

# Robust Sample Preparation on Low-Cost Digital Microfluidic Biochips

Zhanwei Zhong  
Duke University  
Durham, NC, USA  
zz114@duke.edu

Robert Wille  
Johannes Kepler University  
Linz, Austria  
robert.wille@jku.at

Krishnendu Chakrabarty  
Duke University  
Durham, NC, USA  
krish@duke.edu

## ABSTRACT

Sample preparation is an important application for the *digital microfluidic biochips* (DMFBs) platform, and many methods have been developed to reduce the time and reagent usage associated with on-chip sample preparation. However, errors in fluidic operations can result in the concentration of the resulting droplet being outside the calibration range. Current error-recovery methods have the drawback that they need the use of on-chip sensors and further re-execution time. In this paper, we present two dilution-chain structures that can generate a droplet with a desired concentration even if volume variations occur during droplet splitting. Experimental results show the effectiveness of the proposed method compared to previous methods.

## KEYWORDS

Sample preparation, digital microfluidics, robust

## 1 INTRODUCTION

*Digital microfluidic biochips* (DMFBs) are now being used for biochemical applications such as high-throughput DNA sequencing, point-of-care clinical diagnostics, and immunoassays [4]. These devices manipulate liquid as discrete droplets of picoliter volumes based on the principle of electrowetting-on-dielectric under voltage-based electrode actuation [4]. DMFBs have been commercially adopted for sample preparation [22] and infectious disease testing [14]. Techniques have been developed for high-level synthesis [7, 10], module placement [23], and droplet routing [2, 11].

Sample preparation plays an important role in biochemical protocols. According to [5], sample preparation accounts for 90% of the cost and 95% of the analysis time in biochemical experiments. In sample preparation, intermediate droplets are repeatedly split and mixed with reagent and buffer droplets until the desired target concentration is obtained. A physical model underlying droplet mixing and splitting operations is investigated in detail [4]. To date, several sample-preparation algorithms have been developed for DMFBs. The “bit-scanning” method (BS, [21]) guarantees a minimum number of dilution operations and therefore the minimum amount of time. The “dilution and mixing with reduced wastage algorithm” (DMRW, [19]) and the “improved dilution and mixing algorithm” (IDMA, [20]) utilize droplet-sharing techniques to reduce the overall droplet usage. On the other hand, the “reactant minimization algorithm” (REMI, [12]) constructs a dilution forest that minimizes the reactant droplet usage. The work in [3] presents an optimal dilution tree based on a minimum-cost maximum-flow model. Note that all of the above algorithms use the (1:1) mixing model (i.e., two droplets of equal size are mixed). The algorithm in [1] extends the (1:1) mixing model to the (1:2) mixing model, but it requires dedicated mixers, which limits the benefits of reconfigurability that can be derived from DMFBs.

However, because of the inherent uncertainty associated with fluidic operations, errors may occur even if the chip is fully tested [13]. It has been reported that 80% of errors in sample preparation can be attributed to volume variations in droplets [18]. These errors are mostly caused by errors in splitting operations. If the accumulated error is large enough to make the resulting concentration fall outside the calibrated range, the product droplet is deemed to be erroneous. In order to address that, error-recovery methods have been developed to ensure correct outcomes in on-chip sample preparation [13]. However, there are two scenarios where we can not apply such an error-recovery method: (1) For low-cost DMFBs without integrated sensors [8], the bioassay runs in an open-loop manner with no sensor feedback; (2) error-recovery methods proposed thus far require large recovery times, which can lead to outcome-droplet invalidation. For example, fibronectin is known to degrade within 10 s [6]; therefore, protocol re-execution is not feasible for this reagent. These two scenarios motivate the need for a *robust* sample preparation method for DMFBs. Such a method must be capable of achieving the desired target concentrations without the need of sensors and re-executions.

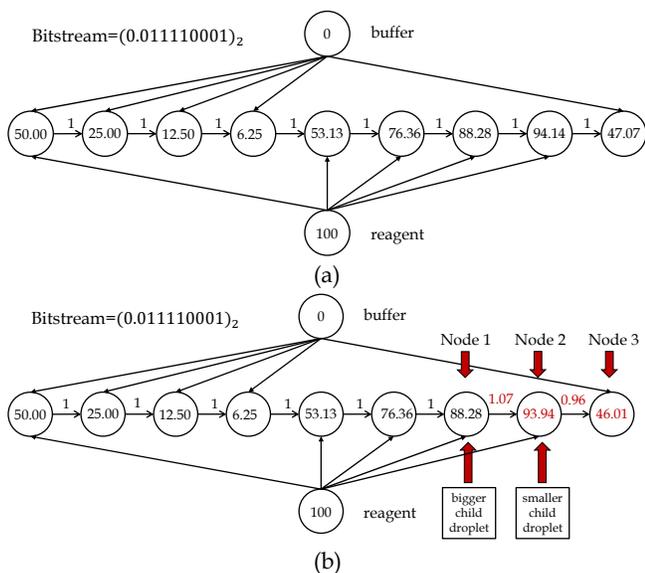
In this paper, we present a sample preparation method that meets these objectives. In contrast to previously proposed sample preparation methods, we do not attempt to address this problem through re-executions (requiring sensors and increasing the execution time), but we deal with volume variations directly during the sample preparation protocol. This is accomplished through the use of two dilution-chain structures that can generate a droplet with concentration in a narrow range even if there are volume variations during droplet splitting. Based on these structures, we present an *integer linear programming* (ILP) method to generate an optimal dilution tree (with respect to an amortized droplet cost). This approach leads to droplets with the target concentrations even if volume variations occur during the process. Experimental results show the applicability and effectiveness of the proposed method.

