Earth-Science Reviews 223 (2021) 103848

Contents lists available at ScienceDirect

Earth-Science Reviews

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

Striving to translate shale physics across ten orders of magnitude: What have we learned?

Yashar Mehmani [a,*], Timothy Anderson [b], Yuhang Wang [c], Saman A. Aryana [d], Ilenia Battiato [b], Hamdi A. Tchelepi [b], Anthony R. Kovscek [b]

[a] The Pennsylvania State University, Department of Energy and Mineral Engineering, PA, United States
[b] Stanford University, Department of Energy Resources Engineering, CA, United States
[c] Technische Universiteit Delft, Department of Geoscience & Engineering, The Netherlands
[d] University of Wyoming, Department of Chemical Engineering, WY, United States

ARTICLE INFO

Keywords:
Porous media
Shales
Scale translation
Data translation
Scale separation
Upscaling
Downscaling
Tyranny of scale
Homogenization
Hybrid methods
Numerical upscaling
Multiscale methods
Multiscale finite element
Multiscale finite volume
Mortar domain decomposition
Pore-level multiscale method
Pore network models
Multiscale imaging
Multimodal imaging
Machine learning

ABSTRACT

Shales will play an important role in the successful transition of energy from fossil-based resources to renewables in the coming decades. Aside from being a significant source of low-carbon intensity fuels, like natural gas, they also serve as geologic seals of subsurface formations that may be used to isolate nuclear waste, sequester CO_2, or store intermittent energy (e.g., solar hydrogen). Despite their importance, shales pose significant engineering and environmental challenges due to their nanoporous structure and extreme heterogeneity that spans at least ~10 orders of magnitude in spatial scale. Two challenges inhibit a system-level understanding: (1) the physics of fluid flow and phase behavior in shales are poorly understood due to the dominant molecular interactions between minerals and fluids under confinement, and (2) the apparent lack of scale separation that prevents a reliable (closed) description of the physics at any single scale of observation. In this review, we focus on the latter issue and discuss *scale translation*, which in its broadest sense is transforming data or simulations from one spatio-temporal scale to another. While effective scale translation is not exclusive to shales, but *all* geologic porous media, the need for it is especially acute in shales given their high degree of heterogeneity. Classical theories like homogenization, while indispensable, fail when scales are not separated. Other methods, like numerical upscaling, scale-translate in only one direction: small to large, but not the reverse, called *downscaling*. However, the confluence of advances in three areas are bringing challenging problems such as shales within reach: increased computational power and scalable algorithms; high-resolution imaging and multi-modal data acquisition; and machine learning to process massive amounts of data. While these advances equip geoscientists with a wide array of experimental and computational tools, no individual tool can probe the entire gamut of heterogeneity in shales. Their effective use, therefore, requires an ability to bridge between various data types obtained at different scales. The aim of this review is to present a coherent account of computational and experimental methods that may be used to achieve just that, i.e., to perform scale translation. We provide a broader definition of scale translation, one that transcends classical homogenization and upscaling methods, but is consistent with them and accommodates notions like downscaling and data translation. After a brief introduction to homogenization, we review hybrid methods, numerical upscaling and its recent extensions, multiscale computing, high-resolution imaging, and machine learning. We place particular emphasis on multiscale computing and propose an algorithmic framework to bridge between the pore (micro) and Darcy (macro) scales. Throughout the paper, we draw comparisons between the various methods and highlight their (often hidden) similarities, differences, benefits, and pitfalls. We finally conclude with two case studies on shales that exemplify some of the methods presented.

* Corresponding author.
E-mail address: yzm5192@psu.edu (Y. Mehmani).

https://doi.org/10.1016/j.earscirev.2021.103848
Received 2 November 2020; Received in revised form 29 October 2021; Accepted 31 October 2021
Available online 4 November 2021

1. Introduction

The coming decades are poised to undergo a transition away from fossil-based energy resources towards renewables, driven primarily by ecological and political pressures induced by climate change. The success of this transition, however, is contingent upon the sustainable production of low-carbon fuels from shales as well as the utilization of geological formations, sealed by shale, to sequester anthropogenic CO_2 and to potentially store renewable energy (e.g., hydrogen). During the past decade, shale formations have grown steadily in importance due to their significant contribution to the global energy supply and security. The low-carbon intensity fuels extracted from these formations, called shale gas and tight oil, have already displaced several carbon-intense fuels like coal for electricity generation. The impact has been especially dramatic in the U.S. Over the last decade, the U.S. has gone from scarce supplies of natural gas to abundance, and in 2018 became the world's largest producer of oil since the 1970s (Yergin, 2020). The International Energy Agency (IEA) has even projected that the U.S may become a net oil exporter within another decade (IEA, 2017). The result has been a marked reduction in the U.S carbon emissions and air pollution (Yergin, 2020).

Despite shale's promise, its engineering challenges are immense not only because the well-understood physics of conventional porous media are no longer applicable, but also because shales exhibit a much greater degree of heterogeneity than conventional reservoirs. The term "shale" refers loosely to ultrafine-grained rock types such as mudstones, marlstones, chalks, and others, which are nanoporous and exhibit chemical and structural heterogeneity at scales ranging from a few nanometers to several meters (Loucks et al., 2012) (Fig. 1); altogether accounting for 10 orders of magnitude. The nanoporosity renders shales nearly impermeable, and it is only through the advent of horizontal drilling and hydraulic fracturing that production from shales has been made possible. Their working principle is to increase the accessible surface area between wellbore fluids and the rock, and thereby production. Despite the success of these technologies, only about 5% of the original oil in place and about 25% of the gas in place is recovered, with production rates decreasing markedly after a few months (Patzek et al., 2013). Moreover, questions about the impact of production on the environment (e.g., fugitive methane; Howarth et al., 2011) and groundwater supplies remain open. Addressing these challenges is key to sustainable field operations (Hemminger et al., 2015), and it requires a fundamental understanding of the physics and chemistry of shales, as well as their manifestation at multiple length and time scales. To appreciate the difficulty, consider that the productivity of a given well is directly tied to the rate at which individual gas molecules desorb from the walls of a nanoscale pore and then diffuse into the nearest microcrack whose aperture is sensitive to geomechanical stresses and the presence of liquids. Other interactions between clays and water further confound the picture by introducing swelling (Wang et al., 2017) and mineral reactions (Harrison et al., 2017).

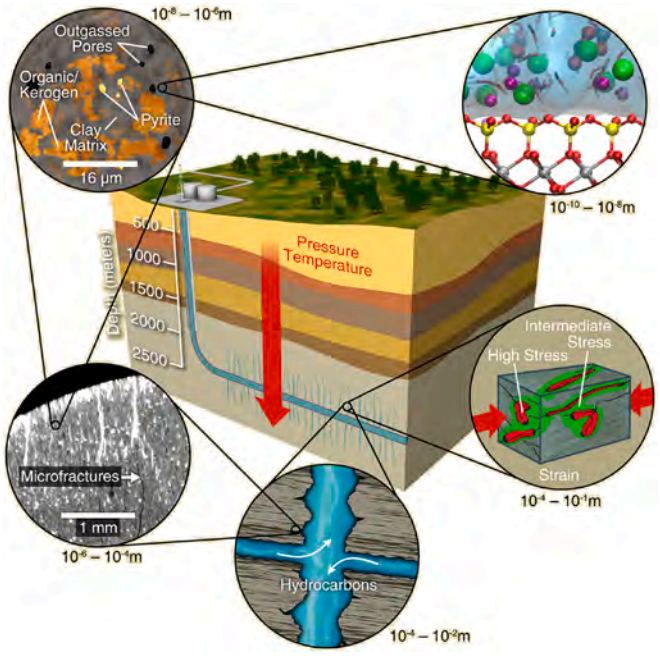

Fig. 1. Schematic of relevant length scales in shale. From the upper right: water/clay interactions in a 10 nm wide slit pore where water is shaded in blue; nano-CT image of shale structure illustrating minerals, kerogen, and outgassed pores; microfractures filled with barite scale; matrix-to-fracture mass transfer; and zones of enhanced ductility.

Nanopores are comparable in size to fluid molecules, rendering classical continuum (or Darcy-scale) descriptions of fluid dynamics and phase behavior invalid; most of which also neglect key molecular forces between minerals and fluids (Jin and Firoozabadi, 2016). What is more, shales seem to lack a clear separation between scales. This makes describing (or closing) the physics at a single scale of observation nearly impossible (Section 2.1). The above challenges can be recast into two broad questions: (1) how do we describe the nanoscale physics and chemistry of shales? And (2) how do we translate such knowledge across spatiotemporal scales? This review explores the latter question, while the former will be addressed in separate publications (Khan et al., 2021; Jew et al., 2020).

We distinguish the "scale" question from the "physics" question because the former is not exclusive to shales, but all geologic porous media. *Scale translation*, which we define as transforming data or information from one observation scale to another, has been a long-standing challenge in geosciences. Shale development has just stirred a more acute need for it. A massive body of literature already exists on scale translation, describing classical methods such as homogenization (Hornung, 1996; Whitaker, 2013; Gray and Miller, 2014) and numerical upscaling (Durlofsky and Chen, 2012; Farmer, 2002; Renard and de Marsily, 1997; Christie, 1996) that allow small-scale measurements or simulations to be transformed into large-scale decisions about the development of a petroleum reservoir or the fate of a contaminant plume. But these methods, while indispensable, have limitations. For one, they translate geologic information in a unidirectional fashion: from small to large, called *upscaling*. Another is that they require spatiotemporal scales associated with structural/chemical heterogeneities of the porous medium, as well as the physical/chemical processes occurring in them, to be separated, which may not hold for shales (Section 2.1).

What is promising about today is that significant progress has been made in three areas since the turn of the century: (1) increase in computational power and the development of algorithms that are scalable on parallel machines, (2) high-resolution imaging instruments (e.g., hyperspectral, X-ray micro/nano-CT) that generate massive amounts of data, and (3) the development of machine learning methods able to recognize patterns among such data. However, these methods have largely evolved in isolation and while geoscientists now have a wider array of experimental and computational tools at their disposal, no single tool can probe the full gamut of heterogeneity present in shales. The gap must, therefore, be filled by an ability to translate one data type into another, and from one spatiotemporal scale to another. As we will describe later, such tasks require not only upscaling, but also *downscaling*, which converts large-scale data into small-scale information; something traditional methods do not address. We believe it is only through such a combination of multi-modal data acquisition and high-powered computation and pattern recognition that a full picture of shales, among other challenging geomaterials, can be constructed.

The aim of this review is to place the foregoing new developments within the context of the more traditional approaches for scale translation and to highlight how they complement, rather than conflict with,

each other. In Section 2, we generalize the definitions of well-established concepts such as *scale translation*, *scale separation*, *tyrannies of prediction*, *upscaling*, *downscaling*, and *data translation*. The goal here is to provide the necessary background for later sections and to broaden the classical definitions used in conventional porous media. In Section 3.1, we provide a brief, but self-contained, review of analytical *homogenization* methods. They consist of mathematical techniques used to derive macroscopic governing equations from their microscopic counterparts. While we discuss several such techniques, we highlight that *all* of them require "scale separation" as a central assumption. Understanding them, however, is crucial for making appropriate use of the computational methods presented next. In Section 3.2, we discuss *hybrid methods*, which aim to model porous-media problems when scale separation is absent. In Section 3.3, we review *numerical upscaling* and its more recent extensions. A discussion of the upsides and pitfalls of numerical upscaling, and the potential avenues for improvement, is presented. Section 3.4 discusses *multiscale computing*, to which a large portion of this review is devoted. First, we discuss older techniques that were developed purely at the Darcy scale, such as multiscale finite element (MsFE), mixed multiscale finite element (MxMsFE), multiscale finite volume (MsFV), and mortar multiscale finite element (MoMsFE). We then discuss straightforward extensions of these methods to the pore (or micro) scale, followed by a presentation of more recent and specialized methods for solving pore-scale problems.

Next, we discuss traditional pore-scale modeling approaches, like pore network models (PNM), and formalize their algorithms by placing them within the context of numerical upscaling and multiscale computing. The main reason for our emphasis on multiscale methods in this paper is that not only do they bear many similarities with hybrid *and* numerical upscaling methods, as outlined throughout the paper, but they also possess additional properties that are computationally attractive such as the ability to downscale and quantify prediction errors. In Section 3.4.4, we propose a new algorithmic framework for bridging between pore- and Darcy-scale physics that combines several of the preceding methods. The anticipated limitations of the framework are also detailed. In Section 3.5, we discuss recent advances and trends in high-resolution and multi-modal imaging that are changing how geologic porous media are being characterized. Such images serve as crucial inputs to either the computational methods covered earlier or the machine learning algorithms discussed in the following Section 3.6.

The exposition style of this paper is pedagogical, and while mathematical details are certainly discussed, they are done so at a sufficiently high level so as not to distract from the main points. We use simple examples to convey the algorithmic details of each method presented. While the governing equations used in the examples describe *conventional* porous media (not shales), the reader should note that the algorithms themselves remain unaltered for shales. The only difference lies in the specific differential operators describing shale physics to which the algorithms would be applied. The validity of these operators, however, is a separate "physics" question actively being explored (Center for Mechanistic Control of Unconventional Formations (CMC-UF)). The paper concludes with Section 4, where we present two case studies that exemplify some of the methods presented.

2. Definitions

2.1. Scale translation

We define *scale translation* as the process of using data at one spatiotemporal scale to infer needed information at another scale. An example is to use pore-scale data, such as X-ray μCT images, to derive Darcy-scale data, such as permeability (Wildenschild and Sheppard, 2013; Blunt et al., 2013). Another is to use core-scale measurements of organic content to reconstruct millimeter-scale variabilities of thermal conductivity (Mehmani et al., 2016a). Scale translation is bidirectional. If fine-scale data are used to obtain coarse-scale information, the process is called *upscaling* (Section 2.4). By contrast, if coarse-scale data are used to obtain fine-scale information, the process is known as *downscaling* (Section 2.5). In geosciences, physicochemical processes such as flow, transport, geochemistry, mechanical deformation and fracturing occur at the scale of individual grains that comprise a rock. Such processes are typically fast and characterized by short time scales. Despite their microscopic origin, subsurface physics exhibit manifestations at *all* spatiotemporal scales relevant to a geologic formation. Microscale processes occurring within individual pores conspire to produce emergent behaviors at the reservoir scale. Gravity fingering in CO_2 storage is one example, where dissolution dynamics at the pore scale cause instabilities that grow into meter-scale convection columns (Lindeberg and Wessel-Berg, 1997). Another is mixing-induced mineral precipitation, in which a kilometer-scale plume is inhibited from further mixing due to pore-scale mineralization at the plume fringes (Zhang et al., 2010; Johnson et al., 2004). In engineering the subsurface, predictions and forecasts are frequently required at the scale of an entire reservoir (kilometers) or a single well-bore (meters). The data collected about a reservoir, however, can range anywhere from pore-scale images, such as SEM and EDS, to inter-well seismic or tracer experiments. Scale translation is a crucial part of subsurface engineering precisely because of the disparity between the scale of data acquisition and the scale of demand for information. It is frequently thought that upscaling is the primary vehicle for converting lab-scale data into useful reservoir-scale decisions. We show later that downscaling is just as, if not more, important. In Section 2.5, downscaling is discussed as a key step for practical scale translation.

The advent of multiple technologies in the recent decade have created a perfect storm for an unprecedented characterization and forecast of subsurface physics. They include high-resolution micro/macro-imaging (X-ray μ/n-CT, FIB-SEM, hyperspectral), computer hardware and powerful parallel machines, scalable multiscale algorithms, and machine learning techniques to encode patterns among data. The definition of scale translation must therefore evolve accordingly, beyond the engrained notion of upscaling, if these technologies are to provide any benefit other than the isolated probing of geologic samples at an instrument's acquisition scale. While the mathematically formal notion of homogenization (Section 3.1) remains indispensable, it only provides a learning tool for understanding subsurface physics in idealized media (scale separable, periodic). Rocks, and in particular shales, are messy and do not adhere to such theoretical conveniences. Geologic data are acquired in many forms and at different scales, probing the formation with acoustic, neutron, X-ray, tracer, PET, electron beam, and other signals. In isolation, none provides a complete picture of the geology at all scales. The modern definition of scale translation must recognize this practical constraint and utilize tools that map and integrate data from multiple sources and scales. We call the mapping of data from one form to another *data translation*, which is a key component of effective scale translation. In Section 2.5, we show that data translation itself requires not only upscaling but also downscaling.

2.2. Scale separation

Geologic porous media have complex microstructure. Rocks consist of solid and void regions separated by a highly irregular boundary. The void provides a conduit for fluid flow while the solid skeleton bears the load of the overburden and undergoes mechanical deformation. Despite the microstructural irregularity, some porous media exhibit regularity when probed at sufficiently large spatial scales. Put differently, a microscopically heterogeneous porous medium may behave as a macroscopically homogeneous domain. A similar phenomenon may occur with *time*. A physicochemical process marked by rapid and noisy dynamics, such as immiscible drainage, can appear slow and smooth over longer time scales. When this happens, we say that the microscale and the macroscale physics are *separable*. A classic example is porosity (Bear, 2013). Imagine that we calculate the porosity of a sample

contained within a sphere of radius r centered around x. When r is small, the sphere samples either the void or the solid and the calculated porosity is either one or zero. As r is increased, more of the void or the solid are sampled and porosity fluctuates. At very large r, porosity may converge to a single value. If it does, scale separation holds for porosity. The smallest r at which porosity converges (up to a tolerance) defines a representative elementary volume (REV) (Bear, 2013). We denote the corresponding r as r_{REV}. Had we chosen a parameter other than porosity, such as permeability, we would have made a similar observation. Permeability fluctuates at small r and then stabilizes at large r. However, the r_{REV} for porosity is not necessarily the same as that for permeability. Nor is r_{REV} for permeability the same as that for hydrodynamic dispersion coefficient or effective geochemical reaction rate. In other words, r_{REV} depends on the physicochemical quantity of interest. Moreover, there is no guarantee that r_{REV} is even bounded. An unbounded r_{REV} (equal to infinity) means that scales are *not* separable. The REV concept is not limited to static properties that depend on the fine-scale geometry of a porous medium, such as porosity. It also extends to dynamic (or time dependent) variables, such as concentration, phase saturation, and pressure. For such variables, r_{REV} is generally a function of the dimensionless numbers that control the underlying physical process. In unstable displacement of one immiscible fluid by another, for example, the characteristic length associated with the size of each viscous finger can be orders of magnitude larger than the characteristic length associated with the microgeometry (Tomin and Lunati, 2016a). As the viscosity ratio between the two fluids, a dimensionless number controlling the process, is varied, the size of the fingers, and thus r_{REV}, also changes. The above arguments for r_{REV} extend quite naturally to *time*, where the shortest interval at which temporal fluctuations dissipate is denoted by t_{REV}. If t_{REV} is unbounded, then the process to which it corresponds is said to exhibit "memory" and is thus not scale-separable in time. In Section 3.1, we make the definition of scale separation mathematically more precise.

The existence of a bounded r_{REV} or t_{REV}, and thus scale separation, is a key assumption in all homogenization methods discussed in Section 3.1. How well does it hold for geologic media? The answer depends on the local lithology and the burial history of a formation. But at least for organic-rich shales, there is evidence that the assumption may be violated. Fig. 2 shows semi-variograms of the vertical variability of organic content (weight fraction) for three wells in the Green River Formation, USA. If scale separation held, the variograms would converge to a single value beyond a minimum separation distance. Instead, the variograms in Fig. 2 never converge and increase indefinitely with distance. While local "sills" do seem to appear at some intervals, they are most likely due to "small-sampling" effects. Taken together, the three wells exhibit large fluctuations that suggest each variogram is not statistically "converged".

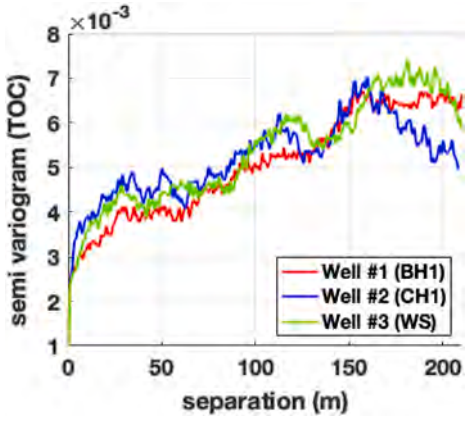

Fig. 2. Semi-variograms of the total organic content (weight fraction) computed for three wells in the Green River formation, USA. The rock is an immature organic-rich shale. Semi-variograms are computed in the vertical direction. A clear separation of scales for the entire probed thickness (~200 m) seems speculative at best. [Plot was produced using well data courtesy of American Shale Oil LLC (AMSO).]

2.3. Tyrannies of prediction

Two challenges have stymied practical scale translation:

1. **Tyranny of scales:** This is the classical difficulty of describing multiscale systems, such as the subsurface, that are governed by processes occurring at a multitude of spatial and temporal scales. The disparity between the scale at which information is needed and the scale at which a phenomenon is studied or understood engenders the challenge for prediction. The difficulty was stated explicitly in the 2006 NSF report (NSF, 2006) in the context of simulation.
2. **Tyranny of characterization:** This is a well-known difficulty specific to geosciences. The tyranny of characterization is a manifestation of the unknowability, due to inaccessibility, of the subsurface. Data are sparse, which requires extrapolation beyond what is measured. When the microstructure is probed, by an X-ray microscope for example, the field of view must shrink, and the larger picture is lost. Conversely, when field-scale data are collected, microscale information are blurred. The implication: mapping the microstructure at the field scale is not feasible, at least deterministically.

While the two tyrannies feed off of each other in making geologic predictions difficult, it is important to keep their contributions separate. The following thought experiment illustrates how. The tyranny of scales says it is hard to come up with a mathematical description of a phenomenon that is purely confined to one scale and needs not be informed of what is happening at other scales. This holds *even if* we had a perfect description of the subsurface down to its intricate microstructure. For example, to describe how large-scale fractures emerge at the field scale, something must be known about how tiny cracks nucleate and coalesce at the micron scale. The degree to which such information is needed from other scales depends on the physics under study. The best-case scenario is when a set of equations can be formulated at a scale of interest that are predictive if the parameters in the equations can be adjusted to capture the physics at some other scale. A worse-case scenario is when single-scale equations can be formulated only partially or not at all. Then, much more than just parameters (whole differential operators describing solute fluxes for example) must be informed by data from other scales. While the tyranny of scales applies to *any* one scale at which a description is sought, it is much more common in geosciences that microscale descriptions are better-understood and more readily available than macroscale descriptions. Because modeling microscale physics over an entire geologic formation is outside the realm of possibility (and interest) for now, single-scale descriptions are frequently sought to capture macroscale physics.

The tyranny of characterization *adds* to this difficulty by saying that even if we had the computational power to solve microscale equations over an entire geologic formation, we still could not. Because the parameters and geometric descriptors needed to solve the equations would simply be unknown over much of the formation. Access points to a formation are limited to wells (and with lesser quality, to outcrops), which are the only sources of high-resolution data like cores, logs, thin sections, and microscopy images. The vast volumes between the wells can only be probed by lower-resolution instruments such as crosswell seismic and resistivity imaging (Neal and Krohn, 2012). A very similar problem exists also in the lab. When imaging a sample with an electron microscope, for example, one must trade off resolution for field of view. This introduces an intrinsic uncertainty to all large-scale predictions and necessitates some degree of extrapolation. What is different today, compared to past decades, is we have a lot more data and better

extrapolation tools.

2.4. Upscaling

Upscaling is the process of compressing fine-scale data in order to deduce coarse-scale information. A simple example is to replace a heterogeneous fine-scale permeability field with a single coarse-scale value, such that the two systems are equivalent in their total fluid throughput. The importance of upscaling in geosciences arises from the disparity between the scale at which geologic systems are understood (pore scale) versus the scale at which engineering decisions are made (field scale). The gap between the two is $O(10^{8-9})$ in conventional reservoirs and even larger in shales $O(10^{9-12})$. Upscaling attempts to bridge this gap by replacing a detailed representation of a reservoir or lab-scale sample with a much simpler description, characterized by fewer degrees of freedom. However, the replacement must preserve certain coarse-scale characteristics of the original system (e.g., energy dissipation). As a result, upscaling depends on the specific "microstructure", or fine-scale heterogeneity, of the domain. For a horizontally layered medium, the upscaled permeability in the horizontal direction is the arithmetic mean of the layer permeabilities. For a vertically layered medium, coarse-scale permeability is the harmonic mean. Numerous studies have demonstrated the intimate relation between microstructure and upscaled properties (Renard and de Marsily, 1997).

In this work, we use the term "upscaling" in a much broader sense than the simple parameter estimation problem discussed above. We define upscaling as the process of obtaining *any* useful information at a scale larger than the one corresponding to the data used as input. The upscaled information may be: (i) a coarse-scale equation derived and parameterized through either homogenization (Section 3.1) or numerical upscaling (Section 3.2), (ii) an algorithmic construct, such as a flux matrix (Section 3.4.4), that represents a fine-scale equation as a lower-dimensional matrix-vector multiplication, or (iii) a machine learned representation between fine-scale data, such as X-ray μCT images, and coarse-scale parameters, such as permeability (Andrew, 2020) (Section 3.6). The utility of such an expansive definition provides a unified framework for studying different methods and drawing useful, and potentially synergistic, comparisons between them.

2.5. Downscaling and data translation

Downscaling is the process of deducing fine-scale information from coarse-scale data. It is the inverse of upscaling, in which fine-scale data are combined to obtain coarse-scale information. Downscaling is more difficult than upscaling because the problem is underdetermined. Fine-scale variabilities reside in a higher dimensional mathematical space than do coarse-scale variabilities. A knowledge of coarse-scale heterogeneity alone is not sufficient to reconstruct the fine-scale heterogeneity. There are two ways with which downscaling can be constrained. The first is to make assumptions about the nature of the fine-scale variability, such as stationarity and Gaussian statistics. Such assumptions are common in geosciences, especially with permeability (Dagan, 1989). They render downscaling into a mathematically well-posed problem but can produce unrealistic reconstructions of the fine scale. The flaw with the approach is that it presupposes how the fine scale should look like without sufficient proof of it being true. The second way to constrain downscaling is to utilize fine-scale measurements of a property other than that being downscaled. An example is to use fine-scale measurements of mineralogical composition of a rock to downscale the spatial variability of thermal conductivity (Mehmani et al., 2016a). For such a method to work, the fine-scale relationship between composition and conductivity must be known or measured beforehand. The approach is very general as no assumptions about the nature of the fine-scale variability (e.g., Gaussianity) are needed. The approach also extends the utility of downscaling to *data translation*, which we now demonstrate with a simple example.

Consider the two-layered domain in Fig. 3b. Layers 1 (top) and 2 (bottom) have thicknesses l_1 and l_2, respectively. Let E and K be two material properties, such as Young's modulus and thermal conductivity, respectively. We use superscripts f and c to denote fine-scale and coarse-scale data. The fine-scale relationship $K^f = f(E^f)$ is generally nonlinear and assumed here to be known from experiments, as shown in Fig. 3a. E^c and K^c correspond to coarse-scale (or averaged) properties in the direction parallel to the layers. Suppose we are given E^c and are asked to compute K^c, i.e., data translation. Since $f(.)$ is nonlinear, we cannot use it directly to map E^c to K^c, i.e., $K^c \neq f(E^c)$. For the layered domain in Fig. 3b, E^c and K^c correspond to the arithmetic averages of the layer properties:

$$E^c = \frac{l_1 E_1^f + l_2 E_2^f}{l_1 + l_2} \qquad K^c = \frac{l_1 K_1^f + l_2 K_2^f}{l_1 + l_2} \tag{1}$$

Graphically, E^c and K^c lie on the secant that connects the layer properties (E_1^f, K_1^f) and (E_2^f, K_2^f) in Fig. 3a (dashed line). Observe that $K^c \neq f(E^c)$. To compute K^c accurately, three steps are required: (1) map the fine-scale value of E^f for each layer, (2) use $K^f = f(E^f)$ in Fig. 3a to compute K^f for each layer, and (3) upscale K^f with Eq. (1) to obtain K^c. Step (1) amounts to downscaling E. The foregoing approach for translating E^c into K^c may be generalized to other petrophysical properties and to heterogeneities much more complicated than Fig. 3b. Mehmani et al. (2016a, 2016b) used it to translate total organic carbon content to the thermal conductivity of organic-rich shales. Step (1) is the most difficult of the three and may be accomplished in several ways. The most common, by far, is to use an imaging instrument such as an X-ray μCT (Wildenschild and Sheppard, 2013) scanner or a hyperspectral camera (Mehmani et al., 2017; Lyder et al., 2010; Speta et al., 2015). By comparison, step (3) is more straightforward as any number of computational upscaling methods, described in Sections 3.3–4, can be used to compute K^c (instead of Eq. (1)).

The overarching idea in data translation is to deduce hard-to-map data, such as permeability, from easy-to-map data, such as porosity. For example, porosity can be easily mapped at the meter scale using fast and non-intrusive X-ray CT images, which correlate strongly with bulk density and thus porosity (Honarpour et al., 1985). Permeability, by contrast, requires more work, either in the form of flow experiments or computationally expensive pore-scale simulations. Neither is able to produce a continuous map of permeability, as each point in the meter-scale domain must be processed in sequence. But if a *fine-scale* relationship between porosity and permeability can be established for a specific geomaterial, then porosity maps can be translated into permeability maps. The coarse-scale permeability follows after a straightforward upscaling step. Downscaling is a crucial step in data translation because without it, the relationship between K^c and E^c would be riddled

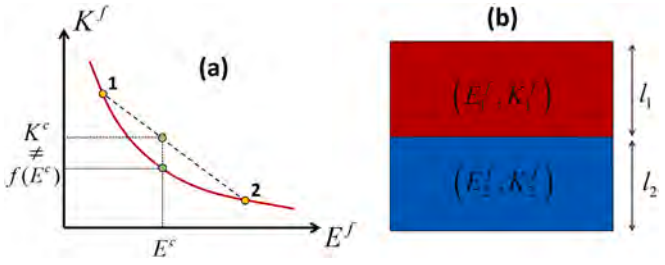

Fig. 3. The importance of downscaling in data translation. (a) Functional relationship $K^f = f(E^f)$ between two fine-scale properties, K^f and E^f. (b) A domain with a two-layer structure at the fine scale. Layers 1 and 2 have fine-scale properties (E_1^f, K_1^f) and (E_2^f, K_2^f), respectively. The coarse-scale (or averaged) properties of the domain are (E^c, K^c). Suppose E^c is known and we want to deduce K^c. If we estimate K^c as $f(E^c)$, using the fine-scale relationship, we incur errors because the function $f(.)$ is nonlinear, i.e., $K^c \neq f(E^c)$. To compute K^c accurately, we must resolve the fine-scale variability of E^f in the domain. Namely, we must map E^f for each layer, use $f(.)$ to compute K^f, and then upscale K^f to obtain K^c. The first step amounts to downscaling.

with "noise" induced by the unresolved fine-scale variability. The observation is ubiquitous in well-log analysis, where permeability is deduced from neutron or density logs of porosity (Ellis and Singer, 2007). The large scatter in such permeability-porosity relations is because logging instruments have a resolution of a few feet, which is much larger than the typical fine-scale correlation length of sedimentary rocks.

3. Methods

3.1. Homogenization

Homogenization is the process of deriving governing equations that apply at the coarse scale from those that apply at the fine scale. The designations "coarse scale" and "fine scale" are relative to each other but arbitrary in absolute terms. For example, they may correspond, respectively, to the range O(10 μm) and O(1 mm) or O(1 cm) and O(10 m). Despite the arbitrariness, a common choice for the fine scale is the *pore scale*, which is the scale at which the porous medium appears as an aggregate of discrete objects and features (e.g., grains, cracks, vugs, fibers). The corresponding coarse scale, where fluctuations associated with such features dissipate, depends on the porous medium and its microstructure. Homogenization is motivated by the fact that porous media (esp. geologic) are often too complex to be modeled at the fine scale and may require prohibitive amounts of computational resources to make predictions. Therefore, instead of resolving every little detail, which may also be redundant, a continuum representation of the physics is sought that effectively reduces the total number of degrees of freedom required to model the problem. Interestingly, but perhaps unsurprisingly, continuum representations of porous media predate the development of most homogenization methods. Continuum equations were originally postulated rather than derived from first principles, a classic example of which is Darcy's law (Darcy, 1856) and its many subsequent derivations (Hornung, 1996; Neuman, 1977; Whitaker, 1986; Tartar, 1980; Keller, 1980). The literature on homogenization is vast and the available tools and methods for it diverse. Our goal is not to provide a comprehensive review of each method, which would be an insurmountable effort for one paper, but to provide a simple pedagogical illustration of what homogenization means and how it differs from, and motivates, the computational methods in later sections. The reader is referred to the review articles (Battiato et al., 2019; Cushman et al., 2002; Davit et al., 2013) and books cited herein for a more detailed exposition. We note that the terminology used in the literature to convey "homogenization" varies widely and includes "upscaling", "coarse graining", and "averaging". We use "homogenization" to differentiate it from other concepts introduced in later sections.

Several homogenization techniques exist for porous media problems, including the method of volume averaging (Whitaker, 2013; Quintard and Whitaker, 1988), multiple scales (or matched asymptotic) expansion (Hornung, 1996; Keller, 1980; Mikelić et al., 2006; Auriault and Adler, 1995; Bensoussan et al., 2011), stochastic averaging (Dagan, 1989; Rubin, 2003; Zhang, 2001), thermodynamically constrained averaging (Gray and Miller, 2014; Gray et al., 2013), and hybrid mixture theory (Bennethum and Cushman, 1996a; Bennethum and Cushman, 1996b; Achanta et al., 1994; Hassanizadeh and Gray, 1990; Marle, 1982) to name a few (see Cushman et al. (2002) for more). In all of them, the starting point is a set of fine-scale equations that are averaged in some sense to obtain coarse-scale governing equations. In the following, we focus on the method of volume averaging, as it is the most intuitive, and proceed by an example to illustrate its various steps. At the end, we provide a brief description of some of the other methods.

