Review of collapse triggering mechanism of collapsible soils due to wetting

Ping Li a, b, Sai Vanapalli a, *, Tonglu Li b

a Department of Civil Engineering, University of Ottawa, Ottawa, ON, K1N 6N5, Canada
b School of Geological Engineering and Geomatics, Chang’an University, Xi’an, Shaanxi, 710054, China

Abstract

Loess soil deposits are widely distributed in arid and semi-arid regions and constitute about 10% of the total land area of the world. These soils typically have a loose honeycomb-type meta-stable structure that is susceptible to a large reduction in total volume or collapse upon wetting. Collapse characteristics contribute to various problems to infrastructures that are constructed on loess soils. For this reason, collapse triggering mechanism for loess soils has been of significant interest for researchers and practitioners all over the world. This paper aims at providing a state-of-the-art review on collapse mechanism with special reference to loess soil deposits. The collapse mechanism studies are summarized under three different categories, i.e. traditional approaches, microstructure approach, and soil mechanics-based approaches. The traditional and microstructure approaches for interpreting the collapse behavior are comprehensively summarized and critically reviewed based on the experimental results from the literature. The soil mechanics-based approaches proposed based on the experimental results of both compacted soils and natural loess soils are reviewed highlighting their strengths and limitations for estimating the collapse behavior. Simpler soil mechanics-based approaches with less parameters or parameters that are easy-to-determine from conventional tests are suggested for future research to better understand the collapse behavior of natural loess soils. Such studies would be more valuable for use in conventional geotechnical engineering practice applications.

1. Introduction

Loess soils are widely distributed and constitute about 10% of the total land area of the world. Several countries including China, Russia, United States, France, Germany, New Zealand and Argentina, have a large area of loess soil deposits (Phien-wej et al., 1992; Rogers et al., 1994; Rogers, 1995; Al-Rawas, 2000; Nouaouria et al., 2008; Ryashchenko et al., 2008; Gaaver, 2012). These soils are typically formed with a loose honeycomb-type meta-stable structure and are susceptible to a sudden decrease in total volume or collapse upon wetting (Feda, 1988; Houston et al., 1988; Lommler and Bandini, 2015). Different types of natural soils may develop a collapsible fabric provided there is an open, potentially meta-stable, partly saturated structure, and a high enough applied stress (Barden et al., 1973; Lawton et al., 1989). In addition, any type of soil compacted at dry of optimum condition is collapsible in nature (Fredlund and Gan, 1995; Kato and Kawai, 2000; Pereira and Fredlund, 2000). Collapse and other collapse associated problems, such as differential settlement, earth cracks, landslides and falls, have contributed to serious damages to the infrastructures that are constructed on loess soils, including loss of human lives in certain scenarios (Derbyshire, 2001; Houston et al., 2001; Peng et al., 2006). During the period of 1974–1975, in China, it was reported that a total of 1505 buildings were damaged and 80 km-long underground pipeline ruptured due to collapse of loess soils (Sun et al., 2013). Fu (2005) reported that typically a quarter of time is required with respect to the entire construction period to treat the collapsible soils prior to placing foundation. Also, the cost associated with collapsible ground preparation is typically about one third of the total cost of the infrastructure. With urban development on the rise in various regions of the world with collapsible soil deposits, there will be more access to water for these soils. As a result, there will be more wetting associated collapse problems. For this reason, it is important to understand the collapse mechanism of these soils.

During the last six decades, many researchers have focused on studying the collapse mechanism of various collapsible soils upon wetting. The discussions on this topic are summarized under three
categories, i.e. traditional approaches, microstructure approach, and soil mechanics-based approaches. Among the traditional approaches, collapsibility was always interpreted considering one sole factor. For example, collapse was attributed to loss of capillary tension or solution of soluble salts (Guo, 1958; Dudley, 1970). In addition, many researchers studied the influence of soil properties, such as the density, clay content, Atterberg limits and grain size distribution, on the collapsibility of loess soils (Sun, 1957; Liu, 1994; Zhao and Chen, 1994; Zhang, 2002; Fan and Guo, 2003; Song and Wang, 2004; Chen et al., 2006; Yuan, 2009), and a large number of empirical equations have been proposed in the literature relating the collapsibility to conventional soil properties (Cleveger, 1958; Feda, 1964, 1966, 1995; Handy, 1973; Basma and Tuncer, 1992, Fan and Guo, 2003; Song and Wang, 2004; Ayadat and Hanna, 2007; Zorlu and Kasapoglu, 2009; Noor et al., 2013). However, most of these empirical equations can be extended only for local soils for which they have been developed. In other words, the proposed empirical equations are not universally valid for use in conventional geotechnical engineering practice (Sun, 1957; Gao, 1979, 1990; Yang, 1988). This is due to soils exhibiting different forms of microstructure and hence presenting different mechanical behaviors in spite of having similar physical properties. For example, a loess soil with few large pores may have a same structural void ratio or density as another loess soil with many small pores; however, their collapse behaviors will be significantly different.

The collapse nature of loess and other soils is interpreted using the information of soil microstructure as a tool (Derbyshire and Mellors, 1988; Delage et al., 1996; Romero and Simms, 2008). Loess soil microstructure can be analyzed in terms of four key factors based on the comprehensive studies on microstructure of loess soils, i.e. particle pattern, contact relation, pore form, and bonding material. These four factors are dependent on each other; however, the pore form and bonding material are suggested as the two dominant factors that have more influence on the collapse behavior (Gao, 1980a, b; Lei, 1983, 1987; Yang, 1988; Zhao and Chen, 1994). Many researchers have been involved in classifying soil pores and distinguishing the water stability of various possible bonding materials, as well as exploring their respective influence on the collapsibility of loess (Lei, 1983, 1987; Derbyshire and Mellors, 1988; Yang, 1988; Zhao and Chen, 1994; Osipov and Sokolov, 1995; Assalay et al., 1997; Smalley et al., 2001; Jiang et al., 2014a, b; Lommler and Bandini, 2015). It is widely accepted that microstructure plays a key role in controlling the collapse behavior; however, it lacks a simple quantitative descriptor for estimating collapse deformations (Alonso et al., 1993). This limitation has been addressed and alleviated to a great extent in recent years by extending image processing techniques which facilitate the quantitative analysis of soil microstructure (e.g. Chen and Sha, 2009a, b; Gu et al., 2011; Fang et al., 2013a, b). These image processing programs are mostly based on the principle of binary grey segmentation to provide reasonably satisfactory results, particularly, for coarse soils. However, their application to fine-grained silts and clays has not been well validated. Gu et al. (2011) indicated that it is difficult to clearly distinguish the grains and pore sizes from digital images using the presently available programs.

