A Network-Flow Based Optimal Sample Preparation Algorithm for Digital Microfluidic Biochips

Trung Anh Dinh[1] Shigeru Yamashita[1] Tsung-Yi Ho[2]

[1]Ritsumeikan University, Japan

[2]National Cheng-Kung University, Taiwan

IEEE/ACM ASP-DAC 2014

Agenda

- Introduction
- Problem Formulation
- Proposed Method
- Experimental Results
- Conclusion

Agenda

Introduction

- Digital Microfluidic Biochips
- Sample Preparation
- Illustrative Example

Problem Formulation

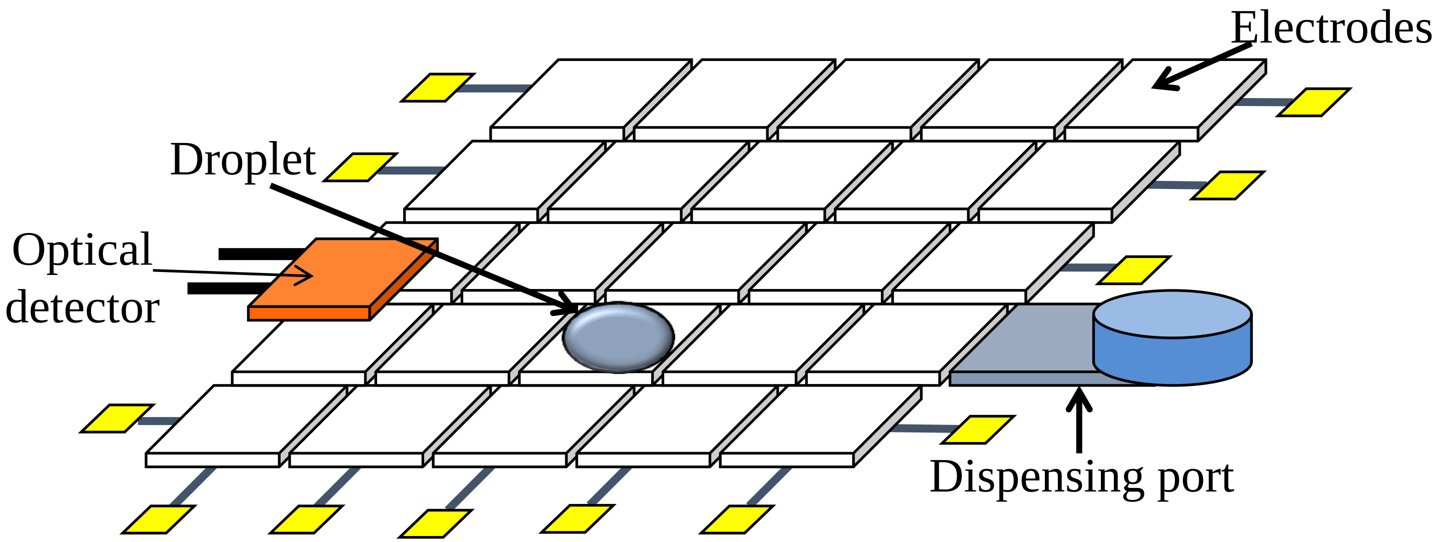
Proposed Method

Experimental Results

Conclusion

Digital Microfluidic Biochips (DMFBs)

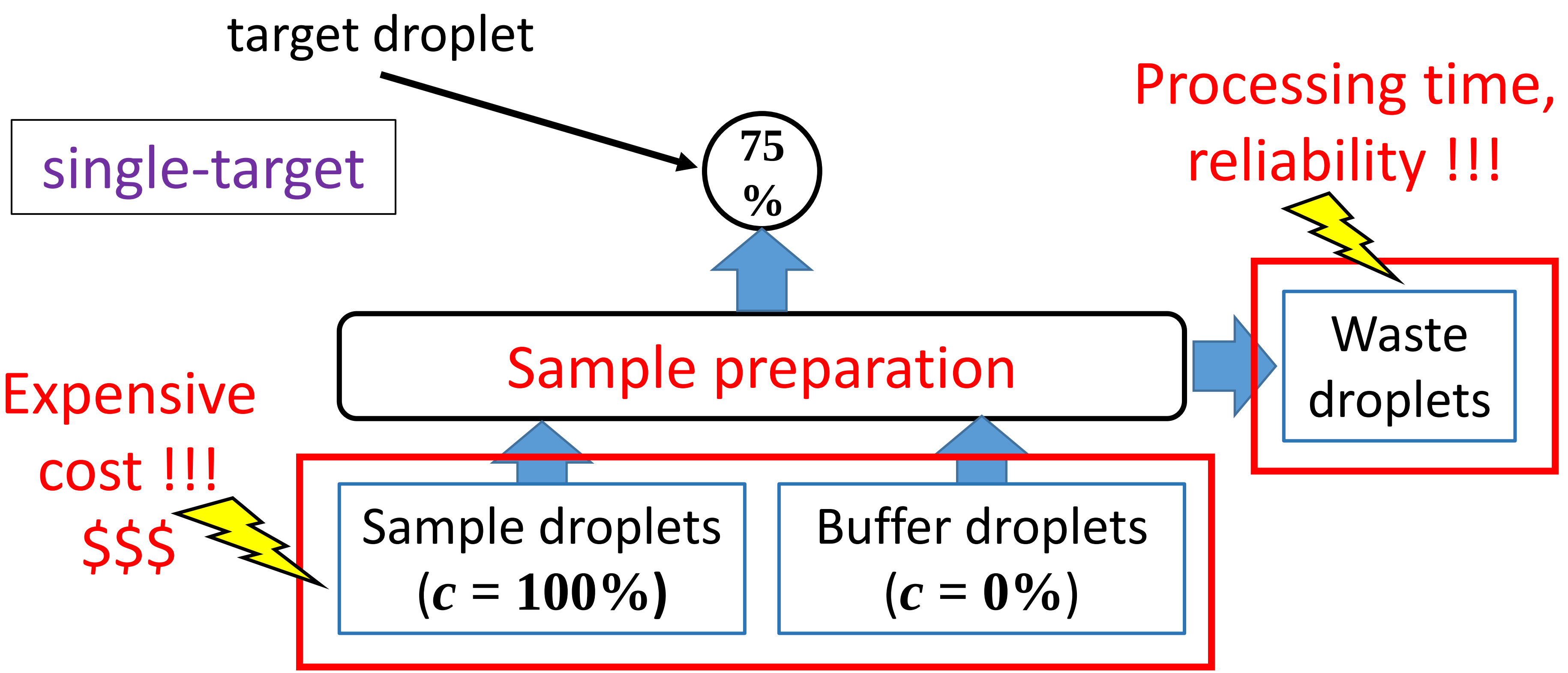
- Architecture of a DMFB:
  - 2D microfluidic array: Basic cells for biological reactions
  - Droplets: Biological samples (picoliter unit)
  - I/O ports, peripheral devices (detector, …)



- Applications: Immunoassay, DNA sequencing, protein crystallization, etc.

