Positive Quartic and Biquartic C^2-Splines and Their Applications to Two-Dimensional Probability Density Estimation

Daichi KITAHARA† Isao YAMADA†

†Department of Communications and Computer Engineering, Tokyo Institute of Technology

Abstract Spline function is a piecewise polynomial and has been widely used for interpolation and smoothing of observed two-dimensional data. In this paper, by using the sufficient condition, derived by Heß and Schmidt, for the nonnegativity of biquartic C^2-splines on square grid, we propose positive biquartic C^2-spline interpolation and smoothing on square grid for estimation of non-negative and twice continuously differentiable functions. Moreover, we newly derive a sufficient condition for the nonnegativity of quartic C^2-splines on triangular grid and also propose quartic C^2-spline interpolation and smoothing on triangular grid. Then we estimate a two-dimensional probability density function (PDF) from its histogram by extending the idea of the positive quartic and biquartic C^2-spline smoothing. Numerical experiments show the effectiveness of the newly derived sufficient condition and the proposed PDF estimator.

1 Introduction

Spline is a function in a class of piecewise polynomials and has been widely used for designs of continuous models in many signal and image processing applications [1], e.g., super-resolution [2], [3], computer aided design [4], [5], and regression analysis [6], [7], due to its flexibility and optimality (see, e.g., Fact 1). On the other hand, designs of positive continuous functions such as probability density function (PDF) [8], [9] and power spectral density [10], [11] are also required in many applications, e.g., pattern recognition [12], [13], quantization [14], filtering [15], data analysis [16], speech enhancement [17], speech recognition [18] and sound source separation [19]. However spline interpolation and smoothing has been hardly applicable to the designs of positive functions because the nonnegativity of splines is not guaranteed in general.

In our previous work [20], by using the sufficient condition [21] for the nonnegativity of univariate splines, we proposed one-dimensional positive spline smoothing and its application to PDF estimation. However, trying to extend this results to higher dimensions, we encounter nonobvious questions even in two-dimensional case, e.g., which grids are suitable for defining bivariate splines and which functionals are suitable as cost of optimization problems.

In this paper, as an extension of [20], we propose two-dimensional positive spline smoothing and its application to PDF estimation. In Section 2, as preliminaries, we introduce two kinds of bivariate spline spaces using square grid or triangular grid (Section 2.1), and define two-dimensional positive spline interpolation and smoothing as optimization problems (Section 2.2). In Section 3, on the basis of the derivation of the sufficient condition [21] for the nonnegativity of bivariate splines over squares (Section 3.1), we newly derive a sufficient condition for the nonnegativity over triangles (Section 3.2). Then we solve the optimization problems under the sufficient condition as quadratic programming problems (Section 3.3). Moreover, by modifying the idea of the positive spline smoothing, we estimate a two-dimensional PDF from its histogram as a bivariate spline (Section 3.4). In Section 4, first we numerically evaluate the effectiveness of the newly derived sufficient condition by experiments for the positive spline interpolation/smoothing (Section 4.1). Second we show the effectiveness of the proposed PDF estimator compared with the kernel density estimation [8], [9], which has been widely used for non-parametric PDF estimation, by experiments for a Gaussian mixture (Section 4.2). In Section 5, we conclude this paper.

Notation Let $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{Z}_+$ be respectively the set of all real numbers, nonnegative real numbers and nonnegative integers. Boldface small and capital letters respectively express a vector and a matrix. The norm of $\boldsymbol{x} := (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ is defined as $\|\boldsymbol{x}\| := \sqrt{\sum_{i=1}^n x_i^2}$.

2 Positive Spline Interpolation and Smoothing

2.1 Bivariate Spline Spaces

Let $\square_{n,m} := \{\mathcal{R}_{i,j} := [x_{i-1}, x_i] \times [y_{j-1}, y_j]\}_{j=1,2,\ldots,m}^{i=1,2,\ldots,n}$ be a set of all squares $R_{i,j}$ on $\Omega := [x_0, x_n] \times [y_0, y_m] \subset \mathbb{R}^2$ s.t. $x_i - x_{i-1} = 1$ for all $i = 1, 2, \ldots, n$ and $y_j - y_{j-1} = 1$ for all $j = 1, 2, \ldots, m$. For $\rho, d\mathbb{Z}_+$ s.t. $0 \leq \rho < d$, define

$$\mathcal{S}_d^\rho(\square_{n,m}) := \{f \in C_\rho^{2\rho}(\Omega) \mid \forall \mathcal{R}_{i,j} \in \square_{n,m} \quad f|_{\mathcal{R}_{i,j}} \in \mathbb{P}_{d,d}\}$$

as the set of all bisplines of degree d and smoothness ρ on $\square_{n,m}$, where $C_\rho^{2\rho}(\Omega)$ stands for the set of all continuous functions $f : \Omega \to \mathbb{R}$ whose partial derivatives $\frac{\partial^{i+j} f}{\partial x^i \partial y^j}$ $(i, j = 0, 1, \ldots, \rho)$ are also continuous over Ω, $f|_{\mathcal{R}_{i,j}} : \mathcal{R}_{i,j} \to \mathbb{R}$ denotes the restriction of f to $\mathcal{R}_{i,j}$, and $\mathbb{P}_{d,d}$ is the set of all bivariate polynomials whose degree is d at most with regard to each variable, i.e., $\mathbb{P}_{d,d} := \{f : (x, y) \mapsto \sum_{p=0}^d \sum_{q=0}^d c_{p,q} x^p y^q \mid c_{p,q} \in \mathbb{R}\}$.

Assume that we observe samples of a twice continuously differentiable function $g : \Omega \to \mathbb{R}$ with additive noise $\epsilon_{i,j} \in \mathbb{R}$ at (x_i, y_j), i.e., we observe $z_{i,j} := g(x_i, y_j) + \epsilon_{i,j}$ $(i = 0, 1, \ldots, n$ and $j = 0, 1, \ldots, m)$. In this situation, a natural[1] bicubic spline $f \in \mathcal{S}_3^2(\square_{n,m})$ is often used to approximate g due to its optimality shown in the following fact [22], [23].

