Intelligent Monte Carlo: A New Paradigm for Inverse Polymerization Engineering

Yousef Mohammadi, Mohammad Reza Saeb, Alexander Penlidis, Esmaiel Jabbari, Philippe Zinck, Florian J. Stadler, Krzysztof Matyjaszewski

"This is an Accepted Manuscript of an article published by Macromol. Theory and Simul., mats.201700106, accepted in January 2018; vol 27, no 3, 1700106 (2018)."

Intelligent Monte Carlo: A new paradigm for inverse polymerization engineering

Yousef Mohammadi[1]*, Mohammad Reza Saeb[2], Alexander Penlidis[3]*, Esmaiel Jabbari[4], Philippe Zinck[5]*, Florian J. Stadler[6], Krzysztof Matyjaszewski[7]*

1. Petrochemical Research and Technology Company (NPC-rt), National Petrochemical Company (NPC), P.O. Box 14358-84711, Tehran, Iran
2. Department of Resin and Additives, Institute for Color Science and Technology, P.O. Box 16765-654, Tehran, Iran
3. Department of Chemical Engineering, Institute for Polymer Research (IPR), University of Waterloo, Waterloo, Ontario N2L 3G1 Canada
4. Biomimetic Materials and Tissue Engineering Laboratory, Department of Chemical Engineering, University of South Carolina, Columbia, SC, USA
5. Unity of Catalysis and Solid State Chemistry, University of Lille, CNRS, Bât C7, Cité Scientifique, 59652 Villeneuve d'Ascq Cédex, France
6. College of Materials Science and Engineering, Shenzhen Key Laboratory of Polymer Science and Technology, Guangdong Research Center for Interfacial Engineering of Functional Materials, Nanshan District Key Lab for Biopolymers and Safety Evaluation, Shenzhen University, Shenzhen 518060, China
7. Department of Chemistry, Carnegie Mellon University, Pittsburgh, USA

To whom correspondence should be addressed:

Dr. Yousef Mohammadi: mohammadi@npc-rt.ir

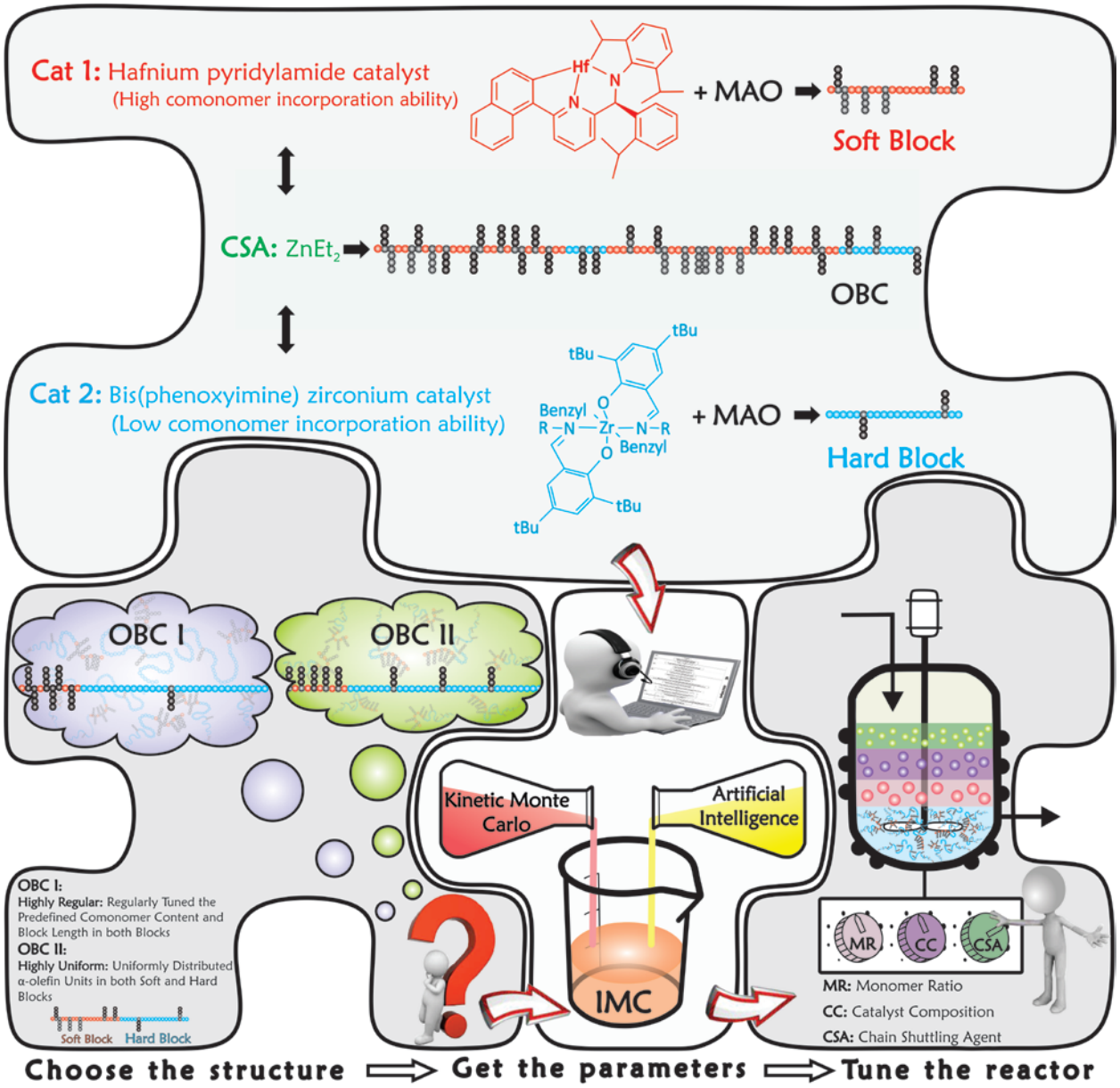
Prof. Alexander Penlidis: penlidis@uwaterloo.ca

Prof. Philippe Zinck: philippe.zinck@univ-lille1.fr

Prof. Krzysztof Matyjaszewski: matyjaszewski@cmu.edu

GRAPHICAL ABSTRACT

Traditional computational methods simulate the microstructure of polymer chains from input polymerization conditions. In this contribution, we introduce the Intelligent Monte Carlo (IMC) approach able to predict optimal recipe/operating conditions for synthesizing complex copolymer molecules with predefined microstructures as input. Chain shuttling copolymerization is chosen as the first test case.



ABSTRACT

Traditional computational methods simulate the microstructure of polymer chains from input reaction conditions, but a need exists for predicting optimum reaction conditions in a computationally-demanding multi-variable space leading to the synthesis of predesigned microstructures and architectures. We introduce herein the Intelligent Monte Carlo (IMC) approach, able to predict optimum reaction conditions for synthesizing copolymers with predefined, complex microstructures as input. This is rendered possible by a combination of Kinetic Monte Carlo (KMC) simulation with Artificial Intelligence concepts, which enables a reasonably enhanced convergence to optimum reactions conditions. Chain shuttling polymerization was chosen as a first test case due to its complexity and the intricate multi-block microstructures that are formed; whose tailoring requires multiple parameters. The IMC approach located optimum reaction conditions for the synthesis of olefinic multi-block copolymers with specific microstructures. This approach provides a new platform for identifying complex reaction conditions to 'produce' and 'tailor-make' materials with precisely predefined microstructures and facilitates the development of meaningful structure-property relationships.

Keywords: Monte Carlo simulation; Chain shuttling polymerization; Artificial Intelligence; Inverse polymerization engineering

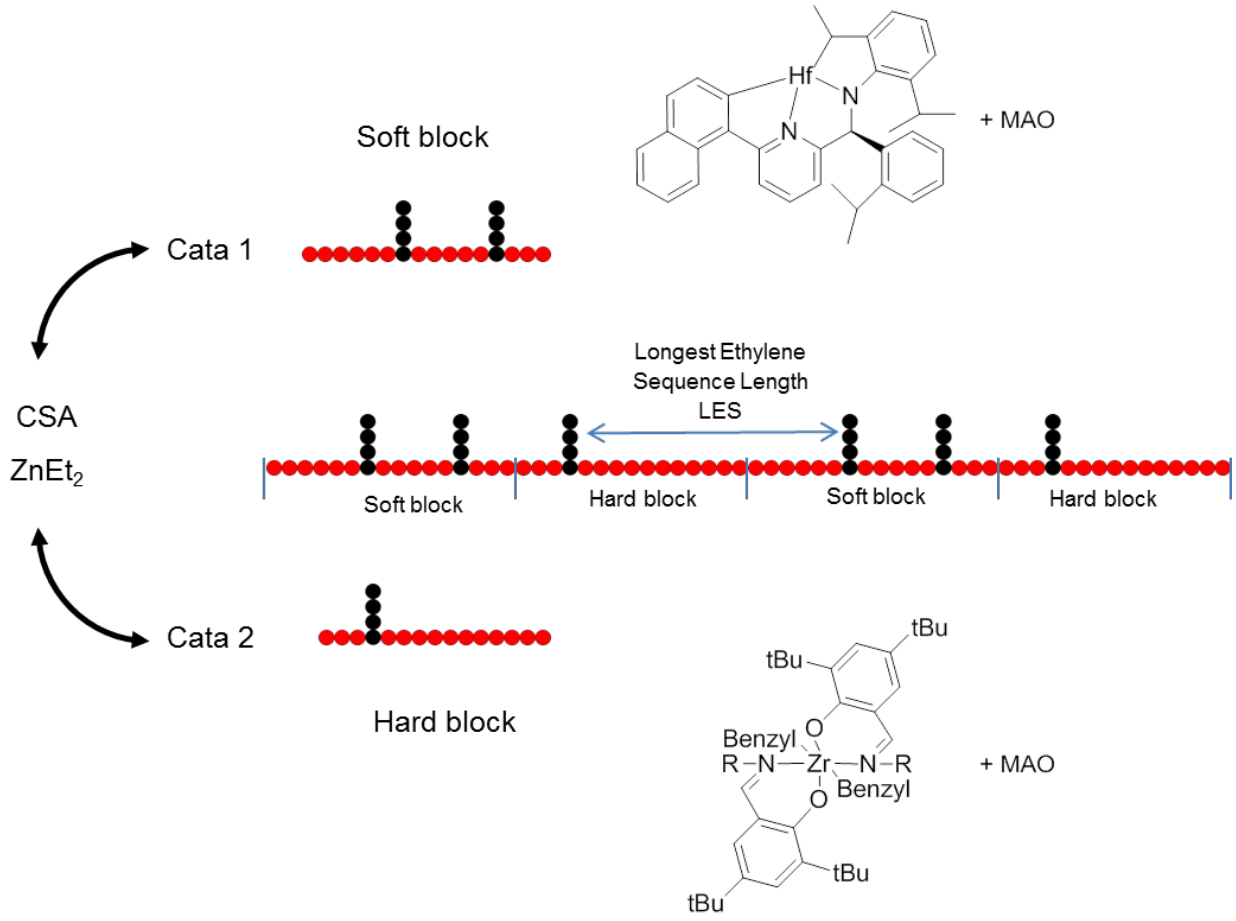
1. INTRODUCTION

Over the last decades, computational chemistry has provided a suite of promising tools for all aspects of process and product design and optimization, namely, for studying reaction mechanisms, designing new compounds and materials, and improving chemical processes [1-4]. In many complex systems including industrial processes, numerical simulations are required to predict the outcome of a certain operation performed with a defined set of input variables. Due to the enormous complexity of the nonlinear problem of linking input factors to output properties, it is not possible to predict readily the former from the latter. However, in industrial processes, if one were able to 'dial in' a certain set of resulting characteristics (desirable product properties) and subsequently, via using some modelling/simulation software tool, obtain feasible suggestions for a set of input factors that would give rise to such properties, would be economically highly important and beneficial. In the chemical industry, for example, a significant number of reactions are frequently occurring simultaneously and influence each other.

