ePlace-3D: Electrostatics based Placement for 3D-ICs

Jingwei Lu$^1$, Hao Zhuang$^2$, Ilgweon Kang$^2$, Pengwen Chen$^3$ and Chung-Kuan Cheng$^2$
$^1$Cadence Design Systems, Inc., $^2$Department of Applied Mathematics, National Chung Hsing University
$^3$Department of Computer Science and Engineering, University of California, San Diego
francesco.liw@gmail.com, hao.zhuang@cs.ucsd.edu, igkang@ucsd.edu, pengwen@nchu.edu.tw, ckcheng@ucsd.edu

Abstract—We extend the two-dimension (2D) ePlace algorithm [20], [22], [28] to a flat, analytic, mixed-size placement algorithm ePlace-3D, for three-dimension integrated circuits (3D-ICs). Nonlinear optimization is applied over the entire cuboid domain. Specifically, we develop (1) eDensity-3D: an electrostatics based 3D placement density function with globally uniform smoothness (2) a 3D numerical solution with improved spectral formulation (3) a 3D nonlinear pre-conditioner for convergence acceleration (4) interleave 2D-3D placement for quality and efficiency enhancement. We integrate all the features into our 3D-IC placement prototype ePlace-3D. Compared to the leading placers mPL6-3D and NTUplace3-3D, our algorithm produces 6% enhancement. We integrate all the features into our 3D-IC placement prototype ePlace-3D. Compared to the leading placers mPL6-3D and NTUplace3-3D, our algorithm produces 6% enhancement. We validate ePlace-3D on the large-scale modern mixed-size (MMS) 3D circuits, which shows high performance and scalability.

I. INTRODUCTION

Placement remains dominant on the overall quality of physical design automation [25]–[27], [31]. Various back-end optimization on timing [24], [45], power [23], [44], routability [8], [38], variability [3], [42] etc. are highly impacted by placement performance. In the recent time, the progress in 3D-IC development [30] challenges the traditional 2D placers [1], [2], [5], [16], [19], [22], [41] to produce 3D circuit layout with minimum total wirelength yet limited vertical interconnects (through-silicon vias (TSVs), monolithic inter-tier vias (MIVs), etc.). Innovations of 3D-IC algorithms for large-scale mixed-size placement become quite desirable, in order to accommodate the advent of the 3D-IC era.

Previous combinatorial 3D-IC placers form two categories. Folding based methods [4] folds the 2D-IC placement layout to produce 3D solution with local refinement. As placement is conducted on the planar domain, such 2D-3D post-transformation is expensive. Partitioning based approaches [7], [17] minimize the usage of vertical resources. Kim et al. [17] partitions the netlist followed by tier assignment, then applies 2D quadratic placement [40] simultaneously over all the tiers. However, as hypergraph partitioning is NP-complete problem, placement quality loss due to sub-optimal partitioning is inevitable.

Analytic placers achieve the best 3D-IC placement performance versus combinatorial algorithms. These placers distribute cells in the 3D domain with the three coordinates of every object simultaneously perturbed. Goplen et al. [6] models the 3D-IC placement by a quadratic framework [5]. Hsu et al. [9] extends the 2D-IC placement prototype [10] and uses Bell-shape function [34] to smooth the vertical dimension. Luo et al. [29] utilizes the 2D algorithm in [1] and relaxes the discrete tiers via Huber function [11]. However, as Bell-shape function and Huber function are only locally smooth, density equalization at the third dimension would be slowed down. Moreover, [9] and [29] are based on hierarchical cell clustering and grid coarsening, which improve the efficiency but degrade the quality [22]. Another issue among all the prior 3D-IC placement works [6], [9], [29] is that the benchmarks [12], [14] are too small. Specifically, there are only 12K to 210K cells in IBM-PLACE circuits [12] used by [6], [9], [29], 11K to 169K cells in IWLS05 [14] used by [9], [17], all of which are too small to represent modern design complexity. Large-scale bookshelf 3D-IC placement benchmarks become desirable.

The algorithmic extension from 2D (planar) to 3D (cuboid) placement is not straightforward. Consider a similar extension from 1D (linear) to 2D (planar) placement, where a perfect solver for the linear arrangement of cells becomes almost useless for the planar distribution of cells. By introducing one more dimension, significant amount of research innovations would be necessary to address the explosion in both problem complexity and solution sub-optimality. The insertion of the third placement dimension substantially expands the optimization space, complicates the placement problem, moreover, fails many mature features and research arts in the state-of-the-art 2D algorithms [1], [2], [16], [18], [40], [41]. As a result, massive research efforts are desirable to address the huge structural and physical differences.

In this work, we extend the 2D mixed-size placers ePlace [20], [22] and ePlace-MS [28] to the 3D domain. A flat, analytic, nonlinear algorithm is developed for mixed-size large-scale 3D-IC placement. Our algorithm is named ePlace-3D and is focused on wirelength minimization and density equalization. Other relevant 3D-IC placement research problems (thermal, power, etc.) are not covered in this work. To the best of our knowledge, this is the first work in literature achieving analytically global smoothness along all the three dimensions. The smoothed cuboid density (i.e., potential) of every single bin is globally correlated with the original density of all the other bins in the placement domain. In contrast, previous analytic works [9], [29] only ensure (partially) local smoothness in their density functions [11], [34], while their less continuous cell movement would slow down placement convergence and cause more penalty on wirelength. We conduct analytic global placement and stochastic legalization in the entire 3D cuboid domain, which maximizes the search space thus further boost the solution quality. ePlace-3D well demonstrates the applicability of the electrostatic density model eDensity [20], [21] in various physical dimensions. Our specific contributions are listed as follows.

