

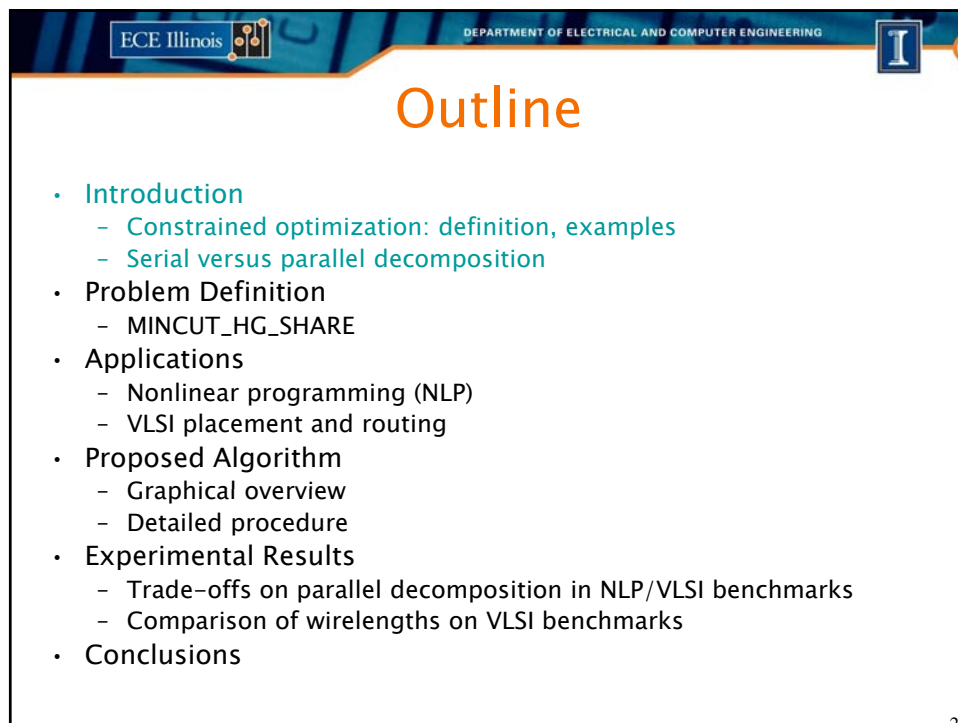
ECE Illinois

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN COLLEGE OF ENGINEERING

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.erhc.uiuc.edu>

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




ECE Illinois

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING


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



DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING



Constrained Optimization

■ **Definition**

$$\min_z f(z)$$

subject to $h_i(z) = 0 \quad i = 1, \dots, m$
 $g_j(z) \leq 0 \quad j = 1, \dots, r$

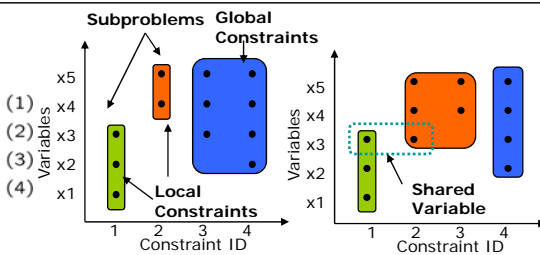
where $z = (x, y)$, $f : \mathbb{R}^n \times \mathbb{Z}^p \rightarrow \mathbb{R}$, $x \in \mathbb{R}^n$ and $y \in \mathbb{Z}^p$

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


■ **Example**

$$\min_x f(x) = x_1x_2 + x_3^2 + x_4x_5$$


subject to $x_1 + x_2 + x_3 = 1$
 $x_4^2 - x_5^2 \leq 2$
 $x_3^2 + x_4^2 + x_5^2 \leq 10$
 $x_2^2 + x_3x_4 + x_5^2 \leq 7$



The diagram consists of two plots. The left plot shows a 5D space with axes x1 to x5. It is divided into subproblems (1) to (4) along the x1 axis. Subproblem (1) has local constraints (1) and (2). Subproblems (2) and (3) have local constraints (3) and (4). Subproblem (4) has no local constraints. Global constraints are shown as a blue box. The right plot shows the same space with a 'Shared Variable' constraint (dotted line) and global constraints (orange and blue boxes).



