A Simple Algorithm for Approximation by Nomographic Functions

Steffen Limmer*, Jafar Mohammadi*, Sławomir Stańczak*†

* Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Einsteinufer 37, 10587 Berlin, Germany.
† Fachgebiet Informationstheorie und theoretische Informationstechnik, Technische Universität Berlin, Einsteinufer 25, 10587 Berlin, Germany.

Abstract

This paper introduces a novel algorithmic solution for the approximation of a given multivariate function by a nomographic function that is composed of a one-dimensional continuous and monotone outer function and a sum of univariate continuous inner functions. We show that a suitable approximation can be obtained by solving a cone-constrained Rayleigh-Quotient optimization problem. The proposed approach is based on a combination of a dimensionwise function decomposition known as Analysis of Variance (ANOVA) and optimization over a class of monotone polynomials. An example is given to show that the proposed algorithm can be applied to solve problems in distributed function computation over multiple-access channels.

Index Terms

Distributed computation, nomographic approximation, compute-and-forward, multiple-access channel

I. INTRODUCTION

Due to the new technological trends such as Internet-of-Things, a designer of future wireless networks will face a dramatic growth of wirelessly connected devices that will generate a huge amount of data. In order to facilitate such a massive connectivity, a paradigm shift in the network design and operation is necessary. In particular, special attention should be attached to the development of novel solutions for data fusion, aggregation and compression. Current approaches usually build upon a separation principle, according to which communication is separated from application needs. Thereby, wireless communication is based on access/multiplexing techniques, in which case individual data-streams are assigned communication resources on a user-by-user basis. In addition to being notoriously difficult to implement in decentralized networks, such techniques are in general highly inefficient for wireless sensor network (WSN) applications, where there is often no need for conveying and processing individual measurements. For instance, a fusion node in a WSN may only be interested in some pre-defined functions of the received data. Examples of such WSN applications include distributed detection, where the objective of a fusion node is to compute a confidence metric for the existence or absence of some phenomenon, and distributed classification, where the task consists in estimating some class labels. To meet the requirements of these applications, the fusion node in general does not need all the individual measurements, which can be processed while being transferred through the network: Communication and computation behind the underlying application are merged.

The idea of joint computation and communication is not new but, recently, there has been an increasing interest in communication strategies that exploit the wireless channel for computation [1], [2]. Such strategies have a potential for huge performance gains expressed in terms of efficiency, complexity and signaling overhead. A key ingredient thereby is that the function to be computed has a suitable nomographic representation that is used to match the process of function computation to the communication channel [1], [3]. The theoretical analysis of functions that can be written in a nomographic form has a long history, which dates back to Kolmogorov [4], Sprecher [5], [6] and Buck [7]. The authors in [5], [6] showed that every function has a nomographic representation, when the outer function can be discontinuous (e.g. a space-filling curve). Implementing such functions in digital signal processing systems, especially when based on space-filling curves as analyzed in [6], is prohibitive and leads to designs that are notoriously susceptible to noise.

1Throughout this paper, we refer to these devices as sensors.
A.2

In doing so, we assume the following:

1. that (in some sense)
2. Therefore, given some function
   inner functions and the outer function is not amenable to implementation on state-of-the-art hardware technologies.

In a nutshell, the proposed method can be used to approximate a class of multivariate functions by means of some nomographic approximations that are composed of a continuous and monotone outer function and continuous inner functions. In addition to being optimal with respect to some approximation metric, the inner and outer functions are more robust to noise and can be implemented in digital signal processing systems in a relatively easy way.

The proposed method can be used in combination with the approach of [2] to approximate a class of multivariate functions using only a single access to the communication channel without the need of transmitting individual measurements. The approach of [2] exploits the superposition property of the wireless channel for computation of some nomographic functions and is of high practical relevance because it is robust under practical impairments such as the lack of synchronization; in addition, there is little need for coordination between different sensor. The recent work [3] demonstrates a hardware implementation of the method presented in [2].

