# Simplicial Multivariable Linear Interpolation 

TR91-002
February, 1991

John H. Halton

The University of North Carolina at Chapel Hill Department of Computer Science CB\#3175, Sitterson Hall Chapel Hill, NC 27599-3175


UNC is an Equal Opportunity/Affirmative Action Institution.

# Simplicial <br> Multivariable <br> LINEAR <br> Interpolation 

John H. Halton<br>Computer Science Department<br>The University of North Carolina at Chapel Hill<br>Chapel Hill, NC 27599-3175

## 1. INTRODUCTION

Given a well-behaved multivariable function $F\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)$, whose values are given at the nodes of a cubic lattice

$$
\begin{equation*}
L=\left\{p_{1} \mu, p_{2} \mu, \ldots, p_{k} \mu \mid(\forall j \mid 1 \leq j \leq k) p_{j} \in \mathbb{Z}\right\}, \tag{1}
\end{equation*}
$$

where $\mu$ is the lattice-constant of $L$ and $\mathbb{Z}$ is the set of all (positive or negative) integers; we consider the problem of efficiently (i.e., with least effort) and effectively (i.e., with greatest accuracy) interpolating these values at any point ( $\xi_{1}, \xi_{2}, \ldots, \xi_{k}$ ).

Let us write

$$
\begin{equation*}
z \mid=\max \{q \in \mathbb{Z} \mid q \leq z\} \quad \text { and }\langle z\rangle=z-\lfloor z \tag{2}
\end{equation*}
$$

(so that $\lfloor z$ is what is usually called the floor function, or the integer part, of $z$, and $\langle z\rangle$ is what is usually called the fractional part of $z$. Clearly, we have that

$$
\begin{equation*}
0 \leq\langle z\rangle<1 \tag{3}
\end{equation*}
$$

Now, let us define

$$
\begin{equation*}
p_{j}=\left\langle\frac{\xi_{j}}{\mu}\right\rangle \quad \text { and } \quad x_{j}=\left\langle\frac{\xi_{j}}{\mu}\right\rangle \tag{4}
\end{equation*}
$$

so that

$$
\begin{equation*}
(\forall j \mid 1 \leq j \leq k) \quad 0 \leq x_{j}<1 \tag{5}
\end{equation*}
$$

and

$$
\begin{equation*}
(\forall j \mid 1 \leq j \leq k) \quad \xi_{j}=\left(p_{j}+x_{j}\right) \mu \tag{6}
\end{equation*}
$$

We limit ourselves to the interpolation of $F$ at $\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)$ as a multivariable linear function of $x_{1}, x_{2}, \ldots, x_{k}$, using only the values

$$
\begin{equation*}
f_{c}=F_{p, \boldsymbol{c}}=F\left(\left(p_{1}+c_{1}\right) \mu,\left(p_{2}+c_{2}\right) \mu, \ldots,\left(p_{k}+c_{k}\right) \mu\right) \tag{7}
\end{equation*}
$$

where

$$
\begin{equation*}
(\forall j \mid 1 \leq j \leq k) \quad c_{j} \in\{0,1\} \tag{8}
\end{equation*}
$$

as coefficients. These are the known values of $F$ at the vertices of the latticecell enclosing the point $\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)$.

The usual 'full' multivariable linear interpolation then takes the form

$$
\begin{align*}
& \Phi_{x}=\Phi\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right) \\
&=\sum_{c_{1}=0}^{1} \sum_{c_{2}=0}^{1} \ldots \sum_{c_{k}=0}^{1}\left\{\prod_{j=1}^{k}\left[c_{j} x_{j}+\left(1-c_{j}\right)\left(1-x_{j}\right)\right]\right\} f_{c} \tag{9}
\end{align*}
$$

Thus, it involves a combination of $2^{k}$ data for each interpolation.
A $k$-dimensional simplex $T$ is defined by $k+1$ vertices,

$$
\begin{equation*}
T=\mathbb{T}^{( }\left(P_{0}, P_{1}, P_{2}, \ldots, P_{k}\right) \tag{10}
\end{equation*}
$$

where

$$
\begin{equation*}
(\forall s \mid 0 \leq s \leq k) \quad P_{s}=\left(\alpha_{1 s}, \alpha_{2 s}, \ldots, \alpha_{k s}\right) \tag{11}
\end{equation*}
$$

$$
\text { and } T=\left\{\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right) \left\lvert\, \begin{array}{cc}
(\forall s \mid 0 \leq s \leq k) & 0 \leq \lambda_{s} \leq 1  \tag{12}\\
& 1=\sum_{s=0}^{k} \lambda_{s} \\
(\forall j \mid 1 \leq j \leq k) & \xi_{j}=\sum_{s=0}^{k} \alpha_{j s} \lambda_{s}
\end{array}\right.\right\} .
$$