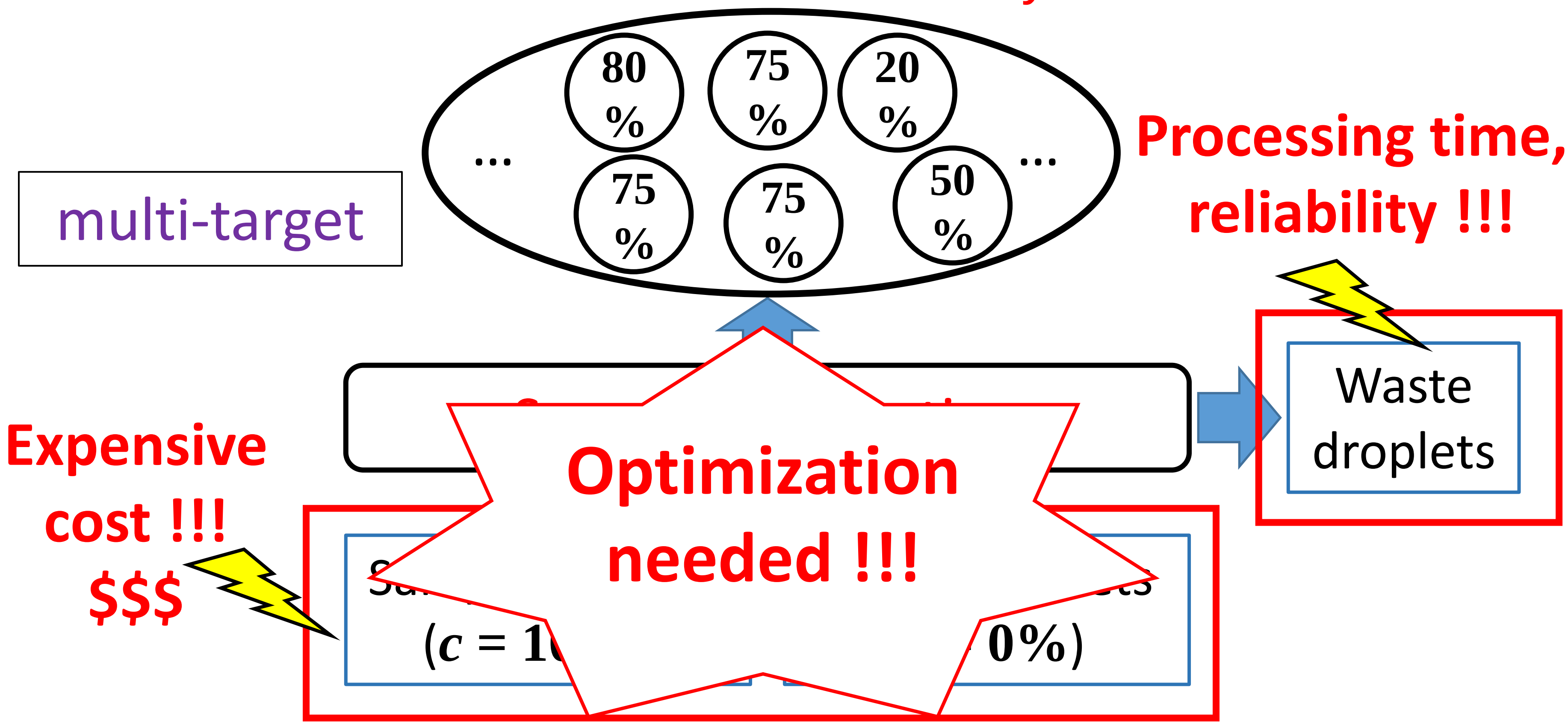
Sample Preparation (1/4)

- To produce droplets of the required concentrations
- A crucial preprocessing step in every application
  - 90% of the cost and 95% of the analysis time



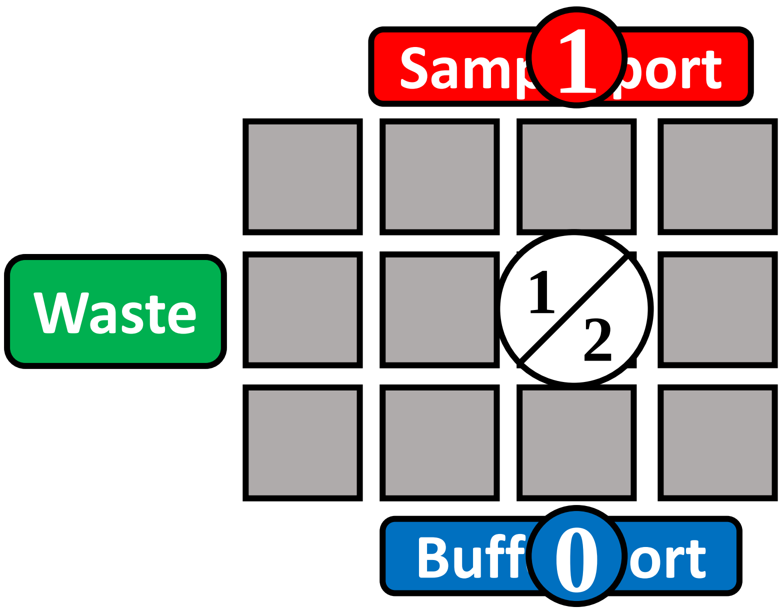
Sample Preparation (2/4)

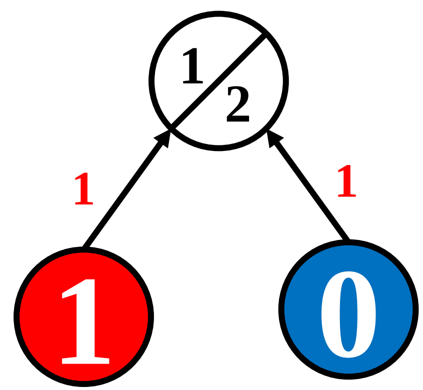
- To produce droplets of the required concentrations
- A crucial preprocessing step in every application
  - 90% of the cost and 95% of the analysis time



Sample Preparation (3/4)

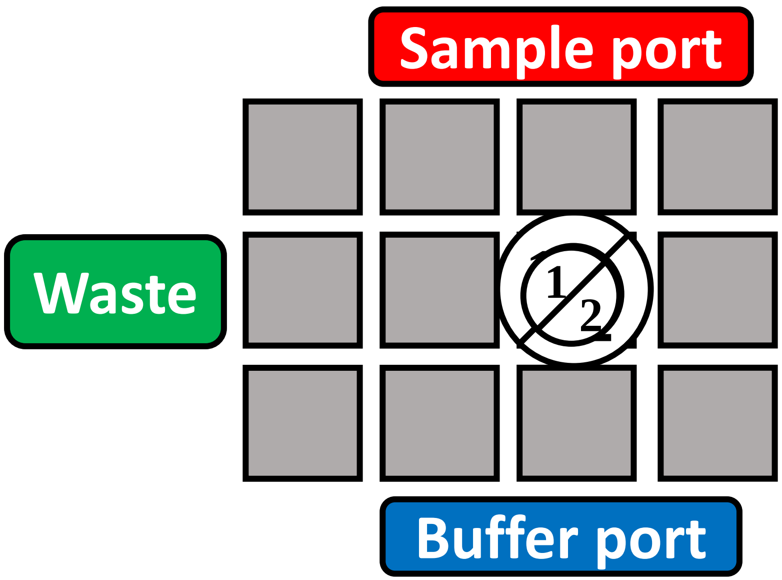
- Dilution method in DMFBs
  - Samples are diluted using 1:1 mixing/splitting ratio
  - E.g., produce a droplet of concentration 25% ($\frac{1}{4}$)

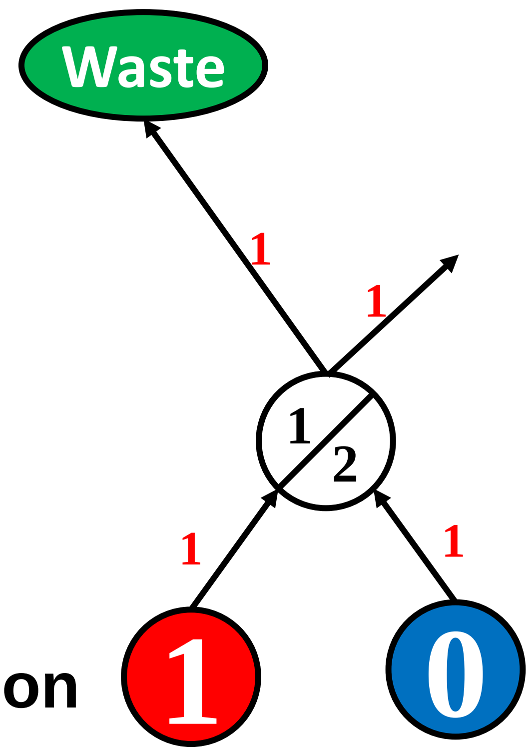




Sample Preparation (3/4)

- Dilution method in DMFBs
  - Samples are diluted using 1:1 mixing/splitting ratio
  - E.g., produce a droplet of concentration 25% ($\frac{1}{4}$)