The third category approaches use the concepts of soil mechanics for explaining the collapse behavior. Collapsible soils are typically unsaturated and significant collapse generally occurs prior to reaching fully saturated condition (El-Ehwany and Houston, 1989). For this reason, the concepts of mechanics for unsaturated soils are more rational for interpreting the collapse behavior (Tadepalli and Fredlund, 1991; Fredlund and Rahardjo, 1993; Fredlund and Gan, 1995; Habibagahi and Mokhberi, 1998; Chen, 1999; Pereira and Fredlund, 2000). The two stress state variables approach proposed by Fredlund and Morgenstern (1977) for describing the mechanical behavior of unsaturated soils has been extended to studies of volume change behavior due to loading and wetting of collapsible soils. Collapse is attributed to the loss of strength associated with suction decrease as a result of wetting (Fredlund and Gan, 1995). For this reason, collapse behavior has been widely investigated using suction-controlled wetting tests (e.g. Chen, 1999; Chen et al., 1999; Sun et al., 2004, 2007a, b; Jotisankasa, 2005), based on which, various models have been proposed for modeling the collapse behavior with respect to varying stress state variable (i.e. matric suction) (Tadepalli et al., 1992; Fredlund and Gan, 1995; Pereira and Fredlund, 1997, 2000). During the last quarter century, a number of elastoplastic models have been developed for modeling the behavior of unsaturated soils (Alonso et al., 1990; Josa et al., 1992; Gens and Alonso, 1992; Wheeler and Sivakumar, 1995; Cui and Delage, 1996; Wheeler, 1996; Wheeler et al., 2003). These models have been extended or modified to interpret the volume change behavior of collapsible soils (e.g. Chen et al., 1999; Kato and Kawai, 2000; Sun et al., 2004, 2007a, b). In these models, collapse is explained as a part of deformation when the stress path crosses the elastic region or yield surfaces. For quantitative analysis, elastoplastic models provide a more precise way to estimate the collapse deformations with respect to varying stress state variables. However, they are too complex as a large number of parameters are required to be determined from cumbersome suction-controlled tests. On the other hand, for addressing some scenarios of collapse, only a few model parameters are required. However, the parameters will be different for the same soils when their initial conditions, such as the initial water content, and the stress state, are different. In other words, when using these models, for each soil sample with a different initial condition, one suction-controlled test should be conducted, from which the model parameters can be determined. In addition, most of these models were proposed based on the experimental results of compacted soils, which may show different collapse behavior from the natural loess soils. For this reason, soil mechanics-based approaches proposed especially for natural loess soils are reviewed in this paper. Loss of structural strength is considered to be a key factor for natural loess soil collapse. The structural strength is mainly influenced by the factors associated with loess soil deposition and initial water content (Hu et al., 2000, 2004). Two models based on the concepts of breakage mechanics proposed by Shen and his group (Shen, 1993, 2003; Shen and Deng, 2003; Shen and Hu, 2003) for modeling the collapse behavior of natural loess soils due to loading and wetting were reviewed. These models have the same limitation as the models described earlier; they are complex, and cumbersome tests are required for determining the parameters, which restricts the application of these models in conventional geotechnical engineering practice. For these reasons, relatively simple models with less parameters or parameters that are easy-to-determine from conventional tests are suggested for advancing research related to collapse behavior of natural loess soils.

The comprehensive summary provided in this paper can be useful for addressing problems associated with collapsible soils and for proposing more efficient approaches in the future.

2. Traditional approaches for interpreting the wetting-induced collapse mechanism

2.1. Loss of capillary tension

Unsaturated soils exhibit apparent strength due to the capillary tension acting along the tangent of the menisci (Terzaghi, 1943). Water and clays suspending in the water retreat towards silt contacts as the soil desiccates, which produces considerably high

capillary tension to maintain the soil structure under unsaturated condition (Fig. 1). However, addition of water destroys the capillary tension, which likely contributes to reduction of soil strength and triggers collapse. Capillary tension concept has been used to explain the collapsibility of loess soils (Alfi, 1984), which however was questioned because of limitations in explaining the phenomenon of identical loess soil samples showing different collapse behaviors when they are treated with different liquids, such as water or saline solutions (Guo, 1958).

2.2. Solution of soluble salts

Crystal or microcrystalline soluble salts are widely considered to act as bonds in unsaturated loess soils. These salts are expected to dissolve as soil water content increases, contributing to a decrease in bonding strength and leading to collapse (Guan, 1986; Houston et al., 1988). This approach however was questioned because soluble salts are found dissolved in pore water even when the water content is as low as 8% (Tan, 1988). In addition, the soluble salts take up typically less than 0.1% in loess soils. Table 1 summarizes various soluble salts contents in loess soils in China (Yang, 1988). Guan (1989) suggested that self-weight collapsible loess soils have a natural water content of 4%–5%, while non-self-weight collapsible loess soils have a value of 21%–24%, which indicates that wetting-induced collapse could not be solely attributed to the solution of soluble salts, especially for non-self-weight collapsible loess soils.

2.3. Shortage of clays

Several researchers described collapse arises in some soils as a result of shortage of clay size fraction (for example, with a clay content less than 10%) (Gao, 1979; Rogers et al., 1994). If a soil has enough clays, it may expand instead of collapse on soaking. Krajew (1969) suggested that limited montmorillonite presence would swell and loosen the primary loose soil structure on soaking, which contributes to collapse. However, Guan (1989) suggested that self-weight collapsible loess soils have a clay content of 23%–39%, and non-self-weight collapsible loess soils have a value of 30%–42%. These loess soils show serious collapsibility upon wetting in spite of their high clay content. In addition, some researchers reported that a monotonic relationship does not exist between the collapsibility of loess soil and its clay content. The maximum collapse occurs at an intermediate rather than the minimum clay content (Hu et al., 1999; Li et al., 2004; Song and Wang, 2004). For this reason, clay content should not be considered as the key factor that triggers the wetting-induced collapse for loess soils.

