Computers and Chemical Engineering 156 (2022) 107578

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Multi-scale sustainable engineering: Integrated design of reaction networks, life cycles, and economic sectors ☆

Vyom Thakker, Bhavik R. Bakshi*

William G. Lowrie Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus, OH 43210, USA

ARTICLE INFO

Article history:
Received 17 May 2021
Revised 19 July 2021
Accepted 15 October 2021
Available online 17 October 2021

ABSTRACT

Efforts toward sustainable engineering have expanded the system boundary to consider the life cycle along with economic and environmental objectives. Many design problems involve selection of reactions from many alternatives, and systematic methods for choosing reaction pathways have also been developed. To prevent burden shifting, selection of reaction pathways needs to consider activities at larger scales of the life cycle and economy. In this work we develop such a framework. The resulting P2P-RNFA framework combines Reaction Network Flux Analysis (RNFA) and Process-to-Planet (P2P) frameworks by means of a novel multi-scale multi-objective optimization formulation. This framework is capable of designing chemical reaction pathways, life-cycle value-chains, and economic cash flows in an integrated manner from a superstructure network containing alternative solutions. Application to synthesizing industrial biofuel, 3-methyl-tetrahydrofuran, demonstrates the method and its features. The P2P-RNFA framework can design novel chemical pathways while being mindful about shifting impacts to larger scales.

1. Introduction

For decades, design and development of novel reaction pathways and products has predominantly focused on the narrow boundary of traditional engineering. Chemistries and mechanisms have been chosen for scale-up on the basis of green chemistry principles (Anastas and Eghbali, 2010) and empirical rules. These include designing for energy efficiency and waste prevention, using safer solvents and catalysis over reagents, reducing the number of reaction steps and derivatives, etc. Although these principles are sound and often effective, they rely on local or reductionist objectives that pertain only to small system boundaries of chemical production facilities. Such localized approaches can lead to unintended harm to the environment, often resulting from shifting impacts outside the system boundary of the chemical plant and its vicinity(Bakshi, 2019). Additionally, it is difficult to evaluate a large set of alternative pathways using analysis and assessment methods. Therefore, finding the best novel reaction pathways requires the following, (i) a systematic approach for ranking alternative pathways and (ii) methods and metrics to perform holistic evaluation of each pathway. The first short-coming has been countered by numerous mathematical models which involve superstructure optimization. Superstructures are a set of all the alternative solutions represented mathematically. These models include and are not limited to Metabolic Flux Analysis (Varma and Palsson, 1994; Lee et al., 2003), Reaction Network Flux Analysis (Voll and Marquardt, 2012), mass-action reaction network modeling (Shinar and Feinberg, 2011), Continuous flow stirred tank reactor (CFSTR) equivalence (Frumkin and Doherty, 2018; 2020), and Nested Evolutionary algorithms for selecting reactions with kinetic considerations (Degrand et al., 2019). In order to select a method for evaluating reactions, one needs to consider the trade-off (Unsleber and Reiher, 2020) between the number of permissible alternative pathways in the superstructure vs. the details of each pathway (such as kinetics and thermodynamics) being considered for evaluation. For instance, initial screening of reactions from a large network can be accomplished using simpler methods such as flux analysis, however detailed plant level design and selection might require dynamic modeling of reaction-separation systems through elaborate methods like mass-action networks, CFSTR, etc. Process systems engineering approaches have also explored design-

☆ This paper is inspired by the pioneering work of Professor George Stephanopoulos on multiscale analysis, modeling, and control of chemical process systems. With gratitude for his leadership, mentorship, and friendship.

* Corresponding author.
E-mail address: bakshi.2@osu.edu (B.R. Bakshi).

https://doi.org/10.1016/j.compchemeng.2021.107578
0098-1354/

ing flow-sheets from complex reaction networks (Ramapriya et al., 2018; Schweiger and Floudas, 1999; Garcia and You, 2015; Kokossis and Floudas, 1991), but their scope is usually designing optimal reaction and separation conditions for a number of pre-selected pathways. The second short-coming, which is related to the need for holistic evaluation is also being tackled by recent academic work. For instance, del Rio-Chanona et al. (2018) have optimized biopolymer reaction networks for the environmental impact objective using molar flux analysis and modeling. The impact objective is made holistic by associating energy requirements with CO_2 impacts along with emissions from reactions. Similarly, Zhuang and Herrgård (2015) have built a multi-scale framework called 'MuSIC' by linking dynamic flux models in metabolic engineering domain with economic models, used to calculate holistic objectives of the study. However, the MuSIC framework relies on scenario analysis and is limited in the number of candidate reaction pathways that can be considered. Expansion of the system boundary for Sustainable Process Design has also been carried out by the Process to Planet (P2P) framework (Hanes and Bakshi, 2015a), by integrating equipment scale design with hybrid life-cycle modeling approaches. These approaches link the process level life-cycle assessment with economic input-output models, thereby expanding the system boundary (Suh, 2004). The motivation behind this work is to build an integrated multi-scale framework to design reaction and value-chain pathways with alternatives from multiple scales. Such an integrated design framework can be very useful for initial ranking of possible reaction pathways along with their value-chain and life-cycle alternatives (Bakshi and Fiksel, 2003; Guillén-Gosálbez et al., 2019). Combining reaction networks with broader life-cycle and economic scale networks, would also facilitate holistic evaluation of alternatives and reduce the chance of shifting impacts across value-chains and spatial scales.

In this paper, we integrate reaction network flux analysis models with the process-to-planet framework to holistically design reaction pathways, life-cycle value-chain decisions and economic cash flows in an integrated manner. This multi-scale approach captures physical, monetary and information flows between various scales, as shown in Fig. 1. The reaction network design scale is a finer and more detailed scale which is added to the process-to-planet (P2P) multi-scale framework. This integration is enabled by the similarity between the frameworks of RNFA and hybrid life-cycle assessment models used in P2P. The original P2P framework is intended for design of process variables, it has to be modified for pathway design. The novel framework created as a result of this integration is called the P2P-RNFA framework. Therefore, P2P performs parametric optimization to find optimal process variables, whereas P2P-RNFA involves topological optimization of superstructure networks and their constituent pathways. As an added benefit of using optimization, the P2P-RNFA framework facilitates systematic evaluation of a large number of possible pathways and their combinations in a reasonable amount of time. When the economy scale consists of a large number of economic sectors, the P2P-RNFA optimization problem can become computationally expensive owing to many possible overlaps of economic sectors with the value-chain activities. To counter this, an approximation algorithm is also proposed which increases the computational tractability of the P2P-RNFA pathway design problem. The deviation of objective values due to the approximation are also examined by means of a simple illustrative example. Details about this example are also provided in the appendix for those readers who may be less familiar with the computational frameworks of LCA, P2P, and economic input-output modeling. Finally, the framework is applied for multiple objectives of sustainability, including environmental impact and economic cost. Integration with the economy scale allows easy formulation of cost objectives, such as life-cycle cost, which can be found as the cash flows from the economy scale to the value-chain, process and reaction scale. This reduces the burden of computing cost objectives based on external data and empirical relations, such as total annualized cost from estimates of energy requirements (Voll and Marquardt, 2012). The framework is also applied to a real-life case study pertaining to the production of a bio-fuel via alternative reaction pathways and raw materials. Various life-cycle based objectives are formulated to select the best pathway from the network. The results from multi-objective optimization are presented in the form of pareto fronts, which quantify the trade-offs between objectives and facilitate exploration of best 'compromise' solutions.

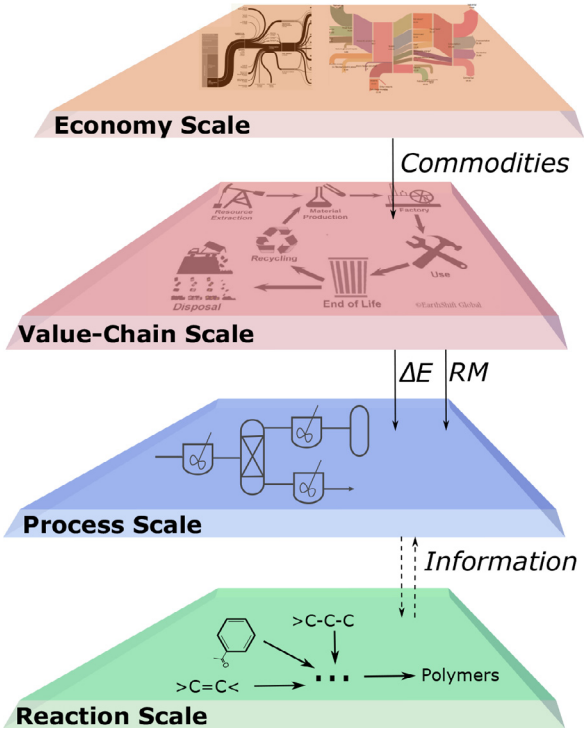

Fig. 1. Multi-scale modeling for sustainability.

The rest of this paper is structured as follows. Section 2 contains a background of methods and tools that are used in development of the P2P-RNFA framework, and insights about their similarities and differences. The background section also includes the reaction network flux analysis model in Section 2.1, and principles of hybrid life-cycle assessment and process-to-planet design framework in Section 2.2. Section 3 elaborates the novel framework developed by integrating the two models using multi-scale matrix representation (Section 3.1.1) of the network and necessary modifications to the P2P framework for pathway design (Section 3.1.2). An approximation algorithm for using the P2P-RNFA framework with large networks is also proposed in Section 3.1.3. Next, the framework is demonstrated for an illustrative example to explore errors from approximation and effects of changing the output from the network. The framework is then applied to a real-life case study in Section 4.1, which contains a reaction network to convert biomass derived Itaconic acid (IA) to a biofuel, 3-methyltetrahydrofuran (3-MTHF). The reaction network is linked to a life-cycle value-chain scale which contains data from life-cycle inventory databases for alternative procurement strategies of IA and energy production. Ultimately, these finer reaction and value-chain scales are connected to the coarser economy scale, built using an environmentally extended economic input-output model based on government databases. Applying the proposed framework to this case study yields optimal pathways, and trade-offs between

environmental and economic objectives, which are described in Section 4.2. This section also compares optimal reaction pathways obtained using RNFA with those from applying the P2P-RNFA framework to an expanded system boundary. Finally the importance of the framework and its usability for holistic reaction design and forecasting novel pathways are conveyed in Section 5.

2. Background

2.1. Reaction network flux analysis

There are several methods (Anastas and Eghbali, 2010; Varma and Palsson, 1994; Unsleber and Reiher, 2020) to rank and evaluate reaction pathways from a superstructure network of alternatives, albeit for localized objectives such as minimizing raw material needed, energy requirement, etc. Many efforts of this type are usually directed towards green chemistry and sustainable process design. Some advanced efforts (Ramapriya et al., 2018; Schweiger and Floudas, 1999) have also designed chemical networks while exploring effects of batch vs. continuous operation, kinetic and thermodynamic limitations, etc. Arguably, for initial vetting of reaction pathways, an elaborate and computationally expensive method is not needed. Instead, a more holistic evaluation with fewer complexities can yield more reliable results for sustainability, thereby eliminating a lot of pathways for further detailed evaluation. Therefore, we use a simple but highly effective linear programming approach called Reaction Network Flux Analysis (RNFA). RNFA was developed by Voll and Marquardt (2012) by building upon the flux balance methods (Schuster et al., 2000; Lee et al., 2000) in the metabolic engineering domain. RNFA is a rapid screening tool for systemically identifying and ranking reaction pathways using optimization. It is based on conservation of molar fluxes using the following equation,

$$\begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix} \begin{bmatrix} s_1^r \\ s_2^r \\ s_3^r \end{bmatrix} = 0 \tag{1}$$

A_2 contains stoichiometric coefficients of reactions, whereas A_1 and A_3 refer to pseudo-reactions required for describing inputs to and outputs from the system, respectively. The vector s^r contains the molar fluxes or reaction extents of all reactions. To separate the outputs, Eq. (1) can be transformed into the following,

$$\begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} s_1^r \\ s_2^r \end{bmatrix} = f^r \equiv A^r s^r = f^r \tag{2}$$

Here, A^r is the matrix of inputs and stoichiometric coefficients, in which the rows correspond to molecules and columns represent reactions. Vector f^r is the net output or final demand from the reaction network whereas vector s^r is a decision variable containing extents of individual reactions. Similarity between RNFA and life cycle assessment is apparent in Table 1. Maximum achievable reaction yield (Y_j) of each reaction j, can be specified as a constraint on reaction extent (s_j^r) for all molecules i in the RNFA framework (Voll and Marquardt, 2012). This yield constraint is formulated as follows,

$$s_j^r \le Y_j \frac{\sum_l A_{i,l}^r s_l^r + \sum_{m,m\neq j} A_{i,m}^r s_m^r}{A_{i,j}^r}, \quad \forall\ i \in \text{ Reactants} \tag{3}$$

The first summation in the above equation runs over all reactions l that consume the ith molecule, and m indicates all reactions, which produce i. This equation can be interpreted as specifying the maximum extent of reaction based on permissible yield for all molecules i. In doing so, the constraint corresponding to the limiting molecule within i will be active, and others will be redundant. $A_{i,j}^r$ denotes the stoichiometric coefficient of molecule i in reaction j. The yield constraint in (3) can be concisely stated as $s_j^r \le Y_j \gamma_i(\cdot)$, where γ_i denotes the molar flux of i divided by the stoichiometric coefficient of i in the reaction j. The flux balance constraint in Eq. (2) and yield constraint from Eq. (3) are constituents of the RNFA model (Voll and Marquardt, 2012). This model can be used to optimize reaction networks for any arbitrary objective $Z(\cdot)$, and is expressed as follows,

$$\begin{aligned} \min_{s^r, f} \quad & z := Z(\cdot) \\ \text{s.t.} \quad & A^r s^r = f^r \\ & s_j{}^r \le Y_j \gamma_i(\cdot)\ \forall\ i \\ & s^r, f^r \ge 0 \end{aligned} \tag{4}$$

Flux balance approaches have wide applicability in screening large reaction networks. For example, del Rio-Chanona et al. (2018), use RNFA to rank reaction pathways in a biopolymer production network. Konig et al. (2019), use RNFA for product and process design for renewable fuels, (iii). However, none of the extensions of RNFA have focused on expanding system boundary and design space to coarser scales, such as life-cycle or economy.