- We propose eDensity-3D: an electrostatics based 3D density function ensuring global smoothness within the cuboid placement domain.
- We propose a 3D numerical solution based on fast Fourier transform (FFT) and improved spectral formulation. Computational complexity is only $O(n \log n)$, where $n$ is the number of placement objects.
- We propose a nonlinear 3D preconditioner to equalize all the moving objects in the optimization perspective.
- We interleave coarse-grained 3D placement with fine-grained 2D placement for a good trade-off between quality and efficiency.
- We implement ePlace-3D, a mixed-size 3D-IC placement prototype, with all the features integrated. ePlace-3D outperforms the leading placer mPL6-3D [29] and NTUplace3-3D [9] with 6.44% and 37.15% shorter wirelength, 9.11% and 10.27%
fewer 3D vertical interconnects, while runs 2.55× faster and 0.30× slower on average of all the ten IBM-PLACE benchmarks [12].

The remainder is organized as follows. Section II introduces the background knowledge. Section III discusses our 3D placement density function eDensity-3D, numerical solution, and nonlinear preconditioner. Section IV provides an overview of ePlace-3D algorithm with strategies on 2D-3D placement interleaving. Experiments and results are shown in Section V. We conclude in Section VI.

II. BACKGROUND

Given a set $V$ of $n$ objects, net set $N$ and 3D cuboid core region $R = \{0, d_x\} \times \{0, d_y\} \times \{0, d_z\}$, global placement is formulated as constrained optimization. The constraint desires all the objects to be accommodated with zero overlap. Let $v$ denote the placement solution, which consists of the physical coordinates of all the objects. The region $R$ is uniformly decomposed into $m_x \times m_y \times m_z$ 3D bins denoted as set $B$. For every bin $b \in B$, the density $\rho_b(v)$ should not exceed the target density $\rho_t$. The objective is to minimize the total half-perimeter wirelength (HPWL) of all the nets. Let $HPWL_v = \max_{b \in B} |x_i - x_j|$ denote the horizontal wirelength of net $e$ (similar for $HPWL_e$), the total 2D HPWL is $\sum_{\forall e \in V} (\beta_e \cdot HPWL_v(e) + \beta_v \cdot HPWL_e(v))$. We use $\beta_e$, $\beta_v$ and $\beta_d$ as dimensional weighting factors. 3D-IC placement needs vertical interconnects, such as through-silicon via (TSV) and monolithic inter-tier via (MIV), to penetrate silicon tiers. Diverse types of connects have different physical and electrical properties. However, ePlace-3D is compatible with any types of connects, which can be reflected on the weight of $\beta_e$. In the remainder of this manuscript, we name all types of 3D vertical interconnects uniformly as VI for simplicity. The number of vertical interconnect units (VI) is computed as how many times silicon tiers have been penetrated, e.g., one vertical connect between tier one and tier three is counted as two VIs. The nonlinear placement optimization is formulated as

$$\min_v \ (HPWL_v + \beta_v \cdot \#VI) \ \text{s.t.} \ \rho_b(v) \leq \rho_t, \ \forall b \in B. \quad (1)$$

Analytic methods conduct placement using gradient-directed optimization. As $HPWL_v$ is not differentiable, we use wirelength smoothing by the weighted-average (WA) model [9].

$$W_{\epsilon} (v) = \frac{\sum_{i \in e} x_i \exp (x_i/\gamma_x)}{\sum_{i \in e} \exp (x_i/\gamma_x)} - \frac{\sum_{i \in e} x_i \exp (-x_i/\gamma_x)}{\sum_{i \in e} \exp (-x_i/\gamma_x)} \quad (2)$$

Here $W_{\epsilon} (v) = \beta_e W_{\epsilon e} (v) + \beta_v W_{\epsilon v} (v) + \beta_d W_{\epsilon d} (v)$ and $W (v) = \sum_{\forall e} W_{\epsilon} (v)$. $\gamma_x$ and $\gamma_y$ control the modeling accuracy. Density function relaxes all the $|B|$ constraints in Eq. (1). Most 2D and 3D quadratic placers [6, 17, 19] follow the linear density force formulation by [5]. Nonlinear placers [1, 9, 10, 29] have their dedicated density functions. NTUplace-3D3 [9] leverages bell-shape curve [34] for local smoothness in 3D domain. mPL6-3D [1] uses Helmholtz function to globally smooth the 2D plane and Huber’s function to locally smooth the vertical dimension. The electrostatics based density function [22] converts objects to charges. By the Lorentz law, the electric repulsive force spreads charges away towards the electrostatic equilibrium state, which produces a globally even density distribution. Let $U (v)$ denote the density cost function, the constraints in Eq. (1) can be relaxed by the penalty factor $\lambda$, while the unconstrained optimization is shown as below.

$$\min_v f (v) = W (v) + \lambda U (v). \quad (3)$$

In this work, we set vertical connects as zero-volumed thus do not consider them in eDensity-3D. Therefore, the optimization of electrostatics will not be affected and can be still achieved based on the movement of netlist objects. Density overflow is used to terminate global placement and denoted as $\tau$, which is computed as

$$\tau = \frac{\sum_{b \in B} \max (V_b^v - \rho_t, V_b^v W_{\epsilon v}, 0)}{V_m} \quad (4)$$

Here $V_m$ is the total volume of all the movable objects, $V_b^v$ is the total volume of objects in the bin $b$, and $V_b^v W_{\epsilon v}$ is the total whitespace in bin $b$. The volume of each cell is computed as its planar area multiplied by the depth of each tier.