The remainder of the paper is organized as follows. Section 2 introduces a motivational example and the proposed dilution-chain structures. Section 3 presents the problem formulation for sample preparation. Section 4 describes the method used to construct a dilution tree based on the two types of dilution-chain structures. Section 5 presents experimental results and comparisons with previous methods. Finally, Section 6 concludes the paper.

## 2 ERRORS IN SAMPLE PREPARATION

In this section, we show how volume variations in splitting operations leads to a droplet with an erroneous concentration in a sample preparation bioassay.

We use the BS method from [21] as an example. A target concentration is obtained by first representing the desired concentration as a binary string. Then, a dilution graph is constructed by scanning the binary string from right to left.



**Figure 1: Illustration of the bit-scanning method (a) without volume variations; (b) with volume variations.**

If a bit “1” is encountered, the intermediate droplet is mixed with a 1X reagent droplet (100% concentration). Otherwise, the intermediate droplet is mixed with a 1X buffer droplet (0% concentration).

**Example 1.** Suppose the desired concentration is  $241/512$  (i.e., 47.07%) and the resolution is  $1/512$ . Here, resolution is the difference between two neighbouring concentrations (e.g., concentrations  $241/512$  and  $242/512$ ). In order to achieve this concentration in the BS method, we first represent the desired concentration in terms of a binary string, i.e.,  $(0.011110001)_2$ . Based on that, a dilution graph is constructed as shown in Fig. 1(a). Here, each node represents a so-called mix-split operation, i.e. two 1X droplets are mixed to form a 2X droplet and, afterwards, the resulting (mixed) droplet is split into two child droplets of 1X again. This way, a sequence of mix-split operations is constructed that starts with a mix-split operation of a reagent and a buffer droplet (yielding an intermediate droplet with concentration of 50% as shown in the left-hand side of Fig. 1(a)). Afterwards, further mix-split operations are employed which mix the resulting intermediate droplet with another buffer (reagent) droplet (if the corresponding position in the binary string occupies a “0” (“1”)).

However, volume variations can occur in each droplet splitting operation. For example, if we split a parent droplet with volume  $V_p$  times the unit volume, it is likely that no droplets of the same size will result, but that two child droplets with volumes of  $V_p \times 0.5(1+x)$  units and  $V_p \times 0.5(1-x)$  units will be generated, respectively. Here, parameter  $x$  is defined as the volume variation of the resulting child droplets. According to [18], the volume variation magnitude  $x$  is within the range  $[-0.07, 0.07]$ .

**Example 2.** Consider again the example from above and assume that two volume errors (with  $x = 0.07$ ) occur in Node 1 and Node 2 (as illustrated in Fig. 1(b)). Node 1 will generate two child droplets (with 88.28% concentration) with

1.07X and 0.93X volumes. If the larger of the two droplet is selected for the next operation (i.e. for conducting the mix-split operations in Node 2), the concentration  $C(2)$  of the intermediate droplet in Node 2 is:

$$C(2) = \left( \frac{88.28 \times 1.07 + 100 \times 1}{1.07 + 1} \right) \times 100\% = 93.94\% \quad (1)$$

The concentration variation is obvious here since a concentration of 94.14% is expected (cf. Fig. 1(a)). Furthermore, the resulting droplet is of volume 2.07X (rather than 2X). Since there is also a 7% volume variation in Node 2, the volume of the two child droplets after splitting is  $2.07 \times 0.5 \times 0.93 = 0.96X$  and  $2.07 \times 0.5 \times 1.07 = 1.11X$ , respectively. In this case, if the smaller droplet is selected and it is mixed with a 1 unit-volume buffer droplet in Node 3, the concentration  $C(3)$  of the mixed droplet in Node 3 is given by:

$$C(3) = \left( \frac{93.94 \times 0.96 + 0 \times 1}{0.96 + 1} \right) \times 100\% = 46.01\% \quad (2)$$

This is clearly off from the expected concentration which was given to be 47.07%. Even considering the acceptable resolution (which, as stated above, is  $1/512 = 0.20\%$  and leads to a range of  $[46.97\%, 47.17\%]$ ), this clearly would lead to a failing sample preparation. ■

It is obvious that, if volume variations occur in the execution of a sample-preparation protocol, the concentration of the resulting droplet is affected, and the sample preparation is likely to fail. Therefore, we need to devise a special dilution-chain structure that is able to generate the correct concentration even in the presence of volume variations.

### 3 OVERVIEW OF PROPOSED SOLUTION

In this section, we introduce the creation of a dilution graph that addresses the problem described above. The general idea rests on not using only one of the droplets generated by a mix-split operation, but both of them. Moreover, by using two dedicated dilution-chain structures, we ensure that volume variations will eventually be cancelled out to a large extent. By this, the resulting concentration variation gets minimized, and we can execute a sample preparation protocol that is robust to volume variations.

#### 3.1 Two Dilution-Chain Structures

Fig. 2(a) and Fig. 2(b) show the proposed dedicated dilution-chain structures (which we refer to as *Type-A* and *Type-B*, respectively). The two leftmost nodes represent the dispensing of buffer droplets (0% concentration) and reagent droplets (100% concentration), respectively. The other nodes represent mix-split operations. Such a structure offers the benefit that both child droplets are used by a subsequent node. With a 1X reagent droplet and two 1X buffer droplets, concentrations of [50%, 25%, 37.5%, 31.25%, 34.375%, ...] can be generated by the *Type-A* dilution-chain structure. Analogously, with two 1X reagent droplets and one 1X buffer droplet, concentrations of [50%, 75%, 62.5%, 68.75%, 65.625%, ...] can be generated by the *Type-B* dilution-chain structure.