Consider the porous domain Ω depicted in Fig. 4a. Suppose the fine-scale governing equation on Ω has the following general form

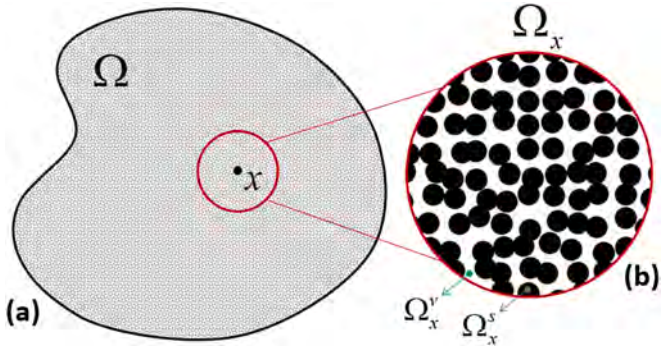

Fig. 4. Schematic of homogenization in porous media. (a) A macroscopic (or large) porous medium, Ω, and (b) its microscopic (or pore-scale) details within a ball Ω_x centered around the point x. In (b), black regions are the solid phase, Ω_x^s, and white regions the void space, Ω_x^v. The goal of homogenization is to derive continuum equations on Ω by "averaging" the details in Ω_x for all x. In the method of volume averaging, the "averaging" is literal. An example is homogenizing the Stokes equation on Ω_x^v to obtain the Darcy equation at x.

$$L_f(\boldsymbol{u};\varsigma_f) = f \tag{2}$$

where $L_f(\cdot)$ is a *linear* fine-scale differential operator, $\boldsymbol{u}$ is the unknown field variable to be solved, and f is the right-hand side forcing term. The symbol ς_f in $L(\boldsymbol{u};\varsigma_f)$ denotes the set of known parameters in the fine-scale problem. As an example, Eq. (2) may correspond to the Stokes flow equation, which describes the mass and momentum conservation of a single-phase Newtonian fluid under isothermal conditions, low Reynolds numbers, and steady state. For this case, $L_f = [-\nabla, \mu\Delta; 0, \nabla\cdot]$ in matrix notation, $\boldsymbol{u} = [p, u]^T$, and $\varsigma_f = \mu$, where Δ, ∇, and $\nabla\cdot$ are, respectively, the Laplace, gradient, and divergence operators. The parameter μ is the fluid viscosity, and p and u are the fluid pressure and velocity, respectively. In Fig. 4, the fine scale corresponds to the pore scale, which is comprised of a solid phase (black) and a void space (white). The Stokes equation is defined over the latter. The goal of the volume averaging method is to derive a coarse-scale equation of the following form

$$L_c(\langle\boldsymbol{u}\rangle;\varsigma_c) = \langle f\rangle \tag{3}$$

where the subscript c in L_c and ς_c means that these entities are defined at the coarse scale. In general, L_c and ς_c differ from their fine-scale counterparts L_f and ς_f. The angular brackets around $\boldsymbol{u}$ and f denote their spatial averages over some predefined support volume Ω_x. Namely,

$$\langle\cdot\rangle|_x = \int_{\Omega_x} \cdot\omega(y)dy \tag{4}$$

where Ω_x is the averaging volume centered at point x, as shown by the ball in Fig. 4b. The averaging kernel $\omega(y)$ is defined on Ω_x and integrates to a constant (i.e., has finite measure; usually unity or $|\Omega_x|^{-1}$). Typically, $\omega(y)$ is the indicator function. The averaged quantities $<\boldsymbol{u}>$ and $<f>$ are coarse-scale field variables, which are defined at each point x of the domain Ω (Fig. 4a). The whole idea behind volume averaging is to convolve $\omega(y)$ with the fine-scale Eq. (2) in order to obtain Eq. (3). This means that one must use Eq. (4) to operate $<\bullet>$ on $L_f(\boldsymbol{u};\varsigma_f) = f$, which yields

$$\langle L_f(\boldsymbol{u};\varsigma_f)\rangle = \langle f\rangle \tag{5}$$

The challenge now is to somehow approximate $<L_f(\boldsymbol{u};\varsigma_f)>$ with $L_c(<\boldsymbol{u}>;\varsigma_c)$ through appropriate definitions of L_c and ς_c. It is crucial to notice here is that $L_c(<\boldsymbol{u}>;\varsigma_c)$, or alternatively Eq. (3), is devoid of any

explicit dependence on $\boldsymbol{u}$ (only $<\boldsymbol{u}>$). This means that to solve Eq. (3), we need not know anything about the fine-scale variability of $\boldsymbol{u}$. Instead, we can solve an equation that is entirely in terms of $<\boldsymbol{u}>$. But how do we get from $<L_f(\boldsymbol{u};\varsigma_f)>$ to $L_c(<\boldsymbol{u}>;\varsigma_c)$? The first step is to use a number of *averaging theorems*, which establish the rules for how integration and differentiation can be commuted. For a comprehensive list of such theorems we refer the reader to Gray et al. (1993), although a very small subset is frequently used in practice (Whitaker, 2013). The next step involves a series of simplifications that eliminate terms that depend explicitly on $\boldsymbol{u}$. Common simplifications include the following assumptions: scale separation (Ω_x is large with respect to pore-scale features in Fig. 4b, i.e., it is an REV), ergodicity (ensemble averages equal spatial averages), stationarity (statistics of the fine-scale variabilities are invariant with $\boldsymbol{x}$), smooth gradients of $\boldsymbol{u}$ (no sharp boundary layers or transition zones exist, e.g., transport is diffusion- rather than advection-dominated). Most of these assumptions are common to other homogenization methods mentioned earlier. The third step is to define the fine-scale deviation field variable as follows

$$\widetilde{\boldsymbol{u}} = \boldsymbol{u}\text{-}\langle \boldsymbol{u} \rangle \tag{6}$$

and substituting it into the simplified form of Eq. (5) obtained from the previous step. The goal is to merely express any remaining terms that depend on $\boldsymbol{u}$ in terms of $\widetilde{\boldsymbol{u}}$. The final step is to formulate a differential equation whose solution is $\widetilde{\boldsymbol{u}}$. Substituting this solution into the averaged equation from step 3 yields Eq. (3) and thus concludes the method.

The crucial step in the above procedure is the formulation and solution of the governing equation for $\widetilde{\boldsymbol{u}}$. To formulate it, one merely needs to subtract the simplified averaged equation obtained from step 3 from the fine-scale Eq. (2). The result often looks like the following

$$L_f\left(\widetilde{\boldsymbol{u}};\varsigma_f\right) = L_s(\langle \boldsymbol{u} \rangle, \nabla\langle \boldsymbol{u} \rangle) \tag{7}$$

where $L_s(\cdot,\cdot)$ is a *linear* operator in terms of coarse-scale arguments $<\boldsymbol{u}>$ and $\nabla<\boldsymbol{u}>$. In general, L_s can depend on any other derivatives of $<\boldsymbol{u}>$ specific to the problem at hand. We have chosen $<\boldsymbol{u}>$ and $\nabla<\boldsymbol{u}>$ as examples. The left-hand side operator of Eq. (7) is identical to that of the fine-scale Eq. (2), which need not be, but is often, the case. Eq. (7) means that the fine-scale deviation $\widetilde{\boldsymbol{u}}$ in the vicinity of $\boldsymbol{x}$ (i.e., inside Ω_x) is controlled by a forcing (or source) term that depends linearly on $<\boldsymbol{u}>$ and $\nabla<\boldsymbol{u}>$. Let us assume for the moment that $<\boldsymbol{u}>$ and $\nabla<\boldsymbol{u}>$ are known and that we want to compute $\widetilde{\boldsymbol{u}}$. Does Eq. (7) mean we must solve a different fine-scale equation on Ω_x for every $\boldsymbol{x}$? The answer is *no*, but only if L_f and L_s are both linear. Because then we can use the principle of superposition to solve only one *fundamental solution* of Eq. (7) and then scale it to obtain $\widetilde{\boldsymbol{u}}$ on Ω_x for every $\boldsymbol{x}$. The fundamental solution is called the *closure variable* and the differential equation associated with it the *closure problem*. More precisely, the closure problem is

$$L_b\left(\boldsymbol{b};\varsigma_f\right) = \begin{bmatrix} L_f\left(\boldsymbol{b}_1;\varsigma_f\right) \\ L_f\left(\boldsymbol{b}_2;\varsigma_f\right) \end{bmatrix} = \begin{bmatrix} L_s(1,0) \\ L_s(0,1) \end{bmatrix} = f_b \tag{8}$$

where $\boldsymbol{b} = [\boldsymbol{b}_1, \boldsymbol{b}_2]$ denotes the closure variable and $L_b(\cdot)$ the closure operator. Given the solution of Eq. (8), the deviation field $\widetilde{\boldsymbol{u}}$ can be reconstructed using the following equation

$$\widetilde{\boldsymbol{u}} = \langle \boldsymbol{u} \rangle \boldsymbol{b}_1 + \nabla\langle \boldsymbol{u} \rangle \cdot \boldsymbol{b}_2 \tag{9}$$

in the neighborhood of any point $\boldsymbol{x}$ (i.e., Ω_x). It is easy to verify that Eq. (9) satisfies Eq. (7)

$$L_f(\widetilde{\boldsymbol{u}}) = \langle \boldsymbol{u} \rangle L_f(\boldsymbol{b}_1) + \nabla\langle \boldsymbol{u} \rangle \cdot L_f(\boldsymbol{b}_2) = \langle \boldsymbol{u} \rangle L_s(1,0) + \nabla\langle \boldsymbol{u} \rangle \cdot L_s(0,1) = L_s(\langle \boldsymbol{u} \rangle, \nabla\langle \boldsymbol{u} \rangle) \tag{10}$$

where the linearity of both L_s and L_f are used. Substituting Eq. (9) into the simplified coarse-scale equation obtained at the end of step 3 above yields the final coarse-scale Eq. (3), which is devoid of any dependencies on $\widetilde{\boldsymbol{u}}$. The coarse-scale operator L_c and the parameter ς_c are determined after some rearrangements. We note that the parameter ς_c is a function of the (averaged) closure variable $\boldsymbol{b}$. The closure problem, Eq. (8), itself is defined at the fine scale and must be solved on a representative elementary volume (REV) of Ω. This REV coincides with Ω_x for some arbitrary $\boldsymbol{x}$.

A few remarks are now in order. Eq. (8) is not the only way to formulate a closure problem, but it is quite general and illustrates the main point. In the above example, the closure problem consists of two sub-problems: one for $\boldsymbol{b}_1$ and another for $\boldsymbol{b}_2$. The boundary conditions of Eq. (8) (not shown) may also depend on $<\boldsymbol{u}>$ or its derivatives, in which case one must construct yet another "basis" or fundamental solution (e.g., $\boldsymbol{b}_3$) that accounts for such non-homogeneities. When $\boldsymbol{u}$ consists of more than one component and $L_s(\cdot,\cdot)$ is a function of component-wise derivatives of $\boldsymbol{u}$, Eq. (8) can yield a redundant number of closure sub-problems. An example is the Stokes equation, the closure problem of which is discussed by Whitaker (1986). To obtain a minimal set of sub-problems, it is probably best to treat each specific fine-scale Eq. (2) separately to exploit its useful peculiarities. Other boundary conditions of Eq. (8) are often chosen to be periodic. The justification is that since scale separation was assumed earlier, the coarse-scale variables must be independent of the specific boundary conditions imposed on the REV to solve the closure problem (see the discussion in Section 2.2).

In the above, we have provided a high-level overview of the key components of the volume averaging method excluding many of its important, but sometimes distracting, details. The main takeaway is that a coarse-scale Eq. (3) is derived by spatially averaging the fine-scale Eq. (2). In so doing, assumptions are made that allow us to *decouple* the fine-scale deviation $\widetilde{\boldsymbol{u}}$ from the coarse-scale trend $<\boldsymbol{u}>$. The decoupling requires the solution of only *one* fine-scale closure problem, whose solution $\boldsymbol{b}$ when scaled by $<\boldsymbol{u}>$ and its derivatives yields $\widetilde{\boldsymbol{u}}$. While the end goal was to obtain a purely coarse-scale (homogenized) equation (Eq. (3)), we can reconstruct an approximate fine-scale solution $\boldsymbol{u}$ to Eq. (2) in the neighborhood of any point $\boldsymbol{x}$ (i.e., within Ω_x). This is done by simply substituting the computed $<\boldsymbol{u}>$ from Eq. (3) and $\widetilde{\boldsymbol{u}}$ from Eq. (9) into $\boldsymbol{u} = <\boldsymbol{u}> + \widetilde{\boldsymbol{u}}$. Such a reconstruction is valid if the fine-scale geometry (Fig. 4b) of Ω_x is *invariant* with respect to $\boldsymbol{x}$. Otherwise different closure problems must be solved at different $\boldsymbol{x}$. If the invariance is deterministic (i.e., the pore-scale structure in Fig. 4b is identical for all $\boldsymbol{x}$), as is the case for periodic media, then the reconstruction may also be regarded as deterministic (i.e., $\boldsymbol{u}$ approximately satisfies Eq. (2)); albeit of a low fine-scale accuracy because several assumptions were made during the averaging process. If the invariance with respect to $\boldsymbol{x}$ is only statistical, as is the case for geologic media, then the reconstruction must also be regarded as statistical as $\boldsymbol{u}$ will not satisfy Eq. (2) pointwise.

In Sections 3.3 and 3.4, we show that all multiscale computing and some numerical upscaling methods have a similar fine-scale reconstruction step. This reconstruction, however, is always deterministic because the goal of such methods is to either solve, or "numerically homogenize", Eq. (2) on a domain Ω in which the fine-scale geometry is fully known (tyranny of characterization is absent). The advantage of multiscale and (some) numerical upscaling methods, over homogenization, is that none of the assumptions made above, such as scale separation or stationarity, are required. The disadvantage is that fine-scale computations must be performed over the entirety of Ω; though in a fully decoupled and parallel fashion. Volume averaging, like all other homogenization methods, capitalizes on certain simplifying properties of a porous medium to restrict fine-scale calculations to a much smaller, but representative, subset of Ω. These calculations are precisely those associated with solving the closure problem on the REV (i.e., Ω_x for some

x). The restriction of fine-scale calculations in homogenization methods dramatically reduces the total number of degrees of freedom of the problem: fine-scale unknowns over Ω versus coarse-scale unknowns over Ω plus fine-scale unknowns over Ω_x. Such a reduction is not possible for porous media that do not exhibit scale separation, ergodicity, and stationarity (Wood, 2009). Shales seem to fall under this category (Fig. 2).

There is one final assumption on which the validity of the coarse-scale Eq. (3) hinges: fine-scale variations in $\boldsymbol{u}$ must be relatively smooth and devoid of any sharp transitions or gradients. If this assumption is false, then additional terms with explicit dependence on $\boldsymbol{u}$ would appear in Eq. (3). In other words, it would not be possible to derive a local[1] coarse-scale equation purely in terms of $<\boldsymbol{u}>$ (Cushman et al., 2002). In this case, the fine-scale variations of $\boldsymbol{u}$ are said to be *coupled* to the coarse-scale variations of $<\boldsymbol{u}>$, or that the scales are *not* separated with respect to $\boldsymbol{u}$. The degree of smoothness of $\boldsymbol{u}$ is controlled by, in addition to the microstructure of the pore space, the fine-scale parameters ς_f in Eq. (2). It is possible to express ς_f as a set of dimensionless numbers that control various regimes of the fine-scale physics. In the context of reactive transport of a solute in a single-phase fluid, several studies (Battiato and Tartakovsky, 2011; Battiato et al., 2009; Boso and Battiato, 2013) have identified sufficient (but not necessary) conditions under which the fine- and coarse-scale equations are decoupled. The conditions are expressed as inequalities in terms of the dimensionless numbers of the problem. In cases where the fine- and coarse-scale equations *are* coupled, Eq. (2) is said to be non-homogenizable and Eq. (3) can no longer be solved in isolation. Instead, some numerical algorithm is necessary that performs computations at both the fine and the coarse scales. Methods capable of such a task include hybrid methods (Section 3.2), recent numerical upscaling approaches (Section 3.3), and multiscale methods (Section 3.4). Each is motivated by and suited for a specific kind of problem with different degrees of coupling across scales. Lastly, our discussion so far has assumed that L_f in Eq. (2) is linear. If nonlinear, several linearization steps and/or assumptions must be introduced to derive the coarse-scale Eq. (3), which are again valid only if the fine-scale gradient of $\boldsymbol{u}$ is small (Whitaker, 1996; Lugo-Méndez et al., 2015). The challenge in homogenizing nonlinear operators is not exclusive to the method of volume averaging.

Let us now briefly discuss three other homogenization methods mentioned earlier. The hybrid mixture theory (HMT) and the thermodynamically consistent averaging theory (TCAT) are both related to the method of volume averaging (MVA), but with a key difference. In MVA, the starting point is a fine-scale differential equation, which is then homogenized through mathematical manipulations and various order-of-magnitude assumptions. While some care is taken to avoid the emergence of unphysical terms (e.g., third-order derivatives; Battiato et al., 2019; Wood and Valdés-Parada, 2013), there is no guarantee that the homogenized equation will be consistent with the second law of thermodynamics. The latter states that entropy production, through dissipative processes in the system, must always be positive. For a process described by the constitutive relation

$$\text{flux} = \text{constant} \times \text{force} \tag{11}$$

the rate of entropy production is proportional to *flux* × *force*. An example of Eq. (11) is Fourier's law of heat conduction, where *flux* is the heat flow and *force* is the (negative) temperature gradient. Both HMT and TCAT ensure that constitutive relations of the form Eq. (11) derived at the *coarse scale* (e.g., Darcy's law) are entropy producing. In MVA, such a constraint is not explicitly enforced. In porous media problems, a fine-scale equation such as Eq. (2) is often a combination of a conservation equation (mass, momentum, energy) and a fine-scale constitutive equation of the form Eq. (11). In MVA, it is this *combination* that is averaged, whereas in TCAT and HMT, only the conservation equation (plus an entropy balance equation) is averaged. The averaging procedure itself is identical to that described earlier for MVA. This leaves the coarse-scale constitutive equation to be determined separately in both HMT and TCAT, which if accomplished, "closes" the coarse-scale problem. In classical mixture theory (Cushman et al., 2002; Bedford and Drumheller, 1983) (MT), the precursor to both HMT and TCAT, one starts by postulating conservation equations at the coarse scale and then manipulates them to formulate coarse-scale constitutive equations of the form Eq. (11). The formulation assumes that thermodynamic equilibrium holds over each averaging volume Ω_x, i.e., local coarse-scale equilibrium. In other words, $\boldsymbol{u}$ and all other thermodynamic variables are nearly constant within each Ω_x but may vary for different x. HMT departs from MT in that the coarse-scale conservation equations are not postulated but derived by averaging their fine-scale counterparts (hence the prefix "hybrid"). Similar to MT, however, local coarse-scale equilibrium is assumed to derive the coarse-scale constitutive relations. TCAT goes one step further and uses averaging to derive not only the coarse-scale conservation (and entropy balance) equations, but also the coarse-scale thermodynamic relations on Ω_x from an assumption of local *fine-scale* equilibrium. The latter is subsequently used to derive the coarse-scale constitutive relations. Local fine-scale equilibrium implies that $\boldsymbol{u}$ does not need to be constant over Ω_x and may exhibit fine-scale fluctuations. Despite improvements over MT, both HMT and TCAT are limited by the same assumptions as MVA: scale separation, stationarity, and ergodicity.

The multiple scales expansion (MSE) method homogenizes Eq. (2) in a different way (Hornung, 1996). Instead of averaging it, the following *ansatz* is defined

$$\boldsymbol{u}(x,y) = \sum_{n=0} \varepsilon^n \boldsymbol{u}_n(x,y) \tag{12}$$

which expresses $\boldsymbol{u}$ as a power series in terms of a small "scale separation parameter" ε. The parameter ε is defined as the ratio of a fine-scale characteristic length (e.g., grain size in Fig. 4b) to a coarse-scale characteristic length (e.g., diameter of Ω in Fig. 4a). Eq. (12) also assumes that $\boldsymbol{u}$ depends on both x (the position vector) and $y = x/\varepsilon$, because then $\boldsymbol{u}(x,y)$ exhibits fluctuations at two frequencies: slow, associated with x or the coarse scale, and fast, associated with y or the fine scale. The coefficients of the power series, $\boldsymbol{u_n}$, must therefore also depend on x and y. Here, we assume Eq. (2) is time independent, otherwise the arguments of $\boldsymbol{u}$ and $\boldsymbol{u}_n$ would contain additional fast and slow time variables (Auriault and Adler, 1995; Battiato, 2016). By inspecting Eq. (12) we see that if we integrate both sides of it with respect to y over some REV (or unit cell) and take the limit $\varepsilon \rightarrow 0$, we obtain $<\boldsymbol{u}>(x) = <\boldsymbol{u_0}>(x)$. This defines the coarse-scale field variable at point x (up to order ε^1), for which we wish to derive a coarse-scale equation (like Eq. (3)). Substituting Eq. (12) into Eq. (2), and collecting like-powers of ε and setting them to zero, yields a coupled sequence of differential equations in terms of the series coefficients $\boldsymbol{u_n}$. The coupling, however, has a short recurrence, allowing low-order coefficients to be solved first and then used to solve higher-order coefficients. The lowest order term often reveals that $\boldsymbol{u_0}$ does not depend on y, confirming $<\boldsymbol{u}>(x) = \boldsymbol{u_0}(x)$ is indeed the sought-after coarse-scale variable. The next order term commonly leads to the formulation of a fine-scale closure problem (similar to MVA) that must be solved on some REV (or unit cell). The term after yields, after integrating it with respect to y, the governing equation for $\boldsymbol{u_0}$. The reader is referred to Hornung (1996) for a more detailed but accessible

[1] In local equations, $<\boldsymbol{u}(x,t)>$ is determined from a finite number of spatial and temporal derivatives of $<\boldsymbol{u}>$ at (x,t). In nonlocal equations (Wood and Valdés-Parada, 2013), one must know $<\boldsymbol{u}>$ and its derivatives at points other than, and distant from, (x,t). Here, we are interested in local coarse-scale equations. Nonlocal equations are computationally undesirable, as they involve integrals in space, time, or both, which lead to dense matrices and large memory requirements. Lagrangian, or particle-based, simulations provide an attractive (and perhaps only viable) alternative for such equations (Berkowitz et al., 2006).

account of MSE applied to various porous media problems. One advantage of MSE is the clarity it provides in identifying the approximations needed to arrive at a coarse-scale equation for $\boldsymbol{u}_0$ (Auriault and Adler, 1995). This is particularly useful in delineating applicability conditions for the coarse-scale equation in terms of the dimensionless numbers that control the physics (Battiato and Tartakovsky, 2011). However, MSE is limited by the same assumptions as all other homogenization methods: scale separation, stationarity, and ergodicity (formally, periodicity is required, but that is relaxed in practice). Lastly, MSE is very similar to perturbation expansion methods used in theoretical physics, where a hard problem (e.g., Schrödinger's equation) is transformed into a sequence of easier problems by introducing a small parameter ε, whose limit is eventually taken to *one* (Bender and Orszag, 1999). The small parameter in MSE, by contrast, connotes scale separation and its limit is taken to *zero*.

3.2. Hybrid computing

Hybrid computing refers to numerical algorithms used to solve problems where the fine-scale and coarse-scale physics are *coupled*. Such problems do not lend themselves to purely coarse-scale descriptions like Eq. (3), because coarse-scale variables, $<\boldsymbol{u}>$, depend on their fine-scale counterparts, $\boldsymbol{u}$. In hybrid methods designed for porous-media problems, "fine scale" invariably refers to the *pore scale* and "coarse scale" to some larger scale called the *continuum scale* that we specify later. To solve nonhomogenizable problems, hybrid methods perform simultaneous fine-scale and coarse-scale simulations that are coupled in a bidirectional fashion. In other words, pore-scale simulations provide input parameters to continuum-scale simulations and vice versa. The term "hybrid" means that not one (continuum-scale) but two (pore- and continuum-scale) models are needed to perform calculations. The freedom to choose a specific pore- or continuum-scale model, from a vast array of available tools, is one reason why hybrid methods are so diverse and application-specific. Another reason is that each hybrid method in the literature is designed to address a specific type of nonhomogenizability of the pore-scale equations. For example, pore-scale equations may be spatially nonhomogenizable but temporally homogenizable. This means that while the spatial fluctuations in $\boldsymbol{u}$ and $<\boldsymbol{u}>$ remain coupled, their temporal dynamics are not. The implication for hybrid methods is that the temporal dynamics in $\boldsymbol{u}$ may be ignored and approximated by a quasistatic process, as they relax much more quickly than temporal dynamics in $<\boldsymbol{u}>$. Another example is a problem where the pore-scale equations are spatially homogenizable throughout a domain except for a very small region (e.g., fracture). In such cases, one may use a pore-scale model for that region and a homogenized continuum-scale model for the rest of the domain. Many such combinations of fully or partially nonhomogenizable problems may be concocted, each of which demands a different hybrid method. We refer the reader to Scheibe et al. (2015a) for a thorough classification of such problems in the context of hydrogeology and their associated hybrid methods. Here, we focus on only two kinds of nonhomogenizable problems, that are among the more common, and discuss a few hybrid methods developed to solve them. A complete survey of the literature is not our intention, for which the reader is referred to recent reviews (Battiato, 2016; Scheibe et al., 2015a; Mehmani and Balhoff, 2015a; E et al., 2007; Oden et al., 2006; Yang, 2013). Instead, our aim is to illustrate the main ideas of hybrid computing and draw useful comparisons to the multiscale methods discussed in Section 3.4. We note that the nomenclature meant to convey "hybrid computing" varies across the literature and often includes the term "multiscale". In this paper, we distinguish between "hybrid" and "multiscale" for clarity.

Consider a porous domain Ω as depicted by Fig. 5a, over which Eq. (2) governs the fine-scale physics. The first kind of nonhomogenizable problem we consider is one where Eq. (2) is homogenizable everywhere on Ω except a very small region Ω_p shown by the red rectangle. In other words, the coarse-scale Eq. (3) applies on $\Omega_c = \Omega \setminus \Omega_p$ but not on Ω_p. Some physical examples include multiphase flow in a fracture (Ω_p) coupled to seepage into the rock matrix (Ω_c) (Hughes and Blunt, 2001); near-well (Ω_p) acidification dynamics coupled to far-field flow in a reservoir (Ω_c) (Golfier et al., 2002); mixing-induced reactions at a contaminant plume's fringes (Ω_p) coupled to a background flow (Ω_c) (Acharya et al., 2007). Hybrid methods used to solve such problems are designated here with the prefix "local". In a local hybrid method, a pore-scale model is used to solve Eq. (2) on Ω_p and a continuum-scale model to solve Eq. (3) on Ω_c. Hence, the *continuum scale*, in this context, is the scale at which Eq. (3) is valid on $\Omega \backslash \Omega_p$. The main factor that distinguishes one local hybrid method from another is the way Eq. (2) and Eq. (3) are coupled. Two categories can be identified: *overlapping* and *non-overlapping* methods.

In an overlapping method (Battiato et al., 2011; Scheibe et al., 2015b; Tartakovsky and Scheibe, 2011), the domain of Eq. (3) is extended from Ω_c to Ω by adding extra, yet unclosed, *source terms* to Eq. (3) that are non-zero only over Ω_p. Coupling is achieved by calculating these terms via pore-scale simulations, which in turn use continuum-scale simulations to determine their initial and/or boundary conditions. From an implementation standpoint, overlapping methods are intrusive as they require existing codes to be modified to account for the extra source terms. Although, significant progress has led to the recent emergence of dedicated parallel codes that circumvent the issue through efficient data management (Scheibe et al., 2015c; Scheibe et al., 2014). In a non-overlapping method (Yousefzadeh and Battiato, 2017; Mehmani and Balhoff, 2014; Tang et al., 2015; Balhoff et al., 2008; Sun et al., 2012a; Balhoff et al., 2007; Roubinet and Tartakovsky, 2013; Tartakovsky et al., 2008), Eq. (3) is solved strictly on Ω_c, which is disjoint from Ω_p. Coupling between pore- and continuum-scale simulations occurs through the interface Γ shared between Ω_c and Ω_p (i.e., $\Gamma = \overline{\Omega}_c \cap \overline{\Omega}_p$). The goal is to ensure continuity of fluxes (mass, momentum, energy) and thermodynamic state variables (concentration, pressure, temperature) across Γ, i.e., the red border in Fig. 5a. Among the many variants of non-overlapping methods in the literature, some adopt a Lagrangian framework to simulate the physics (e.g., smoothed particle hydrodynamics in Tartakovsky et al., 2008), which reduces the coupling to the interaction and/or transfer of particles across Γ. Others adopt an Eulerian framework where the coupling reduces to a set of algebraic constraints that must be solved alongside the differential Eqs. (2) and (3) (Yousefzadeh and Battiato, 2017; Mehmani and Balhoff, 2014; Balhoff et al., 2008; Sun et al., 2012a; Balhoff et al., 2007; Roubinet and Tartakovsky, 2013; Molins et al., 2019; Weishaupt et al., 2019; Baber et al., 2016). A general way to solve such differential-algebraic systems is through the method of Lagrange multipliers (adopted from the optimization literature; Nocedal and Wright, 2006), or called, in this context, *mortar* domain decomposition (Bernardi et al., 1994; Belgacem, 1999).

While the mathematical details of mortars are postponed until Section 3.4.1.4, the main idea is to introduce additional unknowns on Γ that serve as multipliers to a set of predefined basis functions on Γ (Bernardi et al., 1994; Belgacem, 1999; Arbogast et al., 2000; Arbogast et al., 2007; Peszynska et al., 1999). The obvious disadvantage of mortars is they increase the total number of unknowns to be solved in the problem (i.e., the old on Ω_c and Ω_p, and the new on Γ). But this drawback pales against the computational gains made by fully decoupling the calculations on Ω_c and Ω_p. Even more important is that the specific pore- or continuum-scale models used on Ω_c and Ω_p, respectively, can be chosen arbitrarily (Mehmani and Balhoff, 2014; Tang et al., 2015; Balhoff et al., 2008; Baber et al., 2016; Peszynska et al., 1999; Peszynska et al., 2000). This is because mortars act as a buffer between Ω_c and Ω_p, through which boundary information are passed. Regardless of the specific format used to store such information, fluxes and thermodynamic state variables on Γ, in the models for Ω_c and Ω_p, they are *projected* onto the mortar space to enforce continuity. A consequence of this projection is that the enforcement is in a weak (or variational) sense, rather than pointwise. The lack of dependence between the coupling scheme and the

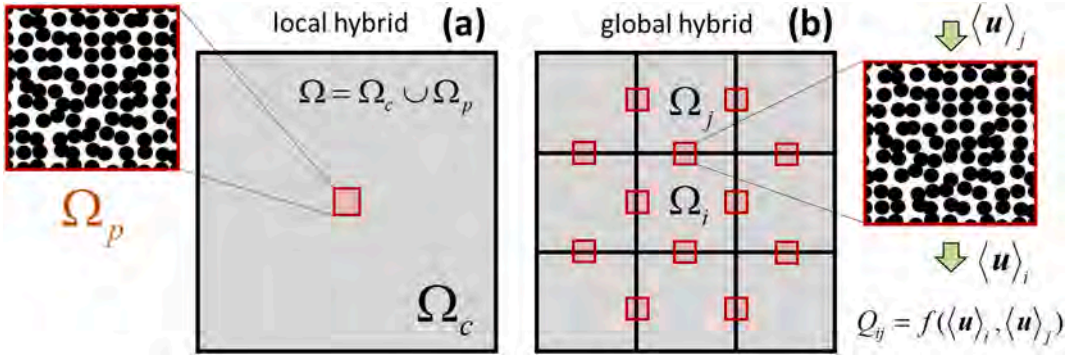

Fig. 5. Schematic of two kinds of hybrid methods. (a) In local hybrid, the fine-scale Eq. (2) is homogenizable over the entire domain Ω except for a small region Ω_p (red box). The coarse-scale Eq. (3) is valid over $\Omega_c = \Omega \backslash \Omega_p$. Pore-scale simulations are performed on Ω_p and continuum simulations on Ω_c. The two are coupled either in a non-overlapping fashion, by exchanging information at the interface (red border), or in an overlapping fashion, by introducing source terms into Eq. (3) that are non-zero only on Ω_p. (b) In global hybrid, either the fine-scale Eq. (2) is nonhomogenizable or the coarse-scale Eq. (3) is unknown over all of Ω. The schematic corresponds to HMM (E et al., 2007). Coarse-scale balance equations are written for each coarse grid (e.g., Ω_i) and combined with coarse-scale constitutive relations determined from fine-scale simulations (between Ω_i and Ω_j). The latter yields the flux Q_{ij} at the interface between Ω_i and Ω_j as a function of the coarse-scale state variables $<\boldsymbol{u}>_i$ and $<\boldsymbol{u}>_j$.

specifics of the subdomain models is particularly welcome in hybrid computing (Mehmani and Balhoff, 2014; Tang et al., 2015; Balhoff et al., 2008; Baber et al., 2016; Mehmani et al., 2012), where pore- and continuum-scale models are bound to differ.[2]

Algorithmically, there are two ways local hybrid problems can be solved with the mortar method: monolithically or sequentially (Mehmani and Balhoff, 2014; Ganis et al., 2014a; Ganis et al., 2014b). The monolithic scheme requires that the residual equations associated with Eq. (3) on Ω_c, Eq. (2) on Ω_p, and the continuity constraints on Γ to be assembled into a single global system, which is then solved in a parallel fashion (often by constructing Schur complements). The sequential scheme consists of formulating a much smaller (interface) system in terms of only the continuity constraints on Γ. To solve it, iterations (with a Newton-Krylov solver) are performed that, in turn, require isolated solutions of sub-problems on Ω_c and Ω_p (hence the name "sequential"). The main advantage of sequential schemes is that both the pore- and continuum-scale models on Ω_c and Ω_p can be treated as black-boxes. No knowledge of their internal workings is required, only their boundary information are needed. The monolithic scheme poses challenges to coupling Lagrangian (particle-based) with Eulerian (grid-based) models, because residual equations of the former are either ill-defined or difficult to integrate into the same system as the latter (Scheibe et al., 2015a). The sequential scheme, by contrast, is devoid of such limitations. On the other hand, the convergence rate of monolithic schemes tends to be much faster (Mehmani and Balhoff, 2014; Ganis et al., 2014a) than sequential schemes; because they account for off-diagonal blocks in their Jacobians. Other non-overlapping methods in the literature (Yousefzadeh and Battiato, 2017) formulate the coupling between Ω_c and Ω_p as a "root-finding problem" for the fluxes on Γ. It is instructive to also construe such methods as a sequential scheme where the mortar space consists of piecewise constant basis functions in the flux (instead of the thermodynamic state) variable.

The second kind of nonhomogenizable problem we consider is one where the fine-scale Eq. (2) cannot be homogenized over all, or a large part, of Ω. In other words, the coarse-scale Eq. (3) is invalid over most of Ω. Hybrid methods designed to solve such problems are prefixed here by the term "global". A notable example is the heterogeneous multiscale method (HMM) (E et al., 2007; Chu et al., 2012; Chu et al., 2013; E and Engquist, 2003; Alyaev et al., 2018), which we discuss next.[3] Consider the domain Ω in Fig. 5b. HMM divides Ω into a number of coarse grids Ω_i delineated by the solid black lines. Even though a coarse-scale equation for $<\boldsymbol{u}>$ is unavailable, we may still write a coarse-scale balance equation for $<\boldsymbol{u}>$ on each coarse grid Ω_i as follows:

$$\frac{\partial \langle \boldsymbol{u} \rangle}{\partial t} = \int_{\partial \Omega_i} flux \tag{13}$$

where *flux* denotes the in/outflow of $\boldsymbol{u}$ through the boundary of Ω_i, or $\partial\Omega_i$. Eq. (13) may represent a balance of mass, momentum, or energy. Here, we ignore various possible source terms for simplicity. The fact that Eq. (2) is nonhomogenizabe means that a closed-form expression for *flux* (i.e., purely in terms of $<\boldsymbol{u}>$) is unavailable. To compute *flux*, HMM places a pore-scale model Ω_{ij}^p (red boxes in Fig. 5b) between every pair of adjacent coarse grids Ω_i and Ω_j. Note that the union of all Ω_{ij}^p does *not* cover Ω, but instead comprises a tiny fraction of Ω ($|\Omega_i| \gg |\Omega_{ij}^p|$). The pore-scale model Ω_{ij}^p is used to compute the *flux* between Ω_i and Ω_j, which we call Q_{ij}. The boundary conditions for Ω^p_{ij} are provided by the coarse-scale variables $<\boldsymbol{u}>_i$ and $<\boldsymbol{u}>_j$ defined as spatial averages of $\boldsymbol{u}$ over Ω_i and Ω_j, respectively. The coarse-scale Eq. (13) and the pore-scale models are solved in tandem to predict the evolution of $<\boldsymbol{u}>$ in space and time.

An interesting analogy between HMM and the homogenization method TCAT, described in Section 3.1, can be made. In TCAT, the coarse-scale constitutive relation for *flux*, namely $Q_{ij} = f(<\boldsymbol{u}>_i, <\boldsymbol{u}>_j)$, is derived formally by homogenizing the fine-scale equations. In HMM, this relation is computed numerically. What motivates HMM is that the *flux* relation cannot be derived by TCAT, or any other homogenization method, for a nonhomogenizable problem. The implicit assumption in HMM, of course, is that pore-scale simulations on Ω_{ij}^p are representative of the whole interface between Ω_i and Ω_j. In effect, pore-scale results on a few sample points are extrapolated over Ω. While limiting, the assumption is far less restrictive than that of homogenization methods, which extrapolate pore-scale simulations on *one* pore-scale domain (the closure problem) to the entirety of Ω. The reader is referred to E et al. (2007) for a lucid presentation of HMM.

We conclude by drawing an analogy between multiscale methods, discussed in Section 3.4, and hybrid methods. The goal of multiscale methods is to solve Eq. (2) either exactly or approximately. The procedure consists of a numerical homogenization step (upscaling) and a fine-scale reconstruction (downscaling) step. The former requires computing a set of numerical basis functions on a collection of subdomains of Ω, the fine-scale details of which can be discarded if deemed redundant (same as homogenization; see Section 3.1). Those parts of Ω

[2] Claims that suggest such a dependence (Battiato, 2016; Yousefzadeh and Battiato, 2017) for mortar methods are unwarranted.

[3] HMM may also be adapted into a local hybrid method to solve problems like Fig. 5a ("Type A" in E and Engquist, 2003)).

where fine-scale details cannot be discarded are nonhomogenizable, which may span either the whole domain (Fig. 5b) or part of it (Fig. 5a). Multiscale methods can localize the downscaling step to these nonhomogenizable regions (adaptively) and reconstruct $\boldsymbol{u}$, while interpolating the coarse-scale variable $<\boldsymbol{u}>$ over the rest of the domain. Both steps are performed using the basis functions.

