An Isochrone-Based Predictive Optimization for Efficient Ship Voyage Planning and Execution

Downloaded from: https://research.chalmers.se, 2024-08-18 04:06 UTC

Citation for the original published paper (version of record):

N.B. When citing this work, cite the original published paper.

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

This document was downloaded from http://research.chalmers.se, where it is available in accordance with the IEEE PSPB Operations Manual, amended 19 Nov. 2010, Sec, 8.1.9. (http://www.ieee.org/documents/opsmanual.pdf).

(article starts on next page)
An Isochrone-Based Predictive Optimization for Efficient Ship Voyage Planning and Execution

Yuhan Chen and Wengang Mao

Abstract—A voyage optimization algorithm is an essential component in a ship’s routing concerning safety, energy efficiency, arrival punctuality, etc. In this study, predictive optimization is integrated with an Isochrone-based voyage optimization algorithm for energy-efficient sailing. Different waypoints generation and grid partition strategies in search spaces are proposed to achieve smooth convergence toward the destination, and costs ahead of the current sailing time stages are estimated in the cost function to avoid the local suboptimization. Based on these measures, this paper introduces the Isochrone-based predictive optimization (IPO) method that can achieve enhanced and robust performance in real-time multi-objective voyage optimization. The unrealistic routes with abrupt turns that occur in the traditional Isochrone and graph search methods are avoided. The IPO method can suggest energy-efficient routes in diverse sailing environments, while complying with punctuality requirements in voyage planning. Meanwhile, it requires a few computational resources that enable online and real-time adjustment during voyage execution, adapting to dynamic sailing environments. Its efficiency and effectiveness are demonstrated by six case study voyages from a chemical tanker with full-scale measurements, and further compared with other widely used voyage optimization methods. The results show that the proposed method can provide smooth routes with subtle turns with 5% fuel reduction on average for all case voyages, with around 40 seconds runtime.

Index Terms—Energy efficiency, Isochrone method, predictive optimization, voyage optimization.

I. INTRODUCTION

The IMO (International Maritime Organization) has enhanced its greenhouse gas (GHG) strategy with an ambition of a 20% emission reduction by 2030 and net-zero emission by 2050 [18]. However, the capability to produce fossil-free fuels required to reach the goal cannot be available so soon [40]. Therefore, the maritime community is actively urging the development and implementation of various shipping energy efficiency measures in the market. Here, energy efficiency refers to minimizing fuel consumption for a ship’s given voyage with pre-defined transport tasks. Since over 80% of global trade is carried out by sea, an average of 1% fuel savings from shipping could contribute to about 8 million MT of CO2 reduction per year [18]. Ship voyage optimization systems are getting great attention due to their capabilities to not only directly assist ship operations to save fuel, but also optimize the utilization of various energy efficiency measures [13]. A voyage optimization system can incorporate a wide range of operation factors, such as weather conditions, energy consumption, arrival time, safety [66]. The core part of such a system is the voyage optimization algorithm, which is investigated in this study.

A. Literature Review on Ship Voyage Optimization Algorithms

Various algorithms were developed for optimal ship voyage planning [60]. Based on if speed variation is considered in the voyage optimization, the algorithms can be divided into two-dimensional (2D routes) and three-dimensional (3D trajectories, i.e., routes with speed/time profiles) methods. Those algorithms can also be categorized into sampling-based and search-based methods. For sampling-based methods, a typical example is Rapidly exploring Random Tree (RRT), which is used more in narrow or multi-obstacle scenarios, e.g., collision avoidance or inland waterway navigation. The other example is probabilistic roadmap (PRM). These sampling-based methods may not be related to this study thus not discussed in the literature review. And the search-based methods can further be categorized as follows:

1) Static Grid-Based Method: In voyage optimization systems available in the shipping market, search-based methods are widely used. For example, dynamic programming, which was initially proposed by Bellman [3] and further developed in [22] and [64], etc. Besides, graph search algorithms are also broadly researched. They discretize the sailing area into sections and time stages by a pre-defined static waypoint grid and associated edges. Two common algorithms using the static grid systems are the Dijkstra [9] and the A * algorithm [17], where A * can be seen as a generalized Dijkstra algorithm with an extra heuristic term in the cost function. Both methods have undergone extensive development. The Dijkstra algorithm was developed in recent research to e.g., include the speed/time domain [35], [36], [37], [57], considering collision risks [67], achieve simultaneous optimization [32], combine with 3D dynamic programming [7], formulate an iterative search [1], and combine with the automatic identification system (AIS) [48]. The A * algorithm was also implemented in e.g., attaining adaptive optimization [41], [47], improvement to 3D
directed in decision-making [58], and artificial network algorithm voyage planning. The particle swarm optimization is also used in multi-objective evolutionary algorithm integrated in [29] for based reinforcement learning [38].

Routing [65]; improved genetic algorithm used in [34]; and algorithm (ACO) to reduce emissions [33]; a 3D ACO in ice solving voyage optimization problems, e.g., the ant colony machine learning (ML) techniques, in recent years more comprehensive navigation monitoring system [14]. These voyage optimization algorithms can be easily implemented; however, their optimization outcomes and computation load closely depend on the grid configuration, e.g., the resolution of the grid and the range of search grid area, etc. Moreover, the final route optimized by those methods often needs to be smoothed for actual ship navigation.

Based on these two standard methods, many variants are also derived to overcome their incompetencies. For example, Dijkstra algorithm led to D* (Dynamic A*), focused D* and D* Lite, and A* spawned bidirectional A*, Theta* (any angle planning), and Lifelong Planning A* [8]. Examples in shipping can be found in [4] that uses multi-criteria Dijkstra’s Shortest Path algorithm in arctic routing. However, these variants are generally more related to pathfinding problems and are more researched for marine autonomous surface ships (MASS).

2) Dynamic Grid-Based Method: Dynamic grid-based methods perform route searching iteratively, and thus do not require the grid to be pre-defined. The waypoints at the following time stage are generated based on the current step, therefore the grid is iteratively progressed, and routes proceed accordingly until reaching the destination. One well-known example is the Isochrone method. An isochrone is a contour consisting of the farthest waypoints, that could be reached following different directions at the same pre-defined time stage. It was originally introduced [21], extended [16] by James, and then modified by Hagiwara [15] for ship voyage planning. The ship performance was later integrated into the Isochrone method [44]. Further, a 3D Isochrone method was proposed [31], and applied in the particle swarm optimization by Lin et al. [30]. The equal sailing time was also replaced with equal fuel consumption as Isopone [23], and equal cost as Isocost method [54] respectively. Similar approaches can also be found in combining the Isochrone method with the generic algorithm [26]. Szlapczynska et al. [53] applied the Isochrone methods in an evolutionary approach for voyage planning, and Sasa et al. [45] further developed this method to include speed loss and ship maneuvering. Another example of dynamic grid-based is the Dividing Rectangles method [25].

3) Advanced Methods: Voyage optimization, as aforementioned, can be sophisticated especially when comprehensively integrated with various components in the navigation system. It often involves processing large volumes of data, dynamic changes, predictions that include uncertainties, etc., thereby making the problem both large in scale and complex. Alongside the emerging trends in artificial intelligence (AI) and machine learning (ML) techniques, in recent years more complex approaches have been proposed. To effectively deal with such challenging problems, advanced methods are utilized in solving voyage optimization problems, e.g., the ant colony algorithm (ACO) to reduce emissions [33]; a 3D ACO in ice routing [65]; improved genetic algorithm used in [34]; and multi-objective evolutionary algorithm integrated in [29] for voyage planning. The particle swarm optimization is also utilized in decision-making [58], and artificial network algorithm based reinforcement learning [38].