In the synthetic polymer chemistry, modern monitoring/measuring technology has given rise to increasingly more complex and elaborated reaction mechanisms and product microstructures and architectures, whose description and simulation is not trivial [5-10]. For instance, the sequence (length) distribution of comonomer units along a macromolecular chain is a key property characteristic that governs several macroscopic properties of a polymer. Kinetic Monte Carlo (KMC) simulations have been successfully used to gain detailed micromolecular level information about chain microstructure, as described, for example, in references [11] to [28]. In traditional simulation approaches, the strategy consists of obtaining microstructural signatures of the copolymer chains mainly in response to input reaction conditions. In many practical applications, however, the desired properties of the polymers can be specified and one should ideally be able to move backwards and define the reaction conditions required to obtain the corresponding microstructures. The reverse simulation going from the macromolecular microstructure (as problem input) to the reaction conditions (as problem output) requires searching a large multi-variable space for the optimum conditions for the synthesis of predesigned copolymer microstructures. This (tedious task for deterministic models/simulations)

can be done by providing typical KMC approaches with the capability to make decisions. KMC should thus be equipped with an appropriate optimization tool to intelligently seek, identify, and screen for molecular patterns in a (potentially huge) copolymerization search space. Artificial intelligence techniques are very versatile and effective stochastic modelling and optimization tools and have recently been revived in many scientific fields including materials science [29-30]. The proposed Intelligent Monte Carlo (IMC) simulation approach is a combination of artificial intelligence with KMC to yield (as problem output) the appropriate reaction conditions in response to predesigned polymeric microstructures (as problem input).

In the current study, chain shuttling polymerization has been chosen as a challenging and intricate polymerization system to test the proposed IMC simulation approach [7, 31-33]. In the process represented in Scheme 1, the growing macromolecular chain is able to shuttle between two catalysts showing a significantly different reactivity vs. ethylene and 1-octene, the two comonomers [5]. The trans-metallation reaction occurs via a chain shuttling agent (CSA). The resulting products, olefinic block copolymers (OBCs), are recently developed representatives of macromolecules with a complex, fine-tuned microstructure composed of blocks of statistical copolymers with two different compositions, one being crystallizable and the other amorphous [34].



Scheme 1. Ethylene/1-octene chain shuttling polymerization leading to olefinic block copolymers (OBCs) [5]. CSA is the chain shuttling agent and MAO methylaluminoxane. R=2-methylcyclohexyl (cata stands for catalyst).

We previously reported on microstructural changes in the semi-batch chain shuttling polymerization of ethylene and 1-octene using KMC simulation [35-38]. The KMC simulation was able to quantify many signature distributional characteristics of ethylene/1-octene copolymerization, such as the block composition, referred to as $C8\%$, the block length represented by the number-average degree of polymerization $\overline{DP}_n$, the relative proportion of the blocks described via the hard block content ($HB\%$), the number of linking points between blocks (LP), ethylene sequence length (ESL), and also the longest ethylene sequence (LES). Although the developed KMC code was capable of precisely simulating the sequence distribution and microstructure of the olefin block copolymer chains (y dependent variables) in a virtual reactor in response to input reaction conditions (x independent variables), in practical

applications, the desired properties of the OBCs can be specified in advance, whereas the reaction conditions such as monomer molar fraction, catalyst package composition, temperature, and CSA concentration (that would yield the OBC desirable properties) are not readily evident and they thus have to be determined. As mentioned above, the reverse simulation going from the output macromolecular properties to the input reaction conditions requires random searches of a large multi-variable space for the optimum conditions for the synthesis of OBCs with predesigned microstructures.

The developed IMC approach is able to utilize synergistically the power of KMC and artificial intelligence to map out the output-to-input variable space of chain shuttling reactions in order to find the optimum reaction conditions irrespective of the number of variables in the input and output search spaces. The IMC approach is faster and more efficient than other available computational tools for predicting and optimizing polymerization reaction conditions for a given copolymer architecture and microstructure, and much quicker (while almost equally reliable) than an empirical approach using pilot-scale reactor experimentation that is usually time-consuming and costly.

2. MODEL DEVELOPMENT

KMC simulation is a unique mathematical tool for virtually synthesizing and visualizing complex polymerization reactions. It has been successfully applied to a variety of macromolecular reactions. Despite its remarkable potential, KMC simulation is limited by its inability to 'match' input variables, i.e. copolymerization recipe and operating condition variables, to outputs (e.g. copolymerization chain topological features). To find the optimum reaction recipe/conditions for the synthesis of predesigned chain microstructures, KMC should be equipped with the appropriate optimization tools to intelligently seek, identify, and screen for molecular patterns in the copolymerization search space. As mentioned earlier, artificial intelligence techniques are very versatile and effective stochastic modeling and optimization tools currently employed successfully in many scientific fields [39-41]. Among different computationally intelligent techniques, Artificial Neural Networks (ANNs, biologically-inspired modeling tools) and Genetic Algorithms (GAs, evolutionary optimization methods) have gained much attention in recent years [42-52]. Intelligent Monte Carlo (IMC) simulation is the application of artificial intelligence within

the KMC framework in order to map the desired chain microstructures to input reaction recipe/conditions. In the following, we describe the IMC approach and demonstrate its use in finding the polymerization conditions warranting the synthesis of predesigned, tailor-made macromolecules.

A two-step modeling and simulation strategy is required for implementing the IMC approach. First, the copolymerization reaction is simulated to virtually synthesize macromolecular chains with complete microstructural details. The first step is accomplished by KMC simulation. In the second step, a multi-objective genetic algorithm is employed to search the copolymerization variable space in an intelligent evolutionary manner for the optimum reaction recipe/operating condition variables. The genetic algorithm generates random copolymerization recipes and essentially recalls the KMC simulator to synthesize and visualize the macromolecular chains for each recipe. By applying this reverse strategy, each copolymerization recipe behaves like a genotype and is precisely related to a set of virtually synthesized macromolecules, which act as phenotypes. The predictions of the KMC simulator are subsequently modeled with artificial neural networks (ANNs) to reduce the frequency of invoking the KMC simulator with long execution times. Hence, the stochastic modeling of the KMC predictions by trained ANNs dramatically reduces the computational time of IMC (examples of typical computational times will be cited later).

2.1. First Step: KMC simulation and its corresponding ANN-based modeling

We previously developed a KMC methodology for simulating copolymerization kinetics and microstructure of the virtually synthesized macromolecules within a relatively large simulation volume [15, 18-19, 21, 25, 28, 35-38]. In this work, complex Olefin Block Copolymers (OBCs) produced by chain shuttling coordination copolymerization reactions have been selected to demonstrate the capabilities of the enhanced IMC approach. OBC copolymerization reactions possess diverse compositional, segmental (hard/soft), molecular, and microstructural properties. In principle, the existence of two catalysts with opposite actions in the chain shuttling copolymerization mechanism involving monomer/comonomer molecules in the presence of a chain shuttling agent, necessitates tracking a massive number of growing macromolecules of different types (identities). We previously tested the KMC simulator by analyzing the

polymerization kinetics and chain microstructure of OBCs including molecular weight distribution (MWD), copolymer composition distribution (CCD), ethylene sequence length (ESL), and longest ethylene sequence (LES) distribution patterns along with their constituent soft and hard blocks in terms of reaction factors, including CSA level, catalyst composition, and monomer molar fraction [35-38]. The simulation algorithm stored the instantaneous characteristics of the dynamic last block along with cumulative information of the static blocks on all virtual chains. The KMC simulator allowed the simulation of a statistically large sample size with a computationally cost-effective execution time. The OBC chains in the reaction volume comprised of soft and hard blocks with fixed length and microstructure, known as static blocks that could not be reorganized. On the other hand, the active terminal blocks of the living or dormant chains were allowed to propagate until experiencing a cross-shuttling or termination to yield a dead chain.

The importance of different reaction variables in tailoring the OBC chains has been discussed in a previous publication [37]. The simulation data on chain characteristics of OBCs can be treated in two ways: (*i*) final properties, which consist of distribution patterns for blocks, chains, and sequences; and (*ii*) time-dependent properties, which are representative of the evolutionary variations in chain microstructure. In the first step, the KMC simulator was used to virtually synthesize tailored OBCs by chain-shuttling copolymerization through an intelligent multi-objective optimization process. In the second step, the KMC simulator was recalled by the genetic algorithm to synthesize new OBCs. Since the KMC simulation is computationally time-consuming and frequently recalled and executed, trained ANNs were used to optimize the microstructure of the OBCs for each intelligently generated copolymerization recipe.

To develop robust ANNs capable of predicting the microstructure of OBCs required a reliable initial dataset on OBCs under different copolymerization conditions to train and test the predesigned ANNs. In that regard, different copolymerization recipes were defined and fed into the KMC simulator based on the independent reaction variables (CSA level, catalyst composition, molar ratio of the monomers) [37]. The CSA level (defined as the logarithm of CSA concentration divided by the initial CSA concentration of 0.27 g/L) ranged from -3 to +1 and determined the length and number of soft and hard blocks per OBC chain [35-38]. The molar fractions of the catalyst package and monomer feed, which ranged from 0.2 to 0.8, determined the weight

fractions of soft and hard blocks and the molar fractions of the monomers in the OBC chains, respectively. The input variables for ethylene/1-octene copolymerization, taken from our previous KMC work, are shown in Table 1.

Table 1. Variables applied in KMC simulation of chain shuttling polymerization (CSP) [36].

Parameter	Value	Unit
Initial concentration of ethylene	2.63	mol L^{-1}
Solvent initial mass	1000	g
Ethylene initial mass	200	g
1-octene initial molar fraction	0.2, 0.4, 0.6, and 0.8	-
Hydrogen initial mass	0.072	g
CSA initial mass	0.27×10^{x}; x ϵ {-3, -2, -1, 0, and 1}	g
Catalyst metal initial mass	1.50×10^{-4}	g
Catalyst 1 initial molar fraction	0.2, 0.4, 0.6, and 0.8	-

Accordingly, the characteristics of 80 virtually synthesized OBCs with diverse microstructural features are shown as Y variables in Table 2. According to Table 2, the properties of the virtually synthesized OBCs were very sensitive to the values of the independent variables and changed over a wide range.

Table 2. Microstructural characteristic of the virtual OBCs synthesized by KMC simulator. X_1, X_2, and X_3 are 1-octene mole fraction, catalyst composition, and log(CSA Level), respectively. Y_1 and Y_2 are average 1-octene content in soft and hard blocks ($\overline{C8\%}$), respectively, Y_3 is the hard block percent ($HB\%$), Y_4 and Y_5 represent average ethylene sequence length in soft and hard blocks ($\overline{ESL}$), Y_6 and Y_7 represent number average degree of polymerization of the soft and hard blocks ($\overline{DP}_n$), and Y_8 and Y_9 are the average numbers of 1-octene units per soft and hard block ($\overline{N}_{C8}$).

