Integrated and Real-Time Quantitative Analysis Using Cyberphysical Digital-Microfluidic Biochips*

Mohamed Ibrahim†, Krishnendu Chakrabarty‡, and Kristin Scott†

†Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708, USA
‡Department of Molecular Genetics and Microbiology, Duke University, DUMC 3053, Durham, NC 27710, USA

{mohamed.s.ibrahim, krishnendu.chakrabarty, kristin.scott}@duke.edu

Abstract—Considerable effort has recently been directed towards the implementation of molecular bioassays on digital-microfluidic biochips. However, today’s solutions suffer from the drawback that multiple sample pathways are not supported and on-chip reconfigurable devices are not efficiently exploited. To overcome this problem, we present a spatial-reconfiguration technique that incorporates resource-sharing specifications into the synthesis flow. This technique is combined with cyberphysical integration to develop the first design-automation framework for quantitative gene expression. The proposed framework is based on a real-time resource-allocation algorithm that responds promptly to decisions about the protocol flow received from a firmware layer. Simulation results show that our adaptive framework efficiently utilizes on-chip resources to reduce time-to-result without sacrificing the chip’s lifetime.

I. INTRODUCTION

According to a recent announcement by Illumina, a market leader company in DNA sequencing, digital-microfluidics (DMF) technology [1] has been transitioned to the marketplace for sample preparation [2]. This significant milestone highlights the emergence of digital-microfluidic biochips (DMFBs) for commercial exploitation and their potential for immunoassays for point-of-care diagnosis [3], proteomic sample processing, and cell-based assays [4]. Using DMFBs, bioassay protocols are scaled down in terms of liquid volumes and assay times and executed by enabling precise control of discrete droplets using a patterned array of electrodes. Therefore, picoliter droplets of samples and reagents can be dispensed, incubated, transported, mixed, split, heated, or electroporated under software control in a cost-effective manner [1]. The flexibility provided by this technology, along with advances in the integration of sensors [1], [5] and droplet monitoring using CCD cameras [6], were exploited in [7] to develop a physical-aware system reconfiguration technique.

Due to the fundamental importance of genomic analysis, major advances have also been reported on miniaturized technology platforms [8]–[10]. With the advent of these platforms, genomic bioassays such as nucleic-acid isolation, DNA purification, and DNA amplification have been successfully realized and integrated on DMFBs. Temperature-cycling of samples for DNA amplification, also referred to as polymerase chain reaction (PCR), has also been demonstrated on a DMFB.

However, since these platforms were intrinsically designed for experimentation on a sample-limited setting, on-chip devices were allocated a priori with respect to the bioassays constituting the protocol; thus the flexibility and reconfigurability of these devices have not been efficiently exploited. For example, today’s DMFBs can be flexibly equipped with a mechanism for temperature-cycling through which a target sample is kept immobile while the temperature of an associated chip region is precisely tuned [9], [10]. This mechanism enables resource sharing, i.e., the option of using such regions for various purposes, such as PCR, thermal cell lysis or even traditional sample processing. Since quantitative analysis protocols include multiple sample pathways that are independently manipulated, there is inherent uncertainty about the order of basic fluidic steps; thus resource sharing is a significant challenge. Inefficient exploitation of on-chip devices for real-life protocols entails significant cost since on-chip devices, such as heaters and sensors, need to be replicated.

On the other hand, the electrodes located within the heating regions are unique in terms of their role since they control the droplets that are subjected to thermal manipulation. Therefore, excessive usage of these electrodes leads to degradation, which significantly impacts biochip lifetime. The same concern applies to other on-chip devices such as sensors and electrodes that are in proximity to magnets used for bead-based assays.

In this paper, we advance cyberphysical integration of DMFBs by introducing the first integrated platform and design-automation solution for quantitative analysis, e.g., the study of gene expression in molecular biology. Based on benchtop experimental work, we present a laboratory-based approach for executing quantitative analysis bioassays. We then present a method for physical-aware resource allocation for multiple sample pathways. The main contributions of this paper are as follows:

• We present a spatial-reconfiguration technique that incorporates resource-sharing constraints into the synthesis flow.
• We describe and evaluate a physical-aware resource-allocation framework that enables tunable resource allocation among bioassays.

