Hypergraph Partitioning for Exploiting Localities in Nonlinear Constrained Optimization

Benjamin W. Wah and Soomin Lee
Department of Electrical and Computer Engineering
and the Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL61801, USA
http://manip.crhc.uiuc.edu

IEEE International Conference on Tools with Artificial Intelligence, 2007
10/29/2007

1

Outline

- Introduction
  - Constrained optimization: definition, examples
  - Serial versus parallel decomposition
- Problem Definition
  - MINCUT_HG_SHARE
- Applications
  - Nonlinear programming (NLP)
  - VLSI placement and routing
- Proposed Algorithm
  - Graphical overview
  - Detailed procedure
- Experimental Results
  - Trade-offs on parallel decomposition in NLP/VLSI benchmarks
  - Comparison of wirelengths on VLSI benchmarks
- Conclusions

2

Constrained Optimization

- Definition

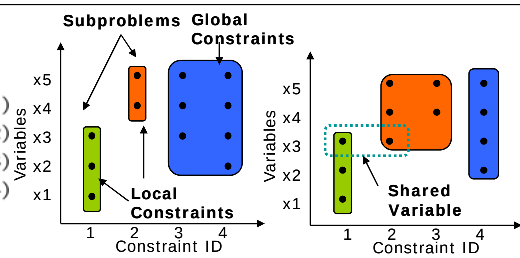
$$\min_z \quad f(z)$$
$$\text{subject to} \quad h_i(z) = 0 \quad i = 1, \ldots, m$$
$$g_j(z) \le 0 \quad j = 1, \ldots, r$$

where $z = (x, y)$, $f : \mathbf{R}^n \times \mathbf{Z}^p \to \mathbf{R}$, $x \in \mathbf{R}^v$ and $y \in \mathbf{Z}^w$

Find an optimal point z^* such that $f(z^*) \le f(z)$ for $\forall z$ satisfying the constraints.

- Example

$$\min_x f(x) = x_1x_2 + x_3^2 + x_4x_5$$
subject to
$$x_1 + x_2 + x_3 = 1 \quad (1)$$
$$x_4^2 - x_5^2 \le 2 \quad (2)$$
$$x_3^2 + x_4^2 + x_5^2 \le 10 \quad (3)$$
$$x_2^2 + x_3x_4 + x_5^2 \le 7 \quad (4)$$



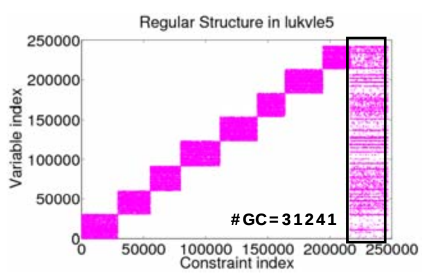
ECE Illinois DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Serial/Parallel Decomposition

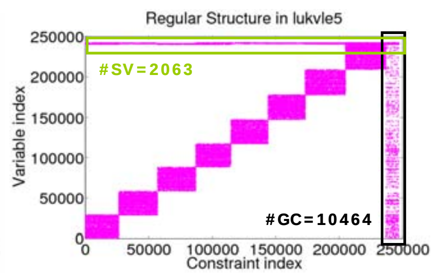
- Decomposition [Guestrin/Gordon]
  - Partition a state space into subproblems
- Serial decomposition
  - Combined state space is the union of the subproblem state spaces
  - Complexity of each subspace is similar to original
- Parallel decomposition
  - Combined state space is the cross product of subproblem state spaces
  - Lead to shared variables and global constraints across partitioned state spaces
  - Exponentially smaller subproblem subspaces

Good Localities ⇒Parallel Decomposition

- Good variable/constraint localities in constrained optimization
  - Example NLP – lukvle5 (250000 variables / 249996 constraints)



