Methods 129 (2017) 60–80

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Synergistic target combination prediction from curated signaling networks: Machine learning meets systems biology and pharmacology

Huey Eng Chua [a,b], Sourav S. Bhowmick [a,b,*], Lisa Tucker-Kellogg [c]

[a] School of Computer Science & Engineering, Nanyang Technological University, Singapore
[b] Complexity Institute, Nanyang Technological University, Singapore
[c] DUKE-NUS Medical School, Singapore

ARTICLE INFO

Article history:
Received 10 December 2016
Received in revised form 4 April 2017
Accepted 18 May 2017
Available online 25 May 2017

Keywords:
Signaling network
Target combination
Synergism
Simulated annealing
Target prioritization
Off-target effects

ABSTRACT

Given a signaling network, the *target combination prediction problem* aims to predict efficacious and safe target combinations for combination therapy. State-of-the-art *in silico* methods use Monte Carlo simulated annealing (MCSA) to modify a candidate solution stochastically, and use the Metropolis criterion to accept or reject the proposed modifications. However, such stochastic modifications ignore the impact of the choice of targets and their activities on the combination's *therapeutic effect* and *off-target effects*, which directly affect the solution quality. In this paper, we present MASCOT, a method that addresses this limitation by leveraging two additional heuristic criteria to minimize *off-target effects* and achieve *synergy* for candidate modification. Specifically, *off-target effects* measure the unintended response of a signaling network to the target combination and is often associated with toxicity. *Synergy* occurs when a pair of targets exerts effects that are greater than the sum of their individual effects, and is generally a beneficial strategy for maximizing effect while minimizing toxicity. MASCOT leverages on a machine learning-based *target prioritization* method which prioritizes potential targets in a given disease-associated network to select more effective targets (better *therapeutic effect* and/or lower *off-target effects*); and on *Loewe additivity theory* from pharmacology which assesses the non-additive effects in a combination drug treatment to select synergistic target activities. Our experimental study on two disease-related signaling networks demonstrates the superiority of MASCOT in comparison to existing approaches.

1. Introduction

Cells use sophisticated biochemical interactions between proteins and other factors in order to perform a variety of information processing functions, collectively known as "cell signaling" in order to perform a variety of functions such as growth, survival, proliferation and development. As signaling proteins rarely operate in isolation through linear pathways, cell signaling can be viewed as a large and complex network. Understanding signal flow in the network is paramount, as alterations of cellular signaling events, such as those that arise by gene mutations or epigenetic changes, can result in various diseases. For example, alterations to the genes that encode canonical signaling proteins `Ras` or `PI3K` are commonly observed in many types of cancers.

Discovery of therapeutic drugs that can target these altered signaling pathways to restore the physiological state of a disease network to normalcy has long been dominated by the "one-target one-drug" paradigm (*i.e.*, identify a single chemical entity that binds to a single target). However, most complex disease states are *polygenic* and are characterized by a combination of interacting genes and their products instead of a single gene. Hence, increasing attention has been shifted to *combination therapy* by targeting multiple molecules simultaneously in a disease-related signaling network [3,41]. Specifically, in this therapy, instead of a single compound interacting with a single target, a concerted pharmacological intervention of several compounds interacting with multiple targets is made. Such a strategy has the potential to yield better benefits compared to a single molecule (referred to as mono-therapy) for complex diseases, because combination therapy offers the potential to achieve equal efficacy using lower doses (i.e., by leveraging *synergism*,[1]) and it may also offer novel opportunities to decrease the frequency at which resistance arises. Examples of a synergistic treatment strategy can be

* Corresponding author at: School of Computer Science & Engineering, Nanyang Technological University, Singapore.
E-mail addresses: hechua@ntu.edu.sg (H.E. Chua), assourav@ntu.edu.sg (S.S. Bhowmick), lisa.tucker-kellogg@duke-nus.edu.sg (L. Tucker-Kellogg).

[1] Synergism has a mathematical definition, as described in Section 3.3, but informally it is understood as giving greater effect than the sum of the individual effects. It is often beneficial for maximizing effect while minimizing toxicity.

http://dx.doi.org/10.1016/j.ymeth.2017.05.015
1046-2023/

found in the combination therapy of AIDS, cancer, and hypercholesterolaemia [41].

The pipeline of *target-based* drug discovery can be boiled down to three essential steps: the *discovery phase*, the *preclinical phase*, and the *clinical phase*. In the *discovery phase*, *targets* are first identified, screened and validated. A *target* is typically an endogenous molecule such as a protein, a gene or a nucleic acid sequence that affects the outcome of a disease or a medical condition.[2] Then, *lead compounds* for the validated targets are identified, screened and optimized. These are chemical compounds that demonstrate positive pharmacological activities that are in line with the desired therapeutic effect and are most likely to be successful in preclinical and clinical trials. Next, in the *preclinical phase*, optimized lead compounds are tested in animal models to establish drug safety and efficacy before proceeding to further testing in human subjects in the *clinical phase*. This pipeline typically spans 12–15 years and costs as much as 1 billion USD to bring a single drug to market [20].

1.1. Limitations of the discovery phase

Unfortunately, the aforementioned pipeline has not resulted in many successful new drugs. One key reason is that novel targets identified during the discovery phase have a low success rate [55]. Only 3% of novel targets identified reach the preclinical phase, or proceed to preclinical studies [61]. The remaining 97% of targets that have not been validated become unattractive to pharmaceutical companies due to the length of time needed for drug discovery and a high chance of failure in clinical trials. Furthermore, the classical approach for combination therapy is generally based on designing combinations based on clinical experience of doctors, knowledge of biological mechanisms, and practical constraints in the design of clinical trials [63]. Most drugs were initially developed as effective single agents and were only later combined clinically. A common assumption in this case is that only drugs that are effective individually should be used as part of a drug combination. Such an assumption excludes effective drug combinations such as anticancer drug combinations that work by leveraging synthetic lethal interactions [11]. Because the effect of drugs depends on the dose, several doses need to be studied for the drug combinations. Consequently, the number of possible combinations can grow exponentially. For instance, a cancer chemotherapy regimen can consist of six or more drugs from more than 100 anticancer drugs. However, investigating all six combinations out of 100 (including partial combinations) at three different doses to determine which combination is effective generates 8.9×10^{11} possibilities. The exhaustive set of all possible combinations is too large for empirical testing, so the design of combination therapies has created a need for scalable methods to prioritize combinations for testing. Although high-throughput screening (HTS) technology [8] allows the testing of pairs of drugs over a range of doses, combinatorial explosion still prevents exhaustive measurement of combinations of more than two drugs. Hence, there is a dire need to discover alternative procedures for the discovery phase to improve the target combination identification process in terms of efficiency of target discovery and efficacy of targets leading to the identification of superior drug combinations.

1.2. Data-driven target combination discovery

Observe that the discovery phase in the drug discovery pipeline for combination therapy broadly consists of two components: (a) identification of biologically-relevant target combinations and (b) development of therapeutic compounds (*e.g.*, drugs) that act on these targets. In this paper, we focus on the first step. Specifically, with the fast accumulation of experimental and omics data from HTS along with growing availability of disease-related signaling networks, *we explore the possibility of the pivotal role that the data analytics community can play in building novel data-driven tools to facilitate early detection of superior target combinations*. Such tools can facilitate discovery of superior drug combinations, contributing to a powerful discovery and pre-screening platform when coupled with other complementary technologies such as HTS. Consequently, it has the potential to reduce the time and dollar costs of the drug discovery pipeline.

1.3. Overview and contributions

In this paper, we present a generic approach called MASCOT (**M**achine Le**A**rning-based Prediction of **S**ynergistic **CO**mbinations of **T**argets) to address the data-driven target combination discovery problem by leveraging *curated* signaling networks. We assume that a signaling network is represented using mass action model where all interactions in the network are represented as reaction equilibriums with kinetics information [3]. Then, given a disease-related curated signaling network G and a desired *therapeutic effect* (*e.g.*, 50% ERKPP down-regulation), MASCOT predicts a set of synergistic target combinations and the required *target activities* (type and extent of perturbations) of these targets that achieve the desired therapeutic effect and have minimum *off-target effects*. Informally, the *therapeutic effect* and the *off-target effects* are measures of the intended and the unintended response, respectively, of a signaling network to the drug combination. Note that each drug effect on a target can be simulated *in silico* by perturbing the network to bring it back from disease state to normal state. Hence, the intended response during this perturbation is the resulting changes to the concentration of the *disease node* (*i.e.,* a downstream node whose activity determines the phenotypes that are manifested in response to signals flowing in the network [51]), whereas the unintended response is the resulting changes to the sum of the concentration of the rest of the nodes in the network.

In order to address the aforementioned problem, it is paramount for MASCOT to address the following key challenges. First, how do we reduce the search space in order to tackle the impact of exponential number of candidate target combinations? Second, how can we ensure that the predicted target combinations are synergistic in nature? Third, how do we quantify the off-target effects of a candidate target combination? The MASCOT algorithm is designed to address these challenges.

Intuitively, the MASCOT algorithm consists of two phases: *preprocessing* and *efficacy-conscious simulated annealing*. In the preprocessing phase, MASCOT first modifies the *reaction dynamics*[3] in G to incorporate actions of *activators* and *inhibitors* as we do not know *apriori* whether a target needs to be inhibited or activated.[4] *Activators* and *inhibitors* of a protein X are drugs/therapeutics that alter the system by increasing or decreasing the function of protein X, respectively. Then, it leverages the Monte Carlo Simulated Annealing (MCSA) technique to compute the target activities of each node in G that are required to achieve the user-defined therapeutic goal, *independent* of other nodes, by perturbing the network parameters. In the efficacy-conscious simulated annealing phase, we utilize the modified G for evaluating the effects of candidate target combinations where the individual target activities are exploited to judiciously guide the selection of appropriate target activities in the candidate combinations. Specifically, this phase realizes two subgoals: (a)

[2] In pathogen-related diseases, the target can sometimes be endogenous to the pathogen, instead of the host. In this paper, our focus is on non-pathogen-related diseases.

[3] Modeled using ordinary differential equations (ODE).

[4] The modified network behaves no differently from the original network when there is no inhibitor or activator activity.

determining the targets in the combination and (b) determining the type of action (*i.e.,* activation or inhibition) and the strength of the action on the targets. In this phase, we select a set of proteins to be the specific targets of the drugs, plus we select the magnitude and sign of each drug's effect on its target (activation or inhibition to a given extent). This selection process (*target selection process*) must ensure that the desired therapeutic effect is achieved while minimizing off-target effects. The selection process (or the subgoals) provide the constraints to be satisfied and the off-target effects provide the optimization problem. Hence, the problem can be modeled as the optimization of a *constraint satisfaction problem* (CSP), which is NP-hard [23]. Due to the hardness of this problem, we deploy a machine learning-driven, simulated annealing-based heuristic solution to realize this phase. Specifically, we first utilize a machine learning-based target prioritization technique [17] to rank target nodes in G for their ability to impact the disease node by considering network topology and network dynamics. Then, higher ranked nodes are *selected* with higher probability for generation of a candidate target combination. Next, MASCOT utilizes the *Loewe additivity theory* [67] from pharmacology to favor the use of more synergistic combinations. Note that synergism implies that lower activity is required to achieve a desired therapeutic goal which in turn is likely to reduce off-target effects. Observe that the target selection process reduces the search space dramatically by filtering antagonistic combinations as well as combinations comprising nodes (targets) in G that potentially do not influence the disease node sufficiently. After MASCOT generates a candidate target combination, it uses an ODE solver (*e.g.,* Copasi [54]) to simulate the candidate combination effect. The resultant concentration-time curves of various nodes in G are then used to compute the therapeutic effect on the disease node and the associated off-target effects. A candidate target combination is accepted if it (a) achieves desired therapeutic goal and has fewer off-target effects or (b) achieves the desired therapeutic goal and satisfies the Metropolis criterion. Lastly, selected target combinations are added to the solution set and ranked based on their off-target effects.

In summary, this paper makes the following contributions:

- We present a *data-driven target combination prediction problem* that aims to predict synergistic target combinations in a disease-related signaling network that can achieve the desired therapeutic effect and have minimum off-target effects. Note that the goal of this work is to identify synergistic combinations of targets with reduced off-target effects and excludes the evaluation of drug compounds that bind and regulate the target molecules.
- We present a simulated annealing-based algorithm called MASCOT that addresses this problem by leveraging machine learning-based target prioritization and *Loewe additivity theory* [67] heuristics to reduce the search space significantly and generate superior quality results.
- We conduct a detailed empirical study applying MASCOT to a set of curated signaling networks that have drug-target data associated with them. This study demonstrates the effectiveness and superiority of MASCOT compared to state-of-the-art network-centric target combination prediction techniques.

The rest of the paper is organized as follows. We discuss related work in Section 2. In Section 3 we introduce the concepts and terminologies necessary to understand this paper. We formally define the data-driven target combination prediction problem in Section 4. We present the MASCOT algorithm in Section 5. Experimental study of our proposed approach is discussed in Section 6. The last section concludes the paper.

2. Related work

In current practice, the design of drug combinations (and the selection of target combinations) is rarely automated. State-of-the-art *in silico* methods can broadly be classified into network-centric *optimization-based* and *association-based* approaches. The *optimization-based* approaches are based on sequential decoding (SD) algorithms [10] or Monte Carlo simulated annealing (MCSA) [31,65]. Stochastic search algorithms such as MCSA are expected to perform better than SD for non-linear problems [10]. MCSA modifies a candidate solution stochastically and the proposed modification is accepted or rejected using the Metropolis criterion. Although stochastic candidate modification effectively covers the search space by producing a wide variety of candidates, it has two key limitations when used for identifying target combinations. First, drug targets in real signaling networks influence the therapeutic and off-target effects differently, due to one or more downstream nodes' involvement in other protein–protein interactions. Ignoring this consideration may yield combinations satisfying the user-desired therapeutic effect, but with excessive off-target effects. Note that in [65] a user needs to specify *a priori* specific side effects (as input to the algorithm) in terms of the ratio of concentration of two relevant nodes. Due to the complexity of biological networks, such a strategy is often impractical as it is highly unlikely for a user to know all *system-wide* side effects ahead of time. A second key limitation is although the target activity affects the combination effects, it is chosen randomly in MCSA. A judicious selection process might improve efficiency of the overall process.

In comparison, TIMMA [59], a network-centric *association-based* approach, utilizes a *target inhibition network* that is constructed from functional data on drugs and their targets obtained from target binding assays and high-throughput drug sensitivity screens. Although TIMMA yields synergistic target candidates that are druggable, it relies on drug treatment data that may not always be publicly available. Moreover, it restricts the target combinations to those targets which are associated with existing drugs. That is, novel targets that are not hit by existing drugs but are beneficial in treating the disease cannot be discovered. In addition, because TIMMA does not include potential off-target effects of target combinations in its analysis, the predicted combinations may be toxic.

Unlike TIMMA, which is designed specifically for cancer, MASCOT is a generic and extensible technique. In addition, it addresses the above-mentioned limitations of optimization-based approaches. In particular, MASCOT utilizes an existing machine learning-based *target prioritization* technique [17], which prioritizes potential targets in a given disease-related network, to select more effective targets in order to reduce off-target effects. It also exploits *Loewe additivity theory* (LOEWE) [67] to assess the interaction effect (*i.e.*, synergism, antagonism) in a combination. LOEWE can then be used to prune the target activity search space, reducing computational cost and ensuring that targets selected are synergistic.

A preliminary version of our work is described in [15] where we propose an algorithm called STEROID for synergistic target combination discovery from curated signaling networks. Although STEROID exploits Loewe additivity theory, the target prioritization technique it utilizes to select targets for a candidate target combination is not driven by machine learning. As we shall see later, this results in worse results compared to MASCOT. Furthermore, in this paper we demonstrate the generic nature of MASCOT by predicting target combinations in two different curated signaling networks instead of a single network as studied in [15].

3. Background

In this section, we briefly describe the graphical representation of the curated signaling networks used in this study. Next, we provide an overview of TAPESTRY [17], a machine learning-based target prioritization framework and the *Loewe additivity theory* (LOEWE) [67] on which the MASCOT algorithm is built. In the rest of the paper, we shall use the heregulin (HRG)-induced MAPK-PI3K signaling network implicated in a variety of cancers (e.g., ovarian [29]) as a running example. Fig. 1 illustrates this network.

3.1. Graph model of signaling networks

A biological signaling network describes the interactions between molecular species involved in the network. Each interaction takes the form of a biochemical reaction. One such reaction is the activation of ERK into phosphorylated ERK (ERKPP) by its kinase, phosphorylated MEK in Fig. 1. Graphically, this reaction is typically represented as a directed hyperedge connecting one set of nodes (*e.g.*, {ERK and MEK}) to another set (*e.g.*, {ERKPP}) [34]. Hence, a signaling network is naturally represented as a directed hypergraph $G = (V, E)$. Analysis of directed hypergraphs is generally more complex than for graphs and many graph algorithms cannot be used directly on hypergraphs. Hence, hypergraphs are often transformed into graphs, for example using methods such as bipartite and substrate graph representation [34] (*e.g.,* Fig. 1). In this paper, we use the method in [22] and we chose the bipartite graph representation as it retains the original structural information of the hypergraphs [34]. Note that the transformed bipartite graph is used to compute the topological features.