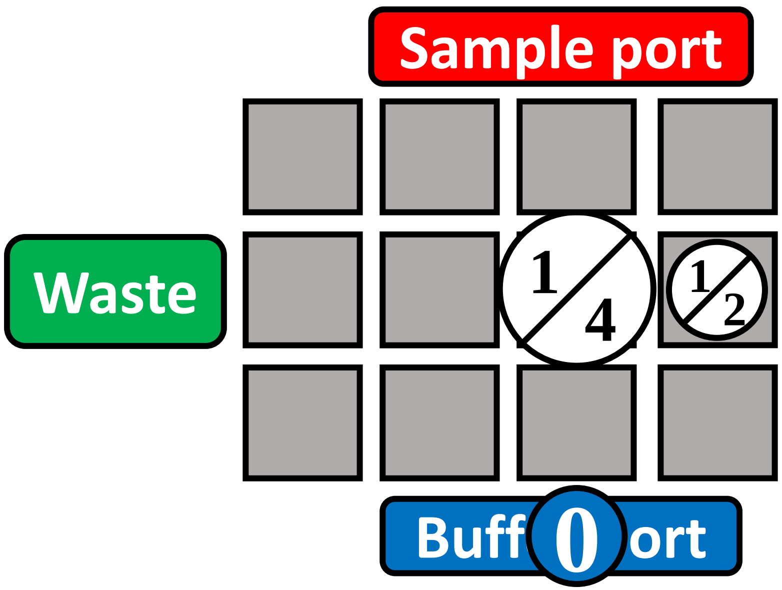


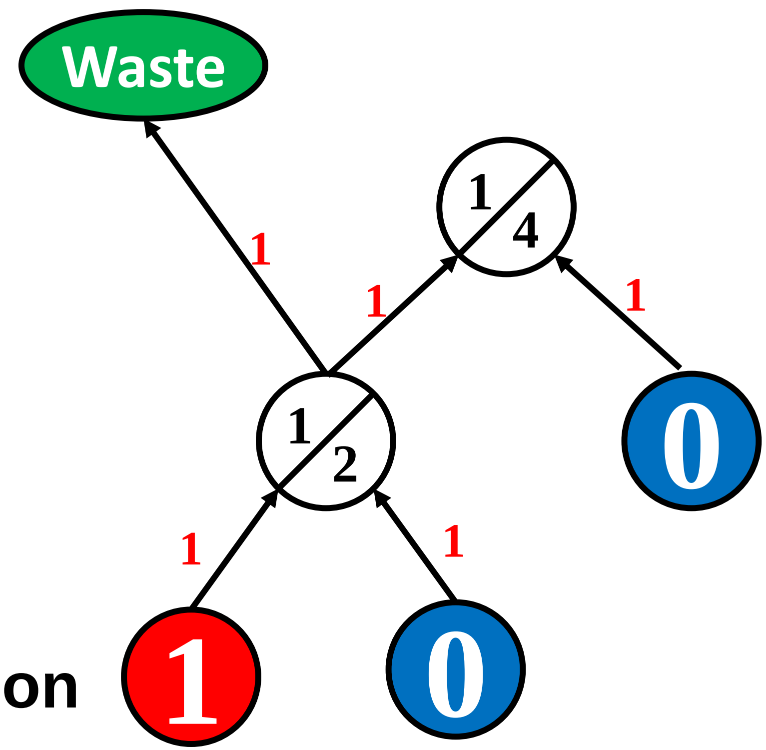


1 sample droplet **1 waste droplet**
1 buffer droplet **1 dilution operation**

Sample Preparation (3/4)

- Dilution method in DMFBs
  - Samples are diluted using 1:1 mixing/splitting ratio
  - E.g., produce a droplet of concentration 25% ($\frac{1}{4}$)



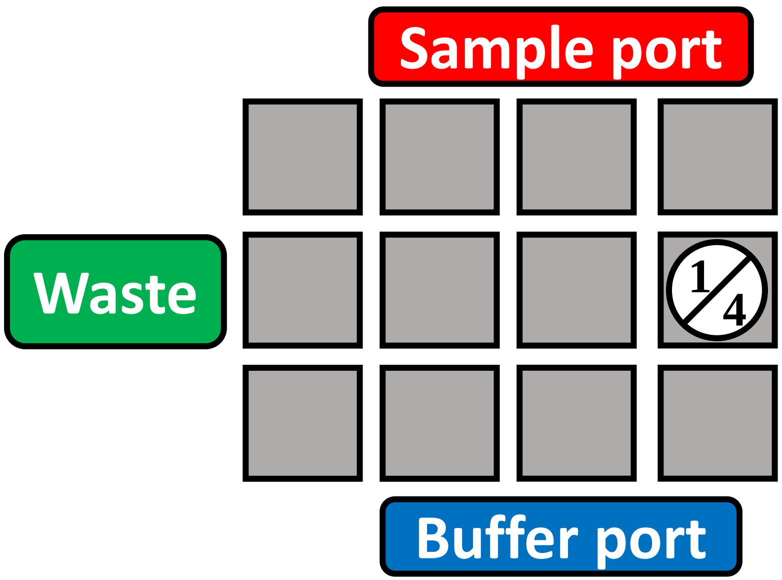


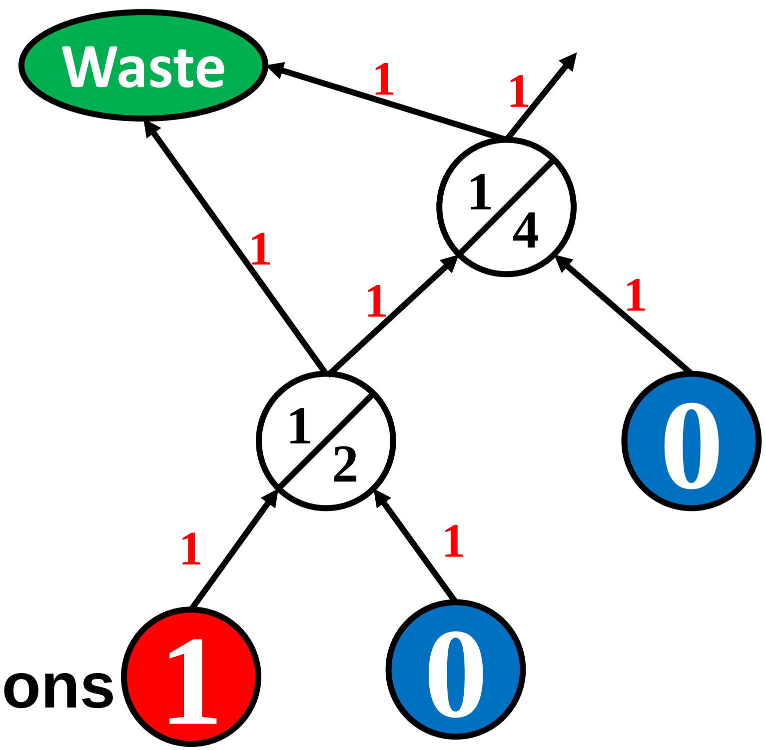
1 sample droplet **1 waste droplet**

2 buffer droplets **1 dilution operation**

Sample Preparation (3/4)

- Dilution method in DMFBs
  - Samples are diluted using 1:1 mixing/splitting ratio
  - E.g., produce a droplet of concentration 25% ($\frac{1}{4}$)

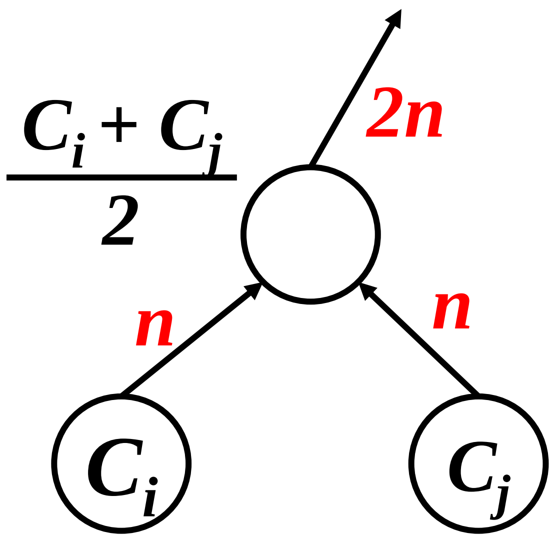




1 sample droplet **1 waste droplet**
2 buffer droplets **2 dilution operations**

Sample Preparation (4/4)