So what then is the difference between multiscale and hybrid methods? In theory, nothing. Both aim to solve or approximate the solution of Eq. (2) and resolve fine-scale details wherever needed. The only technical difference is that hybrid methods solve an *analytically* homogenized Eq. (3) in regions where it is applicable, whereas multiscale methods use *numerically* homogenized basis functions to interpolate the solution in such regions. In practice, however, there is a big difference between the two: multiscale methods rely on the tyranny of characterization being absent and assume that a fine-scale (geometric) description of Ω is available. Neither homogenization nor hybrid methods make this assumption. In geologic media, pore-scale characterizations are available only on very small samples (e.g., core plugs). Modeling performed at the field scale, by either solving homogenized equations or performing hybrid computations, tacitly assumes that the available pore-scale descriptions can be extrapolated to the whole reservoir. Thus a handful of closure problems, in homogenization, or pore-scale models, in hybrid methods, becomes the basis for such extrapolation. But is extrapolation an exclusive feature of hybrid methods? No. In multiscale methods, extrapolation translates to ascribing the same basis functions constructed on a *few* coarse grids (or samples) to *all* the coarse grids in Ω. The basis functions themselves may even be built on a very small portion of each coarse grid (similar to HMM), if scales are separated, and then extrapolated over the whole grid. Hence, the distinction between multiscale and hybrid methods, with the former often identified as a "multiresolution" (Scheibe et al., 2015a) or "traditional multiscale" (E et al., 2007) approach, is, especially in light of recent developments (see Section 3.4.2), no longer sharp. While ignored above, similar arguments hold for problems exhibiting temporal nonhomogenizablity.

3.3. Numerical upscaling

Numerical upscaling was born out of a practical necessity to reduce the complexity of geostatistical models. Such models provide a probabilistic snapshot of the fine-scale variability in the subsurface conditioned to measurements such as cores, well-logs, and seismic. In this context, "fine scale" refers to some continuum scale, not the pore scale. Geomodels often consist of $O(10^7$–$10^8)$ grid cells (Durlofsky and Chen, 2012; Farmer, 2002), which until a decade ago were ~1–2 orders of magnitude larger than what most advanced reservoir simulators could handle. While current hardware capabilities have certainly improved to the point that these numbers are no longer considered "large", upscaling remains important. The reason lies in the inadequacy of a single simulation in producing useful predictions about the subsurface. Multiple simulations are needed because petrophysical descriptions of the subsurface are uncertain. Upscaled models are much faster at running through the thousands of geostatistical realizations at hand than the original fine-scale geomodel. Another recurrent motivator for numerical upscaling is that a few, albeit expensive, simulations performed upfront will save substantial computational cost later. The upfront simulations may be single-phase flow and the subsequent simulations time-dependent two-phase flow. Upscaling accelerates two-phase flow simulations because solving a fine-scale two-phase flow problem is computationally equivalent to solving a fine-scale single-phase flow problem for *every time step*. Similarly, the computational cost of numerically upscaling a few realizations in a geostatistical ensemble can be amortized over thousands of subsequent realizations. Early upscaling involved the use of analytical methods such as power-averaging (Renard and de Marsily, 1997; Deutsch, 1989). Our review focuses only on numerical methods as they are widely accepted to be the only viable option for practically complex problems.

The goal of numerical upscaling is to parametrize a coarse-scale equation whose solution approximates, in an averaged sense, that of a given fine-scale equation. The entire workflow consists of two steps: (1) *assume* a mathematical form for the coarse-scale equations, and (2) compute the parameters that appear in them. Each step is the basis for a different classification of upscaling methods. The first is based on *which* physical parameters, appearing in the coarse-scale equations, are to be upscaled. The second is based on *how* such parameters are actually computed. In homogenization (Section 3.1), hybrid methods (Section 3.2), and multiscale methods (Section 3.4), the first step is absent because no a priori assumptions about the form of the coarse-scale equations are made. Similar to multiscale methods, however, numerical upscaling can be viewed as a kind of "numerical homogenization".

The first classification divides methods into *single-phase parameter* and *two-phase parameter* upscaling; we drop the word "parameter" for brevity. In single-phase upscaling, the fine-scale equation, in its simplest form, is

$$\nabla\cdot(k\nabla p) = 0 \tag{14}$$

and the postulated coarse-scale equation is

$$\nabla\cdot(k^*\nabla p^c) = 0 \tag{15}$$

which has the same mathematical form as Eq. (14). In Eq. (14), k and p are the fine-scale permeability and pressure, respectively. In Eq. (15), k^* is the upscaled (or equivalent) permeability and p^c is the coarse-scale pressure. The goal of single-phase upscaling is to compute k^*. If the fine-scale Eq. (14) were time-dependent, then a fine-scale porosity φ and an equivalent porosity φ^* would appear in Eqs. (14) and (15), respectively. The latter would have to be upscaled alongside k^*. But upscaling φ^* from φ is trivial because porosity is an additive parameter, i.e., φ^* is a weighted average of φ. In two-phase upscaling, the fine-scale equations, in their simplest form, are

$$\nabla\cdot(\lambda(S)\,k\nabla p\,) = 0 \qquad \varphi\frac{\partial S}{\partial t}+\nabla\cdot(uf(S)\,) = 0 \tag{16}$$

and the postulated coarse-scale equations are

$$\nabla\cdot(\lambda^*(S^c)\,k^*\nabla p^c\,) = 0 \qquad \varphi^*\frac{\partial S^c}{\partial t}+\nabla\cdot(u^c f^*(S^c)\,) = 0 \tag{17}$$

where we have neglected fluid compressibility, gravity, and capillarity (Rabinovich et al., 2015). In Eq. (16), S is the fine-scale saturation; $\lambda(S)$ and $f(S)$ are the fine-scale total mobility and fractional flow function, respectively; and u is the fine-scale total velocity. In Eq. (17), S^c is the coarse-scale saturation; $\lambda^*(S^c)$, and $f^*(S^c)$ are the coarse-scale mobility and fractional flow function, respectively; and u^c is the coarse-scale total velocity. The total velocities depend on pressure through Darcy's law: $u = -k\ \nabla p$ and $u^c = -k^*\ \nabla p^c$. In both Eq. (16) and Eq. (17), the left equation is called the *pressure equation* and the right equation the *saturation equation*. The goal of two-phase upscaling is to determine k^*, φ^*, $\lambda^*(S^c)$, and $f^*(S^c)$. The calculation of k^* and φ^* is the same as before, which makes accurate single-phase upscaling a prerequisite to accurate two-phase upscaling (Durlofsky, 2005).

An important question is whether the assumed forms of the coarse Eqs. (15) and (17) are valid? Eq. (15) is indeed on solid theoretical ground as it can be derived from Eq. (14) via homogenization (Bourgeat, 1984; Sáez et al., 1989), assuming spatial scale separation. Eq. (17) on the other hand is less so, because the homogenized form of a hyperbolic saturation equation at the fine scale is generally a non-local equation at the coarse scale (E, 1992; Tartar, 1989; Efendiev et al., 2000). In an interesting analysis by Hou et al. (2006), the nonlocality was captured by a projection operator that integrates fine-scale fluctuations along streamlines. But since non-local equations are undesirable from a computational respect, leading to dense matrices and/or requiring excessive storage, local approximations have been sought (Efendiev and

Durlofsky, 2003). For example, Sáez et al. (1989) derived a local two-saturation equation at the coarse scale for a two-porosity medium under the assumption that the fine- and coarse-scale temporal dynamics are separated. Despite its heuristic form, experience has shown that Eq. (17) leads to acceptable coarse-scale predictions in many geologically complex settings (Durlofsky and Chen, 2012; Chen and Durlofsky, 2006a).

The second classification of upscaling methods, based on the algorithm used to compute coarse-scale parameters, includes *local*, *extended local*, *global*, *local-global* (or *quasi-global*), and *ensemble-level* methods. We describe each method briefly and with reference to the schematic in Fig. 6. For simplicity, we limit ourselves to single-phase upscaling and the computation of k^* in Eq. (15). The ideas that follow can be generalized to two-phase upscaling, for which we provide appropriate references.

Consider the global domain in Fig. 6a consisting of fine grids (thin lines) and coarse grids (thick lines). The dots mark the location of the coarse-scale unknowns p^c. The ratio between the size of a coarse grid to that of a fine grid is called the "coarsening ratio" and is equal to three in Fig. 6a. Our goal is to calculate k^* for the green highlighted coarse grid. In local upscaling (Durlofsky, 1991; Pickup et al., 1994), k^* is obtained by performing fine-scale calculations over the highlighted coarse grid only; see Fig. 6b. In extended local upscaling (Gomez-Hernandez, 1991; Wen et al., 2003), the support of fine-scale calculations is expanded to include one (or more) layers of the surrounding coarse grids; see Fig. 6c. An important issue in both methods is the choice of boundary conditions (BCs) imposed on the local regions. In local upscaling, several such BCs exist. The simplest condition is the so-called "pressure-no flow" BC. The idea is to impose a pressure gradient along one of the coordinate directions and seal off the remaining lateral boundaries (Farmer, 2002; Durlofsky, 2005). The disadvantage is that the computed k^* is always diagonal. This is incorrect if bedding planes are not aligned with the grid lines. Another option is to impose periodic BCs (Durlofsky, 1991; Pickup et al., 1994). This results in a full-tensor k^*. The advantage of periodicity is that k^* is guaranteed to be symmetric and positive definite (see Boe (1994) for proof), consistent with the second law of thermodynamics which precludes upgradient flow. Other BCs are discussed in (Gomez-Hernandez, 1991; Boe, 1994). We note in passing that local BCs are particularly important in two-phase upscaling, for which specialized saturation BCs have been proposed (Wallstrom et al., 2002). Despite the computational efficiency of local upscaling, the BCs imposed on each coarse grid are somewhat arbitrary and account for most of the upscaling error (the rest comes from assumptions made to formulate the coarse-scale Eq. (15)). Extended local upscaling alleviates the shortcoming by adding a border region around each coarse grid, Fig. 6c. But the BCs imposed on the extended region itself are just as arbitrary.

The desire to improve the BCs on coarse grids has motivated the development global (White and Horne, 1987; Holden and Nielsen, 2000) and local-global (Chen et al., 2003; Chen and Durlofsky, 2006b) upscaling methods, in which local BCs are determined from global computations.[4] In global upscaling, the fine-scale Eq. (14) is solved over the entire domain (Fig. 6a). The resulting flow field is then used to derive k^* for each coarse grid. Global upscaling is more accurate than local and extended local methods, but it is also more expensive (Chen and Li, 2009). The high cost of global fine-scale simulations is justified by the fact that the coarse model can be reused later. Even so, for problems that are too large or involve time-dependent global fine-scale simulations, like two-phase flow, global upscaling can be prohibitive.

Local-global methods were conceived to reduce the computational cost of global upscaling while retaining similar accuracy. Fig. 6d depicts the idea. A global *coarse-scale* simulation is performed to provide local BCs to an extended region around the green highlighted coarse grid. A fine-scale problem is then solved over this region yielding k^* for the coarse grid. The k^* for all coarse grids are computed in this manner and fed back into another coarse-scale simulation. The process is repeated for a few iterations (<5; see Chen et al., 2003; Chen and Li, 2009; Kolyukhin and Espedal, 2010) until a "self-consistent" solution is reached. Several variants of local-global upscaling exist, among which *adaptive local-global* methods (Chen and Durlofsky, 2006b; Kolyukhin and Espedal, 2010; Chen and Li, 2010) have proved particularly robust in both single-phase and two-phase upscaling. The advantages of adaptive methods are twofold: computations are localized, which reduces cost, and the occurrence of artifacts in k^* (e.g., negative eigen values) is reduced.

In global and local-global upscaling, flow problems must be solved on the entire domain (Fig. 6a). A key question is then what global BCs should be imposed on the domain? Two options exist (Durlofsky and Chen, 2012): (1) "generic flow", and (2) "specific flow". Generic flow means to impose uniform flow (or pressure gradient) along each of the coordinate directions, resulting in a total of two simulations in 2D and three simulations in 3D. Generic flow produces k^* values that tend to be robust: the upscaled model is reusable under a wide range of flow conditions and global BCs. Specific flow means to impose global BCs that are sufficiently similar to those of a particular problem of interest. For example, we may want to apply the upscaled model to a specific constellation of wells in a petroleum reservoir. Specific flow BCs lead to more accurate, but less robust, k^* compared to generic flow BCs. Although, modest changes in flow conditions seem to be within tolerance (Chen and Durlofsky, 2006b).

Finally, ensemble-level upscaling is designed to accurately reproduce, not the flow field of a specific realization, but the statistics of all plausible realizations of the flow field (Durlofsky and Chen, 2012; Li and Durlofsky, 2016). This is motivated by using upscaling in the context of uncertainty quantification. In ensemble-level upscaling, any one of the local or global methods can be used to calculate upscaled parameters for a small fraction (~10%) of the realizations. Statistical mappings are then constructed that relate upscaled parameters, like k^*, to easily calculated attributes of the coarse grids. An example of such an attribute is the distribution of permeabilities in neighboring coarse grids. Upscaling for the remaining realizations (~90%) is accomplished by simply mapping the precomputed attributes of all the coarse grids to the upscaled parameter of interest.

Eq. (17) is often discretized with the finite volume method over the coarse grids. In finite volume, a quantity called "transmissibility" is assigned to each interface between adjacent coarse grids, which can be thought of as a (nonlinearly) weighted interface permeability. Transmissibilities are used to compute flowrates across grid interfaces. In all the upscaling methods discussed so far, it is possible to compute upscaled parameters over either the *volume* of each coarse grid, like k^*, or the *interface* shared between adjacent coarse grids, see Fig. 6e. The latter yields, in the context of single-phase upscaling, the upscaled transmissibility, T^*, that can be used in subsequent coarse-scale simulations with the finite volume method (Durlofsky and Chen, 2012). While we did not discuss transmissibility upscaling in detail, several studies suggest it to be more accurate than permeability upscaling (Chen et al., 2003; Romeu and Noetinger, 1995).

So what are the limitations and pitfalls of numerical upscaling? The first concerns local and extended local upscaling, and we demonstrate it by invoking a simple example devised by Farmer (2002). Consider the two layered coarse grids i and $i+1$ in Fig. 7, which share an interface. Our goal is to compute k^* for grid i in the horizontal direction. If the layers of grid $i+1$ align with those of grid i at the interface (Fig. 7a), then local upscaling produces the exact value of k^* (arithmetic mean). If the layers do not align (Fig. 7b), then local upscaling does not *see* the discontinuity at the interface and instead of predicting a smaller value of

[4] Interestingly, the development of global upscaling seems to have predated local upscaling (Renard and de Marsily, 1997; White and Horne, 1987). Therefore, global upscaling may have been motivated less by the inadequacies of local upscaling and more by an expectation that local BCs are important for obtaining accurate upscaled parameters.

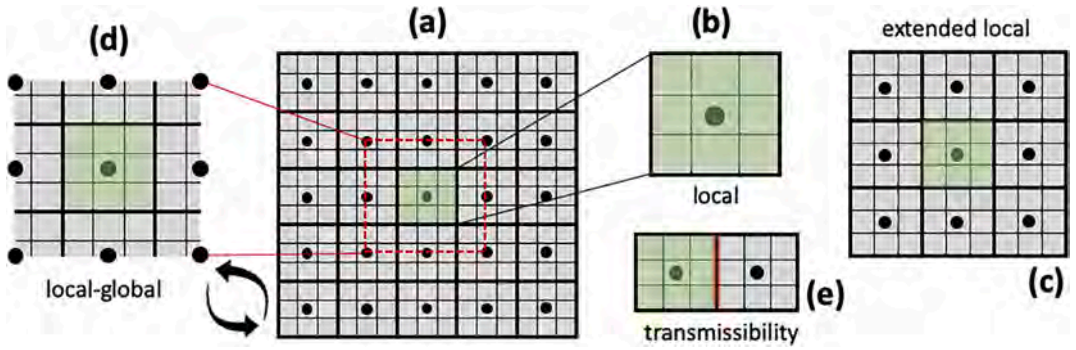

Fig. 6. Schematic of various upscaling methods. (a) Global domain with fine grids (thin lines), coarse grids (thick lines), and coarse-scale unknowns (dots). Green highlights the coarse grid subject to upscaling. (b) Local upscaling performs local fine-scale calculations on the green grid only. (c) Extended local upscaling includes a layer (or more) of neighboring coarse grids into the local fine-scale calculations. (d) Local-global upscaling uses global coarse-scale simulations on (a) to determine local fine-scale BCs on (d). The coarse-grid parameters derived for the green coarse grid are then fed back to another global coarse-scale simulation on (a). (e) Instead of k^*, one may upscale the transmissibility, T^*, at the interfaces (red) between adjacent coarse grids.

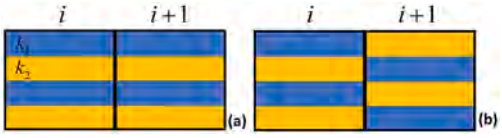

Fig. 7. Schematic of two adjacent coarse grids i and $i+1$. Each block consists of two types of horizontal layers. The permeability of the blue layers is k_1 and that of the yellow layers is k_2. (a) Layers of the two grids align at the interface. (b) Layers do not align at the interface. Local upscaling fails in (b) but not in (a).

k^*, it yields the same value as before (Fig. 7a). Moreover, the error in k^* increases as the contrast between the layer permeabilities, k_1 and k_2, is amplified. Extended local upscaling, by contrast, *does* detect the discontinuity and leads to a better estimate of k^*. But we could then construct a similar example that would cause extended local upscaling to fail too, by simply moving the discontinuity to the border of the extended local region (Fig. 6c). The same argument holds for transmissibility upscaling (Fig. 6e), i.e., the discontinuity could be placed at the center of a coarse grid. Farmer's example highlights a fundamental challenge in (extended) local upscaling: how to *a priori* select the size of the (extended) local coarse grids (Kolyukhin and Espedal, 2010). Flow-based gridding (Durlofsky, 2005) is a promising solution but requires global calculations of the flow field.

A second pitfall of upscaling is that, except for local upscaling (Fig. 6b), all other techniques are susceptible to producing k^* tensors that violate either symmetry or positive definiteness (Durlofsky, 2005). In the case of transmissibility upscaling (Fig. 6e), T^* may be negative, which is unphysical. The reason is most methods determine k^* using a least squares approach that imposes symmetry only weakly (and rarely positive definiteness). In practice, k^* is symmetrized *ex post facto* by simply taking the average of k^* with its transpose. And violations of positive definiteness are treated by replacing anomalous k^* tensors with values obtained from local upscaling (Durlofsky, 2005). Anomalous k^* are observed typically in regions where flowrates are small in magnitude, and thus the mean flow has indefinite direction. Adaptive local-global upscaling excludes such regions from calculations altogether, which has the added benefit of reducing computational cost (Chen and Durlofsky, 2006b).

Another pitfall of numerical upscaling concerns downscaling. With a few exceptions (Chen et al., 2003; Gautier et al., 1999), once the coarse-scale equations are parameterized, they are used to perform simulations with no further recourse to the fine scale. While this is often acceptable for single-phase flow simulations, where the governing equations are either elliptic or parabolic with relatively smooth solutions, it is generally unacceptable for two-phase flow simulations, where the saturation equation is hyperbolic. In Section 3.4.3, we discuss that hyperbolic equations exhibit sharp variations at the fine-scale and thus require some form of downscaling. In numerical upscaling, the saturation equation can be solved in three ways. The first is to perform single-phase upscaling, to obtain k^* and φ^*, and then discretize and solve the *fine-scale* saturation equation in Eq. (16) on the *coarse grid*. The second is to perform two-phase upscaling and then solve the coarse-scale saturation equation in Eq. (17) on the coarse grid. The third, is to perform single-phase upscaling, reconstruct the fine-scale velocity with a method similar to that proposed by Gautier et al. (1999), and then solve the *fine-scale* saturation equation in Eq. (16) on the *fine grid*. The first incurs large errors if sharp saturation gradients exist at the fine scale. The second alleviates these errors to a degree, and the third removes them almost completely. For strongly hyperbolic saturation equations, the third approach may be the only reliable option (see discussion in Section 3.4.3).

Perhaps the most important limitation of numerical upscaling is that there is no built-in mechanism to *a priori* estimate or control errors. Local-global methods are capable of iteratively improving the quality of coarse-scale predictions, but iterations result in solutions that are "self-consistent" not *convergent* (to the averaged fine-scale solution). Several studies (Chen et al., 2003; Kolyukhin and Espedal, 2010) show that an irreducible error persists even after a large number of iterations. Algorithms that provide *a priori* estimates of such errors are valuable, especially when upscaling is applied to cases for which no prior test cases exist, because they allow users to assess the reliability of coarse-scale predictions. If predictions are deemed unreliable, a separate algorithm that can control the errors may be used to guarantee a reliable solution. Multiscale methods (Section 3.4) possess such built-in algorithms (Hajibeygi et al., 2008; Lunati et al., 2011) that allow them to converge arbitrarily close to the exact solution. Considering the main difference between multiscale methods and numerical upscaling is that the latter assumes a specific form for the coarse-scale equations, one may conclude that the lack of convergence in upscaling is rooted in its insistence on computing upscaled *parameters* (not equations). If this requirement were to be lifted, convergence might follow, but then the benefit of obtaining coarse-scale parameterizations (if necessary) is lost. Various "error models" for upscaling have been proposed (Lødøen et al., 2005; Glimm et al., 2001; O'Sullivan and Christie, 2005) and they provide a promising route to quantifying errors. However, such models are *statistical* and, thus, do not provide the means to control errors of a particular realization.

3.4. Multiscale computing

Multiscale computing aims to obtain the *solution* of a partial differential equation (PDE) by performing calculations at two or more

spatiotemporal scales. As opposed to numerical upscaling, where the goal is to parameterize a postulated coarse-scale equation and then use it to obtain an approximate coarse-scale solution, the goal of multiscale methods is to compute the solution directly. No coarse-scale equation is postulated or parameterized. Multiscale methods produce solutions on both a coarse grid, like numerical upscaling, and a fine grid, unlike numerical upscaling. The coarse-scale solution is a consistent representation of the fine-scale solution, with the latter obtained by downscaling the former (made precise later). While multiscale methods are not designed to compute coarse-scale parameters, some do arise naturally from their mathematical formulation as discussed later (e.g., flux matrices). Such parameters allow useful comparisons to be made between multiscale and upscaling methods. As discussed in Section 3.2, multiscale methods also differ (but only in technical details) from hybrid methods. In hybrid methods, a coarse-scale equation is solved together with a fine-scale equation by using separate models for each. The coarse-scale equation is an analytically homogenized PDE. In multiscale methods, no homogenized coarse-scale equations are required. Instead, an algebraic *coarse-scale problem* (or system) is formulated that is consistent with the fine-scale equations. Hence, there is no need for separate models at different scales, which makes multiscale methods algorithmically simpler. Most importantly, multiscale methods have the ability to estimate and control prediction errors, because they have a built-in mechanism for *convergence*. Numerical upscaling and hybrid methods lack this capability, and while their solutions are internally consistent, they do not, and are not meant to, satisfy the fine-scale PDE.

The motivation for using multiscale methods is the same as numerical upscaling: subsurface geomodels are too complex and computationally too expensive to perform conventional single-scale simulationsover a fine grid. The workflow of multiscale methods is as follows: (1) divide the geomodel into a number of coarse grids, each consisting of many fine grids; (2) perform local fine-scale calculations on each coarse grid to construct a set of local basis functions; (3) formulate a coarse-scale problem and solve it to obtain a coarse-scale solution; and (4) reconstruct the fine-scale solution by using the basis functions and the coarse-scale solution from step 3. Step 2 is amenable to parallelism and step 3 is computationally much cheaper than global fine-scale simulations common in single-scale methods. Step 4 is a downscaling step that may be omitted if only a coarse-scale solution is desired (Section 3.4.3).

The literature on multiscale methods is vast but has its origins in the mid-1980s to early 2000s (Babuška and Osborn, 1983; Babuška et al., 1994; Hou and Wu, 1997; Chen and Hou, 2002; Jenny et al., 2003). Our intent is not to cover all of the contributions since, but to provide a representative cross-section of the available ideas to the uninitiated reader and hopefully add new insights as to where the state-of-the-art lies relative to what may be possible in the near future. In Section 3.4.1, we discuss multiscale methods designed to solve continuum- or Darcy-scale problems, among which we focus on: multiscale finite element (MsFE), mixed multiscale finite element (MxMsFE), multiscale finite volume (MsFV), and multiscale mortar finite element (MoMsFE). The important class of variational multiscale methods (Hughes, 1995; Hughes et al., 1998; Arbogast, 2012) is not discussed. Section 3.4.2 is devoted to multiscale methods designed for pore-scale problems, in which we discuss: straightforward extensions of Darcy-scale methods to the pore scale, reinterpretation of pore-network models (PNM) as numerical upscaling, and a recent pore-level multiscale method (PLMM). Section 3.4.3 discusses the role of downscaling in obtaining reliable coarse-scale predictions of subsurface processes. We conclude with Section 3.4.4, in which we introduce the idea of a "flux matrix" and propose an algorithmic way of bridging some of the gap between pore- and Darcy-scale physics.

3.4.1. Darcy scale

In the following sections, we discuss several multiscale methods in the context of a prototypical problem of practical interest in subsurface applications: flow of two immiscible fluids in a porous medium. The governing equations are

$$\phi\frac{\partial S_w}{\partial t}+\nabla\cdot u_w = -q_w \tag{18a}$$

$$\phi\frac{\partial S_o}{\partial t}+\nabla\cdot u_o = -q_o \tag{18b}$$

where

$$\begin{aligned} u_w &= -\lambda_w\nabla p \qquad & \lambda_w &= k k_{rw}/\mu_w \\ u_o &= -\lambda_o\nabla p & \lambda_o &= k k_{ro}/\mu_o \end{aligned} \tag{18c}$$

Eq. 18a-b describes the conservation of mass for water, subscript w, and oil, subscript o. φ denotes porosity, S_w and S_o water and oil saturations, u_o and u_w phase fluxes, and q_w and q_o source terms due to injector or producer wells. Eq. (18c) is a statement of Darcy's law for multiphase flow, where λ_w and λ_o denote water and oil motilities, k absolute permeability, k_{rw} and k_{ro} relative permeabilities of water and oil, and μ_w and μ_o viscosities of water and oil. Eq. (18) neglects effects due to gravity, capillarity, and compressibility. In computer simulations, a different but equivalent formulation of Eq. (18) is preferred, which is given by

$$\nabla\cdot(\lambda\nabla p) = q \tag{19a}$$

$$\phi\frac{\partial S_w}{\partial t}+\nabla\cdot(f_w u) = -q_w \tag{19b}$$

where

$$u = -\lambda\nabla p \qquad \lambda = \lambda_o+\lambda_w \tag{19c}$$

In Eq. (19), λ is the total mobility, u is the total velocity ($u = u_o + u_w$), and q is the total volumetric source term ($q = q_o + q_w$). Eq. (19a) is obtained by adding Eqs. (18a) and (18b), whereas Eq. (19b) is just a restatement of Eq. (18a). Eq. (19) is preferred over Eq. (18) because it decouples the elliptic character of the problem, captured by Eq. (19a), from its hyperbolic character, captured by Eq. (19b). Elliptic equations tend to have smooth solutions that are amenable to coarse-scale approximation. A small perturbation in the parameters of an elliptic equation (e.g., q) affects its solution globally, or over the whole domain. Hyperbolic equations, by contrast, tend to have non-smooth solutions marked by localized discontinuities and steep gradients. They are less amenable to coarse-scale approximation. Moreover, perturbations in the parameters of a hyperbolic equation affects its solution locally, or over part of the domain. For these reasons, the computational cost of solving elliptic (and parabolic) equations is often much higher than that of solving hyperbolic equations, where adaptivity and explicit time stepping may be utilized. Most multiscale methods have therefore been developed to accelerate the solution of elliptic (and parabolic) equations such as Eq. (19a). The fine-scale velocity u obtained from solving Eq. (19a) (and using Eq. (19c)) is subsequently used to solve Eq. (19b) over the *fine grid*; although nothing prevents one from solving it over the coarse grid instead. In Sections 3.4.1.1–4, we focus on Eq. (19a) and describe various multiscale methods developed to solve it. We refer, henceforth, to Eq. (19a) as the *flow equation* and to Eq. (19b) as the *transport equation*.

Multiscale methods are described with reference to Fig. 8, which depicts a global domain (Fig. 8a) consisting of coarse grids (thick lines) and fine grids (thin lines). Each coarse grid contains nine fine grids, though the number is much larger in practice. Both grid types are assumed to be Cartesian, but more complicated unstructured grids are also possible. The coarse-scale pressure unknowns are depicted by filled circles, whose spatial arrangement differs between methods. Our discussion focuses only on the main ideas of each method and ignores all fine-scale discretization details. For ease of reference, we restate the three steps common to all multiscale methods with respect to Eq. (19a)

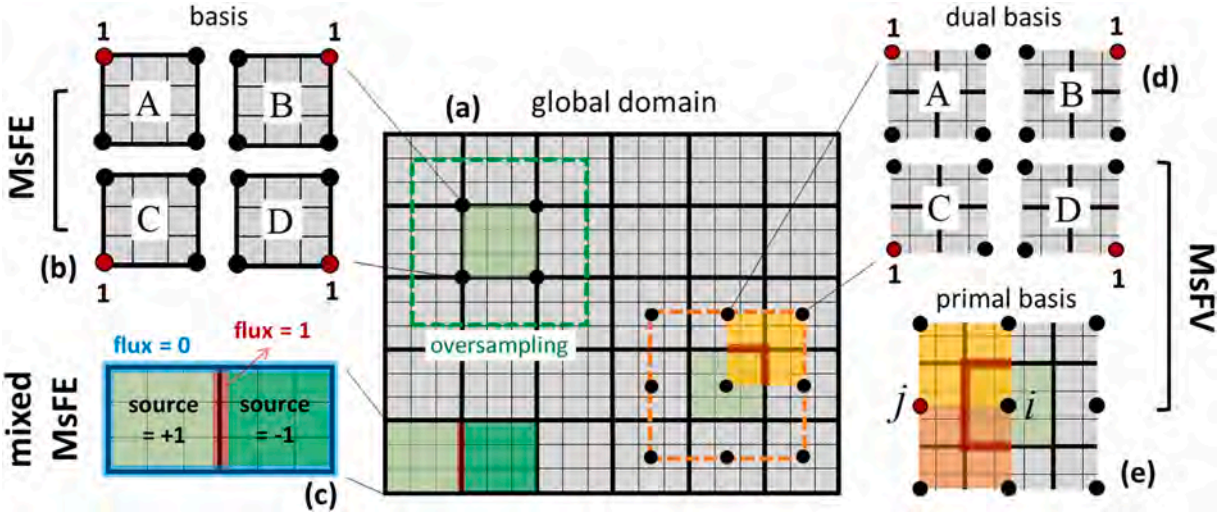

Fig. 8. Schematic of different multiscale methods. (a) Global domain with coarse grids (thick lines) and fine grids (thin lines). (b) MsFE: four bases (A to D) are constructed per coarse grid (green). Pressure is set to one at the red nodes and to zero at the black nodes. Linear pressure, or the 1D solution of Eq. (24), along the edges is imposed as the BCs of Eq. (22). The dashed green box in (a) is the "oversampling" region for the green coarse grid. (c) MxMsFE: bases are associated with the interfaces between adjacent coarse grids. The flux is set to one (left-to-right) at the interface (red) between the two grids and to zero along their remaining boundaries (blue). Uniformly distributed source and sink terms are added, respectively, to the light and dark green grids to ensure solvability of the basis functions. (d-e) MsFV: coarse-scale unknowns are defined at primary grid centers; dots in (a). Dual bases (A to D) are constructed on each dual grid (orange box) in exactly the same way as (b). (e) Primal bases are built by setting the pressure at node *j* (red) to one and at all the other 8 nodes (black) to zero. The reconstructed flux on the red highlighted boundary is then used as a Neumann BC to compute one of the primal bases of grid *i*. The remaining 8 primal bases are obtained by moving *j* to the other notes (back) in (e).

1. Build basis functions for each coarse grid by solving a localized version of Eq. (19a)
2. Formulate and solve a coarse-scale problem to obtain a coarse-scale solution to Eq. (19a)
3. Reconstruct the fine-scale pressure and velocity solutions to Eq. (19a) needed to solve Eq. (19b)

Multiscale methods differ only in the specifics (numerics, gridding, etc.) of each step. Step 1 is the most important as it determines the overall accuracy and convergence rate of a method (Hou et al., 1999). Many have therefore urged further research towards improving the quality of local bases by devising better BCs on the coarse grids (Babuška and Osborn, 1983; Hou et al., 1999). The reconstructed fine-scale velocity, u, in step 3 must be divergence-free to be useful in subsurface applications. If not, large errors in the subsequent solution of the transport Eq. (19b) result. Step 3 is optional if no transport equation is to be solved and only a coarse-scale solution to Eq. (19a) is desired.

3.4.1.1. Multiscale finite element (MsFE). Consider the global domain Ω in Fig. 8a. We outline MsFE by first describing standard finite element (FE) as it would be applied to discretizing and solving Eq. (19a) on the *coarse grid* (ignoring the fine grid). In FE, Eq. (19a) is expressed in variational form by multiplying both sides of it by a sufficiently smooth test function and then integrating it over Ω. Assuming no-flux BCs on $\partial\Omega$ for simplicity, the result is

$$a(p,v) = q(v) \tag{20a}$$

$$a(pv) = \int_\Omega \lambda_{ij}\frac{\partial v}{\partial x_i}\frac{\partial p}{\partial x_j}dx \qquad q(v) = -\int_\Omega q\,v\,dx \tag{20b}$$

where $a(p,v)$ and $q(v)$ are bilinear and linear operators, respectively. Eq. (20a) is equivalent to Eq. (19a) in a "weak" sense, provided that it holds for *all* test functions in an appropriately defined function space. The solution, p, is sought within another function space. Here, both p and v are members of H^1, short for the Hilbert space of all functions with square-integrable first derivatives. To compute numerically the solution of Eq. (19a) on the coarse grid, p and v must be chosen from a *finite dimensional* subspace of H^1 equipped with a finite collection of basis functions φ_i. The solution, p, and all test functions, v, can then be expressed as linear combinations of the bases. In particular, p is written as

$$p = \sum_i p_i^c \varphi_i \tag{21}$$

The summation in Eq. (21) is over all the basis functions. The scalar multipliers, p_i^c, are the coarse-scale (note the superscript) pressure unknowns. Each basis φ_i is non-zero only over one coarse grid, and zero outside of it. Hence, φ_i is said to be defined only on that coarse grid. Let Ω^c denote the green highlighted coarse grid in Fig. 8a and h its size. In FE, φ_i is chosen to be a simple function such as a piecewise polynomial. As an example, let φ_i be a piecewise bilinear function. For the Cartesian Ω^c, four such bases can be defined as shown in Fig. 8b. Namely, φ_i assumes the value *one* at one of the four corners of Ω^c (red dot in Fig. 8b) and *zero* at the remaining corners (black dots in Fig. 8b). The edge values of φ_i vary linearly between the corner values. The coarse-scale unknowns, p_i^c, correspond to the pressure values at the corners of Ω^c. Similar bases and coarse-scale unknowns are defined on all other coarse grids of Ω. A crucial feature of FE is that φ_i is chosen (bilinear here) independently from the underlying fine-scale variability of λ over Ω^c. Let the spatial correlation length of such fluctuations be ε. If $\varepsilon \ll h$, then it is obvious that a polynomial φ_i leads to a bad approximation of the fine-scale pressure over Ω^c (see Eq. (19a)). The only way the approximation would be good is if the *amplitude* of the λ fluctuations is small, i.e., λ is approximately homogeneous over Ω^c. Of course, if $h \ll \varepsilon$, then φ_i is adequate and FE yields an accurate solution. Unfortunately, geologic porous media are heterogeneous and exhibit large oscillations, of several orders of magnitude, in λ at the fine scale. To obtain an accurate solution, FE must use grids that satisfy $h \ll \varepsilon$, which is computationally prohibitive as this is equivalent to solving a fine-scale problem over Ω.