$\text{Overhead} = N \times T_{sub} + T_{GC}$



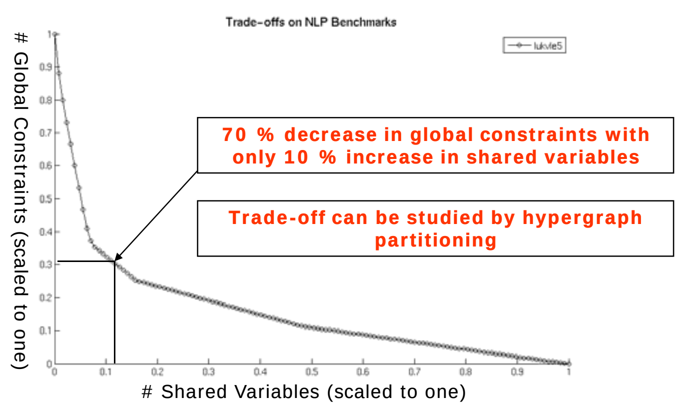
$\text{Overhead} = N \times T'_{sub} + T'_{GC} + T_{SV}$

5

Trade-offs in Parallel Decomposition

Example NLP – lukvle5

250000 variables & 249996 constraints

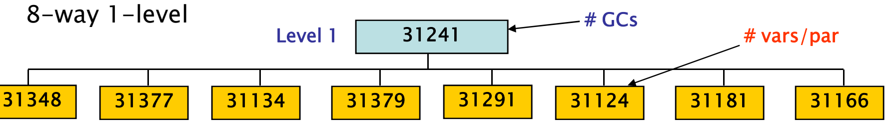


6

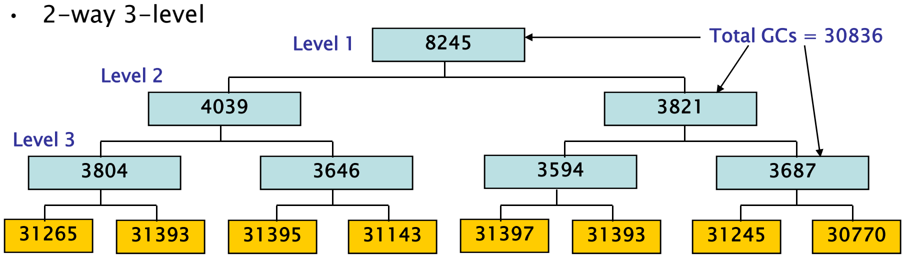
Multiple-Level Parallel Decomposition

- 8-way partitioning
  - Example NLP – lukvle5 (250000 variables / 249996 constraints)

- 8-way 1-level



- 2-way 3-level



- Further decomposition of the global constraints in this problem is not helpful. It is solvable by one-level decomposition.

Outline

- Introduction
- Problem Definition
  - MINCUT_HG_SHARE
- Applications
- Proposed Algorithm
- Experimental Results
- Conclusions

Problem Definition:
MINCUT_HYPERGRAPH_SHARE

Definition. *MINCUT_HYPERGRAPH_SHARE*

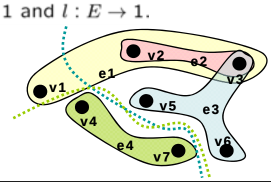
Input: $H = (V, E),\ w : V \rightarrow Z^{+},\ l : E \rightarrow Z^{+}$, and $K, J, M \in Z^{+}$

Property: Partition V into possibly non-disjoint subsets $V_1, V_2, \ldots, V_m$, s.t. $\sum_{v \in V_i} w(v) \leq K$, $\sum_{e \in E''} l(e) \leq J$, and $\sum_{v \in V'} w(v) \leq M$ where E'' has hyperedges that connect strictly two or more subsets, and V' are shared vertices.

Example. *MINCUT_HYPERGRAPH_SHARE*

Partition in 2-way with $K = 5,\ J = 1,\ M = 1,\ w : V \rightarrow 1$ and $l : E \rightarrow 1$.