In the literature, the most common formalisms for signaling network modeling approaches for combination therapy are, namely Bayesian network, logic-based network, and ordinary differential equation (ODE) model [3]. In this paper, we adopt the ordinary differential equation (ODE) model. In this model, each reaction (edge) in a signaling network is associated with an ordinary differential equation (ODE). The ODE model describes the system's behaviour over time by using mass-action kinetics[5] for instance, to model the production and consumption rates of different molecular species [2]. These models are typically constructed by translating prior knowledge of production and consumption rate of different molecular species into differential equations. For example, the ODE model of Fig. 1 can be found in *Biomodels* [47] (BIOMD0000000146). Note that the determination of these reaction kinetics can be technically challenging. Hence, a large proportion of these kinetics are usually estimated using parameter estimation techniques [49]. Despite this uncertainty, these under-determined ODE systems can still model real, observable biological behaviour, providing valuable means for quantitative study.

Lastly, a *disease node* in a signaling network is a molecule that is either involved in some dysregulated biological processes implicated in a disease, or is of interest due to its potential role in the disease. An example of a disease node in the MAPK-PI3K network (Fig. 1) is phosphorylated ERK (ERKPP) [58].

3.2. Machine learning-based target prioritization

Target prioritization is the process of ranking targets according to some criteria such as sensitivity or gene expression level. It is potentially useful in helping to plan experiments since resources are limited and experiments can be costly and time-intensive [43].

In this work, we utilize TAPESTRY [17], a network-centric, machine learning-based approach that prioritizes targets in signaling networks with respect to a disease node using both topological and dynamic features of the network. An unique aspect of TAPESTRY is its ability to leverage knowledge gained from learning of training networks with known targets to first identify predictive features that characterize targets and then use these predictive features to identify targets of the given network with unknown targets. Specifically, it is built on top of TENET, a recently proposed *target characterization*[6] technique [16]. TENET deploys a support vector machine (SVM)-based strategy to learn offline the optimal set of *predictive topological features* for characterizing *known* curated targets in a set of publicly-available signaling networks (referred to as *candidate networks*) and generates a set of *characterization models* based on these features. Then given a disease-related signaling network with unknown targets (referred to as *unseen network*) and the set of characterization models generated by TENET, TAPESTRY prioritizes its nodes (ranks nodes based on topological and dynamic criteria) with respect to a disease node by leveraging a characterization model and network dynamics. Specifically, it selects the "best" characterization model it should adopt as its *prioritization model* from the collection of characterization models of the candidate networks. A *prioritization score* (referred to as *putative target score*) is then derived from the selected model and the dynamics of the unseen network, and used to prioritize candidate targets. Note that candidate targets are those that are upstream of the disease node [60].

Remark. The goal of target prioritization is to rank individual target nodes and hence is different from our goal to identify synergistic combinations of targets with reduced off-target effects.

3.3. Loewe additivity theory (LOEWE)

LOEWE computes the *combination index* as a measure of the interaction effect between the drugs in a combination. The foundation and assumptions of this method have been documented extensively [67]. Given a set of drugs X and therapeutic effect T, let D_x and d_x be the doses of drug $x \in X$ required to achieve effect T when used alone and in combination, respectively. Then, the *combination index* is defined as

$$CI = \sum_{x \in X} \frac{d_x}{D_x}. \quad (1)$$

The combination is *synergistic*, *additive* or *antagonistic* if $CI < 1, CI = 1$ or $CI > 1$, respectively. The isobologram (Fig. 2) provides a visual interpretation of LOEWE. It is a graph with the individual drug doses (D_1 and D_2) as its axes. The "line of additivity" is used to interpret the drug interaction. Synergistic and antagonistic combinations are represented by drug doses that fall below and above the line of additivity, respectively [67]. As we shall see later in Section 5.2, this theory can be adapted to guide selection of synergistic targets.

4. Target combination prediction problem

In this section, we formally define the problem of *target combination discovery*. We begin by introducing several concepts related to *drug target*.

[5] The law of mass-action states that for an elementary reaction where all the stoichiometric coefficients of the reactants are one, the rate of reaction is proportional to the concentrations of the reactants.

[6] *Target characterization* is the process of defining the characteristics of the targets and is useful in drug design for these targets and in the identification of novel targets that share similar characteristics with known targets.

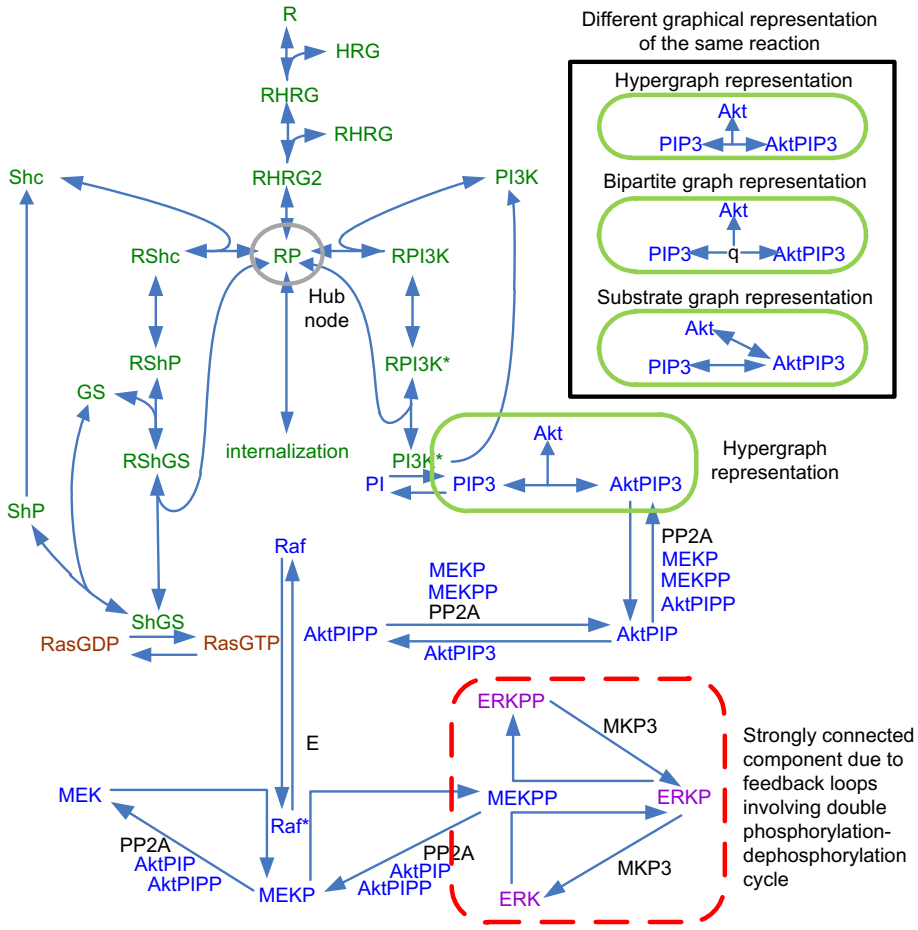

Fig. 1. MAPK-PI3K signaling cascade [29].

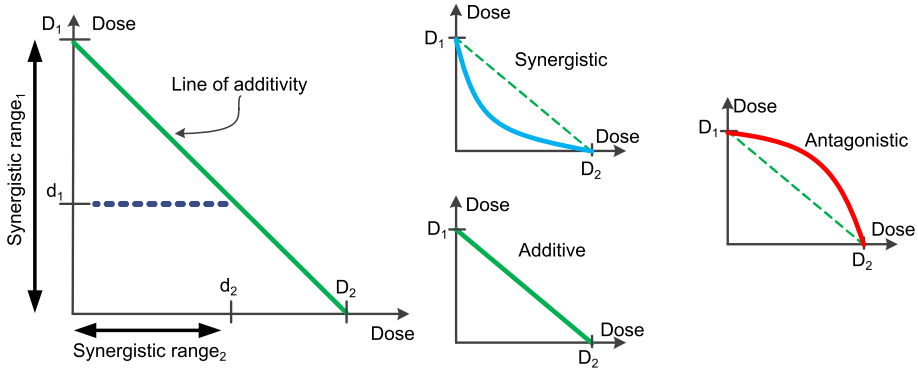

Fig. 2. Isobologram. D_1 and D_2 are the dose of each drug that achieves the desired therapeutic effect if administered alone. d_1 and d_2 together can achieve the same effect.

4.1. Drug target and target activity

First, we present the concept of a *target* and its *activity* (referred to as *target activity*) in the context of signaling networks. A drug asserts its effect on a network through the *target* and the *target activity* is a variable related to the extent of *target* perturbation. The drug effect is typically modeled *in silico* as modulation of the node concentration. The modulation is achieved by modifying a network parameter that controls the concentration of the node associated with the *target*. This parameter can either be the node's edges (typically represented as ODE reactions) [65] or the node itself (initial concentration) depending on whether the node concentration varies with time. We now formally define these two concepts. We first introduce the notion of *reactant-product edge set* to facilitate exposition. Given a signaling network $G = (V, E)$ and a node $u \in V$, the *reactant-product edge set* of u is defined as $\text{RPE}_u = R_u \bigcup P_u$ where $R_u \subset E$ and $P_u \subset E$ are the edge sets involving u as reactants and products, respectively.

Definition 1. Given a signaling network $G = (V, E)$, and node $u \in V$ with concentration time-series profile φ_u and reactant-product edge set RPE_u, the **drug target** of a node u is $c_{fix} = u$ if φ_u is constant, and it is $c_{var} \in \text{RPE}_u$ otherwise.

Definition 2. Given a drug target c perturbed by drug D with dissociation constant K_D, the **target activity** of c is defined as $\Gamma_c = \frac{[D]}{K_D}$ where $[D]$ is the concentration of D.

The ODE modification varies according to the drug type (*e.g.*, activators or inhibitors) and the mechanism of action. Note that there are several different types of activators and inhibitors. In this paper, we model activation using *nonessential activation* [13], and inhibition using *competitive inhibition* [65] for the following reasons. *Nonessential activator* affects the rate of reaction but does not stop the reaction from happening when it is absent. Since drugs are exogenous to the biological system, we expect them to affect the system only when they are present. *Competitive inhibitors*, on the other hand, is the most common type of inhibition. We assume that a target is druggable by both activators and inhibitors due to the lack of readily available information on the type of action (activation or inhibition) that is valid for each target. Note that violation of this assumption will result in an invalid target combination. However, since MASCOT returns a set of acceptable solutions instead of a single solution, invalid solutions can be discarded in favor of the next valid solution. Note that each reversible reaction is made up of an equivalent pair of irreversible reactions (forward and backward reactions). We explicitly convert all reversible reactions to irreversible reactions using [54] in order to clearly distinguish whether it is the forward or the backward reactions that is activated or inhibited.

Given two nodes u and v, an inhibitor I and an activator A where u and v have constant and variable concentration time-series profiles, respectively, let RPE_v be the reactant-product edge set of v. Then, the targets of u and v, denoted as c_{fix} and c_{var}, respectively, are defined as

$$c_{fix} = u \tag{2}$$

$$c_{var} = \frac{V_{max}[S]}{K_m + [S]} \in \text{RPE}_v \tag{3}$$

where V_{max} is the maximum velocity; K_m is the Michaelis–Menten constant; and $[S]$ is the concentration of the substrate S.

The *competitive inhibition* of c_{fix} and c_{var} are given by the following equations:

$$\mathcal{I}(c_{fix}) = \frac{[u]_0}{\frac{[I]}{K_I}} \tag{4}$$

$$\mathcal{I}(c_{var}) = \frac{V_{max}[S]}{K_m(1 + \frac{[I]}{K_I}) + [S]} \tag{5}$$

In the above equations, $[u]_0$ is the initial concentration of u and K_I is the dissociation constant of I. Similarly, let K_A be the dissociation constant of A. The *nonessential activation* of c_{fix} and c_{var} are defined as follows.

$$\mathcal{A}(c_{fix}) = \frac{[A]}{K_A}[u]_0 \tag{6}$$

$$\mathcal{A}(c_{var}) = \frac{V_{max}[S](1 + \frac{[A]}{K_A})}{K_m + [S]} \tag{7}$$

For example, the reaction PIP3 + Akt⟶AktPIP3 whose original ODE is $k[\text{PIP3}][\text{Akt}]$ becomes $k[\text{PIP3}][\text{Akt}](1 + \frac{[A]}{K_{[A]}})$ and $\frac{k[\text{PIP3}][\text{Akt}]}{1+\frac{[I]}{K_I}}$ when modified to simulate nonessential activation and competitive inhibition, respectively.

4.2. Target effects

Next, we formally define the notions of *therapeutic effect* and *off-target effects*. Given a signaling network $G = (V, E)$, a drug target c and the desired therapeutic effect ς_{th}, let $u \in V$ be the node (disease node) associated with effect ς_{th}, and $\text{AUC}(\varphi_u^-)$ and $\text{AUC}(\varphi_u^+)$ be the area under the concentration–time series profile curve of node u before and after c is perturbed, respectively. Then, the *therapeutic effect* $\varsigma_{th(c)}$ and *off-target effects* $\varsigma_{off(c)}$ of c are given by the following equations.

$$\varsigma_{th(c)} = \frac{|\text{AUC}(\varphi_u^-) - \text{AUC}(\varphi_u^+)|}{\text{AUC}(\varphi_u^-)} \tag{8}$$

$$\varsigma_{off(c)} = \sum_{v \in V \setminus u} \left(\frac{|\text{AUC}(\varphi_v^-) - \text{AUC}(\varphi_v^+)|}{\text{AUC}(\varphi_v^-)} \right) \tag{9}$$

Note that $\varsigma_{th(c)}$ and $\varsigma_{off(c)}$ can be determined from *in silico* simulation using *Copasi* [54]. The combination effects are defined similarly whereas $\text{AUC}(\varphi_u^+)$ and $\text{AUC}(\varphi_u^-)$ can be estimated using the linear trapezoidal rule method [12].

For example, the therapeutic effect of Akt is given as $\varsigma_{th(\text{Akt})} = \frac{|\text{AUC}(\varphi^-_{\text{ERKPP}}) - \text{AUC}(\varphi^+_{\text{ERKPP}})|}{\text{AUC}(\varphi^-_{\text{ERKPP}})}$ whereas the off-target effects is given as $\varsigma_{off(\text{Akt})} = \sum_{v \in V \setminus \text{ERKPP}} \left(\frac{|\text{AUC}(\varphi_v^-) - \text{AUC}(\varphi_v^+)|}{\text{AUC}(\varphi_v^-)} \right)$. Note that in practice, the therapeutic effect is dependent on the stage of the disease and is typically measured as inhibition of certain phenotypic response (*e.g.*, cell growth) which may not be linearly correlated with the inhibition of the disease node concentration.

4.3. Problem definition

The goal of the *target combination prediction problem* is to identify targets and their activities that achieve a user-specified therapeutic effect (*e.g.*, to achieve 50% inhibition of ERKPP) while minimizing the off-target effects. Hence, the problem can be modeled as the optimization of a *constraint satisfaction problem* (CSP) which is NP-hard [23]. The CSP is represented as a triple (X, D, C), where X, D and C represent the set of variables, the variables' domain and the set of constraints, respectively. The element X represents the set of drug targets and target activities; D represents the set of candidate targets in a given disease-related network and the target activity range; and C represents the condition that the combination therapeutic effect must match the desired therapeutic effect.

Definition 3. Given a set of target combination $C = \{C_1, \cdots, C_N\}$ and a desired therapeutic effect ς_{th}, let $C_i = \{c_1, \cdots, c_m\}$ where $c_j \in C_i$ is the j^{th} target in the i^{th} combination. Let $\varsigma_{off(C_i)}$ and $\varsigma_{th(C_i)}$ be the off-target effects and therapeutic effect of combination C_i, respectively. Then, the **target combination prediction problem** is defined as

$$C_i = min\{\varsigma_{off(C_i)} | \varsigma_{th(C_i)} = \varsigma_{th}\}$$

5. Predicting target combinations

In this section, we begin by providing the rationale behind the design of MASCOT. Then, we present the two practical heuristics that we shall exploit for modifying candidate solutions. Finally, we describe the MASCOT algorithm.

5.1. Rationale behind the design of MASCOT

MASCOT integrates simulated annealing with machine learning-based target prioritization and LOEWE from pharmacology for

predicting target combinations. Briefly, it consists of two phases, namely, *preprocessing* and *efficacy-conscious simulated annealing*. In the *preprocessing* phase, all reversible reactions in the network are converted to equivalent irreversible pairs of reactions and then the reactions are modified to simulate actions of nonessential activators and competitive inhibitors. Recall from Section 4.1, such preprocessing allow us to disambiguate the reaction that shall be perturbed and the type of perturbation required for our predicted target combination solutions. Note that clarity of our target combination solutions can facilitate drug design activity that follows by pinpointing the exact reactions that the drugs should target and how the drugs should affect the reactions.

