Estimating interconnection lengths in three-dimensional computer systems

Dirk Stroobandt*

Department of Electronics
and Information Systems
University of Ghent
B-9000 Gent, Belgium
dstr@elis.rug.ac.be

Jan Van Campenhout

Department of Electronics
and Information Systems
University of Ghent
B-9000 Gent, Belgium
jvc@elis.rug.ac.be

Abstract— Three-dimensional systems can be used to alleviate the pin limitation problem. The benefits can be characterised by estimating the important layout properties of electronic designs, such as space requirements and interconnection length values. For a two-dimensional placement, Donath found an upper bound for the average interconnection length that follows the trends of experimentally obtained average lengths [1]. Yet, this upper bound deviates from the experimentally obtained value by a factor of approximately 2 which is not sufficiently accurate for some applications. In this paper, we first extend Donath's technique to a three-dimensional placement. We then compute a significantly more accurate estimate by taking into account the inherent features of the optimal placement process.

I. Introduction

Creating a physical layout of an electronic design involves the placement and interconnection of elementary blocks onto a carrier. Important properties of such layouts are area (or space) requirements and interconnection lengths. The ability to predict these properties, without actually having to perform the placement and routing itself, is important for the following reasons:

- Most placement algorithms use estimates for interconnection lengths and area requirements to limit the search in the solution space [2, 3].
- The predictions offer a way to gain a more fundamental insight in the placement of designs on different carriers.

There have been many attempts to predict area requirements and interconnection lengths. A first upper bound for interconnection lengths has been found by Sutherland and Oestreicher [4]. It is based on a random placement and therefore yields excessively large estimates. Donath [1, 5] found that a hierarchical placement technique gives a much better estimate of interconnection lengths and his results have been used by several other researchers [6, 7]. Recently, we have extended Donath's technique to include knowledge about an optimal placement process [8]. This resulted in a significant improvement of the interconnection length estimation for designs placed in a two-dimensional system.

Current research more and more evolves towards the implementation of designs in three-dimensional systems [9, 10, 11, 12]. Chip sizes are scaled down rapidly and the number of transistors per chip increases likewise. The first evolution leads to a smaller available pin count since the perimeter of the chip decreases while the I/O-pad size remains equal. The second one, on the other hand, imposes a much higher need for chip terminals to provide the necessary communication with other components. This leads to highly pin limited designs. The problem can be alleviated by using area-I/O instead of perimeter-I/O [13], thus requiring nearly three-dimensional systems. It is not yet clear what these systems should look like and how they should be implemented (totally electronic or partially optical?). Therefore a theoretical estimation of the most important and necessary features for such systems is badly needed. Accurate estimates can verify the usefulness of a proposed architecture before it is actually built. These estimates can also be used to choose the technology in which to implement the third dimension.

Our research is situated in modelling three-dimensional anisotropic architectures [14, 15, 16]. Until recently we used Donath's technique. In this paper, we show that an extension of this technique leads to more accurate results since knowledge about the behaviour of an optimal placement process is included. In the next section we will explain the key issues of Donath's placement technique, but extended to three dimensions. We will then show a way to improve Donath's method. As a result we will compare the new estimates with those obtained by Donath and with the estimates for a two-dimensional implementation.

*Supported as Research Assistant with the Belgian National Fund for Scientific Research.

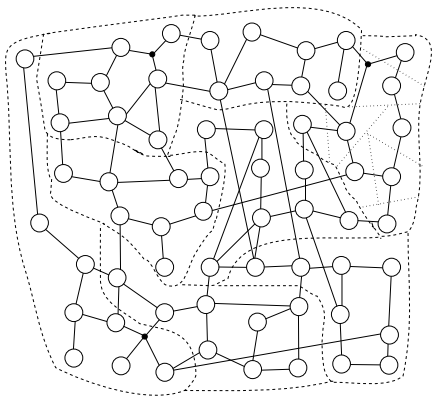

Fig. 1. Recursive partitioning scheme of the design.

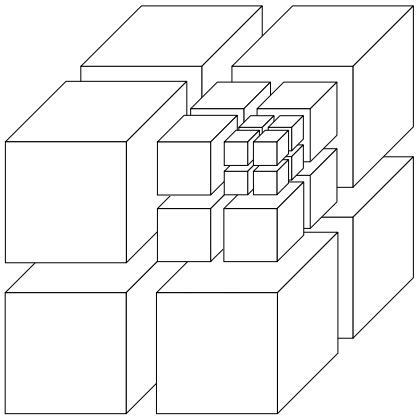

Fig. 2. Recursive partitioning scheme of the physical architecture.

II. Donath's Technique

Donath estimates the average interconnection length using a relationship between the number of elementary blocks B in a module of a partitioned design, and the number of the module's external connections (pins) P, known as Rent's rule [17, 18]:

$$P = CB^r, \tag{1}$$

where C is the average number of interconnections per elementary block, and r is called the Rent exponent. This exponent is a measure of the interconnection complexity of the design. Its value is bounded by 0 and 1, with increasing values for increasing interconnection complexity. Generally, r varies from 0.4 for simple regular designs (such as Random Access Memories), up to 0.8 for complex designs (such as fast full custom VLSI designs). The validity of Rent's rule is a result of the fact that designers tend to build their designs hierarchically, imposing the same complexity on the design at each level of hierarchy.