III. eDensity-3D: Density Function for 3D Placement

In this section, we introduce our novel 3D density function eDensity-3D, a fast numerical solution of spectral methods through 3D frequency decomposition, and an approximated 3D nonlinear preconditioner. The key insight to our 3D approach is, we treat the third dimension equally as the other two dimensions, such that vertical cell movement will be as smooth as the planar movement in 2D placement. The behavior of eDensity-3D is visualized in Figure 1. For more details of our placement flow, please refer to Section IV.

A. 3D Density Function

Extending the planar function eDensity in [22], eDensity-3D models the entire placement instance as a 3D electrostatic field. Every placement object (standard cells, macros and fillers) is converted to a positively charged cuboid. The electric repulsive force spreads all the objects away from the high-density region. The 3D density cost $U$ is modeled as the total potential energy of the system and defined as below

$$U (v) = \sum_{i \in V} U_i (v) = \sum_{i \in V} q_i \Phi_i (v). \quad (5)$$

$q_i$ denotes the electric quantity of the charge $i$ and is set as the physical volume of placement object $i$. $\Phi$ is the electric potential at charge $i$. Charges with high potential will reduce the placement overlap by moving towards the direction of least energy descent. Unlike the spatial density distribution $\rho(x, y, z)$ (Figure 1) which is coarse and non-differentiable, the electric potential distribution $\Phi(x, y, z)$ is globally smooth. We use the potential gradient (thus electric field), $\nabla \Phi (x, y, z) = \mathbf{E} (x, y, z)$, to direct cell movement for density equalization. Given a placement layout $v$, we generate the respective density map $\rho(x, y, z)$, then compute the potential map $\Phi (x, y, z)$ by solving the following 3D Poisson’s equation

$$\begin{align*}
\mathbf{n} \cdot \nabla \Phi (x, y, z) &= -\rho(x, y, z), \\
\hat{n} \cdot \nabla \Phi (x, y, z) &= 0, \quad (x, y, z) \in \partial R, \\
\iint_R \Phi (x, y, z) &= \iint_R \rho (x, y, z).
\end{align*} \quad (6)$$

Here $\hat{n}$ is the outer unit normal of the placement cube $R$. $\partial R$ is the boundary of $R$, which consists of six orthogonal rectangular planes to enclose the placement cube. In the first equation of Eq. (6), the second-order differentiation operator is $\nabla \cdot \nabla \Phi \equiv \partial^2 \Phi / \partial x^2 + \partial^2 \Phi / \partial y^2 + \partial^2 \Phi / \partial z^2$. Neumann condition by the second equation of Eq. (6) requires that when any object $i$ reaches any boundary plane, its density force vector will have the component perpendicular to the plane reduced to

$1$ Practically, vertical connects can never be zero volumed. However, for academic research we are able to simplify the engineering problems to boost scientific innovations. Similarly, state-of-the-art 2D placement academic works [1], [2], [16], [19], [22], [40], [41] target wirelength only and ignore other objectives like timing, power and routability. As vertical connects may be of large volume thus significantly contribute to the placement density, we will put it in our future work.
zero. As a result, object \( i \) could only move on the plane rather than penetrating it. As the third equation of Eq. (6) shows, the integral of density \( \rho(x, y, z) \) and potential \( \Phi(x, y, z) \) within \( R \) are both set to zero to ensure that

- The electric force will eventually drive all the charges towards an even density distribution (rather than pushing them towards infinity), which well matches the placement objective.
- The 3D Poisson’s equation would have a unique solution by satisfying the Neumann condition.

We differentiate the potential \( \Phi_i \) on each charge \( i \) to generate the electric field \( \nabla \Phi_i = E_i = (E_{ix}, E_{iy}, E_{iz}) \). The electric (density) force is \( \nabla U_i = q_i \nabla \Phi_i = q_i E_i \).

**B. 3D Numerical Solution**

Based on the 2D solution in [22], we solve the proposed 3D Poisson’s equation via spectral methods based on frequency decomposition [39]. The integration can be easily realized by switching between sinusoidal and cosine waveforms. To satisfy the Neumann condition, i.e., zero gradients at the boundaries, we use sinusoidal wave to express the electric field \( \mathbf{E}(x, y, z) \). We construct an odd and periodic field distribution by negatively mirroring itself w.r.t. the origin, then periodically extending it towards positive and negative infinities. Electric potential and density distributions are then expressed by cosine waveforms, which are the integration and differentiation of the sinusoidal expression of field. Let \( a_{j,k,l} \) denote the 3D coefficients of the density frequency.

\[
    a_{j,k,l} = \frac{1}{W^3} \sum_{x,y,z} \rho(x, y, z) \cos(w_j x) \cos(w_k y) \cos(w_l z) \tag{7}
\]

eDensity [22] sets \( w_j = \pi x / x \), which equals the discrete index for the \( j \)th frequency component. However, as we are conducting placement in a continuous domain, the multiplication of \( x \) and \( w_j \) induces inconsistency. In this work, we propose improved spectral methods for the 3D placement density function. Specifically, we set \( w_j = \pi d_j / x \), since \( x \) ranges within \( (0, d_j) \). As a result, \( w_j x = \pi j \) well matches the original unit of discrete frequency index, and we have all the frequency indexes defined as

\[
\begin{align*}
    w_j &= \frac{\pi j}{d_j} \\
    w_k &= \frac{\pi k}{d_j} \\
    w_l &= \frac{\pi l}{d_j}
\end{align*}
\tag{8}
\]

As mentioned in Section II, here \( d_x \), \( d_y \) and \( d_z \) represent the dimensions of the cuboid placement core region. Notice that \( d_{\{x,y,z\}} \) can be set as any value since \( w_{\{x,y,z\}} \) would be normalized by \( \frac{\pi}{d_{\{x,y,z\}}^2} \).