	Inputs			Outputs/Responses								
	X_1	X_2	X_3	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9
KMC RUN				$\overline{C8\%}^{Soft}$	$\overline{C8\%}^{Hard}$	$HB\%$	$\overline{ESL}^{Soft}$	$\overline{ESL}^{Hard}$	$\overline{DP_n}^{Soft}$	$\overline{DP_n}^{Hard}$	$\overline{N}_{C8}^{Soft}$	$\overline{N}_{C8}^{Hard}$
1	0.2	0.2	-3	47.608	4.106	58.255	2.339	27.722	650.874	38085.624	309.867	1564.478

2	0.2	0.2	-2	47.608	4.111	58.254	2.339	27.682	591.332	12751.269	281.528	524.735
3	0.2	0.2	-1	47.651	4.153	58.238	2.339	27.308	309.000	1669.303	147.255	69.406
4	0.2	0.2	0	48.002	4.579	58.083	2.339	24.030	54.613	173.840	26.219	7.969
5	0.2	0.2	1	50.721	8.445	56.779	2.346	11.255	7.118	18.735	3.611	1.584
6	0.2	0.4	-3	46.630	3.958	34.176	2.389	28.662	722.854	32876.143	337.071	1302.243
7	0.2	0.4	-2	46.631	3.966	34.180	2.389	28.580	667.499	7416.399	311.278	294.411
8	0.2	0.4	-1	46.653	4.054	34.226	2.389	27.765	376.282	849.138	175.588	34.459
9	0.2	0.4	0	46.871	4.901	34.651	2.392	21.708	71.322	87.369	33.443	4.286
10	0.2	0.4	1	48.532	12.006	37.816	2.426	7.344	9.076	10.040	4.407	1.206
11	0.2	0.6	-3	45.882	3.850	18.806	2.428	29.391	780.176	28914.024	357.963	1114.093
12	0.2	0.6	-2	45.883	3.863	18.813	2.429	29.258	735.608	5218.855	337.546	201.808
13	0.2	0.6	-1	45.893	3.991	18.881	2.429	28.010	468.605	570.842	215.150	22.805
14	0.2	0.6	0	46.012	5.251	19.556	2.433	19.699	102.278	58.866	47.098	3.094
15	0.2	0.6	1	46.960	15.113	24.834	2.472	5.587	12.857	7.172	6.044	1.085
16	0.2	0.8	-3	45.261	3.763	8.013	2.463	30.007	830.011	25255.961	375.675	951.517
17	0.2	0.8	-2	45.261	3.780	8.017	2.463	29.823	804.149	4024.560	364.007	152.286
18	0.2	0.8	-1	45.267	3.954	8.071	2.463	28.093	609.983	428.997	276.352	16.980
19	0.2	0.8	0	45.317	5.617	8.586	2.466	17.982	180.678	44.565	82.041	2.505
20	0.2	0.8	1	45.750	17.900	13.003	2.492	4.587	23.727	5.734	10.884	1.027
21	0.4	0.2	-3	23.056	1.409	58.376	4.701	74.672	830.495	36636.472	191.483	516.521
22	0.4	0.2	-2	23.059	1.411	58.371	4.701	74.424	754.144	12281.551	173.906	173.546
23	0.4	0.2	-1	23.081	1.443	58.321	4.701	71.567	394.779	1608.190	91.139	23.244
24	0.4	0.2	0	23.314	1.748	57.867	4.699	51.992	70.296	166.944	16.395	2.924
25	0.4	0.2	1	24.641	4.473	54.230	4.672	14.518	9.885	17.894	2.438	0.802
26	0.4	0.4	-3	21.790	1.320	33.908	4.967	79.492	948.639	31663.938	206.711	418.301
27	0.4	0.4	-2	21.790	1.327	33.910	4.967	78.783	875.351	7127.176	190.760	94.696
28	0.4	0.4	-1	21.800	1.387	33.920	4.968	72.681	494.301	816.101	107.817	11.336
29	0.4	0.4	0	21.881	1.983	34.004	4.984	41.119	94.599	83.967	20.720	1.667
30	0.4	0.4	1	22.001	6.852	34.387	5.149	8.428	13.101	9.535	2.888	0.654
31	0.4	0.6	-3	20.820	1.256	18.408	5.194	83.325	1049.436	27828.497	218.494	350.158
32	0.4	0.6	-2	20.819	1.265	18.410	5.194	82.269	989.781	5046.993	206.109	63.926
33	0.4	0.6	-1	20.821	1.354	1.354	5.196	72.642	630.248	549.010	131.373	7.447
34	0.4	0.6	0	20.826	2.224	18.856	5.218	33.848	138.577	56.535	28.923	1.259
35	0.4	0.6	1	20.516	8.928	21.749	5.452	6.123	18.752	6.779	3.861	0.606
36	0.4	0.8	-3	20.016	1.205	7.728	5.400	86.761	1141.502	24482.891	228.493	295.360

37	0.4	0.8	-2	20.016	1.217	7.732	5.400	85.150	1104.721	3879.224	221.200	47.286
38	0.4	0.8	-1	20.015	1.336	7.770	5.402	71.897	837.689	412.872	168.038	5.523
39	0.4	0.8	0	19.999	2.469	8.133	5.419	28.696	248.662	42.759	50.000	1.057
40	0.4	0.8	1	19.720	10.881	11.133	5.593	4.904	34.132	5.399	6.786	0.588
41	0.6	0.2	-3	10.596	0.577	52.513	9.877	176.673	1304.194	36429.899	138.193	210.484
42	0.6	0.2	-2	10.597	0.579	52.505	9.876	175.011	1182.713	12110.545	125.344	70.267
43	0.6	0.2	-1	10.604	0.598	52.441	9.869	159.871	619.956	1589.107	65.764	9.532
44	0.6	0.2	0	10.649	0.787	51.776	9.811	86.035	111.210	164.847	11.851	1.301
45	0.6	0.2	1	10.607	2.396	46.827	9.203	16.053	16.279	17.570	1.729	0.423
46	0.6	0.4	-3	9.409	0.515	28.068	11.101	197.450	1553.223	31462.242	146.140	162.264
47	0.6	0.4	-2	9.409	0.519	28.068	11.100	193.268	1432.063	7051.519	134.766	36.655
48	0.6	0.4	-1	9.404	0.555	28.048	11.102	159.573	809.251	806.544	76.193	4.486
49	0.6	0.4	0	9.353	0.904	27.859	11.125	58.754	156.071	82.849	14.627	0.751
50	0.6	0.4	1	8.580	3.586	26.311	11.172	8.847	22.767	9.315	1.961	0.335
51	0.6	0.6	-3	8.513	0.471	14.264	12.260	215.231	1781.103	27252.799	151.642	128.744
52	0.6	0.6	-2	8.513	0.476	14.266	12.260	207.813	1678.297	4963.964	142.934	23.714
53	0.6	0.6	-1	8.506	0.527	14.278	12.267	155.077	1068.159	542.253	91.075	2.867
54	0.6	0.6	0	8.432	1.018	14.393	12.329	44.478	236.343	55.774	20.022	0.569
55	0.6	0.6	1	7.593	4.598	15.104	12.796	6.314	33.574	6.592	2.569	0.304
56	0.6	0.8	-3	7.790	0.437	5.668	13.400	2.194	1993.206	24158.214	155.280	105.936
57	0.6	0.8	-2	7.788	0.444	5.669	13.402	220.333	1931.377	3849.548	150.539	17.127
58	0.6	0.8	-1	7.783	0.509	5.686	13.408	148.506	1463.606	408.405	114.477	2.083
59	0.6	0.8	0	7.732	1.129	5.850	13.463	35.711	434.938	42.163	34.035	0.477
60	0.6	0.8	1	7.191	5.617	7.158	13.953	5.015	61.637	5.228	4.511	0.295
61	0.8	0.2	-3	3.647	0.190	41.899	27.818	524.333	2332.332	36025.802	85.067	68.524
62	0.8	0.2	-2	3.647	0.191	41.888	27.812	508.471	2121.120	12058.470	77.359	23.140
63	0.8	0.2	-1	3.645	0.199	41.824	27.701	397.730	1109.221	1578.188	40.462	3.161
64	0.8	0.2	0	3.626	0.281	41.219	26.652	125.682	198.287	163.772	7.200	0.462
65	0.8	0.2	1	3.408	0.942	37.483	19.003	16.867	27.151	17.388	0.927	0.165
66	0.8	0.4	-3	2.894	0.155	19.711	35.011	637.592	2865.481	31349.508	82.932	48.793
67	0.8	0.4	-2	2.894	0.157	19.709	35.000	595.556	2640.330	7014.604	76.433	11.028
68	0.8	0.4	-1	2.889	0.171	19.691	34.895	359.492	1492.626	802.289	43.216	1.375
69	0.8	0.4	0	2.847	0.307	19.522	33.837	73.250	286.182	82.331	8.176	0.254
70	0.8	0.4	1	2.494	1.301	18.716	25.047	9.038	38.236	9.187	0.958	0.120
71	0.8	0.6	-3	2.407	0.133	9.306	42.095	738.563	3325.084	27202.171	80.038	36.322

72	0.8	0.6	-2	2.406	0.135	9.306	42.088	658.375	3136.358	4967.879	75.528	6.726
73	0.8	0.6	-1	2.403	0.153	9.310	41.986	315.360	1996.809	539.812	48.179	0.832
74	0.8	0.6	0	2.368	0.332	9.351	40.962	51.667	438.618	55.401	10.470	0.185
75	0.8	0.6	1	2.085	1.590	9.842	32.029	6.396	57.326	6.482	1.207	0.104
76	0.8	0.8	-3	2.070	0.118	3.543	48.969	825.428	3736.867	24003.196	77.354	28.493
77	0.8	0.8	-2	2.070	0.120	3.544	48.957	698.307	3618.114	3840.034	74.973	4.642
78	0.8	0.8	-1	2.068	0.143	3.551	48.882	275.143	2739.296	406.292	57.093	0.582
79	0.8	0.8	0	2.050	0.355	3.621	48.086	39.925	811.181	41.888	16.946	0.150
80	0.8	0.8	1	1.893	1.888	4.267	41.109	5.061	109.078	5.127	2.113	0.097

Next, a suitable code was developed to establish and train nine separate ANNs for intelligent stochastic modeling of each KMC output parameter of the OBC reaction. In other words, each ANN was responsible for mapping the complex relationship between the input and output variables and predicting one of the KMC outputs. During the training phase, the parameters of the networks, i.e. weights and thresholds (biases), were adjusted to minimize the prediction errors. The error of a particular configuration of the networks was determined by running all training data through the networks and comparing the generated outputs with the desired targets. After completion of the training process, the ANNs had gained the capability to predict the outputs based on any input data similar to the pattern that they had learned. Prior to entering the input data into the ANNs, the following linear transformation (in the range of -1 to +1) was used to normalize the input and response variables in order to prevent larger numbers overriding smaller ones, which could lead to premature saturation of hidden nodes.

$$X_i = 2 \times \left(\frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \right) - 1 \qquad (1)$$

where X_i is the normalized value of x_i, and x_{min} and x_{max} are minimum and maximum values of x, resspectively. The normalized data were subsequently divided randomly into training and test datasets. 70% of the data (56 scenarios/recipes among the 80 OBCs produced by KMC) were randomly selected for training the ANNs, while the rest were used for testing the trained ANNs. Five-layer ANNs were defined with 4-3-5-2-1 neurons in the hidden and output layers to model

the target microstructural features of the OBC chains with the input layer having three neurons for each independent reaction variable. Figure 1 depicts the architecture of the proposed ANNs.

