Deflection Routing in Hierarchical Mesh NoCs

Gadi Oxman, *Student Member, IEEE*, and Shlomo Weiss, *Senior Member, IEEE*

***Abstract*—Deflection routing has benefits that may make it attractive especially when inexpensive network-on-chip (NoC) implementations are considered. In this letter, we present a proposal of deflection routing in hierarchical mesh NoCs. An evaluation of performance and power, obtained with a cycle accurate simulator and using both synthetic and application benchmarks, is reported. We introduce interleaving and shifting, two techniques that reduce the number of routers with a high radix crossbar without reducing network performance.**

***Index Terms*—Deflection routing, hierarchical topology, network-on-chip (NoC).**

I. Introduction

DEFLECTION routing is a method that resolves contention for an output link of the router by allowing one flit to win arbitration and proceed, while other flits may be deflected away from their shortest path to the destination. Deflection routing is an inherently adaptive algorithm that routes around hot-spots, alleviating congestion, as well as providing fault tolerance by routing around faults [1]. Additional benefits include, for example, the independent way in which routing decisions are made by each router, without consulting the buffer space of nearby routers prior to sending a packet. Deflection routers may be implemented inexpensively, an important consideration in networks-on-chip (NoCs), which are often streamlined versions of computer networks.

To our best knowledge, this is the first proposal of deflection routing in hierarchical mesh NoCs. We further propose interleaving and shifting, two techniques that reduce the number of links in high radix routers without lowering network performance. Finally, we evaluate a hierarchical mesh topology that incorporates deflection routers and report performance and power figures obtained with synthetic and application benchmarks.

II. Related Work

Bufferless deflection routing has been proposed for use in on-chip networks [1]–[5]. Hierarchy was used to improve off-chip multiprocessor networks [6]. With the introduction of on-chip networks, various hierarchical NoCs have been proposed in the literature. They can be roughly divided into two groups. The first group, clustered hierarchical NoCs, consists of a cluster of NoCs [7]–[10]. Modules within a cluster communicate using the lower level cluster NoC, while modules in different clusters can only communicate using the higher levels of the hierarchy. Each cluster can have a different network topology. In contrast, in nonclustered hierarchical NoCs [11]–[15], modules can always connect using the low hierarchy level NoC, and the high levels of the hierarchy only add shorter "express" paths on top of the low level network. The proposed algorithm belongs to the nonclustered family of algorithms. Fallin *et al.* [16] proposed deflection routing in a hierarchical ring topology using bridge routers that have minimal buffering. Our research is different in that its focus is hierarchical deflection routing in the widely used mesh topology, and the use of deflection routing is such that no buffers are required.

Manuscript received September 7, 2015; accepted April 14, 2016. Date of publication April 21, 2016; date of current version May 25, 2016. This manuscript was recommended for publication by A. Gordon-Ross.

The authors are with the School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel (e-mail: gdaliaox@post.tau.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LES.2016.2557353

III. Hierarchical Deflection Routing

One of the reasons for the popularity of the mesh topology is the regular layout of the tiles containing processors and routers. The hierarchical mesh topology preserves the tiling feature, and adds longer wires between routers. We define multiple levels of hierarchy in a $N \times N$ mesh topology. At level 0, each router at coordinates i, j is connected to its neighbors with distance one. On higher levels of the hierarchy we decrease the number of routers along each axis by a factor of s (the step size). Router i, j is part of level l if

$$\Big(i \bmod s^l = 0\Big) \text{ or } \Big(j \bmod s^l = 0\Big). \tag{1}$$

Level l contains $(N/s^l)^2$ routers, in a $(N/s^l) \times (N/s^l)$ configuration. The number of links at level l is given by

$$4\frac{N}{s^l}\left(\frac{N}{s^l} - 1\right). \tag{2}$$

Fig. 1 illustrates an example of a 16×16 hierarchical mesh topology with four levels, and a step size $s = 2$. Using (2), it contains 960, 224, 48, and 8 links at the first, second, third, and fourth levels, respectively. If we denote the link length between neighbor routers by L, level l contains links of length $s^l L$. Therefore in the example, the total wire length overhead is $((960 + 224 * 2 + 48 * 4 + 8 * 8)/960) - 1 = 73.33\%$.

