.$$

Here $d(1, L_+^2(\varphi_{\boldsymbol{x}}))$ is the *minimum mean square (prediction) error (MMSE)*. Subsequently, the filter coefficients of the optimal prediction filter are always be denoted by $\{h_{\mathrm{opt}}[n]\}_{n=1}^{\infty}$ and so the transfer function of the optimal filter is given by

$$h_{\mathrm{opt}}(z) = \sum_{n=1}^{\infty} h_{\mathrm{opt}}[n]\, z^n \;, \qquad z \in \mathbb{D} \;. \tag{7}$$

Remark V.1: Note that $\mathrm{P}_+ : L^2(\varphi_{\boldsymbol{x}}) \to L_+^2(\varphi_{\boldsymbol{x}})$ depends on the spectral density $\varphi_{\boldsymbol{x}}$, i.e. different spectral densities define different projection operators.

We need the closed form expression for h_{opt}, which can be obtained by means of the *spectral factorization* of the $\varphi_{\boldsymbol{x}}$.

Proposition V.1: *Let $\boldsymbol{x}$ be a purely non-deterministic wss stochastic process with spectral density $\varphi \in L^1(\mathbb{T})$. The transfer function of the optimal prediction filter is given by*

$$h_{\mathrm{opt}}(z) = 1 - \varphi_+(0)/\varphi_+(z) \;. \tag{8}$$

with the spectral factor φ_+ of φ, given by

$$\varphi_+(z) = \exp\left(\frac{1}{4\pi}\int_{-\pi}^{\pi} \log\varphi(\mathrm{e}^{\mathrm{i}\theta}) \frac{\mathrm{e}^{\mathrm{i}\theta} + z}{\mathrm{e}^{\mathrm{i}\theta} - z}\, \mathrm{d}\theta\right), \; z \in \mathbb{D}. \tag{9}$$

[1]Usually only the boundary function $h(\mathrm{e}^{-\mathrm{i}\theta})$, $\theta \in [-\pi, \pi)$ is called the *transfer function* of H. However, since the prediction filter is causal, the sum (5) converges for all $z \in \mathbb{D}$ if the boundary function exists.

Remark V.2: Recall that the spectral factor has the following properties: $\varphi_+ \in H(\mathbb{D})$ with $\varphi_+(z) \neq 0$ for all $z \in \mathbb{D}$ and

$$\varphi(\zeta) = |\varphi_+(\zeta)|^2 \quad \text{for almost all } \zeta \in \mathbb{T} \;. \tag{10}$$

B. Non-computable and computable projections

We show that there exist spectral densities $\varphi_{\boldsymbol{x}}$ such that P_+ shows a behavior as in Row 4 of Table I. This follows from results in [20] which are restated as a single statement.

Proposition V.2: *There exists a wss stochastic process $\boldsymbol{x}$ with spectral density $\varphi_{\boldsymbol{x}} \in \mathcal{C}_{\mathrm{c}}(\mathbb{T})$ that is continuously differentiable with $\varphi'_{\boldsymbol{x}} \in \mathcal{C}_{\mathrm{c}}(\mathbb{T})$ and that satisfies Szegö's condition* (4) *such that $h_{\mathrm{opt}} = \mathrm{P}_+(1)$ is not a computable function and $d(1, L_+^2(\varphi_{\boldsymbol{x}}))$ is not a computable number.*

So even with additional smoothness assumptions on $\varphi_{\boldsymbol{x}}$ (continuously differentiable with $\varphi'_{\boldsymbol{x}} \in \mathcal{C}_{\mathrm{c}}(\mathbb{T})$), the MMSE $d(1, L_+^2(\varphi_{\boldsymbol{x}}))$ might not be computable and the projection operator P_+ may map the constant function 1 onto an h_{opt} that is not computable. Nevertheless, choosing other spectral densities, also a positive result [13] can be derived.

