Membrane distillation at the water-energy nexus: limits, opportunities, and challenges

Akshay Deshmukh, Chanhee Boo, Vasiliki Karanikola, Shihong Lin, Anthony P. Straub, Tiezheng Tong, David M. Warsinger and Menachem Elimelech

Energy-efficient desalination and water treatment technologies play a critical role in augmenting freshwater resources without placing an excessive strain on limited energy supplies. By desalinating high-salinity waters using low-grade or waste heat, membrane distillation (MD) has the potential to increase sustainable water production, a key facet of the water-energy nexus. However, despite advances in membrane technology and the development of novel process configurations, the viability of MD as an energy-efficient desalination process remains uncertain. In this review, we examine the key challenges facing MD and explore the opportunities for improving MD membranes and system design. We begin by exploring how the energy efficiency of MD is limited by the thermal separation of water and dissolved solutes. We then assess the performance of MD relative to other desalination processes, including reverse osmosis and multi-effect distillation, comparing various metrics including energy efficiency, energy quality, and susceptibility to fouling. By analyzing the impact of membrane properties on the energy efficiency of an MD desalination system, we demonstrate the importance of maximizing porosity and optimizing thickness to minimize energy consumption. We also show how ineffective heat recovery and temperature polarization can limit the energetic performance of MD and how novel process variants seek to reduce these inefficiencies. Fouling, scaling, and wetting can have a significant detrimental impact on MD performance. We outline how novel membrane designs with special surface wettability and process-based fouling control strategies may bolster membrane and process robustness. Finally, we explore applications where MD may be able to outperform established desalination technologies, increasing water production without consuming large amounts of electrical or high-grade thermal energy. We conclude by discussing the outlook for MD desalination, highlighting challenges and key areas for future research and development.

Broader context

Tackling water scarcity, which currently affects 40% of the global population, is one of the greatest technological challenges of the 21st century. Water desalination technologies, particularly those that treat highly saline or contaminated waters, are energy intensive. Minimizing the energy consumption of desalination processes is especially important given the reciprocal interdependence between energy generation and water production, termed the water-energy nexus. Membrane distillation (MD) is an emerging thermal desalination process that is capable of treating high-salinity waters, including industrial wastewater from unconventional hydrocarbon extraction sites and brines from desalination plants. Although the energy efficiency of MD is low compared to reverse osmosis (RO), its ability to utilize low-grade heat sources, such as waste heat from power and industrial plants, is advantageous. MD desalination systems also have the potential to be more compact and versatile than traditional thermal desalination processes, such as multi-effect distillation and multi-stage flash. In this article, we provide a critical review of MD, exploring the key factors determining the energetic performance and operational efficacy of MD desalination and discussing the limitations that have hindered its large-scale deployment.
1. Introduction

The United Nations estimates that 1.8 billion people currently live in countries experiencing absolute water scarcity.\(^1\) Driven by population growth and industrialization, current trends in water consumption are forecast to lead to a 40% shortfall in freshwater supplies by 2030.\(^1-^4\) Desalination technologies can help alleviate water stress by extracting freshwater from a range of saline or contaminated sources including seawater, brackish groundwater, and wastewater.\(^5,^6\) However, water desalination processes are often energy intensive, consuming large amounts of electrical energy or high-temperature heat.\(^5,^7\) Minimizing the energy consumption of desalination and treatment processes is particularly important given the interdependence between water production and energy generation, the water-energy nexus. Water desalination technologies also play an important role in reducing the environmental impacts associated with the inland disposal of contaminated brines, such as produced water from unconventional hydrocarbon extraction or brines from the desalination of brackish water.\(^8,^9\)

Currently, reverse osmosis (RO), which is the most energy efficient desalination technology, accounts for over 60% of global desalination capacity, while thermal processes, such as multi-effect distillation (MED) and multi-stage flash (MSF), account for 34%.\(^10,^11\) Membrane distillation (MD), which can harvest low-grade or waste heat to desalinate and treat high-salinity waters, is a potentially promising process at the water-energy nexus. In MD, saline feed water is heated before being contacted with cool permeate stream across an air-filled hydrophobic membrane.\(^12\) Water evaporates at the membrane-solution interface on the warmer feed-side of the membrane before diffusing through the air trapped in the membrane pores and condensing at the cooler permeate-side membrane-solution interface.\(^13,^14\) Non-volatile solutes, which are completely rejected at the vapor–liquid interface, remain in the saline feed stream. Although MD is inherently less energy efficient than RO, its ability to utilize low-grade thermal energy rather than electricity to treat high-salinity brines, which cannot be treated by RO, is highly advantageous. In areas where low-grade or renewable heat sources, including waste-heat from industrial sources, are available, MD can be a promising technology for desalination.

Akshay Deshmukh is a PhD Candidate in the Elimelech Research Group, part of the Department of Chemical and Environmental Engineering at Yale University. Prior to Yale, he completed his BA and MEng in Chemical Engineering at the University of Cambridge. His research interests center on understanding processes at the water-energy nexus with research activities focusing on the fundamental modelling of reverse osmosis and membrane distillation desalination.

Chanhee Boo is a PhD candidate in the Department of Chemical and Environmental Engineering at Yale University. He received his BSc from Korea University, Seoul, South Korea and MSc degree from Yale. He is interested in technologies for sustainable water and energy production, and his research activities focuses on engineering material design for the development of advanced desalination and water purification membranes.

Shihong Lin is an Assistant Professor in the Department of Civil and Environmental Engineering at Vanderbilt University. He received his BSc from Harbin Institute of Technology, PhD from Duke University, and postdoctoral training from Yale University. His research interests include water treatment and desalination technologies as well as environmental interfacial science and engineering, with the mission to promote sustainability at the food-water-energy nexus.

Anthony Straub is a Postdoctoral Research Fellow in the Department of Materials Science and Engineering at the Massachusetts Institute of Technology. He received his BSc from the University of Illinois at Urbana-Champaign and PhD from Yale University in Chemical and Environmental Engineering. His research focuses on developing technologies that address challenges at the intersection of water and energy, and he is broadly interested in applications related to desalination, advanced water treatment, and sustainable power generation.
processes or solar thermal collectors, are readily available, MD has the potential to provide sustainable water treatment.\textsuperscript{15–17}

Despite substantial work on MD over the past 30 years, there still exists uncertainty regarding its overall viability and efficacy.\textsuperscript{12} The energy efficiency of MD, which is low compared to other thermal desalination processes such as MED and MSF, has proved a persistent drawback. However, in recent years novel MD membrane designs and process configurations have reduced transmembrane heat loss and increased the proportion of heat recovered from the permeate stream, leading to reductions in energy consumption.\textsuperscript{18–20} Studies have also examined the sources of inefficiencies in the MD process, quantifying the impact of membrane properties and process parameters on MD performance.\textsuperscript{13,21–25} In addition, novel MD membranes with special surface wettability have been developed to tackle issues related to membrane wetting, in which the saline feed water enters the membrane pores compromising the separation process, and fouling.\textsuperscript{30–34} Novel membrane modifications are also required to increase the rejection of volatile contaminants, which can be preferentially transported across an MD membrane and concentrated in the permeate.\textsuperscript{35,36}

In this critical review, we focus on the energy efficiency and robustness of MD desalination. We begin by discussing the inherent limitations to the energetic performance of MD as a thermal separation process. Drawing on previous literature and industrial data, we compare the efficiency of MD with MED and MSF, two widely used thermal desalination processes. Utilizing numerical models from existing literature, we quantify and explore the potential for improving efficiency through improved membrane design and novel process configurations. In particular, we examine the impact of membrane porosity and thickness on entropic losses in MD and discuss how heat recovery and temperature polarization affect the energy efficiency of MD. We then proceed to investigate strategies for improving the robustness of MD membranes by mitigating and controlling fouling and scaling. Finally, we outline the key developments in membrane and process design required to achieve the transformative increases in energy efficiency and robustness that will enable MD to fulfill its potential to desalinate high-salinity brines using low-grade heat.