For every two routers A and B at coordinates x_1, y_1 and x_2, y_2 we define their distance as

$$D(A, B) = |x_1 - x_2| + |y_1 - y_2|. \tag{3}$$

The output links of a router are denoted L_i. A link connects a router to a neighbor N_i. $D(N_i, \text{dst})$ is the distance from the neighbor router through that link to the destination. The hierarchical deflection algorithm that selects the shortest distance is summarized in Fig. 2.

1943-0663 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: POLITECHNIKI WARSZAWSKIEJ. Downloaded on March 14,2020 at 22:03:11 UTC from IEEE Xplore. Restrictions apply.

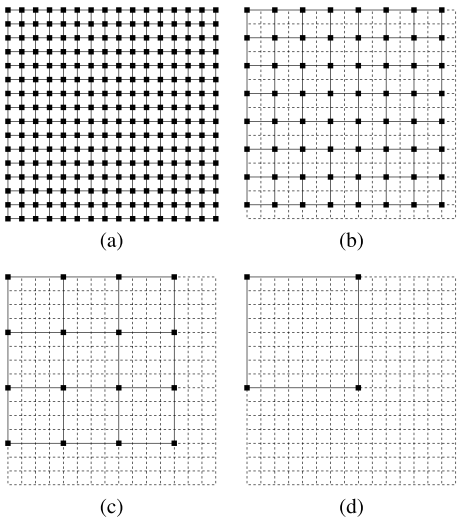


Fig. 1. Hierarchical levels of a 16×16 mesh with step size $s = 2$. The first level contains all the 16×16 routers. The second level consists of 8×8 routers, which are the even numbered routers in both dimensions, skipping over the odd routers. The third level consists of 4×4 routers, skipping four routers in each dimension. The last level contains only four routers, spaced eight routers apart in each axis. (a) Level 0. (b) Level 1. (c) Level 2. (d) Level 3.

1 Mark all the output links as available.
2 Rank the incoming flits in decreasing age order.
3 Iterate over the flits starting from the oldest flit.
4 For each available output link i, calculate its neighbor distance to the destination $D(N_i, dst)$
5 Assign the flit to the available output link with the smallest $D(N_i, dst)$
6 Mark the output link as not available.

Fig. 2. Hierarchical deflection routing algorithm.

We note that the above algorithm can be used for nonhierarchical deflection routing as well. The use of the distance from the neighbor to the destination provides the adjustment to hierarchical routing. Since a hop at level l jumps over s routers in one step, routing over higher hierarchy levels is preferred if they shorten the number of clock cycles to reach the destination; otherwise the level 0 mesh (without hierarchy) is preferred.

IV. Evaluation

We simulate performance using deflection routing network on chip simulator [17], a deflection routing simulator, and power using design space exploration for network tool (DSENT) [18]. We configure DSENT to use 45 nm process, 2 mm interconnect length between adjacent routers, and a 1 GHz operation frequency. We use synthetic uniform random traffic, as well as multithreaded applications from the SPLASH-2 benchmark [19]. The benchmarks are simulated with the SNIPER multicore simulator [20].

For the evaluation we use a hierarchical configuration of a 16×16 hierarchical mesh using 1–4 hierarchy levels. We use bufferless deflection routers, and each clock cycle a single flit