In the *efficacy-conscious simulated annealing*, *target prioritization heuristic* and LOEWE *heuristic* are used for guiding the selection of candidate target combination solutions without exhaustively exploring all possible combinations. In particular, target prioritization heuristic leverages on target prioritization rank to select for more effective targets for the combination whereas LOEWE heuristic is used for selecting targets with synergistic activities. Unlike state-of-the-art MCSA-based approaches that generate random candidate solutions, MASCOT uses these heuristics to effectively reduce the solution search space leading to efficient discovery of potential solutions. Note that the heuristics are chosen based on their relevance to efficacious therapies. Specifically, we advocate that preferentially selecting for prioritized targets (using a machine learning-based target prioritization technique) in the combinations can lead to more effective target combinations. In addition, LOEWE ensures synergistic interaction of targets. Note that in MASCOT, candidate solutions have to satisfy the desired therapeutic conditions and are optimized for off-target effects. Instead of choosing specific off-target effects (*e.g.*, ratio between two proteins), we consider off-target effects in a general sense. That is, deviation of activities of all nodes except the disease node before and after perturbation of the target combination. The rationale behind this is that importance of specific off-target effects are dependent on signaling networks and their associated disease and users may not know at the onset which particular off-target effects are more important.

5.2. Heuristics

Target prioritization heuristic. The goal of using the target prioritization heuristic is to improve the average solution quality by choosing more effective targets with higher probability, thereby minimizing off-target effects. To achieve this, we first translate the prioritization rank (generated by TAPESTRY) to an equivalent *target rank*, then convert the rank to a *selection probability* value which is used to decide if the target will be accepted. We now introduce these two concepts.

Given a signaling network $G = (V, E)$ and a target prioritization method P, let c_{fix} and c_{var} be the targets of nodes u and v, respectively where u and v have constant and variable concentration–time series, respectively, and $u, v \in V$. Then, the *target ranks* of c_{fix} and c_{var}, denoted as $\Psi_{c_{fix}}$ and $\Psi_{c_{var}}$, respectively, are defined as

$$\Psi_{c_{fix}} = \Psi_{P:u} \tag{10}$$

$$\Psi_{c_{var}} = \sum_{w \in W} \Psi_{P:w} \tag{11}$$

where $\Psi_{P:u}$ is the rank of u based on $P, W = X \bigcup Y, X, Y \subset V$, and $c_{var} = (X, Y)$.

The *selection probability* (*sp*) of a target is the likelihood of selecting it. We use the rank-based fitness function to obtain a target's selection probability. The fitness function is based on the individual target ranks and avoids scaling problems associated with using actual objective values. It is defined as

$$sp = \frac{2 - \lambda^+}{|\mathcal{T}|(2 - \lambda^+) + 2(\lambda^+ - 1)} \tag{12}$$

where $\mathcal{T}$ is the set of individual targets in the signaling network and λ^+ is the parameter (called selective pressure) used to control the expected sampling rate of the individual target. Note that λ^+ is typically in the range [1,2] [6].

Observe that the aforementioned heuristic is independent of any specific target prioritization method. However, as we shall see in Section 6, target combination prediction using machine learning-based TAPESTRY typically generates superior quality results compared to non-machine learning-based techniques.

LOEWE heuristic. The effects (recall from Section 4.2) resulting from a drug combination can be interpreted as drugs at particular dosages hitting their targets. This produces certain target activities causing a particular response of the network. Hence, an interaction of multiple targets in a combination can be assessed the same way as drug interactions by replacing the drug doses with target activities. A target combination is guaranteed to be synergistic if its target activities are chosen from values below the line of additivity. Following from Section 3.3, we define the *target interaction* as follows.

Definition 4. Given a therapeutic effect ς_{th} and a target combination $C = \{c_1, \cdots, c_m\}$, let $\Gamma_{0(c_i)}$ and $\Gamma_{(c_i)}$ be the target activities of the i^{th} target in C that achieve ς_{th} when targeted alone and in combination, respectively. Then, the **target combination index** of C is defined as

$$\text{TCI}_C = \sum_{c_i \in C} \frac{\Gamma_{(c_i)}}{\Gamma_{0(c_i)}} \tag{13}$$

The combination is **synergistic**, **additive** or **antagonistic** if $\text{TCI}_C < 1, \text{TCI}_C = 1$ or $\text{TCI}_C > 1$, respectively.

Following from Definition 4, for a 2-target combination $C = \{c1, c2\}$, the *synergistic ranges* of targets $c1$ and $c2$, denoted as sr_{c1} and sr_{c2}, respectively are defined as $sr_{c1} = [0 - \Gamma_{0(c1)})$ and $sr_{c2} = [0 - \Gamma_{(c2)})$, where $\Gamma_{(c2)} \in [0 - \Gamma_{0(c2)})$ and TCI < 1. Graphically, these synergistic ranges can be visualized in Fig. 2 (rightmost isobologram) as "synergistic range$_1$" and "synergistic range$_2$".

5.3. The algorithm MASCOT

For solving CSPs, systematic algorithms (*e.g.*, backtracking) have been proposed that rely on partial instantiation of the candidate solutions to eliminate candidates that violate the constraints. In the target combination prediction problem, full instantiation of the candidate solution is needed to find the combination effects. Hence, these systematic algorithms cannot be applied effectively. Metaheuristics (e.g., simulated annealing (SA)) are used instead to find approximate solutions as they can achieve good performance results for large combinatorial optimization problems [45]. MCSA (a variant of SA) has been proposed for finding drug target combinations [65,31], but suffer from certain limitations as highlighted in Section 2. We shall now present an algorithm called MASCOT (outlined in Algorithm 1) that addresses these limitations by leveraging on machine learning-based target prioritization and LOEWE heuristics for modifying drug target and target activity of candidate solutions.

Algorithm 1: Algorithm MASCOT

Input: Signaling network G, set of prioritized node rank Ψ, therapeutic effect ς_{th} and combination size $\mathcal{S}$.
Output: Solution set $\mathcal{R}$.
1 $\mathcal{R} \leftarrow$INITIALIZE($\mathcal{R}$)
2 (λ^+,ϵ_t,ϵ_a,N,t_0,i_{max})$\leftarrow$SETTODEFAULTS(λ^+,ϵ_t,ϵ_a,N,t_0,i_{max})
3 $t \leftarrow t_0$
4 (G',$\mathcal{T}$,Γ_0)$\leftarrow$PREPROCESSINPUT(G,ς_{th},ϵ_t,t_0,i_{max}) /*Phase 1*/
5 **while** $t \geq 0$ *and* $|\mathcal{R}| \leq N$ **do**
6 **foreach** *iteration* i=*1 to* i_{max} **do**
7 $\mathcal{X} \leftarrow$GETCOMBI($\mathcal{T}$,λ^+,Ψ,ϵ_a,Γ_0,$\mathcal{R}$,$\mathcal{S}$) /*Phase 2.1*/
8 ($\varsigma_{th(\mathcal{X})}$,$\varsigma_{off(\mathcal{X})}$)$\leftarrow$GETEFFECT($G'$,$\varsigma_{th}$,$\mathcal{X}$) /*Phase 2.2*/
9 $\mathcal{R} \leftarrow$ACCEPTCOMBI($\varsigma_{th(\mathcal{X})}$,$\varsigma_{off(\mathcal{X})}$,$\varsigma_{th}$,$\epsilon_a$,$\epsilon_t$,$t$,$\mathcal{R}$) /*Phase 2.3*/
10 $t \leftarrow t-1$

Given a signaling network G, a set of prioritized node rank Ψ generated by TAPESTRY, a desired therapeutic effect ς_{th} and the required combination size $\mathcal{S}$, MASCOT identifies a set of synergistic target combinations $\mathcal{R}$ which satisfies ς_{th} and has minimal off-target effects ς_{off}. The inputs G and Ψ are used to modify the drug targets and target activities. In addition, G is also used to simulate the target combination effects. Several other parameters ($\lambda^+, \epsilon_t, \epsilon_a, N, t_0$ and i_{max}) that are required by MASCOT are set to default values, but can be modified if required (Line 2). The parameter λ^+ is used to compute the selection probability of the target. In practice, it is difficult to achieve the therapeutic effect exactly, and additive target combinations are generally close to the line of additivity without lying exactly on the line. Hence, we specify adjustment factor parameters ϵ_t and ϵ_a to relax the condition for therapeutic effect and additive combination into bound conditions, respectively (*e.g.*, 49.5%–50.5% inhibition of ERKPP and additive if $0.95 \leqslant$ TCI $\leqslant 1.05$). Finally, the parameters N, t_0 and i_{max} are used to configure the SA and they control when the SA terminates: when N solutions are found or when $t_0 \times i_{max}$ iterations are completed.

Algorithm 2: The PREPROCESSINPUT Procedure (Phase 1)

Input: Signaling network $G = (V, E)$, therapeutic effect ς_{th}, adjustment factor for therapeutic effect ϵ_t, initial temperature t_0 and maximum number of iterations per temperature cycle i_{max}.
Output: Modified signaling network G', set of candidate target $\mathcal{T}$ and set of individual target activity Γ_0.
1 ($\mathcal{T}, \Gamma_0$) $\leftarrow$INITIALIZE($\mathcal{T}, \Gamma_0$)
2 **foreach** $e \in E$ **do**
3 **if** ISREVERSIBLE(e)=*true* **then**
4 ($e_{forward}, e_{backward}$) $\leftarrow$GETIRREVERSIBLERXNPAIR(e)
5 $G \leftarrow$REPLACEREVREACTION($e_{forward}, e_{backward}, e, G$)
6 **foreach** $v \in V$ **do**
7 **if** HASCONSTANTTIMESERIESPROFILE(v)=*true* **then**
8 $\mathcal{T} \leftarrow \mathcal{T} \bigcup v$
9 **else**
10 RPE $\leftarrow$GETREACTANTPRODUCTEDGESET(v)
11 $\mathcal{T} \leftarrow \mathcal{T} \bigcup$ RPE
12 **foreach** $c \in \mathcal{T}$ **do**
13 $G' \leftarrow$MODIFYTARGETREACTION(G, c)
14 $\Gamma_{0(c)} \leftarrow$GETITA($G, c, \varsigma_{th}, \epsilon_t, t_0, i_{max}$)
15 $\Gamma_0 \leftarrow \Gamma_0 \bigcup \Gamma_{0(c)}$

Phase 1: Preprocessing. In this phase (Algorithm 2), the reversible reactions in G are first converted into pairs of irreversible reactions using [54] (Lines 4–5). Then, the drug targets c_{fix} (Line 8) and c_{var} (Lines 10–11) of nodes with constant and variable concentration–time series profiles, respectively, are found based on Definition 1. Next, the reactions are modified to simulate the effects of the targets when modulated by non-competitive inhibitors and essential activators (Line 13) according to Eqs. (4)–(7). Finally, the individual target activities (ITA) Γ_0 required to achieve the desired therapeutic effect (*e.g.*, 50% down-regulation of ERKPP, ϵ_t=5%) are found using MCSA configured with the parameters t_0 and i_{max}. Targets that cannot achieve the desired therapeutic effect alone (*i.e.*, complete the maximum number of iterations without finding any target activity that can achieve the desired therapeutic effect) are deemed to have $\Gamma_0 = \infty$.

Phase 2: Efficacy-conscious simulated annealing (ESA). The ESA consists of three subphases which are repeated until either the temperature t reaches zero or the required number of solutions N is found (Line 5). The subphases consist of *target combination generation* (Line 7), *combination effect calculation* (Line 8) and *candidate acceptance test* (Line 9).

Algorithm 3: The GETCOMBI Procedure (Phase 2.1)

Input: Set of candidate target $\mathcal{T}$, selective pressure λ^+, set of prioritized node rank Ψ, adjustment factor for target interaction ϵ_a, set of individual target activity Γ_0, solution set $\mathcal{R}$ and combination size $\mathcal{S}$.
Output: Combination candidate $\mathcal{X} = \{(x_1, \Gamma_1), \cdots, (x_S, \Gamma_S)\}$.
1 $\mathcal{X} \leftarrow \phi$
2 **foreach** *combination candidate component* (x_i, Γ_i)=(x_1, Γ_1) *to* (x_S, Γ_S) **do**
3 **while** ISNULL(x_i) *is* TRUE **do**
4 $\mathcal{A} \leftarrow$SELECTRANDOMTARGET($\mathcal{T}/\mathcal{X}$,$\mathcal{R}$)
5 $x_i \leftarrow$ACCEPTTARGET($\mathcal{A}$,$\mathcal{T}/\mathcal{X}$,Ψ,λ^+)
6 $\Gamma_i \leftarrow$SELECTACTIVITY(x_i,$\mathcal{X}$,Γ_0,ϵ_a,$\mathcal{R}$)
7 $\mathcal{X} \leftarrow \mathcal{X} \bigcup (x_i, \Gamma_i)$

In the GETCOMBI procedure (Algorithm 3), the candidate combination $\mathcal{X}$ consisting of $\mathcal{S}$-target is generated. Lines 3–5 implement the target prioritization- and Line 6 the LOEWE heuristics. The first target $\mathcal{A}$ is randomly selected using SELECTRANDOMTARGET (Line 4) and accepted in ACCEPTTARGET (Line 5) if the probability of selecting $\mathcal{A}$ (selection probability) is greater than a random number in the range [0–1] (*i.e.*, $sp_{\mathcal{A}} >$ RAND$(0,1)$ where RAND$(\cdot)$ is the random operator). Its activity is then selected within the synergistic range (Definition 4) using the SELECTACTIVITY procedure (Line 6). Similar steps are repeated to find subsequent targets and their activities.

Next, the GETEFFECT procedure obtains the therapeutic and off-target effects by first simulating the candidate solution using *Copasi* [54] and then calculating the therapeutic effect and the off-target effects (Section 4.2). Finally, these effects are used to assess the candidate in ACCEPTCOMBI (Algorithm 4) using the Metropolis criterion. A candidate is accepted if it satisfies any one of these conditions:

1. If it is synergistic, achieves the required therapeutic effect and has off-target effects lower than the current solution (*curr*) (Line 2).
2. If it achieves the required therapeutic effect and $e^{-\frac{\varsigma_{off(\mathcal{X})} - \varsigma_{off(curr)}}{t}} \geqslant$ RAND$(0,1)$ (Line 4).

The UPDATESOLUTION procedure updates the solution set and the current solution with $\mathcal{X}$ if the candidate is accepted.

Algorithm 4: The ACCEPTCOMBI Procedure (Phase 2.3)

Input: Combination therapeutic effect $\varsigma_{th(\mathcal{X})}$, combination off-target effects $\varsigma_{off(\mathcal{X})}$, therapeutic effect ς_{th}, adjustment factor for target interaction ϵ_a, adjustment factor for therapeutic effect ϵ_t, temperature t and solution set $\mathcal{R}$.
Output: Solution set $\mathcal{R}$.
1 $\mathcal{M} \leftarrow$GETCURRENTSOLUTION($\mathcal{R}$)
2 **if** $\varsigma_{off(\mathcal{X})} \leq \varsigma_{off(\mathcal{M})}$ *and* $\sum_{x \in \mathcal{X}} \frac{\Gamma_{\mathcal{R}}}{\Gamma_{\mathcal{R}0}} < 1 - \epsilon_a$ *and* $\frac{|\varsigma_{th(\mathcal{X})} - \varsigma_{th}|}{\varsigma_{th}} \leq \epsilon_t$ **then**
3 $\mathcal{R} \leftarrow$UPDATESOLUTION($\mathcal{R}$,$\mathcal{X}$)
4 **else if** $\frac{|\varsigma_{th(\mathcal{X})} - \varsigma_{th}|}{\varsigma_{th}} \leq \epsilon_t$ *and* RAND$(0,1) \leq e^{-\frac{\varsigma_{off(\mathcal{X})} - \varsigma_{off(\mathcal{M})}}{t}}$ **then**
5 $\mathcal{R} \leftarrow$UPDATESOLUTION($\mathcal{R}$,$\mathcal{X}$)

For instance, consider a two-target combination. Let backward reaction 29 (denoted as *r29b*), which represents the reaction AktPIP3→PIP3 + Akt, be the first randomly selected target.

$r29b$ will be accepted if its selection probability $sp_{r29b} > \text{RAND}(0,1)$. The activity of $r29b$ is selected from within the range $[0 - \Gamma_{0(r29b)})$, where $\Gamma_{0(r29b)}$ is the activity of $r29b$ alone that is required to achieve 50% down-regulation of ERKPP with ϵ_t=5%. The second target (*e.g.*, $r13(\text{Raf} \rightarrow \text{Raf}^{\star})$) is randomly selected from the set of candidates excluding $r29b$ (*i.e.*, $\mathcal{T}/r29b$), and will be accepted if $sp_{r13} > \text{RAND}(0,1)$. The activity of $r13$ is selected from the range $[0 - \Gamma_{r13})$ where $\frac{\Gamma_{(r29b)}}{\Gamma_{0(r29b)}} + \frac{\Gamma_{(r13)}}{\Gamma_{0(r13)}} < 1 - \epsilon_a$ and $\Gamma_{(r29b)} \in [0 - \Gamma_{0(r29b)})$. The therapeutic and off-target effects of the combination $c = \{(r29b, \Gamma_{(r29b)}), (r13, \Gamma_{(r13)})\}$ are computed as follows:

$$\varsigma_{th(c)} = \frac{|\text{AUC}(\varphi^{-}_{\text{ERKPP}}) - \text{AUC}(\varphi^{+}_{\text{ERKPP}})|}{\text{AUC}(\varphi^{-}_{\text{ERKPP}})}$$

$$\varsigma_{off(c)} = \sum_{v \in V_{\text{MAPK}} \setminus \text{ERKPP}} \left(\frac{|\text{AUC}(\varphi^{-}_{v}) - \text{AUC}(\varphi^{+}_{v})|}{\text{AUC}(\varphi^{-}_{v})} \right)$$

Finally, the Metropolis criterion is used to assess if the combination will be added into the solution set.