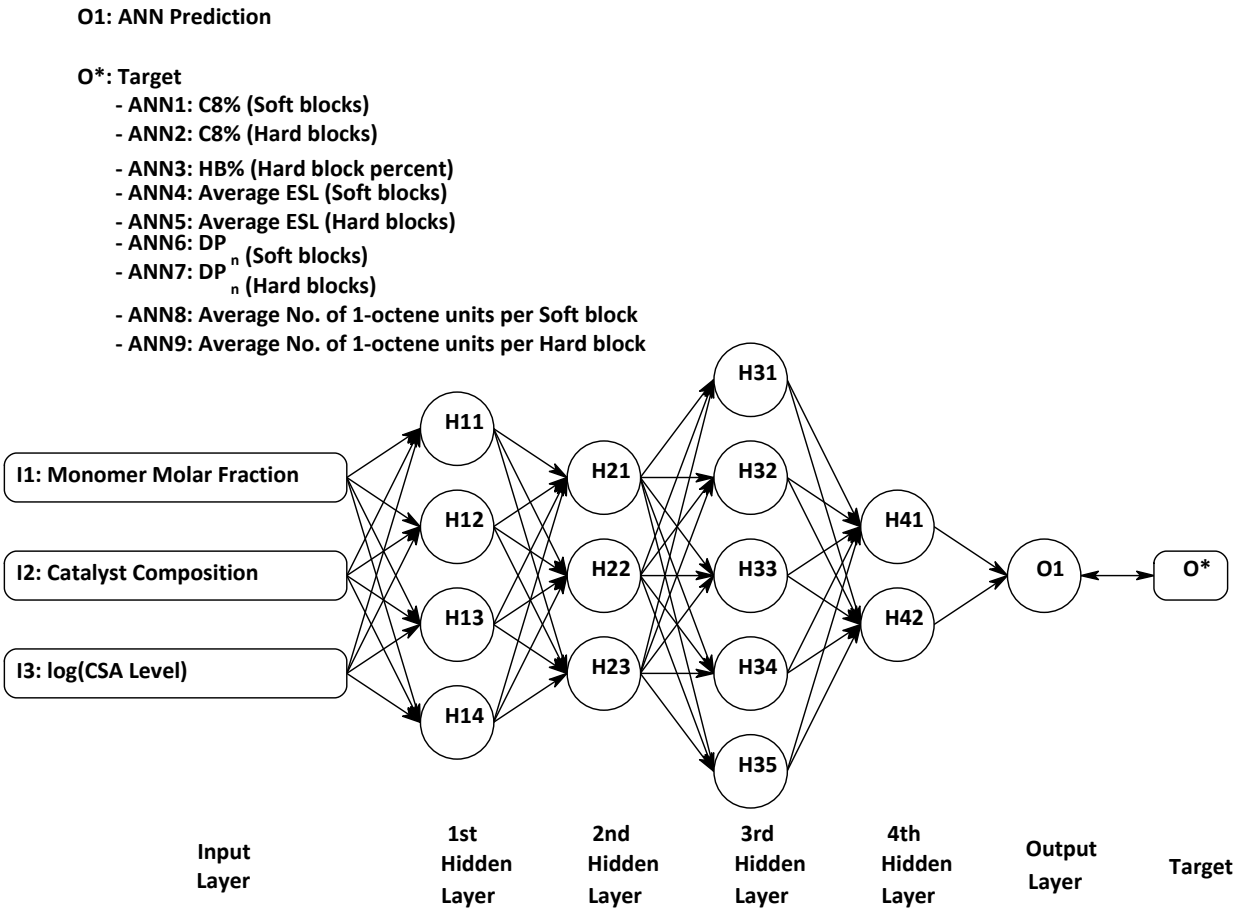


Figure 1. The structure of ANNs used to simulate the microstructure of the OBCs.

The hyperbolic tangent sigmoid function in the hidden and output layers was utilized as the activation transfer function to scale outputs of each neuron within the range of -1 to +1. The following activation function was used:

$$f(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{2}$$

where x is the sum of weighted inputs and the bias and f is the hyperbolic tangent sigmoid activation function. A genetic evolutionary algorithm was used to train the ANNs by coding the

unknown adjustable parameters of the network as chromosomes (Figure 2). Since it is customary to define a bias for a given neuron in the hidden and output layers, the number of unknown biases was equal to the number of neurons. Accordingly, the number of unknown biases and weights were 15 and 51, respectively.

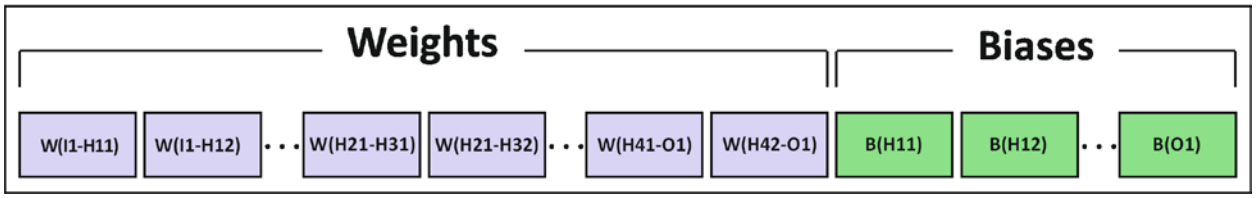


Figure 2. The chromosome structure to codify the adjustable parameters of the proposed ANNs. The weights and biases are represented by W and B, respectively.

It should be noted that each gene within the chromosome structure could take values in the range of -1 to +1. An initial population of 50 chromosomes was randomly generated and the ANNs were trained to scan and explore chromosomes by means of the genetic algorithm to find the best chromosome with minimum errors for both training and test datasets. Next, the selected chromosome *j* was independently entered into the predefined structure of the ANNs for evaluation of the networks based on weights and biases assigned to each chromosome. The normalized input values were separately entered in the network structure to determine the corresponding outputs. Then, the following relationship was used to evaluate the reliability of the ANNs in terms of the mean-squared error (*MSE*) [53]:

$$MSE(j) = \frac{1}{N_{training}} \sum_{i=1}^{N_{tarining}} \left(y_{i,ANN}(j) - y_{i,Target} \right)^2 \quad (3)$$

where $N_{training}$ is the number of data used for training (56 scenarios), and *MSE*(*j*) is the *MSE* for the *j*-th chromosome. Further, $y_{i,ANN(j)}$ is the output of the ANN corresponding to the *i*-th scenario with respect to the *j*-th chromosome, and $y_{i,Target}$ is the target value for the *i*-th scenario. The training and test errors were related to training and test *MSE*s as follows:

$$Error = \frac{\sqrt{MSE}}{Max(Error)} \times 100 = \frac{\sqrt{MSE}}{2} \times 100 \quad (4)$$

where *Max*(*Error*) is the maximum expected error of the networks, which was 2 as the data were normalized between -1 and +1. After determining the error for each chromosome, the criterion of minimum *MSE*s was used to sort chromosomes from lowest to highest error to identify the best chromosome. Then, selection, mating, crossover, and mutation operators were applied to produce a new generation of chromosomes. The new chromosome population was entered in the evaluation unit and the optimization process was continued until the selection of optimum chromosomes in accordance with the criterion of minimum *MSE*s. Merging, sorting, and truncating mechanisms were used for the selection operator, while the roulette wheel mechanism was employed to couple the selected chromosomes. One-point recombination was used to crossover the two parent chromosomes and produce an offspring, and the mutation operator picked randomly one gene from the selected chromosome and stochastically exchanged its value with a new digit in the range of -1 to +1. The mutation rate was fixed at 20.00%, meaning that the mutation operator was applied to 20% of the 'child' chromosomes in each iteration or epoch. The values of the parameters used for evolutionary optimization of the ANNs are given in Table 3.

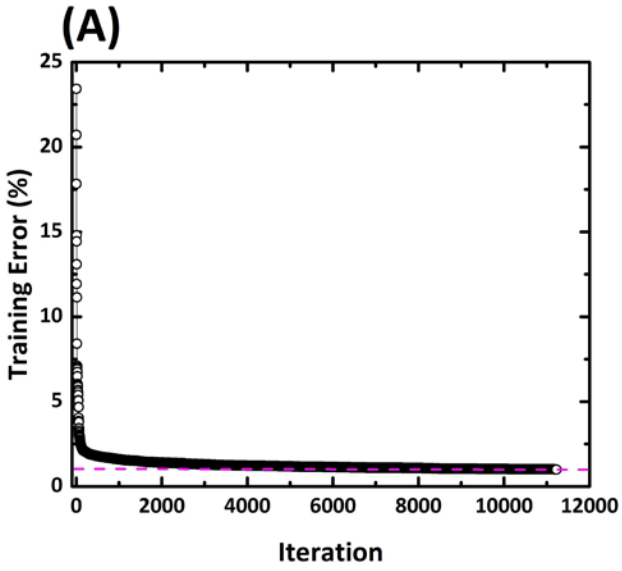
Table 3. The value of the parameters related to single-objective genetic algorithm-based optimization of ANNs.

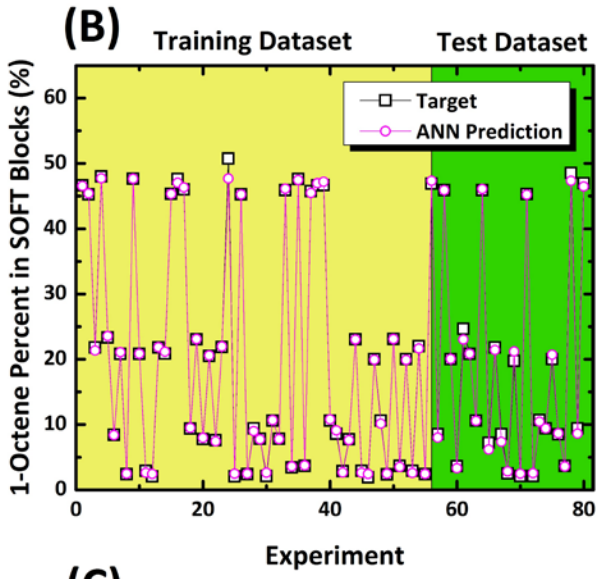
Optimization Parameter	Value
Initial Population Size	50
Selection Mechanism	Merge, Sort, and Truncate
Mating Mechanism	Roulette wheel selection
Crossover Mechanism	Single-point crossover
Mutation Rate	20.00%
Training and Test Errors[a]	1.00% and 2.00%, respectively

[a] The training and test errors were 3.00% and 4.00% in the case of Y3 ($HB\%$), respectively, and 0.50% in the case of Y7 and Y9.

After completion of the training stage, ANNs generated from the genes of the best chromosomes were evaluated by the test dataset. The optimization process was terminated when a test *MSE* of less than or equal to a preset *MSE* was achieved. Otherwise, the optimization algorithm returned to the training stage and the process was repeated.