B. Challenges and Objectives

Advanced algorithms can consider more control parameters and allow for more frequent variations of those parameters along the voyage optimization. However, their high computational demand and constant change in navigation settings can make such voyage optimization systems not applicable to the actual shipping market [49], [56]. For example, ocean-crossing ships are normally too heavy to execute flexibly changing sailing status (speed, heading, power, etc.), as output from those complicated algorithms. Furthermore, frequent status changes require continuous adjustment of ship maneuvering, which also leads to increased fuel consumption, emissions, and navigation risks. Some meta-heuristic approaches, unlike the deterministic Isochrone method, cannot guarantee the return of the result even if one exists.

To assist shipping efficiency, voyage optimization is used at two separated stages of ship operations, i.e., voyage planning and execution. A voyage planning is normally conducted before departure. However, the voyage optimization requires reliable weather forecast inputs, and the forecast contains large uncertainties beyond 2-3 days [63]. Thus, the original planning requires updates during the sailing to adapt to the dynamic environment. The voyage executions may also be influenced by various uncertainties/dynamics [62]. In addition, the voyage execution can also be influenced by commercial dynamics in the shipping market, e.g., fluctuated freight rate, charter rates, fuel prices, blocked route [2], [42], [43], [59], where some ships may have to change her original route, ETA, or destination port [12], [39]. Therefore, the capability of optimization algorithms to conduct fast voyage optimization is essential to allow for real-time updates during voyage executions, to address those weather/commercial dynamics. And computationally efficient algorithms and simple configurations of ship navigation control in voyage optimization are still widely used, such as the 2D Isochrone method.

The computational efficiency of the Isochrone methods has been demonstrated in practice for decades. Furthermore, its characteristics can ensure a more accurate ETA for voyages [15]. In the industrial sector, it is widely acknowledged that Just In Time (JIT) approaches, i.e., precisely adhering to planned schedules, are essential for ships to enhance efficiency and cost-effectiveness [20], [43]. It not only avoids employing excessive speeds to reduce fuel usage, but also ensures prompt operation processes and eliminates long waiting times at port. For ships with flexible destination ports and ETAs, the proposed optimization algorithm can be executed quickly to generate optimal voyages, adapt to those changes in destination ports and ETAs, and ensure punctuality.

However, one disadvantage for the Isochrone method is the so-called “isochrone loop” [61], which is an irregular shape of an isochrone caused by the non-convexity of a ship’s performance at sea. This isochrone loop propagates when the number of isochrones increases and leads to inapplicable results [44]. Furthermore, the current 2D Isochrone method by Hagiwara [15] is outdated with obvious flaws in route convergence at the final stage. Fig. 1 presents a typical example of the optimized route set by [15] where abrupt turns around destination can be clearly observed. Simple treatment of the
convergence problem may lead to significantly reduced search space and locally optimized results [6]. Other researchers developed the Isochrone method by including the speed variation and advanced ML algorithms, but also increased complexity. In this paper, a predictive optimization is proposed in the 2D Isochrone voyage optimization method with at least three innovative improvements: 1) dividing the voyage optimization into two stages with different waypoints/grid partition methods, 2) refining the cost function for searching waypoints to avoid local optimization, and 3) integrating physics-informed machine learning ship performance model for fast and reliable predictive optimization. Furthermore, the proposed IPO voyage optimization method is a deterministic approach that guarantees a result is returned if one exists, while possessing practicality and computation efficiency for real-time applications inherited from the original method. The framework and procedure of the proposed IPO method are described in detail in Section II. In Section III, the case study is presented to demonstrate the capability of the proposed algorithm. The comparisons between the proposed algorithm, actual routes, and other well-developed algorithms, are given in Section IV accompanied by discussions.

1) Limitations of the Scope: It should be noted for delimitations that there are some aspects not covered in this paper. For ship safety, comprehensive consideration requires models to describe safety margins, which is beyond the scope of this study. Safety can be integrated as supplementary components for further practical implementations.

The proposed method addresses the ship weather routing problem and definitions can be referred to [19] and [66]. Thus, it requires information of destination port and ETA as prerequisites, which can be given by decisions from maritime economics and management. Based on the given destination and ETA, it helps seafarers to conduct voyage planning and execution, and optimize the route and speed for a single voyage with safety and increased energy efficiency.

If commercial dynamics lead to any changes when the ship is en route, e.g., a new port of call, and/or new ETA, this method can efficiently update the voyage accordingly. For different destination ports and ETAs, this method can also quickly provide optimal voyages with associated minimized fuel cost, which may be valuable inputs to maritime economics and management for their decision making.

II. THE PROPOSED ISOCHRONE-BASED PREDICTIVE OPTIMIZATION (IPO) METHOD

For a ship’s voyage optimization by an Isochrone method, the voyage should be first divided into a series of time stages, i.e., \( T_i, i = 0, 1, \ldots, n, f \), from the departure \( P_0 \) to the destination \( P_f \). Fig. 2 illustrates the division of generated candidate routes in the Isochrone method with discretized time stages. \( T_0 \) represents the departure time and \( T_f \) represents the expected time of arrival (ETA).

For each \((i^{th})\) time stage, potential sailing waypoints \( P_{i,j}^k \) are generated from the previous \(((i-1)^{th})\) time stage’s waypoint \( P_{i-1,k} \), and denoted by,

\[
P_{i,j}^k = [x, y, T_i].
\]

where \( x \) and \( y \) represent its longitude and latitude at the passing time \( T_i \). The subscript \( j \) indicates that \( P_{i,j}^k \) is the \( j^{th} \) new waypoint/successor generated from the waypoint \( P_{i-1,k} \), and the superscript \( k \) indicates that \( P_{i-1,k} \) is the \( k^{th} \) pre-reserved waypoint at the \((i-1)^{th}\) time stage isochrone \( \{P_{i-1}\} \). The total number of waypoints and time stages are pre-defined as parameters, as they influence the partition of search space in the Isochrone method. It will be further discussed in detail in Section II-D.

The number of newly generated waypoints \( \{P_{i,j}^k\} \) will be a multiple of the current waypoints in \( \{P_{i-1}\} \). As the voyage advances in time, i.e., \( i \) increases, the number of search waypoints will increase exponentially, easily exceeding computational capacities. To avoid this, the Isochrone optimization method pre-reserves a limited number of “optimal” waypoints at the \( i^{th} \) time stage isochrone \( \{P_i\} \), by waypoint selection according to the sailing cost associated with the new waypoints \( \{P_{i,j}^k\} \).