[1]A bispline $f \in \mathcal{S}_3^2(\square_{n,m})$ is *natural* if $\frac{\partial^2 f}{\partial x^2}(x, y) = 0$ $\forall (x, y) \in \{x_0, x_n\} \times [y_0, y_m]$ and $\frac{\partial^2 f}{\partial y^2}(x, y) = 0$ $\forall (x, y) \in [x_0, x_n] \times \{y_0, y_m\}$.

Fact 1 (Bicubic spline as a solution of a variational problem) *There always exists a unique minimizer $f^* \in C_2^4(\Omega)$ of*

$$\sum_{i=0}^{n}\sum_{j=0}^{m}|f(x_i,y_j)-z_{i,j}|^2+\lambda\iint_{\Omega}\left|\frac{\partial^4 f}{\partial x^2\partial y^2}\right|^2 \mathrm{d}x\mathrm{d}y$$

and it is a natural bicubic spline $f^* \in \mathcal{S}_3^2(\square_{n,m})$, where smoothing parameter $\lambda > 0$ controls the trade-off between data fidelity and smoothness.

In order to achieve more flexibility with respect to the resolution of the discretization in Ω, partitioning of Ω into triangles has been studied [24]–[28]. Define a triangle $\mathcal{T}$ on $\mathbb{R}^2$, by specifying three vertices $\boldsymbol{v}_k := (x_k, y_k) \in \mathbb{R}^2$ $(k = 1, 2, 3)$ which are not arranged linearly, i.e., $x_1y_2 - y_1x_2 + x_2y_3 - y_2x_3 + x_3y_1 - y_3x_1 \neq 0$, as

$$\begin{aligned}\mathcal{T} &:= \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\rangle \\ &:= \left\{ r\boldsymbol{v}_1 + s\boldsymbol{v}_2 + t\boldsymbol{v}_3 \in \mathbb{R}^2 \,\middle|\, \begin{array}{l} r,s,t \in [0,1] \\ r+s+t=1 \end{array}\right\}.\end{aligned}$$

Then create four kinds of triangles

$$\begin{cases} \mathcal{T}_{i,j,1} := \langle (x_{i-1}, y_{j-1}), (x_i, y_{j-1}), (\frac{x_{i-1}+x_i}{2}, \frac{y_{j-1}+y_j}{2})\rangle \\ \mathcal{T}_{i,j,2} := \langle (x_{i-1}, y_j), (x_{i-1}, y_{j-1}), (\frac{x_{i-1}+x_i}{2}, \frac{y_{j-1}+y_j}{2})\rangle \\ \mathcal{T}_{i,j,3} := \langle (x_i, y_j), (x_{i-1}, y_j), (\frac{x_{i-1}+x_i}{2}, \frac{y_{j-1}+y_j}{2})\rangle \\ \mathcal{T}_{i,j,4} := \langle (x_i, y_{j-1}), (x_i, y_j), (\frac{x_{i-1}+x_i}{2}, \frac{y_{j-1}+y_j}{2})\rangle \end{cases}$$

by diagonally cutting every square $\mathcal{R}_{i,j}$. Let $\boxtimes_{n,m} := \{\mathcal{T}_{i,j,1}, \mathcal{T}_{i,j,2}, \mathcal{T}_{i,j,3}, \mathcal{T}_{i,j,4}\}_{j=1,2,\dots,m}^{i=1,2,\dots,n}$ be a set of all triangles $\mathcal{T}_{i,j,k}$ on Ω. For $\rho, d \in \mathbb{Z}_+$ s.t. $0 \le \rho < d$, define

$$\mathcal{S}_d^\rho(\boxtimes_{n,m}) := \{f \in C^\rho(\Omega) \,|\, \forall \mathcal{T}_{i,j,k} \in \boxtimes_{n,m} \quad f|_{\mathcal{T}_{i,j,k}} \in \mathbb{P}_d\}$$

as the set of all bivariate splines of degree d and smoothness ρ on $\boxtimes_{n,m}$, where $C^\rho(\Omega)$ stands for the set of all ρ-times continuously differentiable functions over Ω, and $\mathbb{P}_d$ is the set of all bivariate polynomials whose degree is d at most, i.e., $\mathbb{P}_d := \{f : (x,y) \mapsto \sum_{p=0}^{d}\sum_{q=0}^{d-p} c_{p,q}x^py^q \,|\, c_{p,q} \in \mathbb{R}\}$.

Remark 1 *For the above spaces, $\mathcal{S}_d^\rho(\square_{n,m}) \subset \mathcal{S}_{2d}^\rho(\boxtimes_{n,m})$, $C^{2\rho}(\Omega) \subset C_\rho^{2\rho}(\Omega) \subset C^\rho(\Omega)$, and $\mathbb{P}_d \subset \mathbb{P}_{d,d} \subset \mathbb{P}_{2d}$ hold.*

2.2 Two-Dimensional Positive Quartic and Biquartic Spline Interpolation/Smoothing

The problem of our interest is to reconstruct a nonnegative function $g : \Omega \to \mathbb{R}_+$ with the use of its nonnegative samples $z_{i,j} := g(x_i, y_j) + \epsilon_{i,j} \ge 0$. In [21], Heß and Schmidt considered the positive C^2-bispline interpolation, and they shown that the lowest degree is $d = 4$ for guaranteeing the existence of a bispline $f \in \mathcal{S}_d^2(\square_{n,m})$ satisfying $f(x_i, y_j) = z_{i,j}$ $(i = 0, 1, \dots, n$ and $j = 0, 1, \dots, m)$ and $f(x, y) \ge 0$ for all $(x, y) \in \Omega$. However, they did not show which functionals are suitable as cost of optimization problems. In this paper, we employ not $\iint_\Omega |\frac{\partial^4 f}{\partial x^2 \partial y^2}|^2 \,\mathrm{d}x\mathrm{d}y$ in Fact 1 but the energy of local changes $\iint_\Omega [|\frac{\partial^2 f}{\partial x^2}|^2 + 2|\frac{\partial^2 f}{\partial x \partial y}|^2 + |\frac{\partial^2 f}{\partial y^2}|^2] \,\mathrm{d}x\mathrm{d}y$ used in [26] as the cost, and consider the following two problems.