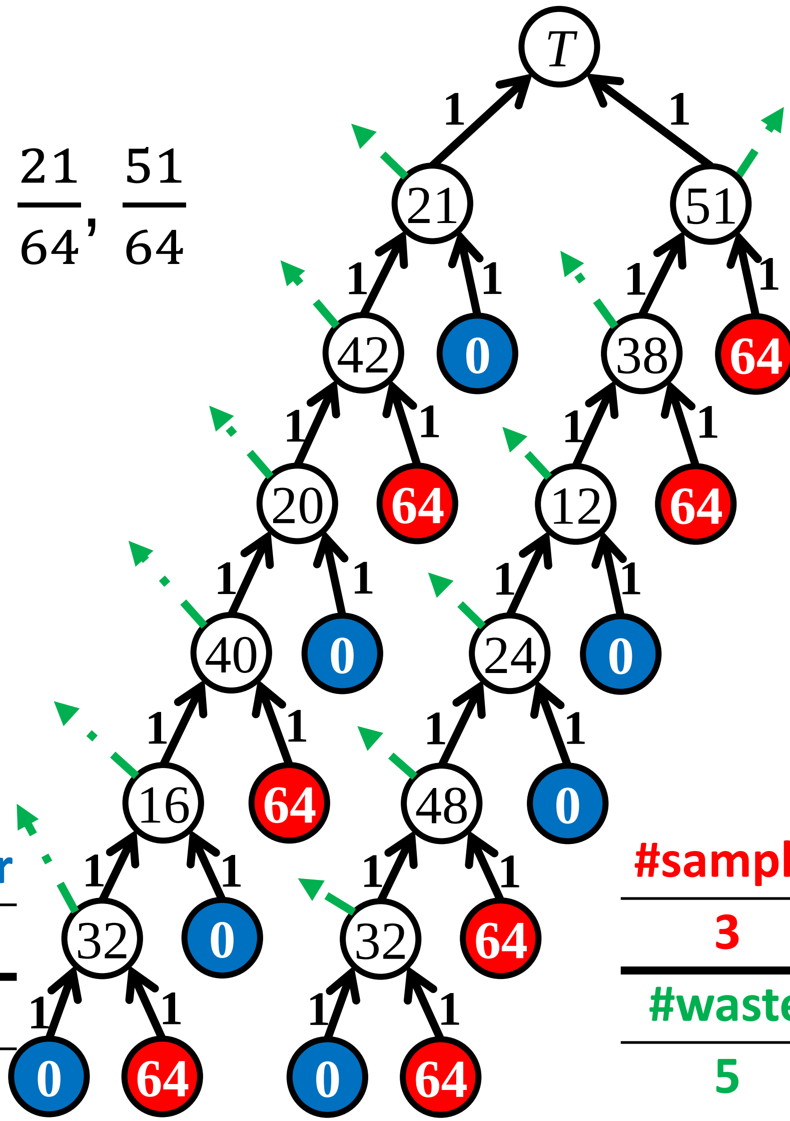
- In general:



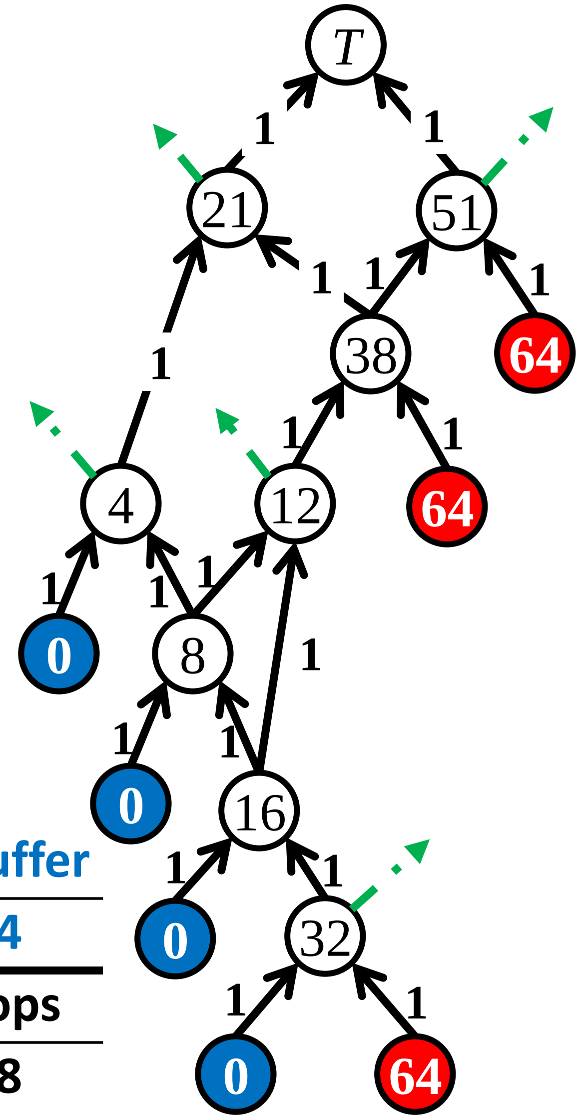
- A concentration value is always expressed by $\frac{c_i}{2^d}$
  - d: precision level of concentration
  - d is given in each problem

Example: Same targets, but different cost

- $d = 6$
- Targets: $\frac{21}{64}, \frac{51}{64}$



#sample	#buffer
7	7
#waste	**#ops**
12	12



#sample	#buffer
3	4
#waste	**#ops**
5	8

Agenda

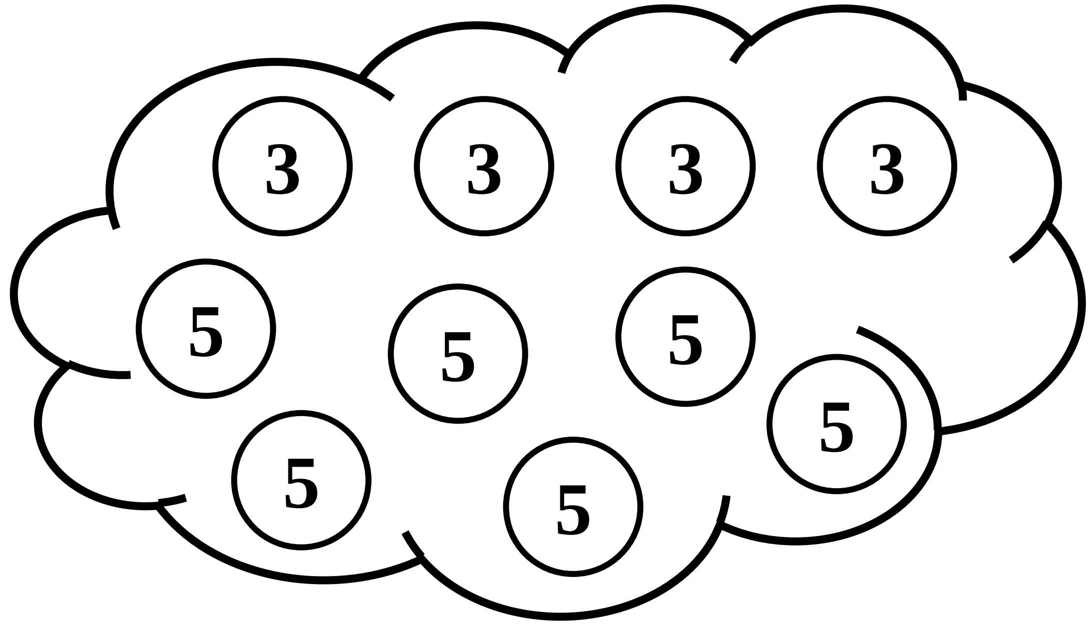
- Introduction
- **Problem Formulation**
  - Problem Formulation
  - Previous Works
- Proposed Method
- Experimental Results
- Conclusion

Problem Formulation (1/2)

- **Given**:
  - Cost of 1 sample droplet: $cost_s$
  - Cost of 1 buffer droplet: $cost_b$
  - Precision level of concentration: d
  - A set of N target concentrations: $TC = \{c_1, c_2, \ldots, c_N\}$
  - A set of the required number of droplets for each target concentration: $S = \{s_1, s_2, \ldots, s_N\}$; $S_R = \sum_{i=1}^{N} s_i$

Problem Formulation (2/2)

- Example:
  - $cost_s = 2$
  - $cost_b = 1$
  - $d = 3$
  - $TC = \{3, 5\}(\frac{3}{8} \& \frac{5}{8})$
  - $S = \{4, 6\}; S_R = 10$



- Output: A valid sample preparation process
- Objective: Minimize cost function

$$\boldsymbol{F = u_s \times 2 + u_b \times 1}$$

Previous Works

- All the previous works are based on heuristics
- All the previous works focus on only one objective optimization

Proposed method: Minimize the cost function

$$F = u_s \times cost_s + u_b \times cost_b$$

- Optimal solution for multiple-target problem
- Flexible to change objective optimization
  - By varying the values of $cost_s$ and $cost_b$
  - E.g., $cost_s = 1$ & $cost_b = 0 \Longrightarrow F = u_s$ (#sample droplets)

(more on this later…)