The cost \( C_p \) corresponding to the waypoint \( P \) can be, e.g., fuel consumption, emissions, fatigue damage, maximum motions, etc., for sailing from the previous waypoint (e.g., \( P_{i-1,k} \)) to the current waypoint \( P_{i,j}^k \) (i.e., along the sub-route denoted by \( R_{i,j}^{k-1} \)). The sailing costs assigned to the waypoint \( P \) can be estimated by different cost functions according to the voyage optimization objectives as,

\[
C_p = f(S(P))
\]

where \( S(P) \) is the sailing state and described by its corresponding sailing parameters and environmental conditions associated with the waypoint \( P \) viz,

\[
S(P) = [V, \theta, S(\omega|H_s, T_c), V_e, V_w].
\]

where \( V \) and \( \theta \) represent a ship’s sailing parameters, i.e., ship speed \( V \) and heading \( \theta \), respectively, and the rest symbols represent a ship’s encountered wave, wind, and ocean current conditions at the waypoint \( P \). Here the encountered wave
environments are described by the ISSC wave spectrum in terms of the significant wave height \( H_s \), the wave period \( T_z \). And wind \( V_w \) and ocean current \( V_w \) are described by horizontal (toward east denoted by \( x \) as subscript) and vertical (toward north denoted by \( y \) as subscript) speeds, as follows,

\[
S(\omega|H_s, T_z) = \left(4\pi^3 H_s^2 / T_z^5 \omega^5\right) \exp\left[-(\omega T_s / 2\pi)^{-4}/\pi\right]
\]

\[
V_x = [V_{cx}, V_{cy}]
\]

\[
V_w = [V_{wx}, V_{wy}]
\]

(4)

Then, the cost function can be formulated to calculate the sailing costs for waypoint selection, according to customized voyage optimization objectives, e.g., minimum fuel, earliest arrival time, etc. The identification of the “optimal” waypoints among all potential waypoints at each time stage is an essential part of the Isochrone optimization algorithm. Since the grid of the Isochrone method is dynamic, all the selected waypoints are temporarily pre-reserved. If one pre-reserved waypoint (e.g. \( P_{j-1, k} \)) cannot proceed further, i.e., none of its successors \( P_{j, k} \) is nominated into the next \((i^{th})\) isochrone \( P_i \), the waypoint \( P_{j-1, k} \) together with all preceding waypoints/sub-routes it connects to will be eliminated for the upcoming time stages. Therefore, the Isochrone optimization method has to face two contradictory phenomena when a voyage approaches its destination, i.e., 1) if the search space expends too wide and scattered as in Fig. 1, the abrupt turns appear in the subsequent time stages leads to most of the candidate voyages not applicable for actual operations; and 2) if the search space converges too narrowly towards the destination, only limited amounts/locally optimized routes will be left for selection. These two issues will be addressed by the proposed method described in the following sections.

When the voyage search advances to its final (i.e., \( n^{th} \)) time stage and all the pre-reserved waypoints \{\( P_{n, k} \)\} at this isochrone \{\( P_n \)\} is found, the entire candidate “optimal” route set can be formed by integrating all the preceding sub-routes as follows:

\[
R^k = \{P_0, R_1^{0\rightarrow k^*}, R_2^{k^*\rightarrow k^*_2}, \ldots, R_{n-1}^{k^*_{n-2}\rightarrow k^*_{n-1}}, R_n^{k^*_{n-1}\rightarrow k^*}, P_f \}.
\]

(5)

where \( R_{n-1}^{k^*_{n-1}\rightarrow k^*} \) is the sub-route between \( P_{n, k} \) and its predecessor waypoint \( P_{n-1, k^*_{n-1}} \) at the \((n-1)^{th}\) time stage isochrone, \( P_{n-1, k^*_{n-1}} \in \{P_{n-1} \} \in \{P_k \} \).

Finally, all the candidate routes \( R^k \) found by the Isochrone method will have similar arrival times. The optimal voyage \( R^* \) is chosen based on the total sailing cost according to the optimization objectives, such as minimum fuel consumption, emissions, etc.

A. Overall Procedure of the Proposed Isochrone-Based Predictive Optimization (IPO)

To initialize the Isochrone-based Predictive Optimization (IPO) algorithm, first the reference speed \( V_s \) should be set:

\[
V_s = D / (T_f - T_0).
\]

(6)

where \( D \) is the geographical distance of the reference route \( GC_{ref} \), which is the great circle (GC) route from \( P_0 \) to \( P_f \).

Due to the involuntary speed reduction under rough weather by restricted engine power, the final sailing time may be longer than the expected \( T_{ETA} = T_f - T_0 \), and \( V_s \) can be set as a bit higher than (6).

In this study, three assumptions are made in the proposed IPO method. Firstly, the ship voyage planning is treated as a 2D optimization problem, which is widely used for practical voyage planning of ocean crossing vessels [56], because of its fast optimization for online updating and execution of voyage planning that can facilitate a dynamic shipping market, such as changes of destination ports, arrival time, and weather forecast [63]. Secondly, the loading conditions and ship status (biofouling, ways of navigations) along the entire voyage are assumed to be the same [55]. Thirdly, when discretizing a voyage into various time stages, the duration of each time stage is set to be small enough to assume stationary sea environment [53], where the ship speed-power performance at each time stage is assumed to be quasi-static [28].

The overall procedure of the proposed IPO method is presented in the flowchart in Fig. 3. In the initialization of the algorithm, the reference speed \( V_s \) and five parameters should be set. Then departing from \( P_0 \), the voyage search is carried out in recursion. Based on the current distance to \( P_f \), the search process is divided into two distinct parts. The main different steps within these two parts are listed for comparison in their blocks in Fig. 3. And finally, when the latest waypoints are closer to \( P_f \) than one time stage’s distance, they are connected to \( P_f \) directly.

To describe the waypoint generation of the IPO method, symbols used for the description, as well as five essential parameters needed to configure the waypoint generations are introduced in Table I.

B. Isochrone of the 1st Half Voyage by the IPO Method

Departing from \( P_0 \), the first half of the voyage search is conducted as follows. The first step is to generate the first isochrone \( \{P_1\} \) based on \( P_0 \), as in Fig. 4(a):
 same steps as above, potential waypoints \( P \), search.

CHEN AND MAO: IPO FOR EFFICIENT SHIP VOYAGE PLANNING AND EXECUTION 5

\[
\Delta \theta = \text{The increment of the heading angle for each current waypoint}
\]

\[
2m+1 = \text{The number of successors for each current waypoint}
\]

\[
2r = \text{The number of subsectors.}
\]

\[
\Delta D = \text{The width of search limit in distance for each local sub-sector.}
\]

\[
\Delta t = \text{The time interval between adjacent time stages}
\]

### TABLE I

**Symbols**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_s )</td>
<td>The departure waypoint of the planned voyage</td>
</tr>
<tr>
<td>( P_f )</td>
<td>The destination waypoint of the planned voyage</td>
</tr>
<tr>
<td>( GC_{ref} )</td>
<td>The reference route from ( P_s ) to ( P_f ) following the GC route.</td>
</tr>
<tr>
<td>( D )</td>
<td>The geographical distance between ( P_s ) and ( P_f ) in ( GC_{ref} ).</td>
</tr>
<tr>
<td>( C_{ref} )</td>
<td>The initial course of ( GC_{ref} ) at ( P_s ).</td>
</tr>
<tr>
<td>( C_{arrd} )</td>
<td>The arrival course of ( GC_{ref} ) at ( P_f ).</td>
</tr>
<tr>
<td>( V_s )</td>
<td>The pre-determined reference sailing speed.</td>
</tr>
</tbody>
</table>

### Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta \theta )</td>
<td>The increment of the heading angle for each current waypoint</td>
</tr>
<tr>
<td>( 2m+1 )</td>
<td>The number of successors for each current waypoint</td>
</tr>
<tr>
<td>( 2r )</td>
<td>The number of subsectors.</td>
</tr>
<tr>
<td>( \Delta D )</td>
<td>The width of search limit in distance for each local sub-sector.</td>
</tr>
<tr>
<td>( \Delta t )</td>
<td>The time interval between adjacent time stages</td>
</tr>
</tbody>
</table>

Fig. 4. The generation of new isochrones in the first half of the voyage search.