2. Is MD an energy efficient process?

2.1 Inherent energy efficiency limitations

From a thermodynamic perspective, desalination technologies can be divided into two broad categories. In the first category, which includes reverse osmosis (RO, Fig. 1B),\textsuperscript{37} electrodialysis (ED, Fig. 1C),\textsuperscript{38} and capacitive deionization (CDI), water remains in the liquid phase throughout the separation process.\textsuperscript{39} Technologies in the second category, which includes multi-effect distillation (MED), multi-stage flash (MSF), mechanical vapor compression (MVC), and MD (Fig. 1A), exploit a phase change to separate water from non-volatile contaminants. Hybrid separation processes such as forward osmosis (FO), may use work-based processes, such as RO, or thermal processes, including low-temperature distillation or MVC, to separate pure water from the diluted draw solution.\textsuperscript{40,41} While RO, ED, and CDI are based on different separation mechanisms and are not equally efficient, the energy consumed by these processes is primarily utilized to separate water and salt, in addition to overcoming parasitic resistances to mass and charge transfer. By contrast, in thermal desalination processes, including MD, a large amount of energy is consumed by the evaporation of water.\textsuperscript{15} Depending on the salinity and temperature, the enthalpy of vaporization, or latent heat ($\Delta H_{\text{vap}}$), varies slightly around 2400 \text{kJ kg}^{-1} (667 \text{kJ m}^{-3}).\textsuperscript{41} This value is two to three orders of magnitude higher than the specific Gibbs free energy of separation ($\Delta G_{\text{sep}}$) for seawater, which is 0.76 \text{kJ m}^{-3} for zero recovery and 1.06 \text{kJ m}^{-3} for 50% recovery with a feed salinity of around 35 g kg$^{-1}$ (35,000 ppm).\textsuperscript{5} At the molecular level, this latent heat reflects the energy required for breaking the hydrogen bonds between water molecules, which is immensely more intense than that for moving the salt and water molecules against their inherent propensity to mix and become homogeneous.

Tiezheng Tong is an assistant professor at the Department of Civil and Environmental Engineering at Colorado State University. He received his BSc from Beijing Normal University, MSc from Tsinghua University, and PhD from Northwestern University in Environmental Engineering. He was a postdoc researcher at Yale University under supervision of Prof Menachem Elimelech. His research interests focus on the design and development of novel membrane materials and processes for sustainable water supply, environmental applications and implications of nanotechnology, and resource recovery from wastewater treatment processes.

Menachem Elimelech is the Roberto Goizueta Professor at the Department of Chemical and Environmental Engineering at Yale University. His research interests include emerging membrane-based technologies at the water-energy nexus, materials for next-generation desalination and water purification membranes, and environmental applications of nanomaterials. Professor Elimelech is a Clarivate Analytics (formerly Thomson Reuters) Highly Cited Researcher in Chemistry and Environment/Ecology. He is a member of the US National Academy of Engineering and a foreign member of the Chinese Academy of Engineering.
State-of-the-art seawater RO consumes about two to three times $\Delta G_{\text{sep}}$ to produce one cubic meter of deionized water.\(^5\) Without recovering latent heat, a thermal process would consume more than six hundred times $\Delta G_{\text{sep}}$. While the specific Gibbs free energy of separation for desalinating high-salinity brines, for which thermal processes including MD are most promising, can be significantly higher than that for seawater desalination, it is still only a small fraction of the latent heat. The orders of magnitude difference between the $\Delta H_{\text{vap}}$ and $\Delta G_{\text{sep}}$ underscores the importance of latent heat recovery from condensation in achieving a reasonable energy efficiency in all thermal desalination processes. Therefore, facilitating latent heat recovery has been the central objective in the development of different configurations for thermal desalination processes, including MED, MSF, and MVC.\(^44,45\) In MD, the permeate stream heats up as it accumulates the latent heat of condensation. This heat stored in the warm permeate stream can be readily recouped to warm up the feed stream by a heat exchanger.\(^46,47\)

In the thermal desalination community, a concept called the gain output ratio (GOR) is widely used to gauge the efficiency of a distillation process. GOR is defined as the kilogram of distillate produced given the energy required to convert one kilogram of water into steam (i.e., the latent heat).\(^48\) In the worst case scenario where a distillation process recovers no latent heat from concentration, GOR = 1. However, if the latent heat can be repeatedly harnessed for preheating the feed solution, a higher GOR can be attained. With MED, a GOR of 16 is practically achievable,\(^49-51\) by lowering the specific energy consumption (SEC), which is defined as the energy consumed to produce one kilogram of product water, to a small fraction of the latent heat.

A recent study on direct contact MD (DCMD) process with heat recovery analyzed the energy consumption in an extremely ideal scenario where water recovery in the MD module was maximized, heat conduction in the MD module was minimized, and heat recovery in the heat exchanger was assumed to be perfect.\(^47\) With these highly ideal conditions, a minimum SEC of 27.6 kJ kg\(^{-1}\) (7.7 kW h m\(^{-3}\)) and a corresponding GOR of about 84 were theoretically obtained with feed and distillate temperatures of 60 °C and 20 °C, respectively. While such a theoretical estimation arguably has little relevance to practical operation and available materials, it sheds light onto important theoretical questions that can help improving our fundamental understanding of distillation.

The $\Delta G_{\text{sep}}$ for the MD desalination system discussed above, which is around 2.9 kJ kg\(^{-1}\) (0.81 kW h m\(^{-3}\)) for an initial feed salinity of 0.6 mol kg\(^{-1}\) and a single-pass water recovery of 6.4%, is a small fraction of the minimum SEC simulated (11%). The reason behind the large difference between $\Delta G_{\text{sep}}$ and the minimum SEC is that not all of the thermal energy input into the system, which is quantified by the minimum SEC, can be converted into work based on the second law of thermodynamics. In that study, the thermal energy was extracted from a generic constant temperature heat source of 60 °C. If a Carnot cycle with a constant temperature heat sink at 20 °C is employed to convert such energy to work, the maximum conversion efficiency is 12%. That means about 3.3 kJ of equivalent work is required to produce one kg of water using the ideal MD process simulated, which is much closer to, but still higher than $\Delta G_{\text{sep}}$. While the simulation of ideal MD operation with perfect heat recovery involved usage of empirical correlations for thermodynamic

---

---

**Fig. 1** Illustrations of the desalination mechanisms for (A) generic thermal process, (B) reverse osmosis (RO), and (C) electrodialysis (ED). In thermal desalination processes, evaporation of water consumes a large amount of energy in the form of enthalpy of vaporization ($\Delta H_{\text{vap}}$), which is near 2400 kJ kg\(^{-1}\) depending on the evaporation temperature and feed salinity. A subsequent condensation process (not shown here) is required to achieve desalination. In RO, water molecules in the pressurized feed solution migrate across a semi-permeable membrane that rejects the ions. The resulting water on the other side of the membrane is nearly salt free. In ED, cations (orange) and anions (green) migrate across the cation exchange membrane and anion exchange membrane, respectively, under the influence of the applied electric field. The center stream thus becomes deionized. In both RO and ED, the Gibbs free energy of separation, $G_{\text{sep}}$, which depends on the composition of the feed, product, and brine streams, represents the minimum energy that is required to separate water and dissolved solutes during desalination. In general, $G_{\text{sep}}$ is at least two orders of magnitudes lower than $\Delta H_{\text{vap}}$. It is closely related to the driving force of the energy driven processes in the desalination system.
properties and may thus lead to uncertainty in quantifying SEC, the appreciable discrepancy between $\Delta G_{\text{sep}}$ and the simulated minimum SEC may reflect a fundamental characteristic of MD and thermal desalination processes in general.

It has been demonstrated in theory that a thermodynamically reversible RO process consumes exactly $\Delta G_{\text{sep}}$. One might expect the same for thermal desalination processes such as MD. However, there exists no hypothetical thermal process, however ideal, that can satisfy the requirements of thermodynamic reversibility. To illustrate this, it is important to realize that any thermal desalination process involves simultaneous mass and heat transfer. If a desalination process were truly thermodynamically reversible, both the driving forces for mass transfer and heat transfer need to be zero simultaneously.

In thermal distillation, the equilibrium condition for mass transfer is an equal partial vapor pressure between feed solution and distillate, as governed by

$$p_v(S, T + \Delta T) = p_v(0, T)$$  \hspace{1cm} (1)

where $p_v(S, T)$ is the partial vapor pressure of water with a salinity of $S$ at temperature $T$, and $\Delta T$ is a small temperature difference between feed solution and distillate. Eqn (1) suggests that the salty feed solution has to be slightly hotter than pure distillate in order to have a zero driving force for mass transfer—a necessary condition for thermodynamic reversibility. However, the driving force for heat transfer in this case is $\Delta T$, which is positive except in the singular situation when the feed solution contains no salt. Therefore, any thermal desalination system working with a salty feed solution does not have conditions that simultaneously satisfy the thermodynamic reversibility for both mass and heat transfer. Consequently, no thermal desalination process, including MD, is thermodynamically reversible.

While the above analysis is of theoretical importance by setting the ultimate limits of energy efficiency for thermal distillation, practical thermal distillation operations can only achieve efficiencies far away from such limits. In MD, energy efficiency is primarily determined by two major factors. The first factor regards how effectively an MD membrane utilizes thermal gradient for vapor transfer as compared to conductive heat loss, which is quantified by the membrane thermal efficiency. The second factor concerns how effectively the system reuses the latent heat of condensation. Both factors will be discussed in detail later on in this review.

2.2 Comparison to other desalination processes

The desalination and treatment of high-salinity brines is inherently energy intensive. Energy typically accounts for a sizeable portion of the operational expenditure of desalination plants that are capable of treating high-salinity waters. Consequently, understanding how the energy efficiency of MD compares with competing thermal desalination technologies, particularly other processes that can utilize low-grade heat such as MED and MSF, is very important. The relatively low energy efficiency of common MD desalination systems and configurations has, in part, hampered their widespread adoption.