\( j \), \( k \) and \( l \) range within \( [0, n-1] \), which is only half of a cosine function period. In contrast, one complete function period centered at the origin is \( [-n, n-1] \). Therefore, we have \( \pi \) rather than \( 2\pi \) in the above frequency index. We set \( d_{0,0,0} = 0 \) to remove the zero-frequency component. The spatial density distribution \( \rho(x, y, z) \) is

\[
    \rho(x, y, z) = \sum_{j,k,l} a_{j,k,l} \cos(w_j x) \cos(w_k y) \cos(w_l z) \tag{9}
\]

To achieve \( \nabla \cdot \nabla \Phi(x, y, z) = -\rho(x, y, z) \), the solution to the potential can be expressed as

\[
    \Phi(x, y, z) = \sum_{j,k,l} a_{j,k,l} \cos(w_j x) \cos(w_k y) \cos(w_l z) \tag{10}
\]

By differentiating Eq. (10), we have the electric field distribution \( \mathbf{E}(x, y, z) = (E_x, E_y, E_z) \) shown as below

\[
\begin{align*}
    E_x(x, y, z) &= \sum_{j,k,l} a_{j,k,l} \sin(w_j x) \cos(w_k y) \cos(w_l z), \\
    E_y(x, y, z) &= \sum_{j,k,l} a_{j,k,l} \cos(w_j x) \sin(w_k y) \cos(w_l z), \\
    E_z(x, y, z) &= \sum_{j,k,l} a_{j,k,l} \cos(w_j x) \cos(w_k y) \sin(w_l z)
\end{align*}
\tag{11}
\]

Let \( |B| = m_x \times m_y \times m_z \) denote the total number of bins in global placement. Instead of quadratic complexity, above spectral equations can be efficiently solved using FFT algorithms [36] with \( O(|B| \log |B|) \) complexity.

**C. 3D Nonlinear Precondition**

Mixed-size placement should tolerate the diverse physical and topological attributes of standard cells and macro blocks. A nonlinear pre conditioner \( H_f \) for 2D placement is proposed in [22],

\[
    H_{\alpha} = \frac{\partial^2 f}{\partial x_i^2} = \frac{\partial^2 W}{\partial x_i^2} + \lambda \frac{\partial^2 U}{\partial x_i^2} \approx |N_i| + \lambda A_i, \tag{12}
\]
here $N_i$ denote all the nets incident to object $i$ and $A_i$ denote the 2D area of $i$. In 3D we use $V_i$ to denote the volume of $i$ instead. The preconditioned gradient $\nabla f_{\bar{v}} f = H^{-1} \nabla f$ helps accelerate placement optimization. In this work, our studies show above approximation. As a result, $|N_i|$ always dominates $\lambda V_i$ making fillers and macros with zero or small $|N_i|$ spread faster than standard cells (Eq. (13)) thus deviate from the preconditioning purpose.

$$\|H_{\bar{v}}^{-1} \nabla f\| = \left\| \frac{\partial f}{\partial x_i} \right\|_{N_i} \approx \left\| \frac{\partial f}{\partial x_i} \right\|_{|N_i| + \lambda V_i} = \|H_{\bar{v}}^{-1} \nabla f\|. \tag{13}$$

Instead, we propose a new preconditioner approximated as below

$$H_{\bar{v}} = \frac{\partial^2 f}{\partial x_i^2} \approx \lambda \frac{\partial^2 U}{\partial x_i^2} \approx \lambda V_i. \tag{14}$$

The noise factors introduced by $|N_i|$ is well resolved. Despite substantial physical and topological differences, all the objects are being equalized in the optimizer’s perspective and simultaneously spread over the entire domain. Experiments show that our 3D preconditioner reduces the global placement iterations by 15% and improves the wirelength by 30% over all the 16 MMS benchmarks.

Theoretically, preconditioning only improves convergence rate. However, as the placement problem is highly nonlinear, non-convex and bad-conditioned, preconditioned gradient reduces the condition number of the nonlinear Hessian matrix thus corrects the search direction. As a result, new search space will be explored while solution quality is expected to be significantly improved.

**D. Complexity**

**Complexity** significantly impacts the placement runtime. In each iteration, we traverse all the bins to reset their density in $O(|B|)$ time, then traverse all the placement objects in $O(n)$ time to update the superimposed density map. By Eq. (7), (10) and (11), five times of 3D FFT computation are invoked to generate the density coefficients $a_{jk}$ in the frequency domain as well as the potential $\Phi(x, y, z)$ and field $E(x, y, z)$ in the spatial domain, which costs $O(5n \log n)$ time. By our grid sizing strategy in Eq. (15), $|B|/n$ depends on the design utilization and is thus limited. As a result, the overall complexity is $O(|B| + n + 5n \log n) \approx O(n \log n)$.

In ePlace-3D, each dimension of the placement region $R$ is normalized to one single unit. The placement domain is geometrically transformed from $R = [0, d_x] \times [0, d_y] \times [0, d_z]$ to $R' = [0, 1] \times [0, 1] \times [0, 1]$. We set the density resolutions $m_s = m_y = m_z = m_{3D}$ to make the placement domain $R'$ uniformly decomposed into $|B| = m_{3D}^3$ cubic bins. Let $V_R$ denote the total volume of $R$ and $V_{c_{avg}}$ denote the average area of all standard cells. The grid sizing is determined as

$$|B| = \frac{V_R}{k \times V_{c_{avg}} \times \rho_t}. \tag{15}$$

Here every $k$ standard cells will be accommodated by exactly one bin. Placement quality (efficiency) can thus be easily improved by setting small (large) $k$, vice versa. In this work, we constantly set $k = 1$.