1) Starting from \( P_0 \), follow the initial headings \( \theta = \theta_{ref} \pm j \Delta \theta \) \( (j = 0, 1, \ldots, m) \) in the GC route to reach waypoints in \( \{ P_1 \} \). \( \theta_{ref} \) is chosen as \( C_{ref} \) at \( P_0 \).

2) Check the feasibility of sailing constraints:

- a) if \( V_s \) can be reached following headings \( \theta \) under the current weather condition at \( P_0 \). If not, update the actual speed \( V \) according to the engine limits.
- b) if it is land-crossing/shallow water, no-go zones, etc.

3) Navigate from \( P_0 \) with headings \( \theta \) and speed \( V \) for \( \Delta t \) hours following the GC route. Waypoints of the first isochrones \( \{ P_1 \} \), i.e., \( \{ P_{1,k}, k = 1, 2, \ldots, 2m+1 \} \) are obtained. Connect \( P_0 \) to each \( P_{1,k} \) with an edge/sub-route directed from \( P_0 \) to \( \{ P_1 \} \).

Next, navigate from each waypoint in \( \{ P_1 \} \) following the same steps as above, potential waypoints \( \{ P_{2,j}, k = 1, 2, \ldots, 2m+1, j = 1, 2, \ldots, 2m+1 \} \) can also be obtained to opt for \( \{ P_2 \} \). To prevent the exponential growth of waypoints and perform selection, the sub-sector is introduced [15], which are sub-areas distributed evenly around \( GC_{ref} \). Thus, starting from \( \{ P_1 \} \), the voyage search is carried out as follows:

1) Repeat the processes above as in Fig. 4(a) at each \((k^{th})\) waypoint \( P_{1,k} \) respectively. The reference heading \( \theta_{ref} \) is the arrival course at \( P_{1,k} \) from \( P_0 \). Each \( P_{1,k} \) leads to \( 2m+1 \) potential waypoints \( \{ P_{2,j}, j = 1, 2, \ldots, 2m+1 \} \). Sub-sectors are defined based on \( 2r+1 \) initial courses \( C_{ref} \pm k \Delta S_i \) \( (k = 0, 1, \ldots, r) \) of the GC route from \( P_0 \), drawn as grey lines in Fig. 4(b).

2) The increment \( \Delta S_i \) \( (i = 2) \), indicating the 2\textsuperscript{nd} time stage) is defined as in [15].

\[
\Delta S_i = c \Delta D / \sin(c d_i), c = \pi / (60 \times 180)
\]

where \( d_i \) \( (i = 2) \) is the expected traveled distance equal to \( i^t \Delta V_s \) \( (i = 2) \). Then, sub-sectors \( \{ S_{i,k} \} \) are given by sub-areas between GC routes with adjacent initial headings, i.e., \( [C_{ref} + (k-r-1) \Delta S_i, C_{ref} + (k-r) \Delta S_i] \) \( (i = 2, k = 1, 2, \ldots, 2r) \).

3) In each \((k^{th})\) sub-sector \( S_{i,k} \), identify the optimal waypoint \( P_{2,k} \) with the optimum cost given by the cost function \( C_p \) in (2), as blue points in Fig. 4(c).

4) Only optimal waypoints \( \{ P_{2,k}, k = 1, 2, \ldots, 2r \} \) are pre-reserved. Connect by directed edges with its predecessor in \( \{ P_1 \} \) respectively, as Fig. 4(d). The second isochrone \( \{ P_2 \} \) is obtained.

Further, based on isochrone \( \{ P_2 \} \), repeat the above steps in recursion: at the \( i^{th} \) time stage, first generate candidate waypoints \( \{ P_{i,j}, k = 1, 2, \ldots, 2r; j = 1, 2, \ldots, 2m+1 \} \), then identify the isochrone \( \{ P_i = \{ P_{i,k}, k = 1, 2, \ldots, 2r \} \) from \( \{ P_{i,j} \} \) in sub-sectors \( \{ S_{i,k} \} \). Subsequent isochrones \( \{ P_i \} \) can be obtained in sequence. Additionally, the number of sub-sectors \( 2r \) controls the capacity of waypoints in each \((i^{th})\) isochrone \( \{ P_i \} \).

The optimization objective for the proposed IPO method in this study is energy-efficient sailing. In the first half of the voyage search, the cost function \( C_p \) is defined to find the waypoint with the shortest distance to \( P_f \). The purpose is to avoid too much deviation from \( P_f \) at the early stages, since it easily leads to a long-distance route, which can be significantly fuel consuming.

### C. Isochrones of the 2\textsuperscript{nd} Half Voyage by the IPO Method

When the distance from the current isochrone \( \{ P_f \} \) to \( P_f \) is less than half of the total distance \((<0.5^*D)\), it comes to the stage where the second half of the voyage search is implemented. In this part, two problems need to be taken special care of, i.e., 1) the convergence of the route towards \( P_f, 2) \) the local optimization of the route.

In the first half of the voyage, the sub-sector \( \Delta S_i \) is defined by a monotonically increasing function with respect to \( d_i \), where \( d_i \) is the expected traveled distance in the \( i^{th} \) stage. Thus, its range will keep growing wide as propagates. If continuing in this way, the range will reach the largest near \( P_f \), and cause some candidate routes to turn sharply to reach \( P_f \) as Fig. 1.
To avoid this problem, in the latter half voyage, sub-sectors are redefined to a monotonically decreasing function. In Section II-B, replace $d_i$, the expected traveled distance from $P_0$, with $d_{1s}$, the expected distance to $P_f$, the reversed and symmetric sub-sectors are generated, which will gradually narrow when approaching $P_f$:

$$\Delta S_{1s} = c \Delta D / \sin (c \cdot d_{1s}), \quad d_{1s} = D - d_i. \quad (7)$$

where $\Delta D$ is a dimensionless coefficient adjusting the maximum width of each sub-sector. Fig. 5 shows the subsectors in both halves of the voyages. Since sub-sectors restrict the waypoints as shown in the right part of Fig. 5, the route gradually proceeds towards $P_f$.

Another problem to consider is local optimization, which is a common concern for 2D voyage optimization methods. For this IPO method, a set of feasible routes from $P_0$ to $P_f$ will be found in the end, and the optimal $R^*$ is chosen among the set. When all the candidate routes significantly overlap, especially, if the overlap starts from some early-stage waypoints, as shown Fig. 6, the result has a certain likelihood of being trapped in a local optimum. This problem can be considered as the optimality of waypoints is not well identified by the cost function. In the latter half of the voyage search, this requires special attention since the reversed sub-sectors gradually grow dense towards $P_f$, and selection within the same sub-sector involves more waypoints. Improper waypoint selection will lead to the elimination of other pre-reserved waypoints while removing routes they connect to, which may potentially be part of the global optimal solution. Thus, the cost function at the latter half of the voyage needs to be carefully formulated.

For energy-efficient sailing, there are two most significant factors influencing fuel usage, i.e., the sailing distance and the encountered weather. Therefore, it should be a mutual consideration for both their impacts in optimization. If only choosing the local optimum at each step, i.e., the waypoint reached with the least fuel, the pathfinding would easily act as a greedy search, which assumes that the integration of local optimums can lead to the global optimum. For complex problems as voyage optimization, this may only generate locally optimized results. Thus, in the 2nd half of the voyage, a predictive optimization is proposed with the scheme presented in Fig. 7.