Donath's theory is also based on a hierarchical placement of the design into a Manhattan grid [1]. We define an *optimal placement* as one that minimizes the total interconnection length. This optimal placement should be correctly modelled by the hierarchical placement method. In this paper we describe Donath's technique in three dimensions. The design will therefore be placed in a cubic-shaped three-dimensional Manhattan grid. For that purpose, the design is partitioned hierarchically into sub-designs. Each sub-design at a hierarchical level consists itself of eight sub-designs (of equal size) at the next (lower) level of hierarchy (fig. 1). We thus assume that the number of gates in the logic design is a power of 8 (there are 8^K gates, with K the number of hierarchical levels).

The partitioning of the design into eight sub-designs of equal size should be done in such a way that the partition satisfies Rent's rule. That is, we want to keep the number of interconnections between the sub-designs as low as possible. This ensures that our placement scheme is a good model for the optimal placement of the design. It is indeed obvious that, in an optimal placement, densely interconnected logic gates are placed as close as possible, resulting in clusters of such gates. Among clusters, there are fewer interconnections. A placement scheme that keeps the number of interconnections between the sub-designs as low as possible leads to many short interconnections and few long ones. This behaviour is modelled accurately by Rent's rule.

The design is placed in a physical architecture which can be modelled as a cubic-shaped Manhattan grid. In this grid, each grid-point corresponds to a location where one logic gate of the design can be placed. The grid-lines correspond to the channels in which the connections between the gates can be routed. All lengths are measured using a Manhattan metric. As is the design, so the cubic-shaped grid is partitioned into eight sub-cubes of equal size (fig. 2).

In Donath's partitioning and placement scheme, each sub-design is paired recursively with a sub-cube until all gates are paired with (and essentially assigned to) exactly one grid location. The recursion levels will be numbered 0 (eight sub-designs consisting of only one logic gate) up to $K-1$ (eight sub-designs that constitute the whole design). Note that external interconnections are not included in the estimations (these interconnections belong to level K).

Given the above model for the design, the physical architecture and the placement technique, we want to find the average interconnection length. We shall do this by calculating the average number of interconnections $\bar{n}_k$ and the average length of the interconnections $\bar{l}_k$ at every hierarchical level k ($0 \leq k \leq K-1$). The average interconnection length $\mathcal{L}$, computed over all hierarchical levels, is then given by

$$\mathcal{L} = \frac{\sum_{k=0}^{K-1} \bar{n}_k \bar{l}_k}{\sum_{k=0}^{K-1} \bar{n}_k}. \tag{2}$$

III. Donath's Average Interconnection Length in Three Dimensions

The expected number of interconnections at each level of the hierarchy can be calculated using Rent's rule [1]. Rent's rule estimates the number of terminals of a sub-design of B gates as (equation 1):

$$T = CB^r \qquad (0 < r < 1).$$

Suppose we have a total of G gates, divided into clusters of size B. The total number of terminals for sub-designs of size B is then given by

$$T_{tot}(B) = CB^r \frac{G}{B} = CGB^{r-1}.$$

For three-dimensional partitioning, B has the value 8^k and $G = 8^K$ where k is the recursion level in the partitioning scheme and K is the number of recursion levels ($0 \leq k \leq K-1$). The number of terminals T_k belonging to recursion level k is given by

$$T_k = T_{tot}(8^k) - T_{tot}(8^{k+1}). \tag{3}$$

Note that terminals of sub-designs of size $8^{(k+1)}$ are also terminals of sub-designs of size 8^k. This means that $T_{tot}(8^k)$ includes all terminals of sub-designs of size $8^{(k+1)}$. Equation 3 expresses the fact that these higher level terminals should be subtracted from $T_{tot}(8^k)$.

The expected number of interconnections (nets) at each level of the hierarchy ($\bar{n}_k$) is a fraction α of the number of terminals T_k. The value of α is 1/2 if each net has just two terminals, and somewhat smaller in the normal case where there are some multi-terminal nets. Since the range of variation is relatively small for α, it is a good approximation to assume that variations of α are insignificant [1]. We then have

$$\bar{n}_k = \alpha T_k = \alpha C 8^K \left(1 - 8^{r-1}\right) 8^{k(r-1)}. \tag{4}$$

We next determine the average interconnection length $\bar{l}_k$ at hierarchical level k. The interconnections belonging to hierarchical level k are those interconnections between gates belonging to the same $(k+1)$-th level hierarchical sub-design, but to different k-th level hierarchical sub-designs. Those interconnections connect two gates placed in different cubes at hierarchical level k. Three different combination classes are possible: either the cubes are adjacent (A) or they are diagonally opposed to a nearby (N) location or they are diagonally opposed to a remote (R) location (fig. 3). We will call a combination of the first class an A-combination, of the second one a N-combination, and of the last one a R-combination[1]. For each of these combinations, we compute the average interconnection length (denoted as $\bar{l}_{k,a}$, $\bar{l}_{k,n}$, and $\bar{l}_{k,r}$ respectively). It is assumed that the starting points and the endpoints of the interconnections between two cubes are uniformly distributed over those cubes.

[1]There are twelve A-combinations, twelve N-combinations and four R-combinations in a cube.

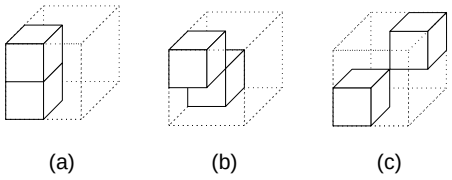

Fig. 3. Three possible combination classes at a hierarchical level k: A-combinations (a), N-combinations (b), and R-combinations (c).