2.4. Under compaction

Under compaction attributes the collapse nature to the characteristic of porous fabric of loess (Houston et al., 1988; Tadepalli and Fredlund, 1991). However, pores in loess soils having different sizes and forms would have completely different influences on the collapsibility. The size and form of pores are controlled by the factors associated with loess soil deposition. Various classifications of loess soil pores taking account of their size and form and their respective effect on the collapsibility are discussed in detail in later sections.

The traditional or conventional approaches listed above were found to be unsatisfactory for universally explaining the collapse behavior of loess soils. However, these approaches have significantly contributed towards better understanding the collapsible loess soils.

3. Microstructure approach for interpreting wetting-induced collapse behavior

Numerous researchers have reported that any soil compacted at dry of optimum condition typically exhibits collapse behavior upon wetting, while soils compacted at wet of optimum condition show no collapsibility (Barden, 1965; Lawton et al., 1989, 1992; Tadepalli and Fredlund, 1991). Collapse is therefore attributed to the metastable nature of the open flocculated structure associated with dry of optimum water content condition. This understanding contributed to interpreting loess soil collapsibility from the microstructure information. Recent advances in techniques and methods, including scanning electronic microscope (SEM), X-ray diffraction and mercury intrusion porosimetry (MIP), have been found to be valuable for understanding the microstructure and its significance on soil mechanical behavior. Several investigators provided reviews on the techniques and methods specifically aimed at characterizing the microstructure of unsaturated soils, among which SEM and MIP have been frequently used (Al-Mukhtar et al., 1996; Simms and Yanful, 2001, 2002; Romero and Simms, 2008). In addition, several image processing techniques have been developed, such as the Scion (Fang et al., 2013a, b), Videolab (Shi et al., 1995) and Mirage (Sha and Chen, 2006; Chen and Sha, 2009a, b), and simple commercial image processing programs such as MATLAB (Wang et al., 2011; Mao et al., 2004) and Photoshop (Zhang et al., 2009). These programs extract quantitative information about soil pores or grains from the digital images, providing strong support for interpreting collapse behavior from microstructure information.

The microstructure of loess soils is analyzed in terms of four factors, i.e. (i) particle pattern, (ii) contact relation, (iii) pore form, and (iv) bonding material.

3.1. Particle pattern

It is widely agreed that loess soils are composed of particles typically in the range of 10–60 μm, of which particles larger than 20 μm account for about 75% of the total solid constituents (Tan, 1988; Dijkstra et al., 1994). During the loess deposition, primary grains may be coated and cemented by clay platelets or calcium carbonate to form aggregates or coagulum (i.e. collection of aggregates). Together with primary grains, they all were regarded as skeleton particles of loess soil structure. Fig. 2 shows the development of loess soil structure from single grain to an element of loess soil structure (Gao, 1980a, b, 1981). All these particle patterns are also illustrated using SEM images (see Fig. 3) which are from natural Malan loess and Lishi loess, respectively, from Lanzhou, China. Both primary grains and aggregates are generally
found in a loess soil. Primary grain dominant soils are more likely to have higher collapse possibility since they have a more open structure compared to aggregate dominant soils. In latter type soils, interspaces are well filled with fine grains, these soils therefore have a more stable structure than the former ones (Gao, 1979, 1980a). In general, the more the aggregates, the weaker the collapsibility of loess would be. The particle pattern in loess soils varies with climate factors (Liu, 1978; Beckwith and Hansen, 1982; Lin and Liang, 1982; Gao, 1984). Typically, the effects of weathering, clayization and calcium carbonate microcrystallization in wet and warm environment are more pronounced compared to those in dry and cold environment. Due to this reason, more small size grains, clays and microcrystalline calcium carbonates are prone to forming in wet and warm areas, which are favorable for aggregates development. In addition, the longer time is available for loess deposition, the formation of more aggregates is likely (see Fig. 3).

On the other hand, particles in loess soils could be classified based on their size. Lei (1983, 1987) suggested grains ranging from 5 μm to 250 μm as skeleton particles, particles larger than 10 μm as coarse grains and those smaller than 10 μm as fine grains. Yang (1988) suggested particles larger than 10 μm, which account for more than 70% of the total solid mass, form the skeleton of loess soils, particles smaller than 2 μm act as bonding materials, those between 2 μm and 10 μm are typically filling materials. Yuan et al. (2007) divided loess soil particles into sand (60 μm < d < 2000 μm), silt (2 μm < d < 60 μm) and clay (d < 2 μm). A loess soil with higher percentage of silts is more favorable to form large pores and shows high collapsibility (Derbyshire and Mellors, 1988; Assallay et al., 1997; Grabowska-Olszewska, 1975, 1988). Handy (1973) stated

### Table 1
Soluble salts contents in loess soils in China [Yang, 1988].

<table>
<thead>
<tr>
<th>Location</th>
<th>NaHCO3</th>
<th>Na2SO4</th>
<th>10H2O</th>
<th>KCl</th>
<th>NaCl</th>
<th>MgSO4·7H2O</th>
<th>MgCl2</th>
<th>Total amount (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zanhuang, Hebei</td>
<td>0.001</td>
<td>0.005</td>
<td>0.001</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.008</td>
</tr>
<tr>
<td>Taiyuan, Shanxi</td>
<td>0</td>
<td>0.036</td>
<td>0.001</td>
<td>0</td>
<td>0.004</td>
<td>0</td>
<td>0.003</td>
<td>0.044</td>
</tr>
<tr>
<td>Linfen, Taiyuan</td>
<td>0</td>
<td>0.068</td>
<td>0.002</td>
<td>0.002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.072</td>
</tr>
<tr>
<td>Luochuan, Shaanxi</td>
<td>0.012</td>
<td>0.027</td>
<td>0.001</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.041</td>
</tr>
<tr>
<td>Tianshui, Gansu</td>
<td>0.003</td>
<td>0.038</td>
<td>0.001</td>
<td>0.004</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.046</td>
</tr>
<tr>
<td>Dingxi, Gansu</td>
<td>0</td>
<td>0.073</td>
<td>0.002</td>
<td>0</td>
<td>0.018</td>
<td>0</td>
<td>0.003</td>
<td>0.096</td>
</tr>
</tbody>
</table>

Fig. 2. Particle patterns in loess soils: (a) primary grain; (b) primary grain coated by clays or calcium carbonate; (c) aggregate; (d) coagulum (collection of aggregates); (e) microstructure element (modified after Gao, 1980a).