### A. Notation

Scalars, vectors, matrices and sets are denoted by lowercase $a$, bold lowercase $a$, bold uppercase $A$ and calligraphic letters $A$, respectively. $(\cdot)^T$, $f \circ g$ and $f^{-1}$ stand for transpose, function composition and function inverse. The sets of natural and real numbers are denoted by $\mathbb{N}$ and $\mathbb{R}$, while $0$ is used to denote the vector of all zeros, where the size will be clear from the context. We use $\mathcal{L}_p(\mathcal{X}^K)$ to refer to the space of $p$-integrable ($1 \leq p < \infty$) real functions, $\mathcal{P}_{K,D}(\mathcal{X}^K)$ to the space of square integrable real polynomials of $K$ variables of degree at most $D$ in each variable, $\mathcal{C}(\mathcal{X}^K)$ to the space of continuous functions, and $\mathcal{N}_C(\mathcal{X}^K)$ to the space of nomographic functions, i.e. functions that can be represented in the form $\psi(\sum_{k=1}^K \varphi_k(x_k))$, and all spaces defined on $\mathcal{X}^K := \mathcal{X} \times \ldots \times \mathcal{X} \subseteq \mathbb{R}^K$, respectively. If in addition the outer and inner functions $\psi$ and $\varphi_k$ fulfill $\varphi_k \in \mathcal{C}(\mathcal{X}) : \mathcal{X} \to \Omega_k \subseteq \mathbb{R}$ and $\psi \in \mathcal{C}(\Omega') : \Omega' \to \Omega$, we denote the corresponding space by $\mathcal{N}_C(\mathcal{X}^K)$.

## II. System Model, Problem Statement and Theoretical Framework

We consider a network consisting of $K \in \mathbb{N}$ sensors indexed by the set $\mathcal{K} := \{1, \ldots, K\}$. The sensors observe measurements $x := [x_1, \ldots, x_K]^T \in \mathcal{X}^K$ and the task of the network is to compute or approximate a multivariate function

$$f : \mathcal{X}^K \to \Omega \subseteq \mathbb{R} \tag{1}$$

at some pre-selected fusion node. The underlying computation and communication scenario is illustrated in Fig. 1.

### A. Problem statement

It was shown in [2], [3] that a nomographic representation of some given function admits an efficient reconstruction or estimation of this function over the wireless channel. However, as already explained in the introduction, although every function has a nomographic representation of the form $\psi(\sum_{k=1}^K \varphi_k(x_k))$, the general construction of the inner functions and the outer function is not amenable to implementation on state-of-the-art hardware technologies. Therefore, given some function $f$ defined by (1), the problem is to find a suitable nomographic representation such that (in some sense)

$$f(x) \approx \psi\left(\sum_{k=1}^K \varphi_k(x_k)\right), \quad \forall x \in \mathcal{X}^K. \tag{2}$$

In doing so, we assume the following:

A.1 $\psi$ is monotone continuous,
A.2 $\{\varphi_k\}_{k \in \mathcal{K}}$ are continuous,
which is based on practical considerations concerning noise robustness and implementability on digital signal processing systems. To emphasize the importance of Assumptions A.1 and A.2 on inner functions $\varphi_k, k = 1 \ldots K$, and outer function $\psi$, let us review some previously known results from literature.

**Fact 1. Representation by nomographic functions**

1) Let $\varphi_k$ be monotone increasing, $\psi$ be possibly discontinuous, then we have \cite{5}, \cite{6}

$$f(x) = \psi\left(\sum_{k=1}^{K} \varphi_k(x_k)\right), \forall f \in C(\mathcal{X}^K). \tag{3}$$

2) Let $\varphi_k \in C(\mathbb{R}), \psi \in C(\mathbb{R})$, then the following holds:
   - $\mathcal{N}_C(\mathcal{X}^K)$ is a nowhere dense subset of $C(\mathcal{X}^K)$ \cite{6}
   - $f(x) = \sum_{i=1}^{2K+1} \psi_i\left(\sum_{k=1}^{K} \varphi_k^{(i)}(x_k)\right)$, i.e. every function can be written as a sum of at most $2K + 1$ nomographic functions \cite{5}.