As shown in Fig. 7, the ship model functions to estimate the sailing cost, which in this study is the fuel usage, for the associated waypoint/sub-routes. Details regarding ship models will be further introduced in Section III-B. For new candidate waypoints, in addition to considering the local/partial cost for reaching them, the future cost that will be consequently consumed after the waypoint is chosen is also predicted and considered in the waypoint optimization. The optimal waypoints are identified and appended to the grid based on the prediction along with the current cost estimation. The voyage search proceeds one step further, and the newly generated routes will be again given to the ship model, for cost estimation at the next iteration. The cost function $C_p$ is modified accordingly. Referring to (2), $C_p$ is augmented with a heuristic term $h(S)$.

$$C_p = \int_{T_0}^{T_1} j(S) \, dt + h(S) \quad (8)$$

where the first term $\int_{T_0}^{T_1} j(S) \, dt$ accumulates the fuel usage from $P_0$ to the current position (e.g., $P_{i,k}$ at the $i$th time stage), i.e., the consumed cost. The second heuristic term $h(S)$ predicts the future fuel, which is expected to be necessary to reach $P_f$ from $P_{i,k}$. It is based on the ship model and the weather forecast, by assuming the ship follows the GC route and considers dynamic weather updates at each time stage. The cost $C_p$ thereby becomes the estimation of the overall fuel consumption of an entire voyage from $P_0$ to $P_f$. Therefore, in the second half of voyage optimization, the optimization performance is improved by including the predictive cost of waypoint, evaluating both impact of weather and distance, rather than only counting the current cost partially.

Thus, the second half of the voyage search is conducted as follows. Assume the current(ith) isochrone is $\{P_i\}$:

1) At each current $(k$th) waypoint $P_{i,k}$, follow reference headings $\theta = \theta_{ref,s} \pm j \Delta \theta$ $(j = 0, 1, \ldots, m)$ to generate the candidate waypoints for the next $(i+1)$th stage. $\theta_{ref,s}$ is the initial course of the GC route from $P_{i,k}$ to $P_f$. Every $P_{i,k}$ leads to $2m+1$ new candidate points $\{P_{i+1,j} \mid j = 1, 2, \ldots, 2m+1\}$ as Fig. 8(a).

2) The reversed sub-sectors are indicated by $2r+i$ GC routes with arrival courses $C_{\text{arr}(r,f)} \pm k \Delta S_{(r+1)s}$ $(k = 0, 1, \ldots, r)$ at $P_f$, as grey lines in Fig. 8(b).
The increment $\Delta S_{i+1,k}$ is calculated by (7), and the sub-sectors $\{S_{i+1}\}$ are sub-areas between adjacent arrival headings at $P_f$, i.e., $[C_{i+1}(r) + (k-\alpha-1) \cdot \Delta S_{i+1}, C_{i+1}(r) + (k-\alpha) \cdot \Delta S_{i+1}], (k = 1, 2, \ldots, 2r).

3) In each $(k)$ sub-sector $S_{i+1,k}$, identify the optimal point $P_{i+1,k}$ with the optimum cost by the augmented cost function $C_p$ in (8), as blue points in Fig. 8(c).

4) Connect optimal points $\{P_{i+1,k}, k = 1, 2, \ldots, 2r\}$ by directed edges with its predecessor in $\{P_i\}$ respectively, as Fig. 8(d).

The next $(i+1)$ isochrone $\{P_{i+1}\}$ is obtained.

Through estimating fuel cost, the weather impact is considered to either avoid severe sea conditions or utilize the wind and ocean current. Further, from $\{P_{i+1}\}$, repeat the above steps, and subsequent isochrones $\{P_{i+2}\}, \ldots, \{P_n\}$ in the 2nd half of the voyage are obtained iteratively.

By employing the reversed subsectors, the width of sub-sectors narrows down rapidly in the area around $P_f$, and sub-sectors becomes compact. Thus, when the distance from the latest isochrone $\{P_{i+1}\}$ to $P_f$ is less than $i^2 \Delta t$, sailing, $\Delta \theta$ would be decreased to $\alpha * \Delta \theta$. In this study, $i$ is chosen as 3, $\alpha = 10\%$. And finally, when $d_p$ is less than $\Delta^2 V_3$, connect all waypoints to $P_f$ directly through the GC route.

A feasible route set $\{R\}$ is obtained, and all candidate routes in $\{R\}$ process approximately the same ETA. For each sub-route, the fuel consumption is based on the local weather conditions at the first waypoint. And the overall fuel consumption is the accumulative cost of all sub-routes, i.e., a series of sub-routes from $P_0$ to $\{P_1\}$, $\{P_1\}$ to $\{P_2\}$, ..., $\{P_n\}$ to $P_f$, as shown in Fig. 9. The optimal $R^*$ will be chosen as the route with minimum accumulated fuel.

It needs to be stated that, in this study voyage division is chosen to be half and half. It is based on the reason that the search area cannot converge too early, as this may result in local minimization in the initial stages. Meanwhile, excessive expansion may lead to difficulties in smooth convergence in later stages.

**Fig. 8.** The generation of new isochrones in the 2nd half of the voyage search.

**Fig. 9.** Examples of the optimal route $R^*$ which consists of a series of discretized sub-routes.

**Fig. 10.** Parameters that require specification in the proposed IPO method for deployment.

### D. Parameters Setting in the IPO Method

To deploy the IPO method, five parameters shown in Fig. 10 should be specified for initialization. These parameters critically influence the algorithm’s performance as they configure dynamic grid generation. The value of parameters varies depending on the specific case, and appropriate values can be obtained for each case based on the following introduction.

- $\Delta \theta$: The increment of heading angles between two adjacent sub-routes, for each current waypoint.
- $m$: Control the number of candidate successors for each current waypoint.
- $\Delta t$: Traveling time between two adjacent stages.
- $\Delta D$: Control the width of a single subsector.

It influences the number of candidate waypoints at each stage. If $m$ is small, the successors might not sufficiently cover the upcoming search area; and if $m$ is large, great computational effort will be required. Generally, it could be chosen around 10 to 20.

- $\Delta t$: Traveling time between two adjacent time stages.

It controls the looseness of the search grid along the direction toward the destination and represents the step size of the iterative pathfinding process. Large $\Delta t$ may lead to divergence towards $P_f$, and small $\Delta t$ might cause the result trapping in local optima at the early stages. Typically, the overall time stages can range around 20-30, and $\Delta t$ can be chosen based on the ETA[h] divided by the number of time stages.

- $\Delta D$: The width of a single subsector.

It influences the width limit of the search grid along the voyage. A small $\Delta D$ gives dense subsectors further forming a narrow and possibly insufficient search area, while a large $\Delta D$ will cover a more extensive space,
making the search process more time-consuming. This is a dimensionless coefficient and can be chosen around 5.

- **r**: Control the number of subsectors.

Since each subsector preserves one optimal waypoint at each stage, increasing the value of r means the number of waypoints in each isochrone is increased. Similar to m, a large r allows more potential candidate waypoints/routes, that may improve the performance of the proposed algorithm but also bring computation burden. Generally, a value of 10 to 20 is sufficient.

### III. Case Study and Measured Voyages

A case study is carried out to validate the performance of the proposed IPO algorithm in energy-efficient voyage optimization. A chemical tanker sailing in the North Atlantic with full-scale measurement is used in the case study. The main particulars of this ship are listed in Table II. Its sailing operation is guided by a conventional weather routing system installed onboard, combined with the ship crew’s experience. Its actual routes have been carefully planned and this case study ship has a certain level of capability in voyage optimization.