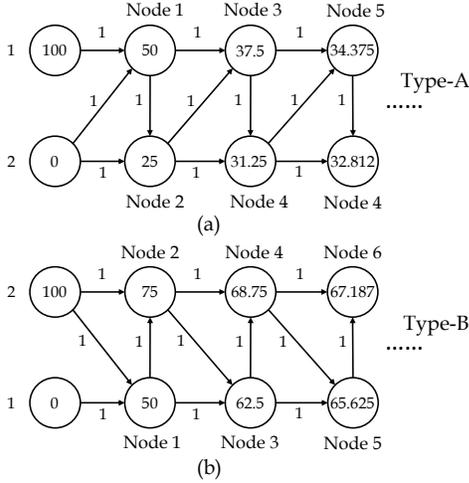


Figure 2: Proposed dilution-chain structures: (a) Type-A; (b) Type-B.

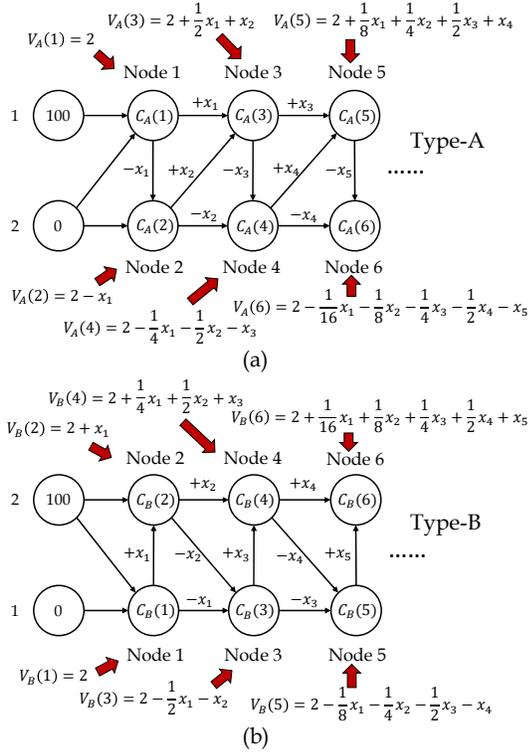


Figure 3: Illustration of the (a) Type-A, and (b) Type-B dilution-chain structures with volume variations.

### 3.2 Stability of Volumes and Concentrations

Fig. 3 illustrates the Type-A and Type-B dilution-chain structures with volume variations. The volume variation at Node  $i$  of a dilution-chain structure is referred to as  $x_i$ . If the volume of the parent droplet at Node  $i$  is  $V_i$ , the volumes of its two child droplets are  $V_i \times \frac{1}{2}(1+x_i)$  and  $V_i \times \frac{1}{2}(1-x_i)$ , respectively. According to [4], the value range of  $x_i$  is  $[-0.07, 0.07]$ . For example, in Fig. 3(a), the volume of the parent

Table 1: Lower and upper bound gaps (percentage values).

	Type-A Structure		Type-B Structure	
	$\Delta_{lo}^A$	$\Delta_{up}^A$	$\Delta_{lo}^B$	$\Delta_{up}^B$
Node 1	0.9067%	0.0000%	0.8454%	0.0000%
Node 2	0.4227%	0.2149%	0.4533%	0.2226%
Node 3	0.4003%	0.1113%	0.3667%	0.1074%
Node 4	0.1833%	0.1204%	0.2001%	0.1257%
Node 5	0.1397%	0.0628%	0.1274%	0.0602%
Node 6	0.0388%	0.0476%	0.0397%	0.0498%
Node 7	0.0291%	0.0176%	0.0285%	0.0174%
Node 8	0.0092%	0.0132%	0.0093%	0.0133%
Node 9	0.0069%	0.0045%	0.0069%	0.0045%

droplet at Node 1 is 2, and the volume variation at this node is  $x_1$ . According to the notation in Fig. 3(a), a child droplet with volume  $(1+x_1)$  is associated with Node 3 while the other with volume  $(1-x_1)$  is associated with Node 2. Note that because the value of  $x_i$  can be either a positive or a negative number, the weights on the edges in Fig. 3 can be either a positive or a negative value as well. For example, if we want the value of  $-x_1$  in Fig. 3(a) to be positive, we just need to assign a negative value to  $x_1$ .

The volume  $V_A(i)$  of the mixed droplet in Node  $i$  in the Type-A dilution-chain structure is given by:

$$V_A(i) = 2 + (-1)^{(i-1)} \times \sum_{j=1}^{i-1} \left(\frac{1}{2}\right)^{(i-1-j)} x_j \quad (3)$$

Likewise, the volume  $V_B(i)$  of the mixed droplet in Node  $i$  in the Type-B dilution-chain structure is given by:

$$V_B(i) = 2 + (-1)^i \times \sum_{j=1}^{i-1} \left(\frac{1}{2}\right)^{(i-1-j)} x_j \quad (4)$$

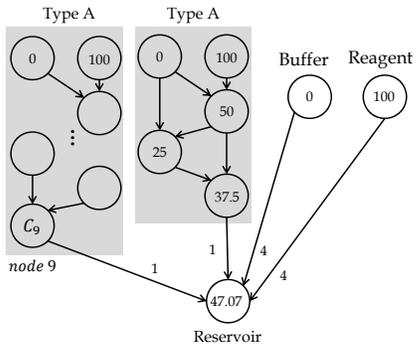
Because the second terms in Equation (3) and Equation (4) are geometric sequences, we obtain identical value ranges for  $V_A(i)$  and  $V_B(i)$  (note that the range of  $x_j$  is  $[-0.07, 0.07]$ ):

$$1.86 \leq V_A(i), V_B(i) \leq 2.14 \quad (5)$$

The concentration of the mixed droplet in each node can be expressed as a function of the volume variations  $x_1, x_2, \dots, x_i$ . The expressions are omitted here for the sake of conciseness, but we can apply numerical methods to obtain upper and lower bounds on these concentrations.