Due to Assumptions A.1-2, we consider the second case and impose an additional constraint of a single function $\psi$.

**B. Theoretical framework: Analysis of variance**

To establish our results, we resort to a general framework for a dimensionwise decomposition of a function $f(x)$ into a sum of lower-dimensional terms. More precisely, we consider the Analysis of Variance (ANOVA) framework \cite{9}, \cite{10}, which has also been considered in the context of many other applications, ranging from chemistry and finance to statistics (see e.g. \cite{10}). The goal of this framework is to decompose a function $f \in L_2(\mathcal{X}^K)$ into a sum of $2^K$ functions $f_S$ that are mutually orthogonal w.r.t. the inner product $\langle f, g \rangle = \int_{\mathcal{X}^K} f(x) \cdot g(x) \, dx$. Here, the function $f$ is decomposed into a sum of lower dimensional functions

$$f(x) = \sum_{S \subseteq K} f_S(x_S), \tag{4}$$

where each function $f_S$ only depends on a subset of variables indexed by the set $S \subseteq K$ and the sum ranges over the power set of $K$. The algorithm to obtain the ANOVA decomposition for a given function $f \in L_2(\mathcal{X}^K)$ is given by Alg. 1.

**Remark 1.** Despite its simple form, the reader should note that a numerical implementation of Alg. 1 is in general not trivial. In fact, the computation of all $2^K$ terms for a full decomposition becomes impracticable for moderate values of $K$ and the involved high-dimensional integrals need to exist and be well-defined. In addition to these requirements, the integrals might still be hard to obtain in analytical form and numerical approximation methods might be necessary. However, for some classes of functions including multivariate polynomials $f \in P_{K,D}(\mathcal{X}^K)$, we can easily obtain a truncated decomposition in a closed form up to moderate values of $K$.

\footnote{Notice that nomographic approximation is used in this paper to refer to a nomographic representation of some function that approximates the function of interest in some pre-defined sense.}

\footnote{Due to the structure of the nomographic representation, every ordinary sensor can act as fusion node.}
The conditions for which a truncated decomposition provides a good approximation of the original function are made precise in the following definition.

**Definition 1.** A function $f$ is said to be of order $d$ if
\[
    f(x) = \sum_{|S|\leq d} f_S(x_S) \Leftrightarrow \sum_{|S|\leq d} \sigma_S^2 = \sigma^2
\]
and of effective superposition dimension $d$ if, for some given sufficiently small $\varepsilon > 0$, there holds
\[
    \sum_{|S|\leq d} \sigma_S^2 \geq (1-\varepsilon)\sigma^2
\]
where $f_S(x_S)$, $\sigma_S^2$ and $\sigma^2$ are obtained by Alg. 1. If a function is of effective superposition dimension 1 (given $\varepsilon > 0$), then we say that the function has a nomographic approximation.

In other words, if the variances of all terms up to order $d$ amounts to almost all of the total variance (e.g. $\varepsilon = 10^{-3}$), the function is well approximated by a function of order $d$. Note that a nomographic representation requires that $d = 1$ (see also [1] below). In turn, this condition ensures that there is no need to exchange information among different sensors as each of them only needs to compute a univariate function of its own measurement (see Fig. 1(b)). The problem is however that if $d$ is too small, then (6) cannot be satisfied for some given $\varepsilon > 0$. So there is a tradeoff between amenability to distributed implementation and approximation accuracy. The former requires small values of $d$, while the latter improves with increasing $d$.