Results for ANN-based modeling of cumulative 1-octene percent in soft blocks (Y1: $\overline{C8}\%^{Soft}$) in ethylene/1-octene OBCs are presented in Figure 3. *MSE* variations of the best chromosomes in the training data showed declining errors as the number of iterations increased. *MSE* variations for $\overline{C8}\%^{Soft}$ in Figure 3 declined from 23.43% to 0.99% within 11,229 iterations. The relatively large number of epochs (iterations) and small errors showed that the developed ANNs accurately predicted the variation pattern of $\overline{C8}\%^{Soft}$.


(A)
Training Error (%)
Iteration
0
5
10
15
20
25
0
2000
4000
6000
8000
10000
12000


(B)
Training Dataset
Test Dataset
Target
ANN Prediction
1-Octene Percent in SOFT Blocks (%)
Experiment
0
10
20
30
40
50
60
0
20
40
60
80

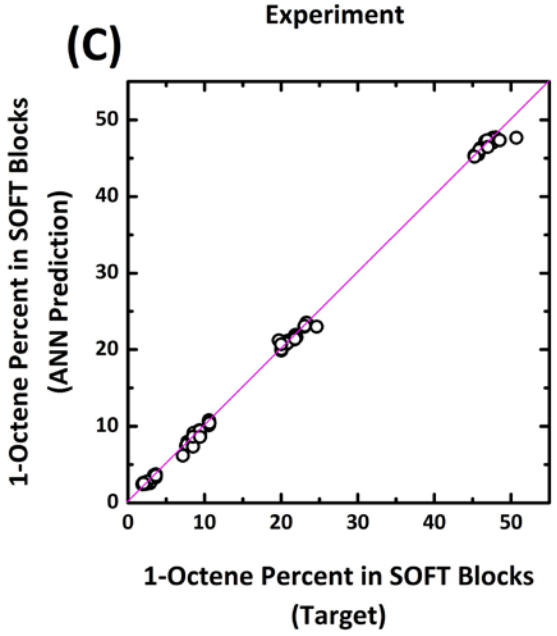

(C)
1-Octene Percent in SOFT Blocks
(ANN Prediction)
1-Octene Percent in SOFT Blocks
(Target)
0
10
20
30
40
50
0
10
20
30
40
50

Figure 3. (A) The dependence of ANN training error on number of iterations; (B) Comparison of the KMC outputs with the predicted values based on ANNs corresponding to training (yellow area) and test (green area) datasets for optimization of the response variable $\overline{C8}\%^{Soft}$; (C) Correlation between the KMC data and ANN-based predictions for the response variable $\overline{C8}\%^{Soft}$.

The performance of the trained ANNs was further tested by comparing the network predictions and the corresponding target values (KMC outputs) for $\overline{C8}\%^{Soft}$ for the entire training and testing datasets (Figures 3B and 3C). The similarity between ANN predictions and target values in Figure 3B is a testament to the accuracy of the developed ANNs in predicting the entire variable space and especially those cases that were not used in the network in the training phase. Further, Figure 3C shows that the trained ANNs were able to predict $\overline{C8}\%^{Soft}$ accurately. Similar results have been obtained for other response variables (Y2 to Y9) and are available upon request. Table 4 lists statistical information about the ANNs for the entire dataset (training and testing), demonstrating the accuracy of the developed ANNs in predicting the microstructural characteristics of OBCs.

Table 4. Results of statistical analyses related to artificial neural networks training and testing.

Response	**Y_1**	**Y_2**	**Y_3**	**Y_4**	**Y_5**	**Y_6**	**Y_7**	**Y_8**	**Y_9**
	$\overline{C8}\%^{Soft}$	$\overline{C8}\%^{Hard}$	$HB\%$	$\overline{ESL}^{Soft}$	$\overline{ESL}^{Hard}$	$\overline{DP_n}^{Soft}$	$\overline{DP_n}^{Hard}$	$\overline{N}_{C8}^{Soft}$	$\overline{N}_{C8}^{Hard}$
Training MSE	0.00039	0.00039	0.00359	0.00039	0.00099	0.00039	0.00039	0.00039	0.00009
Test MSE	0.00078	0.00133	0.00575	0.00125	0.00129	0.00109	0.00083	0.00142	0.00034
Training Error (%)	0.99936	0.99999	2.99926	0.99949	1.57728	0.99987	0.99991	0.99937	0.49875
Test Error (%)	1.39431	1.82189	3.79107	1.77011	1.79838	1.65747	1.44154	1.88695	0.91952
Max Training Error (%)	6.25839 (5)[a]	6.65651 (20)	17.46769 (33)	2.11045 (16)	8.31046 (56)	3.38446 (1)	3.87043 (76)	5.19647 (16)	2.41113 (1)
Max Test Error (%)	3.36469 (25)	5.39016 (5)	9.66778 (32)	4.87684 (80)	5.93603 (71)	5.33410 (67)	3.80190 (19)	3.74983 (28)	2.71168 (26)
R-Squared	0.99897	0.99523	0.98954	0.99829	0.99531	0.99845	0.99811	0.99778	0.99879
CC[b]	0.99948	0.99761	0.99476	0.99914	0.99765	0.99923	0.99906	0.99889	0.99939
CoE	0.99893	0.99517	0.98932	0.99822	0.99506	0.99843	0.99785	0.99771	0.99870
GoF (%)	96.72911	93.04744	89.66348	95.77579	92.97125	96.03834	95.36176	95.21012	96.39783
CoD	0.99897	0.99523	0.98954	0.99829	0.99531	0.99845	0.99811	0.99778	0.99879

[a] The digit within parentheses specifies the assigned number to the KMC RUN results for maximum error in training or testing *MSE*. [b] CC: Correlation Coefficient; CoE: Coefficient of Efficiency; GoF: Goodness of Fit; CoD: Coefficient of Determination.

According to Table 4, the *MSE* values of the optimized networks corresponding to training and testing stages were reasonably low. Another beneficial feature of this simulation is the ability to identify the scenario with maximum error in both the training and test phases. It should be noted that 'Scenario' refers to the number assigned to KMC RUN before modeling (as per the number in Table 2), whereas 'Experiment' refers to the assigned number to KMC RUN in the training and testing data sets. It appears that in some cases the large range of variability in the output responses was dependent on the variation pattern of the input reaction factors (as per Table 2). In spite of that, the coefficient of determination corresponding to all studied responses was close to unity.

2.2. Second Step: Artificial intelligence based multi-objective optimization

The non-dominated Sorting Genetic Algorithm-II (NSGA-II), a multi-objective version of the conventional GA, was utilized to simultaneously regulate several microstructural features of OBC chains generated virtually by the KMC simulator [47-48]. NGSA-II stochastically generates structured chromosomes as genotypes, visualizes as well as analyzes the corresponding macromolecules as phenotypes, and subsequently scans and reports the optimized ones that match the predetermined targets. The following approach was used for the artificial intelligence-based optimization of OBC reactions. First, the three independent variables to be optimized were coded as chromosomes (Figure 4) such that every chromosome had three genes corresponding to the number of independent/input variables. Next, the initial population of chromosomes was generated stochastically to initiate optimization through the evolutionary genetic algorithm. The values of the genes were randomly selected in the normalized range of -1 to +1 (Table 2). Then, the fitness of each chromosome (representing a quality indicator of the OBC chains expected from the corresponding scenario) was quantitatively determined by recalling the KMC simulator and the trained ANNs. The desired chromosome was the one satisfying the generation of OBCs with predetermined properties.

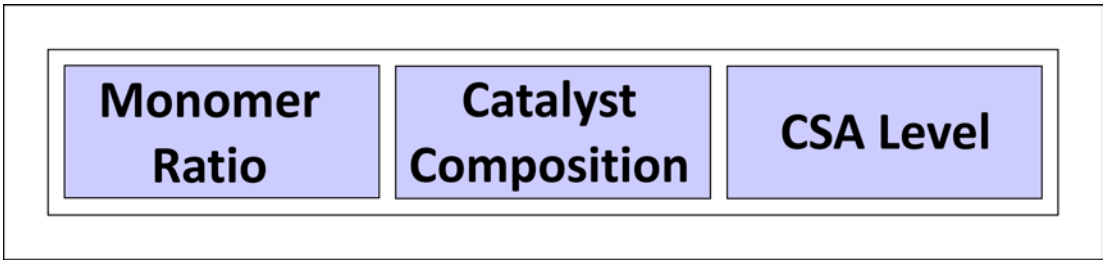


Figure 4. Defined chromosome structure for coding the copolymerization factors.

The principal difference between NSGA-II and conventional GA is in the sorting mechanism. The chromosomes in NSGA-II are sorted based on the concept of domination. Assuming $F_1(i)$, $F_2(i)$, $F_3(i)$, … and $F_n(i)$ as fitness values (or objectives) of chromosome *i*, the non-dominated sorting algorithm was used to determine fitness of the selected chromosome with respect to others. It was essential to sort the chromosomes based on order as well as on quality criteria. The former, known as the primary criterion, is of a higher degree of importance in sorting chromosomes. Accordingly, chromosomes were compared to each other to identify and label the dominated chromosome *i* over *j* with a Yellow Card based on the following condition:

$$i \quad dom \quad j \quad \Leftrightarrow \quad \begin{cases} \forall\, x: & F_x(i) \le F_x(j) \\ \exists\, y: & F_y(i) < F_y(j) \end{cases} \tag{5}$$

In other words, chromosome *i* dominated chromosome *j* if it was not worse than chromosome *j* for all objectives and strictly better than j for at least one objective. For each chromosome, the KMC simulator and the previously trained ANNs were recalled to extract fitness functions of each chromosome. After comparing the chromosomes, the non-dominated ones were screened and labeled as the first Pareto front [54-55]. This procedure was repeated to determine the second, third, and *n*-th Pareto fronts. The *n*-th Pareto front was assigned to the chromosomes that had been dominated (*n*-1) times. After ranking the chromosomes, the secondary criterion known as Crowding Distance, *C.D.*, was implemented for sorting the members of each Pareto front using the following criterion [56]:

$$C.D.(i) = \sum_{x=1}^{N} d_x(i) \quad where: \quad d_x(i) = \frac{\left|F_x(i+1) - F_x(i-1)\right|}{\left|\max(F_x) - \min(F_x)\right|} \tag{6}$$

where N is the number of objectives, $C.D.(i)$ is the crowding distance of chromosome i, and $d_x(i)$ is the crowding distance of chromosome i with respect to objective x. Applying the domination criterion, all chromosomes were sorted from best to worst with the best chromosomes lying on the first Pareto front (thus satisfying the minimum error in predicting the predefined targets). Next, the selection, mating, crossover, and mutation operators were applied on the sorted population to create a new generation of chromosomes. The adjustable genetic algorithm characteristics are listed in Table 5.

Table 5. Values of algorithm characteristics for multi-objective optimization by NSGA-II.

Optimization Parameter	Value
Initial Population Size	500
Selection Mechanism	NSGA-II
Mating Mechanism	Roulette wheel selection
Crossover Mechanism	Single-point crossover
Mutation Rate	15.00 %
Number of Iterations	5000

Using this approach, the KMC simulator benefited from artificial intelligence by hybridization with ANNs and GA. The resulting IMC algorithm is a powerful mathematical tool able to check billions of cases to identify a macromolecular pattern as close as possible to the desired macromolecule. Figure 5 shows the flow chart of the developed IMC algorithm.