Proposition V.3: *Let $\boldsymbol{x}$ be a wss stochastic process with spectral density $\varphi_{\boldsymbol{x}}$ that satisfies*

1) *$\varphi_{\boldsymbol{x}}$ is continuously differentiable*
2) *$\varphi_{\boldsymbol{x}}$ and $\varphi'_{\boldsymbol{x}}$ are both computable continuous functions*
3) $\min_{\zeta\in\mathbb{T}} \varphi_{\boldsymbol{x}}(\zeta) = c_0(\varphi_{\boldsymbol{x}}) > 0$

then $h_{\mathrm{opt}} = \mathrm{P}_+(1) \in \mathcal{C}_{\mathrm{c}}(\mathbb{T})$ and $d(1, L_+^2(\varphi_{\boldsymbol{x}})) \in \mathbb{R}_{\mathrm{c}}$.

Proposition V.3 provides sufficient conditions on the spectral density $\varphi_{\boldsymbol{x}}$ such that the transfer function of the optimal prediction filter is guaranteed to be a computable continuous function and such that the MMSE $d(1, L_+^2(\varphi_{\boldsymbol{x}}))$ is a computable number.

C. Complexity blowup for computing h_{opt}

Even if $h_{\mathrm{opt}} = \mathrm{P}_+(1)$ is computable, its computation might be very complex. Therefore we ask for the complexity of computing h_{opt}, or equivalently for the complexity of computing the filter coefficients $h_{\mathrm{opt}}[n]$, $n \in \mathbb{N}$ of the impulse response $\{h_{\mathrm{opt}}[n]\}_{n\in\mathbb{N}}$. Assume $\varphi_{\boldsymbol{x}}$ satisfies the conditions of Proposition V.3 and $\varphi_{\boldsymbol{x}}$ is additionally polynomial-time computable. Is it then true that every filter coefficients $h_{\mathrm{opt}}[n]$, $n \in \mathbb{N}$ is a polynomial-time computable number? The following theorem gives a negative answer for the first filter coefficient.

Theorem V.4: *Assume $FP_1 \neq \#P_1$ and let $K \geq 2$ be arbitrary. There exists a spectral density φ_* with*

1) *φ_* is K-times continuously differentiable*
2) *φ_* and all of its derivatives up to order K are polynomial-time computable continuous functions*
3) $\min_{\zeta\in\mathbb{T}} \varphi_*(\zeta) = c_0(\varphi_*) > 0$

but such that the first filter coefficient $h_{\mathrm{opt}}[1]$ of the optimal prediction filter (7) *is not polynomial-time computable.*

Sketch of Proof: Let φ be an arbitrary spectral density satisfying Szegö's condition (4). Then the optimal causal prediction filter $h_{\mathrm{opt}}(z)$, given by (8), is analytic for $z \in \mathbb{D}$ and satisfies $h_{\mathrm{opt}}(0) = 0$. So h_{opt} can be written as in

(7). Defining $g(z) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\varphi(\mathrm{e}^{\mathrm{i}\theta})\frac{\mathrm{e}^{\mathrm{i}\theta}+z}{\mathrm{e}^{\mathrm{i}\theta}-z}\,\mathrm{d}\theta$, the spectral factor (9) of φ can be written as $\varphi_+(z) = \exp[g(z)]$. Therewith, (8) becomes

$$\begin{aligned} h_{\mathrm{opt}}(z) &= 1 - \exp\left[g(0) - g(z)\right] = 1 - \exp\left[g_1(z)\right] \\ &= -\sum_{m=1}^{\infty} \tfrac{1}{m!}\left[g_1(z)\right]^m , \quad z \in \mathbb{D} , \end{aligned} \tag{11}$$

with $g_1(z) = g(0) - g(z)$, and where we used the power series expansion of the exponential function to get the last equality. Since $g_1(0) = 0$, the power series expansion of g_1 has the form $g_1(z) = \sum_{n=1}^{\infty} g_1[n]\, z^n$ for $z \in \mathbb{D}$. Now, we insert this power series of g_1 into (11) and compare the coefficients with the power series (7). This shows that $h_{\mathrm{opt}}[1] = -g_1[1]$. Moreover, the first coefficient of the power series expansion of g_1 is given by $g_1[1] = \frac{1}{2\pi}\int_{-\pi}^{\pi} g_1(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}\theta}\,\mathrm{d}\theta$. Inserting the definition of g_1, using that $g = \log\varphi$, and applying (10) yields

$$h_{\mathrm{opt}}[1] = -g_1[1] = \tfrac{1}{4\pi}\int_{-\pi}^{\pi} \log\varphi(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{e}^{-\mathrm{i}\theta}\,\mathrm{d}\theta \;. \tag{12}$$