IV. ePLACE-3D: PLACEMENT STRUCTURE OVERVIEW

ePlace-3D is built upon the infrastructure of ePlace-MS framework [28]. Figure 2 shows the flowchart of our algorithm. Given a placement instance (benchmark), our algorithm minimizes the quadratic wirelength over the 3D domain to produce the initial solution $v_{1P}$. To approach the minimum-wirelength violation-free (optimum) solution in the end, we make $v_{1P}$ be the minimum-wirelength violation-tolerant solution as the initial condition of global placement. Our 3D-IC global placement is visualized in Figure 3. Unconnected fillers [1], [22] are inserted to populate up extra whitespace. All the fillers are equally sized by the average dimensions of all the standard cells, i.e., the vertical dimension of every single filler equals the thickness of one tier. As a result, the optimum solution of 3D global placement will have all the cells, macros and fillers orient towards discrete tiers. Otherwise, some cuboid placement sites will be partially unoccupied (wasted), degrading the solution quality. Figure 3(f) illustrates the beauty of our approach, i.e., the analytic 3D placer is visually approaching density evenness from the vertical dimension. Nonlinear optimization is performed on all the standard cells, fillers and macros simultaneously. We use Nesterov’s method [35] as the nonlinear solver and predicts the steplength by [22].

A multi-tier 2D-IC mixed-size global placement follows by assigning all the macros and standard cells to the closest tiers and separately filling the remaining whitespace on each tier with 2D fillers. As Figure 3(f) shows, our global placement approach is able to remove overlaps and orient cells and macros towards certain discrete tiers, which ensures negligible quality overhead by our tier assignment. Planar placement is conducted simultaneously over all the tiers. As wirelength smoothing is homogeneous over all the tiers (with the same $\gamma$), heterogeneous grid sizing is not feasible since density force is dependent on the resolution (Eq. (11)). As a result, we set all the tiers with the same density resolution $m_{2D}$, which is the maximum of that of all the tiers computed by Eq. (15) with $k = 1$. In practice, $m_{2D}$ is usually much larger than $m_{3D}$, while we have $O(m_{2D}^3) \approx O(m_{3D}^3)$. Figure 4 illustrates the iterative progression, where gradient-guided planar movement resolve the remaining overlaps in finer granularity.

Our 3D-IC macro legalizer generates a legal macro layout with zero macro overlap and small wirelength overhead. The algorithm is stochastic based on simulated annealing. A 3D-IC standard-cell global placement follows to mitigate the quality loss due to suboptimal macro legalization. We assign standard cells to their closest tiers and conduct simultaneous 2D-IC standard-cell placement on all the tiers. The standard-cell layouts of all the tiers are locally refined. Figure 5 shows the respective placement progressions, more details can be found in [28]. The detailed placer from [37] is then invoked for a tier-by-tier standard-cell legalization and detailed placement from the bottom tier (with all the fixed blocks) to the top tier.

In general, we have the post-processing fine-grained 2D placement interleaved with the coarse-grained 3D placement, which achieves a good trade-off between quality and efficiency. On average of all the ten IBM-PLACE circuits, the application of 2D-IC refinement reduces the wirelength by more than 4%. 

![Fig. 2: The flowchart of ePlace-3D.](image-url)
We implement ePlace-3D using C programming language in the single-thread mode and execute the program in a Linux machine with Intel i7 920 2.67GHz CPU and 12GB memory. There is no benchmark specific parameter tuning in our work. #VI are controlled by the weighting factor $\beta_s$ based on capacitance ratio. By [15], one TSV (VI) has the capacitance of $C_{VI} = 30fF$ at 45nm tech-node. ITRS annual reports [13] show that unit capacitance of interconnects at intermediate routing layers is constantly $4pF/cm$ across various tech-nodes. Placement row height is $1.4um$ at 45nm tech-node ($70nm$ M1 half-pitch, ten M1 tracks per row), capacitance becomes $C_{ROW} = 0.3fF$ for 2D interconnect spanning one-row height. Based on the length units for each benchmark, as well as our geometric transformation of the placement core region to be $[0,1] \times [0,1] \times [0,1]$ as discussed in Section III-D, we compute the respective capacitance ratio of one VI versus one unit wirelength and use it as the VI weight. Specifically, we have

$$\beta_s = \frac{\text{# tiers} \times C_{VI}}{\text{# rows} \times C_{ROW}}$$

Notice that the focus of this work is the algorithm framework of 3D placement, not the accurate weight modeling of vertical connects. The weighting factor can be adjusted by VLSI designers for their particular needs, e.g., vertical connects of different electric and physical attributes (TSVs, MIVs, super contacts, etc.).

We conduct experiments on the IBM-PLACE [12] standard-cell benchmarks without macros or blockages, where all the cases are derived from real IC design. We include two state-of-the-art 3D-IC analytic placers, mPL6-3D [29], and NTUplace3-3D [9], in our experiments on IBM-PLACE. As other categories of algorithms (e.g., folding and partition based approaches) have been outperformed by analytic placement in literature, we do not include them in our experiments. We have obtained the binary of NTUplace3-3D from the original authors and executed it on our machine for experiments. mPL6-3D is not available (as notified by the author), so we cite the performance from their latest publication [29]. We use exactly the same benchmark transformation as that by mPL6-3D and NTUplace3-3D. I.e., we insert four silicon tiers into each benchmark, scale down each tier to 6× of the original 2D placement area, add 10% whitespace to each tier, and keep the aspect ratio of each tier to be the same as the original 2D design. As a result, all the experiments on the three placers, including those from [29], are conducted on exactly the same IBM-PLACE-3D benchmarks. As HPWL and #VI are being computed in exactly the same way, the performance comparison among the three placers are fair. The results on IBM-PLACE cases are shown in Table I. On average of all the ten circuits, ePlace-3D outperforms mPL6-3D and NTUplace3-3D with 6.44% and 37.15% shorter wirelength together with 9.11% and 10.27% fewer VI. ePlace-3D runs $2.55 \times$ faster than mPL6-3D but $0.30 \times$ slower than NTUplace3-3D, nevertheless, the improvement on wirelength (37.15%) and #VI (10.27%) is significant.