The key contribution of Babuška and Osborn (1983), Babuška et al. (1994), Hou and Wu (1997), Hou et al. (1999), and many others since, has been to choose φ_i such that it is *informed* by the fine-scale variability of λ over Ω^c. This is accomplished by requiring that each φ_i satisfy

$$\nabla\cdot(\lambda\nabla\varphi_i) = 0 \tag{22}$$

over Ω^c. When λ is oscillatory, Eq. (22) ensures that φ_i is too. This simple but crucial modification to φ_i is the essence of multiscale finite element (MsFE). Each basis, φ_i, is numerically computed by discretizing and solving Eq. (22) over a set of fine grids that comprise Ω^c. Notice the departure from FE through the introduction of fine grids in addition to coarse grids. When $\varepsilon \ll h$, substituting φ_i from Eq. (22) into Eq. (21) results in a much more accurate solution, p, than if φ_i is chosen to be a polynomial. And when $h \ll \varepsilon$, MsFE reduces to FE, as both resolve fine-scale variabilities in λ and are thus equally accurate. To compute p_i^c in Eq. (21), a coarse-scale problem is formulated by simply substituting Eq. (21) into Eq. (20a) and setting the test functions, v, equal to an arbitrary basis function φ_j. The result is a linear system in terms of p_i^c, whose coefficient (or stiffness) matrix consists of entries of the form

$$K_{ij} = \int_\Omega \lambda_{kl}\frac{\partial\varphi_i}{\partial x_k}\frac{\partial\varphi_j}{\partial x_l}dx \tag{23}$$

Despite the apparent simplicity of MsFE, its main challenge, as with all multiscale and numerical upscaling methods, is in the definition of the BCs needed to solve Eq. (22) on Ω^c. Such BCs are inherently approximate, because exact knowledge of them would require the global fine-scale solution itself. The specific choice of BCs on Ω^c is sometimes called *numerical closure* (or localization assumption) and is the only significant source of error in not just MsFE, but all multiscale and upscaling methods. The more accurate the BCs, the more accurate the solution, p. The simplest BCs on $\partial\Omega^c$, the boundary of Ω^c, are the same as the bilinear φ_i described earlier. Namely, one or zero at the corners of Ω^c and linear variations along the edges. A more accurate alternative, suggested by Hou and Wu (1997), is to solve the following 1D problem along each of the four edges of Ω^c subject to the zero/one corner values depicted in Fig. 8b

$$\frac{\partial}{\partial x_t}(\lambda\nabla\varphi_i)_t = 0 \tag{24}$$

Eq. (24) is the 1D version of Eq. (22) along the tangent coordinate, t, to each edge. Eq. (24) assumes that the fluxes normal to $\partial\Omega^c$ do not interact with those tangent to it. In other words, tangential fluxes are conserved independently from normal fluxes on $\partial\Omega^c$. The BCs obtained from Eq. (24) are more accurate than the linear BCs for the same reason that the basis functions obtained from Eq. (22) are more accurate than the bilinear basis functions in FE: Eq. (24) accounts for the fine-sale variability of λ (Hou et al., 1999).

While Eq. (24) has proved successful in many applications (Hou and Wu, 1997; Efendiev and Hou, 2008), it remains an approximation whose errors may be large depending on the variability of λ and the relative magnitudes of h and ε. If scale separation holds for λ (fluctuations are stationarity with finite correlation length), then as $\varepsilon/h \to 0$ the MsFE solution p converges to the solution of the *homogenized* (not pointwise) form of Eq. (19a) (not shown). Hou and Wu (1997) and Hou et al. (1999) showed that the (L_2-)errors scale as $O(\varepsilon/h)$. But this means that if $\varepsilon \approx h$, then the errors are of $O(1)$ even in the limit h and $\varepsilon \to 0$. Hou and Wu (1997) and Hou et al. (1999) called this the *resonance effect*, in which the size of fine and coarse grids conspire to prevent errors from vanishing. Intuitively, the finding implies that if the coarse and fine grids are of comparable size, then the errors can be large. To reduce resonance, Hou and Wu (1997) proposed the idea of "oversampling", in which the BCs of φ_i are determined not from Eq. (24), but from solving Eq. (22) over an enlarged region containing Ω^c; dashed green box in Fig. 8a. The solution, ψ_j, on the oversampling region is used to extract the BCs of φ_i on $\partial\Omega^c$. In practice however, instead of re-solving Eq. (22) on Ω^c with the new BCs, one may obtain φ_i by first cropping ψ_j to fit Ω^c and then rescaling it so that φ_i assumes either one or zero at the corners of Ω^c. Oversampling improves the accuracy of MsFE for the same reason that extended local upscaling improves the accuracy of local upscaling: better local BCs. But it also shares the same drawback as extended local upscaling: how big should the oversampling region be?

We conclude with three remarks. If only the coarse-scale solution, p_i^c, is desired, the fine-scale bases, φ_i, can be discarded right after the coarse-scale stiffness parameters K_{ij} are computed for all ij. This makes MsFE very similar to numerical upscaling. MsFE yields only an approximate solution to Eq. (19a). One way to control these (pointwise) errors is to use MsFE as a global preconditioner to standard FE (Castelletto et al., 2017), a point which we expound on in the context of MsFV (Section 3.4.1.3). Despite the successful application of MsFE in many complex problems (Efendiev and Hou, 2008), its main disadvantage is that the fine-scale velocity, u, is not divergence-free. This is partly because FE, used as the fine-grid solver to compute φ_i from Eq. (22), is not conservative. Another reason is that the p obtained from MsFE (Eq. (21)) has discontinuous first derivatives, and thus u, across shared boundaries between adjacent coarse grids. Both MxMsFE and MsFV, discussed next, address this shortcoming. The reader is referred to Efendiev and Hou (2008) for further details on MsFE.

3.4.1.2. Mixed multiscale finite element (MxMsFE). Similar to Section 3.4.1.1, we outline MxMsFE by first describing standard mixed finite element (MxFE) as it would be applied to discretizing and solving Eq. (19a) on the *coarse grid* (ignoring the fine grid). MxFE can be thought of as a flux-conservative version of FE (Gatica, 2014). Flux conservation is achieved by introducing two sets of unknowns into the discretized problem: one for pressure and another for velocity. The velocity unknowns are defined over shared edges between adjacent grids, and ensure that fluxes are continuous across grid boundaries. The scheme is therefore conservative by construction. A variational form of Eq. (19a) is first derived by splitting it into two first-order equations, one representing Darcy's law and the other mass conservation. Each equation is then multiplied by a test function and integrated over Ω

$$\int_\Omega \lambda^{-1}u\cdot v\,dx - \int_\Omega p\nabla\cdot v\,dx = 0 \tag{25a}$$

$$\int_\Omega \omega\nabla\cdot u\,dx + \int_\Omega q\,\omega\,dx = 0 \tag{25b}$$

Eqs. (25a) and (25b) are, respectively, the "weak" forms of Darcy's law and mass conservation. For simplicity, no-flux Neumann BCs are assumed on $\partial\Omega$. Eq. (25) is equivalent to Eq. (19a) provided that it holds for *all* test functions v and ω chosen from appropriately defined function spaces. The definition of such spaces is not trivial and we refer to Gatica (2014) for details. To numerically solve Eq. (25) on the coarse grid, one must select p, u, v, and ω from *finite dimensional* subspaces of the original function spaces. It is possible to then write p and u as linear combinations of a finite number of pressure, φ_i, and velocity, ψ_j, basis functions as follows

$$p = \sum_i p_i^c\varphi_i \qquad u = \sum_j u_j^c\psi_j \tag{26}$$

where the multipliers p_i^c and u_j^c denote the coarse-scale pressure and velocity unknowns, respectively. In MxFE, simple bases, like polynomials, are selected for φ_i and ψ_j. A popular choice is the lowest order Raviart-Thomas bases, RT0, where φ_i is a piecewise constant function on each coarse grid and ψ_j is a piecewise linear vector function with a constant divergence over each coarse grid. A key feature of MxFE is that the normal components of the velocity bases, ψ_j, match up (or are continuous) along shared coarse-grid boundaries. From the discussion in Section 3.4.1.1, it must by now be clear that if $\varepsilon \ll h$, then polynomial bases like RT0 will not capture the fine-scale variability of λ and can thus lead to large errors.

Multiscale mixed finite element (MxMsFE) replaces bases like RT0 with ones that are informed by the fine-scale variability of λ. Chen and Hou (2002) proposed the following formulation of MxMsFE, based on

modifying MxFE. Consider the pair of light and dark green coarse grids in Fig. 8c. Denote the left grid by Ω_L^c and the right grid by Ω_R^c. Let Γ denote the shared interface between them. Let the pressure bases defined on Ω_L^c and Ω_R^c be φ_L and φ_R, and the velocity bases associated with Γ and defined on Ω_L^c and Ω_R^c be ψ_L and ψ_R, respectively. The pressure bases φ_L and φ_R are chosen to be the same as those in MxFE: constants over Ω_L^c and Ω_R^c. But the velocity bases ψ_L and ψ_R are set equal to $-\lambda\nabla\theta$ on Ω_L^c and Ω_R^c, respectively, where θ is the solution of the following boundary-value problem on $\Omega_L^c \cup \Omega_R^c$

$$\nabla\cdot(\lambda\nabla\theta) = +\frac{1}{|\Omega_R^c|} \quad on \quad \Omega_R^c \tag{27a}$$

$$\nabla\cdot(\lambda\nabla\theta) = -\frac{1}{|\Omega_L^c|} \quad on \quad \Omega_L^c \tag{27b}$$

$$\lambda\nabla\theta\cdot n = 0 \qquad \lambda\nabla\theta\cdot n_\Gamma = -\frac{1}{|\Gamma|} \tag{27c}$$

In Eq. (27c), n is the normal vector on the external boundary of $\Omega_L^c \cup \Omega_R^c$ (blue in Fig. 8c) and n_Γ is the normal on Γ (red in Fig. 8c) pointing from left to right. The intuitive meaning of Eq. (27a-b) is to impose a uniform flux of magnitude $1/|\Gamma|$ (total flowrate equal to one) from left to right across Γ and to seal off the remaining boundaries of $\Omega_L^c \cup \Omega_R^c$ (see Fig. 8c). To balance the imposed interface flux, a spatially uniform source term of magnitude one is added to Ω_L^c, and an identical sink term is added to Ω_R^c (see Fig. 8c). The balance of source, sink, and interface fluxes ensures the solvability of Eq. (27). Eq. (27) yields one velocity basis for Ω_L^c, ψ_L, and another for Ω_R^c, ψ_R, both associated with the interface Γ. Similar equations are solved for all other coarse-grid interfaces in Ω. Once all ψ_j are computed, they are inserted into Eq. (26) along with φ_i, for which no calculations were performed, to obtain p and u. To formulate a coarse-scale problem (or linear system) in terms of p_i^c and u_j^c, Eq. (26) is substituted into Eq. (25) and the test functions v and ω are, respectively, set equal to an arbitrary choice of the velocity and pressure basis functions. The method is thereby complete.

A few remarks are now in order. Note that the basis problem, Eq. (27), is markedly different from MsFE in Section 3.4.1.1. Eq. (27) is a Neumann problem which, unlike MsFE, guarantees that fluxes are continuous across Γ. Also note the fine-scale pressure p, in Eq. (26), is identical to the coarse-scale pressure because φ_i was chosen to be a constant over each coarse grid. Therefore, the above version of MxMsFE captures the impact of fine-scale variabilities of λ on u but not p. Aarnes (2004) and Aarnes et al. (2005) proposed a way to recover the fine-scale variabilities of p. Modifications to Eq. (27) for coarse grids that contain point sources, like wells, were also proposed. Finally, Aarnes (2004) suggested that instead of imposing a uniform flux of $1/|\Gamma|$ across Γ in Eq. (27), a better option might be to impose a non-uniform λ-weighted flux with, again, a total flowrate equal to one.

We note that other MxMsFE formulations related to, but different from, the above have also been proposed in the variational context (Arbogast et al., 1998; Arbogast, 2002). An interesting one is by Arbogast (2002), where two sets of finite element spaces are defined: RT0 for the fine grid and a higher-order Brezzi-Douglas-Duran-Fortin (BDDF1) space (Brezzi et al., 1987) for the coarse grid. Velocities in the RT0 space do not contribute to any fluxes crossing the coarse-grid boundaries, only those in the BDDF1 space do. The two spaces allow p and u to be written as decompositions of a coarse-grid component and a fine-grid component. Substituting the decompositions into the variational form, Eq. (25), and testing the latter with only RT0 basis functions on a coarse grid, yields the basis problem associated with that coarse grid; which differs from Eq. (27). The coarse-scale problem is formulated by testing Eq. (25) with only BDDF1 basis functions. Accurate results were reported for highly variable λ and h/ε values up to 100. The reader is referred to Arbogast (2012) for a detailed yet clear presentation of MxMsFE.

3.4.1.3. Multiscale finite volume (MsFV). Similar to MxMsFE, multiscale finite volume (MsFV) guarantees mass conservation over both the course grid and the fine grid. But unlike MxMsFE, it does so with much fewer degrees of freedom, as pressure is the only unknown compared to pressure and velocity in MxMsFE. Since its inception by Jenny et al. (2003), MsFV has evolved significantly and has been applied to a variety of practical problems in the subsurface including two-phase flow (Jenny et al., 2004, 2006), compressible gas flow (Lunati and Jenny, 2006), black-oil model (Lee et al., 2008), near-well flow (Wolfsteiner et al., 2006), and poromechanics (Castelletto et al., 2019). Here, we confine our attention to the original formulation as it encapsulates the main ideas. More recent milestones such as correction functions (Lunati and Jenny, 2006), iterative correction (Hajibeygi et al., 2008), and algebraic formulations (Zhou and Tchelepi, 2008) are briefly discussed.

In MsFV, the coarse grids depicted by Fig. 8a are called *primary grids*. By connecting the centroids of the primary grids (dots in Fig. 8a), a second set of coarse grids, called *dual grids*, is also constructed (orange in Fig. 8a). Coarse-scale pressure unknowns, p_i^c, are associated with the centroids of the primary grids (or corners of dual grids). Over each dual grid, four basis functions for pressure, φ_i, are constructed in a manner identical to that described for MsFE (see Section 3.4.1.1 and Fig. 8b). Namely, Eq. (22) is solved on each dual grid, Ω^d, subject to either linearly varying BCs over the edges of Ω^d or BCs that correspond to the solution of Eq. (24) over each edge. The value of φ_i is equal to *one* at only one of the four corners of Ω^d (red in Fig. 8d), and *zero* at the remaining corners (black in Fig. 8d). In MsFV, the φ_i are referred to as *dual bases*. The only difference between the bases in MsFE and the dual bases in MsFV is that the latter uses a finite volume method to solve Eq. (22) over the fine grids comprising Ω^d. To reconstruct the fine-scale pressure, p, an equation identical to Eq. (21) is used. The main point of departure in MsFV from other multiscale methods is the way the coarse-scale problem is formulated to obtain p_i^c, and the way the conservative fine-scale velocity, u, is reconstructed (steps 2 and 3 in Section 3.4.1).

The coarse-scale problem is obtained by imposing mass conservation over each primary grid Ω_i^p

$$\int_{\partial\Omega_i^p} \lambda\nabla p\cdot n\, d\Gamma = \int_{\Omega_i^p} q\, d\Omega \tag{28}$$

Eq. (28) is obtained by integrating Eq. (19a) over Ω_i^p and using Gauss's theorem. To convert the left-hand side of Eq. (28) into an expression in terms of p_i^c, we consider the green-highlighted primary grid, Ω_i^p, in the bottom-right corner of Fig. 8a. Notice that four dual grids overlap with Ω_i^p, one of which is highlighted in orange. Since four pressure bases, φ_j, are defined on each dual grid, a total of 16 bases contribute to the net flux crossing $\partial\Omega_i^p$. The contribution is proportional to the p_j^c associated with φ_j, because Eq. (21) is a linear combination. More precisely, the left-hand side of Eq. (28) can be expressed as follows

$$\int_{\partial\Omega_i^p} \lambda\nabla p\cdot n\, d\Gamma = \sum_k T_{ij} p_j^c \tag{29}$$

where T_{ij} is the flux contribution of the jth dual basis, φ_j, on the ith primary grid, Ω_i^p. The parameters T_{ij}, which can be computed during a preprocessing step, are called *coarse-scale transmissibilities*. The initiated reader may recognize the similarity between T_{ij} here and transmissibility matrices (over "interaction regions") in multipoint flux approximation (MPFA) schemes used in finite volume methods (Avatsmark et al., 2008). Eq. (29) expresses the net flux crossing $\partial\Omega_i^p$ as a linear combination of T_{ij} weighted by p_j^c. Substituting Eq. (29) into Eq. (28) yields a linear system in terms of p_j^c, which constitutes the coarse-scale problem. The fine-scale pressure, p, is obtained from Eq. (21).

Unfortunately, the fine-scale velocity u derived from p, using $u = -\lambda\nabla p$, is discontinuous across dual grid boundaries and is therefore not conservative over the fine grid. To address this shortcoming, Jenny et al. (2003) introduced a second set of basis functions ψ_j, this time for

velocity, that are defined on the primary grids and referred to as *primal bases*. Notice the subscript j here indexes the primary grids, unlike Eq. (29) where it indexed the dual grids. To construct ψ_j, let us again consider the green-highlighted primary grid Ω_i^p in Fig. 8e. There are nine coarse-scale nodes (black dots) that correspond to the corners of four neighboring dual grids overlapping Ω_i^p. The goal is to compute the primal basis ψ_j on Ω_i^p corresponding to node j. This is done by first reconstructing an intermediate velocity field on Ω_i^p and then modifying it to obtain ψ_j. The intermediate velocity, u_j^*, is obtained by setting the multiplier p_j^c in Eq. (21) to *one* at node j and to *zero* at the remaining 8 nodes of Fig. 8e. The resulting p from Eq. (21) is then substituted into $u_j^* = -\lambda \nabla p$ to obtain u_j^* (recall u_j^* is discontinuous across dual grid boundaries). We then restrict u_j^* to $\partial\Omega_i^p$, where it is non-zero only on the red-highlighted segment in Fig. 8e, and use it as the BCs for a new local problem on Ω_i^p. This problem is similar to Eq. (19a) and its solution yields ψ_j. The above steps are repeated for all j corresponding to the other eight nodes in Fig. 8e (black dots). Once all nine ψ_j are computed, a *conservative* fine-scale u over Ω_i^p is reconstructed by superposition of the primal bases ψ_j weighted by their associated p_j^c; the latter obtained from solving the coarse-scale problem.

Primal bases are an efficient way to reconstruct a conservative velocity, u, for the time-dependent transport Eq. (19b), provided that changes in λ are either local or small between time steps. If not, dual and primal bases must be recomputed periodically, which can be expensive. For the 2D domain in Fig. 8a, 4 dual bases per dual grid and 9 primal bases per primary grid must be computed. In 3D, these numbers increase to 8 per dual grid and 27 per primary grid. The cost is justified only if the bases can be reused across multiple time steps or if they require updating for only a small fraction of the coarse grids. In the two-phase flow simulations conducted by Jenny et al. (2004; 2006), both conditions held. But in problems that exhibit strong nonlinearity (Lunati and Jenny, 2006), like compressible flow, or complex physics (Lunati and Jenny, 2008), like capillarity and gravity, the primal bases can become cumbersome and diminish the flexibility of MsFV. A better alternative was suggested by Jenny et al. (2006) and developed extensively by Lunati and Jenny (2006). Instead of computing 9 primal bases in 2D, and 27 in 3D, it is possible to compute only one *correction function* per primary grid Ω_i^p, which is designed to account for any non-homogeneous terms arising from the presence of complex physics or the linearization of nonlinear equations. Correction functions (much like ψ_j) are computed by solving a pure-Neumann problem on Ω_i^p that automatically ensures u is conservative on the fine grid. The only drawback presented by corrections functions, compared to primal bases, is that they must be recomputed *every time step*. Even so, the ability to compute them in parallel and the increased flexibility enabled by them to simulate complex physics has made correction functions the default strategy in MsFV.

The above description of MsFV produces an efficient and accurate approximation to Eq. (19a). But the algorithm lacks the ability to estimate or control errors. Hajibeygi et al. (2008) proposed a convenient remedy. Because the only source of error in MsFV comes from the assumed BCs (like Eq. (24)) for the dual bases, one can *iteratively* improve such BCs. The idea is to use the solution from the previous iteration to improve the BCs in the next iteration. But instead of modifying the BCs of the dual bases directly, which would be expensive because all four φ_i per dual grid would have to be recomputed, Hajibeygi et al. (2008) modified the BCs of the correction functions thereby requiring only one calculation per primary grid. The improved BCs are obtained by using a so-called "smoother" (or fine-grid preconditioner), such as incomplete LU or line relaxation, to attenuate high-frequency (or fine-grid) errors across iterations. Intuitively, a smoother can be thought of as solving a diffusion equation for the fine-scale error in p (Hajibeygi et al., 2008), which quickly dissipates the fluctuations of this error on the fine grid but not the coarse grid. The foregoing iterative strategy has equipped MsFV with the ability to assign confidence intervals to predictions and to control them if desired.

We conclude by describing a recent algebraic interpretation of MsFV (Zhou and Tchelepi, 2008; Lunati and Lee, 2009) that has rendered its implementation within existing reservoir simulators non-intrusive. Suppose Eq. (19a) is discretized over the fine grid on Ω (Fig. 8a), resulting in the following global linear system

$$\mathbf{A}\,\mathrm{p} = \mathrm{b} \tag{30}$$

Let N^f and N^c be the total number of fine grids and coarse grids in Ω, respectively. The coefficient matrix $\mathbf{A}$ is $N^f \times N^f$ and the column vectors b and p are both $N^f \times 1$. The first step is to construct a *prolongation matrix*, $\mathbf{P}$, by assembling the discrete forms of the dual bases, φ_i, into the columns of $\mathbf{P}$. The result is a $N^f \times N^c$ matrix with the property that if $\mathbf{P}$ left-multiplies the $N^c \times 1$ vector p^c, containing the coarse-scale pressure unknowns, it yields an approximation to the $N^f \times 1$ vector p

$$\mathrm{p} = \mathbf{P}\,\mathrm{p}^c \tag{31}$$

In other words, $\mathbf{P}$ is the algebraic operator that *downscales* p^c into p. The second step is to construct a *restriction matrix*, $\mathbf{R}$, which does the exact opposite of $\mathbf{P}$. Namely, it *upscales* p into p^c. In MsFV, $\mathbf{R}$ is a $N^c \times N^f$ matrix with only zeros and ones as entries. Left-multiplying a $N^f \times 1$ vector by $\mathbf{R}$ has the effect of summing up all the rows of that vector associated with the fine grids comprising each individual coarse grid. To obtain the coarse-scale problem, Eq. (30) is first left-multiplied by $\mathbf{R}$, which is equivalent to imposing mass conservation on each coarse grid similar to Eq. (28). Next, Eq. (31) is substituted for p, resulting in

$$\mathbf{A}^c\mathrm{p}^c = \mathrm{b}^c \quad s.t. \quad \mathbf{A}^c = \mathbf{RAP} \qquad \mathrm{b}^c = \mathbf{R}\,\mathrm{b} \tag{32}$$

where $\mathbf{A}^c$ is a $N^c \times N^c$ matrix with entries identical to T_{ij} in Eq. (29). Note that $\mathbf{A}^c$ is much smaller than $\mathbf{A}$. Solving Eq. (32) yields the coarse-scale unknowns p^c, and substituting p^c into Eq. (31) yields the fine-scale solution p. If only the former is desired, one may discard $\mathbf{P}$ and $\mathbf{R}$ right after $\mathbf{A}^c$ and b^c are assembled via Eq. (32). The above algebraic formulation is equivalent to the geometric description of MsFV presented earlier and can be used as a coarse-grid preconditioner in iterative solvers like GMRES. To improve convergence, and to estimate and control errors, one may pair this coarse-grid preconditioner with a fine-grid preconditioner (or smoother) like iLU. The pairing allows both high- and low-frequency errors to be reduced simultaneously. The above algebraic formulation of MsFV shares similarities with multigrid methods (Saad, 2003).

3.4.1.4. Multiscale mortar finite element (MoMsFE). We describe MoMsFE (Arbogast et al., 2007) with reference to Fig. 9a-b. The global domain Ω is divided into a number of coarse grids (thick lines), or *subdomains*, which themselves contain a number of fine grids (thin lines). Let Γ_{ij} denote the interface between two adjacent subdomains Ω_i^c and Ω_j^c, and $\Gamma = \cup_{ij}\Gamma_{ij}$ the union of *all* interfaces between the subdomains. The whole idea behind MoMsFE is to compute the pressure field p on Γ. Because if p on Γ is known, then Eq. (19a) can be solved on each subdomain, as its BCs would be known. This, in turn, would yield p over Ω. So the main question is: how do we compute p on Γ? In MoMsFE, the interface Γ is discretized by placing finite element nodes (black dots in Fig. 9a) on it. Associated with each node is a *scalar* coarse-scale pressure unknown p_i^c and a finite element basis *function* μ_i defined strictly on Γ. A common choice for μ_i is the hat function, as depicted for nodes i (red) and j (green) in Fig. 9b. In 2D, μ_i is a 1D function (like Fig. 9b), and in 3D, it is a 2D function (not shown). With such a discretization of Γ, the pressure field on Γ can expressed as

$$p|_\Gamma = \sum_i p_i^c \mu_i \tag{33}$$

where one formally refers to μ_i as a *mortar basis* function (Bernardi et al., 1994; Arbogast et al., 2000) and to p_i^c as its corresponding *Lagrange multiplier*. The main benefit of Eq. (33) is that it transforms our original

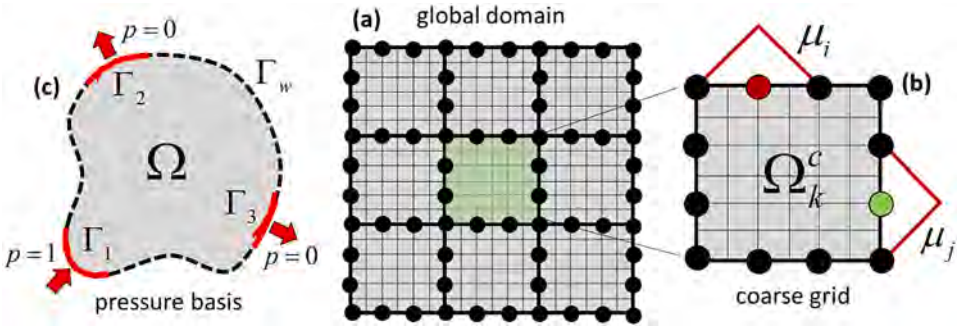

Fig. 9. (a–b) Schematic of MoMsFE and the construction of a flux matrix. (a) A global domain divided into coarse grids (thick lines) and fine grids (thin lines). Coarse-scale unknowns are defined at nodes (dots) along coarse-grid interfaces. (b) Mortar bases (here hat functions) are defined on these interfaces. The mortar pressure basis φ_{ki} corresponding to the coarse grid $\Omega_k{}^c$ is obtained by solving Eq. (19a) on $\Omega_k{}^c$ subject to the BC that $p = \mu_i$ on $\partial\Omega^c_k$. (c) Construction of a general pressure basis. The subdomain Ω has three open boundaries, $\Gamma_{1\text{-}3}$ (red), and one sealed boundary, Γ_w (dashed). A general pressure basis, for the Darcy-scale flow Eq. (19a) or pore-scale Stokes Eq. (49), is obtained by setting the pressure at Γ_1 to $p = 1$, and at Γ_2 and Γ_3 to $p = 0$. Two more pressure bases are obtained by setting $p = 1$ at Γ_2 and another time at Γ_3. For each pressure basis, the flowrates at $\Gamma_{1\text{-}3}$ are calculated and assembled into the columns of a flux matrix $\mathbf{M}$ (Section 3.4.4).

problem of "computing an unknown function p over Γ'' to the much simpler problem of "computing a handful of scalars p_i^c over Γ''. In short, if we know p_i^c, we have solved the problem.

We begin by first constructing a set of *pressure bases* on each subdomain (related to the *flux bases* of Ganis and Yotov (2009)). Let us consider the coarse grid $\Omega_k{}^c$ in Fig. 9b. The idea is to compute a set of basis functions on $\Omega_k{}^c$, one for each of the 12 nodes along $\partial\Omega^c_k$, such that if all the p_i^c corresponding to these nodes are known, we can reconstruct the fine-scale p inside $\Omega_k{}^c$ by superposing the bases. We denote the pressure basis associated with $\Omega_k{}^c$ and the interface node i (red in Fig. 9b) by φ_{ki}. To compute φ_{ki}, we first reduce Eq. (33) to $p|_\Gamma = \mu_i$, by setting the multiplier of node i to one and the rest to zero, and then use it as BCs to solve Eq. (19a) on $\Omega_k{}^c$. The solution is the fine-scale pressure basis φ_{ki}. The process is repeated for all i and k. Intuitively, the pressure basis φ_{ki} captures the effect of the coarse-scale pressure p_i^c over Γ on the fine-scale pressure p over $\Omega_k{}^c$. Put differently, pressure bases allow p_i^c to be downscaled onto the fine grid comprising $\Omega_k{}^c$. To compute p_i^c, we impose

$$\int_\Gamma [\![u(p^c)]\!]\mu_j\, d\Gamma = 0 \tag{34}$$

on Γ, which is a statement of mass conservation across subdomain interfaces. Eq. (34) says that the jump in flux, $[\![u]\!]$, across Γ must be zero in a "weak" sense. Note that Eq. (34) must hold for all j. Because $[\![u]\!]$ depends on the BCs imposed on each coarse grid, and such BCs depend on p_i^c according to Eq. (33), $[\![u]\!]$ is a function of p_i^c. In Eq. (34), we highlight this dependence with the notation $[\![u(p^c)]\!]$. A crucial observation is that the dependence between $[\![u]\!]$ and p_i^c is linear, because both Eq. (19a) and Eq. (33) are linear. We can therefore write

$$[\![u]\!] = \sum_i p_i^c\, \omega_i \tag{35}$$

where

$$\omega_i = \sum_k n_\Gamma \cdot (-\lambda \nabla \varphi_{ki})\Big|_\Gamma \tag{36}$$

The vector n_Γ is the unit normal on Γ. The weight ω_i in Eq. (35) represents the contributions of φ_{ki} for all k to the flux jump $[\![u]\!]$ on Γ. Notice ω_i is a fine-scale function defined on Γ. Substituting Eq. (35) and Eq. (36) into Eq. (34) yields a linear system in terms of p_i^c

$$\sum_i p_i^c \underbrace{\sum_k \left(\int_\Gamma n_\Gamma \cdot (-\lambda \nabla \varphi_{ki})\Big|_\Gamma \mu_j\, d\Gamma \right)}_{M_{ij}^k} = \sum_i p_i^c \underbrace{\left(\int_\Gamma \omega_i \mu_j\, d\Gamma \right)}_{M_{ij}} = 0 \tag{37}$$

Eq. (37) is the coarse-scale problem, and solving it completes the algorithm for MoMsFE.

The coefficients M_{ij} are the entries of a small $N^c \times N^c$ matrix, where N^c denotes the number of coarse-scale unknowns on Γ (dots in Fig. 9a). The parameter M_{ij} is an aggregate of subdomain contributions $M_{ij}{}^k$. We call the matrix $\mathbf{M}^k$, whose entries consist of $M_{ij}{}^k$, the *flux matrix* (same as *flux bases* in Ganis and Yotov (2009)). $\mathbf{M}^k$ is an upscaled representation of Eq. (19a) on $\Omega_k{}^c$. In Section 3.4.4, we elaborate more on flux matrices and discuss their properties. Here, we only mention that while Eq. (37) suggests $\mathbf{M}^k$ is a $N^c \times N^c$ matrix, it is sparse and contains many zero entries. If only the non-zero rows and columns are retained, $\mathbf{M}^k$ reduces to a 12×12 matrix for the subdomain $\Omega_k{}^c$ in Fig. 9b; because $\Omega_k{}^c$ has 12 nodes on its external boundary. The calculation of M_{ij} (or $M_{ij}{}^k$) requires only the mortar basis μ_j and the pressure basis φ_{ki}, both of which can be, respectively, selected or precomputed. Notice the coarse-scale parameter M_{ij} is similar to the coarse-scale stiffness K_{ij} in MsFE (Eq. (23)) and the coarse-scale transmissibility T_{ij} in MsFV (Eq. (29)). MoMsFE can thus be viewed as a form of numerical upscaling because the fine-grid pressure bases φ_{ki} can be discarded after all M_{ij} are computed.

We conclude with a few remarks. MoMsFE is able to estimate and control approximation errors by increasing either the order or the number of mortar basis functions μ_i per coarse-grid edge (Arbogast et al., 2007). For example, we could use 8 instead of 4 nodes along the edges of each subdomain in Fig. 9b, and we could use quadratics instead of linear hat functions. In MoMsFE, the fine grids inside two neighboring subdomains need not match (or conform) along their shared interface. This is very attractive because each subdomain can be discretized independently using different fine grids. In MoMsFE, the coarse-scale problem, Eq. (37), corresponds to a mass balance across subdomain *interfaces*, whereas in MsFV the coarse-scale problem, Eq. (28), corresponds to a mass balance over subdomain *volumes*. An interface balance is usually much easier to formulate than a volume balance, especially for nonlinear problems, because the latter requires careful averaging of the fine-scale equations to ensure consistency between the fine- and coarse-scale problems (see Lunati and Jenny, 2006). No such requirement exists in MoMsFE. One drawback of MoMsFE is that the fine-scale velocity u is only "weakly" conservative across Γ. Only the integral of u over the support of each mortar basis is conserved, but not its pointwise values. The smaller the support, the more locally is mass conserved. Piecewise discontinuous polynomial mortars offer the most flexibility in localizing mass conservation along Γ. An alternative to MoMsFE that ensures pointwise mass conservation on Γ is the enhanced velocity method (Wheeler et al., 2002; Thomas and Wheeler, 2011; Ganis et al., 2019).

3.4.2. Pore scale

At the pore scale, the equivalent of the Darcy-scale Eq. (19) for two-phase flow is, in its simplest form, the following (Popinet, 2018)

$$\nabla\cdot u = 0 \tag{38a}$$
$$\mu\,\Delta u - \nabla p + \kappa\sigma\delta_s = 0 \tag{38b}$$
$$\partial_t\alpha + \nabla\cdot(\alpha u) = 0 \tag{38c}$$

where

$$\kappa\delta_s = \nabla\cdot(\nabla\alpha/\|\nabla\alpha\|)\nabla\alpha \tag{38d}$$

Eqs. (38a) and (38b) denote mass and momentum conservation, respectively. Eq. (38c) describes the evolution of each phase with time. Eq. (38) assumes that the two phases are incompressible and have equal viscosity and density. The variables u, p, μ, σ, and κ denote velocity, pressure, viscosity, surface tension, and interfacial curvature, respectively. The surface Dirac delta, δ_s, ensures that capillary forces are localized at the fluid-fluid interface (Popinet, 2018). The scalar indicator function, α, takes the value 1 in one of the phases and 0 in the other. Across the fluid-fluid interface, α varies smoothly between 0 and 1. The field variable α, much like the saturation S_w in Eq. (19b), specifies how the two phases are spatially distributed at any given moment in time. Capillary forces are computed by inserting α into Eq. (38d). To solve Eq. (38) numerically, a common strategy is to first solve Eq. (38a-b) for u and p, using α from the previous time step, and then solve Eq. (38c) to obtain α at the next time step. By analogy with Section 3.4.1, we call Eq. (38a-b) the *flow equation* and Eq. (38c) the *transport equation*. As in Section 3.4.1, solving the flow Eq. (38a-b) is computationally much more expensive than solving the transport Eq. (38c). Multiscale methods have therefore been developed to accelerate the solution of the flow Eq. (38a-b). In the following, we review several of them. The fine-scale u obtained from these methods must be made conservative (or divergence-free) so that Eq. (38c) can be solved over the fine grid.