2.2. Hybrid life cycle modeling and P2P framework

Process life-cycle assessment (LCA) is a powerful tool to find environmental impact by accounting for the most important processes or value-chain (V) that is operating for satisfying the final demand of a particular product in the society. The computational structure of LCA is also based on linear programming. The fundamental equations of Process-LCA (Heijungs and Suh, 2002) are as follows,

$$\begin{aligned} \underline{A}\ \underline{s} &= \underline{f} \\ \underline{g} &= \underline{B}\ \underline{s} \end{aligned} \tag{5}$$

Here $\underline{A}$ denotes the technology matrix which contains process-level input output information of all processes in a particular value-chain. Rows in technology matrix correspond to products, whereas columns correspond to processes. The scaling vector $\underline{s}$ is the factor by which the technological activities contained in $\underline{A}$ are scaled to meet a final demand of $\underline{f}$, using the first part of Eq. (5). Emissions from the life-cycle of a product are found from $\underline{g} = \underline{B}\ \underline{s}$, where $\underline{B}$ is the intervention matrix containing flows to or from the environment. Each process in $\underline{A}$ and $\underline{B}$ has a basis or reference product, according to which all other elements of the column are scaled. $\underline{A}\ \underline{s} = \underline{f}$ can be thought of as a flow conservation equation for each product across nodes of the value-chain where splitting and merging occurs. The underbar is used to denote the value-chain scale matrices and vectors throughout the paper. This method is called process LCA because it is concerned with value-chain processes, and physical transformations occurring within them. LCA is a well-established method with many applications, and more information about its computational structure is available in the literature (Heijungs and Suh, 2002; da Luz et al., 2018; Azapagic, 1999). Eq. (5) and Table 1 convey that the computational frameworks of RNFA and LCA are very similar, which facilitates a smooth integration. This property is exploited by Motianlifu (2017), who merge the reaction scale with the value-chain scale. However, the life-cycle model is only used for assessment and not topological optimization of value-chains.

Process LCA relies on inventory data about specific processes. Such data may not always be available for all processes, particularly those that are less common. In addition, the infinite size of the full life cycle network means that a finite system boundary needs to be defined by focusing on the most important processes. The approach of hybrid LCA combines process LCA at the value chain scale with environmentally-extended input-output models

Table 1
Mathematical formulation of modeling approaches in Section 2. Underbar and overbar notations refer to the value chain scale and the economy scale, respectively. The notation without bar represents the equipment or reaction scale. The double bar notation refers to multiple scales.

Model	Transaction Equation	Intervention Equation	Comments
Process-LCA	$\underline{A}\underline{s} = \underline{f}$	$g = \underline{B}\underline{s}$	Describes the life-cycle value chain
EEIO	$(I - \overline{A})\overline{s} = \overline{f}$	$g = \overline{B}\overline{s}$	Economy Scale cash flow between sectors
Hybrid LCA	$\begin{bmatrix} I-\overline{A}^* & -\underline{X}_u \\ -\underline{X}_d & \underline{A} \end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\end{bmatrix} = \begin{bmatrix}\overline{f}\\ \underline{f}\end{bmatrix}$	$\overline{\underline{g}} = \begin{bmatrix}\overline{B}^* & \underline{B}\end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\end{bmatrix}$	Integration of life-cycle and economy scales, used for assessment
P2P	$\begin{bmatrix} I-\overline{A}(\{z\})^* & -\underline{X}_u & -X_u^E \\ -\underline{X}_d & \underline{A}(\{z\}) & -X_u^V \\ -X_d^E & -X_d^V & X(\{z\}) \end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\\ s\end{bmatrix} = \begin{bmatrix}\overline{f}\\ \underline{f}\\ f\end{bmatrix}$	$\overline{\underline{g}} = \begin{bmatrix}\overline{B}^* & \underline{B} & B\end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\\ s\end{bmatrix}$	Integrating hybrid LCA with equipment level design, used for design of process variables z
RNFA	$A^r s^r = f^r$	-	Reaction Network design using flux balances
P2P-RNFA	$\begin{bmatrix} \overline{I}-\overline{A}^*(\{\overline{s}\}) & -\underline{X}_u & -X_u^E \\ -\underline{X}_d & \underline{A}^*(\{\overline{s}\}) & -X_u^V \\ -X_d^E & -X_d^V & A^r \end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\\ s^r\end{bmatrix} = \begin{bmatrix}\overline{f}\\ \underline{f}\\ f^r\end{bmatrix}$	$\overline{\underline{g}} = \begin{bmatrix}\overline{B}^*(\{\overline{s}\}) & \underline{B}^* & B\end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\\ s^r\end{bmatrix}$	Integrated reaction, value-chain and economic network design

(EEIO) Leontief (1970) at the economy scale. The EEIO model covers activities at a coarser scale than the value chain scale but is more complete since its system boundary can cover entire nations or even the whole world. Unlike process-LCA, the EEIO model considers monetary flows in between economic sectors using a transactions matrix, $\overline{A}$. The cash flow is conserved using the equation $(I - \overline{A})\overline{x} = \overline{f}$, where I is an identity matrix, $\overline{x}$ is a vector of throughputs to or from an economic sector, and $\overline{f}$ is a vector containing final demand from economic sectors to society. This is supplemented with an interventions equation, $\overline{B}\overline{x} = g$, which is used to calculate emission flows to and from the environment, resulting from meeting the final demand, $\overline{f}$. $\overline{x}$ can be interpreted as the scaling vector of each economy sector in a monetary unit ($\overline{s}$), and thus, $\overline{x}$ and $\overline{s}$ are interchangeable. The bar over matrices $\overline{A}$ and $\overline{B}$ denote that the matrices belong to the economy scale. Noticing the similarities in the linear computational structure, Suh (2004) developed hybrid LCA models which link the LCA models at the value chain scale (process LCA) and economy scale (EEIO LCA). Such integration expands the focus of process-LCA and avoids shifting impacts outside its system boundary. These hybrid methods essentially use the following two concepts for the integration.

1. *Off diagonal cut-off flows*: The hybrid LCA framework is formulated by considering process LCA and EEIO as two tiers of a multi-scale matrix. The technology matrix of process LCA ($\underline{A}$) and Leontief matrix ($I - \overline{A}$) are at the diagonal of this multi-scale matrix. The off-diagonal elements capture flows between the scales, also known as cut-off flows. Hybrid LCA can therefore be represented as follows,

$$\begin{bmatrix} I-\overline{A}^* & -\underline{X}_u \\ -\underline{X}_d & \underline{A} \end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\end{bmatrix} = \begin{bmatrix}\overline{f}\\ \underline{f}\end{bmatrix}$$
$$\overline{\underline{g}} = \begin{bmatrix}\overline{B}^* & \underline{B}\end{bmatrix}\begin{bmatrix}\overline{s}\\ \underline{s}\end{bmatrix} \tag{6}$$

   $\underline{X}_u$ denotes the upstream cut-off flows from economy to life-cycle value-chain scale, whereas $\underline{X}_d$ denotes the downstream cutoffs. Equality constraints in Eq. (6) ensure that inter-scale flows represented by off-diagonal matrix elements do occur for meeting the final demands. The sub-matrices on the diagonal can either be obtained directly as direct requirements (EEIO model) and technology (process-LCA) matrices, or can be computed using make (V) and use (U) matrices (Suh et al., 2010; Konijn and Steenge, 1995). Make and Use matrices contain actual flows within the economy and value-chain as recorded/observed by repository collectors. An introduction to these matrices is provided in the Appendix. Hybrid LCA captures larger amounts of environmental flows than process LCA due to its connection with the economy scale. Typically EEIO models are very elaborate and contain complex connections within the economy, which broadens the system boundary of LCA to a great extent.
2. *Disaggregation equations*: Since the value-chains captured by process-LCA are ultimately part of the economy, simple integration using the tiered approach shown in Eq. (6) can lead to double-counting of emissions and resource use. To rectify this, Suh (2004), suggested disaggregating the economy to remove the parts corresponding to the value-chain. This is done by correcting the make and use matrices of the coarser scale, which contain actual flows recorded from the economy scale. Disaggregation uses permutation matrices to indicate whether processes from the smaller (or finer) scales are double-counted in certain economic sectors. An additional price vector is needed to convert flows in physical units at the value chain scale, to their monetary values in the economy scale. Disaggregation methods are further elaborated in Section 3.1.2, since our framework inherits these disaggregation equations and modifies them for compatibility with the design problem. Details on disaggregation can also be found in the hybrid LCA and P2P literature (Suh, 2004; Suh and Huppes, 2005; Hanes and Bakshi, 2015a).

The approaches stated so far in this subsection have been used widely to 'assess' value-chains holistically (Yue et al., 2016; Gao and You, 2018). Designing for sustainability is a non-trivial extension of assessment because it requires new optimization formulations and carefully constructed design spaces. The Process to Planet (P2P) framework is developed by Hanes and Bakshi (2015a) and it facilitates Sustainable Process Design by integrating hybrid LCA methods with equipment scale engineering models. This results in an additional tier corresponding to the equipment scale added to hybrid LCA. This equipment scale contains flows (material and energy) between various unit operations in a chemical plant and their physico-chemical models. The P2P formulation is shown in Table 1, wherein equipment scale design focuses on process variables ($\{z\}$) constrained by a set of equipment level equations such as material energy-balances, etc. Thus, the P2P framework is capable of obtaining optimal plant level designs while considering interactions with the value-chain and economy scales. It has been implemented for case studies related to manufacturing and sup-

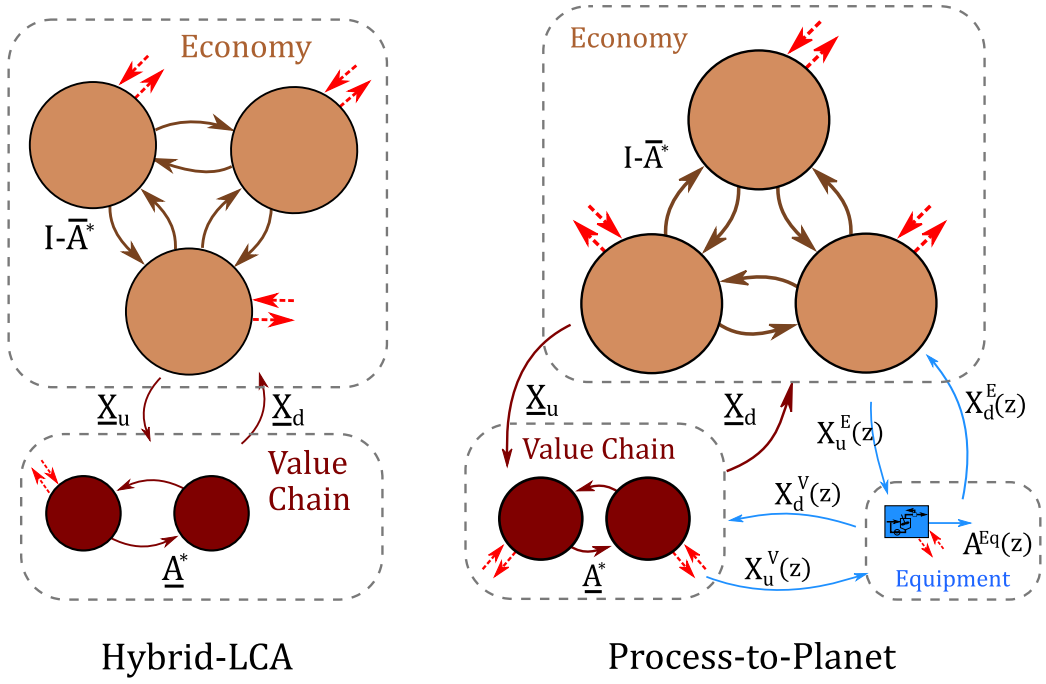

Fig. 2. State-of-the-art multi-scale frameworks for systainability assessment and design (a) Hybrid life-cycle assessment and (b) Process-to-planet (P2P) framework.