The three separate programs based on the computational algorithms and summary flow chart presented in Figure 5 were written in Pascal programming language (Lazarus IDE) and compiled into 64-bits executable FPC 2.6.2. A subroutine based on the Mersenne Twister algorithm was used to generate random numbers for the simulations [57]. The random number generation subroutine satisfied the tests of uniformity and serial correlation with high resolution. The cycle length of the random number generator was $2^{19937}-1$. Simulations were performed on a desktop computer with Intel Core i7-3770K (3.50 GHz), 32 GB of memory (2133 MHz), under Windows 7 Ultimate 64-bit operating system. The run-time was approximately 10 hours, 85 minutes, and 42

seconds for each KMC simulation, including ANN-based modeling of each response (both training and test stages), and heuristic optimization of each case study, respectively. In contrast, the training and test stages of nine different ANNs utilized in the current study took around 12.75 hours. Interestingly, the intelligent optimization step recalls the trained ANNs several million times to handle the optimization problem in less than 42 seconds. Hence, as can be observed, the IMC approach is robust and versatile, and at the same time a computationally cost-effective simulation/optimization tool.

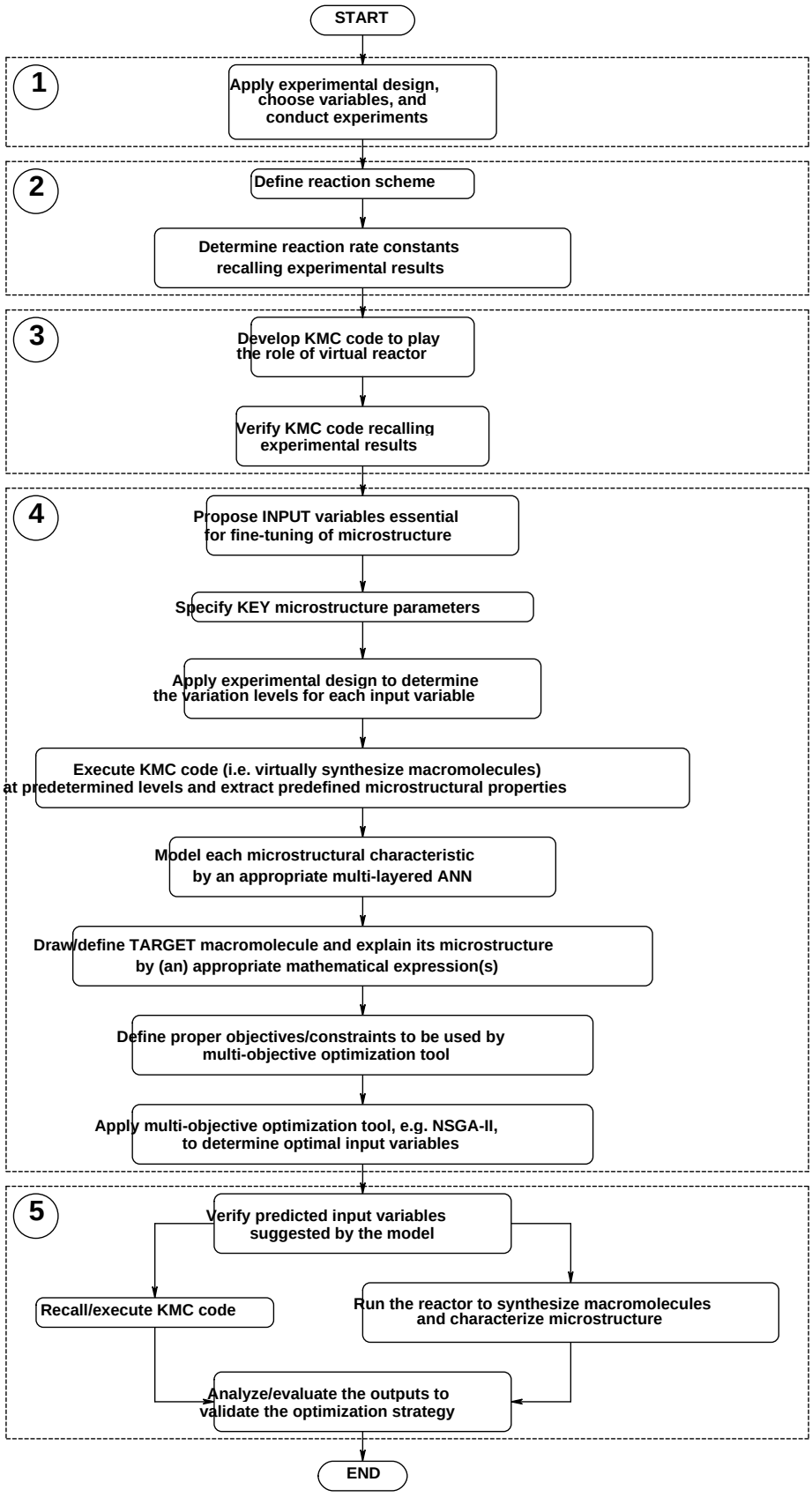


Figure 5. Overall summary flow chart for the IMC simulation approach.

3. RESULTS AND DISCUSSION

Suppose that several microstructural features of the polymer chains in a complex polymerization need to be controlled. This is similar to a situation where one draws on a sheet of paper, as detailed as possible, any hypothetical macromolecule. The IMC simulator acts as an intelligent consultant to bring the virtually drawn ('designed') macromolecule into existence. In this section, the IMC simulator is used to find the reaction conditions for target OBC chains in a quasi-living ethylene/1-octene copolymerization reaction with specified hard-soft segmental transitions and the corresponding changes in sequence length and sequence length distribution of ethylene units. The target OBC chains with hypothetically defined molecular characteristics were named OBC1, OBC2, OBC3, OBC4, and OBC5. The selection of a large gene pool, which depended directly on the intervals assigned to the input variables, allowed for an extended search space. Therefore, the variability of genotypes and the corresponding phenotypes determined the predictability range of the IMC simulator. In the following, the virtual synthesis of two tailored OBCs (OBC1 and OBC2) using the IMC is presented in detail; for the sake of brevity, the results from the synthesis of OBC3, OBC4, and OBC5 (which have a greater number of microstructural constraints) are available upon request.

3.1. OBCs with highly uniform microstructure (OBC1)

Suppose we are faced with the production of highly uniform OBCs in which 1-octene units are distributed evenly in both soft and hard blocks. The microstructure of a highly uniform OBC is schematically illustrated in Figure 6.

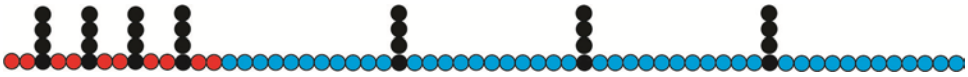


Figure 6. A typical OBC1 macromolecule with uniform distribution of short 1-octene branches (black chains) in soft blocks (red units) and hard blocks (blue units).

In OBC chains with uniform microstructure, the average *ESL* of the soft blocks ($\overline{ESL}^{Soft}$) is equal to the ratio of the average length of soft blocks ($\overline{DP}_n^{Soft}$) to the average number of comonomer

units per soft block ($\overline{N}_{C8}^{Soft}$). A similar definition is applied to the hard blocks. In mathematical terms, this requires minimization of the criteria defined by relations (7) and (8):

$$Objective\ 1: \quad MIN\left|\left(\frac{\overline{DP}_n^{Soft}}{1+\overline{N}_{C8}^{Soft}}\right)-\overline{ESL}^{Soft}\right| \tag{7}$$

$$Objective\ 2: \quad MIN\left|\left(\frac{\overline{DP}_n^{Hard}}{1+\overline{N}_{C8}^{Hard}}\right)-\overline{ESL}^{Hard}\right| \tag{8}$$

The first term in parentheses in equation (7) is the average ESL of an ideal uniform soft block having average block length $\overline{DP}_n^{Soft}$ and average number of 1-octene units per soft block of $\overline{N}_{C8}^{Soft}$. The second term in equation (7) is the simulated average ESL of the soft blocks. The virtual synthesis of OBC1 required concurrent control of six microstructural characteristics of the copolymer chains. Although the IMC simulator was trained and tested with only 80 scenarios, the simulation generated and analyzed billions of macromolecules to find the chain microstructure closest to the target OBC1 with a minimum number of iterations. Figure 7 shows the Pareto front showing the best solutions or the best operating conditions for the synthesis of OBC1.

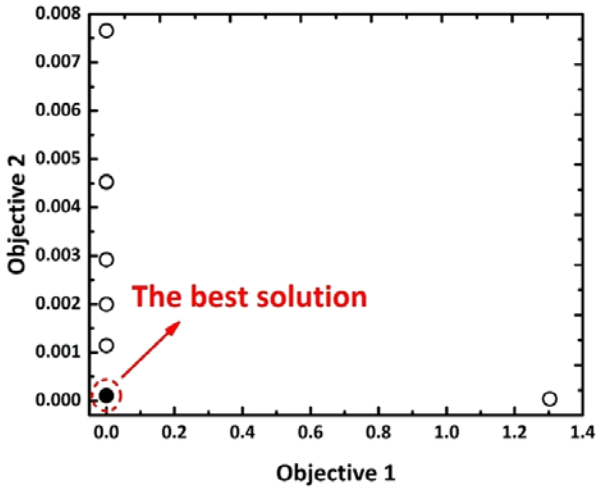


Figure 7. Pareto front for the multi-objective optimization of a highly uniform OBC1.

Objectives 1 and 2 in the abscissa and ordinate of Figure 7 are the criteria defined by relations (7) and (8), respectively. According to Figure 7, multiple solutions are proposed by the IMC simulator, among which the optimum solution (designated by the red circle) had the lowest values for objectives 1 and 2 with values of 1.88×10^{-4} and 1.02×10^{-4}, respectively. Additional details on Pareto fronts and optimal solutions corresponding to the OBC1 case can become available upon request. The optimal 1-octene molar fraction, catalyst composition, and log(CSA Level) for OBC1 synthesis, as proposed by IMC, are 0.37322, 0.65154, and -0.23192, respectively, and the corresponding response variables at these optimal conditions are given in Table 6. To validate the authenticity of the IMC results, the optimal reaction conditions were entered into the KMC simulator to virtually synthesize OBC chains and determine their microstructural characteristic. The KMC outputs were in good agreement with those of IMC as shown in Table 6 (green row).

Table 6. Optimal input and response variables generated by IMC (highlighted in yellow) and KMC (highlighted in green) simulations for the synthesis of OBC1.

Optimal Input Variables

1-octene Molar Fraction:	0.37322
Catalyst Composition:	0.65154
log(CSA Level)	-0.23192

Optimal Responses

	$\overline{C8\%}^{Soft}$	$\overline{C8\%}^{Hard}$	$HB\%$	$\overline{ESL}^{Soft}$	$\overline{ESL}^{Hard}$	$\overline{DP_n}^{Soft}$	$\overline{DP_n}^{Hard}$	$\overline{N}_{C8}^{Soft}$	$\overline{N}_{C8}^{Hard}$
IMC	29.2706	2.06407	11.7927	4.2255	26.9386	242.214	127.389	56.3195	3.72889
KMC	22.9705	2.07426	15.8438	4.74382	40.5258	227.801	88.1609	52.4429	1.83104
LV[a]	1.89253	0.11813	1.35431	2.33881	2.19434	5.12667	7.11763	0.92712	0.09724
HV[b]	50.7209	17.9001	58.3763	48.9687	825.428	38085.6	3736.87	375.675	1564.48
Error (%)	12.9026	0.05734	7.10444	1.11157	1.65046	0.03785	1.05177	1.03445	0.12132

[a] The lowest value of a given response among 80 scenarios presented in Table 2. [b] The Highest value of a given response among 80 scenarios presented in Table 2.