3.4.2.1. Straightforward extensions. Several extensions of the multiscale methods discussed in Sections 3.4.1.3–4 to the pore scale have been proposed. Tomin & Lunati (Tomin and Lunati, 2013) adapted MsFV to solve single-phase flow (Eq. (38a-b) without the $\kappa\sigma\delta_s$ term) and two-phase flow (a more general version of Eq. (38)) at the pore scale. The overall workflow is very similar to that of Section 3.4.1.3, which we briefly outline here. A pore-scale domain is divided into a number of rectangular primary, Ω_i^{p}, and dual, Ω_i^{d}, coarse grids. Four basis functions are constructed over each dual grid by solving

$$\nabla\cdot\psi_{ij} = 0 \tag{39a}$$
$$\mu\,\Delta\psi_{ij} - \nabla\varphi_{ij} = 0 \tag{39b}$$

on Ω_i^{d} for $j\in\{1,2,3,4\}$. The index j enumerates the dual grid corners. Eq. (39) differs from the Darcy-scale Eq. (22) in that the dual bases here consist of both a pressure, φ_{ij}, and a velocity, ψ_{ij}, component (not just pressure). Eq. (39) is solved subject to linearly varying pressure BCs along the edges of Ω_i^{d} with the corner pressures specified as in Fig. 8d. The BCs for velocity consist of setting the normal gradient of ψ_{ij} on $\partial\Omega_i^{\mathrm{d}}$ to zero and imposing no-slip conditions on the fluid-solid interface. Note that Eq. (39) does not include the capillary term, $\kappa\sigma\delta_s$, and corresponds therefore to single-phase flow. To capture the effect of capillarity, a correction function satisfying

$$\nabla\cdot\widetilde{\psi}_i = 0 \tag{40a}$$
$$\mu\,\Delta\widetilde{\psi}_i - \nabla\widetilde{\varphi}_i + \kappa\sigma\delta_s = 0 \tag{40b}$$

is constructed on Ω_i^{d}. The correction function consists of a pressure, $\widetilde{\varphi}_i$, and a velocity, $\widetilde{\psi}_i$, component. Eq. (40) is subjected to the same kind of BCs as Eq. (39) except that they are all homogenous (i.e., equal to zero). Note that Eq. (39) must be solved only once per dual grid, whereas Eq. (40) must be solved every time step provided $\kappa\sigma\delta_s$ is non-zero over the dual grid (i.e., the dual grid is occupied by two phases). Lastly, the fine-scale p and u are reconstructed by superposition

$$p = \sum_j\sum_i p_{ij}^c\varphi_{ij} + \sum_i\widetilde{\varphi}_i \tag{41a}$$
$$u = \sum_j\sum_i p_{ij}^c\psi_{ij} + \sum_i\widetilde{\psi}_i \tag{41b}$$

Eq. (41) assumes that the basis and correction functions are extended by zero outside of the dual grids they are defined on. To compute the coarse-scale pressure unknowns, p_{ij}^c, a coarse-scale problem is formulated by integrating the fine-scale Eq. (38a) over each primary grid Ω_i^{p}. Tomin and Lunati (2013) then postulated that the net flux crossing $\partial\Omega_i^{\mathrm{p}}$ is a *linear* function of the difference between p_{ij}^c and its neighboring coarse-scale pressure values; we call this the *Darcy postulate*. Inserting the postulate into the integrated equation yields a linear system in terms of p_{ij}^c. As in Section 3.4.1.3, the u obtained from Eq. (41b) is not conservative over the fine grid and cannot be used to solve Eq. (38c). To remedy this, the u and p from Eq. (41) are used as BCs for a pure-Neumann problem (not shown) on each primary grid Ω_i^{p}. Solving this problem yields a conservative u.

The accuracy of the above MsFV extension relies on the degree of scale separation between the pore-scale characteristic length, ε, and the coarse-grid size, h. If $\varepsilon/h\sim 1$, the scheme incurs large errors (Tomin and Lunati, 2013) because the Darcy postulate is invalid. This is quite reminiscent of the resonance effect in MsFE (Hou and Wu, 1997) (Section 3.4.1.1). The above extension of MsFV was later recast as an algebraic solver (Tomin and Lunati, 2015) and several adaptivity criteria were explored (Tomin and Lunati, 2016b) to increase its computational efficiency. A particularly important criterion was to switch to a Darcy-scale description when pore-scale changes inside a primary grid became negligible. In drainage and imbibition problems, the criterion allows one to localize computations to the displacement front. Its drawback, however, is that once the switch to the Darcy scale is made, some of the fine-scale information are permanently lost, which may be important in, for example, cyclic flows (drainage followed by imbibition).

A second extension of MsFV to the pore scale was proposed by Khayrat and Jenny (2017) and Khayrat et al. (2018), where the fine-scale solver was a pore-network model (PNM). The PNM, a graph-based method that we explain later, was an adaptation of existing methods (Joekar-Niasar et al., 2010; Thompson, 2002) that crudely approximates Eq. (38). The MsFV extension was shown to accelerate the solution of the flow equation by a factor of 2–10, providing a means to probe samples that are much larger than possible with single-scale PNM. The overall workflow is nearly identical to that of Section 3.4.1.3, with minor differences specific to the PNM solver.

Extensions of MoMsFE to the pore scale have also been reported by Balhoff et al. (2008), Sun et al. (2012a, 2012b), Mehmani and Balhoff (2014) and Mehmani et al. (2012). In these works, the fine-scale solver was a single-phase PNM (two-phase flow was not explored). The overall workflow is very similar to the Darcy scale MoMsFE discussed in Section 3.4.1.4, except that the mortar and pressure bases are defined over discrete domains (because PNMs are graphs). The discrete nature of the pore space introduces specific challenges to defining mortar bases, μ_i in Eq. (33). For example, mortar bases may have supports that lie entirely in the solid phase. Another difficulty is to ensure pointwise flux conservation over subdomain interfaces Γ. Standard polynomial mortars (e.g., hat functions) produce fine-scale velocity fields that are not pointwise conservative. Mehmani and Balhoff (2014) and Mehmani et al. (2012) proposed specially designed piecewise constant mortar bases that rendered the velocity conservative across Γ, and thus useful for simulating solute transport. Despite this advantage, piecewise constant mortars can produce artifacts in simulating solute transport if the flow direction is tangent to Γ (Mehmani, 2014). Piecewise linear mortars would circumvent this issue, but remain underexplored (Mehmani, 2014). Similar MoMsFE extensions have been reported to model biofilm growth (Tang et al., 2015) and fuel cells (Baber et al., 2016).

3.4.2.2. Pore-level multiscale method (PLMM). A multiscale method specifically designed for pore-scale problems was recently proposed by Mehmani and Tchelepi (2018, 2019). The pore-level multiscale method (PLMM) is not an extension of the Darcy-scale methods reviewed in Section 3.4.1 but inspired by several ideas therein. Here, we describe PLMM with reference to Fig. 10a-d. Consider the pore-scale domain in Fig. 10a, where the black color denotes the solid phase. In PLMM, the void space, Ω, is decomposed into a number of coarse grids, Ω_i^c, by using a well-known image-analysis algorithm called *watershed segmentation* (Beucher and Lantuejoul, 1979). A key property of watershed segmentation is that each coarse grid, Ω_i^c, corresponds to a local enlargement of Ω, called a *pore*. And the interface, Γ_{ij}, between two adjacent coarse grids, Ω_i^c and Ω_j^c, corresponds to a local constriction of Ω, called a *throat*.

In PLMM, basis functions are constructed on each coarse grid as follows. Consider Ω_i^c in Fig. 10b, which shares five interfaces (red lines), Γ_{ij}, with its neighboring coarse grids Ω_j^c, where $j \in \{1,2,3,4,5\}$. Eq. (39) is solved on Ω_i^c to obtain the pressure, φ_{ij}, and velocity, ψ_{ij}, basis functions, respectively. The BCs imposed at Γ_{ik}, for $k \in \{1,2,3,4,5\}$, are $\varphi_{ij} = \delta_{jk}$ and $\partial_n \psi_{ij} = 0$, where δ_{jk} is the Kronecker delta and ∂_n is the derivative normal to Γ_{ik}. The BC at the fluid-solid interface is the no-slip velocity condition. Fig. 10d illustrates all five of the basis functions on Ω_i^c. Red dots mark interfaces, Γ_{ik}, where the pressure basis, φ_{ij}, is equal to one, and gray dots mark interfaces, where φ_{ij} is equal to zero. Notice that, similar to Section 3.4.2.1, φ_{ij} and ψ_{ij} are local solutions to the single-phase Stokes equation. To account for capillarity, correction functions for pressure, $\widetilde{\varphi}_i$, and velocity, $\widetilde{\psi}_i$, must be constructed by solving Eq. (40) on Ω_i^c subject to homogeneous BCs (same as Section 3.4.2.1). Once all basis and correction functions are computed, the fine-scale velocity, u, and pressure, p, can be reconstructed using Eq. (41). The coarse-scale unknowns, p_{ij}^c in Eq. (41), correspond to scalar pressure values that are constant along each Γ_{ij}. Fig. 10a illustrates the spatial arrangement of p_{ij}^c by the green dots. To compute p_{ij}^c, a coarse-scale problem is formulated by imposing mass conservation across Γ_{ij}

$$\int_{\Gamma_{ij}} [\![u(p^c)]\!] d\Gamma = 0 \tag{42}$$

where $[\![u]\!]$ is the jump in flux across Γ_{ij}. In PLMM, similar to Section 3.4.1.4, $[\![u]\!]$ is a linear function of p_{ij}^c (hence the notation $[\![u(p^c)]\!]$). Therefore, Eq. (42) can be written as a linear system in terms of p_{ij}^c. Solving this system yields p_{ij}^c, which is then used to reconstruct u from Eq. (41). This u, however, is not *pointwise* conservative (or continuous) across Γ_{ij}. The reason is that Eq. (42) imposes mass conservation only in a "weak" sense. To remedy this shortcoming, u is modified by solving a local problem over a small region containing Γ_{ij}. Such regions are called *throat grids* (not shown in Fig. 10) and constitute a small fraction of Ω; typically <10%. The modification renders u conservative and thus suitable for solving the transport Eq. (38c). Throat grids are also used to iteratively improve PLMM predictions and to estimate and control errors.

PLMM bears some similarities with MoMsFE, MsFV, and two-level Schwarz methods (Dolean et al., 2015). Like MoMsFE, the coarse-scale problem is formulated as an interface balance equation (compare Eq. (42) to Eq. (34)), as opposed to a volume balance equation in MsFV. Interface conditions are more robust and easier to implement, because their mathematical form remains unaltered when applied to flow problems more complicated than Eq. (38a-b) (e.g., nonlinear compressible flow; see Guo et al., 2019). In MsFV, by contrast, one must exercise care to ensure consistency between fine-scale and coarse-scale equations, which may require a priori approximations about the latter (Lunati and Jenny, 2006). Like two-level Schwarz methods (Dolean et al., 2015), PLMM uses an overlapping region around Γ_{ij} to iterate until convergence. But unlike two-level Schwarz, and like MsFV, each iteration involves the calculation of local correction functions that are also used to render u divergence-free.

In PLMM, the basis functions, φ_{ij} and ψ_{ij}, are computed only once per simulation provided that both phases have a constant density and viscosity. The correction functions, $\widetilde{\varphi}_i$ and $\widetilde{\psi}_i$, are recomputed every time step. Careful inspection of Eq. (40), however, reveals that $\widetilde{\varphi}_i$ and $\widetilde{\psi}_i$ are only non-zero when Ω_i^c is occupied by two phases. Fine-scale computations can therefore be confined to the fluid-fluid interface. In displacement processes, this means coarse grids ahead and behind the displacement front are exempt from fine-scale computations. Only coarse grids that lie *at* the front require $\widetilde{\varphi}_i$ and $\widetilde{\psi}_i$ to be computed. A similar adaptivity criterion was used in the first MsFV extension (Tomin and Lunati, 2013; Tomin and Lunati, 2016b) discussed in Section 3.4.2.1. The key difference here is that in PLMM, coarse grids correspond to physical pores, whereas in MsFV, coarse grids are rectangular boxes containing tens to hundreds of pores. Because pore-filling events (e.g., Haines jump, snap-off) occur at the scale of individual pores in a

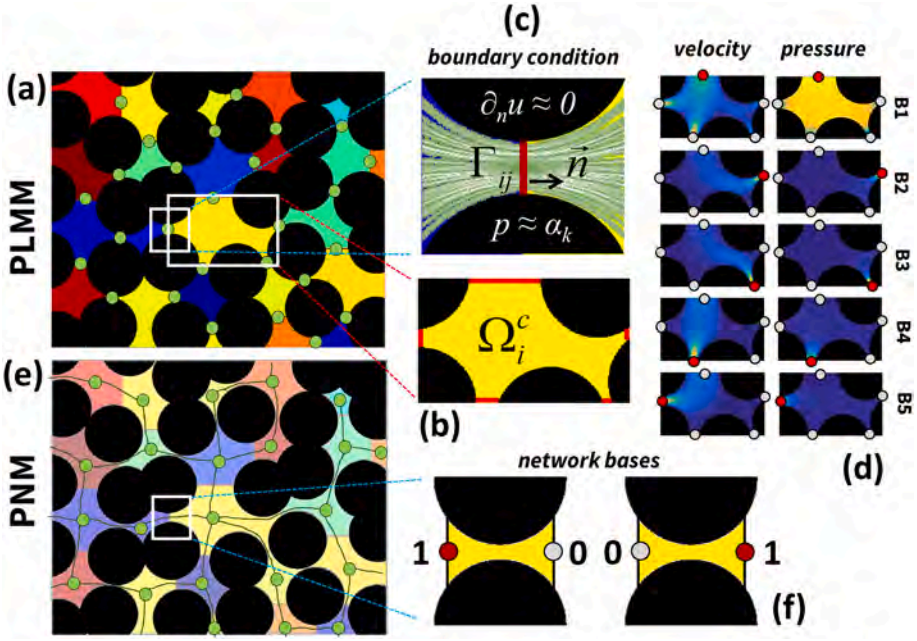

Fig. 10. Schematic of PLMM and PNM. (a) Coarse grids are the randomly colored regions. Green dots show the placement of coarse-scale unknowns at coarse-grid interfaces. (b) The coarse grid Ω_i^c shares five interfaces (red) with its neighbors. (d) The BCs used to build local bases on Ω_i^c are informed by how streamlines (green) meander near each interface Γ_{ij} (red line), i.e., they are optimal. (e) Pressure and velocity bases (B1 to B5) are obtained by setting the pressure equal to one and zero at the interfaces marked by red and gray dots, respectively. The velocity BC is set to have a zero normal gradient across all the interfaces. (f) PNM interpreted as an upscaling method. In PNM, coarse-scale unknowns are placed at the center of each coarse grid (or pore). The fine-scale pressure is represented by a piecewise constant function that is constant over each pore's volume. Two basis functions are calculated for each throat by solving the Stokes equation. Red and gray dots denote one and zero pressure BCs, respectively. The construction of bases in PNM is equivalent to computing throat conductivities.

porous medium, it is far easier to define adaptivity criteria in PLMM. The adaptivity criteria in MsFV (Tomin and Lunati, 2016b) are expressed in terms of integral measures such as changes in average phase saturation. But saturation is not a good measure, because the same value may correspond to different phase distributions at the pore scale. In PLMM, the adaptivity criterion requires a simple check to see whether Ω_i^c is occupied by one or two phases. Even more efficient criteria have been proposed (Mehmani and Tchelepi, 2019).

Differences in the sizes and shapes of coarse grids in PLMM versus MsFV (Tomin and Lunati, 2013) have another implication. Unlike MsFV, the BCs used to construct PLMM basis functions are informed by the local pore-scale geometry. Fig. 10c depicts an interface Γ_{ij} between two adjacent coarse grids. The flow streamlines are drawn in green color. PLMM assumes that the fine-scale pressure along Γ_{ij} is approximately constant and equal to p_{ij}^c. The gradient of velocity normal to Γ_{ij} is also assumed to be zero. Since watershed segmentation ensures Γ_{ij} always coincides with a local constriction, streamlines are bound to exhibit a converging-diverging pattern near Γ_{ij}. It is therefore possible to show (Mehmani and Tchelepi, 2018) that the optimal BCs on Γ_{ij} are the same as those imposed by PLMM. This has been confirmed by highly accurate predictions (error $< 1\%$) reported for different pore-scale geometries (Mehmani and Tchelepi, 2018). By comparison, MsFV imposes linearly varying pressure and zero-gradient velocity BCs over box-shaped coarse grids, regardless of what the underlying pore-scale geometry looks like. This leads to less accurate predictions.

In PLMM, convergence to the fine-scale solution is rapid. The main reason is that instead of using a generic smoother, like incomplete LU, local problems are solved near Γ_{ij} (or on throat grids) to obtain more accurate BCs on Ω_i^c. Most of the error reduction occurs within the first iteration, often by a factor of 10–100 (Mehmani and Tchelepi, 2018; Mehmani and Tchelepi, 2019). For most subsurface applications, either *zero* or *one* iteration is sufficient to obtain an accurate solution. Generalizations of PLMM to compressible flow (Guo et al., 2019) and geomechanics (Mehmani et al., 2021) have also been reported.

3.4.2.3. Pore networks: upscaling or multiscale?. Pore-network models (PNM) are approximate methods for solving flow and transport problems at the pore scale. A PNM represents the complex microstructure of a porous medium with a graph consisting of nodes and links. Fig. 10e shows an example of a pore network. Green dots are the nodes and black lines are the links. Each node lies at the center of a pore, represented in Fig. 10e by the colored regions. Links connect the nodes across the interfaces (or throats) between the pores. All the information about a pore, such as its volume and surface area, are assigned to the corresponding node. And all the information about a throat, such as its inscribed radius and capillary entry pressure, are assigned to the corresponding link. No geometric details of either the pore or the throat are resolved. Only their "upscaled" impact on the physics is captured. The graph is used to simulate flow and transport by solving equations that roughly approximate Eq.38. The simplest example corresponds to single-phase flow, in which mass balance is imposed at each node i

$$\sum_{j=1}^{N_i} g_{ij}(p_j - p_i) = 0 \tag{43}$$

The subscript ij denotes the link between nodes i and j. In Eq. (43), p_i and p_j are the nodal pressures; g_{ij} is the hydraulic conductivity of link ij; and N_i is the number of links connected to node i. Eq. (43) equates the sum of the flowrates, $q_{ij} = g_{ij}(p_j - p_i)$, over all the links connected to node i to zero. The fluid is assumed to be incompressible and Newtonian. Implicit to Eq. (43) are two important assumptions: (1) g_{ij} captures the hydraulic resistance of the microstructure between nodes i and j, and (2) the pressure field inside each pore, colored regions in Fig. 10e, is approximately constant and can therefore be represented by a *scalar*, p_i. To the initiated reader, the formulation of Eq. (43) is identical to the two-point flux approximation (TPFA) scheme in finite volume.

In the 1950–80s, PNMs were used primarily as conceptual tools to provide qualitative understanding of pore-scale physics. Geometric details of the pore space were difficult to measure, and g_{ij} in Eq. (43) was often determined based on intuitive, yet ad hoc, assumptions. From the 1990s onward, X-ray μ-CT imaging provided rich descriptions of microstructural details of porous media. As a result, various PNM algorithms (Lindquist et al., 1996; Sheppard et al., 2006; Dong and Blunt, 2009; Silin and Patzek, 2006) were developed to construct realistic graphs with more accurate parameters, like g_{ij} (Prodanović et al., 2007). To compute g_{ij} in a modern PNM, the Stokes equation is solved on the actual void geometry representing the ij link, as shown in Fig. 10f. A unit pressure difference is imposed across the two ends of the link and the calculated flowrate through the inlet (or outlet) equals g_{ij}. Similar uses of direct numerical simulation (DNS) to compute network parameters have been proposed for solute transport and two-phase flow (Mehmani and Tchelepi, 2017; Raeini et al., 2018).

Using the terminology of this section, modern PNMs require some amount of local fine-scale (DNS) computations to parameterize a global coarse-scale equation, e.g., g_{ij} in Eq. (43). The calculation of g_{ij} is identical to constructing a local basis function. For the throat shown in Fig. 10f, two bases can be built by setting the pressure equal to one at the boundaries marked by red nodes and to zero at the boundaries marked by gray nodes. However, the two bases are *not* independent, because their sum equals one. Using either basis to compute the flowrate through the inlet (or outlet) of the throat yields g_{ij}. This procedure resembles numerical upscaling, but not multiscale computing. Because similar to upscaling, the form of the coarse-scale Eq. (43) is first assumed and then parameterized. And unlike multiscale computing, the ability to downscale fine-scale (or sub-pore) details is absent. Modern PNMs can therefore be considered as a form of numerical upscaling, with all the pros and cons discussed in Section 3.3 attached.

In Section 3.4.2.2, we presented PLMM as a multiscale method, but another interpretation of it is: a more accurate alternative to PNM. Despite the wide-reaching success of PNM in improving our understanding of various pore-scale processes, it has four major drawbacks: (1) it yields poor predictions except in idealized geometries and specific flow regimes[5]; (2) the predictions are not bound by confidence intervals and the ability to control errors is absent; (3) the approximation of a given microstructure by a pore network is often ambiguous and non-unique, but has a large impact on predictions; and (4) the solutions represent *averages* of pressure and velocity within pores and throats with no option to recover sub-pore details. Such details are known to be important in, for example, advection-dominated solute transport (Mehmani and Tchelepi, 2017; Mehmani et al., 2014; Mehmani and Balhoff, 2015b). Multiscale methods that use PNM (Mehmani and Balhoff, 2014; Balhoff et al., 2008; Mehmani et al., 2012; Khayrat and Jenny, 2017; Khayrat et al., 2018) as their fine-grid solver (to construct bases) inherit all of these limitations. PLMM, by contrast, is beleaguered by none.

While not reported here, it is possible to recast PNM as a multiscale method using the MsFV formalism. An outline of such a formulation was presented by Mehmani and Tchelepi (2018), whereby local basis functions are systematically constructed and a coarse-scale problem is formulated by integrating fine-scale equations. Compared to PLMM, however, the approach is expected to be less accurate because the bases are obtained using less accurate BCs (Section 3.4.2.2). Even so, the formulation would place PNM on solid theoretical footing and improve its predictions. Errors would be quantifiable and controllable, and the

[5] PNM predictions are reliable in only a handful of problems: single-phase flow, quasi-static drainage, and diffusive transport. The first and third are characterized by pressure and concentration fields that are smooth, and thus easy to approximate with piecewise constants. Outside these problems, favorable results are often reported for *averaged* quantities such as relative permeability or capillary pressure, which obscure local but potentially large errors.

ability to downscale sub-pore details available if desired.

3.4.3. Multiscale vs. upscaling: to downscale or not?

In Sections 3.3–4, we outlined the differences between numerical upscaling and multiscale methods. The most important were the ability of multiscale methods (1) to estimate and control errors and (2) to downscale fine-scale details. Almost all upscaling methods, including PNM at the pore scale, lack the first capability (Section 3.3) and only some possess the second (Chen et al., 2003; Gautier et al., 1999). Here, we focus on downscaling and discuss whether it is necessary for obtaining reliable predictions of subsurface processes. The question is important because both the cost and accuracy of simulations depend on the amount of fine-scale details resolved.

In all the multiscale methods discussed in Sections 3.4.1–2, downscaling is the last step of the algorithm. Namely, a fine-scale solution is reconstructed by superposing local basis functions weighted by the solution of a coarse-scale problem. Of particular interest is the fine-scale velocity field, which must be conservative (or divergence-free) to solve a transport equation on the fine grid. In subsurface applications, coarse-scale predictions are often more desirable than fine-scale predictions. The reason, aside from lower computational cost, is that engineering decisions require information either at a few select locations, such as flowrates at wells, or in the form of averages over larger regions, such as permeabilities of meter-scale cores. Both upscaling and multiscale methods are capable of providing such coarse-scale information, but the latter also presents the option to downscale. So the question is whether this extra downscaling step is important for obtaining accurate *coarse-scale* predictions. As it turns out, the answer depends on the problem and our tolerance for error. In the following, we assume that we have no desire to estimate or control prediction errors, which is a feature generally absent in numerical upscaling and one that invariably demands downscaling in multiscale methods.

In single-phase flow, downscaling is redundant if all we want is an *approximate* coarse-scale solution to an elliptic (or time-independent) problem without needing to use the solution to solve a subsequent transport problem. Quasi-static solid mechanics, in which a sequence of elliptic problems are solved to capture multiple loading steps, also falls under this category (Castelletto et al., 2017). Recall that all multiscale methods reduce their fine-scale basis functions into some kind of coarse-scale parameter such as the coarse-scale stiffness in MsFE, the coarse-scale transmissibility in MsFV, and the flux matrix in MoMsF (see Section 3.4.1.4). As soon as these coarse-scale parameters are computed, the fine-scale bases can be discarded because the former is the only thing that is required to compute an approximate coarse-scale solution. Two-phase flow, by contrast, is time-dependent and consists of an elliptic pressure equation, Eq. (19a), and a hyperbolic transport equation, Eq. (19b). It is the precise character of the transport equation that dictates whether downscaling is necessary or not. In Section 3.4.1, we mentioned hyperbolic equations have solutions that exhibit localized features such as steep gradients. One example is a shock front formed by displacing one fluid by another inside a porous medium. Such features can only be resolved on the fine grid, requiring that some (localized) downscaling is performed. Now, if we were to add capillarity to Eq. (19b), then the transport equation would have a mixed character: part hyperbolic and part parabolic. If the parabolicity, or capillarity, dominates, then sharp shocks in saturation would give way to smooth and diffuse transition zones that may be resolvable on the coarse grid alone. If so, then downscaling would be redundant; notice this depends on the coarse-grid size. A similar argument holds for single-phase transport of a passive solute in groundwater. If advection dominates over diffusion, then the fine-scale details are important, at least in some parts of the domain, and must be downscaled. Else, one may simulate the physics entirely on the coarse grid. Pore-scale equations, Eq. (38), abide by the same rules. In single-phase Stokes flow, downscaling is redundant if all we want is to compute an approximate macroscopic permeability for a sample. But in two-phase flow and advection-dominated solute transport, downscaling is necessary due to the presence of menisci and sharp boundary layers (Tomin and Lunati, 2013; Mehmani and Tchelepi, 2018; Mehmani and Tchelepi, 2019; Mehmani and Tchelepi, 2017).

What happens if we insist on solving a hyperbolic equation on a coarse grid that cannot resolve its dynamics, i.e., the problem requires downscaling but we ignore it? An example is if we discretize and solve the transport Eq. (19b) on the *coarse grid*. The practice is common in earlier numerical upscaling methods (e.g., not local-global; Chen et al., 2003). While the task can be accomplished by both upscaling and multiscale methods, it is highly error prone and should be avoided. In the case of Eq. (19b), the result would be severe numerical dispersion of the saturation field and the distortion of simulated injection fronts or contaminant plumes (Salamon et al., 2006). The reason is that the homogenized form of a hyperbolic equation defined on the fine grid is not another hyperbolic equation defined on the coarse grid, but rather an integral equation containing convolutions; or very high-order derivatives (E, 1992; Tartar, 1989; Mehmani and Balhoff, 2015b; Berkowitz et al., 2016). The integrals, even local approximations to them (Efendiev and Durlofsky, 2003), require knowledge of the fine-scale solution. The implication: to accurately solve hyperbolic problems on a coarse grid, some form of downscaling is necessary. The good news is that the downscaling can be localized because the solutions of hyperbolic equations exhibit steep gradients on only small parts of the domain, e.g., shock fronts in fluid-fluid displacements (Jenny et al., 2006; Mehmani and Tchelepi, 2019).

The oft-cited assertion that the ultimate goal of multiscale methods, as opposed to numerical upscaling, is to obtain a fine-scale solution is therefore unjustified. Technical feasibility is not the issue, as both methods can produce coarse-scale predictions. The need to downscale is driven less by a desire to obtain a fine-scale solution and more by the underlying character of the governing equations.

3.4.4. Flux matrix: a bridge between pore scale and Darcy scale

We now outline a general and systematic approach for bridging between the pore scale and the Darcy scale. The proposed method is purely algorithmic and combines MoMsFE and PLMM within a nested hierarchical framework. It makes no assumptions about the form of the governing equations at the Darcy scale and does not require scale separation to hold. It also allows for both upscaling and downscaling to be performed. The method revolves around the idea of a *flux matrix*, first encountered in Section 3.4.1.4 as $\mathbf{M}^{\mathrm{k}}$, which is an upscaled representation of the fine-scale equations over a subdomain (or coarse grid). Both MoMsFE and PLMM make use of flux matrices because their coarse-scale unknowns and corase-scale problems are, respectively, defined and formulated over subdomain interfaces (Figs. 9a and 10a). By itself, the idea of a flux matrix is not new (Ganis and Yotov, 2009; Mehmani and Tchelepi, 2018; Sun, 2012), as it has been separately developed at both the Darcy scale and the pore scale. But its use as a scale-bridge in the way outlined here is. In the following, we begin by introducing a more general description of a flux matrix for an arbitrary (pore- or Darcy-scale) subdomain and then discuss how this fits within the proposed scale translation framework.

Consider the subdomain Ω in Fig. 9c, whose boundary $\partial\Omega$ consists of three open segments $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ (solid red) and one closed segment Γ_w (dashed black). For simplicity, let the fine-scale equation defined on Ω be the Darcy-scale flow equation, repeated here for convenience

$$\nabla\cdot(\lambda\nabla p) = 0 \tag{44}$$

The flux matrix associated with Ω is constructed as follows. Set the

pressure, p, equal to one at Γ_1 and to zero at Γ_2 and Γ_3; as shown in Fig. 9c. Impose no-flow BCs at Γ_w. Solve Eq. (44) and compute the total flowrate crossing each open boundary via

$$q_i^1 = \int_{\Gamma_i} u \, d\Gamma = -\int_{\Gamma_i} \lambda \nabla p \, d\Gamma \qquad \forall i \in \{1,2,3\} \tag{45}$$

In Eq. (45), q_i^1 denotes the total flowrate at Γ_i obtained by solving Eq. (44) subject to $p = 1$ at Γ_1 and $p = 0$ at Γ_2 and Γ_3. We next arrange all q_i^1 into the entries of a 3×1 column vector Q^1

$$Q^1 = [q_1^1 \;\; q_2^1 \;\; q_3^1]^T \tag{46}$$

The superscript T denotes transposition. We repeat the above steps by setting $p = 1$ at Γ_2 and $p = 0$ at Γ_1 and Γ_3. We then solve Eq. (44) and use an equation similar to Eq. (45) to obtain q_i^2 for $i \in \{1,2,3\}$. The 3×1 column vector Q^2 is obtained by arranging the q_i^2 into its entries, as in Eq. (46). In much the same way, Q^3 is computed by setting $p = 1$ at Γ_3 and $p = 0$ at Γ_1 and Γ_2, followed by repeating the foregoing steps. The flux matrix associated with Ω (Fig. 9c) is a 3×3 matrix $\mathbf{M}$ with Q^1, Q^2, and Q^3 as its columns

$$\mathbf{M} = [\, Q^1 \;\; | \;\; Q^2 \;\; | \;\; Q^3 \,] \tag{47}$$

Note that each column of $\mathbf{M}$ is a condensed representation of a fine-scale solution of Eq. (44), which we called a *pressure basis* in Section 3.4.1.4. Here, we have used the symbol p instead of φ to denote pressure bases, as it makes the exposition more intuitive. Note that if Ω in Fig. 9c consists of 3000 fine grids, then the coarsening (or upscaling) ratio achieved by $\mathbf{M}$ is equal to 1000, because its construction requires 3 separate solutions of Eq. (44). The most important use of $\mathbf{M}$ is that it can relate arbitrary pressure values imposed on Γ_1, Γ_2, and Γ_3 to corresponding flowrates at these boundaries. To demonstrate, let P and Q in Eqs. (48a-b) denote 3×1 column vectors that contain the pressures, p_i, and total flowrates, q_i, at Γ_i for $i \in \{1,2,3\}$. Eq. (48c) then holds by superposition

$$P = [p_1 \;\; p_2 \;\; p_3]^T \tag{48a}$$

$$Q = [q_1 \;\; q_2 \;\; q_3]^T \tag{48b}$$

$$Q = MP \tag{48c}$$

Namely, given *any* P, left-multiplication of it by $\mathbf{M}$ yields the corresponding Q. No information regarding the details of the fine-scale equation or its discretization on Ω is required. We can generalize this further: if Ω has n open boundaries Γ_i (instead of 3 in Fig. 9c), then $\mathbf{M}$ would be an $n \times n$ matrix satisfying Eq. (48c). The "open boundaries" need not even be disjoint or separated by Γ_w, as is the case in Fig. 9c. In MoMsFE, they correspond (loosely) to the support of each mortar basis on $\partial\Omega$. In Fig. 9a, the flux matrix corresponding to the coarse grid Ω_k^c, $\mathbf{M}^k$, is a 12×12 matrix; because there are 12 nodes along $\partial\Omega_k^c$. The entry M^k_{ij} of $\mathbf{M}^k$ (Eq. (37)) is obtained by first solving Eq. (44) subject to the pressure BC $p|_\Gamma = \mu_i$ on $\partial\Omega_k^c$ and then projecting the boundary flux derived from that solution (i.e., pressure basis φ_{ki}) onto the mortar basis μ_j as outlined by Eq. (37) (see Fig. 9b). Hence, if n mortar bases are defined on $\partial\Omega$, then $\mathbf{M}^k$ is an $n \times n$ matrix. What makes Eq. (48c) so attractive is that an identical relation holds for the (Stokes) flow equation at the pore scale

$$\nabla \cdot u = 0 \tag{49a}$$

$$\mu \Delta u - \nabla p = 0 \tag{49b}$$

In PLMM, a unique $\mathbf{M}$ is calculated for every coarse grid (or pore) shown in Fig. 10a; called $\mathbf{G}$ in Mehmani and Tchelepi (2018, 2019). The procedure is the same as described above, except that Eq. (44) is replaced by Eq. (49) and a homogeneous velocity BC is imposed on all Γ_i (i.e., zero normal gradient). For the coarse grid in Fig. 10b, $\mathbf{M}$ is a 5×5 matrix that upscales the fine-scale bases shown in Fig. 10d. Other flux matrices in the context of PNM have been reported (Sun, 2012). But given the starting point of PNM is Eq. (43), a crude approximation to Eq. (49), the resulting $\mathbf{M}$ is less accurate.

Eq. (48c) bears a strong resemblance to (the discretized) Darcy's law. One may therefore be tempted to assume that $\mathbf{M}$ has properties similar to permeability: symmetry and positive definiteness. $\mathbf{M}$ is indeed symmetric for both Eq. (44) and (49). A simple proof of this is sketched in Appendix A for Eq. (44), where the symmetry of $\mathbf{M}$ is deduced from that of λ and the fine-grid discretization stencil. But unlike permeability, $\mathbf{M}$ is singular. Specifically, $\mathbf{M}$ has a zero column-sum, because the fluid is incompressible and so the boundary flowrates corresponding to each pressure basis must sum to zero. $\mathbf{M}$ also has a zero row-sum, because the pressure bases of Eq. (44) and (49) form a partition of unity over Ω, which is easy to verify. The diagonal entries of $\mathbf{M}$ are positive, because they correspond to inflow through parts of $\partial\Omega$ where $p = 1$ is imposed, and its off-diagonal entries are negative, because they correspond to outflow through parts of $\partial\Omega$ where $p = 0$ is imposed. Hence, $\mathbf{M}$ is diagonally dominant. From symmetry and diagonal dominance follows positive semi-definiteness. Note that if $\mathbf{M}$ is $n \times n$, only the first n - 1 columns need to be computed, because subtracting their summation from one yields the n^{th} column. The above properties of $\mathbf{M}$, in the context of MoMsFE applied to PNM, have been numerically explored by Sun (2012).

Given that Eq. (48c) remains invariant between the pore scale and the Darcy scale, it is possible to coarsen Eq. (49) *recursively*, starting directly from the pore scale. The approach involves a hierarchical, or multilevel, decomposition of a porous medium into a nested sequence of coarse grids. Fig. 11 illustrates the idea. The domain at level 3 is decomposed into three nested coarse grids at levels 2, 1, and 0. Level 0 corresponds to the pore scale and is governed by Eq. (49). The decomposition of level 1 into level 0 is obtained by watershed segmentation as discussed in Section 3.4.2.2. The remaining levels are decomposed into Cartesian coarse grids, although more general coarse grids are also possible. At level 0, an $\mathbf{M}$ matrix is built for each coarse grid by solving Eq. (49). This step is identical to the construction of local bases in PLMM, shown in Fig. 10d, followed by their compression into $\mathbf{M}$ via Eqs. (45)–(47).

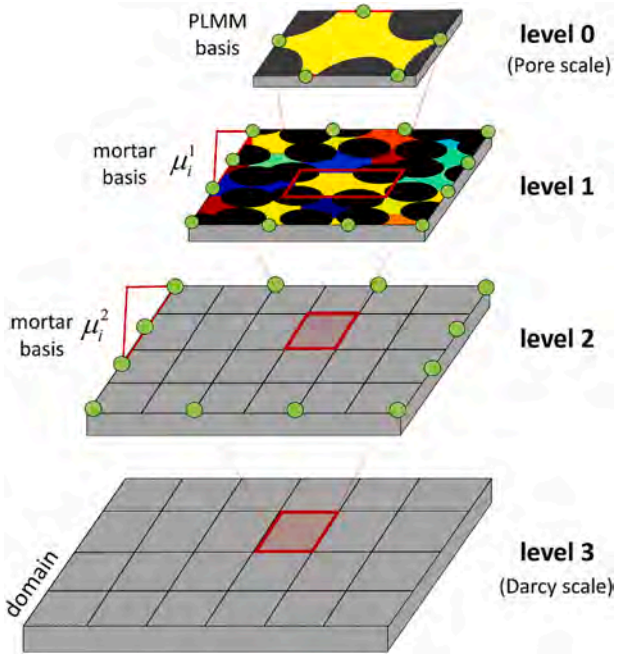

Fig. 11. Multilevel decomposition of a domain (level 3) into a nested hierarchy of coarse grids (levels 0–2). Level 0 corresponds to the pore scale and levels 3 to the Darcy scale.