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

ECE Illinois
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Good Localities \Rightarrow Parallel Decomposition

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

Regular Structure in lukvle5

Variable index

Constraint index

#GC=31241

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

Regular Structure in lukvle5

Variable index

Constraint index

#SV=2063

#GC=10464

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

5

ECE Illinois
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

Trade-offs in Parallel Decomposition

Example NLP - lukvle5

250000 variables & 249996 constraints

Trade-offs on NLP Benchmarks

lukvle5



Global Constraints (scaled to one)

Shared Variables (scaled to one)

70 % decrease in global constraints with only 10 % increase in shared variables

Trade-off can be studied by hypergraph partitioning

6

ECE Illinois  DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING 


Multiple-Level Parallel Decomposition

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

- 8-way 1-level

Level 1: 31241 (# GCs)

vars/par

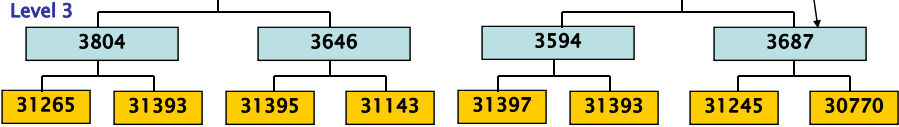


- 2-way 3-level



Level 1: 8245 (Total GCs = 30836)

Level 2: 4039, 3821

Level 3: 3804, 3646, 3594, 3687




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


ECE Illinois  DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING 

Outline

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



DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING



Problem Definition: MINCUT_HYPERGRAPH_SHARE

Definition. MINCUT_HYPERGRAPH_SHARE

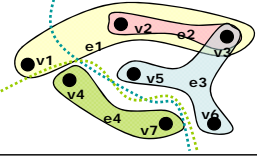
Input: $H = (V, E)$, $w : V \rightarrow Z^+$, $l : E \rightarrow Z^+$, and $K, J, M \in Z^+$

Property: Partition V into possibly non-disjoint subsets V_1, V_2, \dots, 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)




DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING




Outline

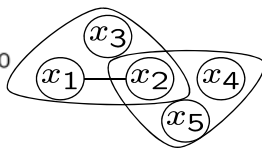
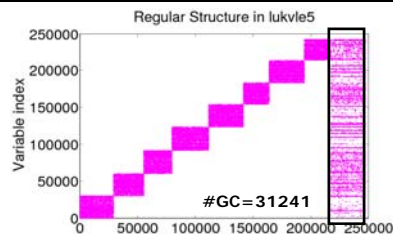
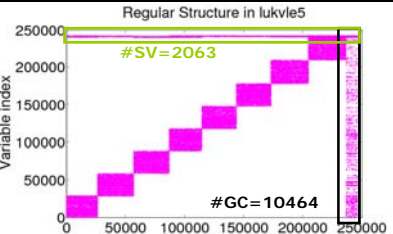
- Introduction
- Problem Definition
- Applications
 - Nonlinear programming (NLP)
 - VLSI placement and routing
- Proposed Algorithm
- Experimental Results
- Conclusions




DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING




Applications: Nonlinear Programming (NLP)