The energy efficiencies of different thermal desalination technologies depend on a broad range of factors, including characteristic mass and heat transfer resistances, inlet and outlet temperatures, system size and configuration, and the efficiency of key process equipment, such as heat exchangers, pumps, and compressors. It is important to note that practical system design often involves sharp trade-offs between energy efficiency and capital expenditure. For example, in smaller scale systems, energetically optimal process configurations may be altered by reducing the number of stages or heat exchangers to lower capital costs. Consequently, the design of thermal desalination systems is heavily influenced by the relative cost of energy to capital inputs.

Given the breadth of variables that strongly influence the energetic performance of thermal desalination processes, unbiased energy efficiency based comparisons between MD, MED, and MSF are challenging. Previous studies that have attempted to compare the efficiency of thermal desalination technologies using a set of shared assumptions, have found that the energetic performance of optimized MD systems can be similar to MSF and within 25% of other phase change based technologies. Studies have also shown that novel MD system designs have the potential to be cost competitive with further research and development. In this study, we use literature values and industrial data to compare the thermal efficiency, represented by gained output ratio (GOR), of state-of-the-art MD, MED, and MSF processes across a wide range of system sizes.

Fig. 2 shows the impact of system size on the GOR of MD, MED, and MSF desalination systems. GOR values were calculated and compiled from previous studies and industrial data from DesalData (Global Water Intelligence, Oxford, UK). Given the relative simplicity and modularity of MD, the range of GOR values achieved by MD desalination systems is largely independent of system size. By contrast, the energetic performance of industrial MED and MSF systems is heavily impacted by system size, as the energy efficiency, cost, and complexity of several key components scale differently with size. For small-scale desalination systems ($<1000 \text{ m}^3 \text{ day}^{-1}$), the economically optimal number of stages ranges from 1 to 4 for MED and 2 to 3 for MSF. Size-constrained systems, such as those used on ships, are often further limited to a single effect or stage. In comparison, larger-scale MED and MSF systems may employ up to 40 and 15 stages, respectively. This drastic reduction of stage numbers with system size leads to larger excess driving forces for mass transport in each stage, lowering desalination efficiency. In addition, small-scale MED and MSF systems often eschew expensive system components such as custom-built heat exchangers, further lowering efficiency. The rate of heat loss from distillation effects and flash chambers also increases with the surface area to volume ratio, which is greater for smaller systems. Consequently, MED and MSF systems are rarely used at sizes smaller than municipal-levels.

While the energetic performance of MD is superior to MED and MSF for small-scale systems ($<1000 \text{ m}^3 \text{ day}^{-1}$), substantial improvements in membrane and system design are needed to compete with large-scale MED and MSF systems.
Hybridizing MD with other thermal technologies may also enable higher process efficiencies than each system alone. Hybridization could take the form of incorporating MD as a heat exchanger into another thermal process such as MVC.

In addition to energy efficiency, there are several other factors that affect the viability of different thermal technologies. Table 1 summarizes the various strengths and weaknesses of MD compared to MED, MSF, RO, and MVC. Notably, relative to other thermal technologies, MD excels at scalability to small sizes and avoiding metal components.

3. How can membrane and system design improve MD performance?

3.1 Improving membrane properties

Membrane properties, which determine the resistance of a membrane to mass and heat transfer, can have a significant impact on the performance of membrane distillation (MD). The water flux ($J$) across an MD membrane is driven by a difference in the partial vapor pressure of water vapor between the feed- and permeate-sides of the membrane ($P_{F} - P_{P}$). This partial vapor pressure difference is generated by a temperature difference across the membrane ($T_{F} - T_{P}$). The water flux across an MD membrane is accompanied by a heat flux ($Q$), which has a convective and a conductive component. The convective component of the heat flux is caused by the evaporation of water on the feed-side of the membrane and the condensation of water vapor on the permeate-side, while the conductive component is driven by the transmembrane temperature gradient. Conductive heat transfer reduces the temperature difference across the membrane and lowers the partial vapor pressure driving force for mass transfer. Ideally, an MD membrane would have a low resistance to the transport of water vapor through the membrane while having a high resistance to conductive heat transfer.

Mass transfer across an MD membrane occurs primarily via molecular and Knudsen diffusion. The permeability coefficient of an MD membrane ($B$), which is inversely proportional to its mass transfer resistance, is defined as the mass flux of water vapor divided by its partial vapor pressure driving force. By solving the Maxwell–Stefan equations for the transport of water vapor in stagnant air, and using the Dusty Gas Model to account the resistance of the porous membrane, the permeability coefficient may be expressed as a function of membrane properties:

$$B = \frac{J}{P_{F} - P_{P}} = \frac{\varepsilon D_{w} p}{\tau_{m} R_{w} T (P_{F} - P_{P})} \ln \left( \frac{1 + \frac{D_{w}}{D_{wK}}}{\frac{P_{F}}{P}} \right) \approx \frac{\varepsilon D_{w}}{\tau_{m} R_{w} T \left( 1 + \frac{D_{w}}{D_{wK}} \right)}$$

Table 1 Qualitative comparison of membrane distillation (MD) with established desalination technologies across a wide variety of desalination performance metrics. Technologies include reverse osmosis (RO), multi-effect distillation (MED), multi-stage flash (MSF), and mechanical vapor compression (MVC). Size data is based on plant data from DesalData (Global Water Intelligence, Oxford, UK), where effective small-scale operation (three stars) refers to a produced water flow rate of < 10,000 m$^3$ day$^{-1}$. Energy efficiency was determined from several reviews as in Fig. 2. For use of low-grade energy, three stars (excellent) refers to < 70 °C, while one star (poor) is > 110 °C or unmeasurable to heat input as in RO and MVC. Minimal pretreatment performance was determined by comparing chemical additive costs relative to RO, where the other technologies shown were 50–66% of RO (two stars) or less than 50% (three stars). Lifetime cost data was included from several sources including DesalData, where three stars is < $1 m^{-3}$ and one star is > $10 m^{-3}$.

<table>
<thead>
<tr>
<th>Metric</th>
<th>MD</th>
<th>RO</th>
<th>MED</th>
<th>MSF</th>
<th>MVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-Salinity Feedwaters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small-Scale Operation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utilizing Low-Grade Energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fouling Resistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rejection of Non-Volatile Neutral Solutes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Pretreatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Lifetime Costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2 Demonstrated gained output ratio (GOR) and corresponding specific energy consumption (SEC) of membrane distillation (MD), multi-effect distillation (MED), and multistage flash (MSF) desalination systems as a function of system capacity. The GOR is defined as the mass flow rate of pure water divided by the rate of heat input into the desalination system, SEC, which is defined as the heat input per unit mass of pure water produced, divided by the enthalpy of vaporization of water divided by the partial vapor pressure driving force. Data were compiled from several reviews listing system size ranges and GOR. DesalData (Global Water Intelligence, Oxford, UK) for system size, and information on number of stages for small systems, which limits maximum GOR. Individual studies were included as well for maximum GOR values. It was assumed that smaller designs could be scaled up with no loss in GOR. Cutoffs of system size were applied where no larger or small system has been reported. Curves are approximate smooth fits to available data and are expected to describe most typical systems in industry.
where $\epsilon$, $\tau$, and $d_m$ are the membrane porosity, tortuosity, and thickness respectively; $D_{mv}$ is the molecular diffusion coefficient of water vapor in air; $D_{wK}$ is the Knudsen diffusion coefficient of water vapor in the membrane matrix; $R_w$ is the specific gas constant of water; and $T$ is the mean membrane temperature. The diffusion coefficient ratio ($D_{mv}/D_{wK}$) is inversely proportional to the nominal pore diameter ($d_p$) of the membrane at a fixed temperature.$^{72}$ The permeability coefficient of an MD membrane may be increased by increasing its porosity or nominal pore diameter, or by reducing its tortuosity or thickness (Fig. 3A).

The heat transfer coefficient of an MD membrane ($h_m$), which is defined as the heat flux divided by the temperature difference across the membrane ($T_F - T_p$), is given by,\textsuperscript{13,21,65}

$$h_m = \frac{Q}{T_F - T_p} = \frac{J\Delta H^\circ(T_P)}{T_F - T_P} + \frac{J\Delta C_P}{\exp\left(\frac{J\Delta C_P}{K}\right) - 1}$$

where $\Delta H^\circ(T_P)$ is the enthalpy of vaporization of water (calculated at the permeate-side temperature), $\Delta C_P$ is the difference between the isobaric specific heat capacity of liquid water and water vapor, and $K$ is the combined thermal conduction coefficient of the membrane. The thermal conduction coefficient of the membrane is defined as the combined thermal conductivity ($\bar{k}$) of the membrane matrix divided by its thickness ($d_m$): $K = \bar{k}/d_m$. The combined thermal conductivity depends on the membrane material, porosity, and structure. It is often approximated by a porosity weighted average of the thermal conductivities of the gas trapped in the membrane pores ($k_g$) and the membrane material ($k_m$): $\bar{k} = \epsilon k_g + (1 - \epsilon)k_m$. The first and second terms on the right-hand side of eqn (3) represent the convective and conductive components of transmembrane heat transfer, respectively. The conductive component of the heat transfer coefficient may be reduced by increasing membrane porosity, reducing the thermal conductivity of the membrane material, or increasing membrane thickness (Fig. 3A).