To validate the scalability of ePlace-3D, we also conduct experiments on the large-scale modern mixed-size (MMS) benchmarks [43] with on average 829K and up to 2.5M netlist objects.  

Fig. 4: Progression of 2D-IC mixed-size global placement on MMS ADAPTEC1 benchmark with three tiers. Standard cells, macros and fillers are denoted by red dots, blue rectangles and cyan dots, respectively.

V. EXPERIMENTS AND RESULTS

We implement ePlace-3D using C programming language in the single-thread mode and execute the program in a Linux machine with Intel i7 920 2.67GHz CPU and 12GB memory. There is no benchmark specific parameter tuning in our work. #VI are controlled by the weighting factor $\beta_s$ based on capacitance ratio. By [15], one TSV (VI) has the capacitance of $C_{VI} = 30fF$ at 45nm tech-node. ITRS annual reports [13] show that unit capacitance of interconnects at intermediate routing layers is constantly $2pF/cm$ across various tech-nodes. Placement row height is $1.4um$ at 45nm tech-node ($70nm$ M1 half-pitch, ten M1 tracks per row), capacitance becomes $C_{ROW} = 0.3fF$ for 2D interconnect spanning one-row height. Based on the length units for each benchmark, as well as our geometric transformation of the placement core region to be $[0,1] \times [0,1] \times [0,1]$ as discussed in Section III-D, we compute the respective capacitance ratio of one VI versus one unit wirelength and use it as the VI weight. Specifically, we have

$$\beta_s = \frac{\text{# tiers} \times C_{VI}}{\text{# rows} \times C_{ROW}}$$

Notice that the focus of this work is the algorithm framework of 3D placement, not the accurate weight modeling of vertical connects. The weighting factor can be adjusted by VLSI designers for their particular needs, e.g., vertical connects of different electric and physical attributes (TSVs, MIVs, super contacts, etc.).
TABLE I: HPWL (e7), #VI (vertical interconnect) (e3) and runtime (minutes) on the IBM-PLACE benchmark suite [12]. Cited results are marked with *. All the experiments are conducted under single-thread mode. The results are evaluated by the same scripts and normalized to ePlace-3D. The best result for each case is in bold-face.

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th># Cells</th>
<th># Nets</th>
<th>HPWL</th>
<th>#VI</th>
<th>CPU</th>
<th>HPWL</th>
<th>#VI</th>
<th>CPU</th>
<th>HPWL</th>
<th>#VI</th>
<th>CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM01</td>
<td>12K</td>
<td>12K</td>
<td>0.34</td>
<td>0.69</td>
<td>0.20</td>
<td>0.26</td>
<td>1.04</td>
<td>2.93</td>
<td>0.25</td>
<td>1.31</td>
<td>0.38</td>
</tr>
<tr>
<td>IBM02</td>
<td>22K</td>
<td>22K</td>
<td>0.76</td>
<td>3.32</td>
<td>0.50</td>
<td>0.59</td>
<td>3.11</td>
<td>3.73</td>
<td>0.55</td>
<td>3.32</td>
<td>1.33</td>
</tr>
<tr>
<td>IBM03</td>
<td>27K</td>
<td>26K</td>
<td>1.00</td>
<td>2.60</td>
<td>0.60</td>
<td>0.81</td>
<td>2.95</td>
<td>6.41</td>
<td>0.74</td>
<td>3.53</td>
<td>1.88</td>
</tr>
<tr>
<td>IBM04</td>
<td>32K</td>
<td>32K</td>
<td>1.30</td>
<td>3.99</td>
<td>0.80</td>
<td>1.05</td>
<td>3.97</td>
<td>6.20</td>
<td>0.92</td>
<td>4.50</td>
<td>2.98</td>
</tr>
<tr>
<td>IBM05</td>
<td>45K</td>
<td>44K</td>
<td>1.92</td>
<td>5.73</td>
<td>1.30</td>
<td>1.59</td>
<td>4.68</td>
<td>8.64</td>
<td>1.50</td>
<td>4.39</td>
<td>3.87</td>
</tr>
<tr>
<td>IBM06</td>
<td>51K</td>
<td>48K</td>
<td>2.08</td>
<td>4.90</td>
<td>1.70</td>
<td>1.71</td>
<td>3.94</td>
<td>11.23</td>
<td>1.54</td>
<td>4.90</td>
<td>4.75</td>
</tr>
<tr>
<td>IBM07</td>
<td>52K</td>
<td>50K</td>
<td>2.39</td>
<td>5.88</td>
<td>1.50</td>
<td>1.45</td>
<td>3.34</td>
<td>14.61</td>
<td>1.40</td>
<td>3.18</td>
<td>5.63</td>
</tr>
<tr>
<td>IBM12</td>
<td>82K</td>
<td>84K</td>
<td>3.69</td>
<td>3.98</td>
<td>2.60</td>
<td>2.88</td>
<td>5.59</td>
<td>19.62</td>
<td>2.67</td>
<td>4.73</td>
<td>8.65</td>
</tr>
<tr>
<td>IBM13</td>
<td>138K</td>
<td>161K</td>
<td>9.16</td>
<td>15.67</td>
<td>7.20</td>
<td>6.79</td>
<td>10.52</td>
<td>46.82</td>
<td>6.39</td>
<td>9.16</td>
<td>40.25</td>
</tr>
<tr>
<td>IBM18</td>
<td>201K</td>
<td>201K</td>
<td>13.41</td>
<td>129.19</td>
<td>13.60</td>
<td>9.16</td>
<td>15.22</td>
<td>53.04</td>
<td>9.45</td>
<td>6.83</td>
<td>63.07</td>
</tr>
<tr>
<td>Avg.</td>
<td>69K</td>
<td>69K</td>
<td>37.15</td>
<td>10.27</td>
<td>0.30</td>
<td>6.44</td>
<td>9.11</td>
<td>2.55</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

TABLE II: HPWL (e6), scaled HPWL (e6, with 1), #VI (vertical interconnect) and runtime (mins) on MMS circuits. Cited results are marked with *. All the experiments are in single-thread mode. The HPWL and CPU results are normalized to the best published 2D placement results [28]. #VI are normalized to # objects.