ply chains of corn ethanol and bio-fuels (Hanes and Bakshi, 2015b; Ghosh et al., 2019), and has been modified to incorporate market effects using the rectangular choice model (Lee et al., 2019). Whether to include economy scales in hybrid models itself is widely disputed (Pomponi and Lenzen, 2018; Yang et al., 2017), and recent work (Ghosh and Bakshi, 2020) has addressed this issue by quantifying trade-offs between complexity and uncertainty while including the economy and value-chain scales. However, this is not within the scope of our study and we do not probe the effects of excluding or aggregating the economy scale. Formulations of hybrid life-cycle assessment and process-to-planet framework are provided in Table 1, and the representative diagrams for the two methods are shown in Fig. 2.

The P2P framework has been used so far for process level designs, and not pathway and reaction network designs. Our previous work(Thakker and Bakshi, 2021) focused on implementing pathway design for value-chain networks, whereas in this work, we have expanded the system boundary to include the finer reaction scale as another tier in the model as shown in Fig. 1. The interaction between the process scale and reaction scale is considered to be through information flows, e.g. choice of reactions which develop the process flowsheet. For ease of demonstration and without loss of generality, we have limited the scope of the process scale to contain only reactors and no separators, heat exchangers, or other equipment. Therefore, the process scale can be substituted with the reaction scale. We call this novel framework 'P2P-RNFA' since it is formulated by convergence of the two methods. The following section on the proposed methodology provides greater insight into the formulation, tractability of the optimization and suggested approximations.

3. Methodology

This section describes the P2P-RNFA formulation, followed by a small illustrative example. Approximation strategies for larger problems are also proposed. The overall approach for pathway design by the P2P-RNFA framework is depicted in Fig. 3. It demonstrates implementation of topological design of multi-scale networks with alternative pathways and the integration of various scales in finding the optimal reaction scheme.

3.1. Integrating RNFA and P2P frameworks

To facilitate multi-scale pathway design for expanding the system boundary of reaction network design, we integrate the Reaction Network Flux Analysis and Process to Plant frameworks. In essence, an additional reaction scale has to be added to the multi-scale network as shown in Fig. 1, wherein information exchange has to occur between the equipment and reaction scales. Equipment scale can contain more unit operations such as separators, distillators, heat exchanging networks, etc. However, in this study we merge the equipment and reaction scales by assuming that the equipment scale contains only a series of reactors corresponding to each reaction. By virtue of the RNFA framework, unconverted reactants and by-products are assumed to be separated by themselves. However, separation processes can be added without much difficulty by replacing RNFA with the Process Network Flux analysis (PNFA) Ulonska et al. (2016), which has a similar computational structure. Therefore, this method can be used for initial vetting of reactions from both holistic life-cycle and product yields perspective. Our focus in this work is on RNFA.

3.1.1. Multi-scale matrix

The transaction or flow conservation equation for the P2P framework (Table 1) is modified to include the RNFA model in the bottom tier, as shown below.

$$\begin{bmatrix} \bar{I}-\bar{A}^*(\{\underline{s},s^r\}) & -\underline{X}_u & -X_u^E \\ -\underline{X}_d & \underline{A}^*(\{\underline{s},s^r\}) & -X_u^V \\ -X_d^E & -X_d^V & A^r \end{bmatrix} \begin{bmatrix} \bar{s} \\ \underline{s} \\ s^r \end{bmatrix} = \begin{bmatrix} \bar{f} \\ \underline{f} \\ f^r \end{bmatrix}$$
$$\equiv \bar{X}(\{\bar{s}\})\ \bar{s} = \bar{f} \tag{7}$$

$$\begin{bmatrix} \bar{B}^*(\{\underline{s},s^r\}) & 0 & 0 \\ 0 & \underline{B}^*(\{s^r\}) & 0 \\ 0 & 0 & B \end{bmatrix} \begin{bmatrix} \bar{s} \\ \underline{s} \\ s^r \end{bmatrix} = \begin{bmatrix} \bar{g} \\ \underline{g} \\ g^r \end{bmatrix}$$
$$\equiv \bar{B}(\{\bar{s}\})\ \bar{s} = \bar{g}$$

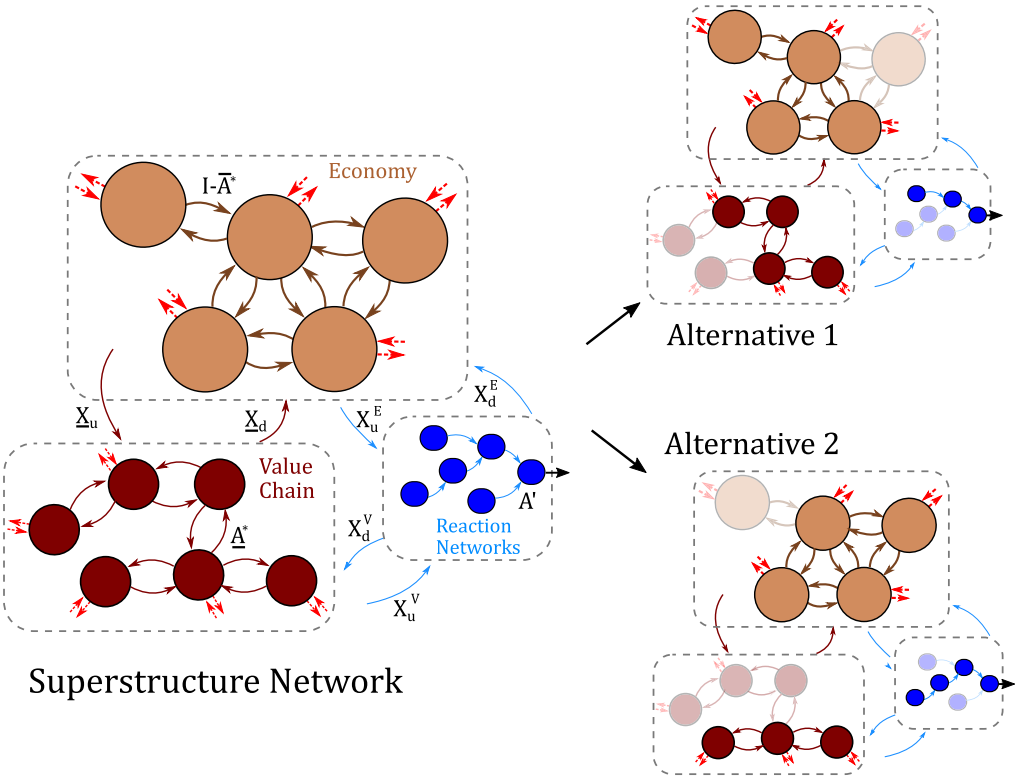

Fig. 3. Pathway design problem with integration of scales: comparing two alternative pathways within a superstructure network. Flows within and between scales are expressed as matrices and are labelled near the arrows representing the flows.

Similar to the P2P framework described in Section 2.2, Eq. (7) ensures intra and inter-scale flow conservation. The diagonal elements contain intra-scale flow coefficients and off-diagonal elements indicate cut-off or inter-scale flows. As in P2P and hybrid LCA, the top two diagonal elements represent the economy scale Leontief matrix (dimensions: $e_1 \times e_2$) and life-cycle technology matrix (dimensions $v_1 \times v_2$). The bottom diagonal sub-matrix corresponds to the reaction scale, and A^r contains the stoichiometric coefficients (dimensions: $r_1 \times r_2$) similar to the RNFA model in Section 2.1. The off-diagonal elements represent the cut-off flows between the three scales. Typically, the upstream cut-offs ($\underline{X}_u$, X_u^E and X_u) are very sparse with only few non-zero elements, indicating the economy cash flows into the value-chain or reaction scales, and raw material coming into the reaction scale from the value-chain scale. The upstream cut-offs to the reaction scale (X_u^V) can be computed by matching value-chain product flows with the input flows present in the A_1 sub-matrix shown in Eq. (1) of the RNFA method. Downstream cut-offs (X_d's) are flows from finer scales to coarser scales and are usually zero matrices for linear networks. This is because outputs of the reaction scale are usually final demands ($\overline{f}$) outside the system and do not alter the life-cycle and economy scale operation. Downstream cut-offs will be non-zero for circular systems such as reaction scale generating by-products that return to the system, or circular economy problems with products of chemical recycling returning back to the value-chain for use. The multi-scale transactions equation will be referred to as $\overline{\underline{X}}(\{\overline{\underline{s}}\})\ \overline{\underline{s}} = \overline{\underline{f}}$ for conciseness. Multi-scale matrices $\overline{\underline{X}}$ and $\overline{\underline{B}}$ are functions of the scaling factor because of conditional disaggregation required for the pathway design problem. This dependence on the scaling vector is elaborated in the next section.

3.1.2. Modified disaggregation equations

For any multi-scale framework, double-counting of environmental impacts and flows is imminent and averted using disaggregation, which has been implemented in hybrid-LCA and P2P frameworks described in Section 2.2. Essentially, this involves removing certain amounts of technological and environmental flows from a coarser scale, when it overlaps with a finer scale. In current state of the art hybrid life-cycle modeling frameworks such as P2P, the amount to be dis-aggregated is known apriori, within the make and use matrices of the value-chain ($\underline{V}$, $\underline{U}$) and equipment scales (V, U). Therefore make ($\overline{U}$) and use ($\overline{V}$) matrices of the economy scale can be modified by subtracting those amounts of flows (Hanes and Bakshi, 2015a; Suh, 2004), as shown in the following equations.

$$\overline{V}^* = \overline{V} - \hat{p}(\underline{P}_P)^{\mathrm{T}}\underline{V}(\underline{P}_F)^{\mathrm{T}} - \hat{p}\left(P_P^E\right)^{\mathrm{T}}V\left(P_F^E\right)^{\mathrm{T}} \tag{8}$$

$$\overline{U}^* = \overline{U} - \left(\hat{p}\underline{P}_F\underline{U}P_P + \underline{p}\underline{P}_F X_d + \underline{X}_u P_P\right) - \left(\hat{p}P_F^E U P_P^E + \hat{p}P_F^E X_d^E + X_u^E P_P^E\right) \tag{9}$$

In these equations, the overlap between scales is removed by the negative terms. The resulting matrices from disaggregation are called corrected matrices and are identifiable through the star (*) superscript. Similar disaggregation can be performed for equipment scale flows that are double-counted in the value-chain make and use matrices. If not available directly, original (non-star) reaction and value-chain scale make and use-matrices can be found by considering only positive and negative values, respectively. In Eqs. (8), (9), P_F are the flow permutation matrices and P_P are process permutation matrices. Permutation matrices with underbars indicate value-chain overlap with economy scales, whereas superscripts E and V indicate equipment scale overlap with economy and value-chain respectively. The interpretation of permutation

matrices is as follows. The cells in the $\underline{P}_P$ matrix are binary variables $\underline{p}_{P^{v_2 \times e_2}}$ which take the value 1 if a value-chain process v_2 is part of the economy sector e_2. Similarly cells of and $\underline{P}_F$ are populated with $\underline{p}_{F^{e_1 \times v_1}}$ which take the value 1 if a value-chain product v_1 is generated by the economy sector e_1. $\hat{p}$ is a square matrix (with dimensions $v_1 \times v_1$) with diagonal elements containing the prices of value-chain products v_1, and off diagonal elements are zero. Once the disaggregated make and use matrices are calculated from Eqs. (8), (9), disaggregated direct requirements matrix and technology matrices are calculated as follows,

$$\begin{aligned} &\overline{A^*} = \overline{U}^* (\overline{V}^{*T})^{-1} && \text{Direct requirements matrix} \\ &\underline{A^*} = \underline{V}^{*T} - \underline{U}^* && \text{Technology Matrix} \end{aligned} \tag{10}$$