Theorem 1. *The worst case time and space complexities of* MASCOT *are* $O(t_0 \times i_{max} \times (|V| + |E|) \times |\varphi|)$ *and* $O(|V|(|E| + |\varphi|))$, *respectively, where* t_0 *is the initial temperature;* i_{max} *is the limit on iterations per cycle;* $|V|$ *and* $|E|$ *are the number of nodes and irreversible reactions, respectively, of the given signaling network; and* $|\varphi|$ *is the number of time points in the concentration time-series profiles used to estimate the target effects. The algorithm converges in finite time on a continuous domain.*

Theorem 2. *The* MASCOT *algorithm converges in finite time.*

The proofs of the above theorems are given in the Appendix.

Remark. Observe that MASCOT can handle large signaling networks efficiently. Consider the time complexity of MASCOT in Theorem 1. In practice, the parameters t_0, i_{max}, and $|\varphi|$ have values ranging from 100 to 500. Hence, the time complexity can be expressed as $O(\alpha \times (|V| + |E|))$ where $\alpha = t_0 \times i_{max} \times |\varphi|$ and tends towards a constant value. Therefore, MASCOT is approximately linear to the size of the input signaling network in practice.

6. Experiments

6.1. Experimental setting

MASCOT is implemented using Java. In addition, we use several publicly available tools and libraries as follows: (a) *lib*SBML library [9] for reading and processing the SBML files of the signaling networks. (b) *Copasi* is a simulation tool for biological networks and offers several functionalities including parameter estimation, sensitivity analysis and steady-state analysis. We use *Copasi* for performing LSA-based target prioritization, and its Java API to simulate the target perturbation effects in STEROID [15] and MASCOT. (c) *NetworkPrioritizer* [33] plugin for *Cytoscape* for performing *Weighted Borda Fuse* (WBF)-based and *Weighted AddScore Fuse* (WASF)-based target prioritization. The experiments are performed on a computer system using a 64-bit operating system with 8GB RAM and a dual core processor running at 3.60 GHz.

Summary of Approaches Studied. Recall that in Section 5.2 we study two heuristics. In particular, the target prioritization heuristic is orthogonal to any specific target prioritization technique. In the comparative study, we investigate different target prioritization approaches, namely, TAPESTRY [17], PANI [14], LSA [28] and *NetworkPrioritizer* [33]. Since *NetworkPrioritizer* provides two prioritization approaches, namely WBF and WASF, we study both approaches. Table 1 summarizes the different approaches studied and the heuristics used in each approach.

Note that modification of the candidate will differ depending on the heuristics used. For MCSA, the targets and activities are selected randomly. For MASCOT-TAPESTRY, MASCOT-PANI, MASCOT-LSA, MASCOT-WBF and MASCOT-WASF, the targets are selected using Algorithm 3 (Lines 3–5) while the activities are selected randomly. For MASCOT-LOEWE, the targets are selected randomly while the activities are selected from within the synergistic range (Algorithm 3, Line 6). For MASCOT, STEROID, MASCOT-LOEWELSA, MASCOT-LOEWEWBF and MASCOT-LOEWEWASF, the targets and activities are selected using Algorithm 3 (Lines 3–6).

Finally, we do not perform comparison with TIMMA [59] as it demands knowledge of target binding assays and high-throughput drug sensitivity screens, which are not publicly-available for the signaling networks that we study.

Datasets. Recall that MASCOT requires signaling networks to be modeled using mass action model. Although many nodes in publicly-available curated signaling networks contain some of that information, few networks have full coverage quite yet. This restricts us to only focus on a small number of curated signaling networks taken from *Biomodels*.[7] Specifically, the MAPK-PI3K network [29] (36 nodes and 34 hyperedges) is used for analysis and the desired therapeutic effect is set to 50% ERKPP down-regulation. We also use the glucose-stimulated insulin secretion network comprising 59 nodes and 45 hyperedges [32] with the therapeutic goal of 25% increase in mitochondrial ATP.[8] Note that it is also not possible to choose large disease-related signaling networks from other sources as the ODE information is not necessarily available for *all* edges in such networks. For example, the largest publicly-available signaling network [21] contains over 6000 nodes. However, it does not contain ODEs.

Performance Metrics. Recall that the goal of MASCOT is to identify target combinations which satisfy the therapeutic effect and have minimized off-target effects. Hence, we assess the different approaches based on the following criteria: (a) off-target effects of target combinations, (b) runtime performance, and (c) number of solutions found.

MASCOT Configuration. Unless otherwise stated, the combination size shall be set to 2 and the following default values shall be used for the rest of the parameters: $\{t_0 = 100,\ i_{max} = 500,\ N = 50,\ \epsilon_t = 5\%,\ \epsilon_a = 5\%,\ \lambda^{+} = 1.8\}$. We used *Copasi* java API for estimating the combination effects and its parameters were set as follows:{*trajectory task* = "Time-Course", *method type* = deterministic, *absolute tolerance* = 1×10^{-12}}. The same configuration is also set for STEROID.

For all tables and discussion that follow, the terms ACT and IN shall denote activators and inhibitors, respectively, whereas forward and backward reactions are marked with superscripts ^{f} and ^{b}, respectively.

6.2. Experimental results

We performed several sets of experiments to compare the different variants of MASCOT against the state-of-the-art approaches. Note that figures in this subsection uses the approach IDs given in Table 1.

Effects of Heuristics. First, we examine the effects of using different heuristics on the MAPK-PI3K network to determine the best

[7] As of April 2017, only 1.27% of the curated networks in *Biomodels* have more than 250 nodes and out of these very few are disease-related networks.

[8] Accumulation of mitochondrial ATP results in activation of insulin granule exocytosis. Hence we set an arbitrary increase of 25% increase in mitochondrial ATP as our therapeutic goal.

Table 1
The different approaches studied. √ indicates the heuristic used by each approach.

Approach	ID	LOEWE	TAPESTRY	PANI	LSA	WBF	WASF
MCSA	1						
MASCOT-LOEWE	2	√					
MASCOT-TAPESTRY	3		√				
MASCOT-PANI	4			√			
MASCOT-LSA	5				√		
MASCOT-WBF	6					√	
MASCOT-WASF	7						√
MASCOT	8	√	√				
STEROID	9	√		√			
MASCOT-LOEWELSA	10	√			√		
MASCOT-LOEWEWBF	11	√				√	
MASCOT-LOEWEWASF	12	√					√

MASCOT variant. We assess different MASCOT variants in terms of off-target effects, actual size[9] of solution sets found and execution time. Fig. 3 summarizes the effects of heuristics. In all our experiments, we use MCSA (Approach 1) as our baseline comparison. We make several observations. First, heuristics can be used to reduce off-target effects. This is illustrated by the boxplot in Fig. 3. The median off-target effects of all MASCOT variants are either the same or lower than that of MCSA. Second, the use of LOEWE heuristic produces fewer outliers, reduces off-target effects and cuts down execution time more effectively than target prioritization heuristic. In addition, compared to approaches that use only target prioritization heuristic, those approaches that use LOEWE heuristic have larger actual solution set sizes. These results suggest that LOEWE heuristic may have a greater role to play in influencing off-target effects, actual solution set size and the execution time required to find target combinations. Third, the use of LOEWE heuristic together with an appropriate target prioritization heuristic may help to reduce the minimum off-target effects. For instance, in Approaches 8 (MASCOT) and 11 (MASCOT-LOEWEWBF), the minimum off-target effects are 1.037 and 1.069, respectively. These minimum values are lower than the corresponding approaches that use only a single heuristic (*i.e.*, MASCOT-LOEWE, MASCOT-TAPESTRY, MASCOT-PANI, MASCOT-LSA, MASCOT-WBF and MASCOT-WASF).

Next, we examine the characteristics of solutions found by using different MASCOT variants. The solutions are characterized based on target interaction (synergistic, additive or antagonistic) and combination type (activators, inhibitors, or mixed activator and inhibitor). We make the following observations from Fig. 4. First, target prioritization heuristic appears to favor identification of antagonistic target combinations. Antagonistic drug combinations are found to be useful in combating drug resistance in various diseases including cancer [5]. Hence, this heuristic may be particularly useful in identifying combinations that can overcome drug resistance. Second, approaches incorporating the LOEWE heuristic tend to yield more solutions that are made up of only inhibitors. Note that development of drugs that inhibit PPI is perceived to be easier than drugs that activate PPI. This is because in the design of drugs that are activators, there is a need to achieve good binding in order to replicate the protein interaction and stimulate increase in target activity [26]. Hence, LOEWE heuristic may also be useful in enriching the solution set with inhibitor-only target combinations.

In particular, Approach 8 (MASCOT) has the lowest median (8.006) and minimum (1.037) off-target effects. In comparison, STEROID performs worse (minimum = 1.073 and median = 12.921). *Hence, machine learning-based* MASCOT *is superior to non-machine learning-based* STEROID.

Biological Relevance of Target Combinations. We shall now examine the biological relevance of the target combinations found using MASCOT and the baseline approach, MCSA.

Target combinations for MAPK-PI3K *network*. First, we examine the ability of both approaches in finding a set of benchmark target combination relevant to ovarian cancer and targeting the MAPK-PI3K network. The benchmark combination set is curated from literature in the *PubMed* repository using "ovarian", "cancer", and "combination" as keywords. Among 5863 *PubMed* records that were returned, we find 3 combinations (Table 2) fitting the criteria. Table 3 shows the targets in the MAPK-PI3K network that match those in Table 2. We examine our solution sets to identify those combinations involving the targets in Table 3. Both solution sets contain one known combination (reported in Table 4). In particular, the target activities and off-target effects of the combination predicted by MASCOT are lower than that predicted by MCSA.[10]

In addition, we observe that a large percentage of solution set found using MASCOT contained at least one target in known drug combinations compared to MCSA (Fig. 5). In particular, the top-10 solutions (Table 5) in MASCOT (top-10 solutions with least off-target effects) are highly enriched with these known targets, particularly the MEK (or ERK) inhibitor. This finding suggests that the MEK inhibitor may be a good target to be used in combinations.

Then, we examine the biological relevance of the top-10 solutions found using both approaches. In the MASCOT solution set, 2 target combinations (combinations ST5 and ST10) involve large target activities (greater than 1000) whereas in MCSA (Table 6), all top-10 solutions (MT1 to MT10) implicate large target activities. Note that large target activity implies either high drug concentration or very small dissociation constant, and is an indication of a high potential in having side effects, especially if the treatment regime requires repeated drug dosing [18]. Hence, MCSA tends to yield target combinations that require large target activities, increasing the potential for side effects. In addition, we perform literature search to look for references that support the usefulness of these combinations. We searched *PubMed* specifically for publications related to these predicted combinations and summarize the findings below.

ST1: Studies reveal that MEK inhibition exhibits the best clinical response in the basal subtype of ovarian cancer where there is no mutation of the oncogenes, RAS and RAF [39]. RAF mutation is common in ovarian cancer and high level of Raf-1 activity was found to correlate with the advanced stage of ovarian cancer [25]. In [36], Li et al. establish the role of Raf kinase inhibitor protein (RKIP) as a metastasis suppressor gene. Based on the above, we reason that combination ST1 consisting of an ERK phosphatase

[9] Note that in the experiments, although we set the required number of solutions to be 50, the actual number of solutions found may be lesser.

[10] STEROID did not predict the known combination. Instead, its predictions can be liken to monotherapy since the resultant combinations are made up of reactions that simulate the behavior of MEK inhibitors.

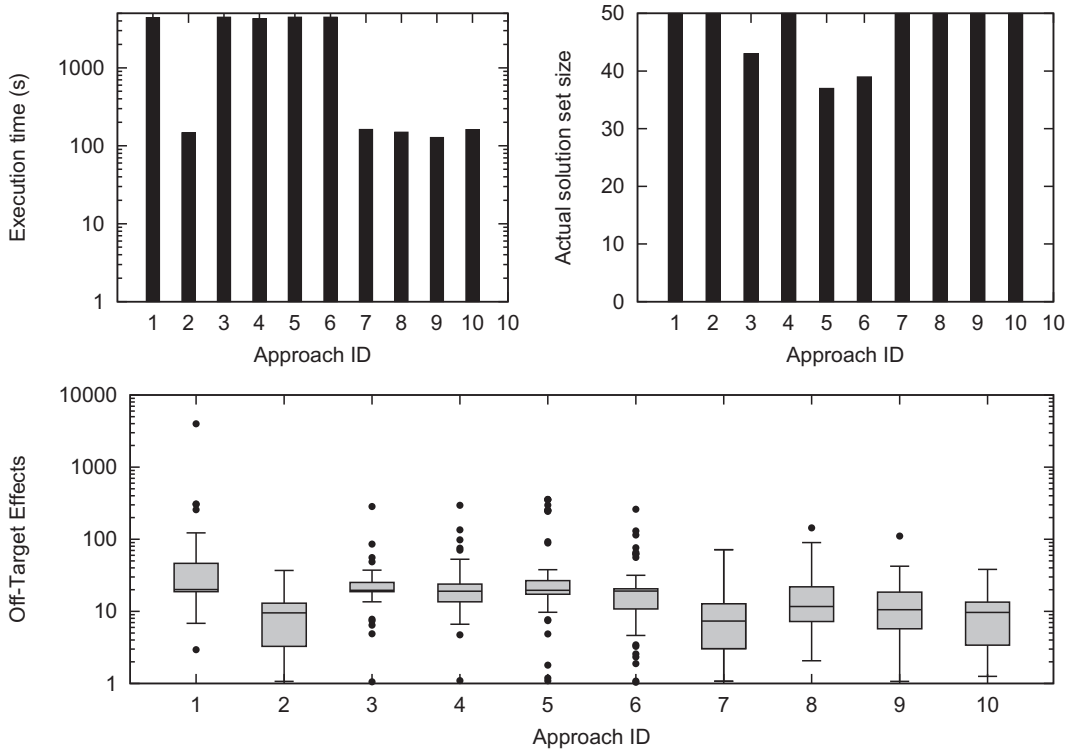

Fig. 3. Effect of heuristics.

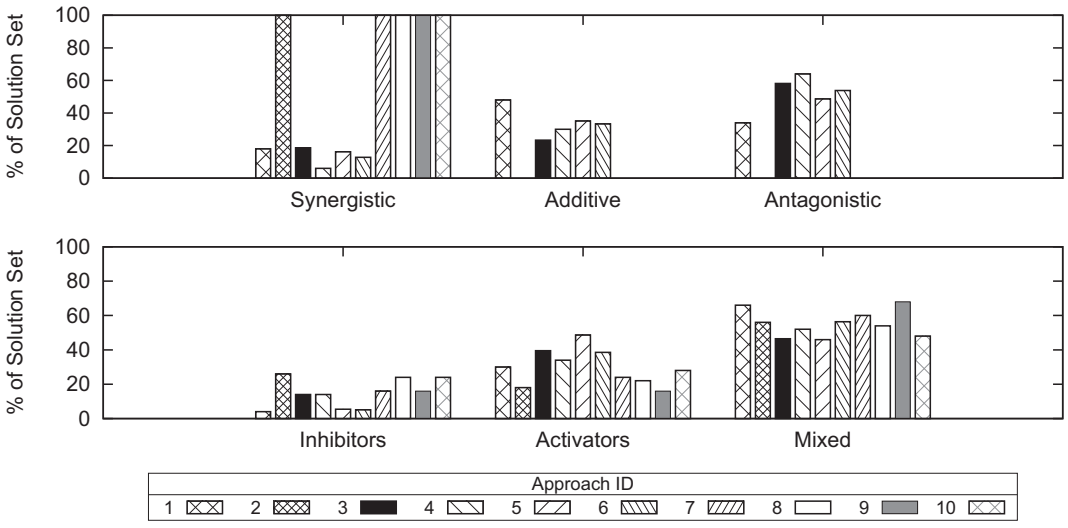

Fig. 4. Characterization of solution sets.

Table 2
PubMed results relevant to ovarian cancer drug combinations targeting the MAPK-PI3K network.

PMID	Target 1	Target 2
22180401, 21062259	Akt inhibitor	MEK inhibitor
21632463	PI3K inhibitor	MEK inhibitor
14675307	PI3K inhibitor	Akt inhibitor

and a Raf inhibitor may be useful in treating the advance stage of ovarian cancer.