<table>
<thead>
<tr>
<th>Benchmarks</th>
<th># Objects</th>
<th># Nets</th>
<th>ePlace-MS* [28]</th>
<th>ePlace-3D w/ 2 tiers</th>
<th>ePlace-3D w/ 3 tiers</th>
<th>ePlace-3D w/ 4 tiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAPTEC1</td>
<td>211K</td>
<td>221K</td>
<td>67.17</td>
<td>5.47</td>
<td>59.51</td>
<td>57.33</td>
</tr>
<tr>
<td>ADAPTEC2</td>
<td>255K</td>
<td>266K</td>
<td>73.77</td>
<td>7.43</td>
<td>73.97</td>
<td>92.69</td>
</tr>
<tr>
<td>ADAPTEC3</td>
<td>451K</td>
<td>466K</td>
<td>164.50</td>
<td>27.23</td>
<td>141.97</td>
<td>5557</td>
</tr>
<tr>
<td>ADAPTEC4</td>
<td>496K</td>
<td>515K</td>
<td>148.38</td>
<td>29.35</td>
<td>126.94</td>
<td>8149</td>
</tr>
<tr>
<td>BIGBLUE1</td>
<td>278K</td>
<td>284K</td>
<td>86.82</td>
<td>7.82</td>
<td>76.06</td>
<td>8222</td>
</tr>
<tr>
<td>BIGBLUE2</td>
<td>557K</td>
<td>572K</td>
<td>130.18</td>
<td>13.79</td>
<td>109.22</td>
<td>2565</td>
</tr>
<tr>
<td>BIGBLUE4</td>
<td>1096K</td>
<td>1126K</td>
<td>72.98</td>
<td>25.17</td>
<td>251.77</td>
<td>24466</td>
</tr>
<tr>
<td>BIGBLUE4  †</td>
<td>2177K</td>
<td>2297K</td>
<td>657.92</td>
<td>204.15</td>
<td>577.98</td>
<td>21263</td>
</tr>
<tr>
<td>BIGBLUE4  †</td>
<td>843K</td>
<td>867K</td>
<td>315.76</td>
<td>48.35</td>
<td>258.18</td>
<td>22705</td>
</tr>
<tr>
<td>NEWBLUE1</td>
<td>330K</td>
<td>338K</td>
<td>62.56</td>
<td>10.87</td>
<td>56.36</td>
<td>90.51</td>
</tr>
<tr>
<td>NEWBLUE2</td>
<td>444K</td>
<td>465K</td>
<td>169.59</td>
<td>62.47</td>
<td>179.82</td>
<td>25571</td>
</tr>
<tr>
<td>NEWBLUE3</td>
<td>494K</td>
<td>523K</td>
<td>304.24</td>
<td>17.53</td>
<td>240.47</td>
<td>7686</td>
</tr>
<tr>
<td>NEWBLUE5</td>
<td>646K</td>
<td>672K</td>
<td>229.95</td>
<td>29.73</td>
<td>197.21</td>
<td>11372</td>
</tr>
<tr>
<td>NEWBLUE6</td>
<td>1233K</td>
<td>1284K</td>
<td>393.21</td>
<td>63.40</td>
<td>344.95</td>
<td>45995</td>
</tr>
<tr>
<td>NEWBLUE7</td>
<td>1255K</td>
<td>1288K</td>
<td>410.04</td>
<td>69.65</td>
<td>379.59</td>
<td>10901</td>
</tr>
<tr>
<td>NEWBLUE7</td>
<td>2507K</td>
<td>2636K</td>
<td>897.81</td>
<td>191.47</td>
<td>814.79</td>
<td>18615</td>
</tr>
<tr>
<td>Avg.</td>
<td>829K</td>
<td>859K</td>
<td>0.000</td>
<td>1.00x</td>
<td>13.67x</td>
<td>2.11x</td>
</tr>
</tbody>
</table>