In the remaining of this paper, we define an *interconnection length distribution* as a collection of values, indicating, for each length l, how many interconnections have length l. The sum of these values over all lengths l equals the total number of interconnections. By multiplying this distribution with a normalisation constant one can make this sum equal to 1. We will then call this a *normalised interconnection length distribution*. A normalised interconnection length distribution denotes, for each length l, the probability that an interconnection has length l.

The normalised interconnection length distributions can be calculated as in [7] but for the three-dimensional cubic-shaped combinations. These distributions are denoted as $P_{k,a}(l)$ for A-combinations, $P_{k,n}(l)$ for N-combinations and $P_{k,r}(l)$ for R-combinations. They are presented here without proof for combinations of cubes with size $\lambda = 2^k$.

$$P_{k,a}(l) = \begin{cases} \frac{1}{30\lambda^6} P^i_{k,a} & (i\lambda \leq l < (i+1)\lambda,\ i = 0,\ldots,3) \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

with

$$\begin{aligned}
P^0_{k,a} &= \left(20\,l^3 + 10\,l\right)\lambda^2 + \left(10\,l^2 - 10\,l^4\right)\lambda + l^5 - 5\,l^3 + 4\,l, \\
P^1_{k,a} &= 74\,\lambda^5 - 260\,l\,\lambda^4 + \left(340\,l^2 - 30\right)\lambda^3 - 3\,l^5 + 15\,l^3 \\
&\quad + \left(110\,l - 180\,l^3\right)\lambda^2 + \left(40\,l^4 - 80\,l^2 + 16\right)\lambda - 12\,l, \\
P^2_{k,a} &= -\left[758\,\lambda^5 - 1420\,l\,\lambda^4 + \left(980\,l^2 - 310\right)\lambda^3\right. \\
&\quad + \left(360\,l - 320\,l^3\right)\lambda^2 + \left(50\,l^4 - 130\,l^2 + 32\right)\lambda \\
&\quad \left. -3\,l^5 + 15\,l^3 - 12\,l\right], \\
P^3_{k,a} &= \prod_{j=-2}^{2} (4\,\lambda - l + j)
\end{aligned}$$

for an A-combination,

$$P_{k,n}(l) = \begin{cases} \frac{1}{60\lambda^6} P^i_{k,n} & (i\lambda \leq l < (i+1)\lambda,\ i = 0,\ldots,4) \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

with

$$\begin{aligned}
P^0_{k,n} &= \left(5\,l^4 - 5\,l^2\right)\lambda - l^5 + 5\,l^3 - 4\,l, \\
P^1_{k,n} &= -\left[25\,\lambda^5 - 105\,l\,\lambda^4 + \left(170\,l^2 - 45\right)\lambda^3\right.
\end{aligned}$$

$$+\left(115\,l-130\,l^3\right)\lambda^2+\left(40\,l^4-90\,l^2+20\right)\lambda$$
$$-4\,l^5+20\,l^3-16\,l\big],$$
$$P^2_{k,n}=775\,\lambda^5-1655\,l\,\lambda^4+\left(1350\,l^2-475\right)\lambda^3$$
$$+\left(605\,l-510\,l^3\right)\lambda^2+\left(90\,l^4-240\,l^2+60\right)\lambda$$
$$-6\,l^5+30\,l^3-24\,l,$$
$$P^3_{k,n}=-\big[3275\,\lambda^5-4555\,l\,\lambda^4+\left(2430\,l^2-1055\right)\lambda^3$$
$$+\left(865\,l-630\,l^3\right)\lambda^2+\left(80\,l^4-230\,l^2+60\right)\lambda$$
$$-4\,l^5+20\,l^3-16\,l\big],$$
$$P^4_{k,n}=\prod_{j=-2}^{2}\left(5\,\lambda-l+j\right)$$

for a N-combination, and

$$P_{k,r}(l)=\begin{cases}\frac{1}{120\lambda^5}P^i_{k,r} & (i\lambda\le l<(i+1)\lambda,\ i=0,\ldots,5)\\ 0 & \text{otherwise,}\end{cases}\tag{7}$$

with

$$P^0_{k,r}=l^5-5\,l^3+4\,l,$$
$$P^1_{k,r}=6\,\lambda^5-30\,l\,\lambda^4+\left(60\,l^2-30\right)\lambda^3+\left(90\,l-60\,l^3\right)\lambda^2$$
$$+\left(30\,l^4-90\,l^2+24\right)\lambda-5\,l^5+25\,l^3-20\,l,$$
$$P^2_{k,r}=-2\,\big[237\,\lambda^5-585\,l\,\lambda^4+\left(570\,l^2-285\right)\lambda^3$$
$$+\left(405\,l-270\,l^3\right)\lambda^2+\left(60\,l^4-180\,l^2+48\right)\lambda$$
$$-5\,l^5+25\,l^3-20\,l\big],$$
$$P^3_{k,r}=2\,\big[2193\,\lambda^5-3465\,l\,\lambda^4+\left(2130\,l^2-1065\right)\lambda^3$$
$$+\left(945\,l-630\,l^3\right)\lambda^2+\left(90\,l^4-270\,l^2+72\right)\lambda$$
$$-5\,l^5+25\,l^3-20\,l\big],$$
$$P^4_{k,r}=-\big[10974\,\lambda^5-12270\,l\,\lambda^4+\left(5340\,l^2-2670\right)\lambda^3$$
$$+\left(1710\,l-1140\,l^3\right)\lambda^2+\left(120\,l^4-360\,l^2+96\right)\lambda$$
$$-5\,l^5+25\,l^3-20\,l\big],$$
$$P^5_{k,r}=\prod_{j=-2}^{2}\left(6\,\lambda-l+j\right)$$

for a R-combination.