![Fig. 3](A) Schematic outlining key membrane properties including membrane porosity, nominal pore diameter, thickness, and thermal conductivity of the membrane material. Water vapor diffuses through stagnant air in the membrane pores from the feed- to the permeate-side. Heat is transferred from the feed- to the permeate-side by convection and conduction. (B) Thermal efficiency and permeability to thermal conduction coefficient ratio as function of membrane porosity for a membrane with nominal pore diameters of 0.1 μm (red curves, correspond to a diffusion coefficient ratio of 1.4) and 0.5 μm (blue curves, diffusion coefficient ratio of 0.29), and material thermal conductivities of 0.15 W m$^{-1}$ K$^{-1}$ (dashed lines) and 0.20 W m$^{-1}$ K$^{-1}$ (solid lines). The feed- and permeate-side temperatures are fixed at 55°C and 45°C, respectively. (C) Exergy efficiency as a function of membrane permeability coefficient for membranes with porosities of 50% (red curve), 70% (green curve), and 90% (blue curve). Exergy efficiency is defined as the ratio of the useful desalination work performed by the system relative to the maximum work that could be performed by a thermodynamically reversible system operating under the same conditions. The corresponding electrical equivalent energy consumption is indicated on the secondary (right) axis. The thickness of each curve represents the change in exergy efficiency as the nominal pore diameter is increased from 0.2 μm at the bottom edge to 0.5 μm at the top edge. Exergy efficiency and permeability coefficient values corresponding to membrane thicknesses of 30 μm (diamonds), 50 μm (squares), and 100 μm (circles) are indicated for each membrane porosity. The MD desalination system has an initial feed salinity of 70 g kg$^{-1}$ (70 000 ppm) and a system-scale water recovery of 50%. The average water flux for the MD system is fixed at $1.5 \times 10^{-3}$ kg m$^{-2}$ s$^{-1}$ (5.4 kg m$^{-2}$ h$^{-1}$) and perfect external heat recovery is assumed. The feed stream is recycled to allow a module-scale recovery of 5% and the initial permeate temperature is 20°C. Note: for an initial feed temperature of 80°C an electrical equivalent energy consumption of 100 kJ kg$^{-1}$ corresponds to an energy consumption of around 690 kJ kg$^{-1}$. 

This journal is © The Royal Society of Chemistry 2018 Energy Environ. Sci.
Thermal efficiency ($\eta_{th}$), which is defined as the convective heat flux across the membrane divided by the total heat flux, $\eta_{th} = J\Delta H(T) / Q$, is widely used as a measure of the efficacy of an MD membrane over an infinitesimal membrane area.\textsuperscript{29,81,82} As $\eta_{th}$ approaches 1, negligible heat is wasted by conduction across the membrane. Membranes with a high thermal efficiency have a relatively low resistance to mass transfer and a relatively high resistance to conductive heat transfer, which results in the transmembrane heat flux being dominated by the convective component. Eqn (2) and (3) may be used to express thermal efficiency as a function of membrane properties, diffusion coefficients, and the mean membrane temperature.\textsuperscript{14,21,54,83}

$$
\eta_{th} = \frac{1}{1 + \frac{K(T_F - T_P)}{B(\rho_F - \rho_P)\Delta H(T)}}
$$

$$
\approx \frac{\varepsilon k(T_F - T_P)}{1 + \frac{D_{wa}}{D_{K}} R_w T} + \frac{\varepsilon D_{wa}(\rho_F - \rho_P)\Delta H(T)}{1 + \frac{D_{wa}}{D_{K}} R_w T}
$$

Fig. 3B shows that the thermal efficiency of transmembrane heat transport may be increased by increasing the porosity, reducing the material thermal conductivity, or increasing the nominal pore diameter of an MD membrane. Current membranes have porosities ranging from around 0.70 to 0.90, tortuositities ranging from around 1.05 to 1.20, material thermal conductivities ranging from 0.15 W m$^{-1}$ K$^{-1}$ to 0.30 W m$^{-1}$ K$^{-1}$, and nominal pore diameters ranging from 0.05 $\mu$m to 0.75 $\mu$m.\textsuperscript{54,84-91} These correspond to a wide range of thermal efficiency values from approximately 29% at the lower end to around 84% at the higher end, assuming feed- and permeate-side temperatures of 55 $^\circ$C and 45 $^\circ$C, respectively.\textsuperscript{43,92,93} While thermal efficiency, which focuses on mass and heat transfer across an infinitesimal membrane area, is a useful indicator of MD performance, it is unable to fully capture the impact of membrane properties on cumulative mass and heat transfer across a finite membrane area.

Recent studies have used module-scale modelling to calculate the energetic performance of MD desalination at the systemscale.\textsuperscript{13,16,24,51,94-97} Module-scale modelling accounts for the variations in the mass and enthalpy flow rates of the feed and permeate streams along a membrane module by integrating mass and heat flux expressions over a finite membrane area.\textsuperscript{22,23} By incorporating feed and permeate stream recycling and heat recovery, system-scale analyses are able to calculate the specific energy consumption, energy efficiency, and exergy efficiency of MD desalination, enabling a deeper understanding of the impact of membrane properties on process performance.\textsuperscript{22,23,98} Fig. 3C shows the impact of the membrane permeability coefficient on the exergy efficiency of DCMD desalination for membrane porosities of 0.50 (red curve), 0.70 (green curve), and 0.90 (blue curve). Exergy efficiency is defined as the minimum heat input that would be required by a thermodynamically reversible desalination process operating between the same thermal reservoirs as the modeled desalination system divided by the heat input required by the modeled system. Membrane thickness decreases from left to right for each membrane porosity, while the nominal pore diameter increases from 0.2 $\mu$m to 0.5 $\mu$m going from the bottom to top edge of each curve. The initial feed salinity is 70 g kg$^{-1}$, the system-scale water recovery ($R$) is 0.50, the module-scale recovery ($r$) is 0.05, the average water flux ($J$) is fixed at 1.5 $\times$ 10$^{-3}$ kg m$^{-2}$ s$^{-1}$ (5.4 kg m$^{-2}$ h$^{-1}$), the initial permeate temperature is 20 as in previous work.\textsuperscript{21} The least work of separation per unit mass of permeate ($\dot{W}_{use}$) is 8.17 kJ kg$^{-1}$ (2.27 kW h m$^{-3}$). The heat input and temperature is calculated as described in a previous study.\textsuperscript{21} Perfect external heat recovery is assumed throughout.

Fig. 3C demonstrates the impact of maximizing porosity while optimizing the permeability coefficient by controlling the thickness of an MD membrane on desalination efficiency. Increasing the porosity or the nominal pore diameter of an MD membrane, increases its permeability coefficient, reducing the entropic losses associated with transmembrane mass transfer thus increasing the exergy efficiency of MD desalination.\textsuperscript{89} Increasing porosity has the additional effect of lowering the thermal conduction coefficient of an MD membrane, which leads to an increase in exergy efficiency by reducing the entropic losses associated with conductive heat transfer. Consequently, increasing membrane porosity yields the greatest enhancement in desalination efficiency.

Although increasing membrane porosity has a large positive impact on the exergetic performance of MD, it can have a detrimental impact on the mechanical robustness of an MD membrane. For highly porous MD membranes, increasing porosity further may lead to a notable reduction in the strength of the membrane matrix, increasing the likelihood of membrane compaction occurring during operation. Membrane compaction can have a significant negative impact on the efficiency of MD by reducing membrane porosity and thickness.\textsuperscript{101,102} Similarly, while increasing the nominal pore diameter of an MD membrane leads to an increase membrane permeability coefficient, it also increases the likelihood of wetting, which is discussed in subsequent sections.\textsuperscript{21,53} Novel membrane materials and architectures may be able to overcome these limitations. For example, a composite membrane with a highly porous inner section sandwiched between two thin dense outer layers could yield a high permeability coefficient without being overly susceptible to wetting.