ST2 and ST3: Profound growth inhibition and apoptosis were observed in ovarian cancer cells treated with CI-1040, a MEK1/2 inhibitor [48]. These ovarian cells carry mutations in KRAS or BRAF and they typically overexpress DUSP4, an ERK phosphatase [56]. Note that there is currently no known inhibitor that acts directly on ERK, and ERK inhibition is typically achieved through MEK inhibitors [53]. Hence, this correlates with our computational prediction of combinations ST2 and ST3 involving ERK phosphatase activator and ERK (or MEK) kinase inhibitor.

ST7: We did not find any supporting evidence that combination ST7 has been performed, successfully or otherwise, in experiments. However, individual components of these combinations have shown efficacy in ovarian cancer [62,44,66]. Hence, they warrant further investigation as a potential target combinations.

MT4: Corresponds to the known target combination consisting of a MEK inhibitor and an Akt inhibitor [4] in Table 2.

Table 3
Mapping between targets in Table 2 and MAPK-PI3K network.

Target in Table 2	Inhibition Mechanism	Corresponding target(s) in MAPK-PI3K network
Akt inhibitor	Disruption of Akt binding to its membrane localizing factor (PIP3) or dephosphorylation of PIP3	Activators of Reaction 29[b], Reaction 30, Reaction 33; Inhibitors of Reaction 29[f], Reaction 31, Reaction 32
MEK inhibitor	MEK dephosphorylation or blockade of MEK phosphorylation. Is also used to achieve ERK inhibition as there is no known ERK inhibitors.	Activators of Reaction 16, Reaction 18, Reaction 20 or Reaction 22; Inhibitors of Reaction 15, Reaction 17, Reaction 19 or Reaction 21
PI3K inhibitor	Inhibits PI3K in ATP-competitive manner	Activators of Reaction 24[b], Reaction 26; Inhibitor of Reaction 24[f]

Table 4
Known combinations in solution sets.

Approach	Target 1 [Activity]	Target 2 [Activity]	ς_{off}	Combination
MASCOT-LOEWETAPESTRY	Reaction 33 ACT [1633.218]	Reaction 16 ACT [185.982]	11.889	Akt inhibitor + MEK inhibitor
MCSA	Reaction 16 ACT [323.857]	Reaction 31 IN [1640.247]	15.534	MEK inhibitor + Akt inhibitor

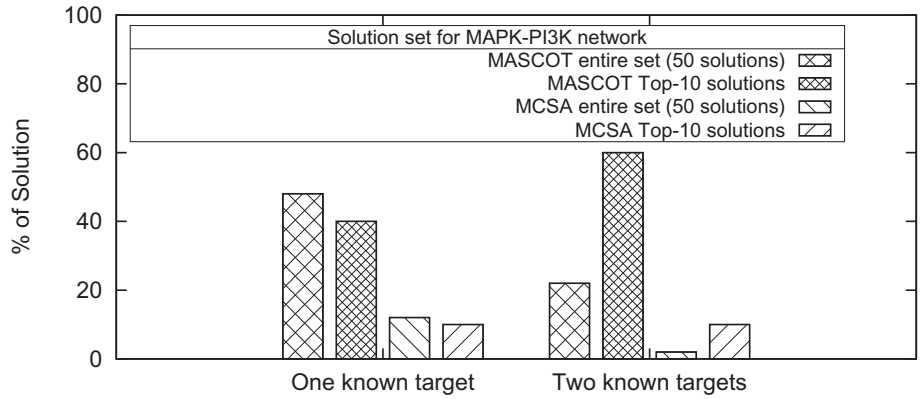

Fig. 5. Percentage of MASCOT and MCSA solution sets with targets in known drug combinations.

Table 5
Top-10 target combinations for MAPK-PI3K network found using MASCOT. Note that targets present in Table 3 are marked with ⋆.

No.	Target 1 [Activity]	Target 2 [Activity]	ς_{off}	Combination	References/Comments
ST1	Reaction 22 ACT⋆ [2.142]	Reaction 14 ACT [0.009]	1.037	ERK phosphatase activator + Raf inhibitor	[39,25,36]
ST2	Reaction 22 ACT⋆ [2.209]	Reaction 15 IN⋆ [0.097]	1.091	ERK phosphatase activator + MEK kinase inhibitor	[48,56]
ST3	Reaction 21 IN⋆ [2.080]	Reaction 22 ACT⋆ [0.203]	1.092	ERK kinase inhibitor + ERK phosphatase activator	[48,56]
ST4	Reaction 21 IN⋆ [2.174]	Reaction 17 IN⋆ [0.068]	1.097	ERK kinase inhibitor + MEK kinase inhibitor	Mono-therapy
ST5	Reaction 22 ACT⋆ [2.150]	Reaction 1[b] IN [7575.148]	1.138	ERK phosphatase activator + Inhibitor of dissociation of heregulin from ErbB4 receptor	–
ST6	Reaction 19 IN⋆ [1.623]	Reaction 21 IN⋆ [0.089]	1.598	ERK kinase inhibitor + ERK kinase inhibitor	Mono-therapy
ST7	Reaction 19 IN⋆ [1.636]	Reaction 12 ACT [0.012]	1.796	ERK kinase inhibitor + Ras inhibitor	[62,44]
ST8	Reaction 18 ACT⋆ [1.299]	Reaction 19 IN⋆ [0.011]	2.411	MEK phosphatase activator + ERK kinase inhibitor	–
ST9	Reaction 17 IN⋆ [1.078]	Reaction 18 ACT⋆ [0.399]	2.458	MEK kinase inhibitor + MEK phosphatase activator	–
ST10	Reaction 18 ACT⋆ [1.288]	Reaction 1[b] IN [7803.570]	2.519	MEK phosphatase activator + Inhibitor of dissociation of heregulin from ErbB4 receptor	-

Note that combinations ST4 and ST6 (resp. MT8 and MT10) are akin to mono-therapies as they involve the same type of drugs (*i.e.*, ERK or MEK kinase inhibitor, and tyrosine kinase inhibitor). In addition, we observe that MCSA produces many combinations (combinations MT1, MT2, MT5, MT6, MT7 and MT9) that involve promoters of known mediators of cancer (*e.g.*, PI3K [19]). Such target combinations, though counter-intuitive, can still achieve the therapeutic goal if the effect of other targets can offset the pro-cancer signals, resulting in an overall anti-cancer signal. However, extra caution should be exercised for these combinations since improper management of the balance between the pro- and anti-cancer signal can easily aggravate the cancer. We did not find any supporting evidence for the remaining combinations (combinations ST8 and ST9). In comparison, only two combinations (MT3 and MT4) of MCSA-based approach has corresponding evidence in the literature. *Hence, compared to* MCSA, MASCOT *is able to identify target combinations that are biologically more relevant.*

Next, we examine the bottom-10 solutions. Table 7 reports the list based on our proposed method (corresponding list based on MCSA-based approach is reported in Table 8). We note that 60% of MASCOT solutions and 100% of MCSA solutions have at least one target with activity greater than 1000. Hence, compared to the top-10 solutions, the bottom-10 solutions have more combinations with large target activities. This implies that combinations with smaller target activities are likely to produce combinations with smaller

Table 6
Top-10 target combinations for MAPK-PI3K network found using MCSA. Note that targets present in Table 3 are marked with ★.

No.	Target 1 [Activity]	Target 2 [Activity]	ς_{off}	Combination	References/Comments
MT1	Reaction 13 ACT [1294.274]	Reaction 14 ACT [7706.958]	2.945	Raf activator + Raf inhibitor	–
MT2	PP2A IN [26.286]	Reaction 3^b IN [6175.158]	6.827	PP2A inhibitor + tyrosine kinase activator	–
MT3	Reaction 16 ACT★ [220.584]	Reaction 4 IN [7444.068]	15.409	MEK phosphatase activator + inhibitor of receptor dimerization	[62,66]
MT4	Reaction 16 ACT★ [656.857]	Reaction 31 IN★ [1640.247]	15.534	MEK phosphatase activator + PIP kinase inhibitor	[4]
MT5	Reaction 26 IN [6535.419]	Reaction 3^b ACT [6033.846]	17.329	PI3K activator + tyrosine kinase inhibitor	–
MT6	Reaction 3^b ACT [5730.172]	Reaction 16 IN [6892.326]	18.375	tyrosine kinase inhibitor + MEK phosphatase inhibitor	–
MT7	Reaction 16 IN [6163.809]	Reaction 3^b ACT [5845.658]	18.455	MEK phosphatase inhibitor + tyrosine kinase inhibitor	–
MT8	Reaction 3^b ACT [5551.488]	Reaction 34^b IN [7065.998]	18.486	tyrosine kinase inhibitor + tyrosine kinase inhibitor	Mono-therapy
MT9	Reaction 3^b ACT [5541.069]	Reaction 1^b IN [8923.956]	18.522	tyrosine kinase inhibitor + inhibitor of dissociation of heregulin from ErbB4 receptor	–
MT10	Reaction 3^b ACT [5758.748]	Reaction 34^b IN [942.042]	18.679	tyrosine kinase inhibitor + tyrosine kinase inhibitor	Mono-therapy

Table 7
Bottom-10 target combinations for MAPK-PI3K network found using MASCOT. Targets present in Table 3 are marked with ★.

No.	Target 1 [Activity]	Target 2 [Activity]	ς_{off}	Combination	Reference/Comments
SB1	Reaction 3^f IN [29.517]	Reaction 18 ACT★ [0.043]	17.866	tyrosine kinase inhibitor + MEK phosphatase	–
SB2	Reaction 3^b ACT [5556.930]	Reaction 5^f IN [0.042]	18.689	tyrosine kinase inhibitor + inhibitor of RP and Shc binding	–
SB3	Reaction 4 ACT [44.387]	Reaction 26 ACT★ [102.262]	19.228	promoter of receptor dimerization + PI3K inhibitor	–
SB4	Reaction 3^f IN [20.701]	Reaction 29^f ACT★ [5784.119]	19.671	tyrosine kinase inhibitor + Akt activator	–
SB5	Reaction 29^f ACT [8541.949]	Reaction 3^b ACT [4481.641]	20.635	PI3K inhibitor + tyrosine kinase inhibitor	[24]
SB6	Reaction 1^f IN [13.847]	Reaction 30 ACT★ [9707.636]	21.885	tyrosine kinase inhibitor + PIP3 phosphatase activator	–
SB7	Reaction 21 IN★ [2.153]	Reaction 1^b ACT [158.909]	22.021	ERK kinase inhibitor + promoter of dissociation of heregulin from ErbB4 receptor	[62,38]
SB8	Reaction 1^f IN [31.578]	Reaction 16 ACT★ [9.148]	35.070	tyrosine kinase inhibitor + MEK phosphatase	–
SB9	Reaction 1^f IN [37.157]	Reaction 9^b ACT [7084.384]	35.086	tyrosine kinase inhibitor + promoter of GS and ShP binding	–
SB10	Reaction 25^b ACT [2859.395]	Reaction 11 IN [164.144]	71.650	promoter of PI3K binding with activated ErbB4 receptor + MEK kinase inhibitor	–

Table 8
Bottom-10 target combinations for MAPK-PI3K network found using MCSA. Note that targets present in Table 3 are marked with ★.

No.	Target 1 [Activity]	Target 2 [Activity]	ς_{off}	Combination	Reference/Comments
MB1	Reaction 34^f ACT [3797.693]	Reaction 24^b IN [5956.664]	66.220	promoter of receptor internalization + PI3K activator	–
MB2	Reaction 23^b ACT [162.172]	Reaction 34^f ACT [3800.193]	66.247	promoter of dissociation of PI3K from RP + promoter of receptor internalization	–
MB3	Reaction 15 ACT [5993.379]	Reaction 6^f IN [148.058]	97.426	MEK kinase activator + Shc kinase inhibitor	–
MB4	Reaction 17 ACT [3168.839]	Reaction 5^b ACT [7628.053]	107.372	MEK kinase activator + promoter of dissociation of Shc from RP	–
MB5	Reaction 11 ACT [20.804]	Reaction 5^f IN [9864.965]	116.578	MEK kinase activator + inhibitor of RP and Shc binding	–
MB6	Reaction 4 IN [8413.204]	Reaction 1^b ACT [6145.133]	122.430	inhibitor of receptor dimerization + promoter of dissociation of heregulin from ErbB4 receptor	[46,1]
MB7	Reaction 30 ACT★ [4506.434]	Reaction 1^b ACT [3407.957]	256.984	PIP3 phosphatase activator + promoter of dissociation of heregulin from ErbB4 receptor	[46,42]
MB8	Reaction 1^b ACT [5221.591]	Reaction 10 IN [5680.547]	303.923	promoter of dissociation of heregulin from ErbB4 receptor + Shc phosphatase inhibitor	–
MB9	Reaction 32 IN★ [5538.346]	Reaction 1^b ACT [5418.050]	308.214	PIP3 kinase inhibitor + promoter of dissociation of heregulin from ErbB4 receptor	[46,30]
MB10	Reaction 1^b ACT [3556.864]	Reaction 28 IN [6870.239]	3998.338	promoter of dissociation of heregulin from ErbB4 receptor + PIP3 phosphatase inhibitor	–

off-target effects. In addition, we note that the targets in the bottom-10 MASCOT solutions are mostly located upstream in the MAPK-PI3K network. This is in contrast with the top-10 MASCOT solutions where targets are positioned further downstream and nearer to the disease node (ERKPP). No particular trends exist for the MCSA solutions.

The literature curation findings for the bottom-10 solutions are summarized as follows:

SB5: Glaysher et al. investigated combinations of EGFR inhibitors and PI3K inhibitors and found that certain combinations (*e.g.*, ZSTK474 with erlotinib and gefitinib) exhibit enhanced synergistic activity [24]. This correlates well with combination SB5.

SB7, MB6, MB7 and MB9: We did not find any supporting evidence that combinations SB7, MB6, MB7 and MB9 have been performed, successfully or otherwise, in experiments. However, individual components of these combination have shown efficacy in ovarian cancer [62,38,46,1,42,30]. Hence, they warrant further investigation as potential target combinations.

We note that several of the targets identified in the combinations (combinations SB3, SB4, SB9, SB10, MB1, MB3, MB4, MB5 and MB10) involve promoters of known mediators of cancer. We did not find any supporting evidence for the remaining combinations (combinations SB1, SB2, SB6, SB8, MB2 and MB8). Although the bottom-10 combinations contain several biologically relevant

combinations, many of them have relatively large off-target effects and involve targets with large activities. In addition, 70% of combinations found using MASCOT contain tyrosine kinase inhibitors (TKI). Clinical studies have found that although treatments with tyrosine kinase inhibitors (TKI) sometimes produce promising results, most treatment lose their effectiveness soon due to resistance often caused by activating mutations in downstream effectors of the tyrosine kinases [57]. Apart from drug resistance, another potential issue of TKI is toxicity (*e.g.*, nephrotic syndrome) due to the disruption of multiple downstream signaling pathways of the tyrosine kinases which are involved in normal organ functioning [27]. Hence, designing combinations involving TKI would require additional considerations such as understanding the idiosyncrasy of a patient's genome in order to select other suitable targets for the combinations which can minimize TKI-resistance. The activity of TKI in the combinations should also be kept low to reduce potential toxicity. However, the predicted combinations that involve TKI generally have high activity level, making them less than ideal as safe target combinations. Hence, target combinations with larger predicted off-target effects may be indicative of less effective and/or more toxic combinations in the real world.

Target combination discovery for insulin secretion network. The application of MASCOT to well-studied MAPK-PI3K network showcases its prediction quality. We now apply it to glucose-stimulated insulin secretion network [32], which has not been extensively studied for target combinations. We set $\{t_0 = 100,\ i_{max} = 500,\ N = 50,\ \epsilon_t = 5\%,\ \epsilon_a = 5\%,\ \lambda^+ = 1.8\}$. Note that MCSA failed to complete execution within 8 days on this network. Hence, we do not present the results.

Table 9
Target combinations for glucose-stimulated insulin secretion network found using MASCOT.