For clarity of exposition, we use the nomenclature "level-x coarse grid" to refer to the single coarse grid depicted by Fig. 11 *at* level x. Similarly, we use "level-x fine grid" to denote the fine grids (or subdomains) that comprise each level-x coarse grid. In Fig. 11, for each level-1 coarse grid, an **M** matrix is constructed by combining the **M** matrices at level 0. The combination is achieved by first defining a number of mortar bases, μ_i^1, on the boundary of each level-1 coarse grid, e.g., hat functions in Fig. 11. Corresponding to each μ_i^1, a pressure basis is computed over the level-1 coarse grid using the MoMsFE procedure outlined in Section 3.4.1.4 (see Fig. 9b). Namely, a unit pressure is first imposed on the boundary of the level-1 coarse-grid using the mortar bases (i.e., $p = \mu_i^1$), then a pressure field inside the coarse grid is computed. This pressure field is expressed in terms of the level-1 fine-grid unknowns, which are the same as the level-0 coarse-grid unknowns. The coarse-scale problem at level 0, Eq. (42) in PLMM, can therefore be thought of as the "fine-scale problem" at level 1. The crucial point is that to solve Eq. (49) on the level-1 coarse grid, only the **M** matrices of the level-0 coarse grids are needed, but none of the pore-scale details (geometric or otherwise) defined on the level-0 fine grid. The pressure basis functions obtained in this manner on the level-1 coarse grid are subsequently condensed into an **M** matrix using Eqs. (45)–(47).

The computation of **M** for a level-2 coarse grid proceeds in a similar vein, requiring all the level-1 **M** matrices contained within the level-2 coarse grid to be combined. First, mortar bases, μ_i^2, are defined on the boundary of the level-2 coarse grid, e.g., hat functions in Fig. 11. The mortar bases are then used to compute a set of pressure bases over the level-2 coarse grid. These bases are expressed in terms of the level-2 fine-grid unknowns, which are the same as the level-1 coarse-grid unknowns. The coarse-scale problem at level 1, Eq. (37) in MoMsFE, can therefore be thought of as the "fine-scale problem" at level 2. Note that to solve Eq. (49) on the level-2 coarse grid, only level-1 **M** matrices are required, but not level-0 **M** matrices or any pore-scale information defined on the level-0 fine grid. The pressure basis functions obtained in this manner on the level-2 coarse grid are subsequently compressed into a level-2 **M** matrix using Eqs. (45)–(47). The final solution at level 3 (Darcy scale) is obtained by solving Eq. (37) over the whole domain, using only **M** matrices from level 2.

The above multi-level strategy has several advantages. At all levels higher than 0, the form of the coarse-scale equation (used to compute **M**) is not assumed (unlike upscaling) but obtained from imposing mass conservation on coarse-grid interfaces (Eqs. (34) and (42)). Furthermore, *at no point does Darcy's law enter into the calculations*. Pore-scale information are propagated directly from the Stokes equation, Eq. (49), up to whatever level they are desired. No assumptions about scale separation or periodicity, common to homogenization, are needed. The number (dots in Fig. 11) and order (linear in Fig. 11) of the mortar bases at each level (1 and 2 in Fig. 11) control the coarsening ratio and the *upscaling error* (PLMM has a different mechanism for controlling errors at level 0; Mehmani and Tchelepi, 2018). The coarsening ratio between levels k and k + 1 is defined as the size of **M** at level k divided by the size of **M** at level k + 1. The higher the number or order of the mortar bases, the lower the coarsening ratio and thus the upscaling error.

The above multilevel strategy (Fig. 11) is a new proposition, but one that is made possible only recently due to extensive progress in the development of multiscale methods at the Darcy scale (Arbogast, 2012; Efendiev and Hou, 2008) and the pore scale (Mehmani and Tchelepi, 2018, 2019). Significant advances in the 3D characterization of porous media, such as X-ray tomography (Wildenschild and Sheppard, 2013), and the development of high-powered parallel machines have served as crucial catalysts. Despite its promise, the approach has limitations. If the pore-scale Eq. (49) is replaced by the two-phase flow Eq. (38), then the calculations at level k not only depend on level k−1, as with single-phase flow, but also on all the levels between 0 and k−1. That said, the dependence on level 0 will be *local*, allowing the use of adaptivity criteria like the ones discussed in Section 3.4.2.2. If Eq. (49) is replaced with a single-phase solute transport equation, then the strength of molecular diffusion will determine the order of recurrence, i.e., whether level k depends on all, or only some, of the lower levels. If diffusion is strong, the concentration field is smooth and the chain of dependence is short (see also the discussion in Section 3.4.3). The tyranny of characterization poses another obstacle by limiting the extent to which a domain can be characterized at level 0. Future advances in instrumentation may open new doors, but for now, probabilistic descriptions are the best bet to extrapolate beyond what is measured.

3.5. Imaging

Images are among the most important data type for probing geologic porous media, because they provide a *continuous* spatial map of how rock properties (e.g., composition, porosity) or thermodynamic variables (e.g., concentration, pressure) are distributed. At the pore scale, microcopy images (μ-CT, FIB-SEM) capture the intricate and complex geometry of the void space and the mineralogical makeup of the solid phase (e.g., EDX). These images are often crucial inputs for many pore- and reservoir-scale simulations. Pore-scale models in particular (e.g., PNM, PLMM) rely heavily on the availability of 3D voxelized images of a few microns to millimeters in size to perform calculations, so much so that such image-based simulations are referred to as *digital* rock physics. But images are useful for two other reasons: (1) scale and data translation and (2) multiscale and multimodal characterization. These are the topics of the following sections.

In Section 3.5.1, we provide an example of how high-resolution images of a meter-scale core, from a shale formation, can be used to: (1) downscale coarse-scale information, like well-logs, to obtain fine-scale information about the rock, like composition; (2) perform data translation to convert one fine-scale data type into another, like composition to thermal conductivity; and (3) how to upscale the translated data back to the coarse scale, where they may be used for reservoir simulation. Note that these are the exact same steps that we outlined in Section 2.5 in order to *define* downscaling and data translation. Here, we apply these concepts for a challenging geomaterial.

In Section 3.5.2, we show how multiple imaging instruments, each probing a centimeter sized sample at a different length scale, can be combined to arrive at a near-complete characterization of the specimen. Multiscale imaging and multimodal imaging refer, respectively, to such combinations of resolution and instrumentation. They are motivated by the fact that no single instrument is capable of resolving the full range of length scales present in challenging geomaterials, like shales. The reason is an inherent tradeoff between resolution and field of view imposed by all imaging instruments. We present a workflow that alleviates this shortcoming.

3.5.1. Image-based scale and data translation

Consider the 100 ft. long core shown in Fig. 12, which is extracted from the Green River Formation (GRF) in the U.S. (Mehmani et al., 2016b). The core is an immature source rock, or shale, high in organic content. The diameter of the core slab in Fig. 12 is ~8 cm. Suppose we are given a well-log of the rock's density, which has a resolution of 0.5–2 ft. Here, we regard the log as coarse-scale data. Since it is possible to deduce organic content from density in the GRF (Mehmani et al., 2016b), we convert the density log into a well-log of average organic concentration, $<c_o>$. The bottom left graph of Fig. 12 shows how $<c_o>$ varies with depth. To produce hydrocarbons from oil shale, one must apply heat. Thermal conductivity is therefore an important petrophysical property of GRF rock. We now pose the following question: is it possible to translate the coarse-scale concentration, $<c_o>$, that is available to us into coarse-scale data on thermal conductivity, $<k>$, that would be useful for engineering? The answer is *yes* and we outline the steps below.

In the GRF, nearly all petrophysical properties, including thermal conductivity, are strong functions of organic content. But if we were to plot experimental measurements of $<k>$ versus the given $<c_o>$, we

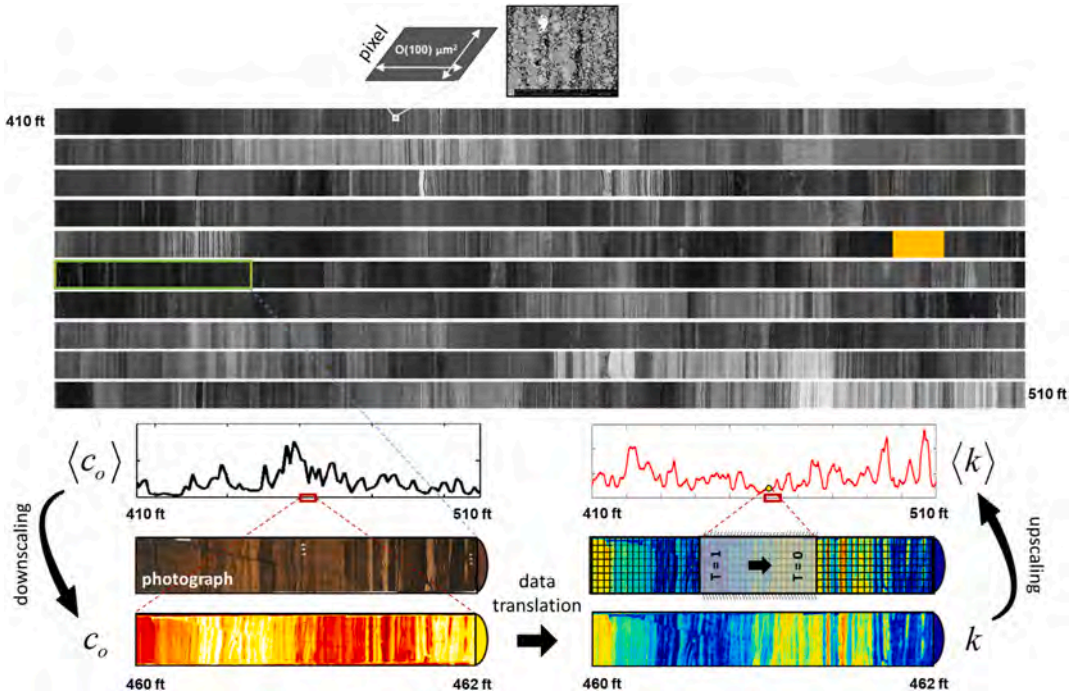

Fig. 12. Schematic of an image-based data translation workflow. Top: A 100 ft. long core from the Green River Formation (GRF), U.S., imaged with an optical camera (Mehmani et al., 2016b). Pixel sizes are $O(100\mu m^2)$, which corresponds to the field of view of the SEM image shown, where minerals (gray) and organic matter (black) are distinguishable. Bottom: coarse-scale (well-log) measurements of organic concentration, $<c_o>$, are downscaled to obtain fine-scale organic concentration, c_o, using a hyperspectral near-infrared scanner (hotter colors mean higher c_o). The photograph and fine-scale map of c_o correspond to the green highlighted segment of the core. The fine-scale c_o is then translated to a fine-scale map of thermal conductivity, k (data translation). There is a one-to-one relation between c_o and k in the GRF (Mehmani et al., 2016a) (not shown). The fine-scale thermal conductivity, k, is subsequently upscaled (using local upscaling; see Section 3.3) to obtain the coarse-scale (or log-scale) thermal conductivity, $<k>$. The fine grids used to perform numerical upscaling correspond to the thin black lines. The above workflow allows deducing $<k>$ from $<c_o>$.

would observe significant scatter. This is because for the same $<c_o>$, there are many more ways for the fine-scale thermal conductivity, k, to be distributed in space. Each distinct distribution corresponds to a different value of $<k>$. To reduce this variability, and thus the scatter, we must first downscale $<c_o>$ to c_o. To do this, Mehmani et al. (2016b, 2017) used both optical photographs and hyperspectral images in the near-infrared to obtain raw images of the rock; like the one shown in Fig. 12. While these images have a very high resolution, $O(100\ \mu m^2)$, their gray-scale values are not equal to c_o. To obtain c_o, the raw images must be calibrated, or *constrained*, to the coarse-scale data $<c_o>$. Once this is done, c_o is recovered for every pixel in the image and thereby downscaling is complete (see bottom-left inset of Fig. 12).

The next step is to map c_o to k, which is illustrated by the horizontal arrow in Fig. 12. A plot of these two fine-scale variables exhibits much less scatter than $<c_o>$ versus $<k>$. This is because the former plot excludes the contribution to scatter from the fine-scale variability within the measurement support of the logging instrument (0.5–2 ft). Mehmani et al. (2016a) compiled millimeter-scale measurements of k versus c_o from the literature and proposed a universal relationship between them. Of course, this relationship is specific to the GRF and there is no reason to believe that a similar one would hold in other formations. But even in those cases, one could perform pore-scale simulations on voxelized images of millimeter-scale samples to numerically "derive" the relationship. Another option would be to downscale not just c_o but also the concentration of other minerals in the rock, which is possible with a hyperspectral scanner, and then link them to k. Machine learning algorithms, discussed in Section 3.6, are very attractive here because they can encode hidden connections between different data types useful for data translation.

The final, and easiest, step is to upscale k to $<k>$ (Mehmani et al., 2016a). It is the easiest step because any one of the numerical upscaling methods, described in Section 3.3, can be used to accomplish this task. In Fig. 12, we discretize the thermal conductivity field k into a number of Cartesian fine grids (thin black lines). We then divide the core into a number of coarse grids, such as the gray interval in Fig. 12, and perform local upscaling. The BCs in the schematic correspond to the pressure-no flow BCs discussed in Section 3.3. In other words, the top and bottom boundaries of the coarse grid are sealed (or insulated) and the left and right boundaries are subjected to a constant temperature gradient. The fine-scale governing equation for heat conduction is identical to Eq. (14), with "k" corresponding to thermal conductivity (not permeability) and "p" to temperature (not pressure). Once $<k>$ is calculated for all coarse grids, a "well-log" of thermal conductivity can be plotted versus depth (Fig. 12). We have thus translated $<c_o>$ to $<k>$.

The above workflow does not imply that such a strategy is always easy or straightforward to implement. Sometimes limitations in imaging instruments or challenges posed by the rock mineralogy (like transition metal-oxides for optical images of GRF) hinder effective downscaling (Mehmani et al., 2016b). Other times a simple relationship between fine-scale variables, like the one between k and c_o in the GRF, is not possible or easy to construct. But despite these challenges, the workflow highlights the importance of (hyperspectral, optical, and other) images for scale and data translation and provides a framework for using them to convert easy-to-acquire data into hard-to-acquire data.

3.5.2. *Multiscale imaging for characterization*

The trend in image-based characterization of nanoporous media has followed a trajectory of increasing resolution to visualize pores and other microstructural details (Loucks et al., 2012; Curtis et al., 2012; Ma et al., 2018; Goral et al., 2019; Frouté and Kovscek, 2020). Such studies have improved our understanding of the distribution of pore shapes, sizes, and fractions within both the organic and inorganic constituents of

shales and other porous media. Imaging instruments such as transmission X-ray microscopy and FIB-SEM, among others, can probe rock samples at different resolutions and produce 3D or quasi-3D images. But despite their flexibility, such instruments are limited to small fields of view and thus small sample sizes. Because the total rock volume imaged is small, the statistical significance of the data may be questionable. Hence, imaging is subject to the tyranny of characterization, and *multiscale imaging* attempts to mitigate that.

Multiscale imaging (Pini and Madonna, 2016) refers to the acquisition of 2D or 3D images of an object at multiple spatial resolutions. When different imaging instruments are employed to accomplish this task, the process is called *multimodal imaging* (Anderson et al., 2020a). Both multiscale and multimodal imaging are powerful tools that provide a wealth of data about a particular physical process (e.g., diffusion) or the microstructural heterogeneity of a geomaterial (e.g., cracks, vugs, nanoporosity, and macroporosity). If used judiciously, multiscale images provide a useful guide for how to sample experimental specimens from a rock at places where the physics are most interesting or relevant to the application at hand (e.g., cracks, kerogen-mineral interfaces). The pitfall one must avoid is to eliminate human bias. This is possible by introducing some level of automation into the sampling workflow. Multiscale and multimodal imaging also provide a unique opportunity to probe the interplay between small-scale features, like kerogen and nanopores, and large-scale features, like fractures and vugs. A single instrument or resolution is often incapable of such a feat.

"Top-down" imaging is a particular application of multiscale imaging, in which a cascade of nested images is acquired to capture successively finer-scale features of a sample. The approach is an efficient way to probe the interrelation between different microstructural features at multiple scales (Vega et al., 2015; Aljamaan et al., 2017). Fig. 13 illustrates the idea for a sample of Barnett shale. The length scales probed range from a few centimeters to a few nanometers, i.e., 7 orders of magnitude. The workflow begins from a core plug (2.5 cm diameter) that is imaged using a conventional X-ray computed tomography (CT) scanner. A typical CT scanner captures large-scale features (e.g., fractures) and has a resolution of ~200 × 200 × 1000 μm^3. The gray-scale values correspond to the density of the rock, from which porosity can be deduced. But to go beyond porosity, and to obtain useful data about the connectivity of the pore space, a penetrant (or contrast agent) like krypton, argon, or xenon is used. The penetrant diffuses into the pore space, if accessible, and attenuates X-rays more strongly than the surrounding matrix (Vega et al., 2014). This allows accessible flow paths of the rock to be visualized. Gaseous penetrants are common because they are easier to inject than liquids (due to their lower viscosity), are nonreactive, and do not cause clay swelling. To increase contrast, two images are subtracted from each other: one of the sample filled with gas and another of the sample emptied of gas. This fluid substitution technique can also be used to visualize temporal evolutions of flow and diffusive mass transfer into a nanoporous matrix (Alnoaimi and Kovscek, 2019).

In the top-down workflow of Fig. 13, CT images of the core plug are used as guides to select a cross section for finer-scale analysis. The selection is driven by features that are of most interest to the researcher. Here, fluid accessibility and conductivity of the pore space were of interest. After a cross section is selected, the plug is physically cut across that cross section, polished, and imaged with an SEM. Since SEM, like any other imaging instrument, has a limited field of view, many smaller images are taken to cover the full cross-sectional area of the core plug. This is done by following a rastering pattern across the sample's surface. The small images are then stitched together to produce a large mosaic over the entire cross section. The mosaic resolves features of *O*(1 cm) to *O*(100 nm), which is much more detailed than the CT image. But because the CT image contains useful information about pore connectivity, it is registered against the SEM mosaic. In Fig. 13, the registration is done with respect to high-gas-concentration regions colored in red. To probe the sample at an even finer scale, the mosaic is used as a guide to select subsamples for transmission X-ray microscopy (TXM or nano-CT). For a cylindrical specimen of 45 μm in diameter, TXM has a resolution of 30 × 30 × 30 nm. This is able to capture the spatial distribution of individual minerals and kerogen pockets within the shale fabric. Here, TXM reveals that microcracks seem to occur preferentially along kerogen-mineral interfaces, suggesting a possible pathway for gas flow. We may probe the sample at an even finer scale by using the TXM images now as a guide to select lamellar subsamples that are then imaged via

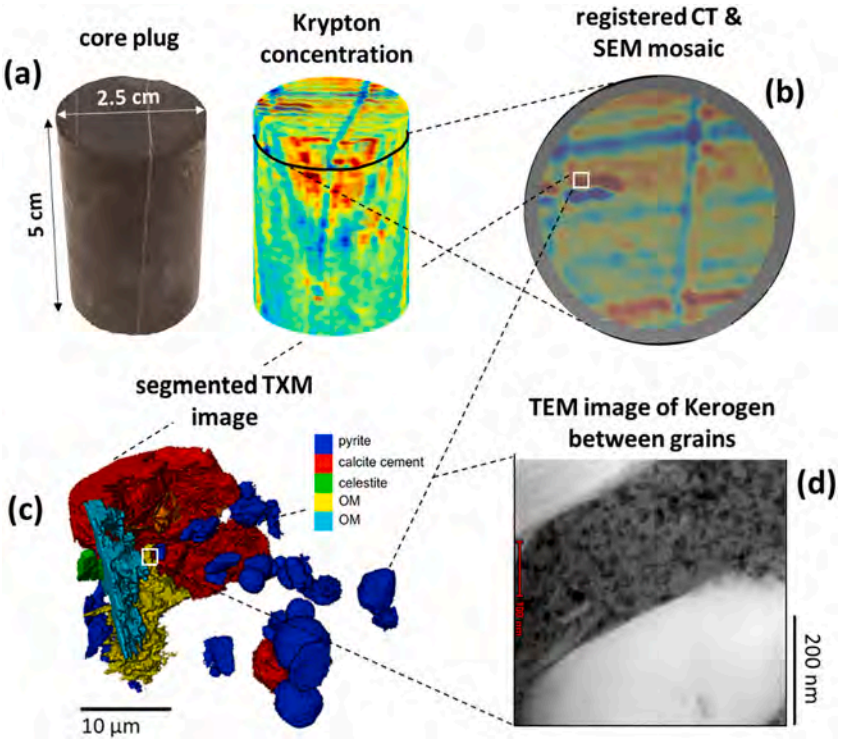

Fig. 13. Schematic of a top-down imaging workflow. Images cascade downward in length scale from *O*(cm) to *O*(nm). (a) Photograph of a Barnett Shale core and its CT image when flooded with an X-ray absorbing gas (Kr). The CT image reveals gas occupancy (hotter colors mean more gas). (b) A 2D cross section of the CT volume is extracted and registered with an SEM mosaic. (c) Subsamples are chosen from high-gas-concentration regions and used in TXM imaging. (c) Segmented TXM reveals mineral and organic constituents. (d) TXM is subsampled again and used in TEM imaging, which shows that the organic matter (kerogen) is sandwiched between two inorganic grains. Darker gray areas correspond to *O* (nm) pores. This image was composed from the works of Frouté and Kovscek (2020), Vega et al. (2015), and Aljamaan et al. (2017).

scanning transmission electron microscopy (TEM). The resolution of TEM is ~1 nm, which reveals nano-structural information such as the distribution and size of pores inside the kerogen.

The foregoing example highlights how multiple imaging instruments, operating at vastly different scales, are used in tandem to probe a geomaterial (e.g., shale) at ever finer detail without losing sight of how it all ties to the largest scale. None of the instruments by themselves is able to probe the sample in full, but together, they provide the flexibility to examine selectively those parts of the sample where the physics beckon for a deeper look. It is also worth noting that for every image acquired at a particular scale of observation, there exists a computational method best-suited for simulating the physics at that scale, and which uses the image as input. One example is the lattice Boltzmann method (LBM) presented in Section 4, designed to simulate gas flow through nanopores.

3.6. Machine learning

Machine learning forms the basis of analyzing and making predictions in data-rich environments and has become popular for characterizing shales. Data-driven scale and data translation differs fundamentally from many of the upscaling and downscaling methods discussed earlier that rely on the computational and mathematical structure of the underlying equations. Machine learning methods aim to assimilate and encode data across different length scales and acquisition modalities by training models that predict one data type from another or capture shared representations between multiple data domains. In characterizing shales, images are the most readily available data type. Accordingly, we focus our attention on advances in machine learning to process image-based data in service of scale and data translation. We note the following exposition neglects, for the sake of brevity, an important and emerging segment of the AI literature called "physics-informed machine learning", which is concerned with using AI to solve partial differential equations. Impressive results (Raissi, 2018; Raissi et al., 2019), mixed with occasional failures (Fuks and Tchelepi, 2020), in this space are enabling the simulation of dynamical systems in a matter of seconds while honoring desirable physical constraints (e.g., mass conservation). These methods certainly do have a place in scale translation, but our focus is directed towards characterization and visualization tasks that are pivotal to experimental workflows.

3.6.1. Upscaling with data-driven proxy models

Supervised machine learning methods are fundamentally function approximators whose forms depend on the properties of the underlying function being approximated (Hastie et al., 2009). Existing methods divide roughly into discriminative and generative models. Discriminative models aim to predict a response value or class label from input data, while generative models seek to encode the underlying probability distribution of the data so that new realizations can be reconstructed. In upscaling, discriminative models are more useful, as they can either predict coarse-grid properties directly from fine-grid inputs or embed data-driven proxy models into a broader simulation workflow (e.g., to accelerate computations). Previous studies on computing upscaled parameters from fine-scale images have used convolutional neural networks (CNNs) to predict diffusivity from synthetic images (Wu et al., 2019a), permeability and velocity distributions from CT images of sandstone (Kamrava et al., 2020; Santos et al., 2020), and shale production data from limited field-scale simulations (Klie and Florez, 2018). As for data-driven proxy models, existing work has embedded neural networks into macroscale simulators. For example, Bao et al. (2020) used a multilayer perceptron as a proxy model to capture device-scale effects of pore-scale flow in batteries, and Wang and Battiato (2021) embedded a recurrent neural network (RNN) into an macroscale framework that simulates flow through a multiscale fracture network. Despite these advances, works on shales remain limited and there is some indication that under certain conditions the unique microstructure of shales can cause existing methods to fail (Wu et al., 2019a). Such challenges warrant continued work with a particular emphasis on microstructurally complex rocks, including shales.

3.6.2. Downscaling through single image super-resolution

Downscaling is underdetermined (Section 2.5) and especially challenging when fine-scale information is scarce or absent (tyranny of characterization). Machine learning can be used to synthesize statistically plausible realizations of the fine scale wherever such information is missing. Generative adversarial networks (GANs) (Goodfellow et al., 2014; Ledig et al., 2017) and convolutional neural networks (CNNs) (Dong et al., 2016) play a particularly important role in this space. The best example of data-driven downscaling in shales is single image super-resolution (SISR), which seeks to construct high-resolution images from low-resolution inputs. Applied to porous rocks, SISR has been used to enhance the "sharpness" of low-resolution X-ray images as part of a Digital Rock Physics workflow, employing algorithms like CNNs (Da Wang et al., 2019), neighbor embedding (Wang et al., 2018), and template matching (Wu et al., 2019b).

While promising results have been achieved with synthetic data, very few studies have applied SISR (or similar downscaling techniques) to multimodal and multiscale experimental data. We believe a key development will be to apply SISR models in experimental settings to overcome challenges and limitations of existing imaging setups. An example is shown in Fig. 14, where paired CT and μCT images are acquired, aligned, and processed to train an SISR model. The model can then be used to convert low-resolution X-ray images of the shale rock to high-resolution virtual μCT maps. A trained SISR model could also be used to increase the spatial resolution of time-dependent experiments captured by a CT scanner, such as the reactive transport of a solute or the diffusion of a gas.

3.6.3. Data translation across imaging modalities

Data translation is another key area where machine learning can be useful for characterizing shales. Here, image-to-image translation (Efros and Freeman, 2001; Isola et al., 2017; Zhu et al., 2017) is of particular value, designed originally to transform images from one domain of acquisition to another, e.g., predicting daytime images from nighttime images. This opens up fertile grounds for so-called "multimodal imaging", where hard-to-acquire images from a potentially destructive instrument, like FIB-SEM, can be linked to easy-to-acquire images from a non-destructive instrument, like nano-CT. Once trained, the image translation model can be used to convert fast-to-obtain but potentially less useful data from the latter instrument into slower-to-obtain but more useful data from the former instrument. An example was presented by Anderson et al. (2020a), who assembled a dataset of paired nano-CT and FIB-SEM images of Vaca Muerta shale and trained a 2D image-to-image translation model (Isola et al., 2017). Future directions in multimodal imaging include generalizing this and similar image translation workflows to 3D-to-3D image prediction and developing algorithms that are trainable on both paired and unpaired images (Zhu et al., 2017). Another opportunity lies in the use of conditional GANs (CGAN) (Mirza and Osindero, 2014) to enhance prediction quality by constraining output images to non-image priors. Yet another is to reconstruct virtual (μCT) images directly from non-image inputs like NMR.

3.6.4. Reconstructing and learning the pore structure

A key part of scale and data translation is the ability to reconstruct synthetic images of porous rocks in such a way that the unique statistics of a geology of interest are honored. In downscaling, generated images can be used to either fill a resolution gap or to augment the training pool of a deep learning algorithm (Kamrava et al., 2019; Shams et al., 2020). In data translation and upscaling, generated images can be used to build structure-property relations that link morphological features of the microstructure to petrophysical properties (Adler et al., 1990). Most methods for reconstructing synthetic images fall under statistical (Caers,

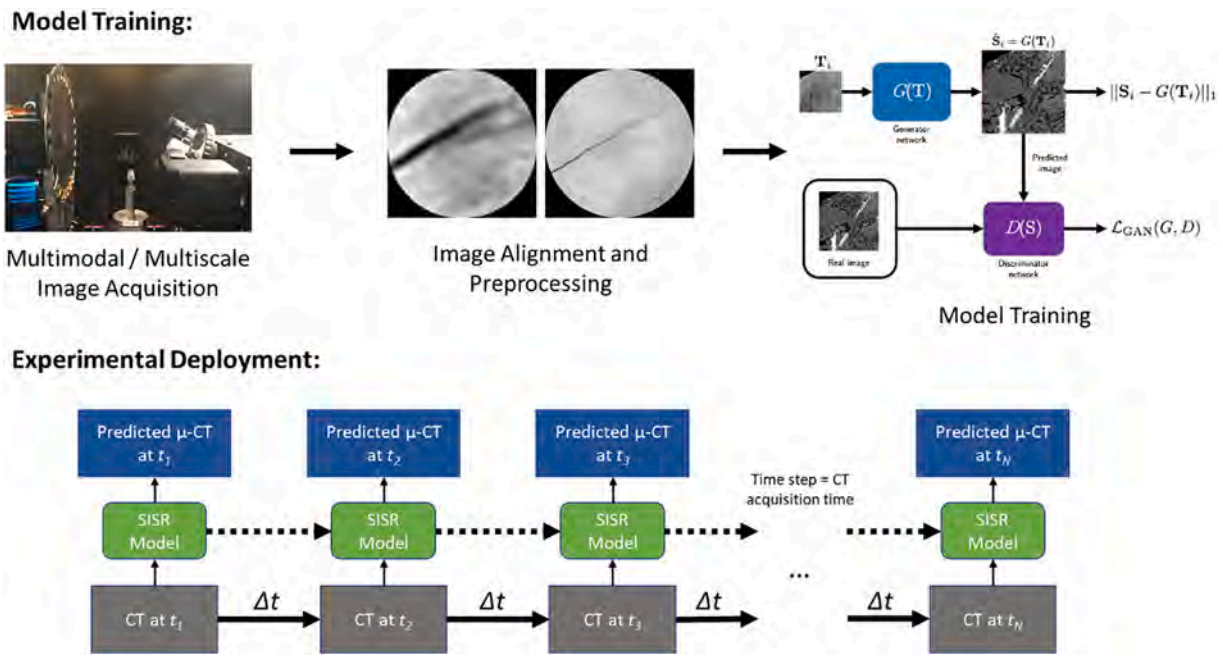

Fig. 14. Diagram of SISR in a shale imaging experimental workflow. Multiscale CT/micro-CT images are acquired and used to train a deep learning SISR model, such as SR-GAN from Ledig et al. (2017). During deployment, high-temporal but low-spatial resolution CT images are acquired and then super-resolved using the trained SISR.

2001; Manwart et al., 2000) or deep learning categories (Mosser et al., 2017). Statistical methods generate realizations by using a collection of statistical measures (e.g., moments, chord functions, variograms) gathered from one or more training images (Roberts, 1997). The realizations are made to have the same measures as the training images. Statistical methods require very little data (Okabe and Blunt, 2004, 2007; Bai and Tahmasebi, 2020) and 3D images can be easily reconstructed from 2D inputs (Tahmasebi et al., 2012) (with certain assumptions about statistical isotropy). The tradeoff is that they are slow in generating large volumes and most are designed to handle only binary images. Shale images are non-binary and thus require methods that can generate grayscale images resolving kerogen, minerals, and the void space. Deep learning methods have shown a lot of promise (Mosser et al., 2017, 2018) but have required 3D inputs to train. This has precluded generating 3D volumes based on 2D training data like SEM. To overcome the challenge, Anderson et al. (2020b) developed a method based on generative flow models (Dinh et al., 2014, 2016; Kingma and Dhariwal, 2018) that synthesizes grayscale rock volumes from 2D inputs. The synthesis is comparable in quality to 3D GANs. When tested on a Bentheimer sandstone, the reconstructions matched the permeability and Minkowski functionals of the training data (Guan et al., 2021). The approach was further generalized to reconstruct multimodal (TXM and SEM) images of shales (Anderson et al., 2020b), making it applicable to other image types.

3.6.5. *Future directions for data-driven scale translation in shales*

Beyond further advances in machine learning methods to processes different combinations of data, scales, and modalities, we see three areas for future research that will enable effective scale and data translation in shales. First is *data curation*: while machine learning is very powerful in assimilating heterogeneous data types, the most difficult step is acquiring and preprocessing paired data across input and output domains. Since predictive models require a large number of paired data points, future work on scale and data translation should accompany developments in single-scale and single-modality algorithms that enable robust data processing, like segmentation and image registration. Such efforts would not only increase the reliability of machine learning algorithms for scale translation, but also accelerate their training and broader experimental deployment in geosciences.

Second is *transfer learning*: in some application domains (e.g., super-resolution) data are much easier to acquire than others (e.g., flow simulations or experiments). Transfer learning refers to repurposing a machine learning model trained in one domain to be deployed in another, where data are scarcer. Images provide a perfect example. The feature maps buried inside a convolutional neural network carry relevant multiscale and multimodal information. Image-to-image deep learning architectures should therefore not be viewed as mere unidirectional operators for upscaling, downscaling, or data translation, but as *links* between different data types, scales, and modalities, on which they are trained. Fig. 15 illustrates an example where an image-to-image super-resolution architecture is repurposed for other tasks performed in scale translation, e.g., data-driven proxy modeling embedded within a simulator. Transfer learning could, therefore, be useful in accelerating progress towards scale and data translation of data-scarce problems.

The third and final area is *autodifferentiation* (AD): a critical step in integrating proxy models into physics-based simulators is the availability of robust AD algorithms. AD allows exact derivatives (or Jacobians) of a function to be calculated, using a computational graph, without the need for finite differences (Griewank and Walther, 2008). Training deep learning models often relies on stochastic gradient descent optimizers (or its variants) (Kingma and Ba, 2014), where the gradient of an objective function must be repeatedly computed. Physics-based simulators equipped with AD could efficiently train proxy models that are embedded into their workflow. Currently, only a limited number of physics-based simulators are equipped with AD (e.g., MRST (Lie, 2019) and AD-GPRS (Younis, 2011)). Therefore, future work should focus on integrating AD capabilities into simulator source codes in such a way that proxy-models can be seamlessly embedded into their framework to accelerate computationally expensive tasks.

4. Case studies

The Center for Mechanistic Control of Unconventional Formations (CMC-UF) is an Energy Frontier Research Center, funded by the U.S.

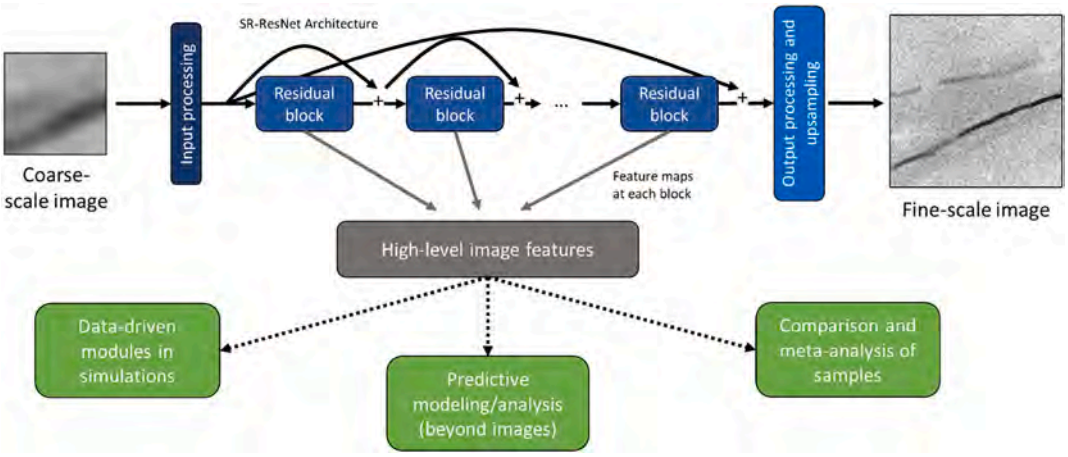

Fig. 15. Workflow showing a super-resolution ResNet architecture (from Ledig et al., 2017) trained to predict fine-scale micro-CT images from coarse-scale input images. The feature maps learned in this architecture could be harnessed for other data-driven scale translation tasks in which data are scarce.

Department of Energy, whose mission is to understand the fundamental interactions between fluids and unconventional source rocks, such as shales. As members of this center, we grapple with challenging questions related to scale and data translation in shales, and part of our activities rely on developing disparate representations of the same physics that correspond to vastly different spatiotemporal scales. They include molecular dynamics (nm), density functional theory (nm-μm), pore-scale modeling (μm-mm), and reservoir simulation (m-km). In accordance with our efforts, we find it instructive to present here two case studies, which at the surface might seem to have little to do with the methods presented in earlier sections. But as we later point out, they are both instances of numerical upscaling.