The average interconnection length for each combination on hierarchical level k is now found as $\bar{l}_{k,C}$ ($C\in\{a,n,r\}$)

$$\bar{l}_{k,C}=\frac{\sum_{l=0}^{6\lambda}l P_{k,C}(l)}{\sum_{l=0}^{6\lambda}P_{k,C}(l)}.\tag{8}$$

We obtain

$$\bar{l}_{k,a}=\frac{5\lambda}{3}-\frac{2}{3\lambda},\qquad \bar{l}_{k,n}=\frac{7\lambda}{3}-\frac{1}{3\lambda},\qquad \bar{l}_{k,r}=3\lambda,$$

with $\lambda=2^k$.

Since there are twelve A-, twelve N-, and four R-combinations, the total average interconnection length $\bar{l}_k$ at the hierarchical level k is given by

$$\bar{l}_k=\frac{12\,\bar{l}_{k,a}+12\,\bar{l}_{k,n}+4\,\bar{l}_{k,r}}{28}.\tag{9}$$
$$=\frac{15\lambda}{7}-\frac{3}{7\lambda}\tag{10}$$

Combining equations 2, 4, and 10 yields

$$\mathcal{L}=\frac{15\,H(K,r,2)-3\,H(K,r,4)}{7\,H(K,r,3)},\tag{11}$$

with

$$H(K,r,x)=\frac{2^{K(3r-x)}-1}{2^{3r-x}-1}.\tag{12}$$

Note that this function should be extended continuously in the singular point $r=x/3$.

IV. Refining Donath's Technique

Donath's technique is primarily based on his hierarchical placement scheme. Every hierarchical level is treated separately without any knowledge about the length of interconnections from other levels of hierarchy. If one calculates the average interconnection length at hierarchical level $k-1$, one doesn't know what happens to the interconnections crossing the border of that level (interconnections belonging to a level k or higher). Donath simply assumes that the starting points of these interconnections are uniformly distributed over the gates of the cubic-shaped grid. It is nevertheless clear that an optimal placement strategy will place interconnected logic gates as close as possible, regardless of the hierarchical level the interconnection belongs to. So it is reasonable to assume that the interconnections crossing the border of level $k-1$ will be as short as possible. This means that an optimal placement procedure will place gates that are interconnected to another cube (at level k) preferably near the border of the two cubes. Donath's technique does not take full account of this information in estimating the average interconnection length.

In [5], simple theoretical considerations are used to indicate that the normalised distribution $f_D(l)$ of interconnection lengths for a good two-dimensional placement in a square Manhattan grid should be of the form[2]

$$\begin{aligned}f_D(l)&\approx g l^{2r-3}\quad(1\le l\le L)\\ &\approx 0\qquad(l>L).\end{aligned}\tag{13}$$

In this distribution g is a normalisation constant, L is a constant related to the size of the square grid, and r is the Rent exponent.

In [8] we showed that the trend of $f_D(l)$ (equation 13) mainly depends on the number of interconnections $\bar{n}_k$ at

[2]The index D is used to denote the fact that the distribution has been derived by Donath

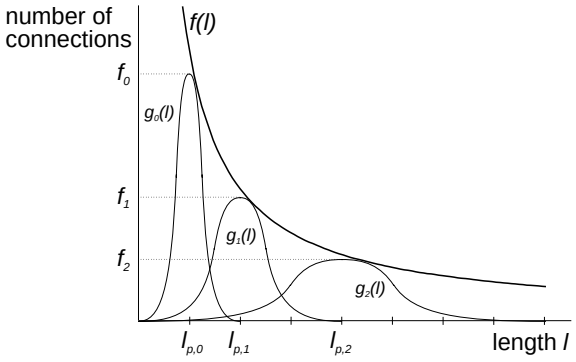

Fig. 4. Interconnection length distributions at every hierarchical level constitute to the global distribution $f(l)$. Only the first three local distributions are shown.

hierarchical level k and not on the way these interconnections are distributed locally. The same reasoning applies to a three-dimensional placement.

Consider the local distribution $g_k(l)$ of interconnection lengths at hierarchical level k ($g_k(l) = 0$ for $l > 2^{k+2}$). Due to Rent's rule, which is a result of the self-similarity within designs, the interconnection complexity is similar at all hierarchical levels. This similar behaviour at each hierarchical level k leads us to the assumption that the local distributions $g_k(l)$ have similar shapes. Now, consider the peak value f_k of $g_k(l)$ and the length $l_{p,k}$ for which $f_k = g_k(l)$. Since the local distributions $g_k(l)$ are similarly shaped, $l_{p,k}$ scales with 2^k (fig. 4). We assume that the peak of the distributions $g_k(l)$ is sufficiently sharp so that every $g_k(l)$ constitutes to the global distribution $f(l)$ ($= \sum_{k=0}^{K-1} g_k(l)$) mainly around $l_{p,k}$. The global distribution $f(l)$ will then be approximated properly by a smooth continuous line through all f_k. Since the total number of interconnections at hierarchical level k equals $\bar{n}_k$, we have

$$\int g_k(l)dl = \bar{n}_k \qquad (0 \leq k \leq K-1).$$

The integral is proportional to the product of the peak value f_k and the length of the domain of $g_k(l)$ by a factor C. Because we assume that the $g_k(l)$ have similar shapes, this factor is the same for all k. We can thus write

$$f_k\, 2^{k+2} = C\, \bar{n}_k$$

for all $k, 0 \leq k \leq K-1$. Using the value of $\bar{n}_k$ (equation 4) this yields

$$f_k = f_0\, 2^{k(3r-4)},$$

and, since the values of $l_{p,k}$ scale with 2^k,

$$f(l) = g l^{3r-4}, \tag{14}$$

where g is the normalisation constant.