#### A. Case Study Ship Voyages Used for Method Verification

Six measured voyages of this ship that took place in 2015 and 2016, are used in this case study as shown in Fig. 11. These voyages comprehensively include eastbound and westbound cases during winter and summer, present various environmental conditions, i.e., sailings in calm, moderate, and severe sea states. The actual voyages and their associated operational data are extracted for comparisons, to demonstrate the proposed algorithms’ capability in voyage optimization.

To show the capability of the proposed algorithm in energy-efficient voyage optimization, the algorithm also needs weather data to describe the sailing environments. These sea environmental data include encountered wind, wave, and current for the estimation of the ship performance model. All the related met-ocean parameters, i.e., the meteorological and oceanographic data were retrieved in 2023, which are historical data from 2015 and 2016 consistent with actual voyages. Wind (speed and direction) with wave (height, direction, and period) are extracted from ECMWF ERA-5 (2023) dataset, and current is obtained from http://marine.copernicus.eu (Copernicus 2023) server. Besides, a ship performance model is essential for the algorithm to estimate the ship’s behaviors at sea, i.e., the speed-power relationship in connection with the environmental conditions. The model is introduced in the following Section III-B.

#### B. Physics-Informed Machine Learning Speed-Power Performance Model

The ship performance model used in this study is based on the work by Lang et al. [24]. For algorithms such as the IPO method, the outcome from the cost function has an essential impact on the optimization result since it supports decision-making. The ship model, which processes a great capability to accurately predict the ship’s behavior at sea, could guarantee that the result of voyage optimization is reliable for real operation. Otherwise, the voyage optimization result can be distorted and inaccurate if implemented in the actual sailing. The ship model employed in this study is constructed using a novel physics-informed machine learning approach. It is a state-of-the-art grey-box model (GBM) that can predict the engine power according to the sailing speed of ocean-crossing ships. For models built by conventional approaches, such as empirical and semi-empirical formulations (named white-box models, i.e., WBM), their accuracies depend on the assumptions and uncertainties implicit in the physical model. And for the data-driven regression/machine learning models (black-box models, i.e., BBMs), their capabilities in interpretability and extrapolation are comparably poor for unseen scenarios beyond the measurement, further leading to significantly inaccurate outcomes. GBMs are developed through the combination of physical principles from WBM and big ship data inferences from BBMs. Therefore, the grey-box ship performance model employed in this study can provide higher accuracy than WBM, while possessing great interpretability and extrapolation capability, avoiding unreasonable results as BBMs.

In this grey-box ship model, a BBM and a physics-informed neural network (PINNs) model are integrated, to describe the relationship between the actual sailing speed over ground $V_g$ and the ship’s power $P_D$. First, the ship speed reduction $\Delta V$ of the specified speed over ground $V_g$ is estimated through a BBM, based on ship operational and sea environmental data. Its associated speed through water $V_w$ is obtained, as $V_w = V_g + \Delta V$. Then, based on this $V_w$ output from the BBM, the required propulsion power $P_D$ can be predicted through the PINNs model. An accurate engine power $P_D$ needed to reach the expected speed $V_g$ is acquired, providing precise predictions in the fuel cost for the decision-making of the algorithm.

Loading conditions are also considered by the ship model, and in this study, the model assumes that the ship is fully

<table>
<thead>
<tr>
<th><strong>TABLE II</strong></th>
<th><strong>PRINCIPAL PARTICULARS OF THE CHEMICAL TANKER SHIP</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Length $L_m$</strong></td>
<td>178.4 m</td>
</tr>
<tr>
<td><strong>Length $L_P$</strong></td>
<td>174.8 m</td>
</tr>
<tr>
<td><strong>Beam $B$</strong></td>
<td>32.2 m</td>
</tr>
<tr>
<td><strong>Depth</strong></td>
<td>17.0 m</td>
</tr>
</tbody>
</table>

Fig. 11. Six actual voyage cases used in this paper for the validation of the proposed IPO voyage optimization method.
loaded with the same draft during the voyage. To further exclude the noises/errors in the measurement and difference from the real operation conditions, only the sailing time and positions (longitude and latitude) of the case study ship are extracted into the ship model for calculating power. For the proposed IPO algorithm, fuel consumption is estimated in the same way. Moreover, provided that the specific speed-power performance model and a reference route are available, the proposed IPO method can also optimize voyages for ship types and trades other than the chemical tanker initially used.

C. Other Voyage Optimization Algorithms as Comparison

Besides the real voyage cases, four voyage optimization methods, as listed in Table III, are also used to compare and demonstrate the effectiveness of the proposed IPO method.

GC routing shows a traditional manual navigation method used by ship crews. It takes the shortest GC route as the fixed route, discretizes it into several time stages based on the ETA, and determines the interval speed according to the local sea conditions to achieve punctuality. Since avoiding unfavored weather also causes detours that may consume more energy, if the encountered sea conditions are not harsh, the result of this shortest-distance routing is generally acceptable.

The Dijkstra algorithm is a well-developed algorithm that has been used widely in today’s voyage optimization applications. Thus, it is taken into comparison in two ways. The 2D Dijkstra algorithm (2DDA) employs constant sailing speed in sub-routes along the whole voyage unless extreme weather is encountered. The 3D Dijkstra algorithm (3DDA) involves speed as an additional control variable, thereby allowing for speed optimization in voyage planning. Compared with 2D algorithms, 3DDA possesses an enhanced capability in voyage optimization due to an extra dimension.

The four approaches listed in Table III are chosen to demonstrate the effectiveness, efficiency, and practical applicability respectively of the proposed method, as an algorithm for real-time voyage optimization. The GC routing is a maritime navigation strategy that is commonly used in industrial practice. Thus, it can verify the potential practicality of the proposed method to be used in real operation. MI is a representative variant based on the traditional Isochrone method, and an important reference where the proposed IPO method derived from. Comparison with MI can further demonstrate the improvement of IPO. Moreover, 2DDA is a widely used voyage optimization method in practice other than the Isochrone types, as reviewed in Section I, well-known for its optimization capability and generalization. And 3DDA is a recent enhancement of 2DDA where the speed optimization is further included. Their results could present the optimization capability of the proposed IPO from a more general level, regardless of the type of the method, for both its effectiveness and efficiency.

IV. RESULTS OF THE OPTIMIZATION FOR VOYAGE CASES

In this study, the proposed IPO algorithm is validated using six voyage cases. The result is evaluated for both optimization effectiveness and computational speed, to demonstrate the proposed algorithm as an online voyage optimization algorithm with robustness and fast response. Computational efficiency is presented and compared in terms of runtime. It is influenced by the method’s complexity when operating in the same environment, but it also slightly varies in each execution for the same algorithm. For each case, the runtime values are used solely for comparison between various methods.

For different voyage cases, each method needs to specify a proper grid to provide a good performance. MI employs the same parameters as the proposed IPO method. Due to involuntary speed reduction, GC routing must test various speeds within a defined range to accurately meet the ETA; the number of attempts is chosen to match the number of candidate routes of the IPO method. 2DDA and 3DDA need to discretize the sailing area with a static grid and enumerate the optimal route within this grid. Their grids are defined to have an equal number of time stages, while having the same number of waypoints in each time stage as the IPO method.

A. Westbound Voyage Optimizations

In the North Atlantic, since storms associated with the prevailing westerlies are normally moving from west to east, ships confront more head-on waves along westbound voyages. It is often more challenging and fuel consuming to navigate west, and requires careful planning and execution to enhance efficiency and safety. Three westbound voyage cases, one during winter and two in summer, are investigated in this section. The optimization results are summarized in Table IV, showing ETA, fuel consumption, sailing distance, average speed, and runtime respectively for each voyage. The optimized routes generated by each method are presented in Fig. 12.