We choose φ to be an even function, i.e. $\varphi(\mathrm{e}^{-\mathrm{i}\theta}) = \varphi(\mathrm{e}^{\mathrm{i}\theta})$ for all $\theta \in [-\pi,\pi]$. Then also $\log\varphi$ is even and (12) becomes

$$h_{\mathrm{opt}}[1] = \tfrac{1}{4\pi}\int_{-\pi}^{\pi} \log\varphi(\mathrm{e}^{\mathrm{i}\theta})\,\cos(\theta)\,\mathrm{d}\theta \;. \tag{13}$$

Next, we construct an even, strictly positive function q_1 that has Properties 1) and 2) of the theorem, such that $q_1(\mathrm{e}^{\mathrm{i}\theta}) = 0$ for all $\theta \in [-\pi, -\pi/2] \cup [\pi/2, \pi]$, and so that

$$\tfrac{1}{4\pi}\int_{-\pi}^{\pi} q_1(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta \tag{14}$$

is not polynomial-time computable. The construction of this q_1 is very similar to the construction in [21, Proof of Theorem VI.1]. Based q_1, we define a second function by

$$q_2(\mathrm{e}^{\mathrm{i}\theta}) = \begin{cases} \frac{q_1(\mathrm{e}^{\mathrm{i}\theta})}{\cos(\theta)} & \text{for} \quad \theta \in [-\pi/2, \pi/2] \\ 0 & \text{for} \quad \theta \in [-\pi, -\pi/2) \cup (\pi/2, \pi] \end{cases} .$$

Since $1/\cos(\theta)$, $\theta \in [-\pi/2, \pi/2]$ is polynomial-time computable and because the product of two polynomial-time computable functions is again polynomial-time computable, it follows that q_2 is polynomial-time computable. Because of (14), it follows that

$$\tfrac{1}{4\pi}\int_{-\pi}^{\pi} q_2(\mathrm{e}^{\mathrm{i}\theta})\,\cos(\theta)\,\mathrm{d}\theta \tag{15}$$

is not polynomial-time computable. Now we distinguish two cases to define an even function q that is polynomial-time computable but such that its zeroth and first Fourier coefficient is not polynomial-time computable.

a) Assume $\frac{1}{4\pi}\int_{-\pi}^{\pi} q_2(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta$ is polynomial-time computable. Then we define $q(\zeta) = q_1(\zeta) + q_2(\zeta)$, $\zeta \in \mathbb{T}$. By this assumption and since (14) and (15) are not polynomial-time computable, it follows that both

$$\tfrac{1}{4\pi}\int_{-\pi}^{\pi} q(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta \quad \text{and} \quad \tfrac{1}{4\pi}\int_{-\pi}^{\pi} q(\mathrm{e}^{\mathrm{i}\theta})\,\cos(\theta)\,\mathrm{d}\theta \tag{16}$$

are not polynomial-time computable.

b) Assume $\frac{1}{4\pi}\int_{-\pi}^{\pi} q_2(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta$ is not polynomial-time computable. Then q_2 already has the desired properties and we simply set $q(\zeta) = q_2(\zeta)$, $\zeta \in \mathbb{T}$.