Problem 1 (Two-dimensional positive spline interpolation) *Find $f^* \in \mathcal{S}_4^2(\square_{n,m})$ (or $f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$) minimizing*

$$\iint_\Omega \left[\left|\frac{\partial^2 f}{\partial x^2}\right|^2 + 2\left|\frac{\partial^2 f}{\partial x\partial y}\right|^2 + \left|\frac{\partial^2 f}{\partial y^2}\right|^2\right]\mathrm{d}x\mathrm{d}y$$

subject to $f(x_i, y_j) = z_{i,j}$ $(i = 0, 1, \dots, n$ and $j = 0, 1, \dots, m)$ and $f(x, y) \ge 0$ for all $(x, y) \in \Omega$.

Problem 2 (Two-dimensional positive spline smoothing) *Find $f^* \in \mathcal{S}_4^2(\square_{n,m})$ (or $f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$) minimizing*

$$\begin{aligned}&\sum_{i=0}^{n}\sum_{j=0}^{m}|f(x_i,y_j)-z_{i,j}|^2 \\ &\quad + \lambda\iint_\Omega \left[\left|\frac{\partial^2 f}{\partial x^2}\right|^2 + 2\left|\frac{\partial^2 f}{\partial x\partial y}\right|^2 + \left|\frac{\partial^2 f}{\partial y^2}\right|^2\right]\mathrm{d}x\mathrm{d}y\end{aligned}$$

subject to $f(x, y) \ge 0$ for all $(x, y) \in \Omega$, where $\lambda > 0$.

In the next section, in order to solve the above problems, we use the sufficient condition for the nonnegativity of $f \in \mathcal{S}_4^2(\square_{n,m})$ in [21]. Moreover, on the basis of the derivation in [21], we newly derive a sufficient condition for the nonnegativity of $f \in \mathcal{S}_4^2(\boxtimes_{n,m})$.

3 Positive Spline Smoothing Under Sufficient Condition and Application to PDF Estimation

3.1 Sufficient Condition for Nonnegativity on Squares

To summarize the discussion in [21], the sufficient condition for the nonnegativity of $f \in \mathcal{S}_4^2(\square_{n,m})$ over Ω was derived as follows. Suppose that a bispline $f \in \mathcal{S}_4^2(\square_{n,m})$ is expressed, over $\mathcal{R}_{i,j}$, as

$$f(x,y) = \sum_{p=0}^{4}\sum_{q=0}^{4} c_{p,q}^{i,j} s^p t^q, \tag{1}$$

where $c_{p,q}^{i,j} \in \mathbb{R}$, $s := x - x_i \in [0, 1]$ and $t := y - y_j \in [0, 1]$. Substitute $s = \frac{\alpha}{1+\alpha}$ and $t = \frac{\beta}{1+\beta}$, which imply that $s, t \in [0, 1]$ if and only if $\alpha, \beta \in \mathbb{R}_+ \cup \{\infty\} =: [0, \infty]$. After some algebra, we can obtain vectors $\boldsymbol{g}_{p,q} \in \mathbb{Z}_+^{25}$ $(p, q = 0, 1, 2, 3, 4)$ satisfying

$$(1+\alpha)^4(1+\beta)^4 f(x,y) = \sum_{p=0}^{4}\sum_{q=0}^{4} \boldsymbol{g}_{p,q}^T \boldsymbol{c}_{i,j} \alpha^p \beta^q,$$

where $\boldsymbol{c}_{i,j} = (c_{4,4}^{i,j}, c_{4,3}^{i,j}, \dots, c_{4,0}^{i,j}, c_{3,4}^{i,j}, \dots, c_{0,0}^{i,j})^T \in \mathbb{R}^{25}$. From $(1 + \alpha)^4(1 + \beta)^4 \ge 1$ and $\alpha^p\beta^q \ge 0$ for all $\alpha, \beta \in [0, \infty]$, the sufficient condition for the nonnegativity of f in (1) is

$$\boldsymbol{g}_{p,q}^T \boldsymbol{c}_{i,j} \ge 0 \quad \text{for all } i,\ j,\ p \text{ and } q. \tag{2}$$

Therefore, by using $\boldsymbol{g}_{p,q}$ $(p, q = 0, 1, 2, 3, 4)$, we can create a matrix $\boldsymbol{G}$ s.t. $\boldsymbol{G}\boldsymbol{c} \ge \boldsymbol{0} \Rightarrow f(x, y) \ge 0$ for all $(x, y) \in \Omega$, where $\boldsymbol{c} := (\boldsymbol{c}_{1,1}^T, \boldsymbol{c}_{1,2}^T, \dots, \boldsymbol{c}_{n,m}^T)^T \in \mathbb{R}^{25mn}$.

3.2 Sufficient Condition for Nonnegativity on Triangles

By utilizing the above discussion, we newly derive a sufficient condition for the nonnegativity of $f \in \mathcal{S}_4^2(\boxtimes_{n,m})$ over Ω. Suppose that a bivariate spline $f \in \mathcal{S}_4^2(\boxtimes_{n,m})$ is expressed, over $\mathcal{T}_{i,j,k}$, as

$$f(x,y) = \sum_{p=0}^{4}\sum_{q=0}^{p} c^{i,j,k}_{p(p+1)/2+q+1} \frac{4!\, r^{4-p} s^{p-q} t^{q}}{(4-p)!(p-q)!q!}, \tag{3}$$

where $c^{i,j,k}_{p(p+1)/2+q+1} \in \mathbb{R}$ and $(r,s,t) \in [0,1]^3$ $(r+s+t = 1)$ is called *barycentric coordinate* of (x,y) with respect to $\mathcal{T}_{i,j,k}$ [24], [25], e.g., the barycentric coordinate with respect to $\mathcal{T}_{i,j,1}$ is