**Example 3.** Suppose the desired concentration for Node  $i$  is  $C(i)$ . Let the lower and upper bounds on this concentration be  $C_{lo}(i)$  and  $C_{up}(i)$ , respectively. Based on this notation, we define the lower bound gap as  $\Delta_{lo}(i) = C(i) - C_{lo}(i)$ , and define the upper bound gap as  $\Delta_{up}(i) = C_{up}(i) - C(i)$ . Table 1 shows these gaps for Node 1 to Node 9 in the two dilution-chain structures. It is obvious by inspection that both bound gaps decrease with an increase in the number of nodes. In other words, as we go deeper in the dilution-chain structure, the lower and upper bounds on the concentration gradually converge to the desired concentration. Moreover, the resulting droplet concentration becomes confined to a narrow range. ■

From Equation (5) and Table 1, we can conclude that the dilution-chain structure provides **volume stability** as well as **concentration stability** for the intermediate droplets. However, concentrations that can be generated from the two dilution-chain structures are limited. For example, the Type-A structure generates concentrations around 33% (see Fig. 2(a)),



**Figure 4: A dilution graph utilizing the proposed idea.**

while the Type-B structure generates concentrations around 66% (see Fig. 2(b)). Therefore, if the desired concentration is not near 33% or 66% (e.g., the concentration 47.07% that we considered earlier), we are not able to generate the desired concentration with only one dilution-chain structure (either Type-A or Type-B).

However, we can first use multiple dilution-chain structures to generate multiple intermediate droplets, and next combine these intermediate droplets with buffer and reagent droplets to achieve the desired concentration.

### 3.3 Illustrative Example

We next show how we can achieve the desired concentration from multiple dilution-chain structures.

Suppose the desired concentration is  $C = 241/512 = 47.07\%$  and the resolution is  $R = 1/512 = 0.20\%$ . In this case, the concentration calibrated range is  $[46.97\%, 47.17\%]$ . Based on the proposed solution, a dilution graph as sketched in Fig. 4 is used to generate a concentration that is within the calibrated range. In this graph, two Type-A dilution-chain structures are used. The left-hand side Type-A structure generates a 1X-volume droplet with concentration of 33.30078125%, while the right-hand side Type-A structure generates a 1X-volume droplet with concentration of 37.5%. Finally, we mix these two droplets with four 1X-volume buffer droplets and four 1X-volume reagent droplets from the reservoir to obtain a droplet with concentration 47.08% (which is within the calibrated range).

However, because of volume variations in each splitting operation, the resulting concentration will deviate from the desired value. Here, we assume that the volume variation  $x$  obeys the uniform distribution (i.e., evenly distributed from -7% to 7%). If we simulate the sample preparation process in Fig. 4 for a total of 100 times, 85 of them generate a concentration that is within the calibration range, and the *probability of success* (POS) is  $85/100 = 0.85$ . In contrast, if we use the BS method reviewed in Fig. 1(a), and simulate the sample preparation process for a total of 100 times, only 2 of them generate a resulting concentration that is within the calibrated range, i.e., the POS is only  $2/100 = 0.02$ .

Suppose the cost of 1X-volume buffer is 0.1 and the cost of 1X-volume reagent is 1. The BS method sketched in Fig. 1(a) consumes a 5X-volume buffer and a 5X-volume reagent. The cost for a one-time run is  $0.1 \times 5 + 1 \times 5 = 5.5$ . According to [16], the number of times ( $T$ ) we need to run the method until we get one success obeys the *geometric distribution*. Because the POS for the BS method is only 0.02, the expectation  $E[T]$  is equal to  $1/0.02$ . Hence, the amortized cost is  $C_{amt} =$

$5.5/0.02 = 275$ . On the other hand, the proposed method shown in Fig. 4 consumes only a 8X-volume buffer and a 6X-volume reagent. Therefore, the cost for a one-time run is  $0.1 \times 8 + 1 \times 6 = 6.8$ , and the amortized cost is  $C_{amt} = 6.8/0.85 = 8$ .

From this example, we conclude that, although the cost for the proposed method is larger than that for the BS method, we achieve a significant reduction in the amortized cost because of the significant higher value of POS.

## 4 DETAILS OF THE PROPOSED SOLUTION

In this section, we propose an ILP formulation to address two problems that arise when realizing the method proposed above: (1) what intermediate droplets must be generated by multiple dilution-chain structures; (2) how many buffer and reagent droplets are needed to derive the desired concentration. We first formulate these two problems and then describe the corresponding ILP model that can address them.

### 4.1 Formulation of the Remaining Problem

We first provide the following definitions and assumption about sample preparation. For a sample preparation protocol, we regard it as a success if the concentration of the resulting droplet is within the *calibrated range* of  $[C - 0.5R, C + 0.5R]$ . Otherwise, it is a failure. Suppose we run sample preparation for a total of  $N$  times. If  $N_s$  of them are successful, the *probability of success* (POS) is defined as  $POS = N_s/N$ . According to [4], dispensing operations are very reliable fluidic operations. Therefore, we assume that there is no volume variation in dispensing.

We now formulate the general problem as follows:

**Input:** (1) Required concentration  $C$ ; (2) Resolution  $R$  (i.e., the difference between two neighbouring concentrations); (3) The cost  $W_b$  of 1X-volume buffer; (4) The cost  $W_r$  of 1X-volume reagent; (5) The volume variation  $x$  in each splitting operation. We assume that it is a random variable that obeys the uniform distribution on  $[-0.07, 0.07]$ .