Agenda

Introduction

Problem Formulation

Proposed Method

Experimental Results

Conclusion

Overview

Input

⇩

Min-cost max-flow (*MCMF*) network model construction

⇩

Integer equal flows problem transformation

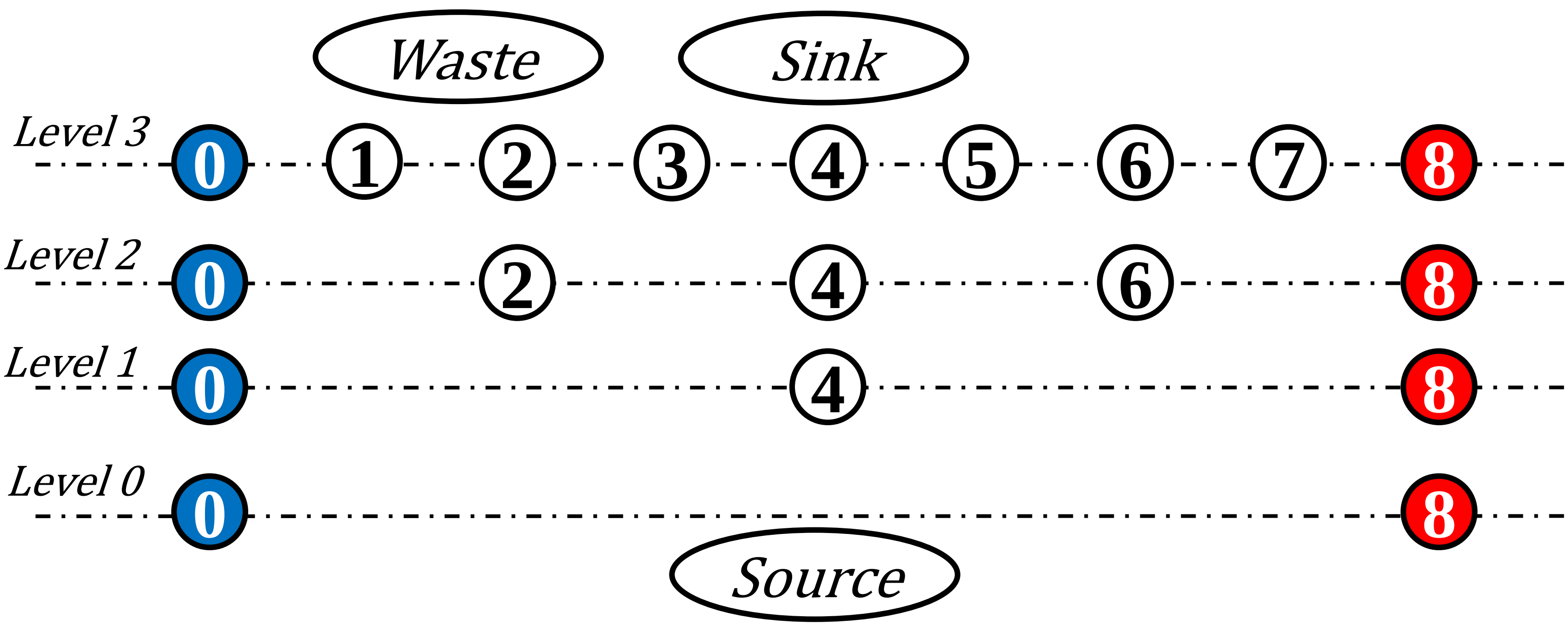
⇩

ILP solver

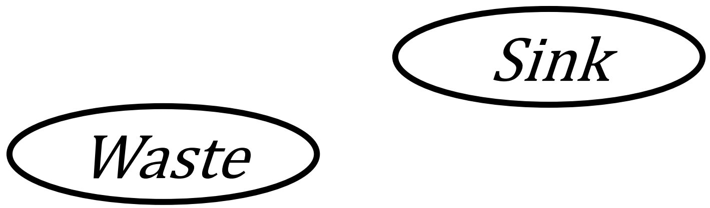
⇩

Output

MCMF Network Model – Set of Vertices V



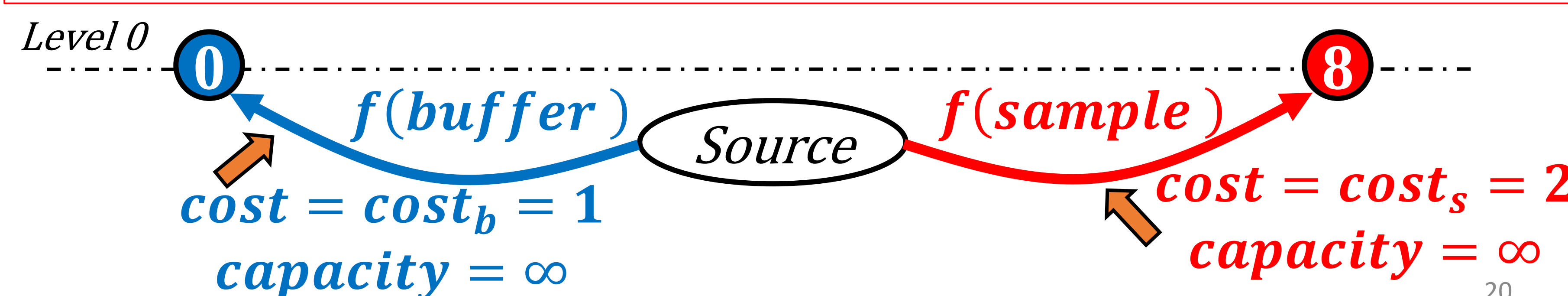
MCMF Network Model – Set of Arcs A



$cost_s = 2$
$cost_b = 1$
$d = 3$
$TC = \{3, 5\}$
$S = \{4, 6\}$

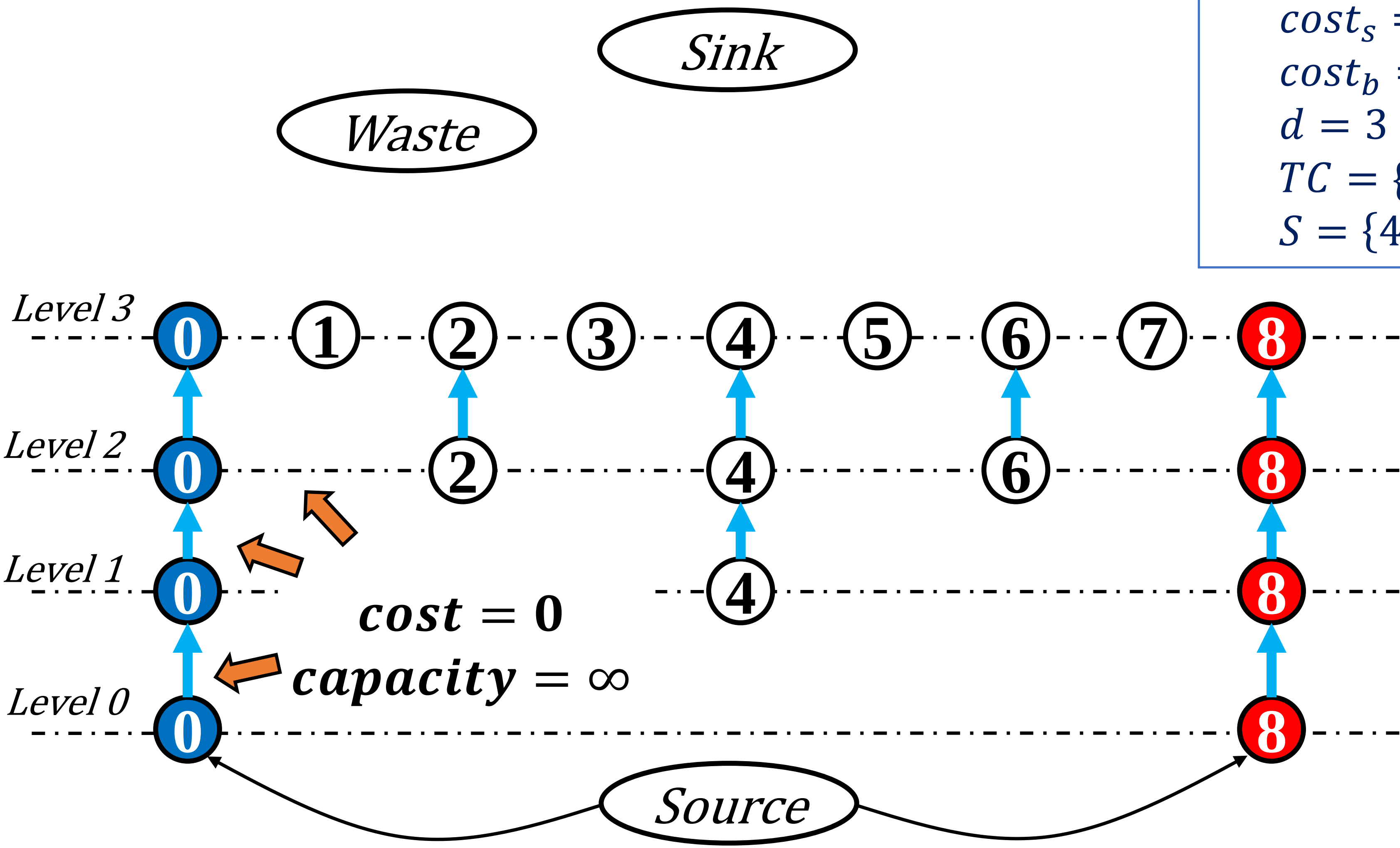
Minimize

$$F = f(buffer) \times cost_b + f(sample) \times cost_s$$



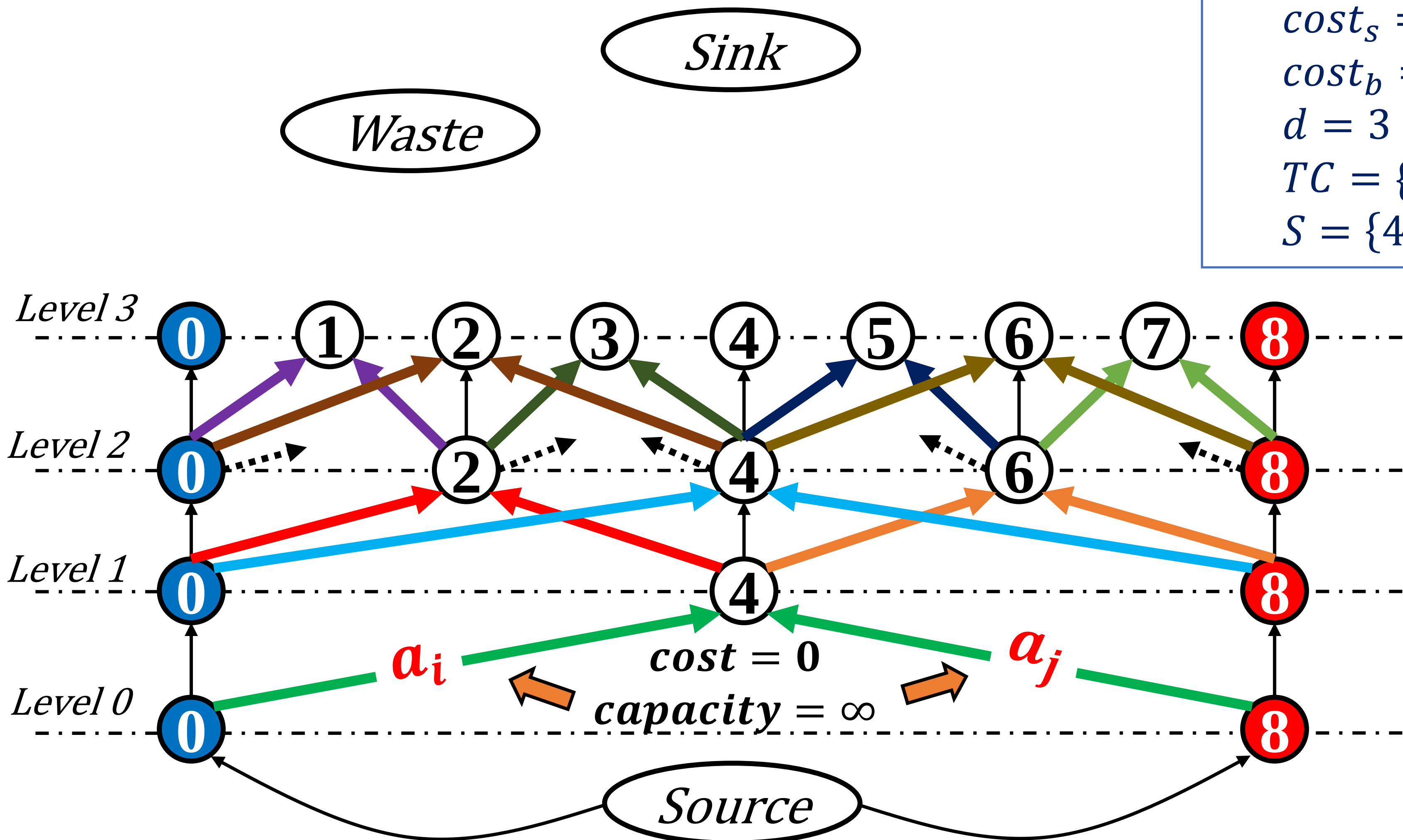
MCMF Network Model – Set of Arcs A

$cost_s = 2$
$cost_b = 1$
$d = 3$
$TC = \{3, 5\}$
$S = \{4, 6\}$



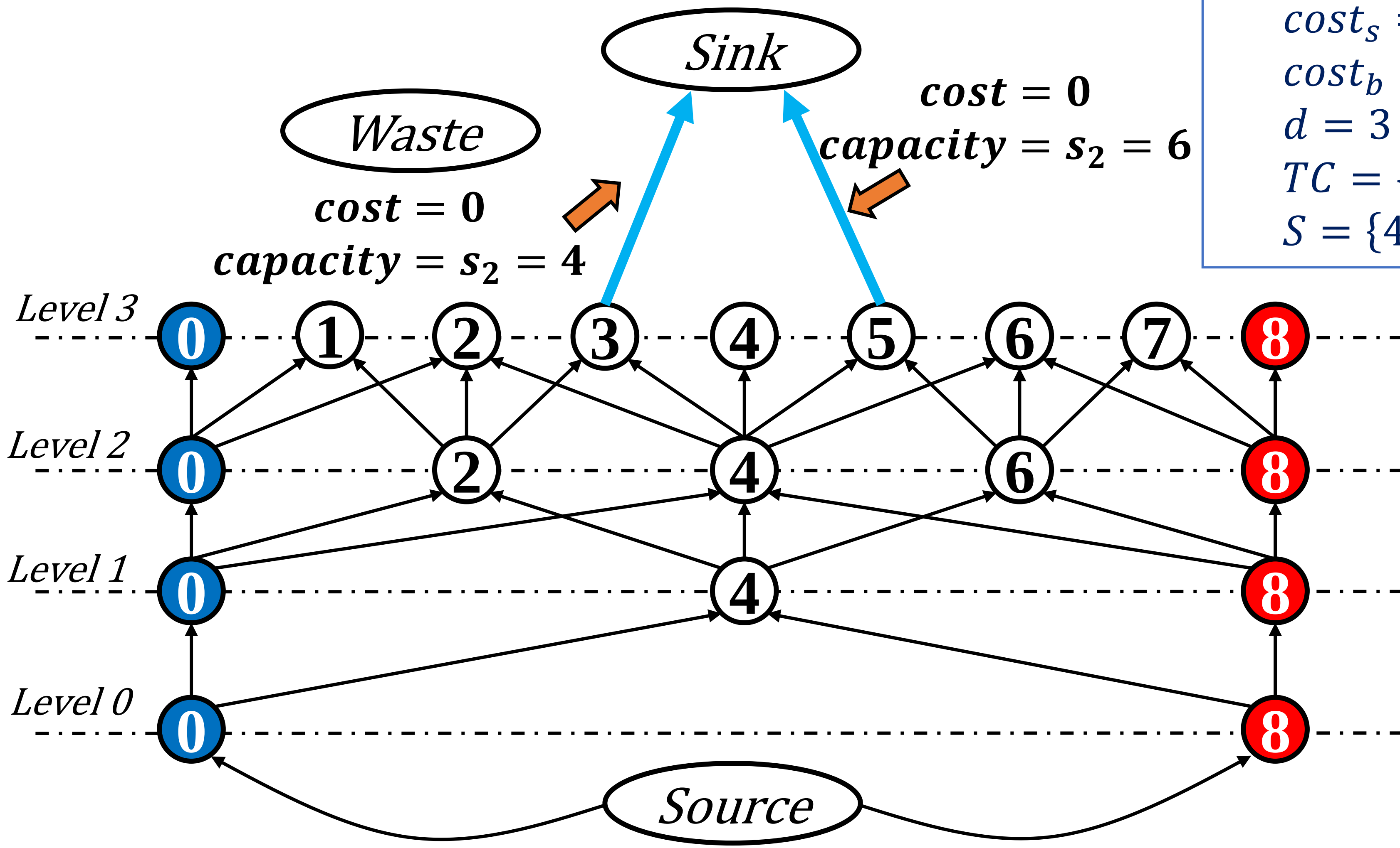
MCMF Network Model – Set of Arcs A

$cost_s = 2$
$cost_b = 1$
$d = 3$
$TC = \{3, 5\}$
$S = \{4, 6\}$



Integer equal flows constraints: $f(a_i) = f(a_j)$

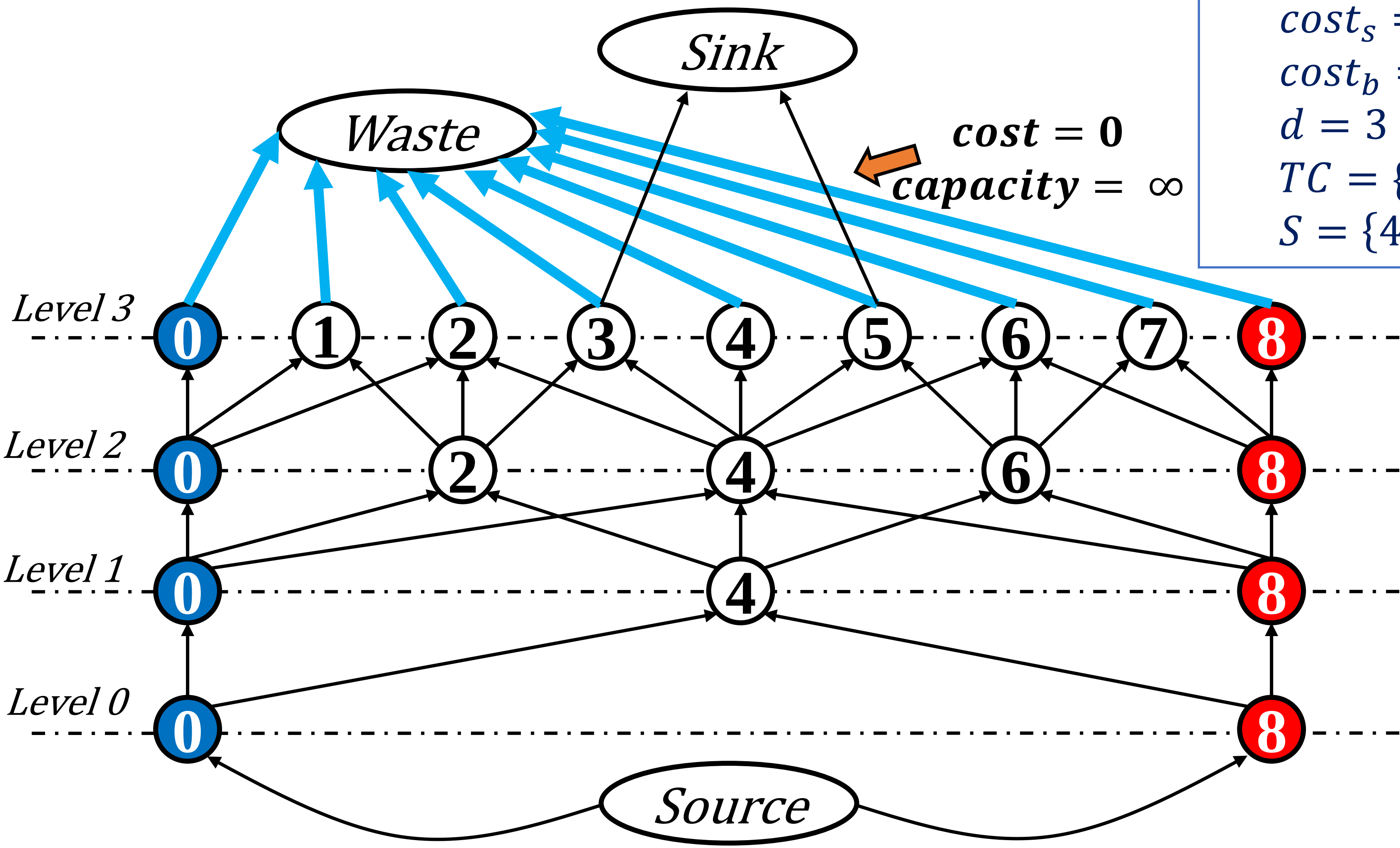
MCMF Network Model – Set of Arcs A



$cost_s = 2$
$cost_b = 1$
$d = 3$
$TC = \{3, 5\}$
$S = \{4, 6\}$

Target concentration constraints: $f(target, Sink) = s_i$