### III. OUTLINE OF THE PROPOSED APPROACH

We start this section with a brief outline of the rationale behind the proposed algorithm. Suppose that $f$ has a nomographic approximation in accordance with Def. 1 so that
\[
    f(x) = (\psi \circ \varphi(x))(x) \approx \psi \left( \sum_{|S|\leq 1} \varphi_S(x_S) \right).
\]
Suppose further that $\psi \in \mathcal{C}(\Omega')$ is one-to-one and onto $\Omega \subset \mathbb{R}$, i.e., it admits a unique functional inverse $\psi^{-1}: \Omega \to \Omega'$ that is also continuous. Then, we can write
\[
    f(x) = (\psi \circ \psi^{-1} \circ f)(x)
\]
which is of the desired nomographic form if
\[
    \varphi(x) = \psi \left( \sum_{|S|\leq 1} \varphi_S(x_S) \right),
\]
or, equivalently, if $\varphi(x) := (\psi^{-1} \circ f)(x)$ is of order 1 in the sense of Def. 1. Similarly, given some $\varepsilon > 0$, a nomographic approximation is obtained if $\varphi(x)$ is of effective superposition dimension 1. The inner approximation of $\varphi(x) \approx \sum_{|S|\leq 1} \varphi_S(x_S)$ is optimal in an $\mathcal{L}_2$ sense [12]. This is summarized in the following Lemma.

---

4For notational convenience, we include the constant function $\varphi_2$ to the set of inner functions. Note that the nomographic representation is retained and the right-hand side of (7) is equivalent to the right-hand side of (2).

5The reader should note however, that due to the usually nonlinear transformation by the outer function $\psi$, this $\mathcal{L}_2$ optimality does not need to hold for the overall approximation error.
Lemma 1. Let \( f : \mathcal{X}^K \to \Omega \in \mathcal{C}(\mathcal{X}^K) \). Then, if \( \psi^{-1} : \Omega \to \Omega' \in \mathcal{C}(\Omega) \) and \( (\psi^{-1} \circ f)(x) \in \mathcal{L}_2(\mathcal{X}^K) \) is such that \( \varphi(x) := (\psi^{-1} \circ f)(x) \) is of order 1, we obtain a nomographic representation \( f = \psi(\sum_{|s| \leq 1} \varphi_s(x_s)) \) by the ANOVA decomposition of \( \varphi(x) \) and the identity \( \psi \circ \psi^{-1} \circ f)(x) = (\psi \circ \varphi)(x) = \psi(\sum_{|s| \leq 1} \varphi_s(x_s)) \). Similarly, if \( (\psi^{-1} \circ f)(x) \) is of effective superposition dimension 1 (given \( \varepsilon > 0 \)), we obtain a nomographic approximation, where the approximation is optimal in an \( \mathcal{L}_2 \) sense (or, equivalently, the minimum variance sense) of the inner approximation problem.

A. A class of monotone polynomials

To obtain a class of continuous inverse outer functions \( \psi^{-1} \), we consider a class of polynomials known as Bernstein polynomials:

Lemma 2. [13] Let \( p(\xi) \in \mathcal{P}_{1,D-1} := \sum_{d=0}^{D-1} z_d \xi^d \) be a real polynomial of degree \( D - 1 \) defined on \( [0,1] \) with real coefficients \( z := [z_1, \ldots, z_D] \). Let \( \tilde{M} \in \mathbb{R}^{D \times D} \) be a lower triangular matrix with entries given by \( [\tilde{M}]_{i,j} = \left( \frac{i}{j} \right)^{(D-1)} \) \( \forall i \geq j \), \( [\tilde{M}]_{i,j} = 0 \forall i < j \). Then \( \min_i [\tilde{M} z]_i \leq p(\xi) \leq \max_i [\tilde{M} z]_i \).

To obtain a suitable class of monotone and continuous functions we proceed by integration of the polynomial \( p(\xi) \) w.r.t. \( \xi \), which is formalized by the following proposition:

Proposition 1. Let \( p(\xi) \in \mathcal{P}_{1,D} \) be a polynomial in \( \xi \) of degree \( D \). Then, \( p(\xi) \) is monotone and continuous on \( \Omega := [0,1] \) if it holds that \( p \in \mathcal{V}(z,c) \subset \mathcal{P}_{1,D} \) with

\[
\mathcal{V}(z,c) := \left\{ \sum_{d=0}^{D} z_d \xi^d + c \mid M z \geq 0, \ z \neq 0 \right\},
\]

with

\[ M := \tilde{M} \text{diag} \left( [1,2,3,\ldots,D]^T \right). \]

Proof: The proof follows from Lemma 2 by integration and comparison of terms.