Fig. 3C shows that increasing the permeability coefficient by reducing membrane thickness (going from the left to right along each curve) initially leads to an increase in exergy efficiency, as the entropic losses associated with mass transfer are reduced. However, as membrane thickness is reduced further, below the 50 $\mu$m to 100 $\mu$m range, the gains in exergy efficiency begin to fall as the increase in the thermal conduction coefficient of the membrane leads to an increase in wasteful conductive heat transfer. Ultimately, increasing membrane permeability beyond around 3.0 $\times$ 10$^{-6}$ kg m$^{-2}$ s$^{-1}$ Pa$^{-1}$ (1080 kg m$^{-2}$ h$^{-1}$ bar$^{-1}$) by reducing membrane thickness leads to a reduction in the exergy efficiency as the increased entropic losses associated with heat transfer outweigh the reduction in entropic losses associated with mass transfer. Optimizing membrane thickness, which balances the benefits of a low mass transport resistance with...
the need for a high conductive heat transport resistance, is therefore an important part of maximizing the exergetic performance of a given MD membrane.\textsuperscript{21,103,104}

The optimal membrane properties described in this section are relevant for MD configurations in which both sides of the membrane are in contact with liquid water, including direct-contact and permeate-gap MD. It is important to note that in air-gap, sweeping-gas, and vacuum MD, minimizing membrane thickness is always energetically favorable, as transmembrane molecular and Knudsen diffusion are no longer the primary resistance to water transport from the feed to the distillate streams.\textsuperscript{25,27,105}

3.2 Optimizing process design and operation

The performance of membrane distillation systems can be improved through optimization of the operating conditions or system configuration. Optimization efforts can enhance the membrane permeability or reduce temperature polarization to achieve a higher flux and thermal efficiency. System-level modifications can also improve recovery of the heat transferred across the membrane to increase the efficiency. In this subsection, a number of approaches are discussed.

Thermal boundary layers on either side of the MD membrane result in a reduced temperature difference at the membrane air–liquid interface as compared to the bulk solution (Fig. 4A).\textsuperscript{14,106–108} The impact of this temperature polarization is highly dependent on the hydrodynamic flow conditions in the feed and permeate channels. The severity of temperature polarization can be reduced by operating the feed and permeate channels with high crossflow velocity or by improving the channel spacer design to enhance hydrodynamic mixing.\textsuperscript{14,21,109} For a high permeability membrane with relatively poor mixing, temperature polarization can reduce the partial vapor pressure driving force by nearly half.\textsuperscript{21} By increasing the crossflow to improve hydrodynamic mixing, the impact of temperature polarization can be reduced such that the partial vapor driving force decreases by less than 10%. It is important to note that reducing the impact of temperature polarization by improving hydrodynamic mixing will inevitably increase the pressure drop in the membrane channel, which leads to a corresponding increase in the pumping energy required in the system.\textsuperscript{110,111} MD modules must therefore balance the need to reduce temperature polarization with the increased pumping energy required.

Innovative methods have been developed to actively heat the membrane surface on the feed side of the membrane to improve the performance of MD (Fig. 4A, yellow curve).\textsuperscript{16,112,113} The surface heating can be used to reduce temperature polarization or achieve higher recoveries than would be possible in a standard MD system. Surface heating can also be used as the sole heat input to reduce the overall size of a system. Two methods have been explored for self-heating membranes: Joule heating and photothermal heating. In Joule heating, the surface of the MD membrane is modified with a conductive material, and an electric current is applied across the conductive layer.\textsuperscript{112} The resistance to electron flow through the conductor causes the surface of the membrane to heat. In photothermal surface heating, UV or solar irradiation of the membrane surface causes a localized heating effect.\textsuperscript{16} To enhance heating, the membrane is modified with a functional material that increases the light absorbance of the membrane. While self-heating membranes may be beneficial to achieve high water recoveries, the energy consumption will be very high (more than two orders of magnitude above the theoretical limit for desalination) unless methods to recover the heat transferred across the membrane are used.\textsuperscript{20}

In the case of Joule heating, high quality electrical energy is required to operate the process, rather than low-grade heat, further increasing the potential cost of such a system. For photothermal systems, it is challenging to effectively irradiate large areas of membrane using UV or solar energy, since membrane modules typically pack membranes very tightly in

Fig. 4  (A) Temperature profiles of a membrane system with low-hydrodynamic mixing, high hydrodynamic mixing, and surface heating. The driving force for vapor permeation is determined by the temperature difference at the air–liquid interfaces of the membrane. \(T_H\) indicates the bulk hot temperature and \(T_C\) indicates the bulk cold temperature. \(T_{bi}\) indicates the bulk hot temperature and \(T_{ci}\) indicates the bulk cold temperature. (B) Direct contact membrane distillation system with heat recovery in the permeate stream via an external heat exchanger. (C) Air-gap and permeate-gap membrane distillation systems that utilize a heat-conductive condensing plate instead of an external heat exchanger.
spiral wound or flat sheet configurations. Thus, the implementation of surface heating techniques will likely be limited to small scale systems where high water recoveries are desired, rather than large-scale desalination systems that must operate at higher efficiencies.

The transport resistances in the membrane are strongly dependent on the bulk diffusion of vapor through air within the membrane pores, and substantial improvements in the vapor permeability may be possible by reducing the air pressure. As the pressure in the membrane pores decreases, the mean free path of the vapor molecules increases, resulting in reduced overall transport resistances. It has been shown that a 50% reduction of gas pressure within the membrane pores can lead to an 80% increase in the vapor permeability of the membrane and a corresponding 20% increase in the exergy efficiency of a model MD system. Operating with reduced pressure is particularly advantageous because an increase in vapor permeability is achieved without a corresponding increase in the thermal conductivity of the membrane, and a higher thermal efficiency can therefore be achieved. Despite the advantages, practical implementation of MD systems that utilize a reduced bulk air pressure is challenging since the feed and permeate streams must be deaerated. This will likely incur substantial capital and operating costs.

As discussed in the previous section, the energy required for thermal distillation is very high due to the transferred latent heat of vaporization. Recovery of the heat transferred across the membrane is therefore critical to enable efficient systems. The most commonly discussed configurations of DCMD employ a separate heat exchanger which transfers heat from the permeate that exits the membrane module to the feed stream that enters the module (Fig. 4B). The incorporation of this heat exchanger improves the efficiency of the system substantially, allowing for a theoretical decrease in the energy requirement of more than an order of magnitude.

Further improvements in the efficiency of MD are possible by using alternative configurations of the system. Air-gap MD is a potential configuration that enables simplified operation and may also improve heat recovery (Fig. 4C). In this configuration, the cold source water is used to cool a condensing plate that is placed on the permeate side of the membrane. After passing along the condensing plate, the source water is heated and flows across the membrane as the feed stream. Water transferred across the membrane will collect on the condensing plate and form the permeate stream. The air-gap MD operation thus preheats the incoming feed stream as water collects on the condensing plate. Studies have shown that air-gap MD can reach similar efficiencies as DCMD systems, with better performance for air-gap MD at higher salinities. Variants of air-gap MD that fill the collecting channel with permeate water or use conductive spacers in the permeate channel have been proposed to improve the performance of the system, and these systems can theoretically achieve higher efficiencies than the standard air-gap MD system at low and moderate salinities. An additional advantage of the air- or permeate-gap MD configurations is that a separate permeate stream is not required at the start of operation. These configurations may also be more compact since separate membrane modules and heat exchangers are not required.

4. Will fouling, scaling, and wetting hinder MD?

4.1 Current strategies of fouling and scaling control in MD

Similar to other membrane processes, such as RO, nanofiltration (NF), and forward osmosis (FO), the performance of MD is largely constrained by membrane fouling and scaling. Organic and biological foulants as well as inorganic scalants accumulate on the membrane surface or within the membrane pores, adversely impacting water productivity and reducing membrane lifespan. Membrane fouling and scaling have been studied extensively for RO, NF, and FO. Herein, we focus on the unique aspects of MD fouling and scaling, which originate from the different membrane properties and operating conditions found in MD compared to hydraulic pressure-driven membrane processes.

MD is generally considered to have a lower fouling propensity than hydraulic pressure-driven membrane processes such as RO and NF, owing to its larger membrane pore size and low operating pressure. However, MD membranes, which can be used in the desalination of high-salinity brines, may be subjected to significantly higher concentrations of foulants and scalants than found in RO. In MD, fouling and scaling lead to both a reduction in the membrane permeability coefficient and an increase in the thermal conduction coefficient, thereby lowering desalination efficiency. Fouling and scaling layers, particularly those with small pore sizes (<50 nm) or low free volumes, impose an additional resistance to mass transfer. In addition, fouling and scaling can also hinder mixing near the membrane–solution interface in the feed stream, exacerbating the detrimental impact of temperature polarization on transmembrane mass transport. Flux decline is not the only consequence of membrane fouling and scaling; pore wetting, which is often caused by surfactants or low surface tension foulants, can lead to a significant deterioration in permeate quality and ultimately process failure.