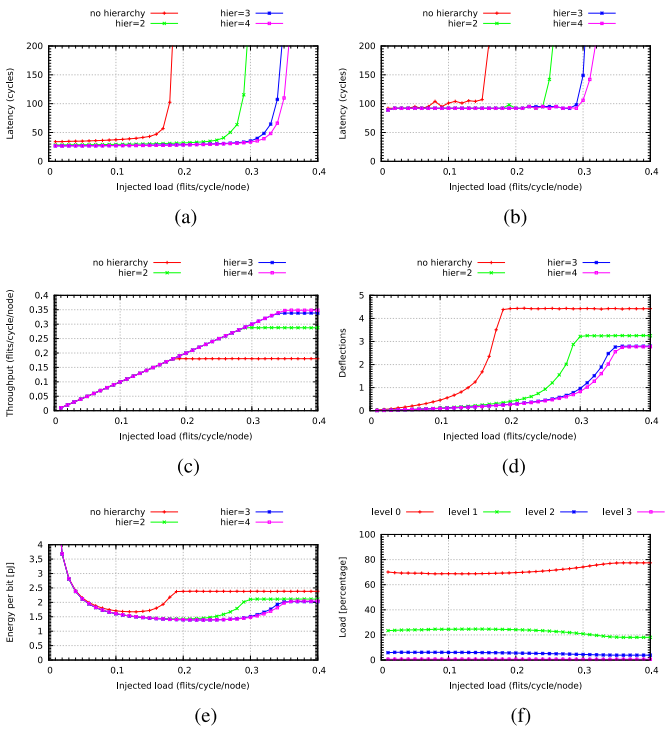


Fig. 3. (a) Average and (b) max latency, (c) throughput, (d) average number of deflections, (e) energy consumption, and (f) load distribution of a hierarchical bufferless deflection router, with varying number of hierarchy levels, for a 16×16 hierarchical mesh as a function of the injected load, for uniform random traffic patterns.

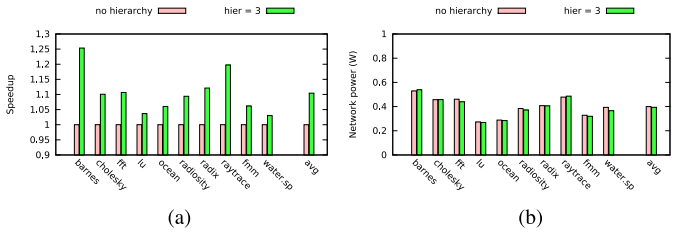
may be presented at each router input. A delay of two clock cycles was simulated for the deflection routers at the lowest hierarchy level, to account for routing computation and crossbar traversal, and an extra clock cycle for the routers at the higher levels of the hierarchy, to account for the slower operation due to the higher crossbar radix. According to [21], in the presence of random process variation for 45 nm technology, the delay of 1 mm long interconnect is between 70 and 140 ps. Given that our clock cycle is 1 ns and the distance between neighbors at level 0 is 2 mm, the link traversal at levels 0 and 1 fit into a single clock cycle, while levels 2 and 3 require two and three clock cycles, respectively. Energy consumption of link traversal is modeled proportional to the link length.

Fig. 3 shows latency, average number of deflections experienced by each flit on its way to destination, as well as the average energy dissipated for each delivered bit, as a function of the injected load for the bufferless deflection router, while varying the number of hierarchy levels from one (no hierarchy) to four levels. The x-axis, injected load, is an independent variable, and is increased during the evaluation beyond the capabilities of the network, to display the behavior near saturation. Adding hierarchy substantially improves the maximum throughput sustained by the network, as seen in Fig. 3(c). Latency is also improved, even at low load, because flits utilize the higher levels of the hierarchy to reach their destinations faster. When the injected load is low, the average energy per bit is high for all the routers due to static power dissipation in a lightly loaded network. As the load increases, energy per bit is reduced due to the more efficient network utilization. Finally, as the network approaches saturation, the average number

Authorized licensed use limited to: POLITECHNIKI WARSZAWSKIEJ. Downloaded on March 14,2020 at 22:03:11 UTC from IEEE Xplore. Restrictions apply.

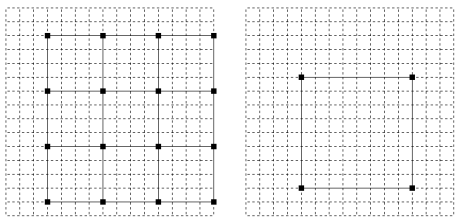
TABLE I
PERFORMANCE AND POWER SUMMARY FOR A 16 × 16 HIERARCHICAL MESH, ON A 45 nm PROCESS AT 1 GHZ OPERATING FREQUENCY, WITH 64-BIT LINKS. THE NUMBERS IN THE PARENTHESES ARE THE INJECTED LOAD (IN FLITS/CYCLE/NODE). LATENCY, POWER, AND ENERGY PER BIT ARE MEASURED BOTH AT AN INJECTED LOAD OF 0.15 AND 0.25