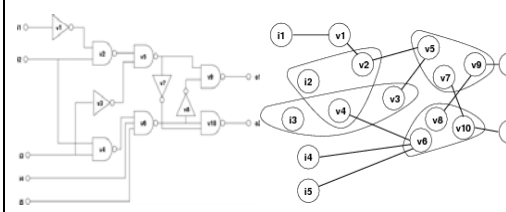
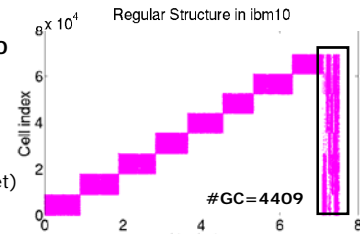
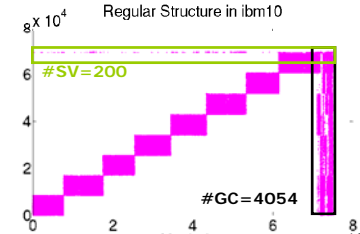
<p>Definition. Nonlinear programming</p> $\min_z f(z)$ <p>subject to $h_i(z) = 0 \quad i = 1, \dots, m$ $g_j(z) \leq 0 \quad j = 1, \dots, p$</p> <p>where $z = (x, y)$, $f : \mathbb{R}^v \times \mathbb{D}^w \rightarrow \mathbb{R}$, $x \in \mathbb{R}^v$ and $y \in \mathbb{D}^w$</p> <p>Find an optimal point z^* such that $f(z^*) \leq f(z)$ for $\forall z$ satisfying the constraints.</p>	<p>NLP and corresponding hypergraph</p> $\min_x f(x) = x_1x_2 + x_2x_3 + x_4x_5$ $x_1 + x_2 = 1$ $x_1^2 - x_2^2 + x_3^2 \leq 2$ $x_3^2 + x_4^2 + x_5^2 \leq 10$ 
<p>Example. lukvle5</p> <p>250000 variables 249996 constraints</p>	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Regular Structure in lukvle5</p>  <p>#GC=31241</p> </div> <div style="text-align: center;"> <p>Regular Structure in lukvle5</p>  <p>#SV=2063 #GC=10464</p> </div> </div>





DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING



Applications: VLSI Placement and Routing



<p>Definition. VLSI placement</p> <ol style="list-style-type: none"> n cells, pads: v_1, v_2, \dots, v_n m nets: e_1, e_2, \dots, e_m $COST = \sum_i WL(e_i)$ <p>Find (x_j, y_j) of each cell v_j on the die without overlaps such that the COST is minimized.</p>	<p>Logic circuit and corresponding hypergraph</p> 
<p>Example. lbm10</p> <p>68685 cells 744 pads on perimeter 75196 nets (up to 41 cells/net) 297567 pins</p>	<div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Regular Structure in lbm10</p>  <p>#GC=4409</p> </div> <div style="text-align: center;"> <p>Regular Structure in lbm10</p>  <p>#SV=200 #GC=4054</p> </div> </div>

ECE Illinois  DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING 

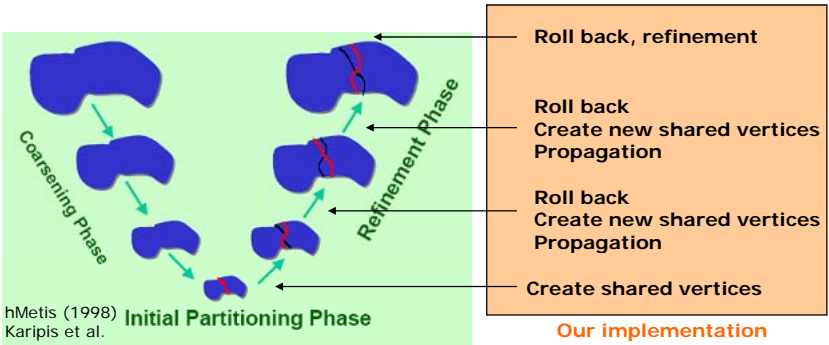
Outline

- 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



hMetis (1998)
Karypis et al.