We now consider the possibility of obtaining an interpolated approximation to $F\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)$ by a formula of the form

$$
\begin{equation*}
\Psi_{x}=\Psi\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)=\sum_{s=0}^{k} \lambda_{s} F\left(P_{s}\right) \tag{13}
\end{equation*}
$$

where the $P_{s}$ are selected from the $2^{k}$ lattice points

$$
\begin{equation*}
\left(\left(p_{1}+c_{1}\right) \mu,\left(p_{2}+c_{2}\right) \mu, \ldots,\left(p_{k}+c_{k}\right) \mu\right) \tag{14}
\end{equation*}
$$

defined as in (4)-(9) above.

## 2. Existence

Our first problem is to determine whether, indeed, given any point ( $\xi_{1}, \xi_{2}, \ldots, \xi_{k}$ ), there always exists a simplex $T$, defined as in (10)-(12), whose vertices are a subset of the vertices of the lattice-cell enclosing the point $\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)$, which contains the given point. Equivalently, we ask whether there is a subset of $k+1$ of the $2^{k}$ points $\boldsymbol{c}=\left(c_{1}, c_{1}, \ldots, c_{k}\right)$ satisfying (8), which form the vertices of a simplex containing any given point ( $x_{1}, x_{2}, \ldots, x_{k}$ ) satisfying (5).

We observe that, by (12), we require that there be $\lambda_{0}, \lambda_{1}, \lambda_{2}, \ldots, \lambda_{k}$, such that $(\forall s \mid 0 \leq s \leq k) 0 \leq \lambda_{s} \leq 1$, and

$$
\left\{\begin{array}{c}
\sum_{s=0}^{k} \lambda_{s}=1  \tag{15}\\
(\forall j \mid 1 \leq j \leq k)
\end{array} \sum_{s=0}^{k} c_{j s} \lambda_{s}=x_{j}\right\}
$$

Dropping the conditions on the $\lambda_{s}$, we see that this is equivalent to the matrix equation

$$
\begin{gather*}
{\left[\begin{array}{ccccc}
1 & 1 & 1 & \ldots & 1 \\
c_{10} & c_{11} & c_{12} & \cdots & c_{1 k} \\
c_{20} & c_{21} & c_{22} & \cdots & c_{2 k} \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
c_{k 0} & c_{k 1} & c_{k 2} & \cdots & c_{k k}
\end{array}\right]\left[\begin{array}{c}
\lambda_{0} \\
\lambda_{1} \\
\lambda_{2} \\
\cdot \\
\cdot \\
\cdot \\
\lambda_{k}
\end{array}\right]=\left[\begin{array}{c}
1 \\
x_{1} \\
x_{2} \\
\cdot \\
\cdot \\
\cdot \\
x_{k}
\end{array}\right], \text { X }} \tag{16a}
\end{gather*}
$$

THEOREM 1. For any $\left(x_{1}, x_{2}, \ldots, x_{k}\right)$ satisfying (5), there is a selection of values $c_{j s}(1 \leq j \leq k, 0 \leq s \leq k)$, such that (16) has a unique solution $\lambda_{0}, \lambda_{1}, \lambda_{2}, \ldots, \lambda_{k}$, and $(\forall s \mid 0 \leq s \leq k) 0 \leq \lambda_{s} \leq 1$.

Proof. We can certainly put the $x_{j} \in\{0,1\}$ into non-increasing order: there is a permutation $\left[r_{1}, r_{2}, \ldots, r_{k}\right]$ of $[1,2, \ldots, k]$, such that

$$
\begin{equation*}
\left\{r_{1}, r_{2}, \ldots, r_{k}\right\}=\{1,2, \ldots, k\} \tag{17}
\end{equation*}
$$

and

$$
\begin{equation*}
1 \geq x_{r_{1}} \geq x_{r_{2}} \geq \ldots \geq x_{r_{k}} \geq 0 \tag{18}
\end{equation*}
$$

Let us now define a corresponding permutation matrix

$$
\begin{align*}
& \boldsymbol{P}=\left[\begin{array}{ccccc}
1 & 0 & 0 & \ldots & 0 \\
0 & \omega_{11} & \omega_{12} & \ldots & \omega_{1 k} \\
0 & \omega_{21} & \omega_{22} & \ldots & \omega_{2 k} \\
. & . & . & . & . \\
. & \cdot & . & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & . \\
0 & \omega_{k 1} & \omega_{k 2} & \ldots & \omega_{k k}
\end{array}\right],  \tag{19}\\
& \omega_{i j}=\left\{\begin{array}{ll}
1 & \text { if }\left\{\begin{aligned}
& \text { either } i=j=0 \\
& \text { or } 1 \leq j=r_{i}
\end{aligned}\right\}
\end{array}\right\} . \tag{20}
\end{align*}
$$