The reader should note, that \( \mathcal{V}(z,c) \) defines a polyhedral cone w.r.t. the variable \( z \) where we have excluded the origin to prevent trivial (constant) polynomials.

B. Nomographic Approximation by cone-constrained Rayleigh quotient optimization

It remains to study the structure of the inner approximation problem. To this end, let \( \psi^{-1} \in \mathcal{V}(z,c) \) with domain \([0,1]\). Then we can establish the following results on \( \sigma^2 \) and \( \sigma_k^2 \).

Lemma 3. Let \( \psi^{-1} \in \mathcal{P}_{1,D} \), \( \varphi := (\psi^{-1} \circ f) \in \mathcal{L}_2([0,1]^K) \) and \( f : [0,1]^K \to [0,1] \). Then, it holds that \( \sigma^2 = \int_{[0,1]^K} \varphi^2(x) \ dx - \left( \int_{[0,1]^K} \varphi(x) \ dx \right)^2 \) can be written in quadratic form \( \sigma^2 = z^T B z \) with \( B := B^{(1)} - b^{(2)} b^{(2),T} \) and

\[
[B]_{i,j}^{(1)} := \int_{[0,1]^K} f(x)^{i+j} \ dx, \ b_i^{(2)} := \int_{[0,1]^K} f(x)^i \ dx
\]

which is independent of the chosen constant \( c \).

Proof: The proof is deferred to Appendix A.

Lemma 4. Let \( \psi^{-1}, \varphi \) and \( f \) be as in Lemma 3 and let the mixed integrals

\[
[A]_{i,j}^{(1)}(k) := \int_{[0,1]} \left( \int_{[0,1]^{K-1}} f(x)^{i} \ dx \right) \ dx_{\mathcal{K} \setminus k}
\]

exist and be finite. Then, we have \( \sigma_k^2 = z^T A_k z \ \forall k \in \mathcal{K} \) with

\[
A_k := A^{(1)}(k) - b^{(2)} b^{(2),T}.
\]
Due to the high complexity of solving (16) directly, we apply a technique known as semidefinite relaxation [14] to the nonconvex semidefinite program by neglecting the (nonconvex) rank constraint (16e) first and solving the resulting convex SDP. Then, a candidate solution to the original problem (16) can be obtained by \( z^* = \sqrt{\lambda_1} q_1 \), where \( \lambda_1 \) and \( q_1 \) denote the largest eigenvalue and eigenvector of \( Z^* \), respectively. The relaxation is tight, i.e. the solution of the SDR coincides with the solution to (16) if \( \text{rank}(Z^*) = 1 \). If the rank constraint is violated, the solution will in general be suboptimal. In this case, we apply a heuristic to obtain a suboptimal feasible solution \( z^* = M^{-1} (M\sqrt{\lambda_1} q_1)_+ \) (cf. (16d)). This heuristic is motivated by our simulation results, which show that applying the projection onto the constraint set after transformation by the matrix \( M \) yields a numerically much more stable solution compared to applying the projection directly by computing \( z^* = \arg\min_{\{z | Mz \geq 0\}} \|z - \sqrt{\lambda_1} q_1\|_2^2 \).