The first case study examines gas flow through nano-sized pores. It discusses the development of a lattice Boltzmann method (LBM) for capturing the diverse transport mechanisms at play. But to capture them correctly, the model must be parameterized, preferably by molecular dynamics simulations. Put differently, finer-scale simulations are required to derive coarse-scale parameters. The model may also be calibrated to other sources of data such as experiments. But regardless of where the calibration data come from, the key point is that the case study involves two steps: (1) postulating an educated guess for the form of the governing equations of nano-scale gas flow, and (2) using fine-scale simulations or other data types to parameterize these equations. Of course, the postulated model in step (1) must be a good one, otherwise the number and sensitivity of the parameters to be calibrated in step (2) increases commensurately. In Section 4.1, we also review recent literature on gas transport in nanoporous media.

The second case study examines dual continuum models, which are yet another instance of numerical upscaling. Dual continuum models resolve the flow in fractures with a different equation than the flow in the matrix. In shales, this is relevant because the formation is either hydraulically fractured or contains natural fractures. The matrix and the fracture communicate, or exchange mass across their shared interface, captured by a transfer function. The coefficients in the transfer function need calibration. Following similar arguments as presented in Sections 3.1–3, one may attempt to derive a homogenized equation for a dual continuum medium from a set of fine-scale equations (Stokes in fractures and Darcy in the matrix). Such a model would inevitably require closure assumptions, one of which is scale separation. Dual continuum models can thus be viewed as either homogenized equations that hold under certain restrictive assumptions, or as postulated equations that need to be parameterized by field-scale data or fine-scale simulations. We provide a brief overview of dual continuum equations below.

4.1. Pore-scale simulation of gas transport in nanopores

Recent characterization of shale samples (Frouté and Kovscek, 2020) using FIB-SEM and STEM techniques have highlighted their complex nature. Shale and tight formations comprise a multitude of nanoscale pores ranging in size from 1 nm to 200 nm. While gas flow in larger submicrometer pores may be reasonably modeled with continuum models, smaller pores present difficulties both in terms of transport and phase behavior (Wang and Aryana, 2021). The flow regime within pores 1–10 nm large is either slip or transitional, where continuum-based descriptions (e.g., the Stokes equation) likely fail. To capture the physics correctly, scale-appropriate modeling paradigms such as Molecular Dynamics (MD) simulation or the Lattice Boltzmann method (LBM) (He and Luo, 1997; Chen and Doolen, 1998) are needed. In MD, the dynamic evolution of a collection of molecules is captured by modeling the precise interactions between them. In the absence of physical observations, MD can provide reliable data for understanding the fundamental physics governing flow and transport (Skoulidas et al., 2002; Firouzi and Wilcox, 2012).

Despite its power, MD is computationally prohibitive at larger spatiotemporal scales. An alternative, is to use a larger-scale simulation method like LBM, which serves as a useful intermediary tool for bridging the gap between the molecular and sub-micron scales. Data obtained from MD simulations (or carefully designed experiments) can be used to constrain the parameters of LBM (i.e., numerical upscaling). The governing equation solved by LBM is the discrete version of the Boltzmann equation given by

$$f_i(\boldsymbol{x}+c\boldsymbol{e}_i\delta t, t+\delta t) - f_i(\boldsymbol{x},t) = \Omega_i(f(\boldsymbol{x},t)) + \delta t F_i(\boldsymbol{x},t) \tag{50}$$

In Eq. (50), the subscript i denotes the index of the discrete velocity; f_i the discrete distribution function along the i-direction; c the lattice speed defined as the ratio of the lattice spacing over the time step (set to 1); Ω_i the collision operator responsible for changes in f_i due to molecular collisions; and F_i the forcing term associated with the body force $\mathbf{G}$. Continuum variables, like density ρ and velocity $\boldsymbol{u}$, are obtained from

$$\rho = \sum_i f_i \qquad \boldsymbol{u} = \frac{1}{\rho}\left(\sum_i \boldsymbol{e}_i f_i + \frac{\delta t}{2}\boldsymbol{G}\right) \tag{51}$$

For flow at large Knudsen numbers (Kn), defined as the ratio of the molecular mean free path over the characteristic length of the flow field, the interactions between molecules and the pore walls dominate. In this case, the transport and phase behavior of the fluid deviate from those at bulk conditions. Fig. 16 shows that three mechanisms are responsible for gas flow through confined pores: viscous flow, surface diffusion, and Knudsen diffusion (Do et al., 2001). Viscous flow is due to interactions

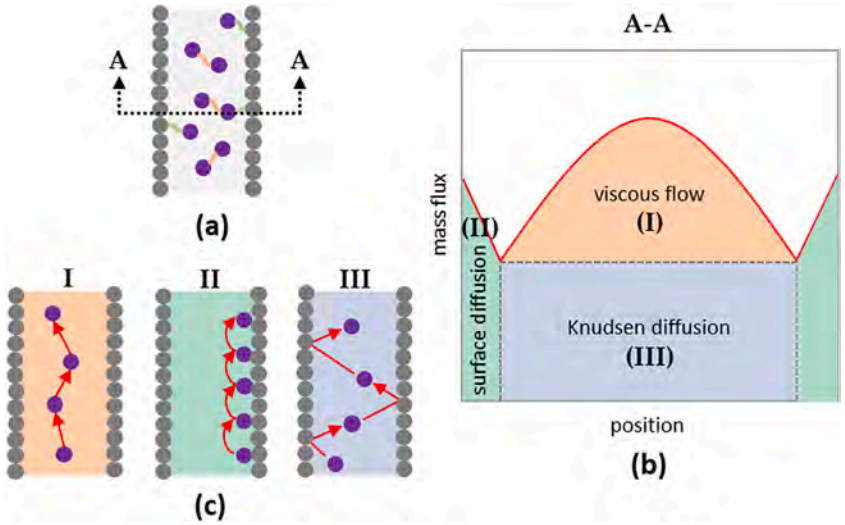

Fig. 16. A schematic of transport mechanisms of gas inside nano-sized slits. (a) Interaction of gas and wall molecules; (b) a cross-sectional mass flux profile. The mass flux profile is partitioned into three sub-regions, representing contributions from three transport mechanisms: viscous flow, surface diffusion, and Knudsen diffusion; and (c) a graphical illustration of each transport mechanism.

between gas molecules and results in a parabolic velocity profile (see Fig. 16b). Surface diffusion accounts for adsorbed gas molecules hopping along the pore wall (see Fig. 16c), which enhances the total flux of the gas. Knudsen diffusion is the additional flux imparted by the collisions between gas molecules and pore walls (see Fig. 16c). This results in a nonzero (slip) velocity near the solid surface. Fig. 16b illustrates the contribution from each mechanism to the overall flux profile. LBM enables accurate characterization and quantification of these transport mechanisms through appropriate boundary conditions and forcing terms near the walls (Eqs. (52)–(53)).

Three types of boundary conditions are commonly employed in LBM simulations: bounce-back (BB), specular reflection (SR), and Maxwellian diffusive reflection (MDR). In BB, molecules colliding with a wall are reflected back in the opposite direction (Nie et al., 2002), emulating no-slip boundary conditions. SR reverses only the velocity component perpendicular (not parallel) to the wall (Lim et al., 2002). Since the implementation of SR depends on the wall geometry and the direction of a molecule's motion, its application is nontrivial for complex geometries. In MDR, molecules that collide with walls lose memory and are scattered back following a Maxwellian distribution (Ansumali and Karlin, 2002). Slip models are often written in terms of continuum velocities, and the second-order slip model below is believed to be capable of describing gas flow at high *Kn* (Zhang et al., 2012; Liehui et al., 2019)

$$u_s - u_w = C_1 Kn \left.\frac{\partial u}{\partial n}\right|_w - C_2 Kn^2 \left.\frac{\partial^2 u}{\partial n^2}\right|_w \tag{52}$$

where C_1 and C_2 are slip coefficients and n is the wall normal coordinate. The subscript w specifies that a variable is defined at the pore wall. To mimic the slip boundary condition given by Eq. (52) in LBM, a combination of LBM boundary conditions are suggested. Prominent examples include BB-SR (Guo et al., 2008), SR-MDR (Tang et al., 2005), and BB-MDR (Chai et al., 2010). The BB-MDR is particularly attractive for simulating flow in complex geometries because of its convenient implementation (Wang and Aryana, 2020).

The interaction between fluid molecules, and between fluid and solid molecules, is incorporated in the form of external forces that act on the density distribution function. A long-range interaction force between the f_i is given by (Shan and Chen, 1993)

$$\boldsymbol{F}_{\mathrm{ff}}(\boldsymbol{x},t) = -c_0 G_{\mathrm{ff}} \psi(\boldsymbol{x},t) \sum_i w_i \psi(\boldsymbol{x}+\boldsymbol{e}_i \delta t, t) \boldsymbol{e}_i \tag{53}$$

where G_{ff} is intermolecular strength, ψ is the pseudopotential function related to the fluid density ρ, w_i is the weighting factor, and c_0 is a constant. Both w_i and c_0 depend on the lattice structure. The interaction between particles and walls is given by a similar form (Sukop and Or, 2004)

$$\boldsymbol{F}_{\mathrm{fw}}(\boldsymbol{x},t) = -G_{\mathrm{fw}} \psi(\boldsymbol{x},t) \sum_i w_i s(\boldsymbol{x}+\boldsymbol{e}_i \delta t, t) \boldsymbol{e}_i \tag{54}$$

where G_{fw} is the adsorptive strength, and s is a switch function equal to 1 for the solid wall and 0 elsewhere. We observe that both $\boldsymbol{F}_{\mathrm{ff}}$ and $\boldsymbol{F}_{\mathrm{fw}}$ rely on ψ. Determination of ψ requires an appropriate equation of state (EOS) (Yuan and Schaefer, 2006). Examples include the van der Waals (vdW)-EOS (Van Der Waals, 1873) and the Peng-Robinson (PR)-EOS (Peng and Robinson, 1976). These two EOSs may not be valid in confined spaces because the thermodynamic behavior of confined fluids deviate from their bulk phase (Sobecki et al., 2019). Yang et al. (2019) proposed an extension of PR-EOS, referred to as EPR-EOS, that incorporates the shift of critical temperature in the model. Wang and Aryana (2021) proposed an additional modification, referred to as mEPR-EOS, in which shifts in critical pressure and temperature are evaluated independently. Numerical solutions of mEPR-EOS are validated against the density phase diagram of methane constructed by MD simulations under different confinement scenarios.

Recent studies have combined LBM and MD to translate observations from molecular scale to mesoscopic or macroscopic scales (Liu et al., 2021). As summarized by Phan et al. (2020), the key to the success of such scale translation is to achieve agreement between LBM and MD in the same domain. This is very similar to the numerical upscaling workflow discussed in Section 3.3. For example, Zhao et al. (2016) determined the adsorption parameters of LBM by matching the adsorption curve from MD. Yu et al. (2017) studied the transport of methane in nano-sized slits, where the flux profiles of LBM and MD were made to agree. Wang and Aryana (2021) revisited this problem by incorporating mEPR-EOS into LBM. Values of adsorptive strength were determined such that the flux profiles agreed with MD (Sobecki et al., 2019).

4.2. *Dual continuum models of fracture-matrix interaction*

The dual continuum approach conceptualizes a porous medium as two superimposed domains: a matrix volume and a fracture volume. The latter does not resolve fractures explicitly. The matrix volume accounts

primarily for fluid storage and the fracture volume for fluid transport (Barenblatt et al., 1960; Warren and Root, 1963). The matrix is often assumed to be much less permeable than the fracture, and the mass exchange between the two is described by transfer functions that can be determined analytically, numerically, or using experimental data. Dual continuum models divide into the categories of dual permeability, dual porosity, and multiple interacting continua (MINC) models (Pruess, 1992). Dual permeability systems account for inter-matrix, inter--fracture, and matrix-fracture mass transfers. Dual porosity is similar except it ignores inter-matrix mass transfer. MINC models resemble dual porosity but may consist of a matrix that is discretized into multiple grid blocks. In both dual porosity and MINC, the matrix communicates only locally (over short distances) with the fractures. The matrix-to-fracture transfer functions incorporate the characteristic lengths and areas of the fractures and the matrix.

The dual continuum approach applied to multiphase systems differs somewhat from that applied to single-phase gas flow in tight media (e.g., shales) with adsorption. Here, we follow Alnoaimi and Kovscek (2019) who studied the latter problem with a MINC model. The computational domain is first divided into two types of overlapping grid blocks: matrix blocks and fracture blocks. The flow within fracture blocks (f) is described, in discretized form, by

$$\Delta\left[T_f \Delta p_f\right] = \frac{V_f}{\Delta t}\Delta_t\left[\frac{\varphi_f}{B_{gf}} + \rho_r\left(1-\varphi_f\right)V_{gf}\right] - \tau_f \tag{55}$$

where T_f is the transmissibility between fracture blocks, Δp_f the change in fluid pressure in the fracture block, ρ_r the rock density, V_f the fracture block volume, φ_f the fracture block porosity, τ_f the matrix-fracture transfer function, V_{gf} the amount of gas adsorbed at a given pressure in the fractures, B_{gf} the gas formation volume factor in the fracture block, and Δ_t the time step.

The flow within a matrix block (m) adjacent to a fracture block is described by

$$\Delta\left[T_m \Delta\left(\frac{p}{Z}\right)_m\right] = \frac{V_m}{\Delta t}\Delta_t\left[\frac{\varphi_m}{B_{gm}} + \rho_r(1-\varphi_m)V_{gm}\right] + \tau_f \tag{56}$$

where T_m is the transmissibility between matrix blocks, Z the real gas compressibility factor, V_m the matrix block volume, φ_m the matrix block porosity, V_{gm} the amount of gas adsorbed on the matrix, and B_{gm} the gas formation volume factor in the matrix block. If there are more than one matrix blocks connected to each fracture block, then the flow within matrix blocks *not* adjacent to the fracture is described by

$$\Delta\left[T_m \Delta\left(\frac{p}{Z}\right)_m\right] = \frac{V_m}{\Delta t}\Delta_t\left[\frac{\varphi_m}{B_{gm}} + \rho_r(1-\varphi_m)V_{gm}\right] \tag{57}$$

The interaction between the matrix and fracture systems in Eqs. (55)–(56) is captured by a source term, or transfer function, that assumes various forms. Transfer functions describe the rate at which the matrix supplies mass to the fractures, or vice versa. For flows with a diffusive driving force, we have

$$\tau_f = \frac{DA_m\varphi_m}{\gamma L_m}\frac{M_g}{RT\rho_{sc}}\left[\left(\frac{p}{Z}\right)_f - \left(\frac{p}{Z}\right)_m\right] \tag{58}$$

where A_m is the interfacial area between matrix and fracture blocks, D the (Knudsen) diffusivity, L_m the characteristic length of the matrix, γ the tortuosity, M_g the molecular mass, R the universal gas constant, ρ_{sc} the density at standard conditions, and T temperature. The transmissibilities T_f and T_m are expressed as follows

$$T_f = \frac{kA}{B_{gf}\mu_g L_f} \tag{59}$$

$$T_m = \frac{DA_m\varphi_m}{\gamma L_m}\frac{M_g}{RT\rho_{sc}} \tag{60}$$

where k is the (apparent) permeability, μ_g the gas viscosity, A_f the interfacial area between fracture blocks, and L_f the fracture block length. Both B_{gf} and B_{gm} are computed using

$$B_g = \frac{\rho_{sc}ZT}{T_{sc}p} \tag{61}$$

where the subscript sc denotes standard conditions. The pressure p in Eq. (61) is equal to the matrix block pressure for B_{gm} and the average pressure between adjacent fracture blocks for B_{gf}.

One strength of the above dual continuum model is the ease with which different physical mechanisms like convection, diffusion, and gas sorption can be incorporated. Another strength is that the model parameters can be determined from first principles through either experiments that measure τ_f by mass balance, images that capture changes in fluid mass within the matrix/fracture, or fine-grid simulations that can serve as calibration data. The division of the heterogeneity and physical processes into fracture and matrix sub-systems provides significant flexibility for experimental characterization. For example, the porosity of the matrix can be measured independently from that of the fracture, and the convective flow within the fracture can be probed separately from the diffusive transport in the matrix.

5. Summary

This review was motivated by the need to describe physicochemical processes in geologically challenging porous media, such as shales, that display large disparity in length and time scales. We showed that such geomaterials generally do not exhibit scale separation and therefore cannot be described with a closed set of equations at only one scale. Instead, the processes observed at large scales are intimately coupled to those occurring at small scales. Homogenization, consisting of a set of mathematical tools to derive analytically closed-form macroscopic equations for scale-separable media, become inadequate. Nonetheless, they provide an important theoretical foundation upon which numerical and experimental methods, required to solve such problems, are based. These methods consist of hybrid computing, numerical upscaling, multiscale computing, high-resolution imaging, and machine learning. Collectively, they enable scale translation, which we defined as using data at one spatiotemporal scale to infer needed information at another, even in the absence of scale separation. After establishing broad definitions for upscaling, downscaling, tyrannies of prediction, and data translation, we provided a pedagogical review of each method and drew instructive comparisons between them. For example, we said that hybrid methods simultaneously solve a homogenized coarse-scale equation alongside a fine-scale equation to establish a two-way coupling between the coarse and fine scales. At the surface, this looks very different from multiscale computing, that performs a series of decoupled local calculations that are then used to formulate and solve a global coarse-scale problem. But upon closer examination, the only difference was that hybrid methods use an analytically homogenized equation to simulate coarse-scale physics, whereas multiscale methods use “numerically homogenized” variables to accomplish the same task. We also noted that while multiscale methods are often thought to apply only when the tyranny of characterization is absent (i.e., the domain can be fully characterized at the fine scale), this need not be the case. The basis functions computed on a few coarse grids may be *assumed* to be identical to those of other coarse grids, for which a fine-scale characterization is unavailable. While not common, there is nothing in the way of using multiscale methods in this fashion. The practice is much like homogenization, in which a closure problem is solved on *one* coarse grid and then assumed to hold across the whole domain. We drew similarly useful comparisons among the methods.

A large portion of the paper was devoted to multiscale methods, because they possess the rare ability to both downscale their coarse-scale solution, and estimate and control their prediction errors; unlike

most hybrid and upscaling methods. We showed that despite some differences in formulation (e.g., interface balance in MoMsFE versus volume balance in MsFV), *all* multiscale methods consist of three steps: (1) constructing a local basis, (2) solving a coarse problem, and (3) reconstructing the fine-scale solution. We noted that step 1 is equivalent to upscaling and step 3 to downscaling. If we eliminate step 3, we have a classical numerical upscaling method regardless of which multiscale method we use. The coarse-scale parameters obtained from each method may come in different forms (stiffness in MsFE, transmissibility in MsFV, and flux matrix in MoMsFE) but they all contain upscaled information about each coarse grid. We also discussed multiscale methods for pore-scale problems including PLMM and recent extensions of MsFV. We highlighted the algorithmic reason for why PLMM is more accurate. We also discussed PNM, one of the earliest pore-scale models, and provided a formal interpretation of it as a numerical upscaling method. Comparing PLMM and PNM, we highlighted the reasons for why the latter incurs larger errors. At the end, we proposed a new framework for algorithmically bridging between the pore and Darcy scales without ever invoking either Darcy's law or scale separation. It consists of performing hierarchically nested computations to propagate pore-scale information upwards, leaving the option open for downscaling and error control. The limitations of the framework were then highlighted reflecting a recurrent theme of the paper: the extent to which a set of fine-scale equations can be reliably upscaled depends strongly on the underlying character of the governing equations (e.g., hyperbolic versus parabolic).

We finally reviewed recent advances in high-resolution imaging and discussed how they are changing the way shales, and other geomaterials, are being characterized. Images, we argued, are a crucial data type that can be used alongside machine learning methods in order to construct useful mappings for the purposes of either downscaling or data translation. They provide much needed inputs upon which the computational methods of earlier sections rely. We concluded the paper by considering two case studies. The first was about modeling the controlling mechanisms of gas flow through nanopores. We argued that describing the physics with a lattice Boltzmann method (LBM) is akin to performing numerical upscaling, where a postulated coarse-scale equation (nano-scale LBM) is parameterized through fine-scale simulations (molecular dynamics). The second was a dual continuum model, for which a similar interpretation was deduced.

We hope that this review provides a broader perspective on scale translation and engenders a deeper appreciation for how new techniques fit alongside, and complement, old ones. It is meant to serve as a useful guide to geoscientists who have a vested interest in understanding and predicting fluid flow, solute transport, reactions, and mechanical deformation across a wide range of spatiotemporal scales in both shales and other geologically challenging porous media.

Declaration of Competing Interest

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

Acknowledgements

This work was supported as part of the Center for Mechanistic Control of Unconventional Formations (CMC-UF), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science under DOE (BES) Award DE-SC0019165.

Appendix A. Symmetry of Darcy-scale flux matrix

Consider Ω in Fig. 9c. If we discretize Eq. (44) on the fine grid using the finite volume method and a two-point flux approximation (TPFA) stencil, we arrive at the following system

$$\mathbf{A}\mathrm{p} = \mathrm{b} \tag{A1}$$

p is the vector of fine-grid pressures and b is the right-hand-side vector. It is well-known that **A** is a symmetric matrix when Eq. (44) is discretized with TPFA. Since $\partial\Omega$ consists of three open boundaries, b has the following structure

$$\mathrm{b} = \begin{bmatrix} O \\ \mathrm{T}_1 \times p_1 \\ \mathrm{T}_2 \times p_2 \\ \mathrm{T}_3 \times p_3 \end{bmatrix} \tag{A2}$$

where the entries are arranged in the following order from top to bottom: interior fine grids, fine grids that share an edge/face with Γ_1, Γ_2, and then Γ_3. O denotes a vector of all zeros for the interior grids of Ω. p_i is the scalar pressure imposed at Γ_i. T_i is the column vector of fine-grid transmissibilities corresponding to Γ_i. $\mathrm{T}_i \times p_i$ denotes multiplication of a scalar by a vector. To prove that **M** is symmetric, we have to show $q_i^j = q_j^i$. We choose $i = 1$ and $j = 2$, as others follow identically. The first pressure basis corresponding to $p = 1$ at Γ_1, in Fig. 9c, is obtained by setting $p_1 = 1$, $p_2 = 0$, and $p_3 = 0$ in Eq. (A2) and solving Eq. (A1). We thus have

$$\mathrm{p}^1 = \mathbf{A}^{-1} \begin{bmatrix} O \\ \mathrm{T}_1 \\ O \\ O \end{bmatrix} \tag{A3}$$

p^1 is the first fine-scale pressure basis over Ω. To compute its associated flowrate at Γ_2, $q_2{}^1$, we can simply left-multiply p^1 by the row vector shown below, where the superscript "T" denotes transposition

$$q_2^1 = \begin{bmatrix} O & O & \mathrm{T}_2^{\mathrm{T}} & O \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} O \\ \mathrm{T}_1 \\ O \\ O \end{bmatrix} \tag{A4}$$

Eq. (A4) is equivalent to computing the flux at Γ_2, by taking the gradient of p^1 and multiplying the result by λ, and then integrating the flux over Γ_2. Notice that since $\mathbf{A}$ is symmetric, so is $\mathbf{A}^{-1}$. Therefore, taking the transpose of both sides of Eq. (A4) leads to the following equality, which completes the proof. Recall $q_2{}^1$ and $q_1{}^2$ are scalars.

$$q_2^1 = \left(\begin{bmatrix} O & O & \mathrm{T}_2^{\mathrm{T}} & O \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} O \\ \mathrm{T}_1 \\ O \\ O \end{bmatrix} \right)^{\mathrm{T}} = \begin{bmatrix} O & \mathrm{T}_1^{\mathrm{T}} & O & O \end{bmatrix} \mathbf{A}^{-1} \begin{bmatrix} O \\ O \\ \mathrm{T}_2 \\ O \end{bmatrix} = q_1^2 \tag{A5}$$

References

Aarnes, J.E., 2004. On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation. Multiscale Model. Simul. 2, 421–439.

Aarnes, J.E., Kippe, V., Lie, K.A., 2005. Mixed multiscale finite elements and streamline methods for reservoir simulation of large geomodels. Adv. Water Resour. 28, 257–271.

Achanta, S., Cushman, J.H., Okos, M.R., 1994. On multicomponent, multiphase thermomechanics with interfaces. Int. J. Eng. Sci. 32, 1717–1738.

Acharya, R.C., Valocchi, A.J., Werth, C.J., Willingham, T.W., 2007. Pore-scale simulation of dispersion and reaction along a transverse mixing zone in two-dimensional porous media. Water Resour. Res. 43.

Adler, P.M., Jacquin, C.G., Quiblier, J.A., 1990. Flow in simulated porous media. Int. J. Multiphase Flow 16, 691–712.

Aljamaan, H., Ross, C., Kovscek, A., 2017. Multiscale imaging of gas storage in shales. SPE J. 22, 1760–1777.

Alnoaimi, K.R., Kovscek, A.R., 2019. Influence of microcracks on flow and storage capacities of gas shales at core scale. Transp. Porous Media 127, 53–84.

Alyaev, S., Keilegavlen, E., Nordbotten, J.M., 2018. A heterogeneous multiscale MPFA method for single-phase flows in porous media with inertial effects. Comput. Geosci. 23, 107–126.

Anderson, T.I., Vega, B., Kovscek, A.R., 2020a. Multimodal imaging and machine learning to enhance microscope images of shale. Comput. Geosci. 145, 104593.

Anderson, T., Guan, K., Vega, B., Aryana, S., Kovscek, A., 2020b. RockFlow: fast generation of synthetic source rock images using generative flow models. Energies 13, 6571.

Andrew, M., 2020. Permeability prediction using multivariant structural regression. E3S Web Conf. **146**, 1–7.

Ansumali, S., Karlin, V., 2002. I. Kinetic boundary conditions in the lattice Boltzmann method. Phys. Rev. E **66**, 026311.

Arbogast, T., 2002. Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow. Comput. Geosci. 6, 453–481.

Arbogast, T., 2012. Mixed multiscale methods for heterogeneous elliptic problems. In: Graham, I.G., Hou, T.Y., Lakkis, O., Scheichl, R. (Eds.), Lecture Notes in Computational Science and Engineering, vol. 83. Springer, Berlin, Heidelberg, pp. 243–283.

Arbogast, T., Minkoff, S.E., Keenan, P.T., 1998. An operator-based approach to upscaling the pressure equation. Trans. Ecol. Environ. 17.

Arbogast, T., Cowsar, L.C., Wheeler, M.F., Yotov, I., 2000. Mixed finite element methods on nonmatching multiblock grids. SIAM J. Numer. Anal. 37, 1295–1315.

Arbogast, T., Pencheva, G., Wheeler, M.F., Yotov, I., 2007. A multiscale mortar mixed finite element method. SIAM Multiscale Model. Simul. 6, 319–346.

Auriault, J.L., Adler, P.M., 1995. Taylor dispersion in porous media: analysis by multiple scale expansions. Adv. Water Resour. 18, 217–226.

Avatsmark, I., Eigestad, G., Mallison, B., Nordbotten, J., 2008. A compact multipoint flux approximation method with improved robustness. Numer. Methods Partial Differ. Equ. 24, 1329–1360.

Baber, K., Flemisch, B., Helmig, R., 2016. Modeling drop dynamics at the interface between free and porous-medium flow using the mortar method. Int. J. Heat Mass Transf. 99, 660–671.

Babuška, I., Osborn, J.E., 1983. Generalized finite element methods: their performance and their relation to mixed methods. SIAM J. Numer. Anal. 20, 510–536.

Babuška, I., Caloz, G., Osborn, J.E., 1994. Special finite element methods for a class of second order elliptic problems with rough coefficients. SIAM J. Numer. Anal. 31, 945–981.

Bai, T., Tahmasebi, P., 2020. Hybrid geological modeling: combining machine learning and multiple-point statistics. Comput. Geosci. 142, 104519.

Balhoff, M.T., Thompson, K.E., Hjortsø, M., 2007. Coupling pore-scale networks to continuum-scale models of porous media. Comput. Geosci. 33, 393–410.

Balhoff, M.T., Thomas, S.G., Wheeler, M.F., 2008. Mortar coupling and upscaling of pore-scale models. Comput. Geosci. 12, 15–27.

Bao, J., Murugesan, V., Kamp, C.J., Shao, Y., Yan, L., Wang, W., 2020. Machine learning coupled multi-scale modeling for redox flow batteries. Adv. Theory Simulat. 3, 1900167.

Barenblatt, G., Zheltov, I., Kochina, I., 1960. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks [strata]. J. Appl. Math. Mech. 24, 1286–1303.

Battiato, I., 2016. Multiscale models of flow and transport. In: The Handbook of Groundwater Engineering. CRC Press, pp. 359–381. https://doi.org/10.1201/9781315371801-22.

Battiato, I., Tartakovsky, D.M., 2011. Applicability regimes for macroscopic models of reactive transport in porous media. J. Contam. Hydrol. 120–121, 18–26.

Battiato, I., Tartakovsky, D.M., Tartakovsky, A.M., Scheibe, T., 2009. On breakdown of macroscopic models of mixing-controlled heterogeneous reactions in porous media. Adv. Water Resour. 32, 1664–1673.

Battiato, I., Tartakovsky, D.M., Tartakovsky, A.M., Scheibe, T.D., 2011. Hybrid models of reactive transport in porous and fractured media. Adv. Water Resour. 34, 1140–1150.

Battiato, I., Ferrero, V.P.T., O'Malley, D., Miller, C.T., Takhar, P.S., Valdés-Parada, F.J., Wood, B.D., 2019. Theory and applications of macroscale models in porous media. Transp. Porous Media 130, 5–76.

Bear, J., 2013. Dynamics of Fluids in Porous Media. Courier Corporation.

Bedford, A., Drumheller, D.S., 1983. Theories of immiscible and structured mixtures. Int. J. Eng. Sci. 21, 863–960.

Belgacem, F.B., 1999. The mortar finite element method with Lagrange multipliers. Numer. Math. 84, 173–197.

Bender, C.M., Orszag, S.A., 1999. Advanced Mathematical Methods for Scientists and Engineers I: Asymptotic Methods and Perturbation Theory. Springer Science & Business Media.

Bennethum, L.S., Cushman, J.H., 1996a. Multiscale, Hybrid Mixture Theory for Swelling Systems - I: Balance Laws, 34, pp. 125–145.

Bennethum, L.S., Cushman, J.H., 1996b. Multiscale, hybrid mixture theory for swelling systems - II: constitutive theory. Int. J. Eng. Sci. 34, 147–169.

Bensoussan, A., Lions, J., Papanicolaou, G., 2011. Asymptotic Analysis for Periodic Structures. American MathematicalSociety.

Berkowitz, B., Cortis, A., Dentz, M., Scher, H., 2006. Modeling non-Fickian transport in geological formations as a continuous time random walk. Rev. Geophys. 44.

Berkowitz, B., Dror, I., Hansen, S.K., Scher, H., 2016. Measurements and models of reactive transport in geological media. Rev. Geophys. 54, 930–986.

Bernardi, C., Maday, Y., Patera, A., 1994. A new nonconforming approach to domain decomposition: the mortar element method. In: Brezis, H., Lions, J.-L. (Eds.), Nonlinear Partial Differential Equations and Their Applications. Longman Sci. Tech., Harlow, UK, pp. 13–51.

Beucher, S., Lantuejoul, C., 1979. Use of watersheds in contour detection. In: International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation (doi:citeulike-article-id:4083187).

Blunt, M.J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., Paluszny, A., Pentland, C., 2013. Pore-scale imaging and modelling. Adv. Water Resour. 51, 197–216.

Boe, O., 1994. Analysis of an upscaling method based on conservation of dissipation. Transp. Porous Media 17, 77–86.

Boso, F., Battiato, I., 2013. Homogenizability conditions for multicomponent reactive transport. Adv. Water Resour. 62, 254–265.

Bourgeat, A., 1984. Homogenized behavior of two-phase flows in naturally fractured reservoirs with uniform fractures distribution. Comput. Methods Appl. Mech. Eng. 47, 205–216.

Brezzi, F., Douglas, J., Durán, R., Fortin, M., 1987. Mixed finite elements for second order elliptic problems in three variables. Numer. Math. 51, 237–250.

Caers, J., 2001. Geostatistical reservoir modelling using statistical pattern recognition. J. Pet. Sci. Eng. 29, 177–188.

Castelletto, N., Hajibeygi, H., Tchelepi, H.A., 2017. Multiscale finite-element method for linear elastic geomechanics. J. Comput. Phys. 331, 337–356.

Castelletto, N., Klevtsov, S., Hajibeygi, H., Tchelepi, H.A., 2019. Multiscale two-stage solver for Biot's poroelasticity equations in subsurface media. Comput. Geosci. 23, 207–224.

Center for Mechanistic Control of Unconventional Formations (CMC-UF). https://efrc-shale.stanford.edu/.

Chai, Z., Shi, B., Guo, Z., Lu, J., 2010. Gas flow through square arrays of circular cylinders with Klinkenberg effect: a lattice Boltzmann study. Commun. Comput. Phys. 8.

Chen, S., Doolen, G.D., 1998. Lattice Boltzmann method for fluid flows. Annu. Rev. Fluid Mech. 30, 329–364.

Chen, Y., Durlofsky, L.J., 2006a. Efficient incorporation of global effects in upscaled models of two-phase flow and transport in heterogeneous formations. Multiscale Model. Simul. 5, 445–475.

Chen, Y., Durlofsky, L.J., 2006b. Adaptive local-global upscaling for general flow scenarios in heterogeneous formations. Transp. Porous Media 62, 157–185.

Chen, Z., Hou, T.Y., 2002. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. Math. Comput. 72, 541–576.

Chen, Y., Li, Y., 2009. Local-global two-phase upscaling of flow and transport in heterogeneous formations. Multiscale Model. Simul. 8, 125–153.