As is easily verified (and well-known),

$$
\begin{equation*}
\boldsymbol{P}^{\top} \boldsymbol{P}=\boldsymbol{P} \boldsymbol{P}^{\top}=\mathbf{I} \tag{21a}
\end{equation*}
$$

where $I$ is the identity matrix and $\boldsymbol{P}^{\top}$ denotes the transpose of $\boldsymbol{P}$; so that

$$
\begin{equation*}
\boldsymbol{P}^{-1}=\boldsymbol{P}^{\top} \tag{21b}
\end{equation*}
$$

Now, from (16),

$$
\begin{equation*}
P M L=P X, \tag{22a}
\end{equation*}
$$

which, with a little thought, reduces to the form

$$
\left[\begin{array}{ccccc}
1 & 1 & 1 & \cdots & 1  \tag{22b}\\
c_{r_{1} 0} & c_{r_{1} 1} & c_{r_{1} 2} & \cdots & c_{r_{1} k} \\
c_{r_{2} 0} & c_{r_{2} 1} & c_{r_{2} 2} & \cdots & c_{r_{2} k} \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
c_{r_{k} 0} & c_{r_{k} 1} & c_{r_{k} 2} & \cdots & c_{r_{k} k}
\end{array}\right]\left[\begin{array}{c}
\lambda_{0} \\
\lambda_{1} \\
\lambda_{2} \\
\cdot \\
\cdot \\
\cdot \\
\lambda_{k}
\end{array}\right]=\left[\begin{array}{c}
1 \\
x_{r_{1}} \\
x_{r_{2}} \\
\cdot \\
\cdot \\
\cdot \\
x_{r_{k}}
\end{array}\right]
$$

Now let us construct the particular matrix

$$
\boldsymbol{P} \boldsymbol{M}=\left[\begin{array}{ccccc}
1 & 1 & 1 & \ldots & 1  \tag{23a}\\
0 & 1 & 1 & \ldots & 1 \\
0 & 0 & 1 & \ldots & 1 \\
. & . & . & . & . \\
. & . & . & . & . \\
. & . & . & . & . \\
0 & 0 & 0 & . . & 1
\end{array}\right]
$$

so that

$$
c_{r_{i} j}=\left\{\begin{array}{ll}
1 & \text { if } \quad j \geq i \geq 1  \tag{23b}\\
0 & \text { otherwise }
\end{array}\right\} .
$$

It then follows from (22), by the usual process of Gaussian elimination, that

$$
\left\{\begin{array}{l}
\lambda_{0}=1-x_{r_{1}}  \tag{2}\\
\lambda_{1}=x_{r_{1}}-x_{r_{2}} \\
\lambda_{2}=x_{r_{2}}-x_{r_{3}} \\
\lambda_{3}=x_{r_{3}}-x_{r_{4}} \\
\cdots \\
\cdots \\
\cdots \\
\lambda_{k-1}=x_{r_{k-1}}-x_{r_{k}} \\
\lambda_{k}=x_{r_{k}}
\end{array}\right\} .
$$

Since, by (18), each of the above $\lambda_{s} \in\{0,1\}$, our theorem follows. 回
It should be observed that, not only have we proved the existence of a suitable simplex in every case, but the proof of the existence theorem shows how the simplex and the coefficients $\lambda_{s}$ can be explicitly constructed.

THEOREM 2. The $k$ ! simplices corresponding to the matrices (23) for all k! possible permutations $\boldsymbol{P}$ or $\boldsymbol{\omega}$ (i.e., $i \rightarrow r_{i}$ ) have no common interior points.

Proof. Consider the simplex described in the proof of Theorem 1. Its vertices are the $k+1$ points $c_{s}=\left(c_{1 s}, c_{2 s}, \ldots, c_{k s}\right)$, with $s=0,1,2, \ldots, k$ (i.e., the columns of $M$, below the first row), satisfying (23). Its interior consists of the points with coordinates satisfying

$$
\begin{equation*}
1>x_{r_{1}}>x_{r_{2}}>\ldots>x_{r_{k}}>0 \tag{25}
\end{equation*}
$$

as is easily seen from the condition that all $0<\lambda_{s}<1$, with (18) and (24).
Now, it is clear that any point $x$ with all its coordinates different can only satisfy an inequality (25) for just one permutation $\omega$, given by (20). In other words, any $\boldsymbol{x}$ interior to one simplex can only be interior to that one simplex and cannot be in the boundary of any such simplex. This proves our theorem. [

Thus we have demonstrated that the simplices defined above constitute a $k$ !-fold dissection of the hypercube.