	latency (load 0.15) (cycles)	latency (load 0.25) (cycles)	max throughput (flits/cycle/node)	power (load 0.15) (Watt)	power (load 0.25) (Watt)	energy/bit (load 0.15) (pJ)	energy/bit (load 0.25) (pJ)	area (mm^2)
flat	43.16	N/A	0.180	4.21	N/A	1.72	N/A	3.09
hier 2	30.31	36.71	0.288	3.59	6.21	1.46	1.52	4.23
hier 3	27.94	30.44	0.339	3.56	5.71	1.45	1.40	4.60
hier 4	27.64	29.88	0.348	3.57	5.72	1.45	1.40	4.67
hier 4+interleave	27.89	30.17	0.350	3.55	5.68	1.45	1.39	4.51



(a) (b)

Fig. 4. SPLASH-2 application speedup and network power of 8 × 8 hierarchical deflection routing, compared to nonhierarchical bufferless routing. (a) Speedup. (b) Power.



(a) (b)

Fig. 5. Hierarchical levels 2 and 3 of a 16 × 16 mesh with step size $s = 2$, along with interleaving and center shift. Levels 0 and 1 are not shown since they are identical to those shown in Fig. 1. (a) Level 2. (b) Level 3.

of deflections increases, along with the dissipated network energy.

Fig. 3(f) shows the load distribution of the network among the four hierarchy levels in a 4-level topology. It can be seen that at low loads, use of the higher levels of the hierarchy is higher than at saturation, due to preference of flits to utilize the higher levels rather than the first level to shorten the route to the destination. At saturation, all the links in the network are utilized, and the load distribution equals, for each level, to the ratio of the available links at that level to the total number of links in the NoC.

The normalized speedup and network power of hierarchical deflection routing relative to nonhierarchical bufferless deflection routing for each of the SPLASH-2 benchmarks is shown in Fig. 4(a) and (b), respectively. On the ten benchmarks, the hierarchy improves performance by 10.43%, and reduces network power consumption by 1.52%. While the network's power reduction is marginal, a multicore chip consists of the network and additional components (such as cores and caches), and therefore the faster execution time contributes to the total energy savings by increasing the chip's idle time.

Table I summarizes latency, throughput, power, energy per bit, and area on the 16 × 16 hierarchical mesh, with up to four levels of hierarchy, under uniform random traffic.

V. INTERLEAVING

Although hierarchical deflection routing improves performance over nonhierarchical deflection routing, it increases the crossbar radix for routers that belong to multiple hierarchy levels. For example, the central router at position $(8, 8)$ in the 16 × 16 mesh is incorporated in all hierarchy levels, and therefore has 14 bidirectional connections. High radix crossbars consume more power, and are also slower, hence placing a limitation on the maximum operating frequency of the routers.

One way to reduce the crossbar radix, while still providing the benefits of hierarchical routing, is to use interleaving. In the 16 × 16 four levels example, level 0 contains all 256 routers. Level 1 contains 64 routers that are also included in level 0, and therefore have a radix of up to 8. Level 2 contains 16 routers with radix up to 12, and finally level 3 contains four routers with a radix up to 14. With interleaving, for levels 2 and 3 we use routers from the 192 remaining routers at level 0 that were not used at level 1. A reduction by a factor of two along each dimension creates four sets of 64 routers, where a router at coordinates (i, j) belongs to the set defined by the following equation:

$$\text{set } (m, n) : (i \bmod 2 = m) \text{ and } (j \bmod 2 = n). \tag{4}$$

We can assign the routers that belong to the higher hierarchy levels to different sets. For example, we assign level 1 to set $(0, 0)$, level 2 to set $(0, 1)$, and level 3 to set $(1, 0)$, leaving set $(1, 1)$ not utilized. This way each router is connected to level 0, and only one additional level, limiting the crossbar radix to a maximum of 8. Beginning from level 2, we can add to the higher levels of the grid a shift that is a multiple of the step size in each direction, without changing the assignment of routers to sets. Fig. 5 shows the interleaved topology, with a shift of $2^{\text{level}-1}$.

Fig. 6 shows the latency when adding interleaving and grid shift to the base hierarchical routing described earlier. Also shown for reference is the performance without hierarchy. Although interleaving without shifting reduces the maximum throughput slightly, it is still significantly higher than routing without hierarchy. However, adding grid shift in addition to interleaving slightly surpasses the performance of the base hierarchical routing.