**Output:** A dilution graph based on the two dilution-chain structures. In this dilution graph,  $N_b$  1X-volume buffer and  $N_r$  1X-volume reagent droplets are used. The POS for this solution is obtained by simulation.

**Objective:** Minimize the amortized cost defined by:

$$C_{amt} = (W_b \times N_b + W_r \times N_r) \times \frac{1}{POS} \quad (6)$$

The first term in Equation (6) represents the cost of one run. If a sample preparation run fails, we need to run it again until success. According to [16], the number of runs until we get a final success obeys the *geometric distribution*, and the expectation number of runs is  $1/POS$ . Therefore, Equation (6) represent the amortized cost needed to obtain the desired resulting concentration.

### 4.2 Proposed ILP Model

As shown in Fig. 5, we use  $N$  *dilution-chain* (DC) structures to construct the dilution tree. Each DC structure  $DC_i$  can either be of Type-A or Type-B. Suppose that the deepest node we can reach is node  $M$ . We define two binary variables: (1)  $a_{i,j}$  is equal to 1 if  $DC_i$  generates a droplet with concentration  $C_A(j)$  (see Fig. 6). Otherwise,  $a_{i,j}$  is equal to 0; (2)  $b_{i,j}$  is equal to 1 if  $DC_i$  generates a droplet with concentration  $C_B(j)$  (see Fig. 6). Otherwise,  $b_{i,j}$  is equal to 0. Because

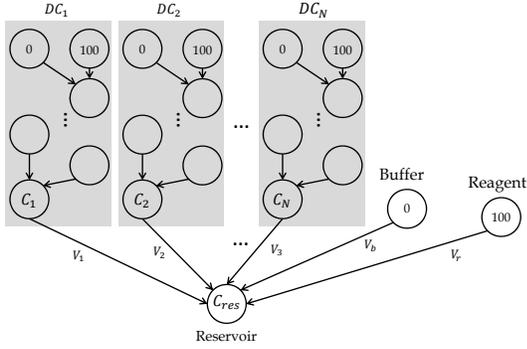


Figure 5: Structure of the ILP problem.

- 
- 1:  $C$ : the resulting droplet concentration;
  - 2:  $R$ : the concentration resolution;
  - 3:  $W_b$ : the cost of one unit-volume buffer droplet;
  - 4:  $W_r$ : the cost of one unit-volume reagent droplet;
  - 5:  $C_A(j)$ : the concentration of node  $j$  in the Type-A DC structure;
  - 6:  $C_B(j)$ : the concentration of node  $j$  in the Type-B DC structure;
  - 7:  $\Delta_{lo}^A(j)$ : the lower bound gap for Node  $j$  in the Type-A DC structure;
  - 8:  $\Delta_{up}^A(j)$ : the upper bound gap for Node  $j$  in the Type-A DC structure;
  - 9:  $\Delta_{lo}^B(j)$ : the lower bound gap for Node  $j$  in the Type-B DC structure;
  - 10:  $\Delta_{up}^B(j)$ : the upper bound gap for Node  $j$  in the Type-B DC structure;
  - 11:  $N$ : the number of DC structures used to construct the dilution graph.
  - 12:  $M$ : the maximum node index we can use in a DC structure.
- 

Figure 6: Variables used in the ILP model.

each DC structure generates only one droplet, we have the following constraints for all  $i = 1, \dots, N$ :

$$\sum_{j=1}^M (a_{i,j} + b_{i,j}) = 1 \quad (7)$$

The concentration of the droplet generated by  $DC_i$  ( $i = 1, \dots, N$ ) can be expressed as:

$$C_i = \sum_{j=1}^M (a_{i,j} \times C_A(j) + b_{i,j} \times C_B(j)) \quad (8)$$

Likewise, the lower bound gap and the upper bound gap (see Table 1) for the concentration in  $DC_i$  ( $i = 1, \dots, N$ ) are:

$$L_i = \sum_{j=1}^M (\Delta_{lo}^A(j) \times a_{i,j} + \Delta_{lo}^B(j) \times b_{i,j}) \quad (9)$$

$$U_i = \sum_{j=1}^M (\Delta_{up}^A(j) \times a_{i,j} + \Delta_{up}^B(j) \times b_{i,j}) \quad (10)$$

Suppose that  $V_i$  ( $i = 1, \dots, N$ ) represents the volume of a droplet generated by  $DC_i$ . Because each node in a DC structure represents a 2X-volume droplet, the value of  $V_i$  can be either 1 or 2. If  $V_i$  is equal to 1, the generated droplet is split before it is sent to the reservoir.<sup>1</sup> Otherwise, it is mixed with other droplets in the reservoir. In addition, the volume of buffer and reagent droplets dispensed to the chip are  $V_b$

<sup>1</sup>Note that the droplet splitting here can also introduce volume variation. However, we do not consider this in the ILP formulation. But it is considered when we calculate the POS for an ILP solution (see Section 4.3).

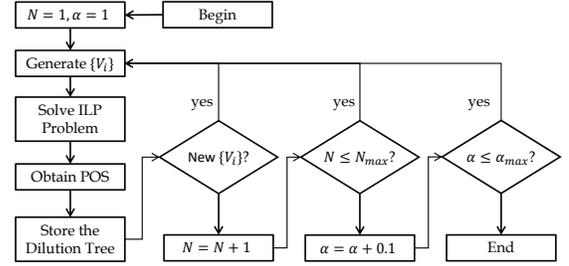


Figure 7: Overall flow of the proposed method.

and  $V_r$ , respectively. Note that the value of  $V_i$  ( $i = 1, \dots, N$ ) is predefined for this ILP problem. However,  $V_b$  and  $V_r$  are free variables, i.e., the ILP solver is supposed to assign them so that the target concentration  $C$  is achieved.