## 3. The Interpolation

The next question we consider is the accuracy of the simplicial interpolation (13), as compared with the full linear interpolation (9). We shall suppose that the function $F\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)$, and the derived function $f\left(x_{1}, x_{2}, \ldots, x_{k}\right)$ in the lattice cell enclosing the given point $\left(\xi_{1}, \xi_{2}, \ldots, \xi_{k}\right)$, have Taylor expansions of sufficient length: by (6) and (7),

$$
\begin{align*}
f_{x} & =f\left(x_{1}, x_{2}, \ldots, x_{k}\right)=F\left(p_{1} \mu+x_{1} \mu, p_{2} \mu+x_{2} \mu, \ldots, p_{k} \mu+x_{k} \mu\right) \\
& =\left\{1+\mu\left(\sum_{i=1}^{k} x_{i} \frac{\partial}{\partial \xi_{i}}\right)+\frac{\mu^{2}}{2}\left(\sum_{i=1}^{k} x_{i} \frac{\partial}{\partial \xi_{i}}\right)^{2}+\frac{\mu^{2}}{3!}\left(\sum_{i=1}^{k} x_{i} \frac{\partial}{\partial \xi_{i}}\right)^{3}+\ldots\right\} f_{0} \tag{26}
\end{align*}
$$

where

$$
\begin{equation*}
f_{0}=f(0,0, \ldots, 0)=F\left(p_{1} \mu, p_{2} \mu, \ldots, p_{k} \mu\right) \tag{27}
\end{equation*}
$$

It follows that

$$
\begin{equation*}
f_{c}=\left\{1+\mu\left(\sum_{i=1}^{k} c_{i} \frac{\partial}{\partial \xi_{i}}\right)+\frac{\mu^{2}}{2}\left(\sum_{i=1}^{k} c_{i} \frac{\partial}{\partial \xi_{i}}\right)^{2}+\frac{\mu^{2}}{3!}\left(\sum_{i=1}^{k} c_{i} \frac{\partial}{\partial \xi_{i}}\right)^{3}+\ldots\right\} f_{0} \tag{28}
\end{equation*}
$$

whence, the full interpolation (9) yields

$$
\begin{align*}
& \Phi_{x}=\sum_{c_{1}=0}^{1} \sum_{c_{2}=0}^{1} \ldots \sum_{c_{k}=0}^{1}\left(\prod_{j=1}^{k}\left[c_{j} x_{j}+\left(1-c_{j}\right)\left(1-x_{j}\right)\right]\right) \\
& \times\left\{1+\mu\left(\sum_{i=1}^{k} c_{i} \frac{\partial}{\partial \xi_{i}}\right)+\frac{\mu^{2}}{2}\left(\sum_{i=1}^{k} c_{i} \frac{\partial}{\partial \xi_{i}}\right)^{2}+\frac{\mu^{2}}{3!}\left(\sum_{i=1}^{k} c_{i} \frac{\partial}{\partial \xi_{i}}\right)^{3}+\ldots\right\} f_{0} \tag{29}
\end{align*}
$$

We are interested in calculating the error in this interpolation,

$$
\begin{align*}
& \Phi_{x}-f_{x}=\sum_{t=0}^{\infty} \frac{\mu^{t}}{t!}\left\{\sum_{c_{1}=0}^{1} \sum_{c_{2}=0}^{1} \ldots \sum_{c_{k}=0}^{1}\right. \\
& \left.\left(\prod_{j=1}^{k}\left[c_{j} x_{j}+\left(1-c_{j}\right)\left(1-x_{j}\right)\right]\right)\left(\sum_{i=1}^{k} c_{i} \frac{\partial}{\partial \xi_{i}}\right)^{t} f_{0}-\left(\sum_{i=1}^{k} x_{i} \frac{\partial}{\partial \xi_{i}}\right)^{t} f_{0}\right\} . \tag{30}
\end{align*}
$$