Initial Partitioning Phase


Coarsening Phase

Refinement Phase


- Roll back, refinement
- Roll back
- Create new shared vertices
- Propagation
- Roll back
- Create new shared vertices
- Propagation
- Create shared vertices

Our implementation

14



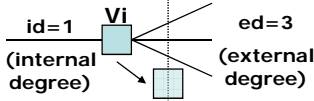
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING



Proposed HGP Algorithm: Detailed Procedure

<Cut measurement and creating shared vertices>
Share vertex with maximum external degree (= maximum reduction in cut)

id=1
(internal degree)



ed=3
(external degree)

Priority Queue

Retrieve V_i with max ed

<Propagation to the next level>
Propagate shared vertices to the upper level

<Rollback computation>
Some vertices need not to be shared

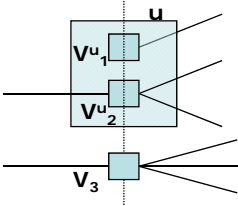
u: Coarser level vertex

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



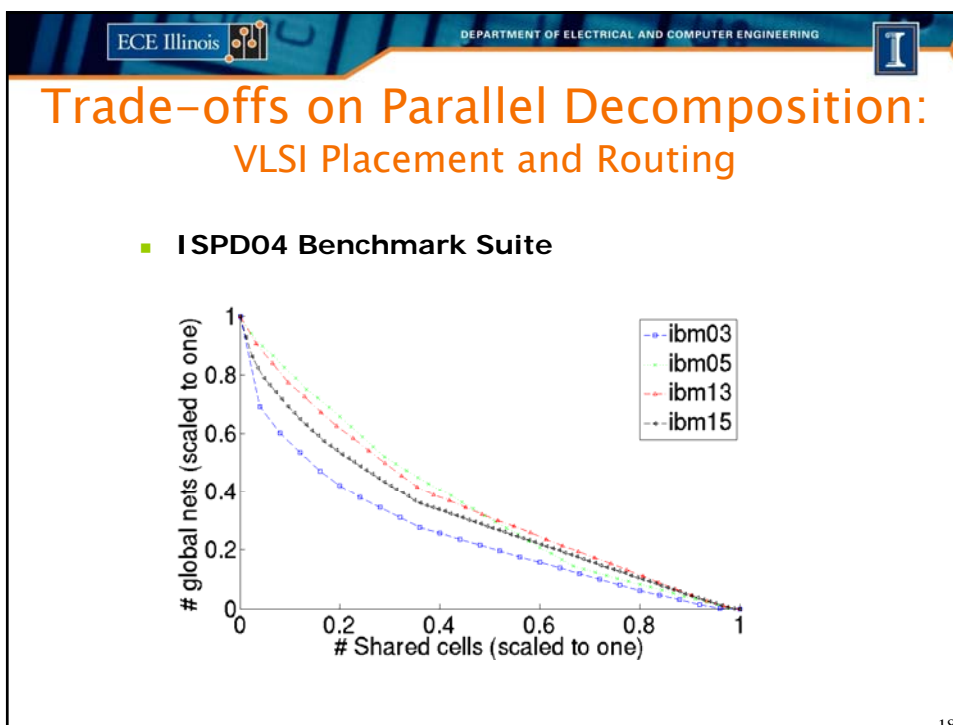
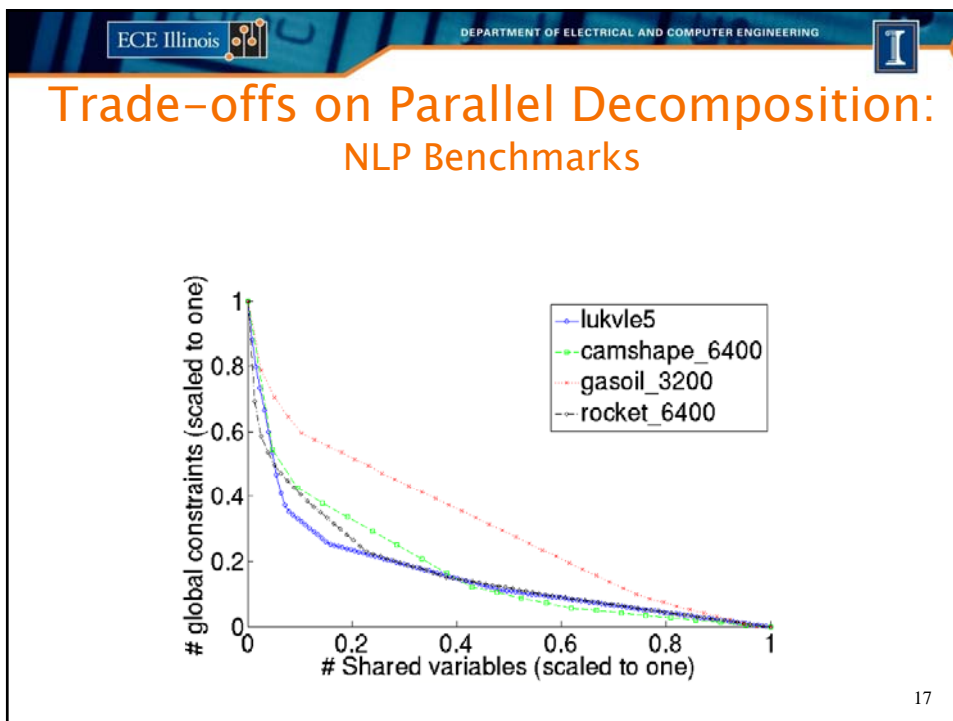
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