Fig. 3. The SEM photographs of (a) natural Malan loess and (b) natural Lishi loess in Lanzhou, China (modified after Dijkstra et al., 1994).
that collapsibility is inversely proportional to clay content for loess soils. However, several researchers stated that there is an optimum clay content corresponding to the maximum collapsibility for loess soils (Hu et al., 1999; Li et al., 2004; Song and Wang, 2004). This value varies depending on loess soil type and its initial water content.

Present understanding of particle pattern in loess soils is mostly based on particle size and origin; however, classifications were also proposed based on particle shape (including sphere, flake, tube, needle and floccules) (Krinsley and Smalley, 1973). Studies on particle shape and the influence of particle shape on soil structure stability are more meaningful for granular soils. It is however difficult to identify the shape of aggregates or coagulum in loess soils. Particle shape is also a significant indicator of the mechanical behavior of clays which are mostly fillings or bonds in loess soils; however, loess particle shape has not been comprehensively investigated (Dijkstra et al., 1995).

### 3.2. Contact relation

For most loess soils, significant collapse starts to occur at a critical water content, which varies depending on the initial water content and the vertical stress under which the soil is wetted. As the soil water content increases beyond the critical value, bonding strength fails to support the soil fabric continuously (Zheng and Zhang, 1989). Subsequently, soil structure fails and a fundamental particle rearrangement takes place, transforming an initial stable open structure into a remolded close structure, which is considered as the collapse process (Dijkstra et al., 1995). Therefore, contact relation is an important indicator for determining whether or not it is easy for particles to slip over each other. Particles typically have two main contact relations in loess soils depending on the amount of bonds at the contact, i.e. point-contact and face-cementation relation, as illustrated in Fig. 3 (Gao, 1980a, 1981). Particles in point-contact relation are mostly naked and there are only a few bonds at the contact. On the contrary, particles in face-cementation relation are typically thickly coated, and there are large amount of bonds at the contact, making it look like face-to-face contact relation.

Conditions are favorable for maintenance of point-contact relation in dry cold areas like northwest China for several reasons: (i) weak weathering makes it unfavorable for grains to be compressed; (ii) a little precipitation makes it difficult for calcium carbonate microcrystallization and clayization. Conversely, in wet and warm areas, frequent circles of precipitation and evaporation as well as active biological chemical action result in sufficient microcrystalline calcium carbonate, clays and small size grains. Generally, two contact relations coexist in loess with different proportions, the variation in terms of contact relation is consistent with climate variation (Gao, 1980a, 1981; Guan, 1989). In addition, point-contact relation is more likely to be in grain dominant soils, while aggregate dominant soils have more face-cementation relation.

Contact relation in loess soils was also classified into overhead, interlocking and dispersed contact (Lei, 1983, 1987). These kinds of contact relations were defined taking into account the pore form and bonds state. Overhead contact is for surrounding particles of overhead pores; interlocking contact is a result of interlocking particle arrangement, which is commonly observed in any kind of soil; dispersed contact is the scenario that soil particles are finely divided by bonds.

A kind of contact relation “clay-bridge” was found in loess soils (Derbyshire and Mellors, 1988). The “ideal bridge” was found to comprise clay particles peeling away from the surface of one grain linking up with similarly arranged clays on the surface of an adjacent grain. Clay-bridge was found to vary in thickness from site to site depending on the degree of clayization and climate condition.

### 3.3. Pore form

Different particle patterns and contact relations result in different pore forms. Pores in loess soils were divided into macro-pores, spaced pores, intergranular pores and intragranular pores from the pore size with respect to surrounding particles by Gao (1980a, 1981), as shown in the sketches in Fig. 4. All forms in loess soil except for macro-pores are illustrated in a SEM image modified after Derbyshire (2001) (Fig. 5), which is from natural Malan loess in Lanzhou region in China. Macropores usually are root holes, wormholes and holes made by rodents' action. These pores are commonly found at a relatively shallow depth in collapsible loess soils (i.e. within 1.0 m approximately, see in Lei, 1983), and their walls are heavily covered by reborn calcium carbonates in most cases. For these reasons, macropores are believed to contribute little to overall collapse or particle rearrangement because of their low proportion in spite of individual macropore having a large volume (Gao, 1980a; Lei, 1983).