As mentioned above, the calibrated range is  $[C - 0.5R, C + 0.5R]$ . However, this range is too restrictive to ensure a feasible ILP solution. We need to introduce a user-defined parameter  $\alpha$  to make this range a little wider, and the resulting adjustable calibrated range is therefore  $[C - 0.5\alpha R, C + 0.5\alpha R]$  ( $\alpha > 1$ ).

If the concentration of each  $DC_i$  achieves its maximum value of  $C_i + U_i$ , the resulting concentration also achieves its maximum value. We require this value to be equal to or smaller than  $C + 0.5\alpha R$ , i.e., we have:

$$\frac{100\% \times V_r + \sum_{i=1}^N (C_i + U_i) \times V_i}{V_b + V_r + \sum_{i=1}^N V_i} \leq C + \alpha \times \frac{R}{2} \quad (11)$$

If the concentration of each  $DC_i$  achieves its minimum value of  $C_i - L_i$ , the resulting concentration also achieves its minimum value. We require this value to be equal to or larger than  $C - 0.5\alpha R$ , i.e., we have:

$$\frac{100\% \times V_r + \sum_{i=1}^N (C_i - L_i) \times V_i}{V_b + V_r + \sum_{i=1}^N V_i} \geq C - \alpha \times \frac{R}{2} \quad (12)$$

The ILP problem needs to minimize the following cost:

$$C = W_b \times (V_b + \sum_{i=1}^N \sum_{j=1}^M 2a_{i,j} + \sum_{i=1}^N \sum_{j=1}^M b_{i,j}) + W_r \times (V_r + \sum_{i=1}^N \sum_{j=1}^M a_{i,j} + \sum_{i=1}^N \sum_{j=1}^M 2b_{i,j}) \quad (13)$$

Note that in this ILP model, we optimize only the (not amortized) cost. The amortized cost is optimized in the control flow discussed in the following subsection.

### 4.3 Overall Flow

Fig. 7 shows the overall flow that minimizes the amortized cost of a dilution tree. Suppose  $N_{max}$  represents the maximum number of DC structures we can use, and let  $\alpha_{max}$  be the maximum value of the user-defined variable  $\alpha$ . First, we initialize both  $N$  and  $\alpha$  to 1. Next, for each value of  $N$ , we generate all possible combinations of  $V_i$  ( $i = 1, \dots, N$ ). For example,  $(V_1, V_2, V_3) = (1, 2, 1)$  and  $(V_1, V_2, V_3) = (1, 1, 1)$  are different combinations, while  $(V_1, V_2, V_3) = (1, 2, 1)$  and  $(V_1, V_2, V_3) = (2, 1, 1)$  are the same combination. For each combination, we formulate and solve the ILP problem as described above, and obtain the resulting dilution tree. Following this step, we carry out Monte Carlo simulation to

obtain the corresponding POS and amortized cost (which are stored in the memory). This process is repeated until all  $(N, \alpha)$  values are iterated, resulting in a substantial number of dilution trees. We select the one with the minimum amortized cost as the optimal dilution tree.

The total number of value combinations for  $V_i$  ( $i = 1, \dots, N$ ) is  $N + 1$ , therefore we need to solve  $(N + 1)$  ILP problems for each innermost loop in Fig. 7. When  $N$  is equal to  $1, 2, \dots, N_{max}$ , we need to solve  $2, 3, \dots, (N_{max} + 1)$  ILP problems, respectively. Therefore, we need to solve  $((N_{max} + 1 + 2)N_{max}/2)$  ILP problems for each middle loop in Fig. 7. Finally,  $\alpha$  is increased from 1 to  $\alpha_{max}$  with a step size of 0.1, so there are  $((\alpha_{max} - 1)/0.1 + 1)$  values of  $\alpha$  iterated for the outermost loop in Fig. 7. Based on the this analysis, the total number of ILP problems  $N_{ilp}$  we need to solve in Fig. 7 can be expressed as:

$$N_{ilp} = \frac{(N_{max} + 3)N_{max}}{2} \times \left[ \frac{(\alpha_{max} - 1)}{0.1} + 1 \right] \quad (14)$$

In practice, we find that when  $N_{max} = 5$  and  $\alpha_{max} = 1.5$ , we are able to obtain a good dilution tree with low amortized cost. The total number of ILP problems we need to solve for this setting is 120, and it takes about 10 s to complete the computation on a computer with a Core i3 3.7 GHz CPU and a 8 GB memory. The solution space for the ILP problem formulated in Section 4.2 is  $O(2^{M \times N})$ , but it only takes about 0.09 s to find the solution using a commercial ILP solver [15].

## 5 EXPERIMENTAL RESULTS

In this section, we compare the proposed method with four prior methods, namely BS [21], DMRW [19], REMIA [12] and FLOW [3]. For these methods, we inject volume variations into each splitting operation, and inspect the performance of each method.

### 5.1 Metrics

We first define the metrics that are used to evaluate the performance of each method. In a sample preparation protocol, several buffer and reagent droplets are used to generate the expected concentration. However, because of the existence of waste droplets, not all the available buffers and reagents are used for the resulting droplet. If there are many waste droplets, the corresponding sample preparation method is not cost-efficient. Therefore, we define a metric called *Droplet Utilization* to evaluate the extent of cost-saving of a sample preparation method, i.e.,

$$UT = \frac{W_b \times N \times (1 - C) + W_r \times N \times C}{W_b \times N_b + W_r \times N_r}, \quad (15)$$

where  $W_b$  is the cost of a 1X-volume buffer,  $W_r$  is the cost of a 1X-volume reagent,  $N_b$  is the volume of the input buffer,  $N_r$  is the volume of the input reagent,  $N$  is the volume of the resulting droplet, and  $C$  is the desired concentration. This metric represent the ratio of (1) the cost of reagent and buffer **remaining** in the resulting droplet versus (2) the cost of the input reagent and buffer. It is obvious that the method reduces cost if the value of this metric is high.