So we constructed a polynomial-time computable continuous function q such that the zeroth and the first Fourier coefficients of q are not polynomial-time computable. Therewith, we finally define $\varphi_*(\zeta) = \exp\left[q(\zeta)\right]$, $\zeta \in \mathbb{T}$. Since q has Properties 1) and 2) of the theorem and because q is non-negative, it easily follows that φ_* has Properties 1)–3) of the theorem. Moreover, (13) shows that $h_{\mathrm{opt}}[1]$ is equal to the first Fourier coefficient of q, which is not polynomial-time computable by our construction. ■

D. Application: FIR approximation

Let (7) be the transfer function of an optimal prediction filter associated with a wss stochastic process $\boldsymbol{x}$ with spectral density $\varphi_{\boldsymbol{x}}$. Assume a sequence $\{h_N\}_{N\in\mathbb{Z}}$ of FIR approximation filters is given, where h_N has the form

$$h_N(z) = \sum_{n=1}^{N} h_N[n]\, z^N , \qquad z \in \mathbb{D} ,$$

with $\lim_{N\to\infty} h_N[n] = h_{\mathrm{opt}}[n]$ for any fixed $n \in \mathbb{N}$, i.e. the coefficients of the FIR approximation converge to the coefficients of h_{opt}. Assume further that for $N \in \mathbb{N}$, all coefficients $h_N[n]$, $n = 1, \dots, N$ are polynomial-time computable numbers. Then h_N is a polynomial-time computable continuous function on $\mathbb{T}$.

Now we ask for a sequence $\{N(M)\}_{M\in\mathbb{N}} \subset \mathbb{N}$ of FIR filter lengths such that

$$\left|h_{\mathrm{opt}}[1] - h_{N(M)}[1]\right| < 2^{-M} , \quad \text{for all } M \in \mathbb{N} \;. \tag{17}$$

Thus, for an arbitrary precision $M \in \mathbb{N}$, we want to determine the filter degree $N = N(M)$ that guarantees that the approximation $h_{N(M)}[1]$ of $h_{\mathrm{opt}}[1]$ is sufficently close in the sense of (17). Since $h_{\mathrm{opt}}[1]$ is a computable number, the sequence $\{N(M)\}_{M\in\mathbb{N}}$ is a computable sequence and we ask whether it is polynomial-time computable.

Corollary V.5: *Let $\{h_N\}_{N\in\mathbb{Z}}$ be an arbitrary sequence of FIR approximation filters of h_{opt}. Then every sequence $\{N(M)\}_{M\in\mathbb{N}} \subset \mathbb{N}$ of FIR filter lengths that satisfies* (17) *is not a polynomial-time computable sequence.*

Proof: If $\{N(M)\}_{M\in\mathbb{N}}$ would be polynomial-time computable, then $\left\{h_{N(M)}[1]\right\}_{M\in\mathbb{N}}$ would be a polynomial-time computable sequence of polynomial-time computable numbers. Consequently its limit $h_{\mathrm{opt}}[1]$ would be a polynomial-time computable number, which contradicts the statement of Theorem V.4. ■

So even if (for a given $N \in \mathbb{N}$) the individual FIR filter coefficients $h_N[n]$ are polynomial-time computable, there are cases where it is impossible to find in polynomial time the degree $N \in \mathbb{N}$ that guarantees the error bound (17).

VI. Complexity of Polynomial Approximation

The non-computability behavior of the orthogonal projection P_+ from Section IV can not occur if the range $\mathcal{U}$ of the projection operator is a finite dimensional subspace (of trigonometric polynomials). Because then $\mathrm{P}_+ f$ will always be computable as long as f is a computable function. In this case, one may suppose that if f is very easy to compute then also $\mathrm{P}_+ f$ will be easy to compute. However, this section will show that this is not necessarily true.

We consider the problem of approximating continuous functions on $\mathbb{T}$ by trigonometric polynomials. Let $f \in \mathcal{C}_{\mathrm{c}}(\mathbb{T})$

be a computable continuous function on $\mathbb{T}$. Then f can be written as a trigonometric Fourier series in $L^2(\mathbb{T})$

$$f(\mathrm{e}^{\mathrm{i}\theta}) = \frac{a_0(f)}{2} + \sum_{n=1}^{\infty} a_n(f)\cos(n\theta) + b_n(f)\sin(n\theta)$$

with Fourier coefficients (1). For an arbitrary $N \in \mathbb{N}$, we consider the mapping $\mathrm{P}_N : \mathcal{C}_\mathrm{c}(\mathbb{T}) \to \mathcal{P}_N$ given by

$$(\mathrm{P}_N f) = \frac{a_0(f)}{2} + \sum_{n=1}^{N} a_n(f)\cos(nt) + b_n(f)\sin(nt)\ .$$