The resulting overall approximation algorithm is described in Alg. 2:\(^6\)

\begin{algorithm}
\caption{Approximation of \( f \) by \( \psi(\sum_{|S| \leq 1} \varphi_S(x_S)) \)}
\begin{enumerate}
  \item Compute \( A, B, M \) using Lemma 3 4 and 2.
  \item Compute \( z^* \) by solving SDR of (16);
  \item Compute ANOVA (Alg. 1) for \( \psi_{z^*}^{-1} \circ f \);
  \item Compute \( \psi \) using numerical inversion of \( \psi^{-1} \);
\end{enumerate}
\end{algorithm}

\(^6\)The corresponding MATLAB implementation will be made available upon acceptance on one of the authors websites.
$$\sigma_2^2 = 0 \quad \sigma_{(1)} = 0.0168 \quad \sigma_{(2)} = 0.0168$$

$$\sigma_{(1,2)} = 0.0043 \quad \sigma^2 = 0.038 \quad (\sigma_{(1)}^2 + \sigma_{(2)}^2) \sigma^{-2} = 0.88$$

| $z_1$, ..., $z_5$ | $z_6$, ..., $z_{10}$ | $z_{11}$, ..., $z_{15}$ | $z_{16}$, ..., $z_{20}$ |
|-----------------|-----------------|-----------------|-----------------|
| 1.2803          | -134.14         | 442644.0        | -366688.0       |
| -12.162         | 2637.0          | -697011.0       | 145299.0        |
| 72.975          | -21534.0        | 874766.0        | -37288.0        |
| -236.66         | 84667.0         | -862822.0       | 8220.6          |
| 334.42          | -222633.0       | 652977.0        | -247.38         |

TABLE I
DIRECT EVALUATION OF ALG. 1 FOR GIVEN TEST FUNCTION.

| $z^*$ obtained by Alg. 2 and $D = 20$. |
|----------------------------------------|
| $z^*$                                  |
| 1.2803                                 |
| -134.14                                |
| 442644.0                               |
| -366688.0                              |
| -12.162                                |
| 2637.0                                 |
| -697011.0                              |
| 145299.0                               |
| 72.975                                 |
| -21534.0                               |
| 874766.0                               |
| -37288.0                               |
| -236.66                                |
| 84667.0                                |
| -862822.0                              |
| 8220.6                                 |
| 334.42                                 |
| -222633.0                              |
| 652977.0                               |
| -247.38                                |

IV. NUMERICAL RESULTS

To evaluate the performance of the proposed algorithm, we evaluate Alg. 2 for an exemplary polynomial given by

$$f(x_1, x_2) = \frac{1}{9}(x_1 + x_1 x_2 + x_2)^2,$$

which admits a nomographic approximation using $D = 20$ and $\varepsilon = 10^{-3}$. The parameters of the obtained polynomial $\psi^{-1}$ and the resulting functions and approximation error are given in Table II and Fig. 2, where it can be seen in Fig. 2(f) that the resulting overall approximation error is bounded by $|f - \psi(\sum_{|S| \leq 1} \varphi_S)| \leq 6 \times 10^{-3}$. For comparison, direct analytical evaluation of Alg. 1 for the given test function, i.e., without applying the nonlinear transform $\psi^{-1}$, are given in Table I.

V. CONCLUSION

In this paper, we studied the problem of nomographic approximation with continuous monotone outer function and continuous inner functions, which we expect to be of high practical relevance due to the superposition property of wireless communication channels. By using a certain function class based on Bernstein polynomials, we obtain a nomographic approximation with prescribed properties with respect to a defined distortion metric. The optimized approximation is obtained by the maximization of a cone-constrained Rayleigh-quotient. Since the problem is nonconvex, we consider its semidefinite relaxation. Though a precise characterization of the class of functions approximable in nomographic form with prescribed error metric still remains an open problem, we can see some interesting applications of the presented results, ranging from distributed learning and optimization to compressed classification.