$$(r,s,t) = (x_{i-1}-x+y_j-y, x-x_{i-1}+y_{j-1}-y, 2y-2y_{j-1}).$$

Substitute $r = \frac{\alpha}{1+\alpha}$, $s = \frac{\beta}{1+\beta}$, and $t = 1-r-s = \frac{1-\alpha\beta}{(1+\alpha)(1+\beta)}$, which imply that $r,s,t \in [0,1]$ if and only if $\alpha, \beta \in [0,\infty]$ and $\alpha\beta =: \chi \in [0,1]$. Then, we have

$$\begin{aligned}(1+\alpha)^4(1+\beta)^4 f(x,y) &= c_1^{i,j,k}\alpha^4 + c_{11}^{i,j,k}\beta^4 \\ &+ \sum_{p=1}^{3} P^{i,j,k}_{4-p}(\chi)\alpha^p + \sum_{q=1}^{3} Q^{i,j,k}_{4-q}(\chi)\beta^q + R_4^{i,j,k}(\chi),\end{aligned}$$

where $P^{i,j,k}_{4-p}(\chi)$, $Q^{i,j,k}_{4-q}(\chi)$ and $R^{i,j,k}_4(\chi)$ are univariate polynomials of degree $(4-p)$, $(4-q)$ and 4, respectively, as shown in (5) (the indices i, j and k are omitted for simplicity). Therefore, the sufficient condition for the nonnegativity of f in (3) is

$$\begin{aligned}&c_1^{i,j,k} \geq 0,\ \ c_{11}^{i,j,k} \geq 0,\ \ P^{i,j,k}_{4-p}(\chi) \geq 0,\ \ Q^{i,j,k}_{4-q}(\chi) \geq 0 \\ &\text{and } R_4^{i,j,k}(\chi) \geq 0 \ \text{ for all } i,\ j,\ k,\ p,\ q \text{ and } \chi \in [0,1].\end{aligned} \tag{4}$$

Finally, by substituting $\chi := \frac{\gamma}{1+\gamma}$ and computing the coefficients of γ^δ $(\delta = 0,1,2,3,4)$ in $(1+\gamma)^{4-p}P^{i,j,k}_{4-p}(\chi)$, $(1+\gamma)^{4-q}Q^{i,j,k}_{4-q}(\chi)$ and $(1+\gamma)^4 R^{i,j,k}_4(\chi)$, we can create a matrix $\boldsymbol{G}$ s.t. $\boldsymbol{Gc} \geq \boldsymbol{0} \Rightarrow f(x,y) \geq 0$ for all $(x,y) \in \Omega$, where $\boldsymbol{c}$ is defined, with the use of $\boldsymbol{c}_{i,j,k} := (c_1^{i,j,k}, c_2^{i,j,k}, \ldots, c_{15}^{i,j,k})^T \in \mathbb{R}^{15}$, as $\boldsymbol{c} := (\boldsymbol{c}_{1,1,1}^T, \boldsymbol{c}_{1,1,2}^T, \ldots, \boldsymbol{c}_{n,m,4}^T)^T \in \mathbb{R}^{60mn}$.

3.3 Problem 1 and Problem 2 under Sufficient Condition

In Problem 1, the cost $\iint_\Omega [|\frac{\partial^2 f}{\partial x^2}|^2 + 2|\frac{\partial^2 f}{\partial x \partial y}|^2 + |\frac{\partial^2 f}{\partial y^2}|^2]\, \mathrm{d}x\mathrm{d}y$ can be written by a quadratic form $\boldsymbol{c}^T\boldsymbol{Qc}$, where $\boldsymbol{c}$ is the coefficient vector of $f \in \mathcal{S}_4^2(\square_{n,m})$ (or $f \in \mathcal{S}_4^2(\boxtimes_{n,m})$) in Section 3.1 (or Section 3.2) and $\boldsymbol{Q}$ is a symmetric positive semidefinite matrix [28], [29]. Moreover, the conditions $f \in C_2^4(\Omega)$ (or $f \in C^2(\Omega)$) and $f(x_i, y_j) = z_{i,j}$ $(i = 0,1,\ldots,n$ and $j = 0,1,\ldots,m)$ are respectively expressed as $\boldsymbol{Hc} = \boldsymbol{0}$ and $\boldsymbol{\mathcal{I}c} = \boldsymbol{z} := (z_{0,0}, z_{0,1}, \ldots, z_{n,m})^T \in \mathbb{R}^{(m+1)(n+1)}$ with the use of certain sparse matrices $\boldsymbol{H}$ and $\boldsymbol{\mathcal{I}}$ [28], [30]. Therefore under the sufficient condition $\boldsymbol{Gc} \geq \boldsymbol{0}$, based on (2) (or (4)), for the nonnegativity of f, Problem 1 and Problem 2 are expressed as follows.

Problem 1.S (Problem 1 under the sufficient condition) *Find $\boldsymbol{c}^* \in \mathbb{R}^{25mn}$ (or $\boldsymbol{c}^* \in \mathbb{R}^{60mn}$) minimizing*

$$\boldsymbol{c}^T\boldsymbol{Qc}$$

subject to $\boldsymbol{Gc} \geq \boldsymbol{0}$, $\boldsymbol{Hc} = \boldsymbol{0}$ and $\boldsymbol{\mathcal{I}c} = \boldsymbol{z}$.

Problem 2.S (Problem 2 under the sufficient condition) *Find $\boldsymbol{c}^* \in \mathbb{R}^{25mn}$ (or $\boldsymbol{c}^* \in \mathbb{R}^{60mn}$) minimizing*

$$\|\boldsymbol{\mathcal{I}c} - \boldsymbol{z}\|^2 + \lambda \boldsymbol{c}^T\boldsymbol{Qc}$$

subject to $\boldsymbol{Gc} \geq \boldsymbol{0}$ and $\boldsymbol{Hc} = \boldsymbol{0}$, where $\lambda > 0$.

Problem 1.S and Problem 2.S are quadratic programming problems and can be solved in polynomial time [31]–[33].