When the KMC simulator was applied to determine other characteristics of OBC1 chains at optimal conditions (columns not highlighted in Table 6), the results were consistent with those of IMC as shown in Table 6. Figure 8 shows the microstructural characteristics of OBC1 chains;

this is essentially the copolymer identification card showing both the evolution and final chain distribution of the intelligently synthesized OBCs. We previously found a broad distribution pattern for the average ESL of the OBC chains which made it nearly impossible to evenly distribute 1-octene units along the copolymer chains [35-37]. This was expected as there was only one average LES for each copolymer chain while $\overline{ESL}$ reflected the average length of many ethylene sequences for a given OBC chain. In other words, average LES is determined from the right-hand tail of the ESL distribution curve, whereas the peak of the distribution is the average ESL. Figure 8 shows that the ESL distribution curve is relatively narrow even though it was difficult to produce OBCs with similar ESL and LES distributions. The ESL and LES distribution curves for soft and hard blocks in Figure 8E and 8F, respectively, show uniform distribution of 1-octene units in hard and soft blocks. Therefore, IMC intelligently maintained the average LES and ESL for soft and hard blocks close to each other to the extent possible, as governed by the underlying polymerization kinetics.

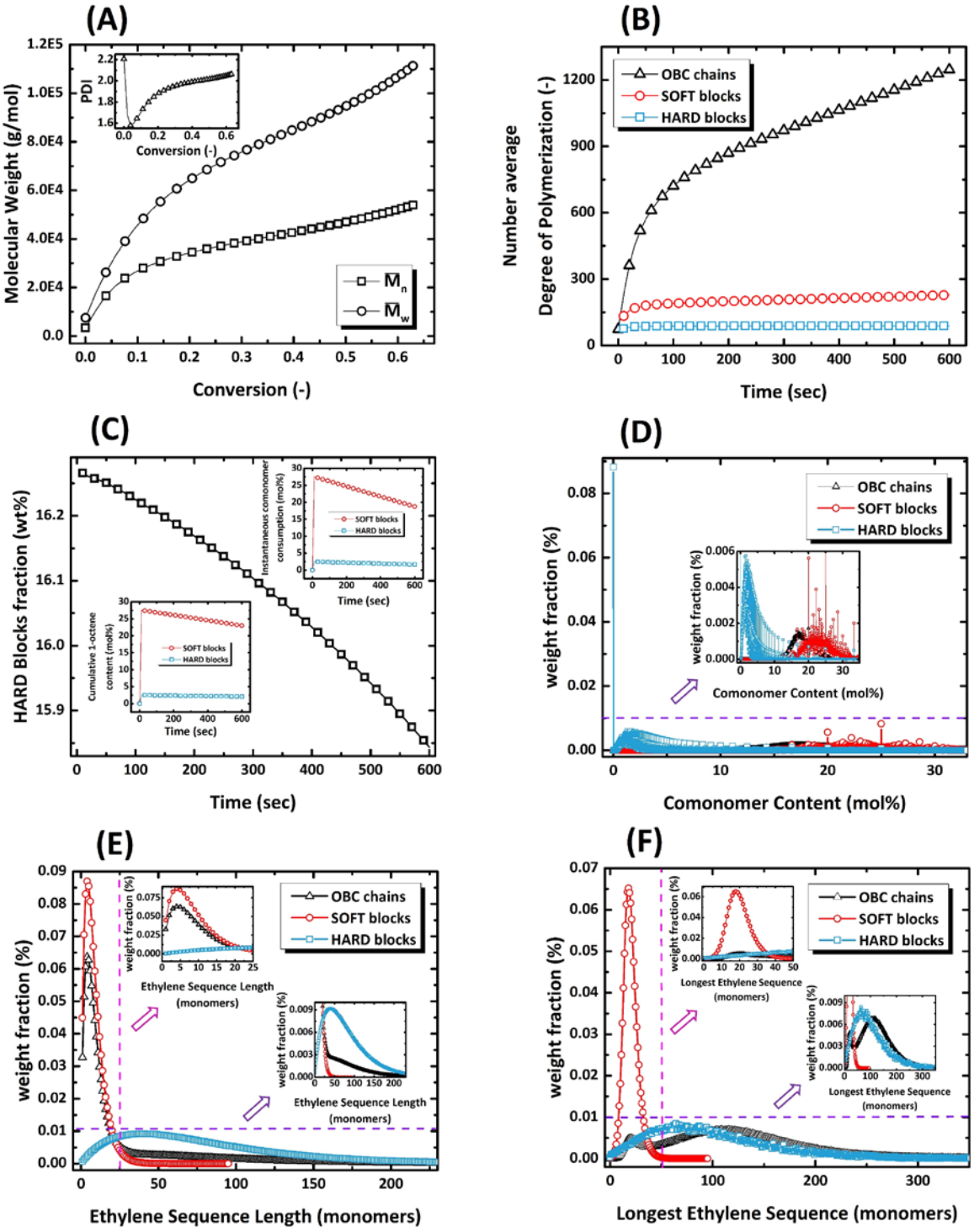


Figure 8. Instantaneous and final characteristics of OBC1.

3.2. OBCs with highly regular microstructure (OBC2)

IMC was applied to virtually synthesize highly regular OBC chains in which the comonomer ratio of soft to hard blocks and the ratio of the length of those blocks had predefined values. Those ratios for OBC2 were 30 and 0.2, respectively. This was equivalent to synthesizing OBC chains in which hard blocks were on average 5 times longer than soft blocks; in addition, the comonomer content of the soft blocks was 30 times higher than that of the hard blocks. The microstructure of a highly regular OBC2 is schematically illustrated in Figure 9.

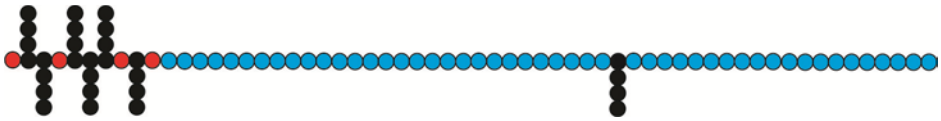


Figure 9. Chain microstructure of a highly regular OBC2 with red units as the soft block, blue units as a hard block, and black short chain branches as 1-octene comonomer.

From a mathematical standpoint, the following objective functions should be simultaneously minimized to produce the highly regular OBC2:

$$Objective\ 1: \quad MIN\left|\left(\frac{C_8\%^{Soft}}{C_8\%^{Hard}}\right)-30\right| \tag{9}$$

$$Objective\ 2: \quad MIN\left|\left(\frac{\overline{DP}_n^{Soft}}{\overline{DP}_n^{Hard}}\right)-0.2\right| \tag{10}$$

IMC was applied to simultaneously optimize four molecular characteristics of OBCs. The Pareto front in Figure 10 shows that the errors in synthesizing the target OBC2 were quite small for the best solutions or the best operating conditions (6.10×10^{-4} and 4.95×10^{-4}, respectively).

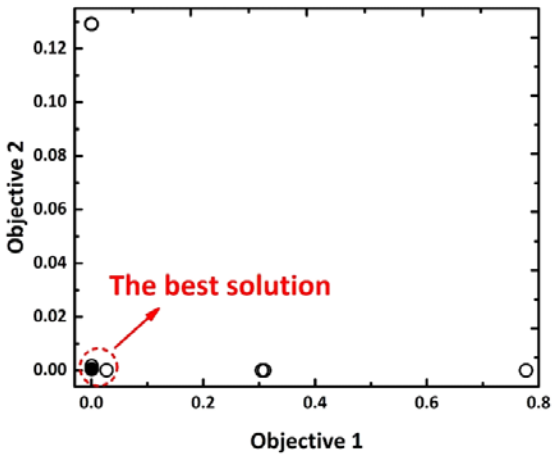


Figure 10. Pareto front for the multi-objective optimization of a highly regular OBC2.

The optimal input variables proposed by IMC for OBC2 synthesis and the corresponding response variables are listed in Table 7.

Table 7. Optimal input variables proposed by IMC for the synthesis of highly regular OBC2 and optimal responses by IMC (highlighted in yellow) and KMC (green).

Optimal Input Variables

1-octene Molar Fraction:	0.67484
Catalyst Composition:	0.20759
log(CSA Level):	-1.76063

Optimal Responses

	$\overline{C8\%}^{Soft}$	$\overline{C8\%}^{Hard}$	$HB\%$	$\overline{ESL}^{Soft}$	$\overline{ESL}^{Hard}$	$\overline{DP_n}^{Soft}$	$\overline{DP_n}^{Hard}$	$\overline{N}_{C8}^{Soft}$	$\overline{N}_{C8}^{Hard}$
IMC	7.90254	0.26342	48.7163	13.3843	271.828	1493.43	7448.71	90.7229	32.7738
KMC	7.48723	0.40181	48.0343	13.8085	246.898	1359.41	7629.85	101.801	30.7512
LV*	1.89253	0.11813	1.35431	2.33881	2.19434	5.12667	7.11763	0.92712	0.09724
HV**	50.7209	17.9001	58.3763	48.9687	825.428	38085.6	3736.87	375.675	1564.48
Error (%)	0.85056	0.77824	1.19608	0.90975	3.02829	0.35192	4.85654	2.95602	0.12929

Note: LV and HV as per notes of Table 6.

Figure 11 shows the microstructural characteristics of OBC2 chains (the copolymer identification card obtained by KMC using optimal polymerization conditions).