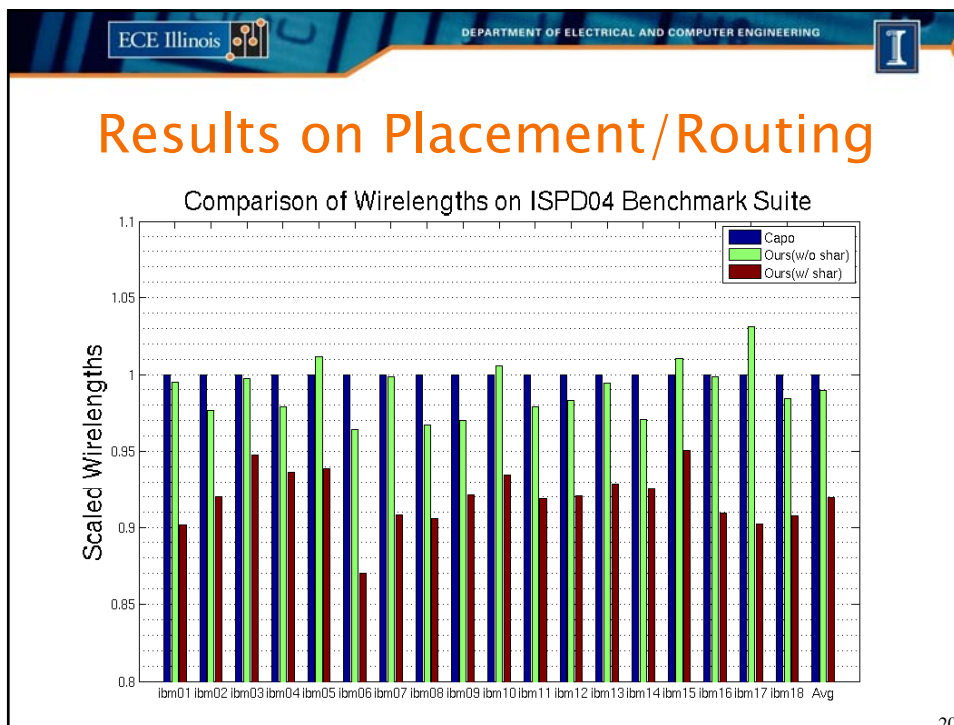
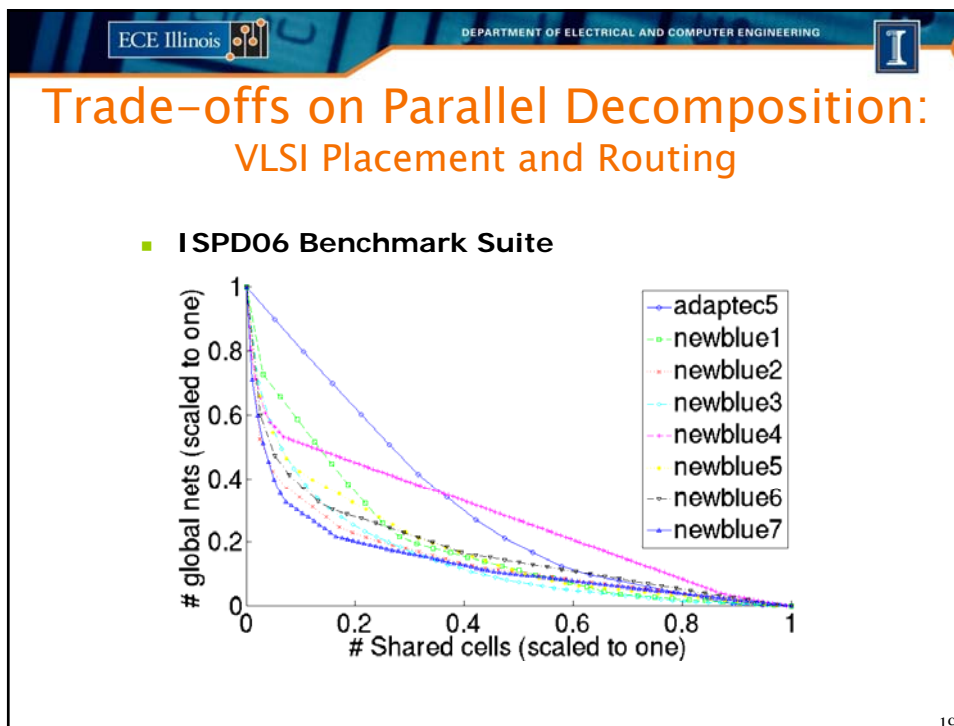