$V_1 = \{v_1, v_4, v_7\},\ V_2 = \{v_1, v_2, v_3, v_5, v_6\}$
$|V_1|,\ |V_2| \leq 5 (= K)$
$E'' = \emptyset,\ |E''| \leq 1 (= J)$
$V' = \{v_1\},\ |V'| \leq 1 (= M)$



MINCUT_HYPERGRAPH_SHARE is NP-complete (see proof in paper)

Outline

- Introduction
- Problem Definition
- Applications
  - Nonlinear programming (NLP)
  - VLSI placement and routing
- Proposed Algorithm
- Experimental Results
- Conclusions

Applications:
Nonlinear Programming (NLP)

Definition. Nonlinear programming

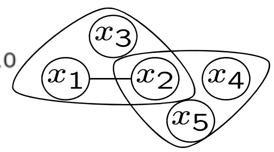
$$\min_z \quad f(z)$$
$$\text{subject to} \quad h_i(z) = 0 \quad i = 1, \ldots, m$$
$$g_j(z) \le 0 \quad j = 1, \ldots, p$$

where $z = (x, y)$, $f : \mathbf{R}^v \times \mathbf{D}^w \to \mathbf{R}$, $x \in \mathbf{R}^v$ and $y \in \mathbf{D}^w$

Find an optimal point z^* such that $f(z^*) \le f(z)$ for $\forall z$ satisfying the constraints.

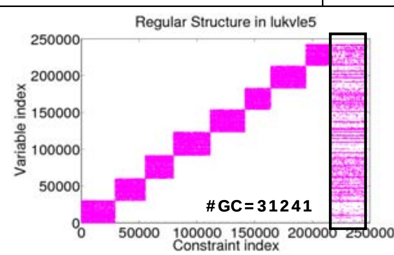
NLP and corresponding hypergraph

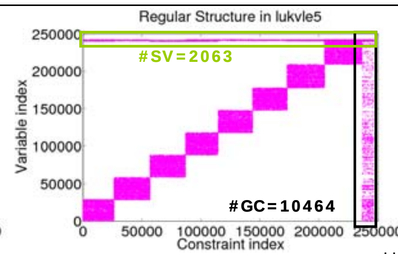
$$\min_x f(x) = x_1x_2 + x_2x_3 + x_4x_5$$
$$x_1 + x_2 = 1$$
$$x_1^2 - x_2^2 + x_3^2 \le 2$$
$$x_2^2 + x_4^2 + x_5^2 \le 10$$



Example. lukvle5

250000 variables
249996 constraints





11

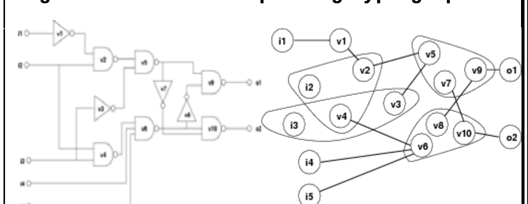
Applications:
VLSI Placement and Routing

Definition. VLSI placement

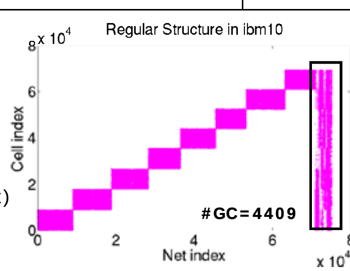
1. n cells, pads: $v_1, v_2, \ldots, v_n$
2. m nets: $e_1, e_2, \ldots, e_m$
3. COST $= \sum_i WL(e_i)$

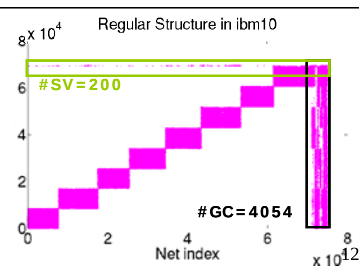
Find (x_j, y_j) of each cell v_j on the die without overlaps such that the COST is minimized.