Table II summarizes the results for a 16 × 16 mesh. Interleaving allows us to reduce the number of output links to

Authorized licensed use limited to: POLITECHNIKI WARSZAWSKIEJ. Downloaded on March 14,2020 at 22:03:11 UTC from IEEE Xplore. Restrictions apply.

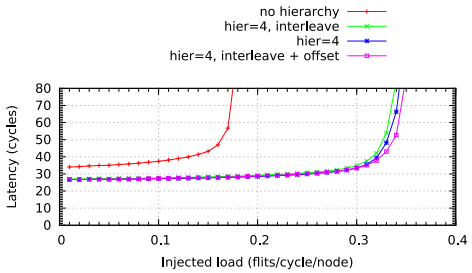


Fig. 6. Latency effects of interleaving and center shift, under uniform random traffic.

TABLE II
SUMMARY OF CROSSBAR POWER AND AREA BENEFITS OF INTERLEAVING

Max crossbar area savings (center router)	61.55%
Max crossbar power savings (center router)	61.87%
Number of routers with more than 8 neighbors	13 (5%)
Total network crossbar power savings	3.47%
Total network crossbar area savings	3.24%

the neighbors of each router. This results in crossbar power, area, and timing savings compared to the baseline topology, while still preserving the performance benefits of hierarchical deflection routing. On a 16×16 mesh, 13 routers out of 256 are affected (5%). The maximum savings occur at the center router, in which the number of neighbor connections is reduced from 14 to 8, the crossbar area is reduced by 61.87%, and power is reduced by 61.55%. Although interleaving substantially reduces a router's crossbar area and power by reducing the number of links the router is connected to, the number of affected routers is small and therefore the overall reduction in power is also small (3.47%). The main benefit of interleaving is to reduce the crossbar radix of the high levels of the hierarchy, simplifying router design, without negatively affecting network performance.

VI. CONCLUSION

We proposed and evaluated hierarchical deflection routing in on-chip mesh topologies. Adding hierarchy levels increases the crossbar size in the routers that participate in higher hierarchy levels. We showed that most of the increase in the crossbar size can be reduced by using two techniques, interleaving and shifting. As in flat mesh networks, hierarchical deflection routing also makes local routing decisions, allowing inexpensive router implementations. We evaluated a 16×16 hierarchical mesh with four levels of hierarchy, using both synthetic and application benchmarks, and demonstrated improvement in throughput, latency, network congestion, and energy consumption.

REFERENCES

[1] C. Feng, Z. Lu, A. Jantsch, M. Zhang, and Z. Xing, "Addressing transient and permanent faults in NoC with efficient fault-tolerant deflection router," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 21, no. 6, pp. 1053–1066, Jun. 2013.

[2] Z. Lu, M. Zhong, and A. Jantsch, "Evaluation of on-chip networks using deflection routing," in *Proc. 16th ACM Great Lakes Symp. VLSI*, Philadelphia, PA, USA, 2006, pp. 296–301.

[3] M. Millberg, E. Nilsson, R. Thid, S. Kumar, and A. Jantsch, "The Nostrum backbone—A communication protocol stack for networks on chip," in *Proc. 17th Int. Conf. VLSI Design*, Mumbai, India, 2004, pp. 693–696.

[4] F. Chaochao, L. Zhonghai, A. Jantsch, and M. Zhang, "A 1-cycle 1.25 GHz bufferless router for 3D network-on-chip," *IEICE Trans. Inf. Syst.*, vol. 95, no. 5, pp. 1519–1522, May 2012.

[5] C. Fallin, C. Craik, and O. Mutlu, "CHIPPER: A low-complexity bufferless deflection router," in *Proc. IEEE 17th Int. Symp. High Perform. Comput. Archit.*, San Antonio, TX, USA, 2011, pp. 144–155.

[6] G. Ravindran and M. Stumm, "A performance comparison of hierarchical ring-and mesh-connected multiprocessor networks," in *Proc. 3rd Int. Symp. High Perform. Comput. Archit.*, San Antonio, TX, USA, 1997, pp. 58–69.