The rest of the paper is organized as follows. Section II describes related prior work. An introduction to gene-expression analysis and the related protocol are presented in Section III. Section IV explains the requirement of spatial reconfiguration for the efficient realization of protocols. Next, the shared resource-allocation algorithm is presented in Section V. Finally, results of our experimental evaluation are presented in Section VI and conclusions are drawn in Section VII.

*This work was supported in part by the National Science Foundation under grant CNS-1135853.
II. RELATED PRIOR WORK

Early research on design automation for digital microfluidics focused on scheduling of fluidic operations, resource binding, droplet routing, and mapping of control pins to electrodes [11]–[14]. However, these methods are limited to single sample pathways and they cannot handle uncertainties in the order of fluidic steps in bioassay execution. Moreover, the interplay between hardware and software in the biochip platform (cyberphysical system design) was not considered in these early methods.

Design and optimization techniques for cyberphysical DMFBs have thus far considered only error recovery [7], [15]–[17] and termination control of a biochemical procedure such as PCR [18]. These designs do not support real-time decision making for protocols designated for multiple samples, and they do not exploit the potential of DMFBs for quantitative analysis-based biochemical reactions. In [19], Gao et al. proposed a modular design with built-in electronic control to enhance the reliability and robustness of DMFBs. However, this design can only be used for droplet volume measurement and position control. With this computationally expensive method, the control of more than a few droplets transported concurrently on an array becomes impractical.

Therefore, a major limitation of all prior work on design automation and cyberphysical integration is that they are limited to simple droplet manipulation on a chip; however, in order to make DMFBs useful to a biologist, we need a new design paradigm to demonstrate that these chips can be used for actual biomolecular protocols from microbiology.

III. MINIATURIZATION OF GENE-EXPRESSION ANALYSIS

Accurate quantification of the expression level for a target gene requires a multi-assay protocol on sample cells [20]. Fig. 1(a) depicts a flowchart of the protocol for quantitative gene-expression analysis in fission yeast that we have studied in the microbiology lab using a benchtop setup. We have studied the transcriptional profile of a green fluorescent protein (GFP) reporter gene under epigenetic control. Control (GFP gene not under epigenetic control) and experimental strains were analyzed by quantitative PCR (qPCR) following cell lysis, mRNA isolation and purification, and cDNA synthesis. These studies allowed us to assess the quality of the protocol implementation and define the influence of epigenetic chromatin structures on gene expression. Intermediate decision points have been used to control the protocol flow for every sample.

An early non-adaptive implementation of this protocol was realized using digital microfluidics [9]. Our recent benchtop experience with multiple sample pathways highlights the need for incorporating decision-making and adaptation capability. Hence, we have developed the enhanced and miniaturized protocol shown in Fig. 1(b). On-chip operation begins with the dispensing of sample droplets containing cultured cells. The cells are then lysed in order to obtain intracellular materials (DNAs, RNAs, proteins, etc.). Using magnetic beads, enzymes, and a washing step, mRNA can be isolated and then reverse-transcribed into the corresponding complementary DNA (cDNA) with primers and other reverse-transcription reagents. Next, the resulting cDNA samples are subjected to thermal cycling via qPCR to amplify the target gene.

Gene-expression analysis requires the execution of this protocol on two sample droplets, one of which is used to quantify the amplification of the gene-under-investigation, whereas the other droplet is used to quantify the amplification of a reference gene, also known as a housekeeping gene [21]. Since the primers are gene-specific, the two droplets are chemically treated using different types of primers. However, based on the outcomes of the intermediate decision points, the two droplets can be utilized in an unpredictable manner by different bioassays; this unpredictability makes microfluidic control difficult.

Furthermore, with the randomness exhibited in molecular interactions, effective gene-expression analysis requires that the experiment be conducted on at least three replicates. The expression level is first calculated for each replicate, then averaged across the three replicates.

Hence, designing an autonomous digital-microfluidic system for gene-expression analysis requires the concurrent manipulation of independent samples. Utilizing the decision points shown in Fig. 1(b), we incorporate sample-dependent decision-making capability into the cyberphysical system. In addition, the specification of the protocol efficiency and the level of gene expression are included on-chip in the feedback system. Details on determining the protocol efficiency and gene-expression level are omitted due to lack of space.