MCMF Network Model – Set of Arcs A



$cost_s = 2$
$cost_b = 1$
$d = 3$
$TC = \{3, 5\}$
$S = \{4, 6\}$

ILP Model

- Minimize

$$F = f(Source, sample) \times cost_s + f(Source, buffer) \times cost_b$$

- Subject to
  - Capacity constraints:

$$f(v_x, v_y) \leq capacity(v_x, v_y) \;\; \forall (v_x, v_y) \in A$$

  - Network-flow conservation:

$$\sum_{v_i:(v_i,v_x)\in A} f(v_i, v_x) = \sum_{v_o:(v_x,v_o)\in A} f(v_x, v_o)$$

  - **Integer equal flow constraints**
  - **Target concentrations constraints**

Agenda

- Introduction
- Problem Formulation
- Proposed Method
- **Experimental Results**
- Conclusion

Comparative Studies

- Single-target sample preparation problem
  - [W. Thies et al., Natural Computing'08] **[BS]**
  - [S. Roy et al., IEEE/ACM DATE'11] **[DMRW]**
  - [J.-D Huang et al., IEEE/ACM ICCAD'12] **[REMIA]**
- Multiple-target sample preparation problem
  - [J.-D Huang et al., IEEE/ACM ICCAD'12] **[REMIA]**

Parameters Settings

- Cost Function

$$F = u_s \times cost_s + u_b \times cost_b$$

- Waste droplets: $[u_s + u_b - S_R]$
  - S_R: The total number of target droplets

$cost_s$	$cost_b$	F	Optimization Objective
1	0	u_s	#sample droplets ($ours_S$)