Remark 2 *In Problem 1.S and Problem 2.S, the vector $\boldsymbol{c}$ and the matrices $\boldsymbol{G}$, $\boldsymbol{H}$, $\boldsymbol{\mathcal{I}}$, and $\boldsymbol{Q}$ depend on which bivariate spline space ($\mathcal{S}_4^2(\square_{n,m})$ or $\mathcal{S}_4^2(\boxtimes_{n,m})$) we use. In this paper, these dependencies are omitted for readability.*

Remark 3 *We can reduce the size of the above problems by using $\frac{\partial^{p+q} f}{\partial x^p \partial y^q}(x_i, y_j)$ (and $\frac{\partial^{p+q} f}{\partial x^p \partial y^q}(\frac{x_{i-1}+x_i}{2}, \frac{y_{j-1}+y_j}{2})$) $(p,q = 0,1,2)$, instead of $\boldsymbol{c}$, as parameters [21].*

3.4 Two-Dimensional PDF Estimation by Positive Spline

In this subsection, we estimate a two-dimensional probability density function (PDF) $g : \mathbb{R}^2 \to \mathbb{R}_+$ s.t. $g \in C^2(\mathbb{R}^2)$ by extending the idea of the positive spline smoothing. In this situation, we cannot observe values of g directly but construct a histogram from observed samples which are

$$\begin{cases} P_1(\chi) := 4\left\{(c_1+c_2-c_3)\chi + c_3\right\} \\ P_2(\chi) := 2\left\{3(c_1+2c_2-2c_3+c_4-2c_5+c_6)\chi^2 + 2(c_2+3c_3+3c_5-3c_6)\chi + 3c_6\right\} \\ P_3(\chi) := 4\left\{(c_1+3c_2-3c_3+3c_4-6c_5+3c_6+c_7-3c_8+3c_9-c_{10})\chi^3 \right. \\ \qquad\qquad \left. + 3(c_2+c_3+c_4+c_5-2c_6+c_8-2c_9+c_{10})\chi^2 + 3(c_5+c_6+c_9-c_{10})\chi + c_{10}\right\} \\ Q_1(\chi) := 4\left\{(c_7+c_{11}-c_{12})\chi + c_{12}\right\} \\ Q_2(\chi) := 2\left\{3(c_4+2c_7-2c_8+c_{11}-2c_{12}+c_{13})\chi^2 + 2(c_7+3c_8+3c_{12}-3c_{13})\chi + 3c_{13}\right\} \\ Q_3(\chi) := 4\left\{(c_2+3c_4-3c_5+3c_7-6c_8+3c_9+c_{11}-3c_{12}+3c_{13}-c_{14})\chi^3 \right. \\ \qquad\qquad \left. + 3(c_4+c_5+c_7+c_8-2c_9+c_{12}-2c_{13}+c_{14})\chi^2 + 3(c_8+c_9+c_{13}-c_{14})\chi + c_{14}\right\} \\ R_4(\chi) := (c_1+4c_2-4c_3+6c_4-12c_5+6c_6+4c_7-12c_8+12c_9-4c_{10}+c_{11}-4c_{12}+6c_{13}-4c_{14}+c_{15})\chi^4 \\ \qquad\qquad + 4(3c_2+c_3+6c_4-3c_5-3c_6+3c_7-3c_8-3c_9+3c_{10}+c_{12}-3c_{13}+3c_{14}-c_{15})\chi^3 \\ \qquad\qquad + 6(c_4+4c_5+c_6+4c_8-2c_9-2c_{10}+c_{13}-2c_{14}+c_{15})\chi^2 + 4(3c_9+c_{10}+c_{14}-c_{15})\chi + c_{15} \end{cases} \tag{5}$$

generated from g. Hence we reconstruct the PDF g from its histogram based on the observed samples.

Let $\{(u_\ell, v_\ell)\}_{\ell=1}^L$ be samples generated from g. We create a histogram by using $\mathcal{R}_{i,j}$ (or $\mathcal{T}_{i,j,k}$) as bins s.t. $x_0 < \min\{u_\ell\}$, $x_n > \max\{u_\ell\}$, $y_0 < \min\{v_\ell\}$ and $y_m > \max\{v_\ell\}$. Then by defining $L_{i,j}$ (or $L_{i,j,k}$) as the number of (u_ℓ, v_ℓ) in $\mathcal{R}_{i,j}$ (or $\mathcal{T}_{i,j,k}$), we can expect $\iint_{\mathcal{R}_{i,j}} g\,\mathrm{d}x\mathrm{d}y \approx \frac{L_{i,j}}{L}$ (or $\iint_{\mathcal{T}_{i,j,k}} g\,\mathrm{d}x\mathrm{d}y \approx \frac{L_{i,j,k}}{L}$). Moreover, by assuming $g(x,y) = 0$ for all $(x,y) \in \mathbb{R}^2 \setminus \Omega$, i.e., $\iint_\Omega g\,\mathrm{d}x\mathrm{d}y = 1$, we try to estimate g with the use of bivariate splines through the following optimization problem.

Problem 3 (Two-dimensional PDF estimation by splines) *Find $f^* \in \mathcal{S}_4^2(\square_{n,m})$ (or $f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$) minimizing*

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\left|\iint_{\mathcal{R}_{i,j}} f\,\mathrm{d}x\mathrm{d}y - \frac{L_{i,j}}{L}\right|^2$$
$$\left(or \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{4}\left|\iint_{\mathcal{T}_{i,j,k}} f\,\mathrm{d}x\mathrm{d}y - \frac{L_{i,j,k}}{L}\right|^2 \right)$$
$$+\lambda \iint_\Omega \left[\left|\frac{\partial^2 f}{\partial x^2}\right|^2 + 2\left|\frac{\partial^2 f}{\partial x \partial y}\right|^2 + \left|\frac{\partial^2 f}{\partial y^2}\right|^2\right]\mathrm{d}x\mathrm{d}y$$

subject to $f(x,y) \geq 0$ for all $(x,y) \in \Omega$, $\iint_\Omega f\,\mathrm{d}x\mathrm{d}y = 1$ and $\frac{\partial^{p+q} f}{\partial x^p \partial y^q}(x,y) = 0$ $(p,q = 0,1,2)$ for all $(x,y) \in (\{x_0, x_n\} \times [y_0, y_m]) \cup ([x_0, x_n] \times \{y_0, y_m\})$, where $\lambda > 0$.