It is well known that $\mathrm{P}_N f$ is the best (with respect to $L^2(\mathbb{T})$ norm) approximation of f by an element from $\mathcal{P}_N$, i.e.

$$\inf_{p\in\mathcal{P}_N} \|f-p\|^2_{L^2(\mathbb{T})} = \|f - \mathrm{P}_N f\|^2_{L^2(\mathbb{T})} = d(f,\mathcal{P}_N)$$

and where the distance $d(f,\mathcal{P}_N)$ is again the MMSE of approximating f by a trigonometric polynomial of degree N. So P_N is an orthogonal projection from $L^2(\mathbb{T})$ onto $\mathcal{P}_N$.

It is clear that $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ always implies $\mathrm{P}_N f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$, because all Fourier coefficients of $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ are computable numbers and $\cos(n\cdot)$ and $\sin(n\cdot)$ are computable continuous functions. So since $\mathrm{P}_N f$ is already computable, we may ask now for the complexity of computing $\mathrm{P}_N f$. In particular, we ask whether $\mathrm{P}_N f$ is polynomial-time computable for every polynomial-time computable function $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$. The answer is given by the following theorem.

Theorem VI.1: *If $FP_1 \neq \#P_1$ then there exists a polynomial-time computable functions $f_2 \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ such that for every $N \in \mathbb{N}$, $\mathrm{P}_N f_2$ is not a polynomial-time computable continuous function.*

As an immediate consequence of the following proof, we obtain that for every $N \in \mathbb{N}$, the set of $2N$ equidistant sampling values of f_2 are not polynomial-time computable.

Corollary VI.2: *If $FP_1 \neq \#P_1$ then there exists a polynomial-time computable function $f_2 \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ such that for every $N \in \mathbb{N}$, it is not possible to compute the set*

$$\left\{ (\mathrm{P}_N f_2)\left(k\tfrac{2\pi}{2N+1}\right) \ :\ k = 1,2,\ldots,2N \right\}. \qquad (18)$$

in polynomial time.

Proof: It is known (cf., [11], [22], [23]) that there exists a polynomial-time computable function $f \in \mathcal{C}_\mathrm{c}(\mathbb{T})$ such that its integral $a_0(f_2) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_2(t)\,\mathrm{d}t$ is not polynomial-time computable, and we refer to [21, Thm. VI.1] for a more explicit construction of such an f_2. Contrary to the statement of the theorem, assume that there exists an $N \in \mathbb{N}$ such that $\mathrm{P}_N f_2$ is a polynomial-time computable continuous function. This implies in particular that all values in the set (18) are polynomial-time computable. Then the zeroth Fourier coefficient of f_* can be expressed [24] as

$$a_0(f_2) = \frac{1}{2N+1}\sum_{k=0}^{2N} (\mathrm{P}_N f_2)\left(k\tfrac{2\pi}{2N+1}\right)\ .$$

Consequently $a_0(f_2)$ is a polynomial-time computable. However, f_2 was constructed such that $a_0(f_2)$ is not polynomial-time computable. So we obtained at a contradiction proving that $\mathrm{P}_N f_2$ is not polynomial-time computable. ■

VII. Summary

We have investigated the complexity of computing orthogonal projections and presented examples of Hilbert spaces $\mathcal{H}$, $\mathcal{U}$ such that a computable $f \in \mathcal{H}$ does not imply that the orthogonal projection $\mathrm{P}_\mathcal{U} f$ and/or the distance $d(f,\mathcal{U})$ is computable. It was also shown that even if $\mathrm{P}_\mathcal{U} f$ is computable, the computation of $\mathrm{P}_\mathcal{U} f$ may show complexity blowup, i.e. $\mathrm{P}_\mathcal{U}$ may map an f of low computational complexity onto an $\mathrm{P}_\mathcal{U} f$ that has high computational complexity. The given examples came from very common applications like causal projection, polynomial approximation, and Wiener filtering.