1	1	$u_s + u_b$	#waste droplets ($ours_W$)

Single-Target Sample Preparation

- ILP solver: CPLEX
- $d = 10$
- Target concentrations: $\frac{1}{1024} \dashrightarrow \frac{1023}{1024}$
  - Take average values of all 1023 cases

	BS	DMRW	REMIA	$ours_S$	$ours_W$
# sample droplets	5.00	3.52	2.41	**2.22**	2.49
# buffer droplets	4.01	3.50	6.09	9.68	**2.50**
# waste droplets	8.01	6.02	7.50	10.90	**3.99**
# operations	8.01	12.52	10.13	15.80	9.85

Multiple-Target Sample Preparation

- $d = 8, 9, 10$ $N = 10, 20, 50, 100$
- For each pair (d, N), generate 100 random test cases

$\boldsymbol{d = 9}$

	$N = 10$			$N = 100$		
	REMIA	$ours_s$	$ours_W$	REMIA	$ours_s$	$ours_W$
# sample droplets	19.59	8.07	8.95	203.98	80.15	80.43
# buffer droplets	31.19	12.67	7.12	277.60	76.44	73.56
# waste droplets	40.78	10.74	6.07	381.58	56.59	53.99
# operations	60.90	43.21	39.69	654.44	276.67	258.84

$\boldsymbol{d = 10}$

	$N = 10$			$N = 100$		
	REMIA	$ours_s$	$ours_W$	REMIA	$ours_s$	$ours_W$
# sample droplets	17.33	11.17	13.96	182.73	101.99	103.41
# buffer droplets	43.35	19.79	11.15	320.25	157.19	97.04
# waste droplets	50.68	20.96	15.11	402.98	159.18	100.45
# operations	89.77	73.12	62.73	720.64	438.36	384.66