In order to solve the synthesis problem for the new paradigm, we represent the protocol as a control flow graph...
Water A bioassay The set of shared resources that are essential The set of shared, reconfigurable modules The complete set of the chip modules The degradation caused by a bioassay The completion time for a bioassay

Bead snapping Mixing or dilution Thermo-cycling

Fig. 2: A CFG representation of the gene-expression analysis protocol for a single sample pathway.

(CFG), in which every node (referred to here as a supernode) signifies a bioassay, e.g., cell lysis. The bioassay operations, in turn, are represented in a directed sequencing graph, which shows the timing and the interdependencies among these operations. The synthesis tool must be able to synthesize the supernodes of the protocol CFG based on the decisions made for each sample during running time. Fig. 2 illustrates the representation of the protocol used for quantitative analysis of gene expression. Due to the mapping of many bioassays onto this platform, the design of the underlying DMFB must include on-chip devices such as heaters, magnets, and sensors that are required to complete the execution of the protocol. Existing biochip synthesis methods [11], [13], [14] must be extended to handle dynamic adaptation for multiple sample pathways and the inherent uncertainties associated with them.

IV. SPATIAL RECONFIGURATION

Gene-expression analysis using DMFBs relies on the concurrent manipulation of a collection of droplets. Since there is uncertainty about the order of execution of bioassays, and the allocation of on-chip devices among the samples is not known a priori, there is a need for a reconfiguration technique that can map specifications of the bioassays to the space of chip resources; we refer to this as spatial reconfiguration.

The above mapping must consider the degradation caused by a bioassay. In [10], Norian et al. provided a degradation model for the electrodes. An electrode’s lifetime can be divided into three regions, reliable operation, safety margin, and breakdown. In the reliable operation region, the threshold voltage needed for actuation is constant. The threshold voltage increases linearly in the safety margin region. In the breakdown region, a significant increase in the electrowetting voltage is needed in order to transport a droplet. This increase in voltage, however, quickly leads to dielectric breakdown [22].

To explore the space of spatial reconfiguration for a bioassay, analysis is needed to determine the total completion time and the resulting electrode degradation for the critical resources over a varying range of resource availability. We consider three levels of spatial reconfiguration:

Non-reconfigurable scheme: This scheme is adopted by current prototypes; on-chip devices are allocated a priori [9], [10]. Inter-bioassay communication is achieved by passing droplets between the dedicated areas.

Restricted resource sharing: The restriction here is in terms of the reconfigurability of the shared devices among bioassays. For instance, a heat-detect device can be shared between qPCR and cell lysis for the purpose of thermal manipulation or even sample processing, but it cannot be used by other bioassays.

Unrestricted resource sharing: In this scheme, no restriction is imposed on resource sharing. Therefore, heat-detect modules cannot only be used for thermal manipulation or sample processing in cell lysis or qPCR, but they can also be utilized by sample processing operations in all bioassays.

Based on the above definition, it is apparent that unrestricted resource sharing achieves minimum completion time, but it reduces chip lifetime. On the other hand, restricted resource sharing decelerates degradation of the chip, but the completion time is higher. Therefore, there is a need for a resource-allocation scheme that combines the best of both worlds; i.e., lower completion time and less degradation of the chip. We refer to this scheme as degradation-aware resource allocation and describe it in the next section.

V. SHARED-RESOURCE ALLOCATION

This section formulates the resource-allocation problem and describes the proposed solution.

A. Problem Formulation

The notation used in this paper is listed in Table I. The system described here is composed of three types of modules: (1) non-reconfigurable modules (input and output ports), (2) sample-processing modules (mixers), and (3) reconfigurable modules (heaters, detectors, and magnet regions). Unlike the non-reconfigurable and samples processing modules, the reconfigurable modules are shared among the protocol’s bioassays, and access control is managed by the resource allocator. In addition, we make the following important observations:

- Each bioassay $b_i \in B$ is mapped to a local space of non-reconfigurable and sample processing modules.