References

[1] J. S. Meditch, "Orthogonal projection and discrete optimal linear smoothing," *J. SIAM Control*, vol. 5, no. 1, pp. 74–89, 1967.

[2] F. Ticozzi, L. Zuccato, P. D. Johnson, and L. Viola, "Alternating projections methods for discrete-time stabilization of quantum states," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 819–826, Mar. 2018.

[3] Y. C. Eldar, A. V. Oppenheim, and D. Egnor, "Orthogonal and projected orthogonal matched filter detection," *Signal Processing*, vol. 84, no. 4, pp. 677–693, Apr. 2004.

[4] P. Astrid, S. Weiland, K. Willcox, and T. Backx, "Missing point estimation in models described by proper orthogonal decomposition," *IEEE Trans. Autom. Control*, vol. 53, no. 10, pp. 2237–2251, Nov. 2008.

[5] C. K. Chui and G. Chen, *Kalman Filtering*. Berlin: Springer, 1999.

[6] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[7] M. Fält and P. Giselsson, "Line search for generalized alternating projections," in *Proc. American Control Conference (ACC)*, Seattle, WA, USA, May 2017, pp. 4637–4642.

[8] K. Zhang and K. H. Johansson, "An algorithm for computing explicit expressions for orthogonal projections onto finite-game subspaces," in *IEEE Conf. Decision and Control (CDC)*, Miami, FL, USA, Dec. 2018, pp. 3196–3202.

[9] M. B. Pour-El and J. I. Richards, *Computability in Analysis and Physics*. Berlin: Springer-Verlag, 1989.

[10] K. Weihrauch, *Computable Analysis*. Berlin: Springer-Verlag, 2000.

[11] K.-I. Ko, *Complexity Theory of Real Functions*. Basel: Birkhäuser, 1991.

[12] S. Arora and B. Barak, *Computational Complexity: A Modern Approach*. Cambridge: University Press, 2009.

[13] H. Boche, V. Pohl, and H. V. Poor, "Characterization of the complexity of computing the minimum mean square error of causal prediction," *IEEE Trans. Inf. Theory*, vol. 70, no. 9, pp. 6627–6638, Sep. 2024.

[14] H. Boche and V. Pohl, "On the algorithmic solvability of spectral factorization and applications," *IEEE Trans. Inf. Theory*, vol. 66, no. 7, pp. 4574–4592, Jul. 2020.

[15] J. L. Doob, *Stochastic Processes*. New York: John Wiley, 1953.

[16] Y. A. Rozanov, *Stationary Random Processes*. San Francisco: Holden-Day, 1967.

[17] A. N. Shiryaev, *Probability*, 2nd ed. New York: Springer, 1996.

[18] N. Bingham, "Szegö's theorem and its probabilistic descendants," *Probab. Surv.*, vol. 9, pp. 287–324, 2012.

[19] H. Boche, V. Pohl, and H. V. Poor, "Recent progress in computability for prediction and Wiener filter theory," *Trans. A. Razmadze Math. Inst.*, vol. 176, no. 3, pp. 323–344, Dec. 2022.

[20] H. Boche, V. Pohl, and H. V. Poor, "The Wiener theory of causal linear prediction is not effective," in *Proc. 62nd IEEE Conf. on Decision and Control (CDC)*, Singapore, Dec. 2023, pp. 8223–8228.

[21] H. Boche and V. Pohl, "Complexity blowup in simulating analog linear time-invariant systems on digital computers," *IEEE Trans. Signal Process.*, vol. 69, pp. 5005–5020, Aug. 2021.

[22] H. M. Friedman, "The computational complexity of maximization and integration," *Adv. Math.*, vol. 53, no. 1, pp. 80–98, Jul. 1984.

[23] K.-I. Ko, "Approximation to measurable functions and its relation to probabilistic computation," *Ann. Pure Appl. Logic*, vol. 30, no. 2, pp. 173–200, Feb. 1986.

[24] A. Zygmund, *Trigonometric Series*, 3rd ed. Cambridge, UK: Cambridge University Press, 2002.