For these three westbound cases, their actual voyages encountered calm and mild sea weather environments, and the significant wave heights \( H_s \) have not exceeded 4 meters. From Fig. 12, the actual routes do not deviate much from the GC route, therefore the sailing distances are also close to the shortest achievable length. Both the weather and distance factors indicate that the fuel consumption of the actual routes is not high, and they have been well-planned owing to the onboard navigation system assisted by the ship crews. However, the optimization result could still present notable differences in fuel consumption, compared with the actual routes.

For Voyage 2015.07.21, 2DDA and 3DDA show the least consumptions, however, with 7- and 2-hour arrival delays. Considering punctuality, IPO provides the most fuel reduction at 7.3%. For Voyage 2016.07.19, IPO and 3DDA result closely at 3.0% as the most reductions with accurate ETAs. And
for Voyage 2016.11.08, IPO and 2DDA give the most fuel savings at around 3.0%. However, 2DDA again fails to meet the ETA. In general, for three voyages, IPO provides the most energy-efficient route with on-time arrivals in voyage optimization. Compared with the more powerful method 3DDA, it could result similarly in fuel cost. However, IPO is around 90 times faster than 3DDA, and 2 times faster than 2DDA in terms of the runtime. 2DDA can suggest significant fuel savings, however, it is hard to ensure the ETA, and it tends to suggest comparatively long sailing distances. GC routing presents no significant improvements in energy efficiency based on the actual routes, and it is worth noticing that MI also does not perform promisingly. It shows similar fuel consumption as the actual routes, and abrupt turnings can be noticed in its voyage around the destination for all three cases as in Fig. 12.

To present more details of the optimization process, two ordinary and calm sea sailing case, i.e., Voyage 2015.07.21 and Voyage 2016.11.08, are shown in Fig. 13 and Fig. 14 with encountered $H_s$ and engine shaft power during the process. In Voyage 2015.07.21, the engine shaft power and encountered $H_s$ are given in Fig. 13. Its sea environment during the voyage is ordinary with the highest $H_s$ around 3.5 meters. The actual route follows the Great Circle routes to save distance, thus, its actual fuel cost is also similar to GC routing. Other optimized routes head slightly northern, and it is seen from Fig. 13 that the reason is to avoid the high waves in the vicinity of the GC route. Specifically, 2DDA shows the greatest deviation in routes, and its encountered $H_s$ is mostly the lowest. However, the scheduled ETA for this voyage is tight, and such a detour in 2DDA results in a nearly 7-hour delay in arrival.
A similar situation appears in the route of MI, as its route also avoids waves by sailing north. Nevertheless, it shows the consideration of ETA by employing the maximum engine power during the voyage to catch the time. When reaching the destination, its route turns sharply by the end due to the lack of route convergence and smooth process. The outcome of IPO achieves improvement based on MI, since it is not only shorter in distance with accurate ETA, but also comes across calmer sea environments. Thus, it suggests a route with the most fuel savings at 7.3% adhering to the planned time.

Another calm sea case Voyage 2016.11.08 is also given in Fig. 14. For this sailing, the environment is even peaceful with $H_s$ less than 2.6 meters. For all the routes, the most apparent divergences appear in the area around longitude $-30^\circ$W. The dynamic voyage evolution process within this area is presented as a contour plot of $H_s$ in Fig. 15. High waves appear in the upper north around the GC route, and they gradually diminish towards the south. Consequently, all the optimized routes diverge slightly towards the south to mitigate this weather impact. The southernmost route 2DDA has the lowest fuel consumption, and IPO seconds to it. 2DDA heads toward the calm waves and continues to move in the steadiest sea states. It is similar to the actual route, which also altered its courses twice, thereby sailing in very calm sea environments. However, their route adjustments either result in a long total distance and late ETA as 2DDA, or a higher average speed as in the actual route, leading to higher fuel consumption. In addition, MI’s route prioritizes short distances over the environmental impact, and also results in a higher fuel cost. Besides, the sharp turn appears again by the end of the route of MI. Finally, IPO method better balances the objectives of energy efficiency and ETA, providing a 3% fuel reduction in this very calm sea sailing situation. In this case, it can be noticed that IPO suggests a better result than 3DDA. This can be attributed to the limited resolution of the 3DDA’s discretized static grid. The dynamic grid of IPO allows for free exploration within the search area, therefore can find the potential optimal solution that may be excluded by the discretized static grid.

### B. Eastbound Voyage Optimizations

Eastbound voyages may contend with the prevailing westerlies which can provide speed boosts for ships, but may also be challenging in navigation due to stronger winds and severe sea states. The North Atlantic is also known for its seasonal variability, with winter months bringing more rough weather conditions and increased chances of encountering storms. In this section, the voyage optimization for one winter and two summer eastbound voyage cases is presented. The results are given in Table V, with optimized routes shown in Fig. 16.

For these three eastbound voyages, their actual routes have encountered diverse sailing environments. The winter voyage presents a severe sea sailing case where the ship confronted storms during the process, and $H_s$ reaches more than 9 meters at the highest. The other two summer voyages are rather calm and moderate, with the highest wave around 5 meters as the westbound voyages above. For such eastbound sailings, voyage optimization particularly requires special attention and inefficient results can cause serious consequences. The actual ship has been thoughtfully navigated and the routes in Fig. 16 could demonstrate its attention to such weather from the heading changes. However, the optimization result presented can still show considerable improvements based on the actual routes. Taking the punctuality requirement into account, for the winter Voyage 2016.02.29, IPO and 3DDA could at most help to reduce 8.6% and 11.1%. For Voyage 2015.05.23 and 2016.05.23, the sea states are not severe thus the optimization results for each method is relatively close. Among these, the IPO and 3DDA give the lowest fuel cost at around 3%. Thus, for three eastbound voyages in this section, IPO demonstrates an improved performance compared with MI, and the best voyage optimization capability in the 2D methods which can be comparable to 3DDA. As for computational efficiency, Voyage 2016.02.29 involves dramatic weather changes, and a denser grid is employed to enable a wider search space, thus the computation load is heavier, causing a longer execution time for every method than other cases. But in general, IPO is 80 to 100 times faster than 3DDA, and 2 times faster than 2DDA.

More details of this winter Voyage 2016.02.29 are presented in Fig. 17 and Fig. 18 to show its severe sea sailing process.
During this voyage, two storms passed through the sea areas near the ship’s navigation zone. It can be recognized from the two peaks in Fig. 17 that the first storm occurred around the longitude $-35^\circ$W, and the second one appeared around $-25^\circ$W. To avoid the first storm, the optimized routes suggest two directions and start to diverge as shown in Fig. 16. The route of MI deviates from the other routes and recommends going up north, while the others including the actual route choose to advance south. Consequently, MI confronts rather lower waves by intention during the first storm, while the other routes are influenced more greatly. However, when it comes to the second storm, the other routes effectively bypassed this storm’s most impact, and could maintain their operations with lower engine power from the upper plot in Fig. 17. But MI employs higher power for the first storm avoidance, and then requires more shaft power again during the second storm. Finally, its route turns greatly towards destinations to compensate for its detour, which for the real operation is impractical and also causes time delay.