| $B$ | The complete set of the protocol bioassays |
| $b_i$ | A bioassay $b_i$ |
| $R$ | The complete set of the chip modules |
| $R_{gr}$ | The set of shared, reconfigurable modules |
| $r_i$ | The set of shared resources that are essential for a bioassay $b_i$ |
| $gr_j$ | The set of shared resources that are granted to a bioassay $b_i$ |
| $D_i^j$ | The degradation caused by a bioassay $b_i$ on the shared resources $gr_j$ |
| $T_i^j$ | The completion time for a bioassay $b_i$ when the resources $gr_j$ are used |
A shared module $r \in R_{sh}$ is essential for a bioassay $b_i$ if and only if the absence of this module leads to a failure in the execution of this bioassay. An example is the heater resource for a thermal-cycling bioassay.

A shared module $r \in R_{sh}$ that is granted to a bioassay $b_i$ may not be essential for the execution of the bioassay, i.e., $r_i \subseteq gr_j$. Typically, this module can be used for sample processing.

Our problem formulation is as follows:

**Inputs:** (1) The protocol CFG $G_c = \{V, E\}$, where $V = \{V_1, V_2, ..., V_m\}$ represents the supernodes of $m$ bioassays and $E = \{(V_i, V_j) \mid 1 \leq i, j \leq m\}$ represents data and biological dependency between all pairs of bioassays $b_i$ and $b_j$. A supernode $V_i$ comprises a directed sequencing graph $G_s = \{V_s, E_s\}$, where $V_s = \{V_{s1}, V_{s2}, ..., V_{sn}\}$ represents $n$ bioassay operations and $E_s = \{(V_{si}, V_{sj}) \mid 1 \leq si, sj \leq n\}$ represents dependencies between all pairs of operations $si$ and $sj$ that belong to the bioassay $V_i$.

(2) The digital microfluidic library, which describes the types and locations of the on-chip modules.

(3) The resource preferences of each bioassay $b_i$, which describes the initial resource requirement. The values of the bioassay’s completion time $T_i$ and its degradation level $D_i$ as a function of the granted resources $gr_j$ are also specified.

(4) The resource-allocation constraints (Section V-B).

**Output:** Allocation of chip modules to the bioassays such that the constraints on resource-allocation are satisfied.

### B. Resource-Allocation Constraints and Algorithm

We have developed a shared-resource allocation scheme based on a timewheel that is controlled by the coordinator, as shown in Fig. 3. The sequence of actions is indicated by the numbers 1–6. To facilitate understanding, we describe the resource-allocation using Fig. 3 and a narrative format. A formal pseudocode procedure is not included here due to lack of space.

Whenever the coordinator receives a command from the firmware layer about the decision for a certain pathway, it first stores the command in a global queue until all preceding requests are fulfilled. When the bioassay’s command is ready to be processed (essential resources are available), the coordinator forwards it to the resource allocator agent. The resource allocator, in turn, checks the preferences of the bioassay and the scheme-specific constraints on resource allocation. Then, the resource allocation is determined and transferred to the actor which, in turn, invokes online synthesis for this particular bioassay.

When the restricted resource-sharing scheme is adopted, the resource allocator must ensure that no non-essential shared resource is allocated to the requesting bioassay; i.e., $r_i \in gr_j$. For instance, a bioassay for the preparation of master mixes cannot get access to the heaters, which are not essential for its execution, but it can get access to additional optical detectors to shorten time-to-completion. Therefore, the worst-case computational complexity for allocating resources to a bioassay $b_i$ is $O(|r_i|)$.

Conversely, the resource allocator in unrestricted resource sharing enables the option of using non-essential shared resources. In this case, these resources will be used for sample processing. For instance, an mRNA extraction bioassay can get access to a heater to perform sample processing in order to shorten its completion time. As a result, the worst-case computational complexity for allocating resources to a bioassay $b_i$ is $O(|R_{sh}|^2)$.

### VI. SIMULATION RESULTS

We implemented the proposed resource allocation schemes using C++. All evaluations were carried out using a 2.4 GHz Intel Core i5 CPU with 4 GB RAM. In order to obtain the resource preferences for each bioassay before resource allocation begins, synthesis results were obtained from offline simulation using [23]. The same tool is used for online synthesis when the simulation of real-time quantitative analysis starts. Since this is the first work on optimization for multiple sample pathways, a comparison with prior work is not feasible.