REFERENCES

[1] B. Nazer and M. Gastpar, “Computation over multiple-access channels,” IEEE Trans. Inf. Theory, vol. 53, no. 10, pp. 3498–3516, 2007.
[2] M. Goldenbaum and S Stanczak, “Robust analog function computation via wireless multiple-access channels,” IEEE Trans. Comm., 2013.
[3] M. Goldenbaum, H. Boche, and S. Stanczak, “Harnessing interference for analog function computation in wireless sensor networks,” IEEE Transactions on Signal Processing, vol. 61, no. 20, pp. 4893–4906, 2013.
[4] A. N. Kolmogorov, “The representation of continuous functions of many variables by superposition of continuous functions of one variable and addition,” Proc. Amer. Math. Soc., 1957.
[5] D. A Sprecher, “A representation theorem for continuous functions of several variables,” Proceedings of the American Mathematical Society, vol. 16, no. 2, pp. 200–203, 1965.
[6] D. A. Sprecher, “On computational algorithms for real-valued continuous functions of several variables,” Neural Netw., vol. 59, pp. 16–22, 2014.
[7] R. C. Buck, “Nomographic functions are nowhere dense,” Proc. Amer. Math. Soc., pp. 195–199, 1982.
[8] A. Kortke, M. Goldenbaum, and S. Stanczak, “Analog Computation Over the Wireless Channel: A Proof of Concept,” in Proc. IEEE Sensors, Valencia, Spain, Nov. 2014.
Fig. 2. Simulation results for Alg. 2 applied to the function $f = \frac{1}{9} (x_1 + x_1 x_2 + x_2)^2$.

[9] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems, vol. I: Linear Information*, 2008.
[10] F. Kuo, I. Sloan, G. Wasilkowski, and H. Wozniakowski, “On decompositions of multivariate functions,” *Math. Comp.*, vol. 79, no. 270, pp. 953–966, 2010.
[11] M. Griebel, *Sparse grids and related approximation schemes for higher dimensional problems*, 2005.
[12] S. Rahman, “Approximation errors in truncated dimensional decompositions,” *arXiv preprint arXiv:1310.6769*, 2013.
[13] G. T. Cargo and O. Shisha, “The bernstein form of a polynomial,” *J. Res. Nat. Bur. Stand. Sect. B*, vol. 70, pp. 79, 1966.
[14] Z. Luo, W. Ma, A. M.-C. So, Y. Ye, and S. Zhang, “Semidefinite relaxation of quadratic optimization problems,” *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, 2010.
APPENDIX A

PROOF OF $\sigma^2 = z^T B z$, INDEPENDENCE OF $c$

To show that $\sigma^2$ is independent of $c$ let $p \in \mathcal{P}_{1,D}$, $\varphi(x) := (p \circ f)(x) \in L_2(\mathcal{X}^K)$, $\mathcal{X}^K := [0,1]^K$ and $f := f(x)$.

$$
\sigma^2 = \int_{\mathcal{X}^K} \varphi(x)^2 \, dx - \left( \int_{\mathcal{X}^K} \varphi(x) \, dx \right)^2
$$

$$
\Delta_1 = \int_{\mathcal{X}^K} \left( \sum_{d=1}^D z_d f^d \right)^2 + 2c \sum_{d=1}^D z_d f^d + c^2 \, dx
$$

$$
\Delta_2 = \left( \int_{\mathcal{X}^K} \sum_{d=1}^D z_d f^d \, dx \right)^2 + \int_{\mathcal{X}^K} 2c \sum_{d=1}^D z_d f^d + c^2 \, dx
$$

$$
\sigma^2 = \int_{\mathcal{X}^K} \left( \sum_{d=1}^D z_d f^d \right)^2 \, dx - \left( \sum_{d=1}^D z_d \int_{\mathcal{X}^K} f^d \, dx \right)^2
$$

As $z$ is independent of $f$ it follows by comparison of terms that $\sigma^2 = z^T B z$ with $B := B^{(1)} - b^{(2)} b^{(2),T}$ and

$$
B_{ij}^{(1)} := \int_{[0,1]^K} f(x)^{i+j} \, dx, \quad \{i,j\} \in \{1, \ldots, D\}^2
$$

$$
b_i^{(2)} := \int_{[0,1]^K} f(x)^i \, dx, \quad i \in \{1, \ldots, D\}.
$$