Spaced pores are typically associated with spaced arrangement of aeolian deposits. They are also characterized by the size larger than surrounding particles which are poorly cemented and more likely in point-contact relation (Yang, 1988). Spaced pores contribute to favorable spatial conditions for collapse to occur. In addition, spaced pores contribute to a critical state condition as the bonding strength equals the gravitational potential of overhanging soil particles. Under such a scenario, when the soil gets wet and bonding strength weakens, overhanging particles fall into spaced pores and soil collapses. In many scenarios, intergranular pores formed by interlocking particle arrangement are observed in any kind of soil, including expansive soils (Popescu, 1985; Gins and Alonso, 1992; Saba et al., 2014). Intergranular pores have been experimentally proved to contribute slightly to loess soil collapsibility (Reznik, 2007). Typically, all kinds of pores coexist in a loess soil except for macropores. In arid environment, spaced pores develop since the poorly-coated grains are prone to having higher hardness for transferring forces, which effectively prevents the soil from being compressed under dead-weight (Gao, 1980a, 1981). However, loess soils in humid regions that lie deep beneath the surface predominantly have intergranular and intragranular pores. Compared with Gao’s (1980a, 1981) classification, Lei (1983, 1987) classified loess soil pores into original pores and secondary pores. Original pores include intergranular pores (i.e. overhead pores and interlocking pores) and intragranular pores (i.e. pores within the aggregates). Overhead pores, interlocking pores and intragranular pores correspond to three different soil structures, i.e. overhead structure, interlocking structure and matrix structure, as shown in Figs. 5 and 6. The secondary pores include root holes, wormholes, rodent holes, joints, fissures and Karst caves. The characteristics of all kinds of pores, as well as their influences on the collapse behavior are summarized in Table 2.
than 20 µm. The pore size density function (PSDF) shows two peaks at pore diameters of about 7 µm and 0.025 µm, respectively. This dual characteristic of the PSDF indicates that there are two main pore groups in the soil, i.e., inter-aggregate type and intra-aggregate type (similar to intergranular and intragranular pores defined by several investigators (Gao, 1980a; Al-Mukhtar et al., 1996)). MIP was also used for the microstructure investigation of both natural and compacted loess soils during the stress path tests by Jiang et al. (2012a, 2014a). Changes in the PSDs due to stress path tests show that loading leads to a significant change in inter-aggregate pores while a slight change in intra-aggregate pores, for both natural and compacted loess soils. Recently, Lommler and Bandini (2015) studied a collapsible soil in New Mexico, USA that has “pin holes” of tubular structure of 0.3 mm diameter visible to the naked eye. These pin holes were recognized as indicator of the collapsibility of this soil. However, increasing magnification revealed the presence of honeycomb structure around the pin holes. The honeycomb structure and highly porous matrix provide a more reasonable explanation for the high collapsibility of local soil. It is also indicated that macropores contribute little to collapse. Zhao and Chen (1994) and Zhao et al. (1997) stated that early classifications for

### Table 2

Characteristics of all kinds of pores in loess soils (modified after Lei, 1987).

<table>
<thead>
<tr>
<th>Type</th>
<th>Radius (µm)</th>
<th>Characteristics</th>
<th>Influence on collapsibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary Root holes</td>
<td>&gt;16 (large pores)</td>
<td>Dense walls</td>
<td>Depend on whether the pore walls are cemented or not</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Loose walls</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Large, fragile</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Closed, with rough edges</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Closed, radial, irregular, with smooth walls</td>
<td></td>
</tr>
<tr>
<td>Wormholes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse holes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Joints and fissures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Karst caves</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Intergranular pores</td>
<td>4–16 (mediate pores)</td>
<td>Spaced pores as per Gao (1980a)</td>
<td>Significant</td>
</tr>
<tr>
<td>Interlocking pores</td>
<td>1–4 (small pores)</td>
<td>Intergranular pores as per Gao (1980a)</td>
<td>Insignificant</td>
</tr>
<tr>
<td>Intragranular pores</td>
<td>&lt;1 (micropores)</td>
<td>Numerous, irregular distribution</td>
<td>Extremely insignificant</td>
</tr>
</tbody>
</table>

loess soil pores mislead spaced pores with overhead pores. Overhead pores are different from spaced pores not only in the size but also in the influence on collapsibility. Both spaced and overhead pores are formed by abnormal arrangement of coarse particles with point-contact relation. Spaced pores are typically larger than surrounding particles (i.e. >20 μm). Spaced pores ranging from 20 μm to 80 μm contribute to heavy collapse, while overhead pores ranging from 8 μm to 20 μm contribute to medium or weak collapse. As soil collapses, spaced pores may convert into overhead pores. Due to this reason, loess soils have been found to collapse more than once. This is consistent with the results reported by Yang (1988). In his study, overhead pores and most spaced pores fall in the group of pores smaller than 54 μm. Table 3 shows the characteristics of each kind of pores and the corresponding collapsibility level (modified after Zhao and Chen, 1994). Assalay et al. (1997) divided soil pores into small, medium, large and giant pores, among which only large pores make a contribution to wetting-induced collapse. Classifications of soil pores as summarized in this paper are more fundamental in nature and well defined for rational explanation of collapse behavior of loess soil deposits. These classifications suggest that pores with different sizes and forms have different influences on the collapsibility. For this reason, void ratio, which is not able to represent pore size and form sufficiently, cannot be used as an index for estimating the collapsibility of loess soils. For example, the same soils when compacted at dry and wet of optimum conditions have the same dry density but different collapse behaviors. Extending the same argument, two soils having the same dry density or void ratio may have different collapsibility, structures, such as flocculated and dispersed structures (Vanapalli et al., 1999). This behavior also explains why the relationship between collapsibility parameters and void ratio derived for collapsible loess soil in one region provides unsatisfactory results when used for prediction of the collapsibility of loess soil in a different region.

From the above discussions it is clear that large pores or interaggregate pores (i.e. spaced pores and overhead pores) have a major influence on collapse behavior of loess soils. Information of changes in PSDs due to loading and wetting provides better evidence that different kinds of pores contribute differently to collapse (Yang, 1988; Jiang et al., 2012a, 2014a, b). More studies in this direction would be useful for better understanding of the loess soil microstructure and influence of soil pores on collapsibility.

### Table 3

<table>
<thead>
<tr>
<th>Collapsibility level</th>
<th>Dominant pore type</th>
<th>Collapsibility coefficient</th>
<th>Changes in pores after collapse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-collapse</td>
<td>Overhead and interlocking pores</td>
<td>&lt;0.015</td>
<td>Not obvious</td>
</tr>
<tr>
<td>Weak collapse</td>
<td>Overhead pores</td>
<td>0.015–0.04</td>
<td>Not obvious</td>
</tr>
<tr>
<td>Medium collapse</td>
<td>Overhead pores</td>
<td>0.04–0.07</td>
<td>Obvious decrease in overhead pores and spaced pores</td>
</tr>
<tr>
<td>Strong collapse</td>
<td>Overhead pores and spaced pores</td>
<td>0.07–0.1</td>
<td>Obvious decrease in spaced pores and part of overhead pores</td>
</tr>
<tr>
<td>Extremely strong collapse</td>
<td>Spaced pores</td>
<td>&gt;0.1</td>
<td>Obvious decrease in spaced pores</td>
</tr>
</tbody>
</table>

3.4. Bonding materials

Bonding strength in loess soils can be attributed to several sources, such as capillary tension, surface friction and bonding materials. The influence of bonding materials, which include calcium carbonate and clays in most cases, is described in this section. The other bonding strength sources, such as capillary tension and friction, are dependent on the minerals and modes of formation of a soil (i.e. natural deposition or compaction). Calcium carbonate is regarded as one of the key bonding materials in loess soils since it is found not only as film coating on grains but also because of its concentration at grain contacts. More importantly, loess soils are always characterized by high carbonate content. Recently, loess soils were mixed with calcium oxide (i.e. CaO) of different percentages, and then water and carbon dioxide (i.e. CO2) were added, in order to prepare artificial structural loess soil samples for study of natural loess soil mechanical behavior (Hu et al., 2000, 2004; Jiang et al., 2012b). Table 4 shows the information of carbonate content for loess soils from different regions in China (Derbyshire and Mellors, 1988).