These simple considerations show that the trend of the distribution $f(l)$ (equation 14) only depends on the scaling behaviour of the local distributions $g_k(l)$ and not on the distributions themselves. This also explains why Donath's average interconnection length estimates follow this trend very well even with a uniform distribution of the starting points and endpoints of interconnections in A-, N-, and R-combinations.

The distribution $f(l)$ is called the *global distribution* and is defined as the interconnection length distribution of the whole design. The global distribution contains information about all interconnections together. At each hierarchical level, we can also define a *local distribution*. Such a distribution only contains information about interconnections at a specified hierarchical level.

By enumerating all possible interconnections in every A-, N-, and R-combination on each hierarchical level, one obtains an interconnection length distribution depending only on the physical architecture the design will be placed in. We will call this the *structural distribution*. The distributions $P_{k,a}$, $P_{k,n}$, and $P_{k,r}$ that we calculated in section III for Donath's placement technique, are structural distributions for the cubic-shaped Manhattan grid. We can also assign to each length the probability that the interconnection corresponding with that length would be laid out in a real placement procedure and with a real design with the specified Rent exponent. This leads to a *probability distribution*.

Each local distribution can be factorized as the product of the structural distribution and the probability distribution. Donath assumes a constant probability distribution, but in [8] we showed that this could not be a good model for an optimal placement. In a good model, the probability should decrease with increasing length. Since it is reasonable to assume that this probability follows the trend of the global distribution $f(l)$ ([8]), we multiply the structural distributions with the global distribution $f(l)$ to obtain the local distributions.

V. Mathematical Derivation of the Average Interconnection Length

The local distribution for each combination on hierarchical level k is found by multiplying $P_{k,C}(l)$ ($C \in \{a, n, r\}$) with the global distribution $f(l)$, given by equation 14. Therefore equation 8 becomes

$$\bar{l}_{k,C} = \frac{\sum_{l=0}^{6\lambda} l P_{k,C}(l) f(l)}{\sum_{l=0}^{6\lambda} P_{k,C}(l) f(l)} = \frac{\sum_{l=0}^{6\lambda} P_{k,C}(l) l^{3r-3}}{\sum_{l=0}^{6\lambda} P_{k,C}(l) l^{3r-4}}. \tag{15}$$

The total average interconnection length at hierarchical level k is then given by equation 9:

$$\bar{l}_k = \frac{12\bar{l}_{k,a} + 12\bar{l}_{k,n} + 4\bar{l}_{k,r}}{28}. \tag{16}$$

Note that in equation 16 we still assume that the probability of an interconnection is the same for A-, for N-, and

for R-combinations. In reality, an optimal placement will favour A-combinations slightly over N-combinations, since the average interconnection length is larger in the latter combination. The same applies to N- and R-combinations. We could consider using the global distribution $f(l)$ as a probability measure for interconnections at the whole hierarchical level instead of at each combination separately. However, this requires that our partitioning scheme would have to be different. The adjustment of the partitioning scheme will be a topic of further research.

The sums in equation 15 can not be computed analytically without knowing the value for $\lambda = 2^k$. Yet, if we want to compare our results with those of Donath, both numerically and theoretically, an analytical form of the average interconnection length is needed. A way around this problem is to approximate the discrete distributions by continuous ones. One can easily verify that the continuous form of the equations 5 through 7 is given by

$$P^c_{k,a}(l) = \begin{cases} \frac{1}{30\lambda^6} P^{c,i}_{k,a} & (i\lambda \le l < (i+1)\lambda,\ i = 0,\ldots,3) \\ 0 & \text{otherwise,} \end{cases}$$

with

$$\begin{aligned}
P^{c,0}_{k,a} &= 20\,l^3\,\lambda^2 - 10\,l^4\,\lambda + l^5, \\
P^{c,1}_{k,a} &= 74\,\lambda^5 - 260\,l\,\lambda^4 + 340\,l^2\,\lambda^3 - 180\,l^3\,\lambda^2 \\
&\quad +40\,l^4\,\lambda - 3\,l^5, \\
P^{c,2}_{k,a} &= -\left[758\,\lambda^5 - 1420\,l\,\lambda^4 + 980\,l^2\,\lambda^3 - 320\,l^3\,\lambda^2\right. \\
&\quad \left. +\ 50\,l^4\,\lambda - 3\,l^5\right], \\
P^{c,3}_{k,a} &= (4\,\lambda - l)^5
\end{aligned}$$

for an A-combination,

$$P_{k,n}(l) = \begin{cases} \frac{1}{60\lambda^6} P^{c,i}_{k,n} & (i\lambda \le l < (i+1)\lambda,\ i = 0,\ldots,4) \\ 0 & \text{otherwise,} \end{cases}$$

with

$$\begin{aligned}
P^{c,0}_{k,n} &= 5\,l^4\,\lambda - l^5, \\
P^{c,1}_{k,n} &= -\left[25\,\lambda^5 - 105\,l\,\lambda^4 + 170\,l^2\,\lambda^3 - 130\,l^3\,\lambda^2\right. \\
&\quad \left. +\ 40\,l^4\,\lambda - 4\,l^5\right], \\
P^{c,2}_{k,n} &= 775\,\lambda^5 - 1655\,l\,\lambda^4 + 1350\,l^2\,\lambda^3 - 510\,l^3\,\lambda^2 \\
&\quad +90\,l^4\,\lambda - 6\,l^5, \\
P^{c,3}_{k,n} &= -\left[3275\,\lambda^5 - 4555\,l\,\lambda^4 + 2430\,l^2\,\lambda^3 - 630\,l^3\,\lambda^2\right. \\
&\quad \left. +\ 80\,l^4\,\lambda - 4\,l^5\right], \\
P^{c,4}_{k,n} &= (5\,\lambda - l)^5
\end{aligned}$$