No.	Target 1 [Activity]	Target 2 [Activity]	ς_{off}	Combination	Reference
1	Reaction 7 IN [12.479]	Reaction 38 IN [8483.898]	97.318	glyceraldehyde 3-phosphate dehydrogenase inhibitor + FAD dependent glycerol-3-phosphate dehydrogenase inhibitor	–
2	Reaction 7 IN [12.998]	Reaction 17 IN [5079.313]	115.524	glyceraldehyde 3-phosphate dehydrogenase inhibitor + oxoglutarate dehydrogenase complex inhibitor	–
3	Reaction 10 IN [31.713]	Reaction 45 ACT [3258.442]	126.778	lactate dehydrogenase inhibitor + malate dehydrogenase activator	–
4	Reaction 36 ACT [4371.772]	Reaction 10 IN [45.927]	2470.605	glutathione reductase activator + lactate dehydrogenase inhibitor	–
5	Reaction 38 IN [8974.797]	Reaction 10 IN [50.682]	2531.171	FAD dependent glycerol-3-phosphate dehydrogenase inhibitor + lactate dehydrogenase inhibitor	[37,40,50]
6	Reaction 38 ACT [4537.414]	Reaction 10 IN [50.993]	2576.963	FAD dependent glycerol-3-phosphate dehydrogenase activator + lactate dehydrogenase inhibitor	[37,40,50]
7	Reaction 36 ACT [2953.557]	Reaction 10 IN [52.789]	2583.011	glutathione reductase activator + lactate dehydrogenase inhibitor	–
8	Reaction 21 IN [2739.023]	Reaction 10 IN [52.789]	2613.709	mitochondrion malate dehydrogenase inhibitor + lactate dehydrogenase inhibitor	–
9	Reaction 38 IN [1189.972]	Reaction 7 IN [21.114]	2615.501	FAD dependent glycerol-3-phosphate dehydrogenase inhibitor + oxoglutarate dehydrogenase complex inhibitor	–
10	Reaction 36 ACT [784.932]	Reaction 10 IN [53.747]	2624.339	glutathione reductase activator + lactate dehydrogenase inhibitor	–

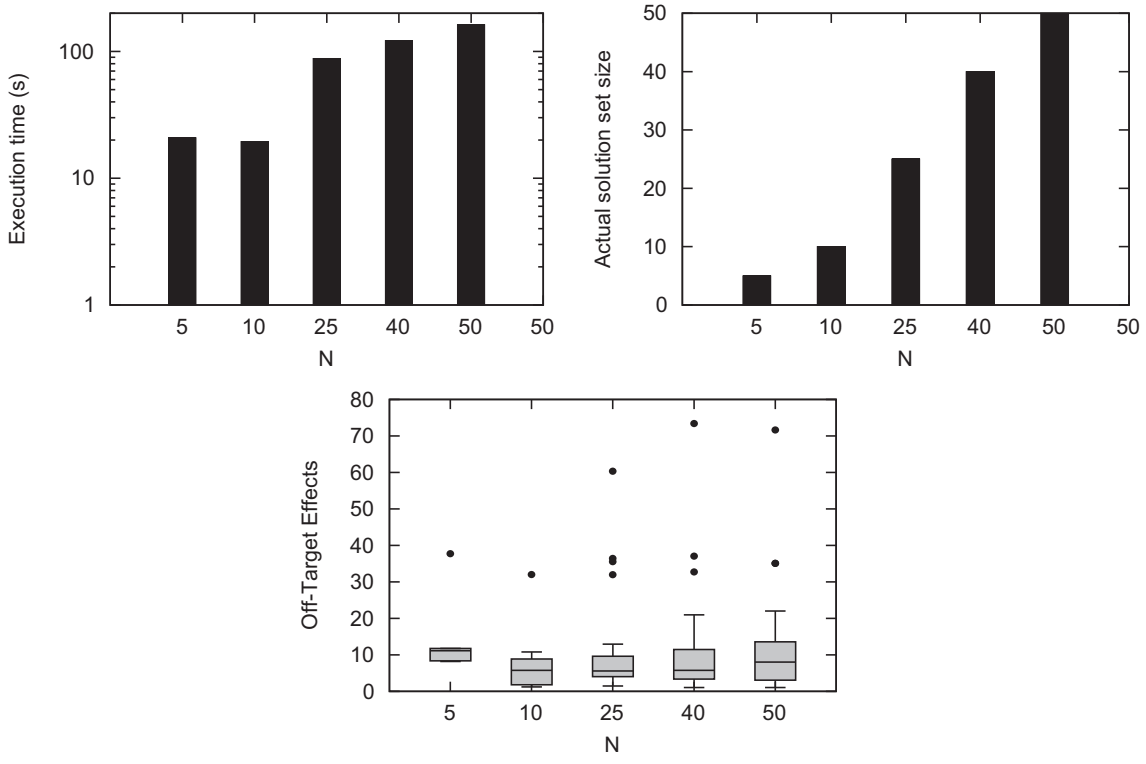

Fig. 6. Effects of solution set size.

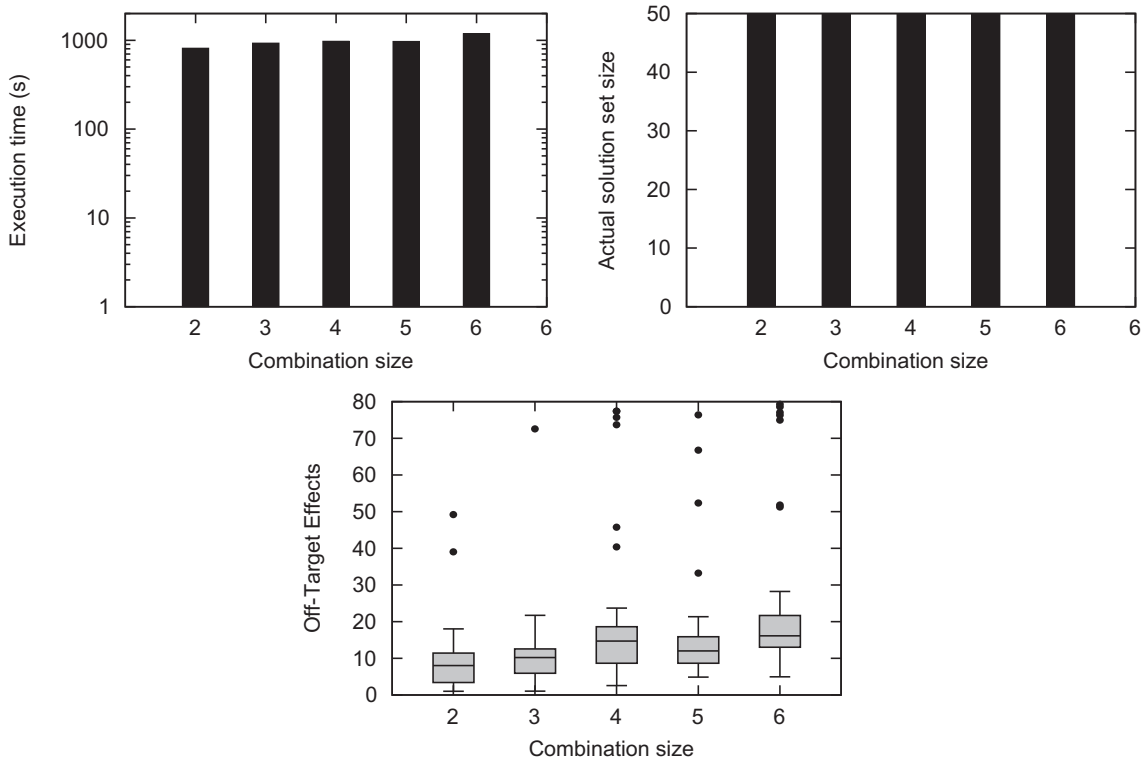

Fig. 7. Effects of combination size.

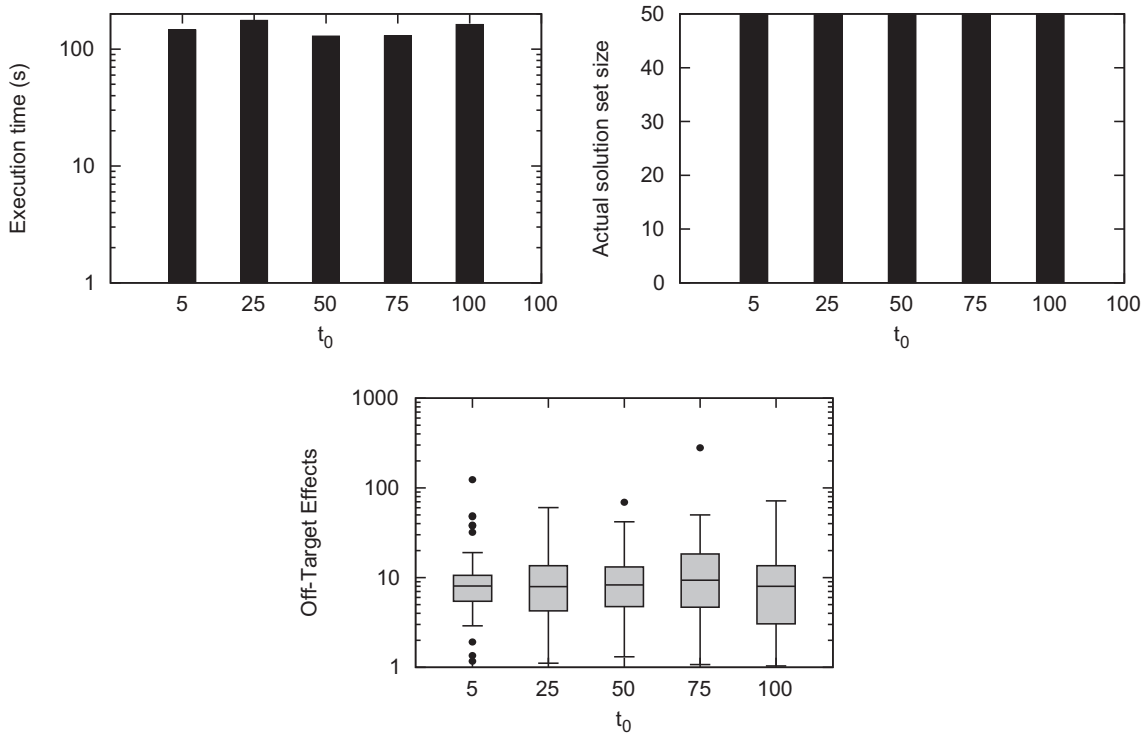

Fig. 8. Effects of initial temperature of simulated annealing.

The solution set is listed in Table 9 and its characteristics are quite different from that of the `MAPK-PI3K` network. First, the solution set of the `glucose`-stimulated `insulin` secretion network contains fewer solutions. Second, the off-target effects of these solutions are generally relatively high (*e.g.*, greater than 1000). Third, for majority (90%) of the solutions, at least one target has activity that exceed 1000. This difference in result could be due to the presence of fewer targets that can effectively

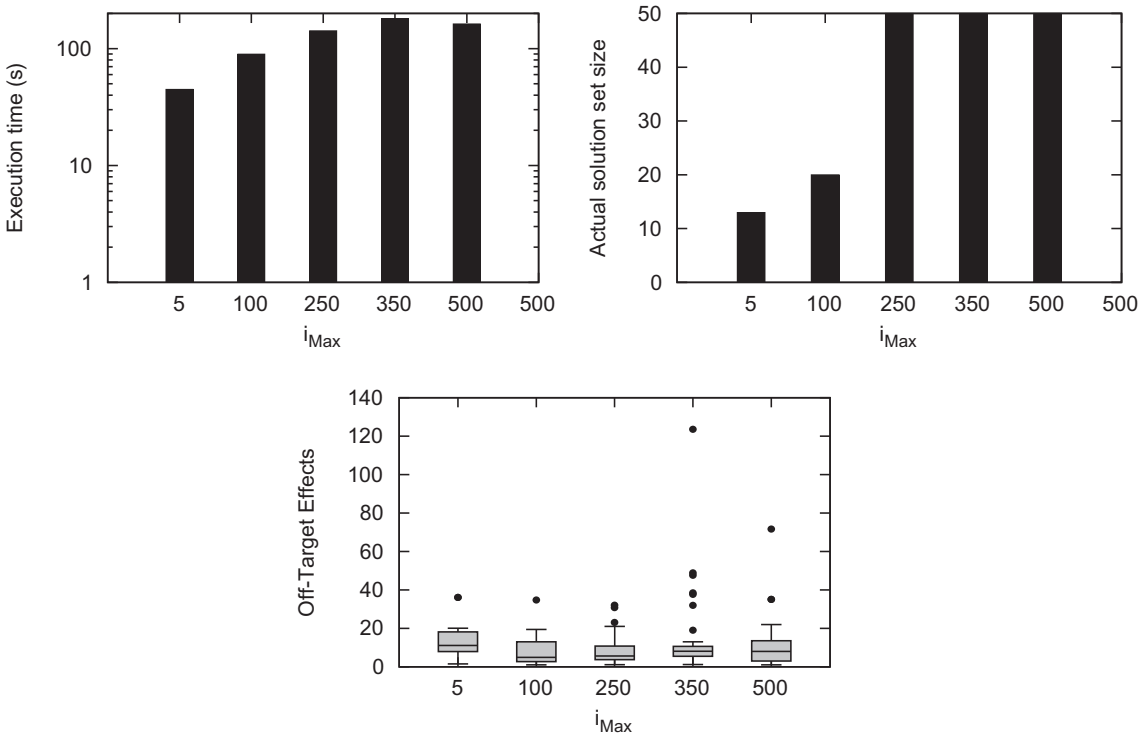

Fig. 9. Effects of limits of iterations per cycle of simulated annealing.

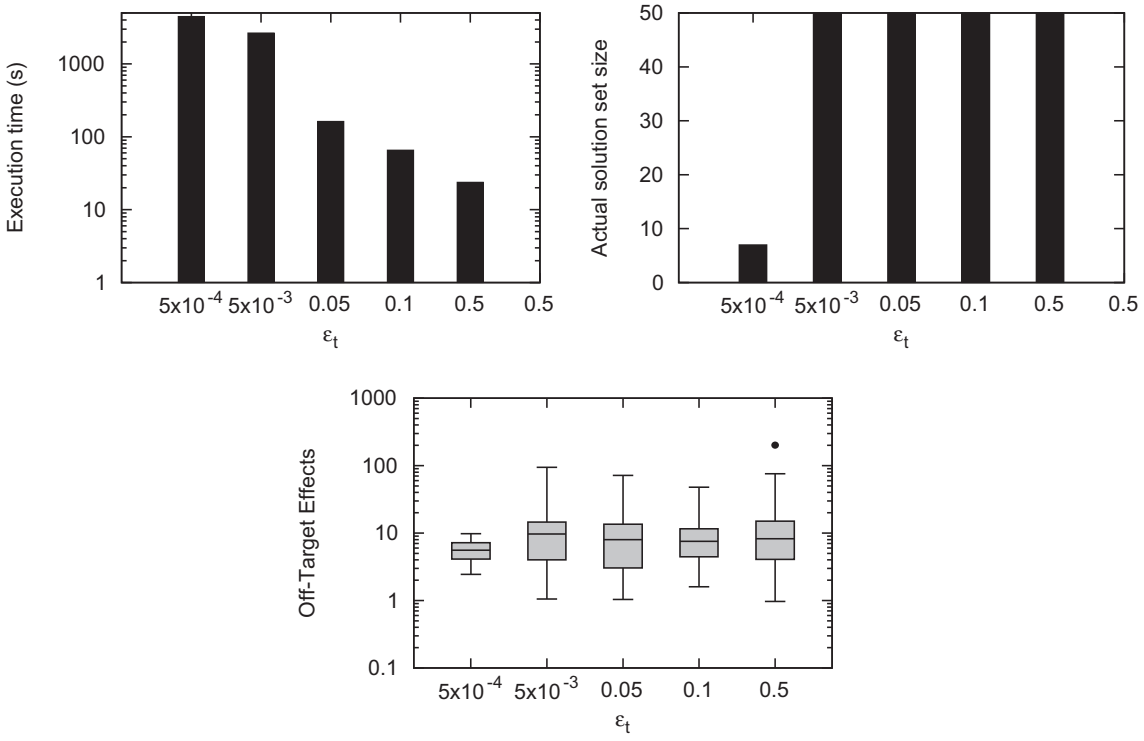

Fig. 10. Effects of the adjustment factor for therapeutic effect.

control the activity of mitochondrial ATP in the glucose-stimulated insulin secretion network. In particular, there were only two targets (4% of total targets) when perturbed individually. This is in contrast to the MAPK-PI3K network in which 64% of the targets can achieve the therapeutic goal when they are perturbed alone. Hence, in comparison to the MAPK-PI3K network, the LOEWE heuristic plays a limited role in the glucose-stimulated insulin secretion network, resulting in fewer solutions found, higher off-target effects and higher target activities in the combinations.

In Table 9, we note an over-representation of two targets that have activities less than 100. They are inhibitors of reactions 7 and 10 which correspond to inhibitors of `glyceraldehyde 3-phosphate dehydrogenase` and `lactate dehydrogenase`, respectively. We explore the possibility of combining these two targets to obtain combinations that have better target activity profile. Indeed, when reactions 7 and 10 are inhibited at target activity of 10 and 39.25, respectively, the desired therapeutic effect was achieved with off-target effects of 2628.29. *Hence, the solution sets of* MASCOT *are useful as a guide for potential targets and target activities to explore in creating new target combinations.*

In addition, we perform literature curation to assess the biological relevance of the solutions in Table 9. A search in *PubMed* using the keywords "type 2 diabetes mellitus drug combinations" did not yield any relevant target combinations for the `glucose-stimulated insulin` secretion network. Hence, we perform further curation by specifically looking for relevant literature pertaining to the predicted combinations in Table 9. Although we did not find any supporting evidence that combinations 5 and 6 have been performed, successfully or otherwise, in experiments, their individual components have shown efficacy in diabetes [37,40,50]. Hence, these combinations warrant further investigation as potential target combinations. We did not find any supporting evidence for the remaining combinations. However, this does not necessarily imply that these combinations are not useful for diabetes. Current research may be focussed on examining other diabetes-related pathways. Hence, these predicted combinations may not have been explored yet. Additional research is necessary to confirm the relevance of these combinations.