We also define a metric to evaluate the unified amortized cost to generate the desired droplet. All prior methods generate a droplet of 2X volume. However, our method generates a droplet with a volume larger than 2X. To ensure fair comparison, we define a second metric called *Unit Amortized Droplet Cost*, i.e.,

**Table 2: Simulation results for previous methods (R=1/256).**

Concentration: 1/256 to 255/256, Resolution: 1/256					
Metrics	$(W_r, W_b)$	BS	DMRW	REMIA	FLOW
$N_r$		4.49	2.98	2.31	2.22
$N_b$		4.02	2.98	4.69	5.93
$N_w$	N/A	6.03	3.96	4.99	6.68
POS		0.15	0.97	0.43	0.28
UT		0.22	0.34	0.40	0.39
COST	(1, 0)	14.97	1.54	2.69	3.96
	(1, 0.1)	16.31	1.69	3.23	5.02
	(1, 1)	28.37	3.07	8.14	14.55
	(1, 4)	68.57	7.68	24.50	46.32

**Table 3: Simulation results for proposed method (R=1/256).**

Concentration: 1/256 to 255/256, Resolution: 1/256						
$(W_r, W_b)$	$N_r$	$N_b$	$N_w$	POS	UT	COST
(1, 0)	4.66	5.67	1.43	0.96	0.96	<b>0.55</b>
(1, 0.1)	4.66	5.67	1.43	0.96	0.96	<b>0.61</b>
(1, 1)	4.89	5.32	1.56	0.94	0.95	<b>1.26</b>
(1, 4)	4.33	4.01	1.21	0.85	0.91	<b>3.36</b>

**Table 4: Simulation results for previous methods (R=1/1024).**

Concentration: 1/1024 to 1023/1024, Resolution: 1/1024					
Metrics	$(W_r, W_b)$	BS	DMRW	REMIA	FLOW
$N_r$		5.00	3.52	2.41	2.22
$N_b$		4.01	3.50	6.09	9.68
$N_w$	N/A	8.01	6.02	7.05	10.90
POS		0.08	0.81	0.19	0.11
UT		0.20	0.28	0.36	0.35
COST	(1, 0)	31.25	2.17	6.34	10.09
	(1, 0.1)	33.76	2.39	7.94	14.49
	(1, 1)	56.31	4.33	22.37	54.09
	(1, 4)	131.50	10.81	70.45	186.09

**Table 5: Simulation results for proposed method (R=1/1024).**

Concentration: 1/1024 to 1023/1024, Resolution: 1/1024						
$(W_r, W_b)$	$N_r$	$N_b$	$N_w$	POS	UT	COST
(1, 0)	8.97	10.11	2.44	0.80	0.96	<b>0.67</b>
(1, 0.1)	8.97	10.11	2.44	0.80	0.96	<b>0.75</b>
(1, 1)	9.45	9.66	2.56	0.77	0.93	<b>1.50</b>
(1, 4)	8.32	7.44	2.13	0.70	0.93	<b>3.99</b>

$$COST = \frac{W_b \times N_b + W_r \times N_r}{POS} \times \frac{1}{N} \quad (16)$$

where  $W_b$ ,  $W_r$ ,  $N_b$ ,  $N_r$  and  $N$  are the same as those in Equation (15), and  $POS$  is the probability of success.

### 5.2 Results

A sample preparation bioassay is typically composed of droplet transportation, droplet splitting, and droplet mixing operations. The physics underlying these operations is described in [4], and experimental results were also reported. In this section, the results of sample preparation methods are generated by simulation. Since the simulations are based on the model in [4], our results reflect the results what is expected for actual biochips.

In our evaluations, we first considered concentrations ranging from 1/256 to 255/256 with a resolution of 1/256. The results are shown in Table 2 and Table 3, where  $N_r$ ,  $N_b$  and  $N_w$  represent the average number of reagent droplets, the average number of buffer droplets, and the average number of waste droplets, respectively. Without loss of generality, we used four  $(W_r, W_b)$  values in our simulations, namely (1, 0), (1, 0.1), (1, 1) and (1, 4). These four values are selected based on the reagents and buffers used in a DMFB-based

bioassay [9]. In this bioassay, reagents with different concentrations are used, and we use four of these reagents to extract the cost values: (1) (1, 0) for trichloroacetic acid in DI water; (2) (1, 0.1) for Phuronic F127 in Borate; (3) (1, 1) for acetonitrile in chloroform; (4) (1, 4) for acetone in chloroform.

As shown in Table 2 and Table 3, the *POS* of BS, REMIA and FLOW are relatively low (smaller than 0.50) because their dilution graphs are not designed with respect to volume variation in splitting operations. Both DMRW and the proposed method achieve a *POS* higher than 0.9. Note that DMRW does not consider volume variation in the design stage. However, because it has a dilution graph that is similar to the two dilution-chain structures, it achieves nearly the same *POS* as our method. For the other methods, many waste droplets are discarded, therefore, the *droplet utilization (UT)* for these methods is less than 0.50. However, in our method, most of the buffer and reagent droplets are used to construct the final droplet, and the value of *UT* is as high as 0.96. When it comes to parameter *COST*, the value of *COST* is the lowest among all methods, because the proposed method has a high *POS* and a high *UT*. Note that the  $(W_r, W_b)$  entry for  $N_r$  to *UT* is labeled as “N/A” because the change of this value does not have any effect to previous methods. However, as the value of  $W_b$  changes, the *COST* values for previous methods change as well. It is obvious by inspection that the *COST* values for BS, REMIA and FLOW methods are much larger than that in DMRW and the proposed method because their droplet utilization (i.e., *UT*) and *POS* are relatively low.