for a N-combination, and

$$P_{k,r}(l) = \begin{cases} \frac{1}{120\lambda^6} P^{i}_{k,r} & (i\lambda \le l < (i+1)\lambda,\ i = 0,\ldots,5) \\ 0 & \text{otherwise,} \end{cases}$$

with

$$\begin{aligned}
P^{c,0}_{k,r} &= l^5, \\
P^{c,1}_{k,r} &= 6\,\lambda^5 - 30\,l\,\lambda^4 + 60\,l^2\,\lambda^3 - 60\,l^3\,\lambda^2 + 30\,l^4\,\lambda \\
&\quad -5\,l^5, \\
P^{c,2}_{k,r} &= -2\,\left[237\,\lambda^5 - 585\,l\,\lambda^4 + 570\,l^2\,\lambda^3 - 270\,l^3\,\lambda^2\right. \\
&\quad \left. +\ 60\,l^4\,\lambda - 5\,l^5\right], \\
P^{c,3}_{k,r} &= 2\,\left[2193\,\lambda^5 - 3465\,l\,\lambda^4 + 2130\,l^2\,\lambda^3 - 630\,l^3\,\lambda^2\right. \\
&\quad \left. +\ 90\,l^4\,\lambda - 5\,l^5\right], \\
P^{c,4}_{k,r} &= -\left[10974\,\lambda^5 - 12270\,l\,\lambda^4 + 5340\,l^2\,\lambda^3\right. \\
&\quad \left. -\ 1140\,l^3\,\lambda^2 + 120\,l^4\,\lambda - 5\,l^5\right], \\
P^{c,5}_{k,r} &= (6\,\lambda - l)^5
\end{aligned}$$

for a R-combination.

A substitution of the sums in equation 15 by integrals, yields

$$\bar{l}_{k,C} = \frac{\displaystyle\int_0^{6\lambda} P^c_{k,C}(l) l^{3r-3} dl}{\displaystyle\int_0^{6\lambda} P^c_{k,C}(l) l^{3r-4} dl}.$$

This results in $\bar{l}_{k,C} = \lambda R_C(r)$, $C \in \{a, n, r\}$ with

$$R_C(r) = \frac{(r-1)}{(r+1)} \frac{T_C(r)}{N_C(r)},$$

and

$$\begin{aligned}
T_a(r) &= 4^{3\,r+3} - 2\,(r+3)\,3^{3\,r+3} \\
&\quad + \left(9\,r^2 + 51\,r + 66\right) 2^{3\,r+1} - 2\left(9\,r^2 + 24\,r + 17\right) \\
N_a(r) &= 4^{3\,r+2} - 2\,(3\,r+8)\,3^{3\,r+1} \\
&\quad + \left(9\,r^2 + 45\,r + 50\right) 2^{3\,r} - 2\left(9\,r^2 + 18\,r + 10\right), \\
T_n(r) &= 5^{3\,r+3} - (3\,r+23)\,4^{3\,r+2} + 2\,(2\,r+7)\,3^{3\,r+3} \\
&\quad - (9\,r+19)\,2^{3\,r+3} + 12\,r + 17, \\
N_n(r) &= 5^{3\,r+2} - (3\,r+22)\,4^{3\,r+1} + 2\,(6\,r+19)\,3^{3\,r+1} \\
&\quad - (9\,r+16)\,2^{3\,r+2} + 12\,r + 13, \\
T_r(r) &= 6^{3\,r+2} - 5^{3\,r+3} + 10\,4^{3\,r+2} - 10\,3^{3\,r+2} \\
&\quad +5\,2^{3\,r+2} - 1, \\
N_r(r) &= 6^{3\,r+1} - 5^{3\,r+2} + 10\,4^{3\,r+1} - 10\,3^{3\,r+1} \\
&\quad +5\,2^{3\,r+1} - 1.
\end{aligned}$$

As a result,

$$\bar{l}_k = \lambda R(r),$$

with

$$R(r) = \frac{12R_a(r) + 12R_n(r) + 4R_r(r)}{28}. \qquad (17)$$

The sum over all hierarchical levels (equation 2) then yields

$$\mathcal{L} = R(r) \frac{H(K, r, 2)}{H(K, r, 3)}, \qquad (18)$$

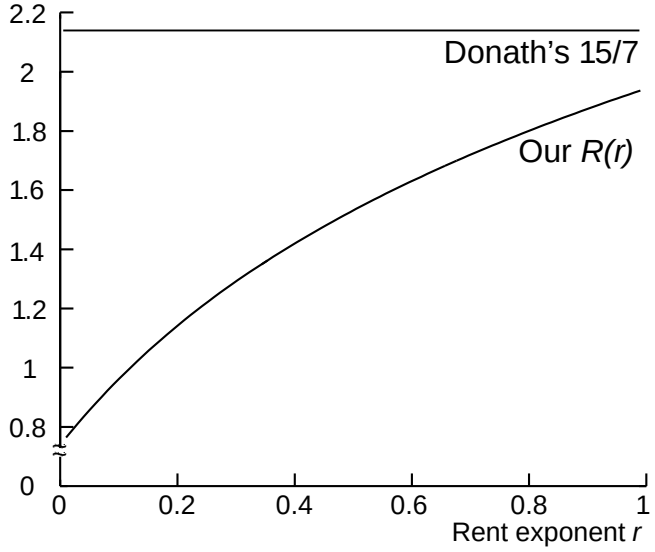

Fig. 5. $R(r)$ (equation 17) versus Donath's factor for $0 < r < 1$.