Effect of Parameters on Solution Set. In this final set of experiments, we shall investigate the effects of various parameters on the solution set of MASCOT using the `MAPK-PI3K` network.

Effect of the solution set size. We vary $N : \{5, 10, 25, 40, 50\}$. We make the following observations from Fig. 6. The whiskers and inter-quartile ranges of the off-target effects boxplot are observed to increase as N increases. This implies that for smaller N, majority of the solutions within the middle 50% of off-target effects have off-target effects values that are closer with each other as compared to those for larger N. This is expected as increased number of solutions invariably results in a higher likelihood of encountering combinations with off-target effects that vary widely. Despite the potential of variability in the off-target effects, the solution

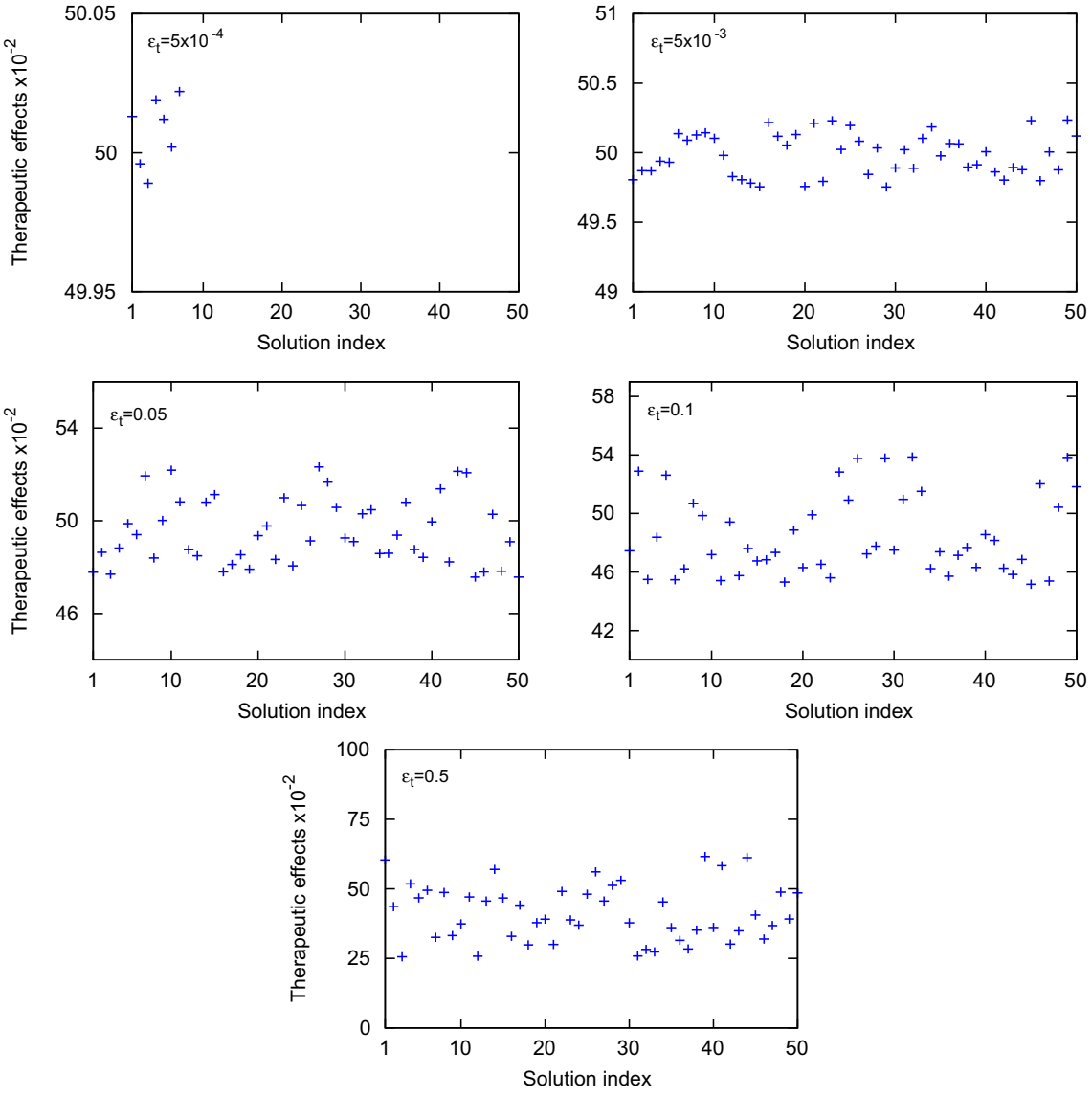

Fig. 11. Solution space when the adjustment factor for therapeutic effect is varied.

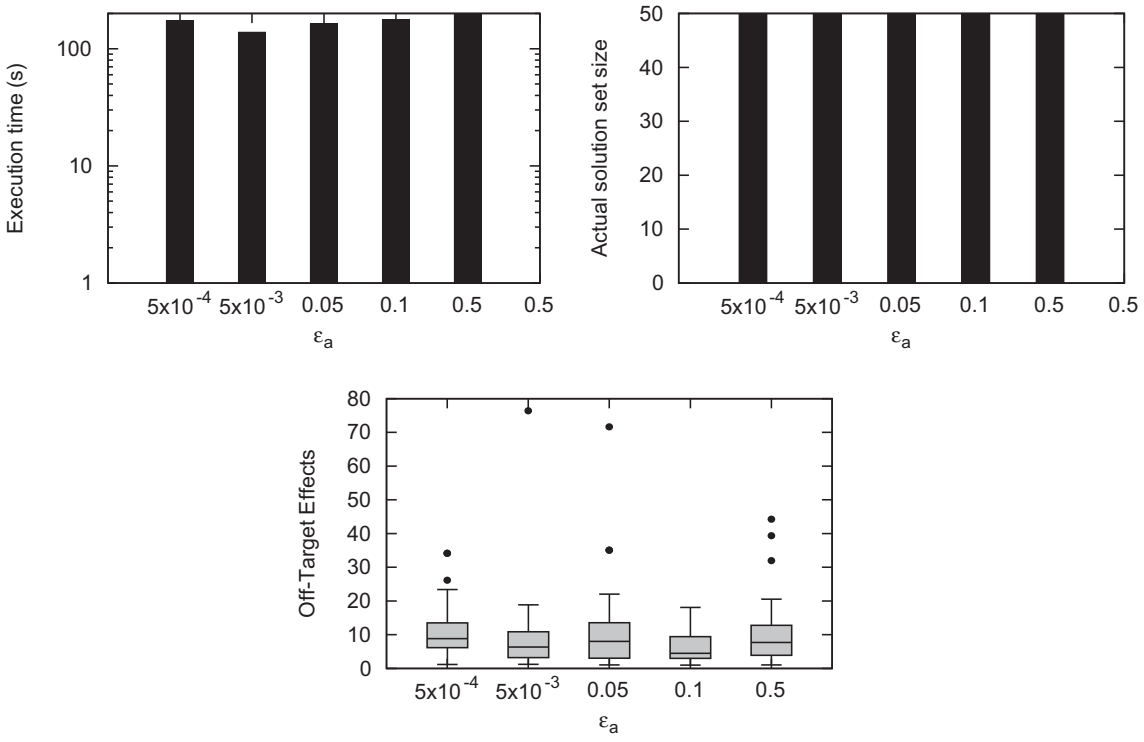

Fig. 12. Effects of the adjustment factor for target interaction.

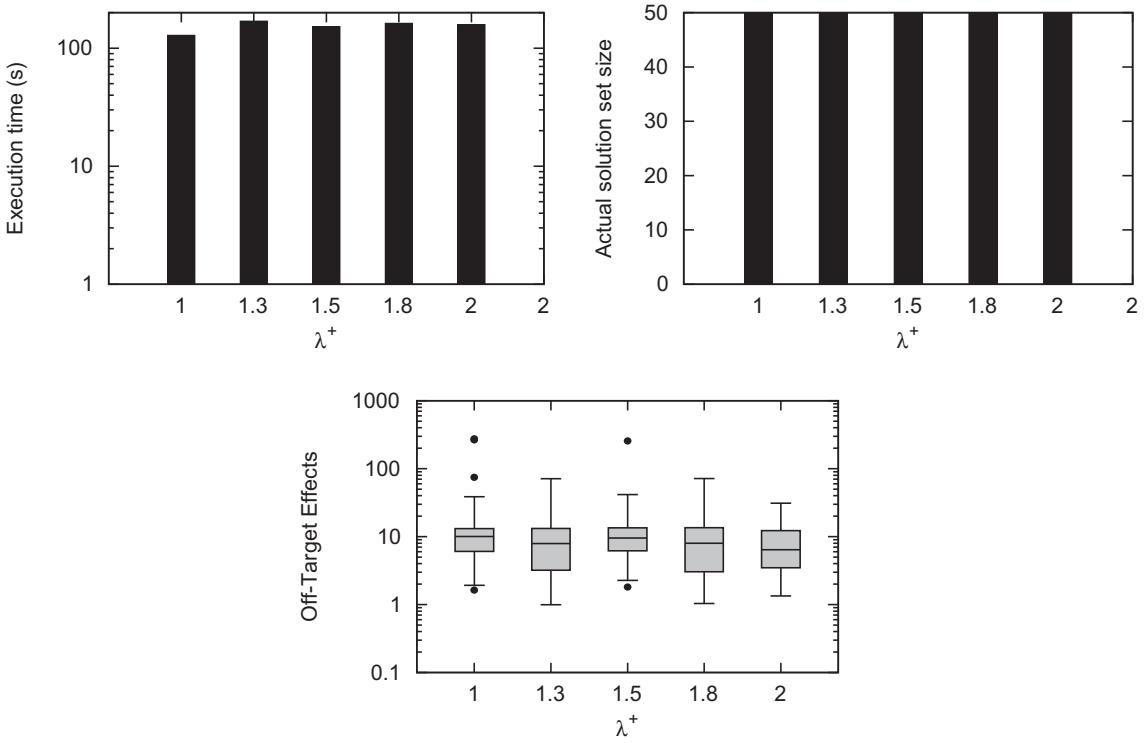

Fig. 13. Effects of the selective pressure.

set is relatively robust as the whiskers of the boxplot (which contains the majority of the solutions) are relatively constant in the range [0–20] when N is greater than 25. In addition, the increase in N demanded longer execution time as more iterations are required to identify the larger number of solutions.

Effect of combination size. Fig. 7 summarizes the effect of varying combination size ($\mathcal{S} : \{2, 3, 4, 5, 6\}$) on off-target effects, execution time and the actual solution set size. We make the following observations. First, off-target effects increased with the size of the combination. Incorporating additional targets into the combination

was likely to cause perturbation of a larger number of downstream nodes of these target. This inevitably increases off-target effects. Second, the increase in combination size also resulted in a slight increase in execution time. This is probably due to additional computation required for the selection of additional targets and their activities. Third, targets closer to the disease node are frequently selected for the combinations. This is evident from the fact that all combinations include either a ERK (or MEK) kinase inhibitor or a ERK (or MEK) phosphatase activator. Fourth, there is a tendency for less desirable targets (*i.e.*, promoter of pro-cancer signals) such as a drug activating the tyrosine kinase to be included in the combination as $\mathcal{S}$ increased. Hence, it may be necessary to consider additional rules such as exclusion of pro-cancer signal promoters when generating combinations of larger sizes.

Effect of initial temperature. In this experiment, we examine the effect of varying t_0. Fig. 8 reports the results. Observe that there is no significant changes in terms of execution time, actual solution set size and off-target effects.

Effect of the number of iteration. Next, we investigate the effects of varying i_{max}. The effects on execution time and actual solution set size that are observed in Fig. 9 are similar to those in Fig. 6. This is because increasing i_{max} increases the maximum number of iterations for the simulated annealing, allowing a larger solution space to be explored. Hence, more execution time is needed and the actual solution set size is larger. The whiskers of the off-target effects boxplot remain relatively constant in the range [0–20]. This corresponds well with our previous observations in Fig. 6. Compared to t_0, the effect of varying i_{max} is more pronounced. This is because, in our design of the simulated annealing algorithm, we have set t_0 to be smaller than i_{max}. Note that t_0 and i_{max} control the outer- and inner-loop of the algorithm, respectively. For instance, when t_0 is set to 5, the maximum number of iterations is $5 \times 500 = 2500$, whereas when i_{max} is set to 5, the maximum number of iterations is reduced further to 100x5 = 500.

Effect of ϵ_t, ϵ_a, and λ^+. Figs. 10 and 11 report the effect of ϵ_t. We note an increase in the actual solution set size and corresponding decrease in execution time as ϵ_t is increased. As ϵ_t is increased, there is further relaxation of the condition for therapeutic effect and this allows more candidates to be accepted. This could also be the reason for an increase in the range of the whiskers of the off-target effects boxplot when ϵ_t increased from 5×10^{-4} to 5×10^{-3} as there is a large increase in solution set size from 7 to 50. Note that although increasing ϵ_t results in a dramatic decrease in execution time, this is at the expense of solutions moving further away from the desired therapeutic effect. In contrast, we do not observe any significant changes in terms of execution time, actual solution set size and off-target effects when we vary ϵ_a (Fig. 12) and λ^+ (Fig. 13).

6.3. Choice of parameter values

The aforementioned experiments demonstrate that several parameters (*e.g.*, t_0, ϵ_a, and λ^+) do not affect the execution time, solution set size and off-target effects significantly. Hence, as a general rule, these parameters can be configured as $t_0 = 100$, $\epsilon_a = 5\%$ and $\lambda^+ = 1.8$. For the remaining parameters, the configuration is largely dependent on a user's goal. In particular, smaller values of ϵ_t yield predicted combinations that are closer to the desired therapeutic effect whereas increasing i_{max} increases exploration of the solution space. A guideline for configuring i_{max} is to use a lower value (*e.g.*, 250) to quickly search the domain space for candidate combinations. In the event that insufficient candidates are returned, then, a higher value (*e.g.*, 500) can be set to increase the exploration space. For example, if the goal is to explore 20 combinations where the desired therapeutic effect must be achieved stringently, then a possible configuration of MASCOT would be $\epsilon_t = 0.5\%$, $N = 20$, $S = 4$ and $i_{max} = 250$.

7. Conclusions

In this work, we describe MASCOT, a generic framework for combination therapy that predicts synergistic target combination based on using simulated annealing. Specifically, it leverages on two heuristics, namely, a machine learning-based target prioritization and LOEWE heuristic from pharmacology. The former heuristic is used for selecting appropriate targets when generating candidate target combinations in the simulated annealing algorithm, whereas the latter heuristic is used to select target activity in order to ensure synergistic target interaction within the combinations. Our results reveal that the heuristics indeed improve execution time and off-target effects of the target combinations, when compared to the state-of-the-art approaches. The solutions found using MASCOT are also enriched with targets in known drug combinations and found to be biologically relevant. Due to the generic and extensible nature of MASCOT, further disease-specific constraints can easily be added on it to improve the result quality with respect to a specific disease. Furthermore, various "omics" data and drug and disease information can be included as heuristics to find target combinations that exclude combinations akin to monotherapies, and that avoid including activators of pro-disease targets as part of the combinations.

Note that MASCOT relies on the specification of the system of ODEs for the input signaling network. Hence, the prediction result is dependent on the accuracy of the given ODEs. The accuracy of the prediction results can be further improved by using signaling networks that are experimental data-driven and context-specific. As part of future work, we intend to leverage additional experimental data (*e.g.*, gene expression data), annotations (*e.g.*, GO process terms), and ontology to address issues of inherent noise in signaling network. In particular, nodes annotated with the same GO process terms are more likely to be located in the same pathway or in pathways that crosstalk. Hence, these nodes are likely to transmit signals to each other. The set of nodes found in protein or gene databases (*e.g.*, *UniProt*) that are annotated with the same GO process terms as targets in the predicted combination may provide an indication of the actual off-target effect due to the target combination. Hence, the existing definition of off-target effects needs to be enhanced to take into consideration such information.

Acknowledgments

Huey Eng Chua and Sourav S Bhowmick were supported by MOE AcRF Tier-1 Grant RGC 1/13.

Appendix A. Proof for complexity analysis

A.1. Time complexity

In Algorithm 1, the PREPROCESSINPUT procedure performs several tasks: (Task 1) convert reversible reactions to irreversible reactions, (Task 2) identify individual drug targets, (Task 3) modify target reactions and (Task 4) finding ITA. In the worst case, Tasks 1, 2 and 3 require $O(|E|), O(|V||E|)$ and $O(\text{MAX}(|E|, |V|))$ time, respectively, where MAX$(\cdot)$ is the maximum operator. MCSA, which is used to find the ITA in Task 4, consists of the following steps: generate random candidate solution ($O(1)$); simulate target combination effect using an ODE solver ($O(|E||\varphi|)$ [52]); estimate the therapeutic and off-target effects ($O(|V||\varphi|)$); and perform candidate acceptance test ($O(1)$), where $|\varphi|$ is the number of time points in the concentration–time series profile curve. Hence, the time complexity of

each MCSA iteration is $O((|V|+|E|)\times|\varphi|)$. In the worst case, no values satisfying the desired therapeutic effect is found when the MCSA terminates on completing all runs ($t_0 \times i_{max}$ iterations) resulting in a time complexity of $O(\text{MAX}(|E|,|V|)\times t_0 \times i_{max} \times (|V|+|E|)\times|\varphi|)$ to obtain the ITA for all targets. The time complexity of PREPROCESSINPUT can be further reduced to $O(t_0 \times i_{max} \times (|V|+|E|)\times|\varphi|)$.