A little algebraic manipulation shows that the first two terms ( $t=0$ and $t=1$ ) vanish identically, and the remaining terms yield

$$
\begin{align*}
\Phi_{x}-f_{x}= & \frac{\mu^{2}}{2}\left\{\sum_{c_{1}=0}^{1} \sum_{c_{2}=0}^{1} \ldots \sum_{c_{k}=0}^{1}\left(\prod_{j=1}^{k}\left[c_{j} x_{j}+\left(1-c_{j}\right)\left(1-x_{j}\right)\right]\right)\right. \\
& \times\left(\sum_{i=1}^{k} c_{i} 2^{2} \frac{\partial^{2} f_{0}}{\partial \xi_{i}^{2}}+\sum_{i=1}^{k .} \sum_{\text {All } h \neq i} c_{h} c_{i} \frac{\partial^{2} f_{0}}{\partial \xi_{h} \partial \xi_{i}}\right) \\
& \left.\quad-\left(\sum_{i=1}^{k} x_{i}^{2} \frac{\partial^{2} f_{0}}{\partial \xi_{i}^{2}}+\sum_{i=1}^{k} \sum_{\text {All } h \neq i} x_{h} x_{i} \frac{\partial^{2} f_{0}}{\partial \xi_{h} \partial \xi_{i}}\right)\right\}+O\left(\mu^{3}\right) \\
= & \frac{\mu^{2}}{2}\left\{\sum_{i=1}^{k}\left(x_{i}-x_{i}^{2}\right) \frac{\partial^{2} f_{0}}{\partial \xi_{i}^{2}}\right\}+O\left(\mu^{3}\right) . \tag{31}
\end{align*}
$$

Here, we make the usual and justifiable assumption that the lattice constant $\mu$ is small.

Now, turning to the simplicial interpolation (13), in the particular case determined by (23) and (24), we see that (for $s=0,1,2, \ldots, k$ ) $P_{s}$ has components ' 1 ' in coordinates with indices $r_{t}(1 \leq t \leq s)$ and components ' 0 ' in all other coordinates [e.g., $P_{0}=(0,0, \ldots, 0) ; P_{1}=(0, \ldots, 1, \ldots, 0)$, with the single ' 1 ' in position $r_{1} ; P_{2}=(0, \ldots, 1, \ldots, 1, \ldots, 0)$, with the two ' 1 ' in positions $r_{1}$ and $r_{2}$; and $\left.P_{k}=(1,1, \ldots, 1)\right]$. Thus, by (28),

$$
\begin{align*}
\Psi_{x} & =F\left(P_{0}\right)+\sum_{s=1}^{k} x_{r_{s}}\left[F\left(P_{s}\right)-F\left(P_{s-1}\right)\right]=f_{0}+\sum_{s=1}^{k} x_{r_{s}}\left(f_{c_{s}}-f_{c_{s-1}}\right) \\
& =f_{0}+\sum_{s=1}^{k}\left\{\mu x_{r_{s}} \frac{\partial f_{0}}{\partial \xi_{r_{s}}}+\frac{\mu^{2}}{2} x_{r_{s}}\left(\frac{\partial^{2} f_{0}}{\partial \xi_{r_{s}}{ }^{2}}+\sum_{\text {All }} \frac{\partial^{2} f_{0}}{\partial \xi_{r_{t}} \partial \xi_{r_{s}}}\right)\right\}+O\left(\mu^{3}\right) \\
& =f_{0}+\sum_{s=1}^{k}\left\{\mu x_{r_{s}} \frac{\partial f_{0}}{\partial \xi_{r_{s}}}+\frac{\mu^{2}}{2} x_{r_{s}}\left(\frac{\partial^{2} f_{0}}{\partial \xi_{r_{s}}{ }^{2}}+2 \sum_{t=1}^{s-1} \frac{\partial^{2} f_{0}}{\partial \xi_{r_{t}} \partial \xi_{r_{s}}}\right)\right\}+O\left(\mu^{3}\right) \tag{32}
\end{align*}
$$

Recalling that the $r_{t}$ form a permutation of $\{1,2, \ldots, k\}$-see (17)-we observe that, by (26), reordered as in (18),

$$
\begin{align*}
& \Psi_{x}-f_{x}=\frac{\mu^{2}}{2} \sum_{s=1}^{k} x_{r_{s}}\left\{\left[\frac{\partial^{2} f_{0}}{\partial \xi_{r_{s}}^{2}}+2 \sum_{t=1}^{s-1} \frac{\partial^{2} f_{0}}{\partial \xi_{r_{t}} \partial \xi_{r_{s}}}\right]\right. \\
& \left.-\left[x_{r_{s}} \frac{\partial^{2} f_{0}}{\partial \xi_{r_{s}}^{2}}+2 \sum_{t=1}^{s-1} x_{r_{t}} \frac{\partial^{2} f_{0}}{\partial \xi_{r_{t}} \partial \xi_{r_{s}}}\right]\right\}+O\left(\mu^{3}\right) \\
& =\frac{\mu^{2}}{2} \sum_{s=1}^{k} x_{r_{s}}\left\{\left(1-x_{r_{s}}\right) \frac{\partial^{2} f_{0}}{\partial \xi_{r_{s}}^{2}}+2 \sum_{t=1}^{s-1}\left(1-x_{r_{t}}\right) \frac{\partial^{2} f_{0}}{\partial \xi_{r_{t}} \partial \xi_{r_{s}}}\right\}+O\left(\mu^{3}\right), \tag{33}
\end{align*}
$$