From Table 3, we can see that the simulation results for value (1, 0) and (1, 0.1) are the same, because in both cases the cost of 1X-volume buffer is negligible. However, when this cost increases (e.g., value (1, 1) and value (1, 4)), the buffer usage is reduced to limit the overall cost. Moreover, *POS* decreases as well because one-time cost dominates the amortized cost. Therefore, solutions with low one-time cost and *POS* are selected by the proposed method.

Next, we consider a total of 1024 concentrations, ranging from 1/1024 to 1023/1024, with a resolution of 1/1024. The results are reported in Table 4 and Table 5. Because the resolution is much smaller (only 1/4 of that in Table 2), the BS, REMIA and FLOW method are not usable due to the low *POS*s (0.08, 0.19 and 0.11, respectively). However, the *POS* for DMRW and our method is acceptable.

## 6 CONCLUSION

We have presented a sample preparation method that provides the desired concentration accuracy despite volume variations in droplet splitting. We attempt to mitigate the volume-variation problem by the use of two dilution-chain structures. These structures can generate a droplet with concentration in a narrow range even if there are volume variations. Based on these structures, we have presented an ILP-based method for the construction of a dilution tree. It generates a resulting droplet with the target concentrations even if volume variations occur during sample preparation. Experimental evaluations demonstrate the applicability and effectiveness of the proposed method. In future work, we will propose a more sophisticated method that utilizes the dilution-chain structure, and compare the area utilizations and number of operations for different sample preparation methods. Besides that, comparisons with error-oblivious sample preparation methods as recently proposed in [17] shall be conducted.

## REFERENCES

- [1] Nilina Bera et al. 2016. Simulation-based method for optimum microfluidic sample dilution using weighted mix-split of droplets. *IET Computers & Digital Techniques* 10, 3 (2016), 119–127.
- [2] Krishnendu Chakrabarty et al. 2010. Design tools for digital microfluidic biochips: toward functional diversification and more than moore. *IEEE Trans. CAD* 29, 7 (2010), 1001–1017.
- [3] Trung Anh Dinh et al. 2014. A network-flow-based optimal sample preparation algorithm for digital microfluidic biochips. In *ASP-DAC*. 225–230.
- [4] Richard B Fair. 2007. Digital microfluidics: is a true lab-on-a-chip possible? *Microfluidics and Nanofluidics* 3, 3 (2007), 245–281.
- [5] Peter RC Gascoyne et al. 2004. Dielectrophoresis-based sample handling in general-purpose programmable diagnostic instruments. *Proc. of the IEEE* 92, 1 (2004), 22–42.
- [6] Amit Gefen. 2009. *Bioengineering Research of Chronic Wounds: A Multidisciplinary Study Approach*. Vol. 1. Springer Science & Business Media.
- [7] Daniel Grissom et al. 2012. Fast online synthesis of generally programmable digital microfluidic biochips. In *CODES+ISSS*. 413–422.
- [8] Daniel T Grissom et al. 2014. A low-cost field-programmable pin-constrained digital microfluidic biochip. *IEEE Trans. CAD* 33, 11 (2014), 1657–1670.
- [9] Mais J Jebraill and Aaron R Wheeler. 2008. Digital microfluidic method for protein extraction by precipitation. *Analytical chemistry* 81, 1 (2008), 330–335.
- [10] Oliver Keszocze et al. 2014. Exact One-pass Synthesis of Digital Microfluidic Biochips. In *DAC*. 1–6.
- [11] Oliver Keszocze et al. 2015. A General and Exact Routing Methodology for Digital Microfluidic Biochips. In *ICCAD*. 874–881.
- [12] Chia-Hung Liu et al. 2015. Reactant minimization in sample preparation on digital microfluidic biochips. *IEEE Trans. CAD* 34, 9 (2015), 1429–1440.
- [13] Yan Luo et al. 2013. Error recovery in cyberphysical digital microfluidic biochips. *IEEE Trans. CAD* 32, 1 (2013), 59–72.
- [14] S Miller et al. 2016. Multiplex detection of respiratory pathogens with Genmark’s eplex sample-to-answer system. *Clinical Chemistry and Laboratory Medicine* 54, 5 (2016), eA5–eA6.
- [15] Gurobi Optimization. 2017. Gurobi optimizer version 7.0. 2.
- [16] Andreas N Philippou et al. 1983. A generalized geometric distribution and some of its properties. *Statistics & Probability Letters* 1, 4 (1983), 171–175.
- [17] Sudip Poddar et al. 2018. Error-Oblivious Sample Preparation with Digital Microfluidic Lab-on-Chip. *IEEE Trans. CAD* (2018).
- [18] Hong Ren et al. 2003. Design and testing of an interpolating mixing architecture for electrowetting-based droplet-on-chip chemical dilution. In *TRANSDUCERS*, Vol. 1. 619–622.
- [19] Sudip Roy et al. 2010. Optimization of dilution and mixing of biochemical samples using digital microfluidic biochips. *IEEE Trans. CAD* 29, 11 (2010), 1696–1708.
- [20] Sudip Roy et al. 2011. Waste-aware dilution and mixing of biochemical samples with digital microfluidic biochips. In *DATE*. 1–6.
- [21] William Thies et al. 2008. Abstraction layers for scalable microfluidic biocomputing. *Natural Computing* 7, 2 (2008), 255–275.
- [22] Lisa R Volpatti et al. 2014. Commercialization of microfluidic devices. *Trends in biotechnology* 32, 7 (2014), 347–350.
- [23] Ping-Hung Yuh et al. 2006. Placement of digital microfluidic biochips using the T-tree formulation. In *DAC*. 931–934.