In the ESA phase, in the worst case, ESA terminates on completing all runs ($t_0 \times i_{max}$ iterations) without finding N solutions. Similar to the MCSA, each ESA iteration consists of the following steps: (Task 5) generate candidate solution, (Task 6) simulate combination effects using an ODE solver, (Task 7) estimate combination effects and (Task 8) perform candidate acceptance test. Task 5 involves target prioritization heuristic for selecting the target. This heuristics accepts higher prioritized targets with higher probability (Section 5.2). In the worst case, the lowest prioritized target (y) from the set of individual targets ($\mathcal{T}$) is always considered. y will be accepted if a randomly generated number is lower than the selection probability of y (sp_y). It takes about $O(|\mathcal{T}|)$ tries in order to accept the worst candidate. Hence, the complexity of generating the candidate solution is $O(|\mathcal{X}||\mathcal{T}|)$ where $|\mathcal{X}|$ is the size of the candidate combination and $|\mathcal{T}|$ is the size of the candidate target. Tasks 6, 7 and 8 take $O(|E||\varphi|)$ [52], $O(|V||\varphi|)$ and $O(1)$ time, respectively. Since $|\varphi| \gg |\mathcal{X}|$ and $O(|E|) = O(|\mathcal{T}|)$, the time complexity of ESA is $O(t_0 \times i_{max} \times (|V|+|E|)\times|\varphi|)$ time in the worst case. Thus, the overall worst time complexity of MASCOT is $O(t_0 \times i_{max} \times (|V|+|E|)\times|\varphi|)$.

A.2. Space complexity

In Algorithm 1, $O(|V|+|E|)$ is needed for storing the input signaling network G and set of prioritized node rank Ψ. Recall that the PREPROCESSINPUT procedure performs several tasks: (Task 1) convert reversible reactions to irreversible reactions, (Task 2) identify individual drug targets, (Task 3) modify target reactions and (Task 4) finding ITA. In the worst case, Tasks 1, 2 and 3 require $O(2|E|), O(2|V||E|)$ and $O(\text{MAX}(|E|,|V|))$ space, respectively, where MAX$(\cdot)$ is the maximum operator. Recall that MCSA, which is used to find the ITA in Task 4, consists of the following steps: generate random candidate solution ($O(1)$); simulate target combination effect using an ODE solver ($O(|V||\varphi|)$); estimate the therapeutic and off-target effects ($O(|V||\varphi|)$); and perform candidate acceptance test ($O(1)$), where $|\varphi|$ is the number of time points in the concentration–time series profile curve. Hence, the space complexity of MCSA is $O(|V||\varphi|)$. Hence, PREPROCESSINPUT has space complexity $O(|V|(|E|+|\varphi|))$.

In the ESA phase, in the worst case, ESA returns N solutions which requires $O(\mathcal{S}\times N)$ storage space. Similar to the MCSA, ESA consists of the following steps: (Task 5) generate candidate solution ($O(\mathcal{S})$), (Task 6) simulate combination effects using an ODE solver ($O(|V||\varphi|)$), (Task 7) estimate combination effects ($O(|V||\varphi|)$) and (Task 8) perform candidate acceptance test ($O(1)$). Hence, the ESA phase requires $O(\mathcal{S}\times N + |V||\varphi|)$ space. Since in most applications, we expect $|V||\varphi| > \mathcal{S}\times N$, the space complexity of the ESA phase can be further reduced to $O(|V||\varphi)$. Thus, the overall worst space complexity of MASCOT is $O(|V|(|E|+|\varphi|))$.

A.3. Convergence analysis

The process of simulated annealing can be interpreted as the change of state of an inhomogeneous Markov chain where the state at the k^{th} step is denoted as θ_k and P_{θ_k} is its probability distribution [35]. The convergence of simulated annealing is established in [35] as follows: if the temperature is kept constant (*i.e.*, $t_k = t$), then the distribution of the state of the chain P_{θ_k} tends to the equilibrium distribution $\pi^{(t)}$. If $t \to \infty$, then $\pi^{(t)}$ tends to the zero-temperature distribution $\pi^{(\infty)}$. Hence, if the cooling schedule tends to infinity (*i.e.*, $t_k \to \infty$), P_{θ_k} "follows" $\pi^{(t_k)}$ which tends to $\pi^{(\infty)}$. The distribution of P_{θ_k} converges to the zero-temperature distribution and coincides with a global optimizer with probability one. For simulated annealing algorithm on a continuous domain, the set of global optimizers has no Lebesgue measure and would result in a set of optimizers with null probability [35]. Note that MASCOT algorithm performs simulated annealing on a system of ODES where the goal is to minimize the off-target effects and to ensure that the therapeutic effect of the predicted combination achieves the desired therapeutic effects. The system of ODES are continuous functions [64]. Hence, MASCOT is a simulated annealing algorithm that runs on a continuous domain and can only find approximate solutions in finite time according to [7]. The approximate global optimizer is defined as follows [35] for a minimization problem:

Definition 5. Let $U:\Theta\to R$ be an optimization criterion where $\Theta\subset R^N$ is bounded, and π_{Leb} be the standard Lebesgue measure. Then θ is an approximate global optimizer of U with value imprecision ϵ and residual domain α if $\pi_{Leb}\{\theta'\in\Theta : U(\theta') < U(\theta)+\epsilon\} \leqslant \alpha\pi_{Leb}(\Theta)$ where $\epsilon \geqslant 0$ and $\alpha\in[0,1]$.

Definition 5 holds with probability of at least $\sigma = \frac{1}{1+\lfloor\frac{1+\delta}{c_1+1+\delta}\rfloor^t\lfloor\frac{1+\delta}{2c_1+\delta}-1\rfloor^{\frac{1+\delta}{\delta}}}$ [64] where $t \geqslant 1$ and $\delta > 0$. Further, for a bounded domain, Definition 5 holds with probability of at least $\sigma - \|P_{\theta_k} - \pi^{(t)}\|_{TV}$ where $\|P_{\theta_k} - \pi^{(t)}\|_{TV}$ is the distance between the distribution of P_{θ_k} and the target distribution $\pi^{(t)}$ and $\|P_{\theta_k} - \pi^{(t)}\|_{TV}$ decreases geometrically to zero as $k\to\infty$. Note that for MASCOT, the domain of the target activity is [0-∞], but can easily be replaced by [0-ψ] where ψ is a arbitrary large value (*e.g.*, 10^5) in order to transform it to a bounded domain.

References

[1] D. Agus et al., Efficacy and safety of single-agent pertuzumab (rhumab 2c4), a human epidermal growth factor receptor dimerization inhibitor, in castration-resistant prostate cancer after progression from taxane-based therapy, J. Clin. Oncol. 25 (6) (2007) 675–681.
[2] B.B. Aldridge et al., Physicochemical modelling of cell signalling pathways, Nat. Cell Biol. 8 (11) (2006).
[3] B. Al-Lazikani et al., Combinatorial drug therapy for cancer in the post-genomic era, Nat. Biotechnol. 30 (7) (2012).
[4] R. Bast, Molecular approaches to personalizing management of ovarian cancer, Ann. Oncol. 22 (Suppl. 8) (2011).
[5] M.V. Blagosklonny, Antagonistic drug combinations that select against drug resistance: from bacteria to cancer, Cancer Biol. Therapy 6 (7) (2007).
[6] T. Blickle et al., A comparison of selection schemes used in evolutionary algorithms, Evol. Comput. 4 (4) (1996).
[7] L. Blum et al., Complexity and Real Computation, Springer-Verlag, New York, 1998.
[8] A.A. Borisy et al., Systematic discovery of multicomponent therapeutics, Proc. Natl. Acad. Sci. 100 (13) (2003).
[9] B. Bornstein et al., Libsbml: an API library for sbml, Bioinformatics 24 (6) (2008).
[10] D. Calzolari et al., Search algorithms as a framework for the optimization of drug combinations, PLoS Comput. Biol. 12 (4) (2008).
[11] D.A. Chan et al., Harnessing synthetic lethal interactions in anticancer drug discovery, Nat. Rev. Drug Discov. 10 (5) (2011).
[12] W.L. Chiou, Critical evaluation of the potential error in pharmacokinetic studies of using the linear trapezoidal rule method for the calculation of the area under the plasma level-time curve, J. Pharmacokinet. Pharmacodynamics 6 (6) (1978).
[13] T.C. Chou et al., A simple generalized equation for the analysis of multiple inhibitions of Michaelis-Menten kinetic systems, J. Biol. Chem. 252 (18) (1977).
[14] H. Chua et al., PANI: A novel algorithm for fast discovery of putative target nodes in signaling networks, ACM BCB, 2011.
[15] H. Chua et al., STEROID: In silico heuristic target combination identification for disease-related signaling networks. ACM BCB, 2012.
[16] H. Chua et al., TENET: topological feature-based target characterization in signaling networks, Bioinformatics 31 (20) (2015) 3306–3314.
[17] H. Chua et al., TAPESTRY: Network-centric target prioritization in disease-related signaling networks. ACM BCB, 2016.
[18] R.A. Copeland et al., Drug-target residence time and its implications for lead optimization, Nat. Rev. Drug Discov. 5 (9) (2009).

[19] K. Courtney et al., The PI3K pathway as drug target in human cancer, J. Clin. Oncol. 28 (6) (2010).
[20] P. Csermely et al., Structure and dynamics of molecular networks: a novel paradigm of drug discovery, Pharmacol. Ther. 138 (2013).
[21] Q. Cui et al., Protein evolution on a human signaling network, BMC Syst. Biol. 3 (1) (2009) 21.
[22] J. Engelfiet et al., A comparison of boundary graph grammars and context-free hypergraph grammars, Inform. Comput. 84 (2) (1990) 163–206.
[23] E.C. Freuder et al., Systematic versus stochastic constraint satisfaction. In IJCAI, 1995.
[24] S. Glaysher et al., Targeting EGFR and PI3K pathways in ovarian cancer, Br. J. Cancer 109 (7) (2013).
[25] J.A. Gollob et al., Role of Raf kinase in cancer: therapeutic potential of targeting the Raf/MEK/ERK signal transduction pathway, Semin Oncol 33 (4) (2006).
[26] S. Gordo et al., Knitting and untying the protein network: modulation of protein ensembles as a therapeutic strategy, Protein Sci. 18 (3) (2009).
[27] K. Gotink et al., Anti-angiogenic tyrosine kinase inhibitors: what is their mechanism of action?, Angiogenesis 13 (1) (2010)
[28] P. Gustafson et al., Local sensitivity analysis, Bayesian Stat. 5 (1996) 197–210.
[29] M. Hatakeyama et al., A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling, Biochem. J. (2003).
[30] A. Howes et al., The phosphatidylinositol 3-kinase inhibitor, px-866, is a potent inhibitor of cancer cell motility and growth in three-dimensional cultures, Mol. Cancer Ther. 6 (9) (2007) 2505–2514.
[31] S. Iadevaia et al., Identification of optimal drug combinations targeting cellular networks: integrating phospho-proteomics and computational network analysis, Cancer Res. 70 (17) (2010).
[32] N. Jiang et al., A kinetic core model of the glucose-stimulated insulin secretion network of pancreatic β cells, Mamm. Genome (2007).
[33] T. Kacprowski et al., NetworkPrioritizer: a versatile tool for network-based prioritization of candidate disease genes or other molecules, Bioinformatics 29 (11) (2013).
[34] S. Klamt et al., Hypergraphs and cellular networks, PLoS Comput. Biol. (2009).
[35] A. Lecchini-Visintini, et al., Simulated annealing: Rigorous finite-time guarantees for optimization on continuous domains. In NIPS, 2007.
[36] H. Li et al., Effects of Raf kinase inhibitor protein expression on metastasis and progression of human epithelial ovarian cancer, Mol. Cancer Res. 6 (6) (2008).
[37] A. Madiraju et al., Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase, Nature (2014).
[38] U. Matulonis et al., Cediranib (azd2171) is an active agent in recurrent epithelial ovarian cancer, J. Clin. Oncol. 26 (15) (2008).
[39] O.K. Mirzoeva et al., Basal subtype and MAPK/ERK kinase (MEK)-phosphoinositide 3-kinase feedback signaling determine susceptibility of breast cancer cells to MEK inhibition, Cancer Res. 69 (2) (2009).
[40] W. Miskimins et al., Synergistic anti-cancer effect of phenformin and oxamate, PloS One 9 (1) (2014).
[41] M.S. Mitchell, Combinations of anticancer drugs and immunotherapy, Cancer Immunol. Immunother. 52 (2003).
[42] C. Mitsiades et al., The AKT pathway: molecular targets for anti-cancer drug development, Curr. Cancer Drug Targets 4 (3) (2004) 235–256.
[43] S. Morgan et al., The cost of drug development: a systematic review, Health Policy 100 (1) (2011).
[44] N. Nakayama et al., KRAS or BRAF mutation status is a useful predictor of sensitivity to MEK inhibition in ovarian cancer, Br. J. Cancer 99 (12) (2008).
[45] A. Nolte et al., A note on the finite time behavior of simulated annealing, Math. Oper. Res. 25 (3) (2000).
[46] N. Normanno et al., Target-based agents against ErbB receptors and their ligands: a novel approach to cancer treatment, Endocrine-Related Cancer 10 (1) (2003) 1–21.
[47] N. Le Novère et al., BioModels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems, Nucleic Acids Res. 34 (2006) D689–91.
[48] G. Pohl et al., Inactivation of the mitogen-activated protein kinase pathway as a potential target-based therapy in ovarian serous tumors with KRAS or BRAF mutations, Cancer Res. 65 (5) (2005).
[49] J.O. Ramsay et al., Parameter estimation for differential equations: a generalized smoothing approach, J. R. Stat. Soc.: Ser. B Stat. Methodol. 69 (5) (2007).
[50] A. Ratsimamanga et al., Sodium oxamate for the treatment of diabetic conditions. US Patent 6,300,375, 2001.
[51] J.C.J. Ray et al., Non-transcriptional regulatory processes shape transcriptional network dynamics, Nat. Rev. Microbiol. 9 (11) (2011).
[52] A. Reibman et al., Numerical transient analysis of markov models, Comput. Oper. Res. 15 (1) (1988).
[53] P. Roberts et al., Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer, Oncogene (2007).
[54] S. Sahle et al., Simulation of biochemical networks using COPASI: a complex pathway simulator. In Proc. of WSC, 2006.
[55] F. Sams-Dodd, Target-based drug discovery: is something wrong?, Drug Discovery Today 10 (2) (2005)
[56] N. Sieben et al., Differential gene expression in ovarian tumors reveals Dusp 4 and Serpina 5 as key regulators for benign behavior of serous borderline tumors, J. Clin. Oncol. 23 (29) (2005).
[57] J. Sierra et al., Molecular mechanisms of acquired resistance to tyrosine kinase targeted therapy, Mol. Cancer 9 (1) (2010).
[58] R. Steinmetz et al., Mechanisms regulating the constitutive activation of the extracellular signal-regulated kinase (ERK) signaling pathway in ovarian cancer and the effect of ribonucleic acid interference for ERK1/2 on cancer cell proliferation, Mol. Endocrinol. 18 (10) (2004).
[59] J. Tang et al., Target inhibition networks: predicting selective combinations of druggable targets to block cancer survival pathways, PLoS Comput. Biol. 9 (9) (2013).
[60] M.H. Ung et al., Integrative analysis of cancer genes in a functional interactome, Sci. Rep. 6 (2016). article no. 29228.
[61] H. Vogel, Drug Discovery and Evaluation: Pharmacological Assays, Springer Science & Business Media, 2008. vol. 1.
[62] F. Warnberg et al., Effect of a farnesyl transferase inhibitor (r115777) on ductal carcinoma in situ of the breast in a human xenograft model and on breast and ovarian cancer cell growth in vitro and in vivo, Breast Cancer Res. 8 (2) (2006).
[63] R.A. Weinberg, The Biology of Cancer, Garland Science, New York, 2007.
[64] E.W. Weisstein, Ordinary Differential Equation, 2004.
[65] K. Yang et al., Finding multiple target optimal intervention in disease-related molecular network, Mol. Syst. Biol. 4 (2008).
[66] R. Zeineldin et al., Targeting the EGF receptor for ovarian cancer therapy, J. Oncol. 2010 (2009).
[67] L. Zhao et al., Evaluation of combination chemotherapy, Clin. Cancer Res. 10 (23) (2004) 7994–8004.