where we note that, by (18),

$$
\begin{equation*}
(\forall t \mid t<s) \quad 0 \leq 1-x_{r_{t}} \leq 1-x_{r_{s}} \leq 1 . \tag{34}
\end{equation*}
$$

Comparing (31) with (33), we see that the error produced by the simplicial interpolation is of the same order as that produced by the full interpolation on the lattice-cell, namely, $O\left(\mu^{2}\right)$, and if we have a bound

$$
\begin{equation*}
(\forall i, j \mid 1 \leq i \leq k, 1 \leq j \leq k)\left|\frac{\partial^{2} f_{0}}{\partial \xi_{i} \partial \xi_{j}}\right|<K, \tag{35}
\end{equation*}
$$

on the second derivatives of $F$, then, by (31), since, as is easily verified,

$$
\begin{gather*}
\max _{0 \leq x \leq 1} x(1-x)=\frac{1}{4},  \tag{36}\\
\left|\Phi_{x}-f_{x}\right|<\frac{K \mu^{2}}{2} \sum_{i=1}^{k} x_{i}\left(1-x_{i}\right)+O\left(\mu^{3}\right) \leq \frac{k K \mu^{2}}{8}+O\left(\mu^{3}\right) ; \tag{37}
\end{gather*}
$$

while, by (33), with (36),

$$
\begin{align*}
\left|\Psi_{x}-f_{x}\right| & <\frac{K \mu^{2}}{2} \sum_{s=1}^{k} x_{r_{s}}\left[\left(1-x_{r_{s}}\right)+2 \sum_{t=1}^{s-1}\left(1-x_{r_{t}}\right)\right]+O\left(\mu^{3}\right) \\
& \leq \frac{K \mu^{2}}{2} \sum_{s=1}^{k} x_{r_{s}}\left[\left(1-x_{r_{s}}\right)+2 \sum_{t=1}^{s-1}\left(1-x_{r_{s}}\right)\right]+O\left(\mu^{3}\right) \\
& =\frac{K \mu^{2}}{2} \sum_{s=1}^{k}(2 s-1) x_{r_{s}}\left(1-x_{r_{s}}\right)+O\left(\mu^{3}\right) \\
& \leq \frac{k^{2} K \mu^{2}}{8}+O\left(\mu^{3}\right) \tag{38}
\end{align*}
$$

Summing up our results in (31), (33), (37), and (38) we have:
THEOREM 3. Applying the bounds (35), we have that

$$
\begin{equation*}
\left|\Phi_{x}-f_{x}\right|=\frac{\mu^{2}}{2}\left|\sum_{i=1}^{k}\left(x_{i}-x_{i}^{2}\right) \frac{\partial^{2} f_{0}}{\partial \xi_{i}^{2}}\right|+O\left(\mu^{3}\right)<\frac{k K \mu^{2}}{8}+O\left(\mu^{3}\right) \tag{39}
\end{equation*}
$$

and

$$
\begin{align*}
\left|\Psi_{x}-f_{x}\right| & =\frac{\mu^{2}}{2}\left|\sum_{s=1}^{k} x_{r_{s}}\left\{\left(1-x_{r_{s}}\right) \frac{\partial^{2} f_{0}}{\partial \xi_{r_{s}}^{2}}+2 \sum_{t=1}^{s-1}\left(1-x_{r_{t}}\right) \frac{\partial^{2} f_{0}}{\partial \xi_{r_{t}} \partial \xi_{r_{s}}}\right\}\right| \\
& <\frac{k^{2} K \mu^{2}}{8}+O\left(\mu^{3}\right) \tag{40}
\end{align*}
$$

The conclusion is, therefore, that we will obtain comparable accuracy with the much quicker simplicial interpolation if we scale the lattice constant $\mu$ by a factor $\sqrt{k}$ :

$$
\begin{equation*}
\mu_{\mathrm{SIMPLEX}}=\frac{\mu_{\mathrm{LATTICE}}}{\sqrt{k}} \tag{41}
\end{equation*}
$$

Of course, if this means actually computing $k^{k / 2}$ times as many data, there is little advantage in the simplicial interpolation, but this is not always a requirement; it may be that already-assembled data suffice to yield the required accuracy, and then, as we shall see, the simplicial interpolation is faster to execute.