with $H(K, r, x)$ given by equation 12

$$H(K, r, x) = \frac{2^{K(3r-x)} - 1}{2^{3r-x} - 1}.$$

VI. Results

The average interconnection length computed in the previous section (equation 18) has the same scaling behaviour[3] as the average interconnection length computed by Donath (equation 11). This is a necessary consequence of the fact that we used the same hierarchical placement technique with the number of interconnections at every hierarchical level estimated by Rent's rule. However, the multiplication factor $R(r)$ in equation 18 is smaller than the factor 15/7 found by Donath, for all possible values of r (fig. 5). Unlike Donath's factor, our factor $R(r)$ increases with increasing r, corresponding to the fact that more complex designs (with a higher Rent exponent r) tend to have longer interconnections. The more complex the design, the more difficult it is for a placement program to place interconnected logic gates as close as possible. Note that this effect is supplementary to the scaling behaviour in the factor

$$H(K, r) = \frac{H(K, r, 2)}{H(K, r, 3)} \qquad (19)$$

(which also expresses the fact that complex designs need more interconnection length). For complex designs the scaling factor $H(K, r)$ will be high due to the fact that there will be relatively more interconnections at higher levels of the hierarchy (and thus at higher length) than is the case with designs of low complexity. This can be seen in the equation for $\bar{n}_k$ (equation 4). The factor $R(r)$ on the other hand increases with r because of the fact that, for complex designs, it is difficult to place all logic gates connected to the outside of a cube on a hierarchical level, close to the border of that cube. For designs with a low Rent exponent, nearly all such gates can be placed near the border. Since Donath assumes that these gates are uniformly distributed over the cube, he describes the case where the placement program can not decide where to place the gates. Therefore, Donath's average interconnection length values are overestimated, especially for designs with low complexity. In all cases ($0 < r < 1$) our estimate of the average interconnection length is more accurate than the one found by Donath.

[3]This means that they have the same behaviour if the number of hierarchical levels increases to infinity ($K \to \infty$, i.e. the number of gates $G = 8^K$ increases to infinity).

TABLE I

Average interconnection length of some benchmark designs: Donath's analysis ($\mathcal{L}_D$) versus our analysis ($\mathcal{L}$) in two and three dimensions. The number of gates is shown in the column N_g, N_c is the number of connections, and r is the Rent exponent. The circuits numbered 1 through 5 are those used by Donath in [1], the others are the same as those used in [8].

No.	N_g	N_c	r	$\mathcal{L}_{exp}^{2D}$	$\mathcal{L}_D^{2D}$	$\mathcal{L}^{2D}$	$\mathcal{L}_D^{3D}$	$\mathcal{L}^{3D}$	$\frac{\mathcal{L}_D^{3D}}{\mathcal{L}_D^{2D}}$	$\frac{\mathcal{L}^{3D}}{\mathcal{L}^{2D}}$
1	528	1007	0.59	2.15	4.02	2.88	3.11	2.61	0.77	0.91
2	576	1111	0.75	2.85	5.26	4.13	3.73	3.31	0.71	0.80
3	671	1670	0.57	2.63	4.07	2.89	3.12	2.59	0.77	0.89
4	1239	2687	0.47	2.14	3.76	2.45	2.94	2.30	0.78	0.94
5	2148	7302	0.75	3.50	7.37	5.74	4.53	3.96	0.61	0.69
6	1024	3047	0.40	1.96	3.28	2.02	2.71	2.03	0.83	1.01
7	1024	2979	0.50	2.21	3.79	2.60	3.00	2.39	0.79	0.92
8	1024	2893	0.60	2.58	4.61	3.32	3.36	2.81	0.73	0.85

In order to verify our theoretical average interconnection length estimation, we compare it with Donath's theoretical estimation for the benchmark designs used in [8] (five of those benchmark designs are also used by Donath in [1]). The results are shown in table I. In this table $\mathcal{L}_{exp}^{2D}$ denotes the experimentally measured value of the average interconnection length for a two-dimensional placement. The theoretical values found by Donath are denoted as $\mathcal{L}_D^{xD}$ and those presented in this paper by $\mathcal{L}^{xD}$. They are shown both for the two- and for the three-dimensional case ($x = 2$ or $x = 3$). One can see that our estimates are much lower than Donath's in both cases.

An experimental verification still requires a lot of research to find a good placement and routing algorithm for three-dimensional placement and routing of the benchmark designs. The comparison with experiments in two dimensions (column $\mathcal{L}_{exp}^{2D}$ in table I, adopted from [8]), does show that our extension to Donath's technique gives estimates that are a lot more accurate than Donath's estimates. The same result is to be expected for three-dimensional systems.