Remark 4 *Actually, for any histogram, we can design an optimization problem like Problem 3, but here we employ $\mathcal{R}_{i,j}$ (or $\mathcal{T}_{i,j,k}$) as bins. This is because $\iint_{\mathcal{R}_{i,j}} f\,\mathrm{d}x\mathrm{d}y$ (or $\iint_{\mathcal{T}_{i,j,k}} f\,\mathrm{d}x\mathrm{d}y$) can be easily computed by using the coefficients $c_{p,q}^{i,j}$ (or $c_{p(p+1)/2+q+1}^{i,j,k}$).*

With the use of $\boldsymbol{\zeta} := \frac{1}{L}(L_{1,1}, L_{1,2}, \ldots, L_{n,m})^T \in \mathbb{R}^{mn}$ (or $\boldsymbol{\zeta} := \frac{1}{L}(L_{1,1,1}, L_{1,1,2}, \ldots, L_{n,m,4})^T \in \mathbb{R}^{4mn}$) and certain sparse matrices $\widetilde{\boldsymbol{H}}$ and $\widetilde{\boldsymbol{\mathcal{I}}}$, Problem 3 under the sufficient condition $\boldsymbol{Gc} \geq \boldsymbol{0}$ is written as the following convex quadratic programming problem.

Problem 3.S (Problem 3 under the sufficient condition) *Find $\boldsymbol{c}^* \in \mathbb{R}^{25mn}$ (or $\boldsymbol{c}^* \in \mathbb{R}^{60mn}$) minimizing*

$$\|\widetilde{\boldsymbol{\mathcal{I}}}\boldsymbol{c} - \boldsymbol{\zeta}\|^2 + \lambda \boldsymbol{c}^T \boldsymbol{Q}\boldsymbol{c}$$

subject to $\boldsymbol{Gc} \geq \boldsymbol{0}$, $\widetilde{\boldsymbol{H}}\boldsymbol{c} = \boldsymbol{0}$ and $\boldsymbol{1}^T\widetilde{\boldsymbol{\mathcal{I}}}\boldsymbol{c} = 1$, where $\lambda > 0$.

4 Numerical Experiments

4.1 Experiments for Problem 1.S and Problem 2.S

Let $\{\widetilde{z}_{i,j}\}_{j=0,1,\ldots,5}^{i=0,1,\ldots,5}$ be generated from the standard normal distribution $\mathcal{N}(0,1)$. Define $(x_i, y_j) := (i,j)$ and $z_{i,j} := |\widetilde{z}_{i,j}|$ $(i,j = 0,1,\ldots,5)$. Then solve Problem 1.S and Problem 2.S with $\lambda = \frac{1}{50}, \frac{1}{250}, \frac{1}{500}$ for two bivariate spline spaces $\mathcal{S}_4^2(\square_{5,5})$ and $\mathcal{S}_4^2(\boxtimes_{5,5})$.

Figure 1 shows an example of the results of Problem 1.S. Figures 1(a), 1(b) and 1(c) respectively depict $z_{i,j}$, $f^* \in \mathcal{S}_4^2(\square_{5,5})$ and $f^* \in \mathcal{S}_4^2(\boxtimes_{5,5})$. In this example, the proposed sufficient condition based on (4) constructs a smoother spline in Fig. 1(c) compared with the existing condition [21] based on (2). Table 1 shows the mean values of the minimum costs of Problems 1.S and 2.S in 1000 times. From Table 1, $\mathcal{S}_4^2(\boxtimes_{5,5})$ is suitable for Problem 1.S and Problem 2.S with small λ, and $\mathcal{S}_4^2(\square_{5,5})$ is suitable for Problem 2.S with large λ. This is because the influence of the sufficient condition is dominant for small λ, and the influence of $\mathbb{P}_d \subset \mathbb{P}_{d,d}$ is dominant for large λ.

4.2 Experiments for Problem 3.S

Let $\{(u_\ell, v_\ell)\}_{\ell=1}^L$ be samples generated from a Gaussian mixture

$$g(\boldsymbol{x}) := \sum_{i=1}^{2} \frac{w_i}{2\pi\sqrt{|\boldsymbol{\Sigma}_i|}}\, e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_i)},$$

where $w_1 = w_2 = 0.5$, $\boldsymbol{\mu}_1 = (1,4)^T$, $\boldsymbol{\mu}_2 = (6,7)^T$, $\boldsymbol{\Sigma}_1 = \left(\begin{smallmatrix} 3 & 2 \\ 2 & 3 \end{smallmatrix}\right)$ and $\boldsymbol{\Sigma}_2 = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$. Define $x_0 := \lfloor \min\{u_\ell\} \rfloor$, $x_n := \lceil \max\{u_\ell\} \rceil$, $y_0 := \lfloor \min\{v_\ell\} \rfloor$ and $y_m := \lceil \max\{v_\ell\} \rceil$, where $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ are respectively the floor and ceiling functions. We compare the performances of Problem 3.S using $\mathcal{S}_4^2(\square_{n,m})$ and $\mathcal{S}_4^2(\boxtimes_{n,m})$ with that of the kernel density estimation [8], [9] using Gaussian kernels. The kernel density estimation constructs an estimate of g as

$$f_{\mathrm{KDE}}(x,y) = \frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{2\pi h_x h_y}\, e^{-\left\{\frac{(x-u_\ell)^2}{2h_x^2} + \frac{(y-v_\ell)^2}{2h_y^2}\right\}}$$

with the use of the bandwidth (h_x, h_y) selected by [9].

Figure 2 shows an example of the results of the proposed method and the kernel density estimation from 1000 samples $\{(u_\ell, v_\ell)\}_{\ell=1}^{1000}$. Figure 2(a) depicts the true PDF g. Figure 2(b), 2(c) and 2(d) depict the estimates f_{KDE}, $f^* \in \mathcal{S}_4^2(\square_{n,m})$ and $f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$, respectively. From Figs. 2(c) and 2(d), the proposed method constructs smoother estimates and achieves the lower ℓ_1-norm errors compared with the kernel density estimation. Table 2 shows the mean values of the ℓ_1-norm errors in 100 times. From Table 2, the proposed method using $S_4^2(\boxtimes_{n,m})$ with $\lambda = \frac{1}{75}$ achieves the best performance due to more flexibility of histograms and splines based on triangular grid.