## 4. Timing Estimates

At this point, we assume that the values of the function $F$ at the lattice points are already computed and given. The full interpolation (9) then takes three operations:
(i) Determination of the $p_{j}$ and $x_{j}$, for $j=1,2, \ldots, k$, by (4); i.e., identification of the particular lattice cell in which the point $\left(\xi_{1}, \xi_{1}, \ldots, \xi_{k}\right)$ lies.
(ii) Computation of the $2^{k}$ coefficients

$$
\begin{equation*}
\prod_{j=1}^{k}\left[c_{j} x_{j}+\left(1-c_{j}\right)\left(1-x_{j}\right)\right] \tag{42}
\end{equation*}
$$

as defined in (8) and (9). The factors are each either $x_{j}$ or $1-x_{j}$, taken in all combinations.
(iii) Evaluation of the interpolated value $\Phi_{\boldsymbol{x}}$ from (9).

Let a denotes the time required for an addition or subtraction, $\mathbf{m}$ denotes the time required for a multiplication or division, and t denotes the relatively short time required for a test, shift, store, or retrieve operation. Then the operation (i) requires, for each coordinate, two retrievals, one division, two split operations, and two storage operations (obtaining the integer and fractional parts), for a total time

$$
\begin{equation*}
\mathbb{T}_{k \text {-CUBE }(\mathrm{i})}=k(\mathbf{m}+6 \mathbf{t}) \tag{43}
\end{equation*}
$$

To perform the operation (ii), we must, for greatest efficiency, proceed inductively. First, we retrieve and store $x_{1}$, subtract it from 1 and store $1-x_{1}$ : this takes time a +3 t. If the total time for the operation (ii) for $k=j-1$ is $\mathbb{T}_{(j-1)-\operatorname{CUBE}(i i)}$, then we know that $\mathbb{T}_{1-\operatorname{CUBE}(\mathrm{ii})}=\mathbf{a}+3 \mathbf{t}$. Also, to
get $\mathbb{T}_{j \text {-CUBE(ii) }}$, we must first retrieve $x_{j}$, subtract it from 1 , and, for each of the $2^{j-1}$ stored coefficients, multiply it by $1-x_{j}$ and store the result; this takes time $2^{j-1} \mathbf{m}+\mathbf{a}+\left(2^{j-1}+1\right)$ t. Then we must retrieve $x_{j}$ again and repeat the multiplications by it, taking time $2^{j-1} \mathbf{m}+\left(2^{j-1}+1\right) t$ more. Thus,

$$
\left.\begin{array}{r}
\mathbb{T}_{1-\mathrm{CUBE}(\mathrm{ii})}=\mathbf{a}+3 \mathbf{t}  \tag{44}\\
\mathbb{T}_{j \text {-CUBE }(\mathrm{ii})}=\mathbb{P}_{(j-1)-\mathrm{CUBE}(\mathrm{ii})}+2^{j \mathrm{~m}}+\mathbf{a}+\left(2^{j}+2\right) \mathbf{t}
\end{array}\right\}
$$

It follows that

$$
\begin{align*}
\mathbb{T}_{k-\mathrm{CUBE}(\mathrm{ii})} & =\sum_{j=2}^{k}\left[2^{j} \mathrm{~m}+\mathbf{a}+\left(2^{j}+2\right) \mathbf{t}\right]+\mathbf{a}+3 \mathbf{t} \\
& =\left(2^{k+1}-4\right) \mathbf{m}+k \mathbf{a}+\left(2^{k+1}+2 k-3\right) \mathbf{t} \\
& \sim 2^{k+1}(\mathbf{m}+\mathbf{t}) \tag{45}
\end{align*}
$$

Finally, to perform the operation (iii) takes additional time

$$
\begin{align*}
\mathbb{T}_{k-\text { CUBE }(\mathrm{iii})} & =2^{k}(\mathbf{m}+\mathbf{t})+\left(2^{k}-1\right)(2 \mathbf{t}+\mathbf{a}) \\
& \sim 2^{k}(\mathbf{m}+3 \mathbf{t}) \tag{46}
\end{align*}
$$

as is easily verified. Thus, the total time for full interpolation is

$$
\begin{align*}
\mathbb{T}_{k \text {-CUBE }} & =\mathbb{T}_{k \text {-CUBE }(\mathrm{i})}+\mathbb{T}_{k \text {-CUBE }(\mathrm{ii})}+\mathbb{T}_{k \text {-CUBE }(\mathrm{iii})} \\
& =2^{k}(3 \mathrm{~m}+5 \mathbf{t})+k(\mathrm{~m}+\mathbf{a}+8 \mathbf{t})-(4 \mathrm{~m}+3 \mathbf{t}) \\
& \sim 2^{k}(3 \mathrm{~m}+5 \mathbf{t}) \tag{47}
\end{align*}
$$