Two chip arrays have been utilized to simulate the quantitative protocol; these chips are labeled $\mathcal{C}_{NR}$ (non-reconfigurable chip) and $\mathcal{C}_R$ (reconfigurable chip). In both arrays, the chip has a dedicated port for each liquid and each bioassay also has its own sample processing (SP) region. However, reconfigurable modules, namely heaters (H), magnet (M), CCD camera region (CCD), and optical detectors (OD), are integrated in the shared area of $\mathcal{C}_R$. This area is accessible to droplets from any

![Fig. 3: The components of the shared-resource allocator.](image)
bioassay via a ring electrode-bus. The chip sizes are $18 \times 18$ and $17 \times 17$ for $C_{NR}$ and $C_R$, respectively; these represent the lower bounds on the size required to execute the bioassays. We expect that increasing the array size for $C_{NR}$ and $C_R$ will result in a proportional decrease in completion time. Performance assessment for different chip sizes is left for future work.

We evaluate four resource-allocation schemes: (1) Non-reconfigurable scheme (NON); (2) Restricted resource sharing (RR); (3) Unrestricted resource sharing (NR); (4) Adaptive degradation-aware resource sharing (DA). Using these schemes, we run simulations on the quantitative analysis protocol described in Section III. We consider three samples being concurrently subjected to fluidic operations. The samples are $S_1$ (GFP gene-targeted sample), $S_2$ (YFP gene-targeted sample, and $S_3$ (actin gene-targeted sample). Using the notation in Table II, we consider three different cases in terms of the sample pathways: (1) Short homogeneous pathways (Case I): an optimistic case in which all the three samples follow the same shortest pathway ($CL-mE-mP-MM-SD-TC$); (2) Long homogeneous pathways (Case II): a pessimistic case in which all the three samples follow the same long pathway ($CL-mE-mP-CL-mE-mP-MM-SD-TC$); (3) Heterogeneous pathways (Case III): a realistic case in which the pathways of these samples are different. The considered pathways are as follows: $CL-mE-mP-MM-SD-TC$, $CL-mE-mP-CL-mE-mP-MM-SD-TC$, and $CL-CL-mE-mP-MM-MM-SD-TC$, respectively. Table II lists the minimum resource requirements for the bioassays.

The metrics of comparison include: (1) The total completion time of the protocol execution (in time steps); (2) the occupancy time (degradation level) for the shared resources (heaters and magnet modules). A time-step refers to the clock period, typically in the range of 0.1 to 1 second [1]. For fair comparison between non-reconfigurable and resource-sharing schemes, the chip array used with the non-reconfigurable scheme (i.e., $C_{NR}$) is designed such that the resources are sufficient only for a single pathway; therefore, the resources are not replicated with the number of samples. In addition, due to the absence of resource sharing in the non-reconfigurable scheme, the occupancy time is not considered in our analysis. Table III compares the four scheme based on the chip configuration. As expected, due to the absence of resource sharing in the non-reconfigurable scheme, the chip modules are replicated. For example, three heaters are sufficient for the resource-sharing schemes (in $C_R$) for thermal manipulation in different bioassays. For $C_{NR}$, both cell lysis and thermal cycling bioassays require their own heaters, thus increasing chip fabrication cost.

Based on our simulations, the CPU time, which includes

Table II: Bioassay notation and resource requirement.

<table>
<thead>
<tr>
<th>Bioassay</th>
<th>Notation</th>
<th>Minimum resource requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Lysis</td>
<td>CL</td>
<td>SP, H, and CCD</td>
</tr>
<tr>
<td>mRNA Extraction</td>
<td>mE</td>
<td>SP and M</td>
</tr>
<tr>
<td>mRNA Purification</td>
<td>MP</td>
<td>SP and OD</td>
</tr>
<tr>
<td>RT Master Mix</td>
<td>MM</td>
<td>SP and OD</td>
</tr>
<tr>
<td>Serial Dilution</td>
<td>SD</td>
<td>SP</td>
</tr>
<tr>
<td>Thermal Cycling</td>
<td>TC</td>
<td>H</td>
</tr>
</tbody>
</table>

Table III: Number of on-chip modules and array electrodes in $C_{NR}$ and $C_R$.