The dynamic progress of the voyage encountering the second storm is presented as contour plots of $H_s$ in Fig. 18. The storm starts to be observed in the south and in front of the four routes that sailed south first. It moved northward subsequently and gradually intensified, coinciding precisely with the route of MI as well as the GC routing. And when the other four routes, i.e., IPO, 2DDA, 3DDA, and the actual route reached the vicinity, they narrowly missed the storm’s center. Thus, the MI and GC routes are not as efficient as the others in this case. In addition, for the actual route, its maneuvers led to an increased distance and higher sailing speeds, and it was significantly affected by the first storm. Although it considers evading both the two storms, it does not offer fuel savings as substantial as achieved by 3DDA and IPO.

The same issue of MI also arises in the other summer voyage 2016.05.23. During this voyage, the most adverse sea conditions appear near longitude $-25^\circ$W along the GC route. And the weather confronted by each route within this area is the major reason leading to different final fuel costs. Although the wave reaches around 5 meters, the overall encountering duration is not long thus the final consumption is rather close for each method. The actual route basically follows the GC route; however, it varies in speed thus slightly fuel saving than GC routing. MI guides the route to avoid the middle area but fails in smooth convergence, again resulting in arrival delay due to the increased length. IPO, 2DDA, and 3DDA closely provide the most fuel reduction. And for another summer voyage 2015.05.23, the ship sailed at the most peaceful sea states, where $H_s$ is below 2 meters along the way. All optimization methods suggest routes that mostly overlapped with the GC routes, opting for the shortest distance sailing in this sea condition. The optimization result given by each method is similar as well. For these two calm and ordinary sea sailing cases, the grid does not need to cover a wide range to search for more flexibility and variations. Thus, the computational loads of voyage optimization for such cases are also reduced compared to the winter voyage 2016.02.29.

C. Discussions

Six voyages measured at diverse sea environments are used as case studies to validate the proposed IPO method for ship
voyage optimization. The results show that the IPO method possesses an enhanced capability and effectiveness than the original Isochrone method (MI). For example, routes generated by the MI method would easily present abrupt turns and increased distance as shown in Fig. 1, due to the lack of the proper route convergence. Thus, when encountering harsh weather conditions in the middle of a voyage, MI attempts to avoid them but cannot smoothly converge back to the destination.

When comparing with other well-developed voyage optimization algorithms, the routes optimized by the 2DDA method show good energy efficiency for all cases, however, similar route detours and arrival delays are also observed. Especially, the 2DDA method is in general hard to ensure and specify the ETA adhering to the schedule, due to its static grid specifications. GC routing is used as a typical routing method in operations, and this method presents similar results as the actual routes in most voyage cases. The GC routing does not consider weather impact, but more focuses on a single objective optimization, i.e., generating the shortest distance ship voyages. Therefore, its results only show more effectiveness in calm sea sailing cases.

Finally, the proposed IPO method balances both the impact of weather and distance on fuel consumption. The GC routing method has demonstrated that considering only the shortest distance factor is not sufficient for energy efficient optimization, since the weather conditions can influence the fuel cost significantly even in ordinary sailing environments. However, counting the current fuel cost would easily behave as a greedy search and lead to local optimums in complex optimization problems. The predictive optimization procedure in the proposed IPO method can effectively avoid local optimization by estimating the future cost, which is caused consequently by the current move, yielding good energy efficiency with an accurate time of arrival.

As for computation efficiency, unlike many static graph search algorithms, the proposed IPO method can provide adequately effective optimization results without the need for high grid resolution, thereby avoiding heavy computational efforts. It is an iterative optimization method, i.e., in each step, it dynamically explores and updates the optimal search direction to proceed toward the destination. These search directions are identified through the cost function by evaluating the optimal waypoints in the current isochrone. And the proceeding step size is specified by its pre-defined parameter, e.g., $\Delta r$. For iterative optimization methods, smaller step sizes do not always benefit the optimization performance. Instead, it may also lead to being trapped in local optimums and divergence. Thus, choosing the suitable parameters for the IPO method could just bring both moderated computation load and satisfying optimization performance. By further facilitating parallel computing, its runtime can be short enough to allow for real-time usage.

For static grid algorithms such as Dijkstra method, they discretize the overall search space and enumerate the optimal solution within this static grid. Thus, enhancing its grid resolution always contributes to approximate near-globally optimized solutions. And to achieve a satisfying result, it requires a grid with a certain resolution while also covering enough range of the sailing area, which inevitably brings a burden to the computation. According to the results in the above sections, the runtime of 2DDA is at least two times longer than IPO in general for all voyage cases.

V. Conclusion

The paper proposed an Isochrone-based predictive method (IPO) for energy-efficient real-time voyage optimization. Six voyages from full-scale measurements of a chemical tanker are used as the case study to verify the effectiveness and efficiency of the proposed IPO method. The case study voyages include diverse sea conditions and time for the verification purposes. Four well-recognized ship voyage optimization algorithms are also used for comparison. Some conclusions can be drawn from the verification analysis as follows:

1) Firstly, the impractical routes with abrupt turns that appear in the traditional Isochrone method and graph search methods are efficiently resolved by the IPO method. Through the voyage division with different waypoint grid partition strategies, the optimized route could converge smoothly toward the destination.
2) By refining the cost functions to include a predictive optimization process, the result shows that the IPO method can give energy-efficient routes in various sailing environments while ensuring an accurate ETA. Compared with the actual route, the proposed IPO method could reduce 9% for severe sea sailing, at least 3% for calm sea sailing, and on average 5% fuel for all cases. And the result is close to the 3DDA algorithm which includes additional speed variations in voyage optimizations.
3) The IPO method is a deterministic approach that ensures obtaining a result, if one exists. Moreover, it requires a few computational resources and allows for parallel computing, which enables online usage and real-time adaptation to the changing environments. The overall process of IPO method takes around 40 seconds and can be further enhanced by utilizing advanced industrial computers.
4) The constant speed employed in the method avoids frequent adjustments in engine and maneuvering, which is preferred by seagoing vessels and ensures the method’s practicality.

A. Outlook of Further Development

For a given voyage, the proposed IPO method allows to specify a required ETA and suggests an energy-efficient route in voyage planning. Further, the pre-planned voyage can be fast updated during the execution. Thus, the proposed IPO method can be both effective in voyage planning before departure, and execution during the voyage for real-time adjustment, to mitigate the divergence due to uncertainty, and adapt to dynamic changes.

However, the performance of the IPO method is sensitive to its parameter setting, and the parameter’s value varies for each individual voyage. Thus, systematic analysis to identify
the proper values more efficiently for each voyage can further be considered for the IPO method. Besides, the constant speed setting can also be replaced by the most direct input of engine, e.g., RPM or power, to further enhance the practicality of the method in operation. Also, some parameters indicating search step sizes, e.g., \( \Delta t \) and \( \Delta \theta \), can be changed from pre-determined constants to variables to achieve adaptive optimizations, which may further improve the optimization performance in the future works.

To assist its decision-making, this paper adopts state-of-art machine learning modeling techniques in ship performance prediction. Further, when integrating with more components for a comprehensive navigation system in applications, the optimization problems can grow large-scale and more challenging. The algorithm's capabilities in e.g., big data classification and processing, online execution, dynamic and adaptive optimization will be beneficial. Thus, in addition to the proposed method, advanced optimization algorithms can also be effective to both show enhanced optimization results and achieve fast computational speeds in such cases. The advanced algorithms, e.g., the use of hybrid heuristics, metaheuristic [51], and hyper heuristics [50], adaptive and self-adaptive algorithms [5], or machine learning algorithms have achieved promising results in various domains such as transportation [10], online scheduling [11], multi-objective optimization [46], etc. In future research, their potential within voyage optimization is also worth investigating and comparing with the proposed method, or integrated in the proposed method for further enhancement.

REFERENCES