[7] C. Puttmann, J.-C. Niemann, M. Porrmann, and U. Ruckert, "GigaNoC—A hierarchical network-on-chip for scalable chip-multiprocessors," in *Proc. 10th Euromicro Conf. Digit. Syst. Design Archit. Methods Tools (DSD)*, Lübeck, Germany, 2007, pp. 495–502.

[8] R. Das, S. Eachempati, A. K. Mishra, V. Narayanan, and C. R. Das, "Design and evaluation of a hierarchical on-chip interconnect for next-generation CMPs," in *Proc. High Perform. Comput. Archit. (HPCA)*, Raleigh, NC, USA, 2009, pp. 175–186.

[9] D. Matos *et al.*, "Floorplan-aware hierarchical NoC topology with GALS interfaces," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Seoul, South Korea, 2012, pp. 652–655.

[10] X. Leng, N. Xu, F. Dong, and Z. Zhou, "Implementation and simulation of a cluster-based hierarchical NoC architecture for multi-processor SoC," in *Proc. IEEE Int. Symp. Commun. Inf. Technol. (ISCIT)*, vol. 2. Beijing, China, 2005, pp. 1203–1206.

[11] U. Y. Ogras and R. Marculescu, "'It's a small world after all': NoC performance optimization via long-range link insertion," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 14, no. 7, pp. 693–706, Jul. 2006.

[12] M. Winter, S. Prusseit, and P. F. Gerhard, "Hierarchical routing architectures in clustered 2D-mesh networks-on-chip," in *Proc. Int. SoC Design Conf. (ISOCC)*, Seoul, South Korea, 2010, pp. 388–391.

[13] S. Bourduas and Z. Zilic, "Latency reduction of global traffic in wormhole-routed meshes using hierarchical rings for global routing," in *Proc. IEEE Int. Conf. Appl. Spec. Syst. Archit. Process. (ASAP)*, Montreal, QC, Canada, 2007, pp. 302–307.

[14] R. Manevich, I. Cidon, and A. Kolodny, "Handling global traffic in future CMP NoCs," in *Proc. Int. Workshop Syst. Level Interconnect Prediction*, Austin, TX, USA, 2012, pp. 40–47.

[15] R. Manevich, I. Cidon, and A. Kolodny, "Dynamic traffic distribution among hierarchy levels in hierarchical networks-on-chip (NoCs)," in *Proc. 7th IEEE/ACM Int. Symp. Netw. Chip (NoCS)*, Tempe, AZ, USA, 2013, pp. 1–8.

[16] C. Fallin, X. Yu, G. Nazario, and O. Mutlu, "A high-performance hierarchical ring on-chip interconnect with low-cost routers," Comput. Archit. Lab, Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. 2011-007, 2011.

[17] G. Oxman and S. Weiss, "DNOC: An accurate and fast virtual channel and deflection routing network-on-chip simulator," in *Proc. IEEE Int. Symp. Perform. Anal. Syst. Softw. (ISPASS)*, Philadelphia, PA, USA, 2015, pp. 193–202.

[18] C. Sun *et al.*, "DSENT—A tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling," in *Proc. IEEE/ACM 6th Int. Symp. Netw. Chip*, Copenhagen, Denmark, 2012, pp. 201–210.

[19] S. C. Woo, M. Ohara, E. Torrie, J. P. Singh, and A. Gupta, "The SPLASH-2 programs: Characterization and methodological considerations," *ACM SIGARCH Comput. Archit. News*, vol. 23, no. 2, pp. 24–36, May 1995.

[20] T. E. Carlson, W. Heirman, and L. Eeckhout, "SNIPER: Exploring the level of abstraction for scalable and accurate parallel multi-core simulation," in *Proc. Int. Conf. High Perform. Comput. Netw. Stor. Anal.*, Seattle, WA, USA, 2011, pp. 1–12.

[21] I. D. Benito, "Global interconnects in the presence of uncertainty," M.S. thesis, Dept. Elect. Comput. Eng., Univ. Massachusetts Amherst, Amherst, MA, USA, p. 85, 2008.

Authorized licensed use limited to: POLITECHNIKI WARSZAWSKIEJ. Downloaded on March 14,2020 at 22:03:11 UTC from IEEE Xplore. Restrictions apply.