Anti-fouling membranes, which are designed to increase the energetic and kinetic barriers to foulant attachment, form an important part of fouling mitigation strategies in RO, NF, and FO. However, at present, efforts to prepare anti-fouling MD membranes are relatively limited. Optimizing surface wettability, which has a large impact on the fouling propensity of an MD membrane in addition to preventing membrane failure by wetting, is currently the primary aim of membrane modification. Superhydrophobic and omniphobic membranes have been fabricated via biomimetic approaches by tailoring both membrane surface chemistry and morphology. The introduction of low surface energy materials, such as fluoralkylsilane coatings, and re-entrant structures have enabled the design of MD membranes with exceptional wetting resistance. The resultant membranes have demonstrated stable performance in MD
desalination of feedwater with low surface tension contaminants, including industrial wastewater from shale gas production.\textsuperscript{31–34} More recently, an in-air hydrophilic, or underwater oleophobic, coating layer has been incorporated on the MD membrane surface to reduce fouling by hydrophobic foulants such as oils.\textsuperscript{10,138} However, to date, less progress has been made on the design of MD membranes resistant to mineral scaling. An improved understanding of the fundamental relationship between membrane surface chemistry and scaling is essential to the further development of anti-scaling MD membranes.

Beyond membrane design, several process-based strategies have also been employed to mitigate fouling and scaling in MD. Feedwater pretreatment by physicochemical processes, including softening, coagulation, filtration, pH adjustment, and degasification, has been used to reduce MD membrane fouling and scaling.\textsuperscript{121,122,139,140} While these mature pretreatment processes have previously been successfully applied to RO and NF,\textsuperscript{141} careful selection based on feedwater chemistry is imperative. Other studies have focused on developing and implementing novel operational techniques to mitigate fouling and scaling. For example, bubbling gas through the feed stream has been shown to improve MD performance by increasing mixing and thus reducing temperature polarization while also forming a foulant-blocking air layer near the membrane solution interface.\textsuperscript{142–146} In another study on the MD desalination of hypersaline water from the Great Salt Lake, which has a salinity of \( \geq 150 \text{ g kg}^{-1} \), reversing the salinity and temperature difference between feed and permeate channels was used to control scaling and maintain stable water flux.\textsuperscript{147} In addition, chemical cleaning has been used to restore MD water productivity by removing scaling and fouling layers from the membrane surface.\textsuperscript{148,149} Acidic chemical agents are particularly effective at dissolving certain inorganic scales such as calcite (\( \text{CaCO}_3 \)) and iron or aluminum oxides.\textsuperscript{149–151} However, the use of strong acids, such as 5 wt\% hydrochloric, citric, and sulfuric acid solutions, for cleaning was shown to damage polypropylene and PTFE membrane integrity.\textsuperscript{148–151}

The use of anti-scalants is the most common strategy for scaling control in MD desalination.\textsuperscript{123} Anti-scalants inhibit scale formation by disrupting crystallization through delaying the onset of nucleation or retarding the growth of formed crystals.\textsuperscript{123} Commercially available anti-scalants can prolong the induction time for both calcite and gypsum without affecting the water flux and permeate conductivity in MD.\textsuperscript{152}

However, the performance of anti-scalants in MD for scaling control requires further evaluation. For example, it was reported that polyphosphate-based anti-scalants hindered the formation of \( \text{CaCO}_3 \) crystals in MD, but led to the formation of an amorphous, non-porous scale layer on the membrane surface that reduced water flux.\textsuperscript{150} As with other membrane processes, anti-scalants increase the operational cost of desalination and may result in organic and biological fouling.\textsuperscript{153} Hence, reducing the use of anti-scalants by developing membranes with an improved resistance to scaling or new MD scaling mitigation techniques is highly desirable to promote the sustainability of MD.\textsuperscript{154} However, improving membrane resistance to scaling is challenging as the relationship between MD membrane properties and scaling potential has yet to be established.

### 4.2 Surface wettability plays a critical role in determining membrane robustness

Conventional MD membranes comprise symmetric hydrophobic pores of sub-micrometer size. They are typically fabricated from low surface energy fluoropolymers such as polyvinylidene fluoride (PVDF) and polytetrafluoroethylene (PTFE).\textsuperscript{36} Membrane surface properties can be tailored to increase process robustness by improving the non-wetting features that are essential for MD separation.\textsuperscript{135} Surface charge and hydrophilicity are two main characteristics determining fouling, scaling, and wetting behaviors of MD membranes. However, for high-salinity applications, where the effect of electrostatic interactions between foulants and membrane surface becomes negligible, surface charge may not play a significant role.\textsuperscript{155} Thus, research has focused on engineering surface wettability, including the development of MD membranes with superhydrophobic, omniphobic, and hydrophilic surfaces, for enhanced MD performance (Fig. 5).\textsuperscript{31,115–158}

Until recently, the fabrication of high performance MD membranes has mainly focused on increasing wetting resistance to water by constructing superhydrophobic surfaces. Roughening of hydrophobic surfaces, followed by coating with low surface energy materials, is a typical approach for the development of superhydrophobic MD membranes. Inorganic nanoparticles, such as silica and titanium oxide, have been employed to obtain MD membrane surfaces with high roughness.\textsuperscript{136,159} Superhydrophobic MD membranes can offer several advantages over conventional hydrophobic MD membranes, including long-term wetting resistance and lower fouling propensity.\textsuperscript{145,150,161} Despite the enhanced wetting resistance to water, superhydrophobic MD membranes are susceptible to wetting and fouling by low surface tension substances. Similar to conventional hydrophobic MD membranes, application of superhydrophobic MD membranes to treat challenging wastewaters that contain diverse low surface tension contaminants is limited.\textsuperscript{31,32,34} For example, oil readily fouls membrane pores, substantially compromising separation efficiency of superhydrophobic MD membranes.\textsuperscript{162–164} Amphiphilic surfactants adsorb to a superhydrophobic surface via attractive hydrophobic–hydrophobic interactions due to the presence of hydrophobic tail moieties, and subsequently render the membrane surface hydrophilic. Surfactants also significantly lower the surface tension of saline solutions due to their accumulation at the air–liquid interface, resulting in wetting of superhydrophobic MD membranes.\textsuperscript{33}

Omniphobic MD membranes that repel both water and oil provide enhanced robustness for MD separation. Like superhydrophobic membranes, omniphobic membranes require surfaces with ultralow surface free energy. However, omniphobicity cannot be achieved by simply lowering the surface energy, because wetting by low surface tension liquids remains thermodynamically favorable for low energy surfaces. Specifically, omniphobic surfaces require a re-entrant structure to develop a local kinetic barrier for
transition from the meta-stable Cassie–Baxter state to the fully wetting Wenzel state for low surface tension liquids.\textsuperscript{33,165,166}

Electrospinning has been demonstrated to be an effective strategy to fabricate omniphobic MD membranes.\textsuperscript{34} Cylindrical electrospun fibers with a re-entrant structure provide an ideal platform for the development of omniphobic membrane substrates.\textsuperscript{167–169} Further decorating the fibers with spherical nanoparticles was shown to effectively increase the level of re-entrant structure, further enhancing wetting resistance to low surface tension substances.\textsuperscript{33,34,137} More recently, coaxially electrospun fibers have been used to fabricate omniphobic MD membranes with multilevel re-entrant structure.\textsuperscript{170}

Omniphobic MD membranes offer excellent wetting resistance to low surface tension solutions including waters containing organic solvents, alcohols, and surfactant.\textsuperscript{143} However, non-polar contaminants such as oil are attracted to omniphobic surfaces, which have an ultralow surface energy, via attractive hydrophobic–hydrophobic interactions. Although the re-entrant structure of an omniphobic surface provides a kinetic barrier to the penetration of non-polar contaminants into membrane pores, the accumulation of these contaminants on the membrane surface blocks pores, reducing membrane permeability.\textsuperscript{30}

Recent studies have demonstrated that hydrophilic surface modifications are an effective way of enhancing the anti-wetting and antifouling properties of MD membranes.\textsuperscript{157,164,171} These surface modifications focus on developing a thin hydrophilic layer on top of a hydrophobic substrate to hinder the transport of non-polar contaminants to the membrane surface without reducing membrane permeability or exacerbating the impact of temperature polarization substantially. The hydration layer that forms on a hydrophilic surface effectively repels non-polar contaminants, lowering the wetting propensity of MD membranes. However, hydrophilic layers provide a limited kinetic barrier to the transport of low surface tension substances that are miscible with or soluble in water. For example, alcohols and surfactants are readily transported across the hydrophilic layer and are subsequently able to wet the underlying hydrophobic substrate.\textsuperscript{30}

The successful application of MD membranes with special surface wettability requires further development. Such membranes need to be evaluated in long-term MD operation to ensure high desalination performance and wetting resistance.\textsuperscript{149,172} The effects of fouling, scaling, and wetting on separation efficiency of MD membranes are often highly synergistic in practical systems where diverse inorganic and organic contaminants are present in the feed stream. A systematic performance evaluation of MD membranes involving various feedwater compositions is required to ensure robust process operation is maintained across a wide range of applications. Lastly, development of scalable fabrication methods of anti-wetting membranes is essential for their implementation in an industrial scale.