Conclusion

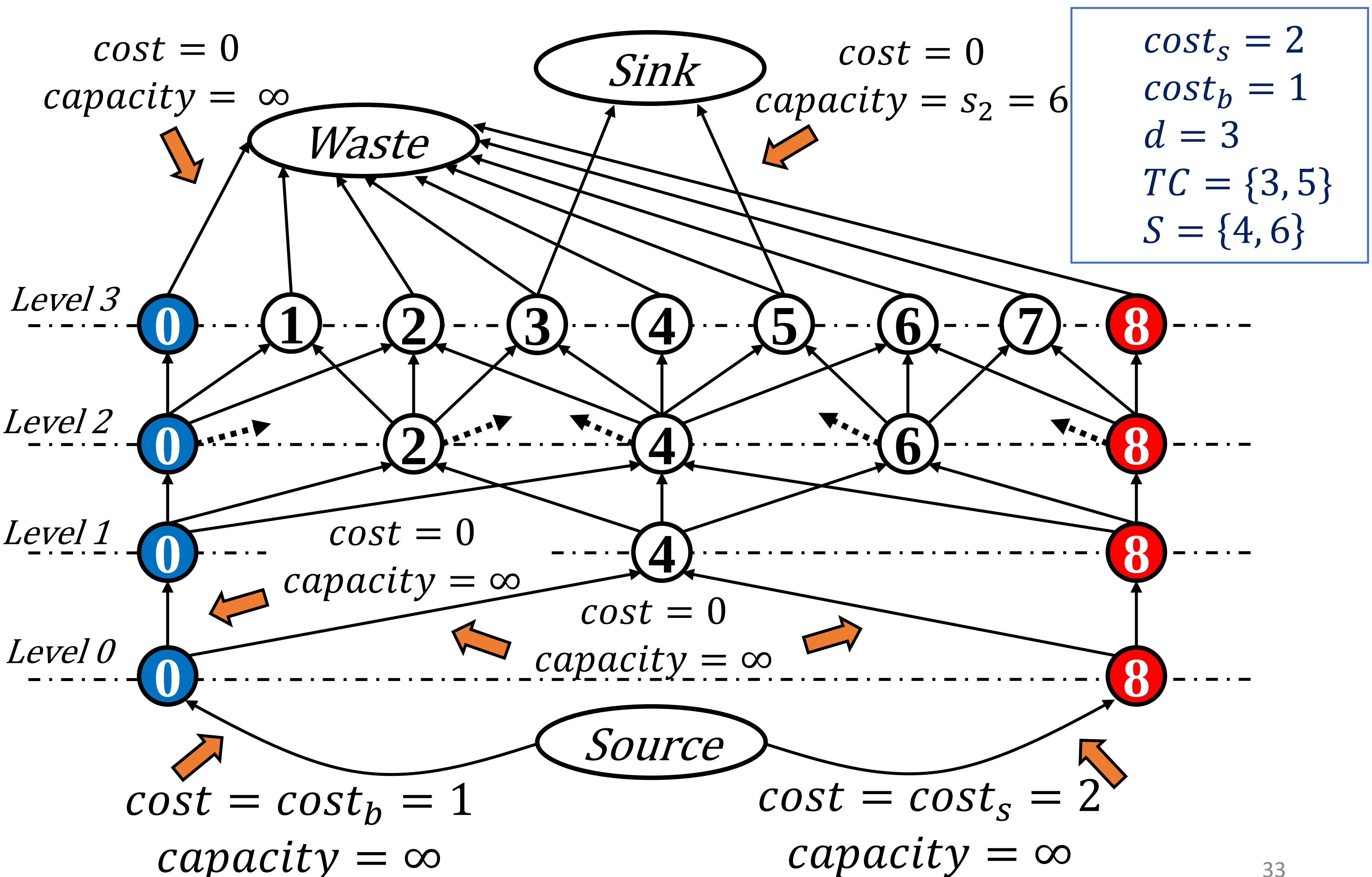
- Sample preparation
  - Pivotal role in every assay, laboratory, and application in biomedical engineering and life science
- The first optimal sample preparation algorithm is proposed
  - Based on a minimum-cost maximum-flow model
  - Reduce the numbers of **sample/buffer/waste** droplets & **dilution operations** significantly

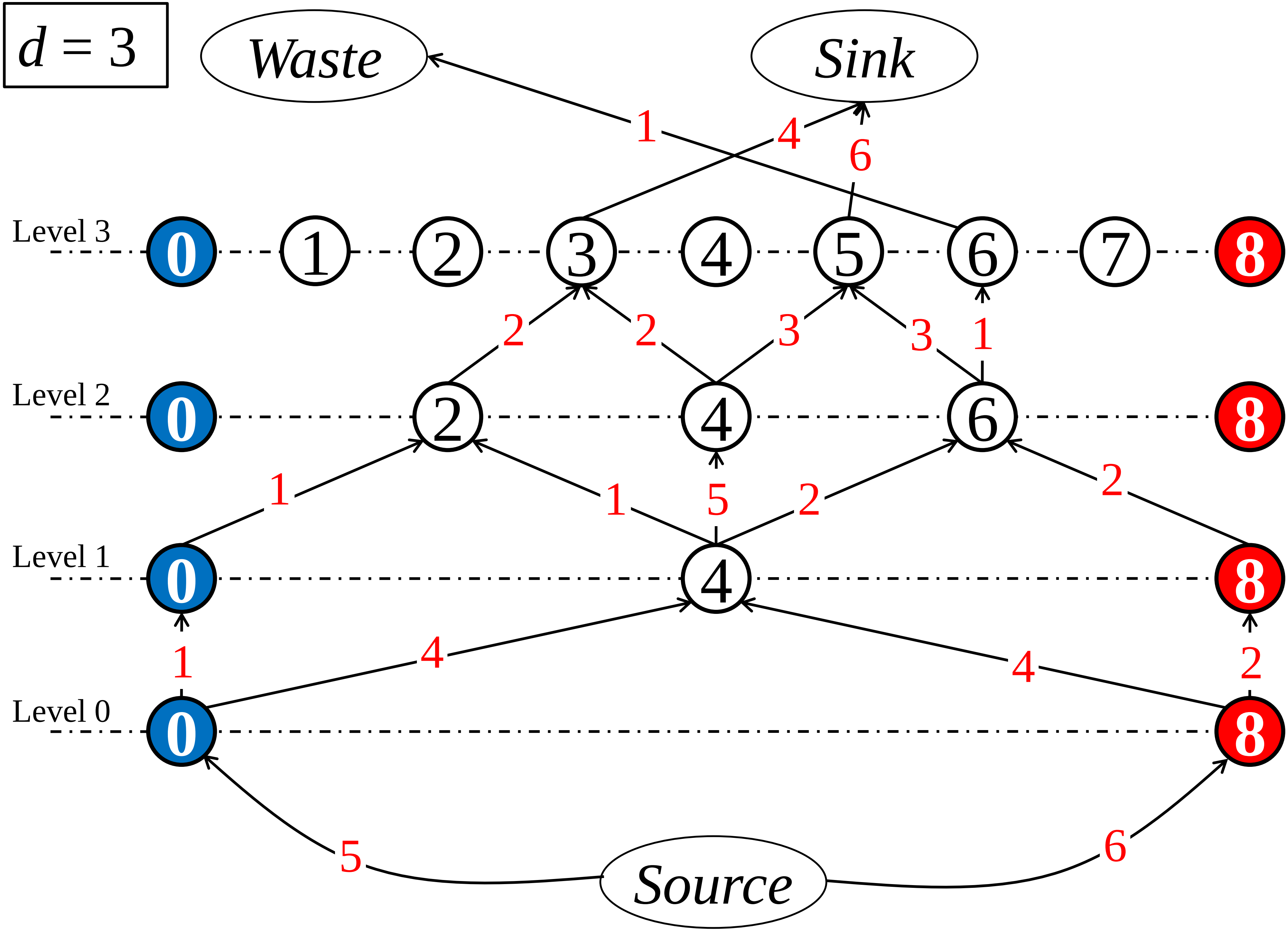
  (~**60%/70%/85% & 60%,** respectively)

Thank you for your attention!!!

Appendix

MCMF Network Model – Set of Arcs A





d = 3
Waste
Sink
Source
Level 3
Level 2
Level 1
Level 0
0
1
2
3
4
5
6
7
8

Problem Formulation

- Hybrid cost function:

$$\boldsymbol{F = u_s \times cost_s + u_b \times cost_b}$$

- The number of waste droplets: $\boldsymbol{u_s + u_b - S_R}$

$cost_s$	$cost_b$	F	Minimization
1	0	$\boldsymbol{u_s}$	#sample droplets
1	1	$\boldsymbol{u_s + u_b}$	#waste droplets
practical values		$\boldsymbol{u_s \times cost_s + u_b \times cost_b}$	practical cost

- The cost function F is *flexible*

Experiment Environment

- Implemented by C++
- ILP Solver: CPLEX
- Linux server
  - Intel® Core(TM) i7 CPU 920 2.67GHz
  - 24 GB Memory

- Largest test case $(d = 10, N = 100)$
  - 352,608 variables
  - 175,374 constraints
  - Computation time: ~30 minutes

Digital Microfluidic Biochips (DMFBs) (2/2)

- Advantages:
  - High portability
  - High throughput
  - Low sample volume consumption
  - Less human intervention errors
- Applications: immunoassay, DNA sequencing, protein crystallization, etc.