Turning to simplicial interpolation (13), we see that there are also three operations:
(i) Determination of the $p_{j}$ and $x_{j}$, for $j=1,2, \ldots, k$, by (4), exactly as for full interpolation.
(ii) Ordering of the coordinates $x_{j}$, as indicated in (18).
(iii) Evaluation of the interpolated value $\Psi_{x}$ from (13).

Proceeding as for full interpolation, we first see that

$$
\begin{equation*}
\mathbb{T}_{k \text {-SIMPLEX }(\mathbf{i})}=k(\mathbf{m}+6 \mathbf{t}) . \tag{48}
\end{equation*}
$$

It is well-known that, if we use an efficient sorting method (e.g., heap-sort, merge-sort, or quick sort), we can achieve [see D. E. Knuth, The Art of Computer Programming, Vol. 3, Sorting and Searching (Addison-Wesley, Reading, MA, 1973) p. 149]

$$
\begin{equation*}
\mathbb{T}_{k \text {-SIMPLEX(ii) }}=\left(18 \log _{2} k+38\right) k \mathbf{t} . \tag{49}
\end{equation*}
$$

Finally, the first line of (32) indicates that

$$
\begin{equation*}
\mathbb{T}_{k \text {-SIMPLEX(iii) }}=k(\mathbf{m}+2 \mathbf{a}+5 \mathbf{t}) \tag{50}
\end{equation*}
$$

Thus, the total time for simplicial interpolation is

$$
\begin{align*}
\mathbb{T}_{k \text {-SIMPLEX }} & =\mathbb{T}_{k \text {-SIMPLEX }(\mathbf{i})}+\mathbb{T}_{k \text {-SIMPLEX }(i i)}+\mathbb{T}_{k \text {-SIMPLEX }(i i i)} \\
& =18 k\left(\log _{2} k\right) \mathrm{t}+k(2 \mathrm{~m}+2 \mathrm{a}+49 \mathrm{t}) \\
& \sim 18 k\left(\log _{2} k\right) \mathrm{t} \tag{51}
\end{align*}
$$

Summing up these results, we have:
THEOREM 4. With timing constants, $\mathbf{t}, \mathbf{a}$, and $\mathbf{m}$, defined as above, we have that

$$
\begin{align*}
\mathbb{T}_{k \text {-CUBE }} & =2^{k}(3 \mathbf{m}+5 \mathbf{t})+k(\mathbf{m}+\mathbf{a}+8 \mathbf{t})-(4 \mathbf{m}+3 \mathbf{t}) \\
& \sim 2^{k}(3 \mathbf{m}+5 \mathbf{t}) \tag{52}
\end{align*}
$$

and

$$
\begin{align*}
\mathbb{T}_{k \text {-SIMPLEX }} & =18 k\left(\log _{2} k\right) \mathbf{t}+k(2 \mathbf{m}+2 \mathbf{a}+49 \mathbf{t}) \\
& \sim 18 k\left(\log _{2} k\right) \mathbf{t} \tag{53}
\end{align*}
$$

Evidently, for large values of $k$, the simplicial interpolation is far more efficient than the full lattice-cell interpolation. Some comparative examples, for somewhat typical timing constants $t=1, \mathbf{a}=8$, and $\mathbf{m}=24$, and for dimensions $k=2,3,4,6,8,16,32,128,1024$, are given below.

| $k$ | $\mathbb{T}_{k \text {-CUBE }}$ | $\mathbb{T}_{k \text {-SIMPLEX }}$ | RATIO |
| :---: | ---: | ---: | ---: |
| 2 | 289 | 262 | 0.907 |
| 3 | 637 | 424.6 | 0.667 |
| 4 | 1,293 | 596 | 0.461 |
| 6 | 5,069 | 957.2 | 0.189 |
| 8 | 19,933 | 1,336 | 0.067 |
| 16 | $5,046,813$ | 2,960 | $5.865 \times 10^{-4}$ |
| 32 | $3.307 \times 10^{11}$ | 6,496 | $1.964 \times 10^{-8}$ |
| 128 | $2.620 \times 10^{40}$ | 30,592 | $1.168 \times 10^{-36}$ |
| 1024 | $1.384 \times 10^{310}$ | 300,032 | $2.168 \times 10^{-305}$ |