Chen, Y., Li, Y., 2010. Incorporation of global effects in two-phase upscaling for modeling flow and transport with full-tensor anisotropy. In: ECMOR XII - 12th European Conference on the Mathematics of Oil Recovery. European Association of Geoscientists & Engineers. https://doi.org/10.3997/2214-4609.20144920.
Chen, Y., Durlofsky, L.J., Gerritsen, M., Wen, X.H., 2003. A coupled local-global upscaling approach for simulating flow in highly heterogeneous formations. Adv. Water Resour. 26, 1041–1060.
Christie, M.A., 1996. Upscaling for reservoir simulation. JPT: J. Pet. Technol. 48, 1004–1010.
Chu, J., Engquist, B., Prodanović, M., Tsai, R., 2012. A multiscale method coupling network and continuum models in porous media I: steady-state single phase flow. Multiscale Model. Simul. 10, 515–549.
Chu, J., Engquist, B., Prodanović, M., Tsai, R., 2013. A multiscale method coupling network and continuum models in porous media II-single- and two-phase flows. Adv. Appl. Math. Model. Comput. Sci. 66, 161–185.
Curtis, M.E., Sondergeld, C.H., Ambrose, R.J., Rai, C.S., 2012. Microstructural investigation of gas shales in two and three dimensions using nanometer-scale resolution imaging. Am. Assoc. Pet. Geol. Bull. 96, 665–677.
Cushman, J.H., Bennethum, L.S., Hu, B.X., 2002. A primer on upscaling tools for porous media. Adv. Water Resour. 25, 1043–1067.
Da Wang, Y., Armstrong, R.T., Mostaghimi, P., 2019. Enhancing resolution of digital rock images with super resolution convolutional neural networks. J. Pet. Sci. Eng. 182, 106261.
Dagan, G., 1989. Flow and Transport in Porous Formations. Springer Science & Business Media. https://doi.org/10.1007/978-3-642-75015-1.
Darcy, H., 1856. Les fontaines publiques de la ville de Dijon: exposition et application. Victor Dalmont.
Davit, Y., Bell, C.G., Byrne, H.M., Chapman, L.A.C., Kimpton, L.S., Lang, G.E., Leonard, K. H.L., Oliver, J.M., Pearson, N.C., Shipley, R.J., Waters, S.L., Whiteley, J.P., Wood, B. D., Quintard, M., 2013. Homogenization via formal multiscale asymptotics and volume averaging: how do the two techniques compare? Adv. Water Resour. 62, 178–206.
Deutsch, C., 1989. Calculating effective absolute permeability in sandstone/shale sequences. SPE Form. Eval. 4, 343–348.
Dinh, L., Krueger, D., Bengio, Y., 2014. NICE: non-linear independent components estimation. In: arXiv.
Dinh, L., Sohl-Dickstein, J., Bengio, S., 2016. Density estimation using real NVP. In: arXiv.
Do, H.D., Do, D.D., Prasetyo, I., 2001. Surface diffusion and adsorption of hydrocarbons in activated carbon. AICHE J. 47, 2515–2525.
Dolean, V., Jolivet, P., Nataf, F., 2015. An Introduction to Domain Decomposition Methods: Algorithms, Theory, and Parallel Implementation. SIAM. https://doi.org/10.1137/1.9781611974065.
Dong, H., Blunt, M.J., 2009. Pore-network extraction from micro-computerized-tomography images. Phys. Rev. E 80, 036307.
Dong, C., Loy, C.C., He, K., Tang, X., 2016. Image super-resolution using deep convolutional networks. IEEE Trans. Pattern Anal. Mach. Intell. 38, 295–307.
Durlofsky, L.J., 1991. Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. Water Resour. Res. 27, 699–708.
Durlofsky, L.J., 2005. Upscaling and gridding of fine scale geological models for flow simulation. In: 8th International Forum on Reservoir Simulation, pp. 1–59.
Durlofsky, L.J., Chen, Y., 2012. Uncertainty quantification for subsurface flow problems using coarse-scale models. In: Graham, I., Hou, T., Lakkis, O., Scheichl, R. (Eds.), Numerical Analysis of Multiscale Problems, vol. 83. Springer, Berlin, Heidelberg, pp. 163–202.
E, W., 1992. Homogenization of linear and nonlinear transport equations. Commun. Pure Appl. Math. 45, 301–326.
E, W., Engquist, B., 2003. The heterogeneous multiscale methods. Commun. Math. Sci. 1, 87–132.
E, W., Engquist, B., Li, X., Ren, W., Vanden-Eijnden, E., 2007. The heterogeneous multiscale method: a review. Commun. Comput. Phys. 2, 367–450.
Efendiev, Y., Durlofsky, L.J., 2003. A generalized convection-diffusion model for subgrid transport in porous media. Multiscale Model. Simul. 1, 504–526.
Efendiev, Y., Hou, T.Y., 2008. Multiscale Finite Element Methods, Theory and Applications. Springer Science & Business Media. https://doi.org/10.1007/978-0-387-09496-0.
Efendiev, Y., Durlofsky, L.J., Lee, S.H., 2000. Modeling of subgrid effects in coarse-scale simulations of transport in heterogeneous porous media. Water Resour. Res. 36, 2031–2041.
Efros, A.A., Freeman, W.T., 2001. Image quilting for texture synthesis and transfer. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 2001. SIGGRAPH, pp. 341–346. https://doi.org/10.1145/383259.383296.
Ellis, D.V., Singer, J.M., 2007. Well Logging for Earth Scientists. Springer.
Farmer, C.L., 2002. Upscaling : a review. Int. J. Numer. Methods Fluids 40, 63–78.
Firouzi, M., Wilcox, J., 2012. Molecular modeling of carbon dioxide transport and storage in porous carbon-based materials. Microporous Mesoporous Mater. 158, 195–203.
Frouté, L., Kovscek, A.R., 2020. Nano-imaging of shale using electron microscopy techniques. In: SPE/AAPG/SEG Unconventional Resources Technology Conference. Unconventional Resources Technology Conference (URTEC). https://doi.org/10.15530/urtec-2020-3283.
Fuks, O., Tchelepi, H.A., 2020. Limitations of physics informed machine learning for nonlinear two-phase transport in porous media. J. Mach. Learn. Model. Comput. 1, 19–37.
Ganis, B., Yotov, I., 2009. Implementation of a mortar mixed finite element method using a Multiscale Flux Basis. Comput. Methods Appl. Mech. Eng. 198, 3989–3998.
Ganis, B., Juntunen, M., Pencheva, G., Wheeler, M.F., Yotov, I., 2014a. A global jacobian method for mortar discretizations of nonlinear porous media flows. SIAM J. Sci. Comput. 36, A522–A542.
Ganis, B., Kumar, K., Pencheva, G., Wheeler, M.F., Yotov, I., 2014b. A global Jacobian method for mortar discretizations of a fully implicit two-phase flow model. Multiscale Model. Simul. 12, 1401–1423.
Ganis, B., Pencheva, G., Wheeler, M.F., 2019. Adaptive mesh refinement with an enhanced velocity mixed finite element method on semi-structured grids using a fully coupled solver. Comput. Geosci. 23, 149–168.
Gatica, G.N., 2014. A Simple Introduction to the Mixed Finite Element Method: Theory and Applications. Springer Science & Business Media. https://doi.org/10.1007/978-3-319-03695-3.
Gautier, Y., Blunt, M.J., Christie, M.A., 1999. Nested gridding and streamline-based simulation for fast reservoir performance prediction. Comput. Geosci. 3, 295–320.
Glimm, J., Hou, S., Kim, H., Lee, Y., Sharp, D.H., Ye, K., Zou, Q., 2001. Risk management for petroleum reservoir production: a simulation-based study of prediction. Comput. Geosci. 5, 173–197.
Golfier, F., Zarcone, C., Bazin, B., Lenormand, R., Lasseux, D., Quintard, M., 2002. On the ability of a Darcy-scale model to capture wormhole formation during the dissolution of a porous medium. J. Fluid Mech. 457, 213–254.
Gomez-Hernandez, J.J., 1991. A Stochastic Approach to the Simulation of Block Conductivity Fields Conditioned Upon Data Measured at a Smaller Scale. Stanford University.
Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. In: Advances in Neural Information Processing Systems 27 (NIPS 2014), pp. 2672–2680.
Goral, J., Walton, I., Andrew, M., Deo, M., 2019. Pore system characterization of organic-rich shales using nanoscale-resolution 3D imaging. Fuel 258, 116049.
Gray, W.G., Miller, C.T., 2014. Introduction to the Thermodynamically Constrained Averaging Theory for Porous Medium Systems. Springer.
Gray, W.G., Leijnse, A., Kolar, R.L., Blain, C.A., 1993. Mathematical Tools for Changing Spatial Scales in the Analysis of Physical Systems. Taylor & Francis Group.
Gray, W.G., Miller, C.T., Schrefler, B.A., 2013. Averaging theory for description of environmental problems: what have we learned? Adv. Water Resour. 51, 123–138.
Griewank, A., Walther, A., 2008. Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation. SIAM.
Guan, K., Anderson, T., Kovscek, A., 2021. Reconstructing porous media using generative flow networks. Comput. Geosci. 156, 104905.
Guo, Z., Zheng, C., Shi, B., 2008. Lattice Boltzmann equation with multiple effective relaxation times for gaseous microscale flow. Phys. Rev. E 77, 036707.
Guo, B., Mehmani, Y., Tchelepi, H.A., 2019. Multiscale formulation of pore-scale compressible Darcy-Stokes flow. J. Comput. Phys. 397, 108849.
Hajibeygi, H., Bonfigli, G., Hesse, M.A., Jenny, P., 2008. Iterative multiscale finite-volume method. J. Comput. Phys. 227, 8604–8621.
Harrison, A.L., Jew, A.D., Dustin, M.K., Thomas, D.L., Joe-Wong, C.M., Bargar, J.R., Johnson, N., Brown, G.E., Maher, K., 2017. Element release and reaction-induced porosity alteration during shale-hydraulic fracturing fluid interactions. Appl. Geochem. 82, 47–62.
Hassanizadeh, S.M., Gray, W.G., 1990. Mechanics and thermodynamics of multiphase flow in porous media including interphase boundaries. Adv. Water Resour. 13, 169–186.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag.
He, X., Luo, L.S., 1997. Theory of the lattice Boltzmann method: from the Boltzmann equation to the lattice Boltzmann equation. Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top. 55, 6811–6820.
Hemminger, J.C., Sarrao, J., Crabtree, G., Flemming, G., Ratner, M.. Challenges at the Frontiers of Matter and Energy: Transformative Opportunities for Discovery Science. http://www.osti.gov/servlets/purl/1283188/.
Holden, L., Nielsen, B.F., 2000. Global upscaling of permeability in heterogeneous reservoirs; the Output Least Squares (OLS) method. Transp. Porous Media 40, 115–143.
Honarpour, M.M., Cromwell, V., Hatton, D., Satchwell, R., 1985. Reservoir rock descriptions using computed tomography (CT). In: SPE Annual Technical Conference and Exhibition. OnePetro. https://doi.org/10.2118/14272-MS.
Hornung, U., 1996. Homogenization and Porous Media, vol. 6. Springer Science & Business Media.
Hou, T.Y., Wu, X.H., 1997. A multiscale finite element method for elliptic problems in composite materials and porous media. J. Comput. Phys. 134, 169–189.
Hou, T.Y., Wu, X.-H., Cai, Z., 1999. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. Math. Comput. 68, 913–943.
Hou, T.Y., Westhead, A., Yang, D., 2006. A framework for modeling subgrid effects for two-phase flows in porous media. Multiscale Model. Simul. 5, 1087–1127.
Howarth, R.W., Santoro, R., Ingraffea, A., 2011. Methane and the greenhouse-gas footprint of natural gas from shale formations. Clim. Chang. 106, 679–690.
Hughes, T.J.R., 1995. Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. Comput. Methods Appl. Mech. Eng. 127, 387–401.
Hughes, R.G., Blunt, M.J., 2001. Pore-Scale modeling of multiphase flow in fractures and matrix/fracture transfer. SPE J. 6, 126–136.
Hughes, T.J.R., Feijóo, G.R., Mazzei, L., Quincy, J.B., 1998. The variational multiscale method—a paradigm for computational mechanics. Comput. Methods Appl. Mech. Eng. 166, 3–24.

IEA, 2017. World Energy Outlook.
Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A., Research, B.A., 2017. Image-to-image translation with conditional adversarial networks. In: Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1125–1134.
Jenny, P., Lee, S.H., Tchelepi, H.A., 2003. Multi-scale finite-volume method for elliptic problems in subsurface flow simulation. J. Comput. Phys. 187, 47–67.
Jenny, P., Lee, S.H., Tchelepi, H.A., 2004. Adaptive multiscale finite-volume method for multiphase flow and transport in porous media. SIAM Multiscale Model. Simul. 3, 50–64.
Jenny, P., Lee, S.H., Tchelepi, H.A., 2006. Adaptive fully implicit multi-scale finite-volume method for multi-phase flow and transport in heterogeneous porous media. J. Comput. Phys. 217, 627–641.
Jew, A., Druhan, J.L., Kovscek, A., Bargar, J.R., Alvarado, V., Brown, G., 2020. Chemical and reactive transport processes associated with hydraulic fracturing of unconventional oil/gas shales. Chem. Rev. Submitt.
Jin, Z., Firoozabadi, A., 2016. Phase behavior and flow in shale nanopores from molecular simulations. Fluid Phase Equilib. 430, 156–168.
Joekar-Niasar, V., Hassanizadeh, S., Dahle, H., 2010. Non-equilibrium effects in capillarity and interfacial area in two-phase flow: dynamic pore-network modelling. J. Fluid Mech. 655, 38–71.
Johnson, J.W., Nitao, J.J., Knauss, K.G., 2004. Reactive transport modelling of CO2 storage in saline aquifers to elucidate fundamental processes, trapping mechanisms and sequestration partitioning. Geol. Soc. London Spec. Publ. 233, 107–128.
Kamrava, S., Tahmasebi, P., Sahimi, M., 2019. Enhancing images of shale formations by a hybrid stochastic and deep learning algorithm. Neural Netw. 118, 310–320.
Kamrava, S., Tahmasebi, P., Sahimi, M., 2020. Linking morphology of porous media to their macroscopic permeability by deep learning. Transp. Porous Media 131, 427–448.
Keller, J.B., 1980. Darcy's law for flow in porous media and the two-space method. In: Sternberg, R.L., Kalinowski, A.J., Papadakis, J.S. (Eds.), Nonlinear Partial Differential Equations in Engineering and Applied Science. Marcel Dekker, Inc., pp. 429–443. https://doi.org/10.1201/9780203745465-27
Khan, H.J., Spielman-Sun, E., Jew, A.D., Bargar, J., Kovscek, A., Druhan, J.L., 2021. A critical review of the physicochemical impacts of water chemistry on shale in hydraulic fracturing systems. Environ. Sci. Technol. 55, 1377–1394.
Khayrat, K., Jenny, P., 2017. A multi-scale network method for two-phase flow in porous media. J. Comput. Phys. 342, 194–210.
Khayrat, K., Epp, R., Jenny, P., 2018. Approximate multiscale flow solver for unstructured pore networks. J. Comput. Phys. 372, 62–79.
Kingma, D.P., Ba, J.L., 2014. Adam: a method for stochastic optimization. In: arXiv.
Kingma, D.P., Dhariwal, P., 2018. Glow: generative flow with invertible 1×1 convolutions. In: *Advances in Neural Information Processing Systems 31 (NIPS* 2018). Curran Associates, Inc.
Klie, M., Florez, H., 2018. Data-driven modeling of fractured shale reservoirs. In: *16th European Conference on the Mathematics of Oil Recovery, ECMOR* 2018, vol. 2018. European Association of Geoscientists and Engineers, EAGE, pp. 1–18.
Kolyukhin, D., Espedal, M., 2010. Modified adaptive local–global upscaling method for discontinuous permeability distribution. Comput. Geosci. 14, 675–689.
Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi Twitter, W., 2017. Photo-realistic single image super-resolution using a Generative Adversarial Network. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4681–4690.
Lee, S.H., Wolfsteiner, C., Tchelepi, H.A., 2008. Multiscale finite-volume formulation for multiphase flow in porous media: black oil formulation of compressible, three-phase flow with gravity. Comput. Geosci. 12, 351–366.
Li, H., Durlofsky, L.J., 2016. Ensemble level upscaling for compositional flow simulation. Comput. Geosci. 20, 525–540.
Lie, K., 2019. An Introduction to Reservoir Simulation Using MATLAB/GNU Octave: User Guide for the MATLAB Reservoir Simulation Toolbox (MRST). Cambridge University Press.
Liehui, Z., Baochao, S., Yulong, Z., Zhaoli, G., 2019. Review of micro seepage mechanisms in shale gas reservoirs. Int. J. Heat Mass Transf. 139, 144–179.
Lim, C.Y., Shu, C., Niu, X.D., Chew, Y.T., 2002. Application of lattice Boltzmann method to simulate microchannel flows. Phys. Fluids 14, 2299–2308.
Lindeberg, E., Wessel-Berg, D., 1997. Vertical convection in an aquifer column under a gas cap of CO2. Energy Convers. Manag. 38, S229–S234.
Lindquist, W.B., Lee, S.-M., Coker, D.A., Jones, K.W., Spanne, P., 1996. Medial axis analysis of void structure in three-dimensional tomographic images of porous media. J. Geophys. Res. Solid Earth 101, 8297–8310.
Liu, L., Wang, Y., Aryana, S.A., 2021. Insights into scale translation of methane transport in nanopores. Journal of Natural Gas Science and Engineering 96, 104220.
Lødøen, O.P., Omre, H., Durlofsky, L.J., Chen, Y., 2005. Assessment of uncertainty in reservoir production forecasts using upscaled flow models. In: Geostatistics Banff 2004. Springer, Dordrecht, pp. 713–722. https://doi.org/10.1007/978-1-4020-3610-1_72.
Loucks, R.G., Reed, R.M., Ruppel, S.C., Hammes, U., 2012. Spectrum of pore types and networks in mudrocks and a descriptive classification for matrix-related mudrock pores. Am. Assoc. Pet. Geol. Bull. 96, 1071–1098.
Lugo-Méndez, H.D., Valdés-Parada, F.J., Porter, M.L., Wood, B.D., Ochoa-Tapia, J.A., 2015. Upscaling diffusion and nonlinear reactive mass transport in homogeneous porous media. Transp. Porous Media 107, 683–716.
Lunati, I., Jenny, P., 2006. Multiscale finite-volume method for compressible multiphase flow in porous media. J. Comput. Phys. 216, 616–636.
Lunati, I., Jenny, P., 2008. Multiscale finite-volume method for density-driven flow in porous media. Comput. Geosci. 12, 337–350.
Lunati, I., Lee, S.H., 2009. An operator formulation of the multiscale finite-volume method with correction function. SIAM Multiscale Model. Simul. 8, 96–109.
Lunati, I., Tyagi, M., Lee, S.H., 2011. An iterative multiscale finite volume algorithm converging to the exact solution. J. Comput. Phys. 230, 1849–1864.
Lyder, D., Feng, J., Rivard, B., Gallie, A., Cloutis, E., 2010. Remote bitumen content estimation of Athabasca oil sand from hyperspectral infrared reflectance spectra using Gaussian singlets and derivative of Gaussian wavelets. Fuel 89, 760–767.
Ma, L., Slater, T., Dowey, P.J., Yue, S., Rutter, E.H., Taylor, K.G., Lee, P.D., 2018. Hierarchical integration of porosity in shales. Sci. Rep. 8, 11683.
Manwart, C., Torquato, S., Hilfer, R., 2000. Stochastic reconstruction of sandstones. Phys. Rev. E 62, 893–899.
Marle, C.M., 1982. On macroscopic equations governing multiphase flow with diffusion and chemical reactions in porous media. Int. J. Eng. Sci. 20, 643–662.
Mehmani, Y., 2014. Modeling Single-Phase Flow and Solute Transport Across Scales. The University of Texas at Austin.
Mehmani, Y., Balhoff, M.T., 2014. Bridging from pore to continuum: a hybrid mortar domain decomposition framework for subsurface flow and transport. SIAM Multiscale Model. Simul. 12, 667–693.
Mehmani, Y., Balhoff, M.T., 2015a. Mesoscale and hybrid models of fluid flow and solute transport. Rev. Mineral. Geochem. 80, 433–459.
Mehmani, Y., Balhoff, M.T., 2015b. Eulerian network modeling of longitudinal dispersion. Water Resour. Res. 51, 8586–8606.
Mehmani, Y., Tchelepi, H.A., 2017. Minimum requirements for predictive pore-network modeling of solute transport in micromodels. Adv. Water Resour. 108, 83–98.
Mehmani, Y., Tchelepi, H.A., 2018. Multiscale computation of pore-scale fluid dynamics: single-phase flow. J. Comput. Phys. 375, 1469–1487.
Mehmani, Y., Tchelepi, H.A., 2019. Multiscale formulation of two-phase flow at the pore scale. J. Comput. Phys. 389, 164–188.
Mehmani, Y., Sun, T., Balhoff, M.T., Eichhubl, P., Bryant, S., 2012. Multiblock pore-scale modeling and upscaling of reactive transport: application to carbon sequestration. Transp. Porous Media 95, 305–326.
Mehmani, Y., Oostrom, M., Balhoff, M.T., 2014. A streamline splitting pore-network approach for computationally inexpensive and accurate simulation of transport in porous media. Water Resour. Res. 50, 2488–2517.
Mehmani, Y., Burnham, A.K., Tchelepi, H.A., 2016a. From optics to upscaled thermal conductivity: Green River oil shale. Fuel 183, 489–500.
Mehmani, Y., Burnham, A.K., Vanden Berg, M.D., Gelin, F., Tchelepi, H., 2016b. Quantification of kerogen content in organic-rich shales from optical photographs. Fuel 177, 63–75.
Mehmani, Y., Burnham, A.K., Vanden Berg, M.D., Tchelepi, H.A., 2017. Quantification of organic content in shales via near-infrared imaging: Green River Formation. Fuel 208, 337–352.
Mehmani, Y., Castelletto, N., Tchelepi, H.A., 2021. Multiscale formulation of frictional contact mechanics at the pore scale. J. Comput. Phys. 430, 110092.
Mikelić, A., Devigne, V., van Duijn, C.J., 2006. Rigorous upscaling of the reactive flow through a pore, under dominant Peclet and Damkohler numbers. SIAM J. Math. Anal. 38, 1262–1287.
Mirza, M., Osindero, S., 2014. Conditional generative adversarial nets. In: arXiv.
Molins, S., Trebotich, D., Arora, B., Steefel, C.I., Deng, H., 2019. Multi-scale model of reactive transport in fractured media: diffusion limitations on rates. Transp. Porous Media 128, 701–721.
Mosser, L., Dubrule, O., Blunt, M.J., 2017. Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys. Rev. E 96, 43309.
Mosser, L., Dubrule, O., Blunt, M.J., 2018. Stochastic reconstruction of an oolitic limestone by generative adversarial networks. Transp. Porous Media 125, 81–103.
Neal, J., Krohn, C., 2012. Higher resolution subsurface imaging. J. Pet. Technol. 64, 44–53.
Neuman, S.P., 1977. Theoretical derivation of Darcy's law. Acta Mech. 25, 153–170.
Nie, X., Doolen, G.D., Chen, S., 2002. Lattice-Boltzmann simulations of fluid flows in MEMS. J. Stat. Phys. 107, 279–289.
Nocedal, J., Wright, S., 2006. Numerical Optimization, Series in Operations Research and Financial Engineering. Springer, New York.
NSF, 2006. In: Oden, T.J. (Ed.), Simulation-Based Engineering Science: Revolutionizing Engineering Science through Simulation.
O'Sullivan, A.E., Christie, M.A., 2005. Solution error models: a new approach for coarse grid history matching. In: SPE Reservoir Simulation Symposium, pp. 307–315. https://doi.org/10.2118/93268-MS.
Oden, J.T., Prudhomme, S., Romkes, A., Bauman, P.T., 2006. Multiscale modeling of physical phenomena: adaptive control of models. SIAM J. Sci. Comput. 28, 2359–2389.
Okabe, H., Blunt, M.J., 2004. Prediction of permeability for porous media reconstructed using multiple-point statistics. Phys. Rev. E 70, 066135.
Okabe, H., Blunt, M.J., 2007. Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics. Water Resour. Res. 43.
Patzek, T.W., Male, F., Marder, M., 2013. Gas production in the Barnett Shale obeys a simple scaling theory. Proc. Natl. Acad. Sci. 110, 19731–19736.
Peng, D.Y., Robinson, D.B., 1976. A new two-constant equation of state. Ind. Eng. Chem. Fundam. 15, 59–64.
Peszynska, M., Lu, Q., Wheeler, M., 1999. Coupling different numerical algorithms for two phase fluid flow. In: Whiteman, J.R. (Ed.), Mathematics of Finite Elements and Applications X. Elsevier.
Peszynska, M., Lu, Q., Wheeler, M.F., 2000. Multiphysics coupling of codes. In: Computational Methods in Water Resources - Volume 1 - Computational Methods For Subsurface Flow and Transport, pp. 175–182.
Phan, A., Fan, D., Striolo, A., 2020. Fluid transport through heterogeneous pore matrices: multiscale simulation approaches. Phys. Fluids 32, 101301.

Pickup, G.E., Ringrose, P.S., Jensen, J.L., Sorbie, K.S., 1994. Permeability tensors for sedimentary structures. Math. Geol. 26, 227–250.
Pini, R., Madonna, C., 2016. Moving across scales: a quantitative assessment of X-ray CT to measure the porosity of rocks. J. Porous. Mater. 23, 325–338.
Popinet, S., 2018. Numerical models of surface tension. Annu. Rev. Fluid Mech. 50, 49–75.
Prodanović, M., Lindquist, W.B., Seright, R.S., 2007. 3D image-based characterization of fluid displacement in a Berea core. Adv. Water Resour. 30, 214–226.
Pruess, Karsten, 1992. Brief guide to the MINC-method for modeling flow and transport in fractured media. Lawrence Berkeley Laboratory, CA, USA.
Quintard, M., Whitaker, S., 1988. Two-phase flow in heterogeneous porous media: the method of large-scale averaging. Transp. Porous Media 3, 357–413.
Rabinovich, A., Itthisawatpan, K., Durlofsky, L.J., 2015. Upscaling of CO2 injection into brine with capillary heterogeneity effects. J. Pet. Sci. Eng. 134, 60–75.
Raeini, A.Q., Bijeljic, B., Blunt, M.J., 2018. Generalized network modeling of capillary-dominated two-phase flow. Phys. Rev. E 97, 023308.
Raissi, M., 2018. Deep hidden physics models: deep learning of nonlinear partial differential equations. J. Mach. Learn. Res. 19, 1–24.
Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707.
Renard, P., de Marsily, G., 1997. Calculating equivalent permeability: a review. Adv. Water Resour. 20, 253–278.
Roberts, A.P., 1997. Statistical reconstruction of three-dimensional porous media from two-dimensional images. Phys. Rev. E 56, 3203–3212.
Romeu, R.K., Noetinger, B., 1995. Calculation of internodal transmissivities in finite difference models of flow in heterogeneous porous media. Water Resour. Res. 31, 943–959.
Roubinet, D., Tartakovsky, D.M., 2013. Hybrid modeling of heterogeneous geochemical reactions in fractured porous media. Water Resour. Res. 49, 7945–7956.
Rubin, Y., 2003. Applied Stochastic Hydrogeology. Oxford University Press.
Saad, Y., 2003. Iterative Methods for Sparse Linear Systems. SIAM. https://doi.org/10.1137/1.9780898718003.
Sáez, A.E., Otero, C.J., Rusinek, I., 1989. The effective homogeneous behavior of heterogeneous porous media. Transp. Porous Media 4, 213–238.
Salamon, P., Fernàndez-Garcia, D., Gómez-Hernández, J.J., 2006. A review and numerical assessment of the random walk particle tracking method. J. Contam. Hydrol. 87, 277–305.
Santos, J.E., Xu, D., Jo, H., Landry, C.J., Prodanović, M., Pyrcz, M.J., 2020. PoreFlow-Net: a 3D convolutional neural network to predict fluid flow through porous media. Adv. Water Resour. 138, 103539.
Scheibe, T.D., Yang, X., Schuchardt, K., Agarwal, K., Chase, J., Palmer, B., Tartakovsky, A., 2014. A many-task parallel approach for multiscale simulations of subsurface flow and reactive transport. In: Proceedings 7th ACM Workshop on MTAGS.
Scheibe, T.D., Murphy, E.M., Chen, X., Rice, A.K., Carroll, K.C., Palmer, B.J., Tartakovsky, A.M., Battiato, I., Wood, B.D., 2015a. An analysis platform for multiscale hydrogeologic modeling with emphasis on hybrid multiscale methods. Groundwater 53, 38–56.
Scheibe, T.D., Schuchardt, K., Agarwal, K., Chase, J., Yang, X., Palmer, B.J., Tartakovsky, A.M., Elsethagen, T., Redden, G., 2015b. Hybrid multiscale simulation of a mixing-controlled reaction. Adv. Water Resour. 83, 228–239.
Scheibe, T.D., Yang, X., Chen, X., Hammond, G., 2015c. A hybrid multiscale framework for subsurface flow and transport simulations. Procedia Comput. Sci. 51, 1098–1107.
Shams, R., Masihi, M., Boozarjomehry, R.B., Blunt, M.J., 2020. Coupled generative adversarial and auto-encoder neural networks to reconstruct three-dimensional multi-scale porous media. J. Pet. Sci. Eng. 186, 106794.
Shan, X., Chen, H., 1993. Lattice Boltzmann model for simulating flows with multiple phases and components. Phys. Rev. E 47, 1815–1819.
Sheppard, A., Sok, R., Averdunk, H., Robins, V., Ghous, A., 2006. Analysis of rock microstructure using high-resolution X-ray tomography. In: Proceedings of the International Symposium of the Society of Core Analysts.
Silin, D., Patzek, T., 2006. Pore space morphology analysis using maximal inscribed spheres. Phys. A Stat. Mech. Appl. 371, 336–360.
Skoulidas, A.I., Ackerman, D.M., Johnson, J.K., Sholl, D.S., 2002. Rapid transport of gases in carbon nanotubes. Phys. Rev. Lett. 89, 185901.
Sobecki, N., Nieto-Draghi, C., Di Lella, A., Ding, D.Y., 2019. Phase behavior of hydrocarbons in nano-pores. Fluid Phase Equilib. 497, 104–121.
Speta, M., Rivard, B., Feng, J., Lipsett, M., Gingras, M., 2015. Hyperspectral imaging for the determination of bitumen content in Athabasca oil sands core samples. Am. Assoc. Pet. Geol. Bull. 99, 1245–1259.
Sukop, M.C., Or, D., 2004. Lattice Boltzmann method for modeling liquid-vapor interface configurations in porous media. Water Resour. Res. 40.
Sun, T., 2012. Upscaling and Multiscale Simulation by Bridging Pore Scale and Continuum Scale Models. The University of Texas at Austin.
Sun, T., Mehmani, Y., Balhoff, M.T., 2012a. Hybrid multiscale modeling through direct substitution of pore-scale models into near-well reservoir simulators. Energy Fuel 26, 5828–5836.
Sun, T., Mehmani, Y., Bhagmane, J., Balhoff, M.T., 2012b. Pore to continuum upscaling of permeability in heterogeneous porous media using mortars. Int. J. Oil, Gas Coal Technol. 5, 249–266.
Tahmasebi, P., Hezarkhani, A., Sahimi, M., 2012. Multiple-point geostatistical modeling based on the cross-correlation functions. Comput. Geosci. 16, 779–797.
Tang, G.H., Tao, W.Q., He, Y.L., 2005. Lattice Boltzmann method for gaseous microflows using kinetic theory boundary conditions. Phys. Fluids 17, 1–4.
Tang, Y., Valocchi, A.J., Werth, C.J., 2015. A hybrid pore-scale and continuum-scale model for solute diffusion, reaction, and biofilm development in porous media. Water Resour. Res. 51, 1846–1859.
Tartakovsky, A.M., Scheibe, T.D., 2011. Dimension reduction numerical closure method for advection-diffusion-reaction systems. Adv. Water Resour. 34, 1616–1626.
Tartakovsky, A.M., Tartakovsky, D.M., Scheibe, T.D., Meakin, P., 2008. Hybrid simulations of reaction-diffusion systems in porous media. SIAM J. Sci. Comput. 30, 2799–2816.
Tartar, L., 1980. Incompressible fluid flow in a porous medium-convergence of the homogenization process. In: Appendix of Non-homogeneous media and Vibration Theory, Lecture Notes in Physics, vol. 127.
Tartar, L., 1989. Nonlocal effects induced by homogenization. In: Colombini, F., Marino, A., Modica, L., Spagnolo, S. (Eds.), Partial Differential Equations and the Calculus of Variations, vol. 2. Springer Science & Business Media, LLC, pp. 925–938.
Thomas, S.G., Wheeler, M.F., 2011. Enhanced velocity mixed finite element methods for modeling coupled flow and transport on non-matching multiblock grids analysis and applications to transport of reactive species in multiphase flow through porous media. Comput. Geosci. 15, 605–625.
Thompson, K.E., 2002. Pore-scale modeling of fluid transport in disordered fibrous materials. AICHE J. 48, 1369–1389.
Tomin, P., Lunati, I., 2013. Hybrid multiscale finite volume method for two-phase flow in porous media. J. Comput. Phys. 250, 293–307.
Tomin, P., Lunati, I., 2015. Local–global splitting for spatiotemporal-adaptive multiscale methods. J. Comput. Phys. 280, 214–231.
Tomin, P., Lunati, I., 2016a. Investigating Darcy-scale assumptions by means of a multiphysics algorithm. Adv. Water Resour. 95, 80–91.
Tomin, P., Lunati, I., 2016b. Spatiotemporal adaptive multiphysics simulations of drainage-imbibition cycles. Comput. Geosci. 20, 541–554.
Van Der Waals, J., 1873. On the Continuity of the Gaseous and Liquid States. Universiteit Leiden.
Vega, B., Dutta, A., Kovscek, A.R., 2014. CT imaging of low-permeability, dual-porosity systems using high X-ray contrast gas. Transp. Porous Media 101, 81–97.
Vega, B., Ross, C.M., Kovscek, A.R., 2015. Imaging-based characterization of calcite-filled fractures and porosity in shales. SPE J. 20, 810–823.
Wallstrom, T.C., Hou, S., Christie, M.A., Durlofsky, L.J., Sharp, D.H., Zou, Q., 2002. Application of effective flux boundary conditions to two-phase upscaling in porous media. Transp. Porous Media 46, 155–178.
Wang, Y., Aryana, S.A., 2020. Pore-scale simulation of gas flow in microscopic permeable media with complex geometries. J. Nat. Gas Sci. Eng. 81, 103441.
Wang, Y., Aryana, S.A., 2021. Coupled confined phase behavior and transport of methane in slit nanopores. Chem. Eng. J. 404, 126502.
Wang, Z., Battiato, I., 2021. Upscaling reactive transport and clogging in shale microcracks by deep learning. Water Resour. Res. **57**, e2020WR029125.
Wang, L.L., Zhang, G.Q., Hallais, S., Tanguy, A., Yang, D.S., 2017. Swelling of shales: a multiscale experimental investigation. Energy Fuel 31, 10442–10451.
Wang, Y., Rahman, S.S., Arns, C.H., 2018. Super resolution reconstruction of μ-CT image of rock sample using neighbour embedding algorithm. Phys. A Stat. Mech. Appl. 493, 177–188.
Warren, J.E., Root, P.J., 1963. The behavior of naturally fractured reservoirs. SPE J. 245–255.
Weishaupt, K., Joekar-Niasar, V., Helmig, R., 2019. An efficient coupling of free flow and porous media flow using the pore-network modeling approach. J. Comput. Phys. X 1, 100011.
Wen, X.H., Durlofsky, L.J., Edwards, M.G., 2003. Use of border regions for improved permeability upscaling. Math. Geol. 35, 521–547.
Wheeler, J.A., Wheeler, M.F., Yotov, I., 2002. Enhanced velocity mixed finite element methods for flow in multiblock domains. Comput. Geosci. 6, 315–332.
Whitaker, S., 1986. Flow in porous media I: a theoretical derivation of Darcy's law. Transp. Porous Media 1, 3–25.
Whitaker, S., 1996. The Forchheimer equation: a theoretical development. Transp. Porous Media 25, 27–61.
Whitaker, S., 2013. The Method of Volume Averaging. Springer Science & Business Media.
White, C., Horne, R.N., 1987. Computing absolute transmissibility in the presence of fine-scale heterogeneity. In: SPE Symposium on Reservoir Simulation. https://doi.org/10.2118/16011-MS.
Wildenschild, D., Sheppard, A.P., 2013. X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. Adv. Water Resour. 51, 217–246.
Wolfsteiner, C., Lee, S.H., Tchelepi, H.A., 2006. Well modeling in the multiscale finite volume method for subsurface flow simulation. SIAM Multiscale Model. Simul. 5, 900–917.
Wood, B.D., 2009. The role of scaling laws in upscaling. Adv. Water Resour. 32, 723–736.
Wood, B.D., Valdés-Parada, F.J., 2013. Volume averaging: local and nonlocal closures using a Green's function approach. Adv. Water Resour. 51, 139–167.
Wu, H., Fang, W.Z., Kang, Q., Tao, W.Q., Qiao, R., 2019a. Predicting effective diffusivity of porous media from images by deep learning. Sci. Rep. 9, 20387.
Wu, Y., Tahmasebi, P., Lin, C., Ren, L., Dong, C., 2019b. Multiscale modeling of shale samples based on low- and high-resolution images. Mar. Pet. Geol. 109, 9–21.
Yang, A., 2013. On the common conceptual and computational frameworks for multiscale modeling. Ind. Eng. Chem. Res. 52, 11451–11462.
Yang, G., Fan, Z., Li, X., 2019. Determination of confined fluid phase behavior using extended Peng-Robinson equation of state. Chem. Eng. J. 378, 122032.
Yergin, D., 2020. The New Map: Energy, Climate, and the Clash of Nations. Penguin Press.

Younis, R.M., 2011. Modern Advances in Software and Solution Algorithms for Reservoir Simulation. Stanford University.
Yousefzadeh, M., Battiato, I., 2017. Physics-based hybrid method for multiscale transport in porous media. J. Comput. Phys. 344, 320–338.
Yu, H., Chen, J., Zhu, Y.B., Wang, F.C., Wu, H.A., 2017. Multiscale transport mechanism of shale gas in micro/nano-pores. Int. J. Heat Mass Transf. 111, 1172–1180.
Yuan, P., Schaefer, L., 2006. Equations of state in a lattice Boltzmann model. Phys. Fluids 18, 042101.
Zhang, D., 2001. Stochastic Methods for Flow in Porous Media: Coping with Uncertainties. Elsevier.
Zhang, C., Dehoff, K., Hess, N., Oostrom, M., Wietsma, T.W., Valocchi, A.J., Fouke, B.W., Werth, C.J., 2010. Pore-Scale study of transverse mixing induced CaCO3 precipitation and permeability reduction in a model subsurface sedimentary system. Environ. Sci. Technol. 44, 7833–7838.
Zhang, W.M., Meng, G., Wei, X., 2012. A review on slip models for gas microflows. Microfluid. Nanofluid. 13, 845–882.
Zhao, J., Yao, J., Zhang, L., Sui, H., Zhang, M., 2016. Pore-scale simulation of shale gas production considering the adsorption effect. Int. J. Heat Mass Transf. 103, 1098–1107.
Zhou, H., Tchelepi, H.A., 2008. Operator-based multiscale method for compressible flow. SPE J. 13, 267–273.
Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., Research, B.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks monet photos. In: Conference on Computer Vision (ICCV), 2017, pp. 2223–2232.