5. Where does MD outperform other desalination technologies?

5.1 MD can utilize low-grade thermal energy

As we discussed earlier, the energy efficiency of membrane distillation is relatively low compared to other desalination systems such as reverse osmosis. However, MD can benefit

\[ \text{Fig. 5 Schematic representing fouling and wetting behaviors of MD membranes with (A) hydrophobic, (B) superhydrophobic, (C) hydrophilic, and (D) omniphobic surfaces. MD membrane fouling and wetting with feed waters containing low surface tension contaminants (e.g., oil and alcohol) and surface-active reagent (e.g., surfactants) are depicted to describe how surface wettability impacts MD separation efficiency.} \]
from using low-value heat sources that are abundant and underutilized. By employing low-grade heat sources, the MD process can reduce the required electricity inputs. Several low-grade heat sources have been considered for the membrane distillation process, including geothermal reservoirs, low-temperature solar thermal collectors, and waste heat sources such as power plants and industrial facilities.\textsuperscript{173,174} Due to inherent inefficiencies in thermal desalination processes, direct uses of low-grade heat, such as district heating for residential and commercial sectors, should be considered before using heat energy for other applications such as desalination.\textsuperscript{75,76,175,176}

Waste heat from industrial sources and power plants represents a possible source of low-grade heat for MD. In the United States, it is estimated that power plants discharge more than 5000 TW h per year as waste heat based on an average efficiency of around 30\%;\textsuperscript{15,177,178} and the industrial sector is estimated to discharge another 3000 TW h per year as waste heat.\textsuperscript{179,180} While the total quantity of this waste heat is massive, the quality or temperature of individual sources is still uncertain. For waste heat from thermal power plants, it is estimated that most of the heat discharged (around 95\%) is below 42 °C, while some of the remaining energy may be appropriate as a source for MD.\textsuperscript{15} Industrial sources may be more promising low-grade heat sources, but efforts to quantify the quantity and temperature of available heat have been limited.\textsuperscript{180} Using waste heat in MD can lead to additional environmental benefits since the MD process helps to lower discharge temperatures to meet regulations and protect receiving water bodies.\textsuperscript{181}

Low-temperature solar thermal collection systems, which use relatively simple components at a low cost, could potentially be implemented for MD.\textsuperscript{47,113,182–185} The scalability, lack of emissions, and locational availability of solar energy make it a promising possible source of heat energy for MD. However, practical implementation of solar systems will require overcoming several barriers. The physical footprint of solar collectors that are able to supply enough energy for reasonable water production may be prohibitively large.\textsuperscript{17,181} Assuming 5 kW h per square meter per day of solar irradiation, direct distillation would result in only around 40 liters of water per square meter of solar collector area per day.\textsuperscript{186} While an MD system with heat recovery will be able to produce more distillate, it is still apparent that a very large solar collector area will be required for such a system to have a suitable desalination capacity. This can result in prohibitively high capital costs or restrict solar MD implementation to areas with a large available land area.\textsuperscript{184} The intermittency of solar energy is also a challenge as a constant heat supply is preferred for the system to operate at peak efficiency.

Geothermal low-grade heat sources offer a stable energy source for MD.\textsuperscript{187–190} Geothermal temperatures less than 150 °C are spatially abundant and available from relatively shallow wells (less than 6.5 km deep in most of the continental United States).\textsuperscript{191} Use of geothermal heat energy could be particularly promising for MD systems treating shale gas produced water, which already require drilling to depths up to around 3 km.\textsuperscript{155} The major challenges to geothermal energy are the relatively high capital cost of well drilling and potential constraints on the location of the geothermal well.

5.2 MD has a low areal footprint and is a modular process
As described previously, MD is inherently less energy efficient than desalination processes in which water does not undergo a phase change, such as RO and electrodialysis (ED). Previous studies have shown that, for larger-scale systems, the energetic performance of MD is lower than that of multi-effect distillation (MED) and multi-stage flash (MSF) (see Fig. 2 and associated discussion). However, as a membrane-based separation process, MD has a relatively small areal footprint and is highly suited to modular system configurations.\textsuperscript{192–194} The compactness and modularity of MD desalination may be highly advantageous in niche, small-scale desalination applications treating water that is too saline for RO desalination.

MD desalination systems, which can operate at much lower hydraulic pressures than RO and require less sophisticated heat transfer equipment than MED and MSF, can be developed with simple, inexpensive, and lightweight process equipment.\textsuperscript{195,196} For example, the cost of piping and pumps is expected to account for less than 15\% of the total installation cost of a small-scale (< 500 L day\textsuperscript{-1}) MD desalination system.\textsuperscript{197,198} The small areal footprint and relatively simple process design of MD are particularly advantageous for portable desalination systems that can be deployed in off-grid applications. The processing capacity or treatment rate required by wastewater desalination systems may vary widely, depending on the application and operating conditions. For instance, the volume of brine produced by the shale gas industry ranges from 1200 m\textsuperscript{3} to 6000 m\textsuperscript{3} per well, with the flow rate from each well varying significantly over its lifetime.\textsuperscript{199–201} MD systems, which are modular and scalable, are well suited to match the dynamic demands of wastewater desalination.

Compared to RO, MD is more resistant to fouling due to its large membrane pores and the absence of an applied hydraulic pressure.\textsuperscript{121,202} Its lower fouling propensity allows MD to have minimal feed pretreatment and less frequent membrane cleaning. Fewer pretreatment and cleaning requirements facilitate simpler and more versatile process designs that can be readily adapted for different applications. In addition, novel MD membranes with high resistance to wetting and fouling can offer unprecedented advantage in reducing or even in eliminating the need for feed pretreatment and membrane chemical cleaning. Such MD membranes will likely be the key element for successful implementation of small-scale, off-grid MD desalination systems.

Recent developments of self-heating MD membranes can strengthen the potential of MD as a small-scale, off-grid, and portable desalination system.\textsuperscript{47} Such systems can enhance desalination performance of MD by incorporating plasmonic\textsuperscript{48} and nanophotonic\textsuperscript{49} materials in a thermal desalination membrane and irradiation by UV light or sunlight to provide local heating of the membrane surface.\textsuperscript{203} These MD systems would be advantageous in locations and situations where prime electrical energy supply is limited, such as developing countries, forward operating bases, and remote unconventional hydrocarbon
5.3 MD excels with challenging hypersaline wastewaters

The treatment of hypersaline wastewaters is a challenging task that is attracting increased interest. For example, shale gas extraction generates large volumes of produced water with salinities or total dissolved solids (TDS) up to 360 g kg\(^{-1}\).\(^{115}\) The high costs and environmental risks associated with disposal of high-salinity brines necessitate low-cost and energy-efficient treatment technologies. Recently, zero liquid discharge (ZLD) applications have expanded globally as an important wastewater management strategy.\(^{8,204-209}\) The solid wastes produced by ZLD processes, which can be disposed of at landfill sites or used as a feedstock for the production of salt byproducts, are less environmentally hazardous than hypersaline brines.\(^{8}\) While RO is the most efficient desalination or brine concentration technology, its current salinity limit of around 80 g kg\(^{-1}\) limits its use to the initial stages of the ZLD process. Phase-change based desalination processes, such as MD, MED, MSF, MVC, and thermolytic FO, are required to desalinate or concentrate brines above the salinity limit of RO.\(^{115}\)

As a phase-change based desalination process, MD is able to concentrate wastewaters to salinity levels that are similar to other thermal desalination technologies such as MED, MSF, MVC, and FO coupled with thermolytic draw solute recovery.\(^{40,210-212}\) Compared to established processes such as MED, MSF, and MVC, the capital costs associated with MD are lower due to its simpler heat transfer equipment and lower operating temperatures.\(^{213}\) The capability of MD to leverage low-grade thermal energy, as discussed earlier, can reduce operational expenditure, by utilizing cheaper heat sources while also reducing the carbon footprint of the desalination system. In addition, the modular nature of an MD system bolsters its versatility, allowing it to adapt to the salinity variation and dynamic demand for wastewater treatment. These features enable MD to be used in an on-site and easy-to-deploy wastewater treatment system, which is particularly advantageous given the remote locations where hypersaline wastewaters may be generated, such as shale gas extraction sites.\(^{155,214}\)

A simplified scheme of an MD-integrated ZLD system is presented in Fig. 6.\(^8\) In such a system, RO is used as the first step to treat the incoming wastewater until its salinity limit. An MD unit is followed to further concentrate the RO brine. The hypersaline MD brine is then sent to a crystallizer where the remaining water is recovered. Due to the low water recovery rate of a single-pass, single module MD unit (6.4% for initial feed and permeate temperatures of 60 °C and 20 °C, respectively),\(^{13}\) a multi-stage configuration\(^{57}\) or brine recycling (not shown in Fig. 6) is required to improve water productivity and energy efficiency. For hypersaline wastewaters that RO cannot treat, the RO step will be removed, leaving MD the sole treatment unit before the brine crystallizer. Utilizing low-grade thermal energy as the energy source for MD is critical to increase the economic viability of MD-integrated ZLD systems. A techno-economic assessment has been recently performed for the treatment of Marcellus shale produced water by MD, showing that the use of waste heat can potentially reduce the treatment cost by around 85% from $8.6 to $1.1 per m\(^3\) of water produced (or $5.7 to $0.7 per m\(^3\) of feedwater treated).\(^{215}\)

However, the large-scale implementation of MD for the treatment of hypersaline wastewaters faces several challenges. Currently, hydrophobic microfiltration membranes, rather than specifically designed commercial MD membranes are often used in MD studies.\(^{216,217}\) The relatively low average water fluxes of currently available membrane modules hinder the use of MD to desalinate brines from brackish water RO.\(^{215}\) MD membranes with a low resistance to water vapor transport and a high resistance to conductive heat transfer need to be developed, as suggested in previous sections, to enhance the energetic performance of MD desalination. Furthermore, the effectiveness of MD in desalinating hypersaline wastewaters is highly constrained by inorganic scaling.\(^{218}\) Hence, lowering the scaling propensity of MD membranes by employing approaches described earlier in this review is of paramount importance to enhance the feasibility of MD in hypersaline wastewater treatment. To date, MD desalination of hypersaline wastewaters has been mostly limited to laboratory-scale studies.\(^{35,147,173,218-220}\) Future studies should be carried out to evaluate the performance of MD at pilot- and full-scale systems.