Unlike P2P or hybrid LCA, through this P2P-RNFA framework we are solving a pathway design problem to find optimal alternatives from a multi-scale superstructure network as shown in Fig. 3. Thus, the amount to be dis-aggregated due to double-counting depends on whether the pathway is actually operational, i.e. has a non-zero scaling factor. As a result, the extent of disaggregation from the economy scale direct requirements matrix ($\overline{A}$) must depend on value-chain scaling factors and reaction extents, and similarly the technology matrix of the value-chain must depend on reaction extents. This dependence is also reflected in multi-scale matrices from Eq. (7), wherein the diagonal elements of the coarser scale are a function of scaling factors of the finer scale. These dependencies are modeled within the disaggregation equations, shown below.

$$\overline{V}^*(\{\underline{s}\}) = \overline{V} - \hat{p}\left(\underline{P}_P \hat{\underline{s}}\right)^{\mathrm{T}} \underline{V} (\underline{P}_F)^{\mathrm{T}} - \hat{p}\left(P_P^E \hat{\underline{s}}\right)^{\mathrm{T}} V \left(P_F^E\right)^{\mathrm{T}} \tag{11}$$

$$\begin{aligned} \overline{U}^*(\{\underline{s}\}) = \overline{U} - &\left(\underline{P}_F \hat{p} \underline{U}\ \hat{\underline{s}} \underline{P}_P + \underline{P}_F \overline{P} \hat{\underline{s}} \overline{P}^T\ \underline{p} \underline{X}_d + \underline{X}_u \hat{\underline{s}} \underline{P}_P\right) \\ &- \left(P_F^E \hat{p} U \hat{s^r} P_P^E + P_F^E P^r \hat{s^r} P^{rT} \hat{p} X_d^E + X_u^E P_P^E \hat{s^r}\right) \end{aligned} \tag{12}$$

Each subtracted term in the disaggregation equations denotes the overlapping or double-counted flow removed from the economy scale. These are similar to the original hybrid-LCA Eqs. (8)–(10). However, in these modified equations, the amount subtracted depends on the scaling factor of the process being disaggregated due to the diagonalized scaling vectors, $\{\underline{s}, s^r\}$. The interpretation of permutation matrices and price vectors remain the same as in Eq. (8). An additional value-chain process-product permutation matrix $\underline{P}$ is also required for proper mapping of overlapping processes with corresponding products in the value-chain scale. Each element of this matrix, $\underline{P}_{v_1, v_2}$ is 1 only if the value-chain product v_1 is an output of the value-chain process v_2. Typically, for value-chains built using life-cycle inventory $\underline{P}$ matrix is expected to be an identity matrix, unless there are any by-products and recycle flows. The permutation matrices and product-process matrices with superscripts E (P_P^E, P_F^E and P^r) are used for disaggregation of the reaction scale from the economy scale, and can be found correspondingly. For example, $p_{P^{r_2 \times e_2}}^E$ is 1 only if the reaction scale product r_2 is also produced by the economy sector e_2.

In a case where reaction scale elements overlap with the value-chain, another set of disaggregation equations are required, and they can be modeled in the same manner as Eqs. (11) and (12). However, for this tier of disaggregation, only three terms corresponding to reaction scale double-counting, will be subtracted from the value-chain make and use matrices. Additionally, the permutation matrices will have superscript 'V' and a different interpretation, i.e. reaction elements being disaggregated from value-chain processes. Interventions from the economy scale (represented as $\overline{R}$ matrix) consisting of emissions and resource use from economic sectors also need to be disaggregated to remove overlaps with value-chain and reaction scales. Similarly, interventions from value-chain scale also need to be disaggregated to remove reaction scale overlaps. This is done in a manner similar to Eqs. (11) and (12), as shown below.

$$\begin{aligned} &\overline{R}^*(\{\underline{s}, s^r\}) = \overline{R} - \underline{B} P_P \hat{\underline{s}} - B P_P^E \hat{s^r} \\ &\overline{b}_{e_2}^*(\{\underline{s}\}) = \frac{1}{\left(\overline{V}^{*T}(\{\underline{s}\}) 1\right)_{e_2}} \overline{r}^*(\{\underline{s}\})_{e_2}, \quad e_2 \in \text{Economy sectors} \\ &\underline{B}^*(\{s^r\}) = \underline{B} - B P_P^V \hat{s^r} \end{aligned} \tag{13}$$

The make, use and interventions matrices, one corrected for disaggregation using Eqs. (11)–(13), are then used to compute transaction matrices using the following equation, which in-turn populate the multi-scale matrix $\overline{\underline{X}}(\{\overline{\underline{s}}\})$.

$$\begin{aligned} &\overline{A}^*(\{\underline{s}, s^r\}) = \overline{U}^*(\{\underline{s}, s^r\})[\overline{V}^{*T}(\{\underline{s}, s^r\})]^{-1} && \text{Direct requirements matrix} \\ &\underline{A}^*(\{s^r\}) = \underline{U}^*(\{s^r\}) - [\underline{V}^*(\{s^r\})]^T && \text{Technology Matrix} \\ &A^r = V^{rT} - U^r && \text{Stoichiometric Matrix} \end{aligned} \tag{14}$$

3.1.3. P2P-RNFA optimization formulation

The P2P-RNFA framework is intended for integrated multi-scale pathway design as shown in Fig. 3. The decision variable is the scaling vector, $\overline{\underline{s}} = [\overline{s} \quad \underline{s} \quad s^r]^T$, which denotes the scale-up and selection of each economic sector, value-chain process and reactions. Flow conservation within and across scales is ensured by specifying multi-scale transactions represented by Eq. (7) as a constraint. In addition, the net interventions or impact, found as $\overline{\underline{g}} = \overline{\underline{B}}(\{\overline{\underline{s}}\})\ \overline{\underline{s}}$, is also specified as a constraint. Additional physico-chemical governing equations such as material-energy balances, separation efficiencies, policies, human behaviors, etc. are also set up as constraints on scaling vectors (decision variables) i.e., $H(\{\overline{\underline{s}}\}) \geq 0$. Chemical reaction yields for each reaction (r_2) in set R, are used to constrain reaction extents, according to the general form of Eq. (3), as shown below.

$$s^r{}_{r_2} \leq Y_{r_2} \frac{\sum_{r \in R} A_{i,r}^r s^r{}_r}{1 + A_{i,r_2}^r Y_{r_2}} \ \forall\ i \in r_1 \tag{15}$$

This equation has been formulated by rearranging the terms in Eq. (3) so as to club stoichiometric coefficients together, and facilitate easier constraint formulation. In this equation, Y_{r_2} is the reaction yield of reaction r_2 with respect to each reactant $i \in r_1$. The value of Y_j is obtained from experimental studies on reaction kinetics or through thermodynamic equilibrium assumptions. A^r is a matrix of reaction stoichiometries with rows (indexed by r_1) indicating molecules/reagents and columns indicating reactions (indexed by r_2). By introducing the yield constraint (generally stated as $s_j^r = Y_j \gamma_i(\cdot)$), we limit the reaction extent by its maximum possible yield. Further, the modified

Table 2
Algorithm to apply P2P-RNFA framework with an approximation.

P2P-RNFA Approximation Algorithm	
1	Initialize: Pathway counter $k = 0$ and Upper-bound vector $u^k = \overline{0}$.
2	Build the multi-scale matrices $(\overline{X}, \overline{B})$ without disaggregation; e.g. $\overline{A}^* = \overline{A} = \overline{U}\,\overline{V}^{T^{-1}}$.
3	Set a dummy objective $z = \text{const}$; solve the P2P-RNFA problem using Eq. (17).
4	Obtain $\underline{s}^\dagger$ as optimal value of scaling vector which denotes a feasible pathway solution.
5	Update upper bound of scaling factor of each column j using $u_j^{k+1} = \max(u_j^k,\ \underline{s}_j^\dagger)$.
6	Identify column index in (v_2, r_2) with highest value in u^{k+1}; $j^{\max} = \text{argmax}(u^{k+1})$.
7	Perform parametric disaggregation for pathway k, using Eqs. (11)–(13), with $m = m^\dagger$ and $s^r = s^{r\dagger}$ from the optimal scaling vector $\underline{s}^\dagger$; thereby obtaining $\overline{\underline{X}}_k$.
8	Run the optimization with added constraint $\underline{s}_{j^{\max}} = 0$.
9	If infeasible, return to step 6 and find index in u without the element at index $j^{\max}$.
10	If solution obtained, return to step 4. Continue 4–9 until $u^{k+2} \approx u^{k+1} \approx u^k$, with a convergence of $\epsilon = 5\%$.
11	Assign binary variables y_k corresponding to each pathway k; to indicate whether a process has been disaggregated or not.
12	Add constraint for every binary variable, $y_k \times u_j^k \geq \underline{s}_j$.
13	New optimization formulation with assumption that disaggregated multi-scale transactions matrix are additive: $\min_{\underline{s},y}\ z := Z(\cdot)$ s.t. $\frac{\sum_k y_k \overline{\underline{X}}_k}{\sum_k y_k}\ \underline{s} = \overline{\underline{f}}$; $\frac{\sum_k y_k \overline{\underline{B}}_k}{\sum_k y_k}\ \underline{s} = \overline{\underline{g}}$; $y_k \times u_j^k \geq \underline{s}_j$; $H(\{\underline{s}\}) \geq 0$; $s_j^r \leq Y_j \gamma_i(\cdot)\ \forall\ i$; $\underline{s} \geq 0$ (18)
14	Here, $\overline{\underline{X}}_k$ & $\overline{\underline{B}}_k$ are parameters and the mixed integer non-linear programming problem with decision variables $\underline{s}$ and y, is much smaller and easier to solve.

disaggregation equations need to be set-up as constraints because the pathway design problem involves conditional disaggregation only if overlapping process in the finer scale is actually selected in the optimal pathway, as explained in Section 3.1.2). Eqs. (11)–(13), are expressed concisely as follows,

$$\begin{aligned} &\overline{U}^* = d_1(\overline{U}, \underline{s}, s^r), \quad \overline{V}^* = d_2(\overline{V}, \underline{s}, s^r), \quad \overline{B}^* = d_3(\overline{V}^*, \overline{R}, \underline{s}, s^r) \\ &\underline{U}^* = d_4(\underline{U}, s^r), \quad \underline{V}^* = d_5(\underline{V}, s^r), \quad \underline{B}^* = d_6(\underline{B}, \underline{s}, s^r) \end{aligned} \tag{16}$$

This concise notation brings out the dependence of corrected make and use matrices on finer scale scaling factors and reaction extents. The P2P-RNFA framework is built by including all the aforementioned constraints in the optimization formulation. The objective $(Z(\cdot))$ can be formulated from multi-scale flows, such as net product yield, life-cycle impact, life-cycle cost, chemical plant cost, social value etc. The resulting optimization formulation of the P2P-RNFA framework is as follows.

$$\begin{aligned} \min_{\underline{s}}\ & z := Z(\cdot) \\ \text{s.t.}\ & \overline{\underline{X}}(\{\underline{s}\})\ \underline{s} = \overline{\underline{f}} \\ & \overline{\underline{B}}(\{\underline{s}\})\ \underline{s} = \overline{\underline{g}} \\ & H(\{\underline{s}\}) \geq 0 \\ & s^r{}_j \leq Y_j \gamma_i(\cdot)\ \forall\ i \in r_1 \\ & \text{Disaggregation Eqs. (11)–(14)} \\ & \underline{s} \geq 0 \end{aligned} \tag{17}$$

This formulation is inherently a non-linear problem because $\overline{\underline{X}}(\{\underline{s}\})$ is a function of the scaling vector as shown in Eqs. (11)–(13); and is multiplied by the scaling vector in the first constraint. Additionally, the disaggregation equations contain reciprocals of the economy scale matrix shown in Eqs. (10) and (13), making the constraints highly non-linear. Due to multiple matrix operations such as inversion and multiplication on variable make and use matrices, it is expected that the design space of this NLP is non-convex. However, the complex nature of transformation prohibits us from computing the Hessian matrix over the domain to prove this hypothesis. We observe that despite the non-linearities, global optimization solvers such as BARON are able to find feasible solutions, albeit only for relatively small superstructure networks such as the illustrative example in Section 3.2 which has a multi-scale matrix $(\overline{\underline{X}})$ of dimensions 5×6. However, even with the most powerful NLP solvers, the framework takes large amounts of time to find solutions for real-life case studies wherein the entire economy scale is included as a direct requirement matrix, and has size > 380 rows and columns. It is expected that using aggregated economy scale matrices could make this problem solvable, but that can lead to loss of information of various economic sectors and increase uncertainty. Therefore, we propose a simple approximation algorithm in Table 2 for increasing the computational tractability of the problem.