Gao (1980a, 1990, 1994) stated that different calcium carbonate states (i.e. granular, re-born crystalline and microcrystalline calcium carbonate) have different effects. However, only microcrystalline calcium carbonate contributes to bonding strength and affects the collapsibility of loess soils. For this reason, the calcium carbonate content cannot be regarded as an index to estimate collapse behavior. Cabrera and Smalley (1973) proposed the concepts of long- and short-range bonding. Calcium carbonate bonding may be either long- or short-range bonding depending on the amount of calcium carbonate bonds at contact. It is the short-range bonding that has a significant effect on the collapsibility of loess. Guo (1958) found that loess soils wetted by hot water, NaCl and CaCl2 solution respectively show a gradually decreasing collapsibility. It was suggested in his study that salt dissolution weakens bonding strength and induces collapse upon wetting by regarding calcium

![Fig. 7. The MIP results of natural loess soil in Shaanxi, China (modified after Jiang et al., 2014a).](image-url)
carbonate as soluble and Ca\(^{2+}\) as the main ion in pore water. It is unreasonable to regard calcium carbonate as soluble (Smalley, 1971; Guo et al., 2004, 2005, 2008) and to attribute collapse to calcium carbonate leaching on soaking. Calcium carbonate leaching is certainly a slow process that requires many cycles of wetting and drying, which cannot explain the collapse phenomenon of sudden decrease in volume upon wetting. This reasoning implies that there are other non-water-stable materials contributing to bonding strength in unsaturated condition except for calcium carbonate.

Clay minerals (i.e. < 2 \(\mu m\)) have been regarded as dominant bonding agents because of their electrically charged characteristic and flocculated or flowing-gelatinous form (Barden et al., 1973; Osipov and Sokolov, 1995; Assallay et al., 1997; Smalley et al., 2001). Capillary tension and electrostatic attraction can also be attributed to the presence of clay minerals. Considerable amount of clays is typically found in loess soils from China (Derbyshire and Mellors, 1988) (Fig. 8). The clay grade minerals were also found to have three major forms in loess soils, i.e. (i) coatings, (ii) clay bridge as described above, and (iii) buttresses (Derbyshire and Mellors, 1988; Derbyshire, 2001). The clay coatings are commonly observed lying parallel to the surface of silts. These coatings were neither continuous nor of uniform thicknesses. Clay buttresses comprising silt-size clay aggregations were observed between quartz grains (Derbyshire and Mellors, 1988). Similarly, Barden et al. (1973) found that the form of clays in collapsible loess soils varies depending on the geologic origins and stress history experienced by the soil. If the clays originally suspend in pore water, then gradual evaporation would cause the retreat of clay plates with pore water to the contacts, giving a buttress type of support to silts. There are many possible forms of clay materials in loess soils, the nature of clay bonding is complex, and it is not clear how much is due to electro-chemical effect and due to capillary effect, respectively (Barden et al., 1973).

Yang (1988) divided bonding materials into water-stable and non-water-stable materials. The former includes calcium carbonate, calcium sulfate, iron oxide, iron hydroxide, water-stable secondary mica and secondary zeolite, and the latter includes soluble salts, clay materials and secondary mica. Since contents of both soluble salts and secondary mica are typically less than 0.1% in loess soils, clay bonding has a significant contribution to collapse upon wetting. However, if clay grade minerals are the only bonding materials in loess soils, collapse deformation upon wetting should exhibit a degree of plasticity instead of a sudden failure. In summary, calcium carbonate and clay minerals should be realized as the key bonding materials that control the collapsibility of loess soils.

As per the microstructure information summarized, collapse behavior of loess soils is influenced by all four structural factors (i.e. particle pattern, contact relation, pore form, and bonding materials). All the four factors are dependent on each other; for example, spaced pores and point-contact relation always reinforce each other. The pore form and bonding material could be suggested as the two main important factors that have more influence on the collapse behavior of loess soils compared to the other two factors as per the discussions summarized in the paper. However, the comprehensive behavior of a collapsible loess soil cannot be fundamentally characterized considering one single factor. Further investigations on all the four structural factors are required for rational interpretation of collapse behavior of loess soils.

4. Soil mechanics-based approaches for interpreting the collapse behavior

Microstructure is widely acknowledged as important information required in explaining the collapse mechanism. However, as discussed earlier, it lacks a simple quantitative descriptor, which makes the analysis more subjective (Alonso et al., 1993; Chen et al., 2006; Tian et al., 2011). This has been addressed by the recently developed image processing programs, which facilitate the qualitative analysis of soil microstructure (Gu et al., 2011). The image processing programs are mostly based on the principle of binary grey segmentation and may provide satisfactory results for granular soils. Their application for fine-grained soils such as silts and clays has not been well validated.

The measurable mechanical behaviors which include shear strength and volume change can be described in terms of the stress state in the soil for both saturated and unsaturated soils. The effective stress equation proposed by Terzaghi (1943) is widely used for interpretation of saturated soils. The same equation has been extended by Bishop and Blight (1959), Bishop (1960) for unsaturated soils. Fredlund and Morgenstern (1977) introduced two stress state variables, i.e. net normal stress and matric suction, for rational interpretation of unsaturated soils. The volume change of collapsible soils has been attributed to changes in both stress state variables (Alonso et al., 1990; Chen, 1999; Chen et al., 1999; Sun et al., 2004, 2007a, b).