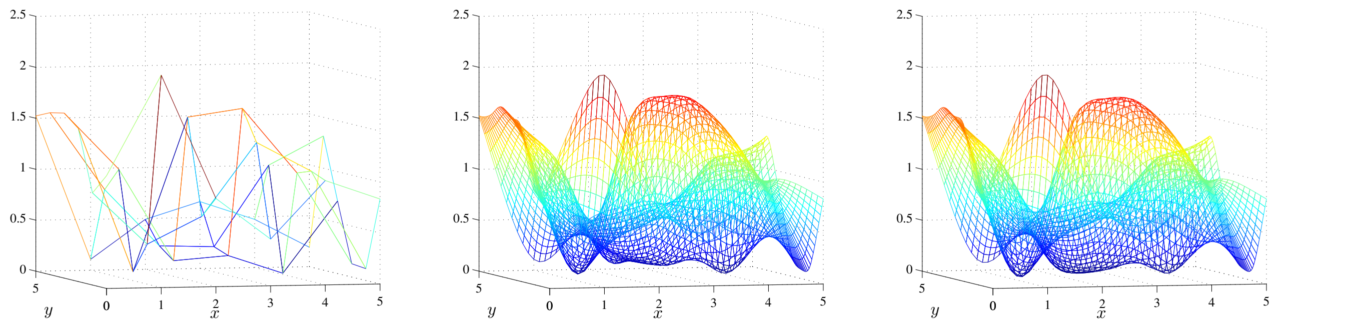
5 Conclusion

In this paper, first we have newly derived a sufficient condition for the nonnegativity of $f \in \mathcal{S}_4^2(\boxtimes_{n,m})$ by utilizing the derivation of the existing sufficient condition for the nonnegativity of $f \in \mathcal{S}_4^2(\square_{n,m})$. Second we solved two-dimensional positive spline interpolation and smoothing under the sufficient condition as quadratic programming problems. Third we estimated two-dimensional PDFs as positive bivariate splines by using the idea of the positive spline smoothing. Numerical experiments show the effectiveness of the newly derived sufficient condition and the proposed PDF estimator compared with the existing condition and the kernel density estimation, respectively.

As future work, we plan to apply the positive spline smoothing to two-dimensional spectral analysis [10], [11] which is especially important in image and speech processing applications.

Acknowledgment

This work was supported in part by JSPS Grants-in-Aid Grant Numbers 26-920 and B-15H02752.

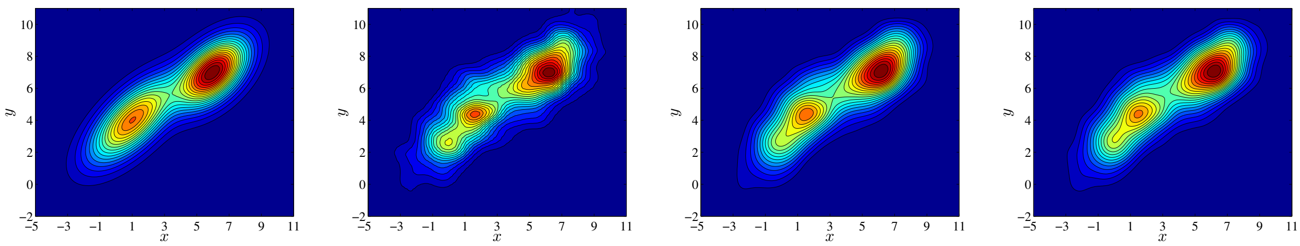


(a) $z_{i,j}$ (b) $f^* \in \mathcal{S}_4^2(\Box_{5,5})$ ($\boldsymbol{c}^{*T}\boldsymbol{Q}\boldsymbol{c}^* = 148.14$) (c) $f^* \in \mathcal{S}_4^2(\boxtimes_{5,5})$ ($\boldsymbol{c}^{*T}\boldsymbol{Q}\boldsymbol{c}^* = \mathbf{143.57}$)

Figure 1: Examples of solutions $f^* \in \mathcal{S}_4^2(\Box_{5,5})$ and $f^* \in \mathcal{S}_4^2(\boxtimes_{5,5})$ of Problem 1.S and their costs $\boldsymbol{c}^{*T}\boldsymbol{Q}\boldsymbol{c}^*$.

Table 1: Mean values of minimum costs $\boldsymbol{c}^{*T}\boldsymbol{Q}\boldsymbol{c}^*$ (Problem 1.S) and $\|\boldsymbol{\mathcal{I}}\boldsymbol{c}^* - \boldsymbol{z}\|^2 + \lambda\boldsymbol{c}^{*T}\boldsymbol{Q}\boldsymbol{c}^*$ (Problem 2.S) in 1000 times.

	Problem 1.S	Problem 2.S ($\lambda = \frac{1}{500}$)	Problem 2.S ($\lambda = \frac{1}{250}$)	Problem 2.S ($\lambda = \frac{1}{50}$)
$f^* \in \mathcal{S}_4^2(\Box_{5,5})$	243.60	0.44205	**0.81872**	**2.7522**
$f^* \in \mathcal{S}_4^2(\boxtimes_{5,5})$	**242.60**	**0.44019**	0.81878	2.7633



(a) g (b) f_{KDE} ($\|g - f_{\mathrm{KDE}}\|_1 \approx 0.1402$) (c) $f^* \in \mathcal{S}_4^2(\Box_{n,m})$ ($\lambda = \frac{1}{25}$) ($\|g - f^*\|_1 \approx 0.1238$) (d) $f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$ ($\lambda = \frac{1}{75}$) ($\|g - f^*\|_1 \approx \mathbf{0.1186}$)

Figure 2: Examples of estimates f_{KDE}, $f^* \in \mathcal{S}_4^2(\Box_{n,m})$ with $\lambda = \frac{1}{25}$, and $f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$ with $\lambda = \frac{1}{75}$ from $\{(u_\ell, v_\ell)\}_{\ell=1}^{1000}$ and the ℓ_1-norm errors ($\|g-f\|_1 := \iint_{\mathbb{R}^2} |g-f| \,\mathrm{d}x\mathrm{d}y \approx \sum_{i=0}^{10n}\sum_{j=0}^{10m} 0.01|g(x_0+0.1i, y_0+0.1j) - f(x_0+0.1i, y_0+0.1j)|$).