MMS benchmarks was first published in DAC 2009. The circuits inherit the same netlists and density constraints \( \rho_t \) from ISPD 2005 [33] and ISPD 2006 [32] benchmarks but have all the macros freed to place. The original planar placement domain is geometrically transformed to be of 2, 3 and 4 silicon tiers, each tier is equally downsized to keep both the aspect ratio and total silicon area unchanged. All the standard cells and macros keep their original dimensions and span only one tier. MMS circuits have all their fixed objects with zero area (volume) and outside the placement boundaries, and we geometrically transform them to the boundary of the bottom (first) tier. Also, as macros are all free to move, we skip the geometrical transformation of the fixed macro from 2D to 3D, which is sub-optimal and usually causes quality loss. Similar to mPL6-3D [29] and NTUplace3-3D [9], we add 10% extra whitespace to each tier, in order to relieve the placement dilemma due to the increased area ratio between large macros and silicon tiers. There are benchmark-dependent target density \( \rho_t \) for eight out of the sixteen MMS circuits. By the benchmark protocol [32], exceeding \( \rho_t \) will scale up the wirelength by \( s_{HPWL} = HPWL \times (1 + 0.01 \times \tau_{avg}) \), where \( \tau_{avg} \) denotes the scaled density overlap per bin and \( s_{HPWL} \) is the penalized wirelength. Detailed circuit statistics can be found in Table 1 of [43]. We create evaluation scripts to compute the total wirelength, number of vertical interconnects, legality and density overflow of the produced 3D-IC placement solution. The results on the MMS benchmarks are shown in Table II. The binary of NTUplace3-3D does not work with these benchmarks, while the binary of mPL6-3D is not available for use. As a result, we compare the 3D DM placement solutions with the best published (golden) 2D results in literature [28]. By using two, three and four tiers, ePlace-3D outperforms the golden 2D placement with on average (13.67%, 20.50% and 27.54% shorter wirelength. On the other side, the average ratio between the number of vertical interconnect units versus the number of placement objects (standard cells and macros) are only 2.17%, 4.30% and 6.10%, respectively. These vertical connect ratios are much smaller than the average VI ratio on IBM-PLACE, which are more than 9% for all the three placers in Table I. Due to the introduction of the third dimension, the search space of placement optimization is substantially enlarged. However, the runtime increase is just 3x, which indicates high efficiency of ePlace-3D.

We also study the trends of HPWL and #VI by linearly sweeping the number of tiers and exponentially sweeping the VI weight. We select eight out of the sixteen MMS benchmarks (ADAPTEC1, ADAPTEC4, BIGBLUE1, BIGBLUE2, BIGBLUE3, BIGBLUE4, NEWBLUE6, NEWBLUE7), all of which could accommodate the maximum macro block after inserting ten tiers. Keeping the same aspect ratio, the area of each tier is scaled down by ten times with the insertion of 10% extra whitespace. Figure 6 shows that ePlace-3D is able to reduce the total 2D wirelength by up to 40% (with the insertion of up to ten tiers), while #VI is roughly scaled up by the
Fig. 5: Post-placement progression on the MMS ADAPTEC1 benchmark with three tiers. Standard cells, macros and fillers are denoted by red dots, blue rectangles and cyan dots, respectively. Om denotes the total macro overlaps.

number of tiers. VI weight sweeping is conducted on all the sixteen circles with 2.2M and 2.5M cells, and they consume

(a) 3D macro legalization: iter=0, HPWL=4.99e7, #VI=7.96e3, Om=9.03e5.
(b) 3D macro legalization: iter=4, HPWL=5.10e7, #VI=9.10e3, Om=0.
(c) 3D standard-cell global placement: iter=0, HPWL=5.10e7, #VI=9.10e3, Om=0.
(d) 2D standard cell global placement: iter=0, HPWL=5.10e7, #VI=9.10e3, τ = 8.1%.
(e) 2D standard cell global placement: iter=4, HPWL=5.08e7, #VI=9.10e3, τ = 66.7%.
(f) 2D standard cell detailed placement: iter=394, HPWL=5.08e7, #VI=9.10e3, τ = 14.8%.

Fig. 6: Average HPWL and #VI of eight selected MMS benchmarks with linearly swept number of silicon tiers.

MMS benchmarks. Figure 7 shows the trends of average HPWL and #VI by dividing the normal VI weight by up to 32 times (i.e. × 0.03125). The total 2D HPWL saturates at the reduction of 7%, while #VI is scaled up by roughly 2.5×.

Our 3D-IC placement algorithm shows significant quality improvement while limited runtime overhead. BIGBLUE4 and NEWBLUE7 are the largest circuits with 2.2M and 2.5M cells, and they consume the longest runtime on the 3D-IC placement. However, compared to the respective golden 2D placement solutions, the runtime ratio is upper-bounded by 2.5×, which is still less than the average runtime ratios of 3.13× for two tiers, 3.03× for three tiers and 2.94× for four tiers, respectively. To this end, ePlace-3D shows good scalability and acceptable efficiency on the large cases.

In this work, we do not test ePlace-3D on circuits with fixed macros, as geometrically transforming the 2D floorplan into 3D is difficult and usually error-prone. However, ePlace-3D shows high performance and scalability on MMS benchmarks with lots of movable large macros, which is more difficult to place than fixed-macro layouts. As a result, we are confident on the performance of ePlace-3D on any circuits with fixed macros. The advantage of 3D tier insertion vanishes if there are large macros to accommodate (BIGBLUE3, NEWBLUE3, etc.). Transformation of 2D planar macros into 3D cuboid macros would resolve this issue and ensure the consistent benefits by inserting more and more tiers. However, it is beyond this work and will be studied in our future research.

VI. Conclusion

We propose the first electrostatics based placement algorithm ePlace-3D, which is effective and efficient for 3D-ICs with uniform exploration over the entire 3D space. Our 3D-IC density function leverages the analogy between placement spreading and electrostatic equilibrium, while global and uniform smoothness is realized at all the three dimensions. Our balancing and preconditioning techniques prevent solution oscillation or divergence. The interleaved 3D coarse-grained optimization followed by 2D fine-grained post processing effectively obtains a good trade-off between quality and efficiency. The experimental results validate the high performance and scalability of our approach, indicating the benefits of placement smoothness to the overall quality. In future, we will develop 3D density function to be aware of the volume of vertical connects.

VII. Acknowledgment

The authors would like to acknowledge (1) Prof. Dae Hyun Kim and Prof. Sung Kyu Lim for providing the 3D-IC flow scripts and ISPD testcases (2) Dr. Meng-Kai Hsu and Prof. Yao-Wen Chang for providing the binary of NTUplace3-3D (3) Dr. Guojie Luo and Prof. Jason Cong for providing the binary of mPL6-3D (4) the support of NSF CCF-1017864.

REFERENCES