It is always believed that collapsible soils experience the maximum settlement as their degree of saturation approaches 100%. For this reason, the collapse settlement at saturated condition of soil is of concern and interest for geotechnical engineers. Several testing methods have been suggested in the literature for estimating collapsibility such as single- and double-oedometer tests (Jennings and Knight, 1957; Knight, 1963). However, in general, collapsible soils undergo significant volume change prior to reaching saturated state. El-Ehwany and Houston (1989) performed a one-dimensional (1D) infiltration test by applying water with a small head at one end of the sample. Their test results suggest that for degrees of saturation above 65%–70%, essentially full collapse occurs for the tested soil. The relationship between degree of saturation and ratio of partial collapse to full collapse is shown in Fig. 9. This result is reasonable because even the soils within the wetted zone do not reach fully saturated condition (El-Ehwany and Houston, 1989).

Fig. 8. Grain size distributions of loess soil samples collected from different regions in China (modified after Derbyshire and Mellors, 1988).
Fig. 9. Partial collapse due to partial wetting (modified after El-Ehwany and Houston, 1989).

4.1. Understanding collapse behavior from the two stress state variables approach

Several early researchers extended effective stress principle to interpret the mechanical behavior of unsaturated soils (Aitchison, 1960; Bishop, 1960; Jennings, 1961; Newland, 1965; Richards, 1966; Allam and Sridharan, 1987; Khalili and Khabbaz, 1998). However, studies, notable as Jennings and Richards, 1966; Allam and Sridharan, 1987; Khalili and Khabbaz, 1998), have shown that a single-valued effective stress is not capable of completely describing unsaturated volumetric behavior. The two stress state variables approach proposed by Fredlund and Morgenstern (1977) provides a more meaningful description because it is applicable to both shear strength and volume change behavior of unsaturated soils and does not use other empirical properties or parameters in the description of the soil stress state. From the perspective of two stress state variables approach, the triggering mechanism for collapse is attributed to the loss of strength due to reduction in matric suction as a result of wetting. In other words, collapse occurs when there is a change in stress state of the soil as it goes from unsaturated condition towards a saturated condition (Fredlund and Gan, 1995).

The stress and deformation state variables can be combined using suitable constitutive relations proposed by Fredlund and Morgenstern (1977) for the soil structure, air phase and water phase. Generally, the constitutive relations for the soil structure and water phase are used in volume change behavior analysis. The proposed constitutive relations can be expressed in a compressibility form, which is consistent with the soil mechanics principles (i.e. Eqs. (1) and (2)). They can also be presented graphically by plotting the deformation state variables with respect to two independent stress state variables in the form of constitutive surfaces (i.e. the three-dimensional (3D) constitutive surfaces for an unsaturated soil, as shown in Fig. 10). The constitutive surfaces provide rational explanation of volume change behavior of unsaturated soils in terms of two stress state variables, which is consistent with the continuum mechanics principles. In addition, they also provide a way to estimate the volume change due to varying stress state variables. For an isotropic, linear elastic material, the coefficients of volume change (i.e. $m_1$ and $m_2$, $m'_1$ and $m'_2$) under various loading conditions could be quickly calculated using the elasticity modulus of the material. However, the estimation of these coefficients is much more complex as the soil yields (Alonso et al., 1990). In addition, these test results also are influenced by soil type, initial soil properties and the soil stress state (Fredlund and Gan, 1995). Therefore, simple procedures for estimating these coefficients from conventional tests would be valuable for using such constitutive relationships in modeling the collapse behavior.

$$d\varepsilon_v = m_1^v d(\sigma_{\text{mean}} - u_a) + m_2^v d(u_a - u_w)$$

$$d(\theta) = m_1^w d(\sigma_{\text{mean}} - u_a) + m_2^w d(u_a - u_w)$$

where $\varepsilon_v$ is the volumetric strain; $\theta$ is the volumetric water content; $m_1$ and $m_2$ are the coefficients of total volume change with respect to changes in net normal stress and matric suction, respectively; $m'_1$ and $m'_2$ are the coefficients of water volume change with respect to changes in net normal stress and matric suction, respectively; $\sigma_{\text{mean}} - u_a$ is the net mean normal stress; $u_a - u_w$ is the matric suction; $\sigma_{\text{mean}}$ is the mean normal stress; $u_a$ is the pore-air pressure; $u_w$ is the pore-water pressure.

Tadepalli and Fredlund (1991), Tadepalli et al. (1992), Fredlund and Gan (1995) conducted a series of tests on a statically compacted silt using a modified oedometer apparatus. Their results show a one-to-one relationship between matric suction and total volume change due to collapse, as shown in Fig. 11. S1M, S2M, S3M and S4M represent four specimens that were tested. Since volume changes are predominantly confined within the wetted zone, only

Fig. 10. 3D constitutive surfaces for an unsaturated soil: (a) soil structure constitutive surface; (b) water phase constitutive surface (modified after Fredlund and Rahardjo, 1993).
Based on the experimental observations, Pereira and Fredlund (1997, 2000) proposed a model for predicting the volume change due to collapse at a given net mean stress. Six curve-fitting parameters including $d_1$, $d_2$, $e_i$, $e_r$, $(u_a - u_w)_{i}$, $(u_a - u_w)_{r}$ were involved (see Fig. 12), where $d_1$ is the slope of volumetric deformation in the pre-collapse phase, $d_2$ is the slope of volumetric deformation in the collapse phase, $e_i$ is the initial void ratio (i.e. before the wetting), $e_r$ is the final void ratio (i.e. after the complete saturation), $(u_a - u_w)_{i}$ is the critical matric suction below which the soil structure starts collapsing, $(u_a - u_w)_{r}$ is the final matric suction below which the soil structure stops collapsing, and $e_t$ is the saturated volumetric water content. Although Pereira and Fredlund’s (1997, 2000) model provides a smooth void ratio state surface with respect to matric suction, the relationships between parameters of the model and net mean stress were mostly derived using the fitting method. This method may not be suitable for higher net mean stress since some assumptions are only valid for net mean stress values lower than 200 kPa. However, an insight into the shape of Pereira and Fredlund’s (2000) model indicates that the collapse behavior in terms of void ratio with respect to changes in matric suction is reverse to the SWCC, especially the wetting curve (see Fig. 12). The three phases of collapse, i.e. pre-collapse phase, collapse phase and post-collapse phase would correspond to the three zones on the wetting SWCC (i.e. residual zone, transition zone and capillary saturation zone) (after Vanapalli et al., 1996, 1999). For this reason, it is very likely that relationships can be found between the parameters in Pereira and Fredlund’s (2000) model and special values on the wetting SWCC. Such a modified model may have wider application in engineering practice.