3.1.4. P2P-RNFA approximation algorithm

This subsection describes the approximation algorithm to solve the P2P-RNFA framework for problems with large number of economic sectors and value-chain activities. This has been done by finding discrete pathways within finer scales, and obtaining corrected matrices based on disaggregation for that particular pathway meeting the prescribed final demand $\overline{\underline{f}}$. The approximation algorithm is provided in Table 2.

The first part of the approximation is focused on identifying distinct pathways of the reaction and value-chain scales, and indexing them using k. In steps 2–3 of the approximation, we optimize the multi-scale network without disaggregation for a dummy objective (e.g., $z =$ Constant). This assumes independent scales without any overlaps between them. The scaling vector corresponding to the optimal pathway

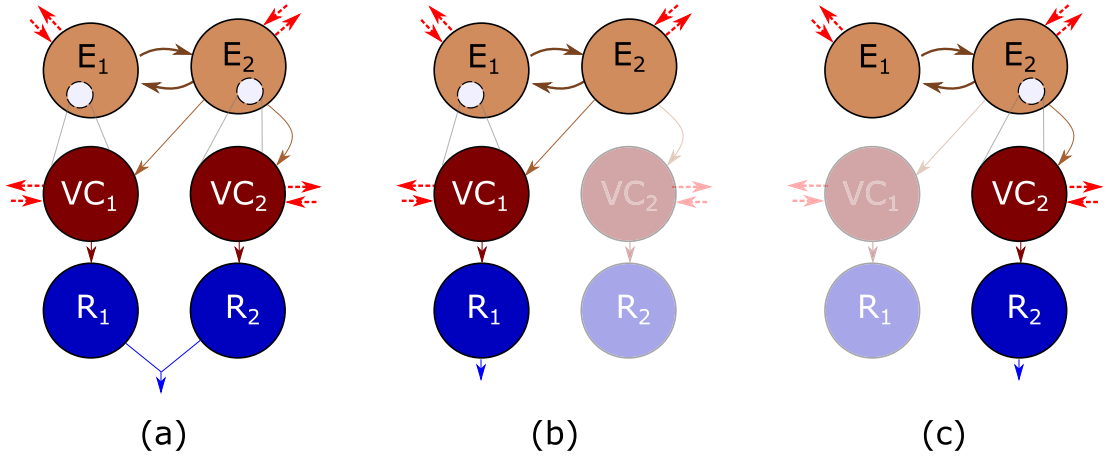

Fig. 4. Illustrative Example. (a) Superstructure Network containing both solutions. (b) Alternative 1 pathway. (c) Alternative 2 pathway.

is stored in $\underline{\bar{s}}^{\dagger}$, and the vector containing upper-bounds (u) of all scaling factors is updated. This upper-bound represents the maximum capacity at which the finer scale process operates and therefore the amount to be disaggregated from the coarser scale. Components of $\underline{\bar{s}}^{\dagger}$ ($m^{\dagger}$ and $s^{r\dagger}$) are used to find disaggregated multi-scale matrices for the k^{th} pathway, i.e. $\underline{\bar{X}}_k$. This is done using Eqs. (11)–(14) with $m = m^{\dagger}$ and $s^r = s^{r\dagger}$, thereby yielding parametric multi-scale matrices without any variables in them. Next, we iteratively remove the dominant scaling factor associated with the k^{th} pathway (by setting $\bar{s}_{j\max} = 0$) and re-optimize for the dummy objective (steps 6–10 of the algorithm). This exercise of finding all pathways ends when the upper-bound vector does not change despite iterative elimination of scaling factors, $\bar{s}_j$. After iterations, the algorithm yields disaggregated parametric multi-scale matrices ($\underline{\bar{X}}_k$) for each pathway k, and an upper-bound on all scaling factors stored in u^k. Finally, a new optimization formulation is constructed with additional binary decision variables y_k which take the value 1 if the k^{th} pathway is active, and 0 otherwise. This is ensured by adding the upper-bound constraint, $y_k \times u_j^k \geq \bar{s}_j, \ \forall j$.

The multi-scale transactions matrix is modified to be a weighted sum of parametric multi-scale matrices ($\underline{\bar{X}}_k$ and $\underline{\bar{B}}_k$) with weights y_k. The resulting constraint is $\frac{\sum_k y_k \underline{\bar{X}}_k}{\sum_k y_k} \bar{s} = \bar{f}$. The optimization problem hence posed in Eq. (18) in Table 2 is a Mixed-Integer Quadratically constrained problem (MIQCP). It is noteworthy that the number of variables in the formulation reduces dramatically from the order of $O[(e_2 + v_2 + r_2)^2]$ in Eq. (17) to $O[k + (v_2 + r_2)^2]$ in Eq. (18), which is a drastic difference, especially for large economy scales. Typically the number of pathways k are not more than $30 - 40$, whereas the size of un-aggregated economy scales from the EEIO model have sizes around 300×300. Thus, this approximation can be particularly fruitful for large case-studies, where one needs to include the economy scale for holistic evaluation, but also the disaggregated flows from finer scales are significant and cannot be ignored. For relatively small studies with less than $15 - 20$ pathways, it is also possible to solve the MIQCP posed by the approximation algorithm using widely available solvers such as Gurobi instead of BARON. However, in this study we use the BARON solver which uses branch and reduction methods to efficiently reach the global optimal solution, even for relatively large problems.

3.2. Illustrative example

In this illustrative example, the P2P-RNFA framework is applied to the system represented in Fig. 4. It serves the purpose of developing a conceptual model, and exploring the effect of approximation and disaggregation. In the illustrative network from Fig. 4, there are two reactions (R_1 and R_2) which convert two types of raw materials obtained from value-chain scale processes (VC_1 and VC_2), to a single product. Both the value-chain processes require inputs from two different economy sectors (E_1 and E_2). Additionally each of the value-chain processes is also present in one of the economy sectors and needs to be disaggregated (VC_1 from E_1, and VC_2 from E_2). We apply both the P2P-RNFA method and the approximation algorithm for this illustrative example.

Make and use matrices of the three scales, non-zero cut-off matrices and parameters such as permutation matrices, process-product matrices and price vectors are listed below,

$$
\begin{array}{lllll}
\text{Economy Scale} & \overline{V} = \begin{bmatrix} \$500 & 0 \\ 0 & \$700 \end{bmatrix} & \overline{U} = \begin{bmatrix} \$150 & \$200 \\ \$300 & \$100 \end{bmatrix} & \overline{R} = \begin{bmatrix} 200 & 800 \end{bmatrix} \\
\text{Value-chain Scale} & \underline{V} = \begin{bmatrix} 35 & 0 \\ 0 & 50 \end{bmatrix} & \underline{U} = \begin{bmatrix} 15 & 7 \\ 5 & 10 \end{bmatrix} & \underline{B} = \begin{bmatrix} 4 & 20 \end{bmatrix} \\
\text{Reaction Scale} & V = \begin{bmatrix} 1 \\ 2 \end{bmatrix} & U = \begin{bmatrix} 0 & 0 \end{bmatrix} & B = \begin{bmatrix} 0 & 0 \end{bmatrix} \\
\text{Cut-offs} & \underline{X}_u = \begin{bmatrix} \$4 & \$5 \\ 0 & 0 \end{bmatrix} & X_u^V = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.02 \end{bmatrix} & X_u^E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\
\text{Permutation matrices} & \underline{P}_F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \underline{P}_P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & P_F^E = P_P^E = [\,] \\
\text{Parameters} & p = \begin{bmatrix} \$0.01 & \$0.01 \end{bmatrix} & \overline{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & P^r = [\,]
\end{array}
\tag{19}
$$

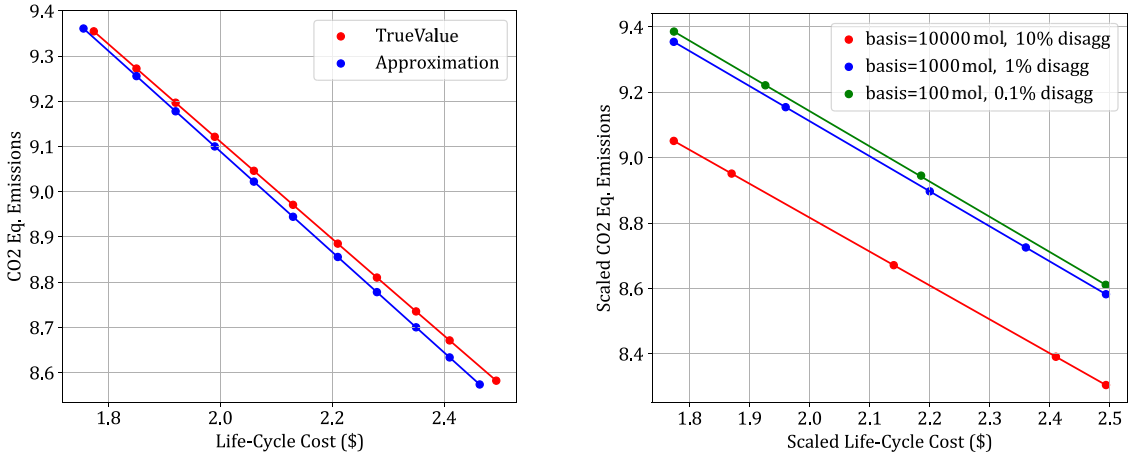

Fig. 5. Illustrative example results (a) Exploring the effect of approximation: Deviation of approximate pareto front from true pareto front is not large (b) Effect of changing basis on pareto front: Larger basis leads to more amount of disaggregation from economy leading to non-linear scaling of problem.

These sub-matrices and parameters are explained in greater detail with their interpretations by labeling flows of the superstructure network in Appendix 1. The multi-scale transaction matrix is obtained by using Eq. (7), and is a function of the scaling vector due to disaggregation Eqs. (11)–(13). Because there are only two alternative solutions of this illustrative superstructure network, they can be distinguished easily, as evident in Fig. 4 (b) and (c). The basis of a problem is defined as the net output required from the system, and is also known as 'functional unit' in the life cycle assessment literature (Bakshi, 2019; Heijungs and Suh, 2002). Setting a basis of 1000 moles (f_{r_1}) output from the reaction scale, the upper bounds of scaling vectors are found using the algorithm in Table 2. These are used to generate the multi-scale matrices corresponding to the two alternatives,

$$\begin{bmatrix} \text{Pathway 1 - Fig. 4 (b):} & \underline{\overline{X}}_1 = \begin{bmatrix} 0.7005 & -0.2857 & -4 & -5 & 0 & 0 \\ -0.5960 & 0.8576 & 0 & 0 & 0 & 0 \\ 0 & 0 & 20 & -7 & -0.01 & 0 \\ 0 & 0 & -5 & 40 & 0 & -0.02 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \\ \text{Pathway 2 - Fig. 4 (c):} & \underline{\overline{X}}_2 = \begin{bmatrix} 0.7000 & -0.2858 & -4 & -5 & 0 & 0 \\ -0.5992 & 0.8590 & 0 & 0 & 0 & 0 \\ 0 & 0 & 20 & -7 & -0.01 & 0 \\ 0 & 0 & -5 & 40 & 0 & -0.02 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \end{bmatrix} \tag{20}$$

The matrices corresponding to two alternative pathways are then used within the P2P-RNFA approximation formulation stated in equation (18). The value-chain and reaction scale sub-matrices, which are found in the bottom four rows and columns are identical, because they don't undergo disaggregation, whereas the economy scale sub-matrix is different for the two alternative pathways due to conditional disaggregation. The two objectives considered for this illustrative example are minimizing (i) CO_2 emissions and (ii) life-cycle cost. Every node in the illustrative example network emits a certain amount of CO_2, expressed in the $\underline{\overline{B}}$ matrix. Net CO_2 emissions from the network are calculated as $\underline{\overline{g}} = \sum_j \underline{\overline{B}}_{1,j}\, \underline{\overline{s}}_j$ and is the first objective. The second objective considered for minimization is life-cycle cost. It is possible to formulate cost objectives such as net-present value, total annualized cost, etc., but that would require additional data collection and empirical relations to calculate capital and operating cost. In the P2P-RNFA framework, life-cycle cost can be calculated easily as the sum of up-stream cut-off flows from the economy scale to the life-cycle system boundary ($\overline{X}_u \overline{s} + X_u{}^E s + X_u{}^E s$). This system boundary also includes the reaction scale, thereby allowing a systematic cost objective to rank reaction pathways.