Logic circuit and corresponding hypergraph



Example. Ibm10
68685 cells
744 pads on perimeter
75196 nets (up to 41 cells/net)
297567 pins





12

ECE Illinois — DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

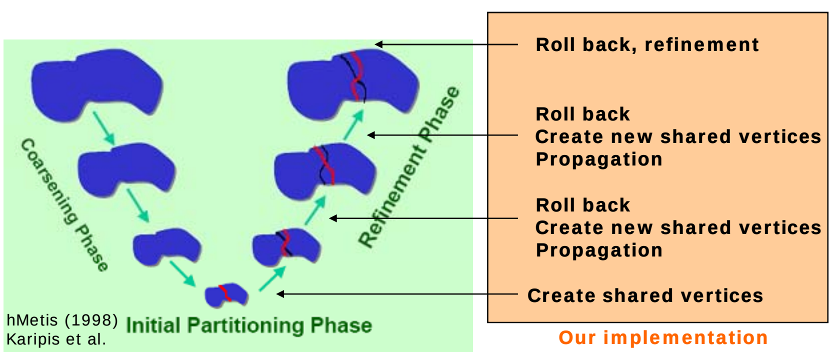
Outline

- Introduction
- Problem Definition
- Applications
- Proposed Algorithm
  - Graphical overview
  - Detailed procedure
- Experimental Results
- Conclusions

13

ECE Illinois — DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Proposed HGP Algorithm: Graphical overview



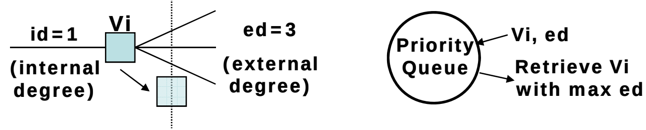
Our implementation

14

Proposed HGP Algorithm:
Detailed Procedure

<Cut measurement and creating shared vertices>

Share vertex with maximum external degree (= maximum reduction in cut)



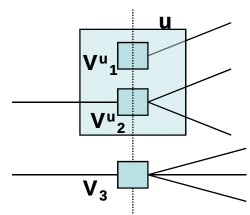
<Propagation to the next level>

Propagate shared vertices to the upper level

<Rollback computation>

Some vertices need not to be shared

- **u**: Coarser level vertex
- V^u_1, V^u_2 : The next finer level vertices of u
- V^u_1 : Roll back to its original partition
- V^u_2 : Keep it as a shared vertex
- V_3 : Add one more shared vertex at this level



15

ECE Illinois
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Outline

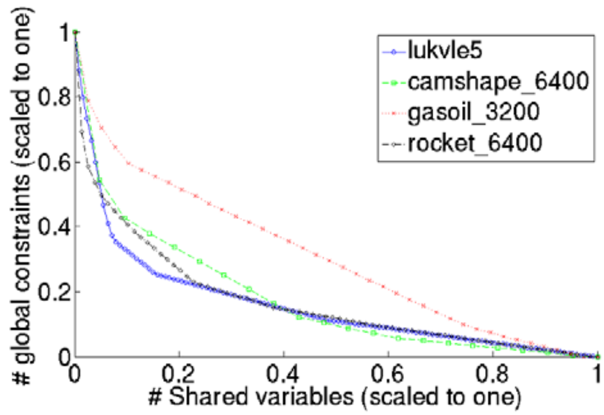
- Introduction
- Problem Definition
- Applications
- Proposed Algorithm
- Experimental Results
  - Trade-offs on parallel decomposition in NLP/VLSI benchmarks
  - Comparison of wirelengths on VLSI benchmarks
- Conclusions

16

ECE Illinois

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Trade-offs on Parallel Decomposition:
NLP Benchmarks



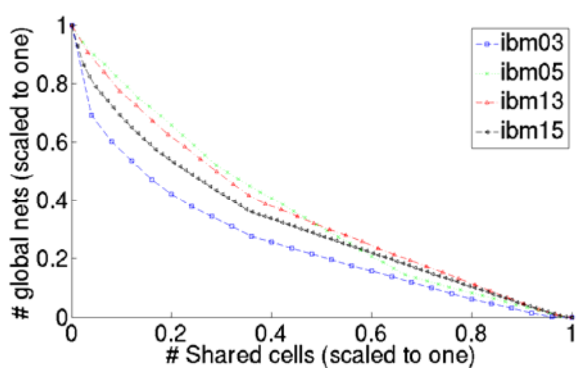
17

ECE Illinois

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Trade-offs on Parallel Decomposition:
VLSI Placement and Routing