5.4 Economic viability of MD desalination is uncertain

Economic considerations are especially important for the viability of MD as a sustainable desalination process. Without the availability of sufficient waste heat, the energetic costs associated with MD, and other phase-change based desalination processes, are very high compared to RO and ED.\(^{5,11}\) However, from an economic perspective it is also important to consider capital expenditure, which can account for around 50% of the total water cost in traditional thermal desalination processes.
such as MED and MSF. Unlike RO, MD does not require high-pressure pumps or membrane module casings. In addition, compared to other thermal desalination processes, including MED, MSF, and MVC, MD does not require specialized process equipment such as custom-built compressors, heat exchangers, or flash chambers. Consequently, the capital costs associated with MD can be relatively low, particularly for small-scale desalination systems, as discussed previously. Economically, the modular nature of MD is also advantageous, as it enables desalination capacity to be increased and decreased rapidly, without a significant financial penalty, to match transient water treatment demand.4,54,213,215

Given the significance of capital costs and operational expenditure that is not directly linked to energy consumption, such as maintenance, it is important to note that maximizing energy efficiency of an MD desalination system does not equate to minimizing water cost.53 In fact, the correlation between energy efficiency and cost remains poorly understood. Energy efficiency and economic cost values reported in the literature are highly variable, spanning as much as four orders of magnitude.27,44,52–54,115,143,177,178,192,212–218 These large variations in energy consumption and water cost are partly due to a wide variety of operating conditions and the lack of standardized cost calculation methodologies. For example, when modeling solar-powered, off-grid desalination, for which MD is particularly well suited, cost calculations often neglect to account for factors such as unsteady or transient operation, which is important given the time variability of solar irradiance.184,186,198 Furthermore, the relative absence of industrial data restricts reliable cost comparisons between MD and other desalination processes.55

MD has the potential to be economically competitive when desalinating brines with salinities in excess of 80 g kg⁻¹ that are too saline for RO.58 From an economic perspective, the scalability and modularity of MD, combined with the relatively standard process equipment required, are also highly advantageous. However, there is a critical need for further information on a range of factors that can have a significant impact on the cost of water produced by an MD desalination system, including pretreatment schemes, optimal process configurations, and long-term membrane performance.721 Finally, further work is required to develop cost calculation and comparison techniques to facilitate more reliable cost-based membrane and process optimization.

6 Conclusions and outlook

In this critical review, we have examined the key factors affecting the energy efficiency and robustness of MD desalination and detailed how future membrane design and process development could improve MD performance. The energy efficiency of MD, which exploits a liquid–vapor phase change to separate pure water from non-volatile solutes, is limited by the need to recover a high proportion of the latent heat of vaporization transferred from the feed stream to the distillate stream by the product water. Incomplete recovery of the latent heat, which is typically two or more orders of magnitude larger than the thermodynamic minimum energy of separation, dramatically reduces desalination efficiency. Consequently, MD and other phase-change based desalination processes, such as multi-effect distillation (MED) and multi-stage flash (MSF), are inherently less efficient than reverse osmosis (RO) and electrodialysis (ED), in which water remains in the liquid phase throughout. By analyzing previous literature and industrial data, we showed that system size can have a significant impact on energy efficiency. Currently, we find that MD is less efficient than MED and MSF for large-scale desalination systems with a water production rate greater than 1000 m³ day⁻¹. However, this trend is reversed for small-scale systems in which the number of stages used in MED and MSF is limited.

Membrane properties and process design can have a significant impact on the energy efficiency of MD. Using numerical modelling, we demonstrated that increasing membrane porosity and optimizing membrane thickness can drastically improve the energy efficiency of an MD desalination system. For direct-contact MD (DCMD), membrane thicknesses ranging from 70 μm to 100 μm achieve the optimal balance between maximizing the membrane permeability coefficient while minimizing the thermal conduction coefficient. Novel membrane architectures are required to achieve high porosities without compromising mechanical integrity or wetting resistance.

Process configuration plays a key role in maximizing latent heat recovery while ensuring that a consistent and sufficient partial vapor pressure driving force is maintained across the entire membrane area. Innovative process designs, such as permeate-gap MD, and variants including conductive-gap MD, have increased MD efficiency by directly capturing the heat released by the condensing product water to preheat the feed stream. Future system designs in which the feed stream is heated along the entire membrane module also have the potential to increase desalination efficiency by reducing the initial feed temperatures required to achieve a specified average water flux. Further investigation is required to determine the potential for multi-stage MD, which could be configured within a single membrane module, to further improve energetic performance.

Membrane wetting, which leads to a catastrophic loss of selectivity, limits the robustness of MD desalination and water treatment systems. Novel MD membranes with tailored surface wettability will play a key role in overcoming such limitations. For example, omniphobic membranes, which are able to resist wetting by both water and oil, can expand the application of MD to the treatment of challenging wastewaters containing low surface tension substances. The detrimental impact of fouling and scaling further constrains MD performance. Composite MD membranes comprising a hydrophilic layer on top of a hydrophobic substrate can significantly enhance oil and organic fouling resistance. Fouling and scaling are particularly problematic in the treatment of high-salinity wastewaters, which often contain a high concentration of potent foulants and scalants. Further innovation in membrane materials and surface modification is required to enhance membrane robustness for such
complex and challenging brine conditions. In addition, systematic understanding of membrane wetting, fouling, and scaling behaviors and underlying mechanisms will better guide strategies for high performance MD membrane design. Finally, current MD membranes should be evaluated in long-term operation to ensure consistent desalination performance across a wide range of practical applications.

MD is capable of treating hypersaline brines, with substantially greater salinities than the current RO treatment limit of around 80 g kg$^{-1}$, using low-grade or waste heat. These include brines from inland brackish water desalination and high-salinity wastewaters, such as produced water from unconventional hydrocarbon extraction. Low-grade heat sources, including low-temperature solar thermal collectors, geothermal reservoirs, and waste heat from power stations and industrial plants, can be cheaper and more sustainable than electricity or high-grade heat, both of which require more sophisticated infrastructure. However, future analyses should consider competing uses for low-grade heat sources, such as district heating and low-temperature electricity generation, when evaluating the potential for MD desalination and water treatment.

The inherent modularity and scalability of MD systems is highly advantageous in wastewater treatment applications where the flow rate and salinity of water fluctuate significantly. Similarly, relatively simple construction of MD desalination units, which do not require expensive materials or high-pressure process equipment, make them highly suitable for small-scale, off-grid applications. While the areal footprint of MD is low relative to MED and MSF, it remains high relative to RO. Future work is required to determine whether innovative MD process configurations are capable of reducing areal footprint further, by increasing average water fluxes at the module-level, without drastically lowering energy efficiency. In addition, coupled energy and cost analyses are imperative to identify applications in which MD desalination systems would provide clear economic benefits and the process parameters that are key to maximizing economic performance.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the NSF Nanosystems Engineering Research Center for Nanotechnology-Enabled Water Treatment (EEC-1449500). We thank the Agnese Nelms Haury Program in Environment and Social Justice at the University of Arizona for supporting Dr Vasiliki Karanikola, the Swiss National Science Foundation (P2SKP2-178166) for supporting Dr Anthony Straub through an Early Postdoctoral Mobility Fellowship, the National Science Foundation for supporting Prof Shihong Lin (CBET-1705048), and the Purdue School of Mechanical Engineering for supporting Dr David Warsinger.

References

111 H. Geng, H. Wu, P. Li and Q. He, Desalination, 2014, 334, 29–38.
118 W. Gao, H. Liang, J. Ma, M. Han, Z. Lin Chen, Z. Huang Han and G. Bai Li, Desalination, 2011, 272, 1–8.
206 R. Bond and S. V. Veerapaneni, American Water Works Association, 2008, **100**, 76.