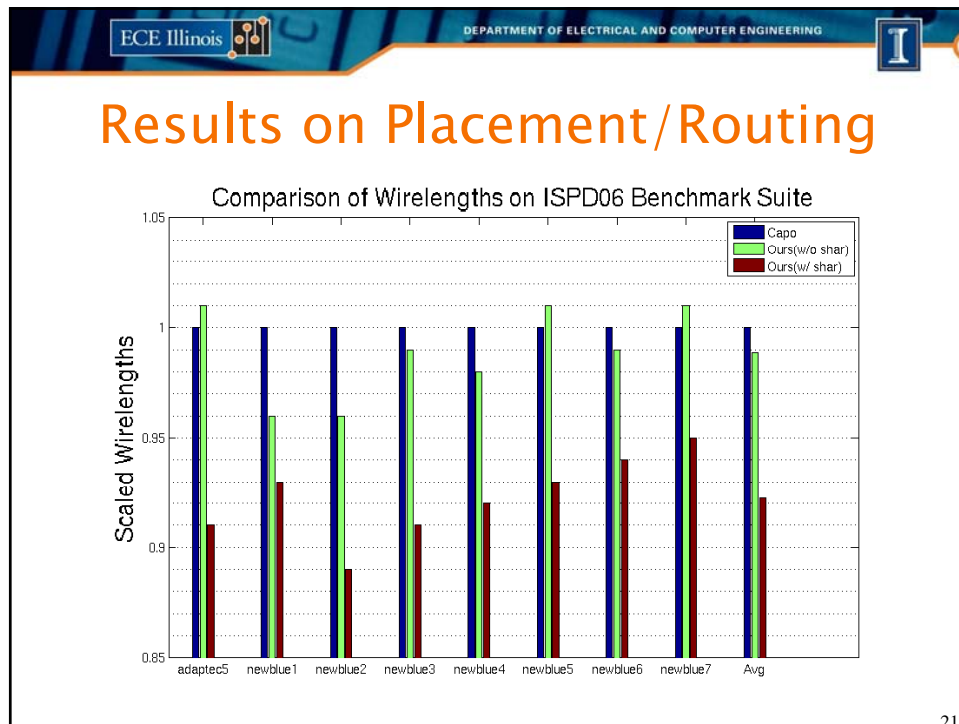
Outline

- 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

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