- **ISPD04 Benchmark Suite**



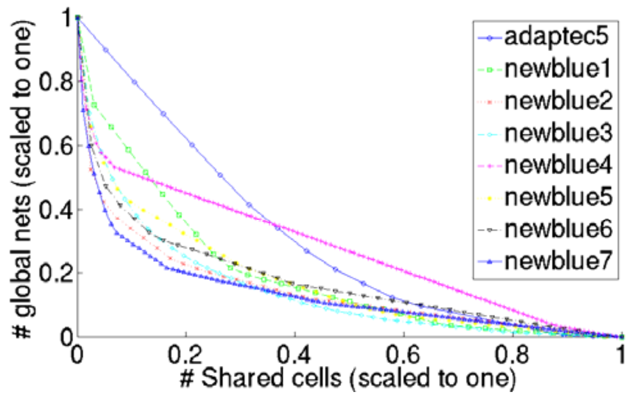
18

ECE Illinois — DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Trade-offs on Parallel Decomposition:

VLSI Placement and Routing

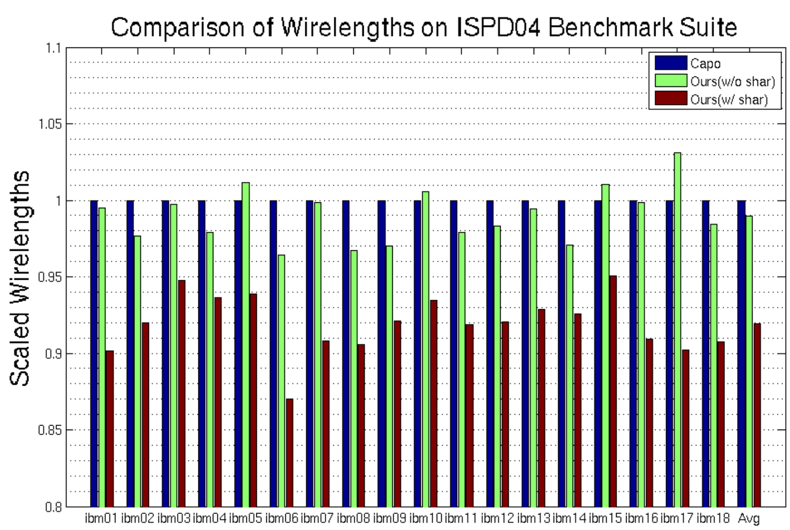
- **ISPD06 Benchmark Suite**



19

ECE Illinois — DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

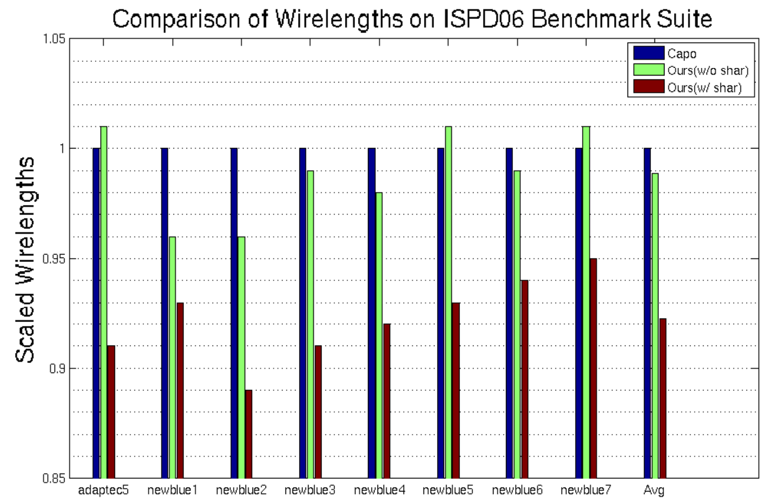
Results on Placement/Routing



20

ECE Illinois — DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Results on Placement/Routing



21

ECE Illinois — DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Conclusions

- Variable/constraint localities in constrained optimization
- Illustrated trade-offs between shared variables and global constraints in parallel decomposition
- Developed a new hypergraph partitioning algorithm
- Demonstrated improvements on wirelengths in the VLSI placement and routing problem

22