Table 2: Mean values of the ℓ_1-norm errors between g and f_{KDE}, $f^* \in \mathcal{S}_4^2(\Box_{n,m})$ and $f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$ in 100 times.

f_{KDE} [9]	$f^* \in \mathcal{S}_4^2(\Box_{n,m})$ ($\lambda = \frac{1}{10}$ / $\mathbf{\frac{1}{25}}$ / $\frac{1}{50}$ / $\frac{1}{75}$ / $\frac{1}{100}$)	$f^* \in \mathcal{S}_4^2(\boxtimes_{n,m})$ ($\lambda = \frac{1}{10}$ / $\frac{1}{25}$ / $\frac{1}{50}$ / $\mathbf{\frac{1}{75}}$ / $\frac{1}{100}$)
0.1607	0.1634 / **0.1415** / 0.1488 / 0.1599 / 0.1696	0.2777 / 0.1849 / 0.1459 / **0.1367** / 0.1400

REFERENCES

[1] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[2] H. S. Hou and H. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 6, pp. 508–517, Dec. 1978.

[3] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 6, pp. 1153–1160, Dec. 1981.

[4] E. Cohen, T. Lyche, and R. Riesenfeld, "Discrete B-splines and subdivision techniques in computer-aided geometric design and computer graphics," *Computer Graphics and Image Processing*, vol. 14, no. 2, pp. 87–111, Oct. 1980.

[5] G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Code*, 4th ed. New York, NY: Academic Press, 1996.

[6] S. Wold, "Spline functions in data analysis," *Technometrics*, vol. 16, no. 1, pp. 87–111, Feb. 1974.

[7] L. C. Marsh and D. R. Cormier, *Spline Regression Models*, ser. Quantitative Applications in the Social Sciences. Thousand Oaks, CA: SAGE, 2001, vol. 137.

[8] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962.

[9] Z. I. Botev, J. F. Grotowski, and D. P. Kroese, "Kernel density estimation via diffusion," *The Annals of Statistics*, vol. 38, no. 5, pp. 2916–2657, Aug. 2010.

[10] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Englewood Cliffs, NJ: Prentice Hall, 2005.

[11] A. Quinquis, *Digital Signal Processing Using MATLAB*. New York, NY: Wiley, Jan. 2010, ch. 10. Power Spectral Density Estimation, pp. 241–278.

[12] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press, 1995.

[13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification* New York, NY: Wiley, 2012.

[14] S. Kar, H. Chen, and P. K. Varshne, "Optimal identical binary quantizer design for distributed estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3896–3901, Jul. 2012.

[15] D. Simon and D. L. Simon, "Constrained Kalman filtering via density function truncation for turbofan engine health estimation," *International Journal of Systems Science*, vol. 41, no. 2, pp. 159–171, 2010.

[16] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures, 4th ed.* New York, NY: Wiley, 2010.

[17] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, Jul. 2001.

[18] H. G. Hirsch and C. Ehrlicher, "Noise estimation techniques for robust speech recognition," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'95)*, 1995, pp. 153–156.

[19] Y. Hioka, K. Furuya, K. Kobayashi, K. Niwa, and Y. Haneda, "Underdetermined sound source separation using power spectrum density estimated by combination of directivity gain," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 6, pp. 1240–1250, Jun. 2013.

[20] D. Kitahara and I. Yamada, "Probability Density function estimation by positive quartic C^2-spline functions," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'15)*, 2015, pp. 3556–3560.

[21] W. Heß and J. W. Schmidt, "Positive quartic, monotone quintic C^2-spline interpolation in one and two dimensions," *Journal of Computational and Applied Mathematics*, vol. 55, no. 1, pp. 51–67, Oct. 1994.

[22] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications*. New York, NY: Academic Press, 1967.

[23] V. Pretlovà, "Bicubic spline smoothing of two-dimensional geophysical data," *Studia Geophysica et Geodaetica*, vol. 20, no. 2, pp. 168–177, Jun. 1976.

[24] G. Farin, "Triangular Bernstein-Bèzier patches," *Computer Aided Geometric Design*, vol. 3, no. 2, pp. 83–127, Aug. 1986.

[25] C. K. Chui, *Multivariate Splines*, ser. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia, PA: SIAM, 1988, vol. 54.

[26] G. E. Fasshauer and L. L. Schumaker, "Minimal energy surfaces using parametric splines," *Computer Aided Geometric Design*, vol. 13, no. 1, pp. 45–79, Feb. 1996.

[27] M. J. Lai and L. L. Schumaker, *Spline Functions on Triangulations*. New York, NY: Cambridge University Press, 2007.

[28] M. J. Lai, "Multivariate splines and their applications," in *Computational Complexity: Theory, Techniques, and Applications*, R. A. Meyers, Ed. New York, NY: Springer, 2012, pp. 1939–1980.

[29] E. Quak and L. L. Schumaker, "Calculation of the energy of a piecewise polynomial surface," in *Algorithms for Approximation II*, M. G. Cox and J. C. Mason, Eds. London, UK: Chapman & Hall, 1990, pp. 134–143.

[30] M. J. Lai, "Geometric interpretation of smoothness conditions of triangular polynomial patches," *Computer Aided Geometric Design*, vol. 14, no. 2, pp. 191–199, Feb. 1997.

[31] Y. Ye and E. Tse, "An extension of Karmarkar's projective algorithm for convex quadratic programming," *Mathematical Programming*, vol. 44, no. 1, pp. 157–179, May 1989.

[32] D. Goldfarb and S. Liu, "An $O(n^3L)$ primal interior point algorithm for convex quadratic programming," *Mathematical Programming*, vol. 49, no. 1, pp. 325–340, Nov. 1990.

[33] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, ser. Studies in Applied and Numerical Mathematics. Philadelphia, PA: SIAM, 1994, vol. 13.