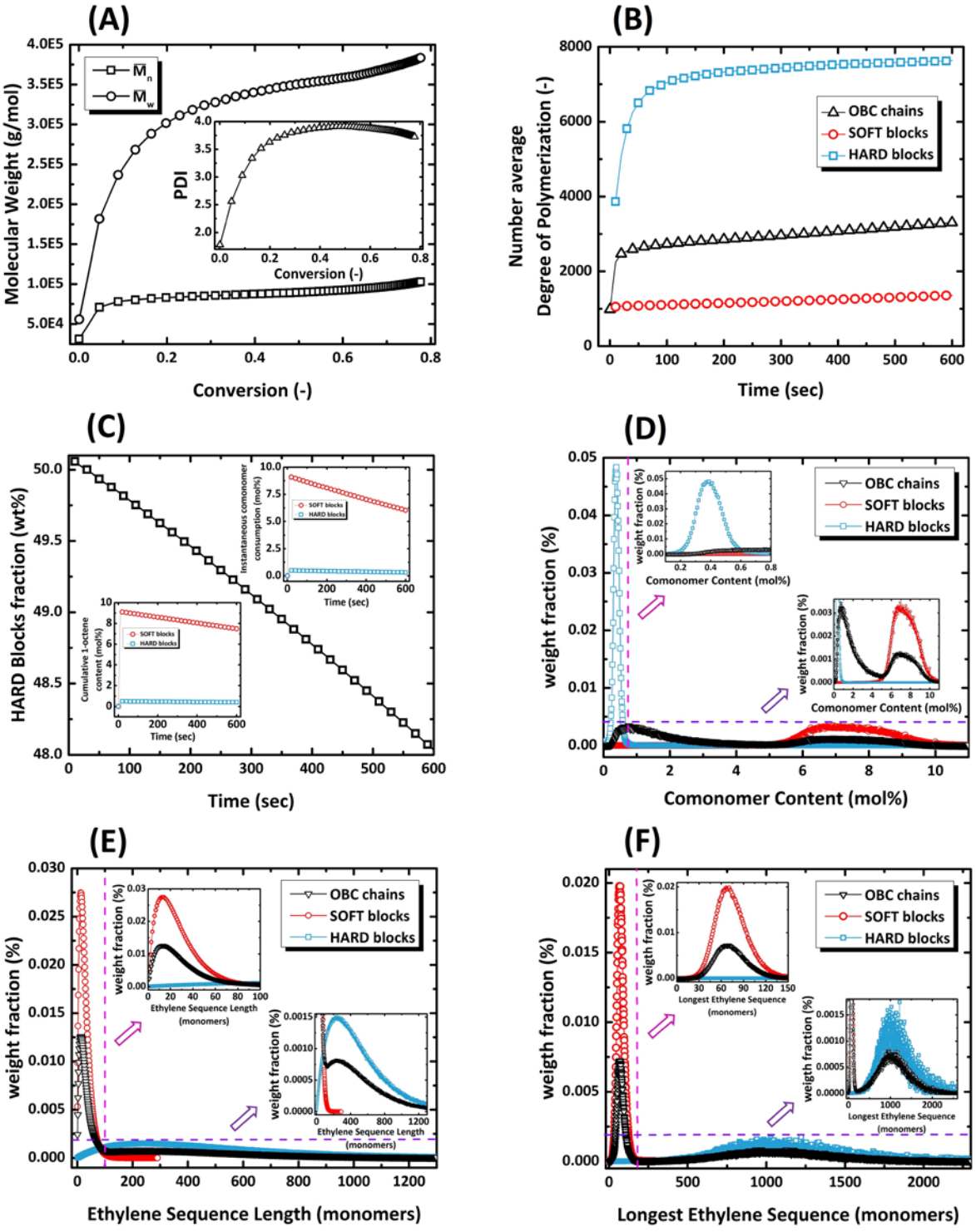


Figure 11. Instantaneous and end-of-batch features of OBC2.

4. CONCLUSION

This work proposes a novel approach for simulating the synthesis of polymers with complex microstructures, namely the Intelligent Monte Carlo (IMC) approach. IMC has the ability to make evolutionary decisions to accurately predict the input operating factors in response to target microstructural characteristics of the polymer chains. To illustrate the ability of the IMC approach, the chain shuttling copolymerization of ethylene and 1-octene was selected, considered a rather complex polymerization system. Macromolecular features describing a highly regular and a highly uniform chain microstructure were defined and simulated by the IMC. The power of IMC in concurrently controlling different microstructural characteristics was demonstrated by negligible errors calculated for the objective functions in all cases as compared to those errors obtained for the more tedious (and with longer time-frames) kinetic Monte Carlo (KMC) approach. IMC can thus be viewed as a comprehensive simulation-optimization package. It starts from the simulation of complex polymerizations, learns and identifies the dependency of microstructural features on operating conditions, finds the closest set of operating conditions that lead to control and optimization of several microstructural parameters simultaneously, and checks for the authenticity of the predictions by synthesizing/'designing' the target macromolecule.

ACKNOWLEDGEMENTS

The Fond Européen de Développement Régional (FEDER) is gratefully acknowledged for funding the France-Wallonie-Vlaanderen Interreg project ELASTOPLAST (PZ).

REFERENCES

1. W. Guo, D.G. Vlachos, Nature Communications, 2015, 6, 1.
2. C.H. Evers, J.A. Luiken, P.G. Bolhuis, W.K. Kegel, Nature, 2016, 534, 364.
3. J.C. Hernández-Ortiz, E. Vivaldo-Lima, M.A. Dubé, A. Penlidis, Macromolecular Therory and Simulations, 2014, 23, 429.
4. J.C. Hernández-Ortiz, E. Vivaldo-Lima, M.A. Dubé, A. Penlidis, Macromolecular Therory and Simulations, 2014, 23, 147.
5. G.W. Coates, R.M. Waymouth, Science, 195, 267, 217.
6. Z. Guan, P. Cotts, E. McCord, S. McLain, Science, 199, 283, 2059.

7. D.J. Arriola, E.M. Carnahan, P.D. Hustad, R.L. Kuhlman, T.T. Wenzel, Science, 2006, 312, 714.
8. K. Matyjaszewski, Science, 2011, 333, 1104.
9. J.-F. Lutz, M. Ouchi, D.R. Liu, M. Sawamoto, Science, 2013, 341, 1.
10. G. Gody, T. Maschmeyer, P.B. Zetterlund, S. Perrier, Nature Communications, 2013, 4, 1.
11. H. Tobita, Journal of Polymer Science Part B: Polymer Physics, 2011, 39, 391.
12. H. Chaffey-Millar, D. Stewart, M.M. Chakravarty, G. Keller, C. Barner-Kowollik, Macromolecular Therory and Simulations, 2007, 16 575.
13. S. Anantawaraskul, P. Somnukguande, J.B. Soares, Macromolecular Symposia, 2012, 312, 167.
14. D. Meimaroglou, C. Kiparissides, Industrial & Engineering Chemistry Research, 2014, 53, 8963.
15. M. R. Saeb, Y. Mohammadi, H. Rastin, T.S. Kermaniyan, A. Penlidis, Macromolecular Therory and Simulations, 2017, 26, DOI: 10.1002/mats.201700041.
16. P. Van Steenberge, D.R. D'hooge, M.-F. Reyniers, G.B. Marin, Chemical Engineering Science, 2014, 110, 185.
17. P. Derboven, D.R. D'hooge, M.-F. Reyniers, G.B. Marin, C. Barner-Kowollik, Macromolecules, 2015, 48, 492.
18. M. Najafi, V. Haddadi-Asl, Y. Mohammadi, Journal of Applied Polymer Science, 2007, 106, 4138.
19. M. R. Saeb, Y. Mohammadi, A.S. Pakdel, A. Penlidis, Macromolecular Therory and Simulations, 2016, 25, 369.
20. E. Mastan, X. Li, S. Zhu, Progress in Polymer Science, 2015, 45, 71.
21. M.M. Khorasani, M.R. Saeb, Y. Mohammadi, M. Ahmadi, Chemical Engineering Science, 2014, 111, 211.
22. E. Mastan, L. Xi, S. Zhu, Macromolecular Therory and Simulations, 2016, 25, 220.
23. H. Tobita, Macromolecular Therory and Simulations, 2017, 26.
24. P.H. Van Steenberge, D.R. D'hooge, Y. Wang, M. Zhong, M.-F. Reyniers, D. Konkolewicz, K. Matyjaszewski, G.B. Marin, Macromolecules, 2012, 45, 8519.
25. M.R. Saeb, Y. Mohammadi, M. Ahmadi, M.M. Khorasani, F.J. Stadler, Chemical Engineering Journal, 2015, 274, 169.
26. L. Wang, L.J. Broadbelt, Macromolecules, 2010, 43, 2228.
27. R. Szymanski, e-Polymers, 2009, 9, 538.

28. Y. Mohammadi, M. Najafi, V. Haddadi-Asl, Macromolecular Therory and Simulations, 2005, 14, 325.
29. J. Stajic, R. Stone, G. Chin, B. Wible, Science, 2015, 349, 248.
30. N. Nosengo, Nature, 2016, 533, 22.
31. P. Zinck, Polymer International, 2012, 61, 2.
32. P. Zinck, Polymer International, 2016, 65, 11.
33. Y. Phuphuak, F. Bonnet, G. Stoclet, M. Bria, P. Zinck, Chemical Communications, 2017, 53, 5330.
34. S. Li, R.A. Register, J.D. Weinhold, B.G. Landes, Macromolecules, 2012, 45, 5773.
35. Y. Mohammadi, M. Ahmadi, M.R. Saeb, M.M. Khorasani, P. Yang, F.J. Stadler, Macromolecules, 2014, 47, 4778.
36. M.R. Saeb, M.M. Khorasani, M. Ahmadi, Y. Mohammadi, F.J. Stadler, Polymer, 2015, 76, 245.
37. M. Ahmadi, M.R. Saeb, Y. Mohammadi, M.M. Khorasani, F.J. Stadler, Industrial & Engineering Chemistry Research, 2015, 54, 8867.
38. M.R. Saeb, Y. Mohammadi, T.S. Kermaniyan, P. Zinck, F.J. Stadler, Polymer, 2017, 16, 55.
39. A. Sfetsos, A.H. Coonick, Solar Energy, 2000, 68, 169.
40. S.A. Kalogirou, Progress in Energy and Combustion Science, 2003, 29, 515.
41. M.D. Fethi, F. Pasiouras, European Journal of Operational Research, 2010, 204, 189.
42. B.G. Sumpter, D.W. Noid, Macromolecular Therory and Simulations, 1994, 3, 363-378.
43. A. d'Anjou, E.J. Torrealdea, J.R. Leiza, J.M. Asua, G. Arzamendi, Macromolecular Therory and Simulations, 2003, 12, 42.
44. R.J. Minari, G.S. Stegmayer, L.M. Gugliotta, O.A. Chiotti, J.R. Vega, Macromolecular Reaction Engineering, 2007, 1, 405.
45. J. Zhang, A.J. Morris, E.B. Martin, C. Kiparissides, Chemical Engineering Journal, 1998, 69, 135.
46. J.C.B. Gonzaga, L.A.C. Meleiro, C. Kiang, R. Maciel, Computers & Chemical Engineering, 2009, 33, 43.
47. S. Garg, S.K. Gupta, Macromolecular Therory and Simulations, 1999, 8, 46.
48. A. Nayak, S.K. Gupta, Macromolecular Therory and Simulations, 2004, 13, 73.
49. S. Garshasbi, J. Kurnitski, Y. Mohammadi, Applied Energy, 2016, 179, 626.

50. R. Azari, S. Garshasbi, P. Amini, H. Rashed-Ali, Y. Mohammadi, Energy and Buildings, 2016, 126, 524.
51. B. Baghaei, M. R. Saeb, S. H. Jafari, H. A. Khonakdar, B. Rezaee, V. Goodarzi, Y. Mohammad, Journal of Applied Polymer Science, 2017, 134(33), 45145.
52. M. Hosseinnezhad, M. R. Saeb, S. Garshasbi, Y. Mohammadi, Solar Energy, 2017, 314.
53. S. Koroglu, N. Umurkan, O. Kilic, F. Attar, Simulation Modelling Practice and Theory, 2009, 17, 1267.
54. S.Y. Liong, S.T. Khu, W.T. Chan, Journal of Hydrologic Engineering, 2001, 6, 52-61.
55. X. Blasco, J.M. Herrero, J. Sanchis, M. Martinez, Information Science, 2008, 178, 3908.
56. P.C. Chang, S.H. Chen, C.Y. Fan, C.L. Chan, Applied Mathematics and Computation, 2008, 205, 550.
57. M. Matsumoto, T. Nishimura, ACM Transactions on Modeling and Computer Simulation 1997, 8, 3.