In the last two columns of table I the improvement is shown for a three-dimensional placement over a two-dimensional one for both Donath's technique and our extension. One can see that the average interconnection length decreases when highly complex designs are placed in a three-dimensional system. This indicates the usefulness of these three-dimensional systems for the placement of designs with a high Rent exponent. More detailed research on this topic, based on Donath's technique, can be found in [14, 15, 16]. For designs with a low Rent exponent, there is not so much difference between two- and three-dimensional placement. In fact, since designs with a Rent-exponent less than 0.5 can be placed in a two-dimensional system without too much difficulties[4], the placement in three dimensions is useless for such designs. It will only result in a two-dimensional placement, folded up into three dimensions. Comparing the last two columns of table I, we see that our technique indeed gives more or less the same results for designs with $r \leq 0.5$ while Donath's technique still shows an improvement for three-dimensional placement over two-dimensional placement. This indicates that Donath's technique overestimates the average interconnection length while our results clearly confirm the fact that an improvement can only be expected for highly complex designs.

VII. Conclusion

In this paper we have presented the extension of Donath's technique to three-dimensional systems. We have also modified Donath's technique, introducing the behaviour of the global distribution into the local distributions at each hierarchical level. This way, we have obtained a more accurate average interconnection length estimation than Donath's estimation. Our results not only cover the scaling behaviour due to Rent's rule, but they also include the knowledge about an optimal placement process in each hierarchical level.

More experimental results will be obtained once our research on placement and routing algorithms in three-dimensional systems is finished. It can be expected that those results will show that our method surpasses Donath's in accuracy, as did experiments in two dimensions.

Further research will focus on including the external interconnections of the whole design in the placement process and on the adjustment of the partitioning scheme in order to also include a different probability distribution for A-, N-, and R-combinations. It can be expected that the resulting estimates will be extremely accurate.

References

[1] W. E. Donath. Placement and average interconnection lengths of computer logic. *IEEE Transactions on Circuits & Systems*, CAS-26: pages 272–277, 1979.

[2] S. Sastry and A. C. Parker. Stochastic models for wireability analysis of gate arrays. *IEEE Transactions on CAD*, 5: pages 52–65, 1986.

[3] F. J. Kurdahi and A. C. Parker. Techniques for area estimation of VLSI layouts. *IEEE Transactions on CAD*, 8: pages 81–92, 1989.

[4] I. E. Sutherland and D. Oestreicher. How big should a printed circuit board be? *IEEE Transactions on Computers*, C-22: pages 537–542, 1972.

[5] W. E. Donath. Wire length distribution for placements of computer logic. *IBM Journal of Research and Development*, 25: pages 152–155, 1981.

[6] M. Feuer. Connectivity of random logic. *IEEE Transactions on Computers*, C-31: pages 29–33, 1982.

[7] J. E. Cotter and P. Christie. The analytical form of the length distribution function of computer interconnections. *IEEE Transactions on Circuits & Systems*, 38: pages 317–320, 1991.

[8] D. Stroobandt, H. Van Marck, and J. Van Campenhout. An accurate interconnection length estimation for computer logic. In *Proceedings of the Sixth Great Lakes Symposium on VLSI*, pages 50–55. IEEE Computer Society Press, March 1996.

[9] F. T. Leighton and A. L. Rosenberg. Three-dimensional circuit layouts. *SIAM Journal on Computing*, 15-3: pages 793–813, 1986.

[10] C. Val and T. Lemoine. 3-D interconnection for ultra-dense multichip modules. *Components*, 13: pages 814–821, 1990.

[11] R. C. Eden. Capabilities of normal metal electrical interconnections for 3-D MCM electronic packaging. In R. T. Chen and J. A. Neff, editors, *SPIE Proceedings: Optoelectonic Interconnects II*, volume 2153, pages 132–145. SPIE, 1994.

[12] J. Depreitere, H. Neefs, H. Van Marck, J. Van Campenhout, R. Baets, B. Dhoedt, H. Thienpont, and I. Veretennicoff. An optoelectronic 3-D field programmable gate array. In Reiner W. Hartenstein and Michal Z. Servít, editors, *Field-Programmable Logic: Architectures, Synthesis and Applications*, volume 849 of *Lecture Notes in Computer Science*, pages 352–360. Springer-Verlag, September 1994.

[13] J. Depreitere, H. Van Marck, and J. Van Campenhout. A quantitative analysis of the benefits of area-I/O in FPGAs. Unpublished, 1996.

[14] H. Van Marck and J. Van Campenhout. Modeling and evaluating optoelectronic architectures. In R. T. Chen and J. A. Neff, editors, *Optoelectronics II, volume 2153 of SPIE Proceedings Series*, pages 307–314, SPIE, 1994.

[15] H. Van Marck and J. Van Campenhout. Modeling signal delay distribution in optoelectronic architectures. In R. T. Chen and H. S. Hinton, editors, *Optoelectronic Interconnects III*, volume 2400 of *SPIE Proceedings Series*, pages 267–276. SPIE, 1995.

[16] H. Van Marck, D. Stroobandt, and J. Van Campenhout. Interconnection length distributions in 3-dimensional anisotropic systems. In M. H. Hamza, editor, *Proceedings of the Thirteenth IASTED International Conference on APPLIED INFORMATICS*, pages 98–101. IASTED-ACTA Press (Anaheim * Calgary * Zurich), 1995.

[17] B. S. Landman and R. L. Russo. On a pin versus block relationship for partitions of logic graphs. *IEEE Transactions on Computers*, C-20: pages 1469–1479, 1971.

[18] H. Van Marck, D. Stroobandt, and J. Van Campenhout. Towards an extension of Rent's rule for describing local variations in interconnection complexity. In Shuo Bai, Jianping Fan, and Xiaozhong Li, editors, *Proceedings of the Fourth International Conference for Young Computer Scientists*, pages 136–141. Peking University Press, 1995.

[4] For $r < 0.5$ the scaling factor $H(K, r)$ of the average length is smaller than the scaling factor of the size of the square grid.