The optimization framework stated in Eq. (17) and the approximation algorithm described in Section 3.1.3) have been applied for the two objectives, $kgCO_2$ emitted and life-cycle cost. Since, we aim to solve a multi-objective problem, as is the case for most sustainability problems, we have to develop a pareto front to characterize the trade-off between the two objectives. Pareto front development is done using the epsilon constraint method (Mavrotas, 2009). This front contains pareto-optimal points, which are the 'best' compromise or win-lose solutions. Solutions lying below the pareto front are infeasible, and the ones lying above are sub-optimal. The values of objectives depend on the functional unit or the basis of the system, i.e. amount of product required out of the system. Fig. 5 (a) explores the effect of approximation and deviation from the true pareto front and (b) demonstrates the effect of changing the basis or net output from the reaction scale.

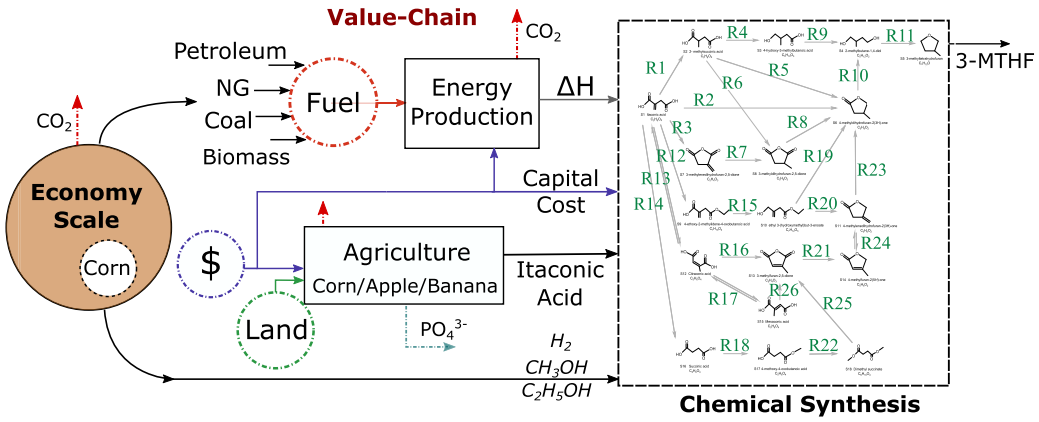

Fig. 6. Case Study, Superstructure of the reaction network with expanded system boundary.

The pareto fronts in Fig. 5 indicate the presence of trade-off between emissions and cost, as typically observed in sustainability problems. It is noticed that different pathways are selected at different extremes of the pareto front; Fig. 4 (b) for lowest CO_2 and (c) for lowest cost. This is because of the values in the make, use and intervention matrices (Eq. (19)), which indicate that the alternative in Fig. 4(c) has higher emissions. It is also evident from Fig. 5 (a) that the approximation algorithm performs really well, and the pareto front is very close to the actual pareto front. For this illustrative example, the computational time for developing these two fronts are not significantly different, however for the case study presented in Section 4.1, the framework without the approximation is not solvable even on a supercomputer due to large number of variables and non-linear constraints. Since, the economic and life-cycle scale matrices can be huge, we expect that the P2P-RNFA framework would have to be used with the approximation algorithm for quicker convergence, albeit with a small error percentage. Fig. 5 (b) presents pareto plots between scaled values of objective functions (scaled according to basis), to demonstrate the importance of disaggregation. For larger values of the basis, the amount of overlap would increase. This is equivalent to increasing size of the hollowed out portion within economy sectors (E_1 and E_2) in Fig. 4. It can be seen that the pareto curve shifts minimally with a ten-fold increase of the basis at lower percentage disaggregation, as compared to ten-fold shift at a higher level of disaggregation. This means that including conditional disaggregation is vital, especially if the finer scale scaling is considerably large in comparison to the coarser scale. This also proves that life-cycle impact is non-linearly scaled for the P2P-RNFA framework, whereas hybrid-LCA and process-LCA assume linear scaling with respect to the basis. This is because amount of disaggregation is a function of the scaling vector in a pathway design problem.

4. Case study

In this section, we demonstrate the P2P-RNFA multi-scale pathway design framework by applying it to a reaction network for converting biomass-based Itaconic acid (IA) to 3-methyltetrahydrofuran (3-MTHF), a potential biofuel. Voll and Marquardt (2012) have applied RNFA to a similar network. However, in this study we expand the system boundary by connecting the reaction scale to life-cycle, value-chain, and economy scales. We are thereby able to find a different set of reactions as compared to the results of Voll and Marquardt (2012). This is due to the more holistic nature of the P2P-RNFA framework. Fig. 6 contains the superstructure of all the alternative solutions within the expanded system boundary. We also perform multi-objective optimization to find optimal pathways for minimizing global warming potential, life-cycle cost, eutrophication potential and land area utilization. Additionally, a pareto front is developed to quantify the trade-off between environmental and economic objectives.

4.1. Description

The case study involves a reaction network containing 26 reactions to convert IA to 3-MTHF via various possible pathways, including 10 discrete options and their infinite fractional combinations. Each reaction has a maximum possible yield, which is set to 0.97 for all reactions arbitrarily due to scarcity of data and to match Voll and Marquardt (2012) for comparison. These reactions are shown in Fig. 6. The life-cycle value chain contains two major sections; energy generation and agriculture for production of IA. The alternative sources for energy production include petroleum, natural gas, coal and biomass. Energy generated from combustion of these fuels is used up by the endothermic reactions in the network. Net energy demand is estimated as the sum of enthalpies of reactions, thereby assuming heat integration. The impact generated and logistical aspects of using these fuels are considered within the value-chain scale. However, since the fuels are procured from the economic sectors, the impact associated with extraction and transport are included in the economy scale. Similarly, the alternatives for harvesting IA are present in the life-cycle scale, which include corn, apple and banana cultivation. Since Corn production is also part of the economy scale (hollowed out portion in Fig. 6), it needs to be disaggregated from the economy. As mentioned in Section 3.1.3, using the P2P-RNFA approximation, disaggregation is performed only when corn is actually chosen as part of the design. Various reagents required in the reaction network include hydrogen, methanol and ethanol, which are procured directly from the economic sectors,“325120 Compressed Gases” and “325190 Basic organic chemicals”. With around 22 types of molecules and 26 reactions, the dimensions of reaction scale matrix (A^r) are 22×26. The life-cycle value-chain scale ($\underline{A}$) matrix is of size 9×12 and is obtained from various life-cycle inventory databases and literature (Islam et al., 2016). The economy scale is represented by the environmentally extended economic input-output model, which is obtained from Input-Output accounts of 385 industries for 2018 provided by the United States Bureau of Economic Analyses (US BEA) U.S. Bureau of Economic Analysis, 2018 (2018). Thus, the Leontief matrix corresponding to the US EEIO model has the dimensions 385×385.

Multi-scale matrices are generated to describe the transactions or flows within the superstructure network shown in Fig. 6. Due

to the large size of these matrices, applying the P2P-RNFA framework (Eq. (17)) without the approximation is impossible. Thus we employ the algorithm proposed in Section 3.1.3 (Table 2), with two binary variables corresponding to whether disaggregation has occurred or not. Since the sum of the two variables is constrained to 1, there exists only one additional degree of freedom over the length of the scaling vector, which arises due to the disaggregation approximation. The length of scaling vector for this case is equal to the number of columns in $\underline{\overline{X}}$, which equals 425. As mentioned in Section 1, sustainability problems often have large numbers of stakeholders, implying that one must consider multiple objectives for the optimization. We have considered the following four objectives: Global Warming potential (GWP) in $kgCO_2eq$, Eutrophication potential (EP) in kgPeq, Land area utilization (Land) in m^2 and Life-cycle cost (LCC) in $. The first three objectives are midpoint indicators (M) of life-cycle impact and can be quantified using intervention flows $\bar{g}$, whereas LCC is the total amount of money sent from the economy scale to the value-chain and reaction scales through up-stream cut-offs. The objectives are expressed in terms of multi-scale flows and interventions, as shown below:

$$\begin{aligned} Z_1 &= \mathrm{M} = \textstyle\sum_l \phi_l^M \bar{g}_l \\ Z_2 &= \mathrm{LCC} = \textstyle\sum_i \sum_j \underline{X}_{ui,j} \bar{s}_j + \sum_i \sum_j X^E_{ui,j} \bar{s}_j \end{aligned} \tag{21}$$

Here ϕ_l^M denotes the characterization factors, which are parameters to convert an intervention l into respective equivalents of global warming potential, eutrophication potential, land area or any other midpoint indicator; listed as a superscript to ϕ_l^M. For example, for methane, global warming potential over 100 years is around 25, which means $\phi^{GWP}_{CH_4}$=25 $kgCO_2eq/kgCH_4$. Life-cycle cost is calculated from upstream cut-off flows from the economy scale to the value-chain and reaction scales. There are several other life-cycle midpoint indicators that can be of interest, such as acidification potential (AP), ozone depletion, ecotoxicity, water scarcity, etc. However, for this study we limit our scope to GWP, EP and land because of the methodological emphasis of this work, and because they seem to be most relevant for biomass-based raw material (Itaconic acid) and bio-fuel as a product. Since corn and fruit cultivation are parts of the value-chain, it is essential to make decisions based on EP and land use, whereas energy requirements for setting up chemical processing plants demands evaluation of emissions and global warming impact. Water use and scarcity are also likely to be important for this case study, and will be considered in future work. Ideally, design of such multi-scale systems must also consider social objectives, such as employment opportunities, urban development, social equity, etc. These considerations are part of on-going work.

The abundance of objectives implies presence of trade-offs and orthogonality. While trade-offs between environmental metrics themselves can be large and therefore crucial in the decision making, we focus on obtaining trade-offs between GWP and LCC due to methodological orientation of this work. These two objectives are also among those that are commonly used for decision making in the sustainability domain. Pareto optimal points capturing this trade-off are found using the epsilon constraint method for the two objectives. These pareto optimal points are the 'best' compromise or win-lose solutions and can be used to generate a pareto front.

4.2. Results and discussions

The results obtained by applying the P2P-RNFA framework to the case study by setting the final demand of 3-MTHF from the network as 1 ton per day; are described in this section. The framework is implemented in the General Algebraic Modeling (GAMS) language (Bussieck and Meeraus, 2004) and optimized using mixed integer non-linear programming (MINLP) solver, BARON v16.8 (Sahinidis, 1996). The optimization yields globally optimal solutions within reasonable time ($\leq$ 1 min) on an Intel Core i7-10510U CPU 2.3 GHz with 16 GB RAM. Results from applying the framework for these objectives are listed in Table 3 and presented in Figs. 7 and 8. Fig. 7 (a) contains the values of objective functions after normalizing by the maximum value among all the optimal solutions from Table 3. The trade-offs between various objectives are very evident. For example, the pathway corresponding to minimum land use has the highest LCC among the four optimal solutions. Similarly, the minimum LCC pathway has the highest GWP, and minimum GWP pathway has the highest eutrophication potential.

We have gone further to quantify the trade-off between the global warming potential and life-cycle cost objectives; thereby obtaining a pareto front shown in Fig. 7 (b). The point on the extreme left has the lowest LCC, but the highest GWP, and the point on the extreme right (bottom) has the lowest GWP, but at a large monetary expense. All the intermediate points on the pareto front (blue line) are 'win-lose' or compromise solutions. Based on the stakeholder's preference, a point on the pareto front can be selected subjectively as satisfactorily optimal in terms of both objectives. All the points above the pareto front are sub-optimal. It can be seen from the graph in Fig. 7 (b) that the pareto front undergoes a knee point at around (1.5E6$, 1.8E6 $kgCO_2eq$). We choose a pareto optimal point (denoted by Δ on the plot) close to the knee point as a compromise solution, wherein life-cycle cost incurred is close to optimal value with a marginal increase in GWP.

The pareto plot in Fig. 7 (b) represents trade-off on the objective space, but pathway decisions are to be made on the design space. Thus, it is crucial to observe the solutions chosen for the two objectives and derive inferences. Fig. 8 contains pathways in the reaction scale, life-cycle value-chain decisions and economy scale cash flows; as obtained for optimal GWP (a), LCC (b) and pareto optimal solution (c).