<table>
<thead>
<tr>
<th>Resource</th>
<th># Modules</th>
<th># Electrodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP</td>
<td>1 (CL), 2 (mE), 2 (mP), 4 (SD)</td>
<td>176, 176</td>
</tr>
<tr>
<td>H</td>
<td>2 (CL), 3 (TC)</td>
<td>3, 3, 3, 3</td>
</tr>
<tr>
<td>M</td>
<td>1 (mE)</td>
<td>1, 1, 1, 1</td>
</tr>
<tr>
<td>OD</td>
<td>2 (mP), 2 (MM)</td>
<td>2, 2, 2, 2</td>
</tr>
<tr>
<td>CCD</td>
<td>1 (CL)</td>
<td>1, 1, 1, 1</td>
</tr>
<tr>
<td>Chip size</td>
<td>22</td>
<td>18, 18 x 18, 17 x 17</td>
</tr>
</tbody>
</table>

Fig. 4: (a)-(f) Comparison between the four resource-allocation schemes—Non-reconfigurable (NON), restricted resource sharing (RR), unrestricted resource sharing (NR), and degradation-aware resource sharing (DA)—executing three pathways. (g)-(h) Performance of the shared resource-allocation schemes with various lengths of homogeneous pathways.

2016 Design, Automation & Test in Europe Conference & Exhibition (DATE)
the time for resource allocation and online synthesis, averaged over the three cases (I, II, and III), is less than 1 ms for all allocation schemes. Thus, compared to the protocol completion time (in the order of minutes), the CPU time is negligible.

A. Case I: Short Homogeneous Pathways

This case arises when all the samples are perfectly grown in a well-controlled medium. In addition, the reagents are contamination-free. We compare the four methods in terms of completion times; see Fig. 4(a). As expected, non-reconfigurable resource allocation leads to the shortest completion time. Restricted resource sharing, on the other hand, shows the worst completion time. We also note that we can use adaptive, degradation-aware resource allocation (DA) to achieve a short completion time, while the chip resources are not severely degraded (degradation is measured in terms of the occupancy time); see Fig. 4(b).

B. Case II: Long Homogeneous Pathways

We next study the case where all the samples are suspended for additional bioassays. As shown in Fig. 4(c) and Fig. 4(d), the proposed schemes show the same profiles of completion times and degradation levels as in Case I, but with higher values. However, the completion time for DA is closer to that obtained with restricted resource sharing.

C. Case III: Heterogeneous Pathways

This is a realistic case that emerges due to the inherent uncertainty about the biological contents of each sample. It introduces the challenge of deciding the allocation of resources among the heterogeneous pathways at runtime. Again, we evaluate the allocation schemes based on the completion time (Fig. 4(e)) and the degradation level (Fig. 4(f)).

D. Convergence of Degradation-Aware Resource Allocation

Finally, we study the convergence of the DA method. We analyze the completion time and the average degradation level of the shared resource-allocation schemes with various lengths of homogeneous pathways, as shown in Fig. 4(g)-(h). We observe that the completion time for DA begins to converge to its counterpart in RR when the length of the pathway is increased. However, due to the restrictions imposed by both RR and DA on resource allocation, the degradation levels remain below a certain limit even when we increase the pathway length substantially, as shown in Fig. 4(h). In real-life scenarios, chip users tend to make optimizations in order to make sure that the protocol is finished as early as possible to avoid droplet evaporation [24]. These scenarios make the DA method especially attractive in practice.

VII. CONCLUSION

We have introduced the first automated design method for a cyberphysical DMFB that performs quantitative gene-expression analysis. This design is based on a spatial-reconfiguration technique that incorporates resource-sharing specifications into the software synthesis flow. A firmware has been developed to collect data from sensors during runtime, perform analysis, and make decisions about the multiple pathways for samples. We have also presented an adaptive scheme for shared-resource allocation that efficiently utilizes on-chip modules. This method manages resource access between the bioassays. The proposed scheme has been evaluated based on the completion time and the electrode degradation level for realistic multi-sample test cases. We have reported results on the performance of the allocation scheme for various pathway lengths.

REFERENCES