We observe from the non-zero reaction extents plotted in Fig. 8 (a) and (b) that the reaction pathways chosen for minimum GWP and minimum LCC are not identical. While the main reaction pathway is $R2 \rightarrow R10 \rightarrow R11$, the side reactions have subtle differences. For example, the minimum GWP pathway (Fig. 8 (a)) has a non-zero reaction extent for reaction id $R13$ whereas the extent for the same reaction is zero in the minimum LCC pathway. On the other hand, reaction id $R14$ has non-zero extent for minimal LCC but zero for minimum GWP. This means that production of 3-MTHF via Citraconic Acid is preferred for minimizing GWP, whereas the route via Succinic acid is preferred for minimizing cost. This situation arises because reaction pathway corresponding to Citraconic acid is self-sustaining in terms of reagents which are regenerated in the pathway (mainly ethanol), whereas the Succinic acid route requires a large amount of methanol and hydrogen from the coarser scales to satisfy stoichiometry. Additionally, the energy requirements (found from enthalpies of reaction) of the two routes are different, affecting the life-cycle cost incurred due to fuel production.

Life-cycle decisions for minimal GWP and LCC are also strikingly different. For instance, fuel used for energy generation is Natural Gas for minimizing GWP, and Coal for minimizing cost. This confirms the notion that Coal is a cheaper but far more polluting source of fuel than Natural Gas. Itaconic acid is sourced from corn to minimize LCC, and banana to minimize GWP. This is also intuitive because industrial farming of corn is known to be more polluting but cost-effective when compared to cultivation of fruits and berries. The economy scale cash flows are also different for the two objectives. Largest cash flows for the minimum GWP pathway in Fig. 8 (a) arise from the utilities sector, probably because procuring natural gas has a tendency of affecting various electricity cogen plants in the utilities sector. In contrast, the largest cash flows

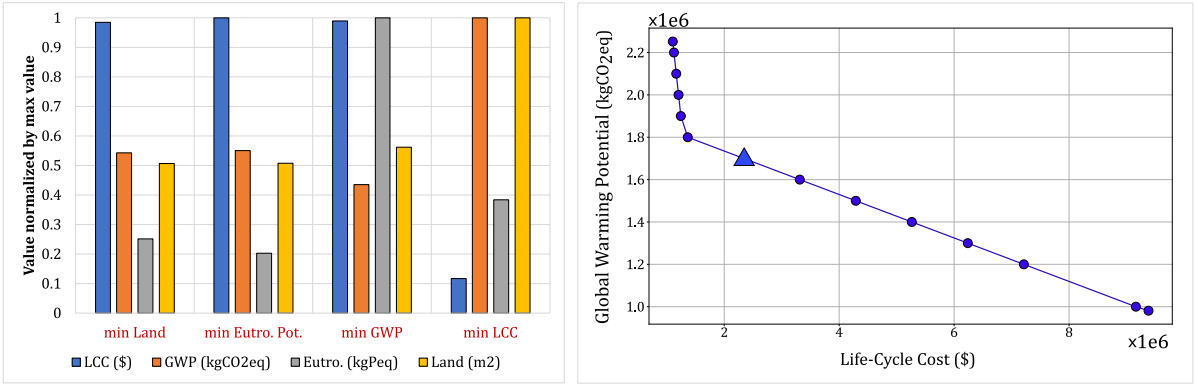

Fig. 7. Case Study Results (a) Normalized objective values (y-axis) for optimal pathways (x-axis) (b) Pareto curve to plot trade-off between GWP and LCC. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

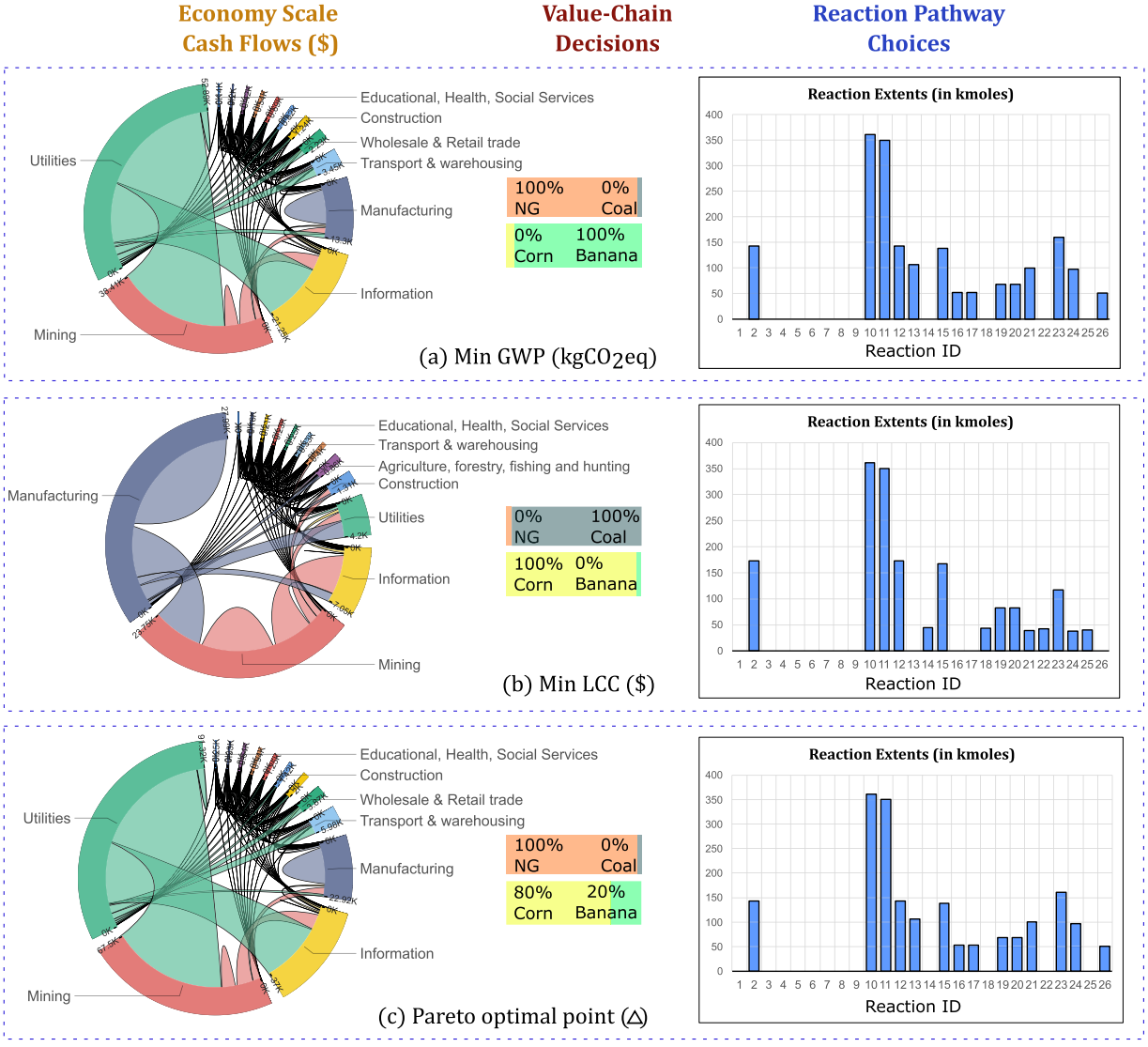

Fig. 8. Economy scale flows (in $), Value-chain decisions and Reaction extents for (a) Min GWP, (b) Min LCC and (c) Pareto Optimal (point marked as Δ in Fig. 7(b)).

Table 3
Values of objective function obtained for various optimal solutions, for basis of 1 ton per day 3-MTHF. Cells with bold entries are optimal objective function values.

Objective	LCC ($)	GWP ($kgCO_2eq$)	EP (kgPeq)	Land Area (m^2)
min Land Area	9,342,244	1,222,620	60.9	**244,137**
min Eutro. Potential (EP)	9,479,247	1,239,534	**49.2**	244,745
min Global Warming Potential (GWP)	9,382,517	**981,170**	242.4	270,820
min Life-cycle cost (LCC)	**1,104,876**	2,252,747	92.9	481,727

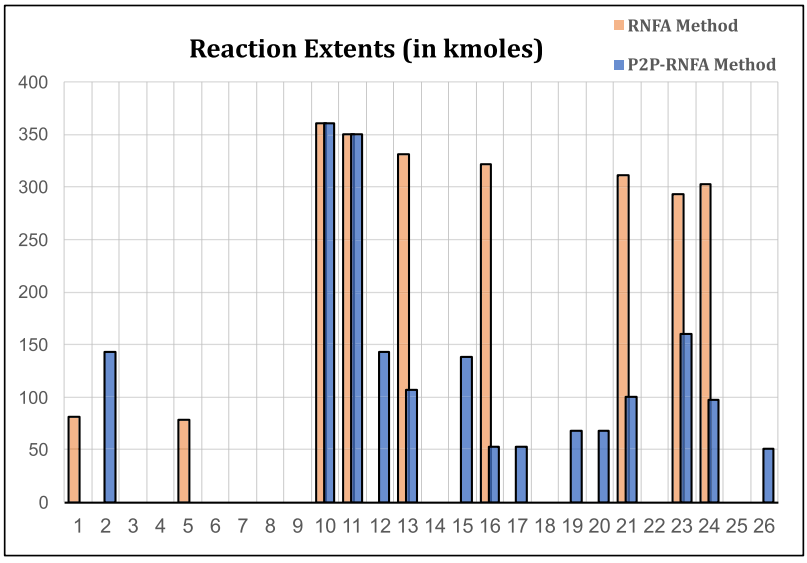

Fig. 9. Comparing optimal reaction pathway for minimal CO_2 objective chosen using simple RNFA method (Voll and Marquardt, 2012) and the P2P-RNFA extension developed in this paper.

corresponding to the minimum LCC pathway in Fig. 8 (b) seem to be arising from chemical manufacturing, and mining sectors of materials such as coal and ores. This is due to the fact that corn cultivation is closely related to various industries such as fertilizers, pesticides, etc.; and coal mining is estimated to require a lot of monetary transactions in the economy.

A pareto front is developed to quantify trade-offs between the two objectives in Fig. 8 (b). We subjectively select a compromise solution (Δ) near the knee point of the pareto front and display the solution in Fig. 8 (c). For this point, we notice that the reaction pathways chosen are identical to the minimal GWP, the cash flows in the economic sectors are all larger than minimal GWP but maintain similar proportions. The value-chain decisions for the pareto optimal are a combination of the decisions in minimal GWP and LCC with 80% IA being sourced from corn, and 20% from banana. It is evident from the contrasting decisions at all the three scales that synthesis of multi-scale pathways is non-trivial and is riddled with many trade-offs. The P2P-RNFA framework is capable of capturing these trade-offs and interactions between various scales, thereby designing chemical reaction pathways from a holistic perspective. However, it is worth mentioning that uncertainty present in life-cycle inventory datasets and EEIO models can affect the pareto optimal solutions. While propagation of uncertainty using analytical methods can become complicated due to disaggregation Eqs. (11)–(14), one can use sensitivity and hotspot analyses to determine relative contributions of uncertainties from different sectors on the objective functions. While we focus only on the methodological aspect of the P2P-RNFA framework in this paper, future work will include practical aspects such as including uncertainty analysis to find pareto optimal designs with appropriate sensitivities and error bars.

Applying the P2P-RNFA framework to the case study, we have demonstrated that it facilitates design of not only reaction pathways but also life-cycle alternatives and economic transactions. Therefore, a user can make decisions with higher confidence of not shifting environmental impact to other parts of the life-cycle and economy. In order to probe the effect of expanding system boundary in reaction pathway design, we compare results from both P2P-RNFA and the RNFA (Voll and Marquardt, 2012) frameworks by applying them to the same reaction network in the case study (Fig. 6). The objective chosen for comparison is the global warming potential. For the RNFA method, GWP is formulated as emissions per kJ generated in the boiler times the net enthalpy of the pathways. We find that the optimal reaction pathways found using the RNFA method are different from P2P-RNFA chosen, as shown in Fig. 9. It is observed that the routes involving $R2$ and [$R12$, $R15$, $R19$] are not chosen in the RNFA method. This proves that a different optimal solution is obtained by including life-cycle and economic implications of reaction pathways. As an added benefit of using P2P-RNFA framework, the cost objective can be formulated easily from the upstream cut-off flows from the economy scale, whereas Voll and Marquardt (2012) have to rely on empirical estimates of total annualized cost to rank pathways based on the cost objective.

5. Conclusions

The novel P2P-RNFA framework described in this paper is shown to be an effective tool to optimize reaction pathways, value-chains and economy cash-flows from a superstructure of alternative solutions in an integrated manner. It is a general framework and can be implemented for any chemical synthesis problem. In addition to an optimal chemical reaction route, it can also be used to design value-chain pathways and raw material life-cycles along with reaction pathways. It may be used for other topological design problems such as sustainable supply chain design. Thus,

the framework expands the system boundary of conventional reaction network design approaches thereby facilitating systematic and holistic viewpoints to find sustainable solutions. The framework also contains an approximation algorithm for case studies with large economy and life-cycle scales. This algorithm follows a hierarchical decomposition approach to find feasible pathways and corresponding disaggregated multi-scale matrices, thereby avoiding the non-linearity associated with conditional disaggregation. Ultimately, the framework can also be used to develop pareto fronts and obtain trade-offs using multi-objective optimization approaches. In conclusion, P2P-RNFA framework introduces a novel methodology which can guide reductionist research on chemical synthesis towards solutions that are sustainable from an holistic perspective, and reduce the chances of shifting impacts across value-chains.

In this paper, we have applied the proposed P2P-RNFA framework to a real-life case study (Fig. 6) for producing 1 ton per day of 3-methyl-tetrahydrofuran, a biofuel. There are over 25 intermediates and numerous pathways to convert biomass-derived itaconic acid (IA) to 3-MTHF. Additionally the value-chains of raw materials and energy sources contain many options, and are linked to the economy scale for inputs. The P2P-RNFA framework is applied to this case-study for the economic objective of life-cycle cost, and numerous environmental objectives, derived from midpoint indicators of life-cycle assessment. The synthesis route via 2-methylbutane-1,2-diol is chosen as the main pathway for all objectives, but the side reactions vary for different objectives. Ultimately trade-offs between the two types of objectives are quantified using pareto fronts, and a compromise solution is proposed (Fig. 8). In this compromise solution, energy is sourced from natural gas combustion, and IA is derived as a by-product from both corn and banana cultivation in the ratio 4:1. It is also noticed that limiting the system boundary only to the reaction network (RNFA method Voll and Marquardt, 2012) yields a different optimal solution results than the expanded system boundary (P2P-RNFA method), indicating the need for holistic design of pathways for sustainability using proposed framework.

In future work, the proposed P2P-RNFA framework may be extended to include market dynamics in the economy scale by substituting the input-output model with general or partial equilibrium models. This will enable the design of reaction networks while considering emerging technology adoption constraints and effects of price change and elasticities. It is also critical to evaluate the uncertainty in the datasets and its impact on pareto optimal designs. Systematic methods to apply tools such as hotspot and sensitivity analyses to the P2P-RNFA framework will also be explored in future work. The value-chain scale may also be extended to include consequential life-cycle models to consider effects of marginal changes in supplies and demands of reagents and raw materials. Future work will also focus on including the kinetics of chemical reactions, separation modules and automated flowsheet generation using reaction scale information.

Notation

Annotation

A bar over a matrix indicates economy scale, e.g. $\overline{A}$
A bar under a matrix indicates value-chain scale, e.g. $\underline{A}$
No bars indicates reaction scale, e.g. A^r
Bars over and below indicate multi-scale model, e.g. $\underline{\overline{A}}$
A hat $\hat{}$ over a vector denotes diagonalization, e.g. $\hat{\underline{s}}$
An asterisk superscript * indicates a disaggregated matrix, e.g. $\overline{A}^*$

Indexes

e_1 = Producing sector of the economy (in the Input-Output Model)
e_2 = Consuming sector of the economy (in the Input-Output Model)
v_1 = Value chain product
v_2 = Value Chain process/activity
r_1 = Reaction scale substance/molecule or reagent
r_2 = Reaction id
k = Pathway id
l = Intervention flow id
i = Superset of rows
j = Superset of columns

Acronyms

RNFA : Reaction network flux analysis model
LCA : Life-cycle assessment model (either process-LCA or hybrid-LCA)
EEIO : Environmentally extended input-output model
P2P : Process-to-planet framework
GWP : Global Warming Potential
LCC : Life-cycle cost
EP : Eutrophication potential

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Vyom Thakker: Conceptualization, Formal analysis, Investigation, Methodology, Software, Writing – original draft. **Bhavik R. Bakshi:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – review & editing.

Acknowledgments

Partial funding for this work was provided by the Global Kaiteki Center at Arizona State University and the National Science Foundation through grant EFMA-2029397.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.compchemeng.2021.107578.

References

Anastas, P., Eghbali, N., 2010. Green chemistry: principles and practice. Chem. Soc. Rev. 39 (1), 301–312.
Azapagic, A., 1999. Life cycle assessment and its application to process selection, design and optimisation. Chem. Eng. J. 73 (1), 1–21.
Bakshi, B.R., 2019. Sustainable Engineering: Principles and Practice. Cambridge University Press doi:10.1017/9781108333726.
Bakshi, B.R., Fiksel, J., 2003. The quest for sustainability: challenges for process systems engineering. AlChE J. 49 (6), 1350–1358.
Bussieck, M.R., Meeraus, A., 2004. General algebraic modeling system (GAMS). In: Modeling Languages in Mathematical Optimization. Springer, pp. 137–157.
Degrand, E., Hemery, M., Fages, F., 2019. On chemical reaction network design by a nested evolution algorithm. In: Proceedings of the International Conference on Computational Methods in Systems Biology. Springer, pp. 78–95.
Frumkin, J.A., Doherty, M.F., 2018. Target bounds on reaction selectivity via Feinberg's CFSTR equivalence principle. AlChE J. 64 (3), 926–939.

Frumkin, J.A., Doherty, M.F., 2020. A rapid screening methodology for chemical processes. Comput. Chem. Eng. 142, 107039.
Gao, J., You, F., 2018. Integrated hybrid life cycle assessment and optimization of shale gas. ACS Sustain. Chem. Eng. 6 (2), 1803–1824.
Garcia, D.J., You, F., 2015. Multiobjective optimization of product and process networks: general modeling framework, efficient global optimization algorithm, and case studies on bioconversion. AlChE J. 61 (2), 530–554.
Ghosh, T., Bakshi, B.R., 2020. Designing hybrid life cycle assessment models based on uncertainty and complexity. Int. J. Life Cycle Assess. 25 (11), 2290–2308.
Ghosh, T., Lee, K., Bakshi, B.R., 2019. Integrating market models and price effects in a multiscale sustainable process design framework. Comput. Aided Chem. Eng. 47, 175–180.
Guillén-Gosálbez, G., You, F., Galán-Martín, Á., Pozo, C., Grossmann, I.E., 2019. Process systems engineering thinking and tools applied to sustainability problems: current landscape and future opportunities. Curr. Opin. Chem. Eng. 26, 170–179.
Hanes, R.J., Bakshi, B.R., 2015a. Process to planet: a multiscale modeling framework toward sustainable engineering. AlChE J. 61 (10), 3332–3352.
Hanes, R.J., Bakshi, B.R., 2015b. Sustainable process design by the process to planet framework. AlChE J. 61 (10), 3320–3331.
Heijungs, R., Suh, S., 2002. The Computational Structure of Life Cycle Assessment, 11. Springer Science & Business Media.
Islam, S., Ponnambalam, S., Lam, H.L., 2016. Review on life cycle inventory: methods, examples and applications. J. Clean. Prod. 136, 266–278.
Kokossis, A.C., Floudas, C.A., 1991. Synthesis of isothermal reactor-separator-recycle systems. Chem. Eng. Sci. 46 (5–6), 1361–1383.
Konig, A., Ulonska, K., Mitsos, A., Viell, J., 2019. Optimal applications and combinations of renewable fuel production from biomass and electricity. Energy Fuels 33 (2), 1659–1672.
Konijn, P.J., Steenge, A.E., 1995. Compilation of input-output data from the national accounts. Econ. Syst. Res. 7 (1), 31–46.
Lee, D.-Y., Yun, H., Park, S., Lee, S.Y., 2003. Metafluxnet: the management of metabolic reaction information and quantitative metabolic flux analysis. Bioinformatics 19 (16), 2144–2146.
Lee, K., Ghosh, T., Bakshi, B.R., 2019. Toward multiscale consequential sustainable process design: including the effects of economy and resource constraints with application to green urea production in a watershed. Chem. Eng. Sci. 207, 725–743.
Lee, S., Phalakornkule, C., Domach, M.M., Grossmann, I.E., 2000. Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. Comput. Chem. Eng. 24 (2–7), 711–716.
Leontief, W., 1970. Environmental repercussions and the economic structure: an input-output approach. Rev. Econ. Stat. 262–271.
da Luz, L.M., de Francisco, A.C., Piekarski, C.M., Salvador, R., 2018. Integrating life cycle assessment in the product development process: a methodological approach. J. Clean. Prod. 193, 28–42.
Mavrotas, G., 2009. Effective implementation of the ε-constraint method in multi-objective mathematical programming problems. Appl. Math. Comput. 213 (2), 455–465.
Motianlifu, M., 2017. Expansion of Reaction Network Flux Analysis toward including Life Cycle Assessment and Ecosystem Services. The Ohio State University Master's thesis.
Pomponi, F., Lenzen, M., 2018. Hybrid life cycle assessment (LCA) will likely yield more accurate results than process-based LCA. J. Clean. Prod. 176, 210–215.
Ramapriya, G.M., Won, W., Maravelias, C.T., 2018. A superstructure optimization approach for process synthesis under complex reaction networks. Chem. Eng. Res. Des. 137, 589–608.
del Rio-Chanona, E.A., Zhang, D., Shah, N., 2018. Sustainable biopolymer synthesis via superstructure and multiobjective optimization. AlChE J. 64 (1), 91–103.
Sahinidis, N.V., 1996. Baron: a general purpose global optimization software package. J. Glob. Optim. 8 (2), 201–205.
Schuster, S., Fell, D.A., Dandekar, T., 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nat. Biotechnol. 18 (3), 326–332.
Schweiger, C., Floudas, C., 1999. Synthesis of optimal chemical reactor networks. Comput. Chem. Eng. 23, S47–S50.
Shinar, G., Feinberg, M., 2011. Design principles for robust biochemical reaction networks: what works, what cannot work, and what might almost work. Math. Biosci. 231 (1), 39–48.
Suh, S., 2004. Functions, commodities and environmental impacts in an ecological–economic model. Ecol. Econ. 48 (4), 451–467.
Suh, S., Huppes, G., 2005. Methods for life cycle inventory of a product. J. Clean. Prod. 13 (7), 687–697.
Suh, S., Weidema, B., Schmidt, J.H., Heijungs, R., 2010. Generalized make and use framework for allocation in life cycle assessment. J. Ind. Ecol. 14 (2), 335–353.
Thakker, Vyom, Bakshi, Bhavik R., 2021. Toward sustainable circular economies: A computational framework for assessment and design. Journal of Cleaner Production 295, 126353. doi:10.1016/j.jclepro.2021.126353.
Ulonska, K., Skiborowski, M., Mitsos, A., Viell, J., 2016. Early-stage evaluation of biorefinery processing pathways using process network flux analysis. AlChE J. 62 (9), 3096–3108.
Unsleber, J.P., Reiher, M., 2020. The exploration of chemical reaction networks. Annu. Rev. Phys. Chem. 71, 121–142.
U.S. Bureau of Economic Analysis, 2018, 2018. Interactive access to industry economic accounts data: Input-output. Technical Report. Available at: https://apps.bea.gov/iTable/itable.cfm?reqid=58&step=1, Accessed on: 01/28/2021.
Varma, A., Palsson, B.O., 1994. Metabolic flux balancing: basic concepts, scientific and practical use. Bio/technology 12 (10), 994–998.
Voll, A., Marquardt, W., 2012. Reaction network flux analysis: optimization-based evaluation of reaction pathways for biorenewables processing. AlChE J. 58 (6), 1788–1801.
Yang, Y., Heijungs, R., Brandão, M., 2017. Hybrid life cycle assessment (LCA) does not necessarily yield more accurate results than process-based LCA. J. Clean. Prod. 150, 237–242.
Yue, D., Pandya, S., You, F., 2016. Integrating hybrid life cycle assessment with multiobjective optimization: a modeling framework. Environ. Sci. Technol. 50 (3), 1501–1509.
Zhuang, K.H., Herrgård, M.J., 2015. Multi-scale exploration of the technical, economic, and environmental dimensions of bio-based chemical production. Metab. Eng. 31, 1–12.