4.2. Understanding collapse behavior from the elastoplastic modeling

The widely used model that was firstly presented in a quantitative form by Alonso et al. (1990) is referred to as the Barcelona basic model (BBM) in the literature. Since then, several modified forms of the BBM were proposed (Gens and Alonso, 1992; Josa et al., 1992; Wheeler and Sivakumar, 1995; Sun et al., 2003). More recently, some models were proposed in which the SWCC was incorporated into the stress-strain constitutive relationships to simulate the nonlinear variation of unsaturated soil properties (Gallipoli et al., 2003; Wheeler et al., 2003; Sun et al., 2007c; Thu et al., 2007a, b). The validity of all these elastoplastic models was tested using several suction-controlled compression tests conducted on unsaturated soils (Cui and Delage, 1996; Chen, 1999; Chen et al., 1999; Rampino et al., 2000; Sun et al., 2004, 2007a, b). The extensions or modifications of the BBM make it more convenient to apply the BBM for general problems associated with unsaturated soils (Sheng et al., 2008). In the BBM, a yield curve representing the locus of yield points in the ($p$, $s$) space was formulated which was labeled as LC (after loading-collapse) yield curve, where $p$ stands for either net normal or net mean stress depending on the type of test considered, and $s$ stands for suction. When the stress path crosses the yield curve, irreversible volumetric strain or collapse occurs. The irreversible volumetric strains may be due to suction decrease or to load increase; both suction decrease and load increase will have a similar effect on the moving LC yield curve. Such a behavior is consistent with the observations of Matyas and Radhakrishna (1968) who suggested that wetting has a two-fold effect on soil structure, i.e. a reduction in intergranular stress and a reduction in the rigidity. Collapse happens when the volume decrease due to rigidity decrease exceeds the volume increase due to intergranular stress decrease; however, the reverse situation results in swelling. The yield locus with respect to changes in suction is referred to as SI (after suction increase) yield

Fig. 12. Volume change behavior of collapsible soils during the wetting process (modified after Vanapalli et al., 1996 and Pereira and Fredlund, 2000).

Fig. 11. Variation of total volume with respect to matric suction for four different soils by Fredlund and Gan (1995).
Therefore, the complete framework for describing the volumetric behavior of an unsaturated soil should be represented by two yield curves LC and SI bounding the elastic region (see Fig. 13). The irreversible strains or collapse deformations due to crossing the SI or LC surface would indicate a change in the soil structure that affects the position of the LC or SI curves. Alonso et al.’s (1990) model is versatile and capable of estimating volume change during wetting (swell or collapse depending on the magnitude of initial suction and net mean stress); however, a large number of parameters are required from suction-controlled tests. Due to this reason, such a model has limitations for wider use in conventional geotechnical engineering practice.

Chen (1999) and Chen et al. (1999) conducted a series of wetting tests on an unsaturated, compacted loess soil from China. The shape of the LC curve obtained from the isotropic compression tests (Fig. 14) supports the concept of the LC model in Alonso et al. (1990). The compressibility indices associated with isotropic compression tests with respect to soil structure are relatively constant with increasing suction. However, the compressibility indices corresponding to water volume change vary significantly in the low suction range. As per the test results, yield due to increasing suction was found to be dependent on the initial soil density and net mean stress. If the initial density is low, then yield may occur at a low suction during drying. However, higher suction will be needed to cause yield by drying for soil specimens having high initial density. Similarly, for wetting case, it is easier for collapsible loess soils with lower initial density to collapse than that with higher initial density. In other words, a larger matric suction decrease is required for loess soils with higher initial density to collapse by wetting. In addition, it is found that if the compressibility of soil structure decreases significantly due to LC yielding, the soil will be insensitive to subsequent increase in suction and may exhibit elastic behavior. In other words, a much higher yield suction is required, which could be attributed to the expansion of yield surface. Consistently, for wetting case, if the soil structure is compressed significantly due to loading, then the soil becomes denser and the pores smaller, making it insensitive to suction decrease due to wetting. For this reason, it can be postulated that suction increase and suction decrease (i.e. drying and wetting) have similar effect on the structure of loess soils.

Kato and Kawai (2000) were among the first to study the collapse behavior using suction-controlled triaxial tests. They found the dependence of collapse deformation on stress path could be interpreted using the expansion of the yield surface due to wetting or loading. The loading-wetting sequence may influence the collapse deformation (collapse here is not limited to wetting-induced collapse) because of the difference between both elastic and plastic stiffness parameters for changes in net mean stress and that for changes in matric suction. For example, soil specimen that is loaded first and then wetted may yield due to wetting, while the soil specimen wetted first and then loaded may yield due to loading. The difference between stiffness parameters results in the difference between collapse deformations induced under two stress paths. In general, the elastic stiffness parameters associated with changes in net mean stress and in matric suction are assumed to be independent of matric suction and net mean stress, respectively (Alonso et al., 1990; Wheeler and Sivakumar, 1995). Due to this reason, collapse deformation may not be affected by the loading-wetting sequence if the stress state is confined to elastic region. However, many researchers have experimentally proved that the volume change is independent of stress path for natural loess soils, where volume change is not limited to that induced by wetting (Zheng and Zhang, 1989; Jiang et al., 1999).

While Kato and Kawai (2000) tests were limited to single density, Sun et al. (2004, 2007a) performed a series of wetting tests on a compacted clay with different initial densities. They found that the wetting-induced collapse deformations mainly depend on the initial void ratio and net mean stress under which the soil is wetted, irrespective of the matric suction that was initially imposed on the soil. The collapse behavior of compacted clay can be explained using the BBM. Later, Sun et al. (2007b) investigated the effects of initial density and stress states on the wetting-induced collapse of compacted clay. The collapse behavior with respect to initial density from isotropic compression tests is shown in Fig. 15. The collapse behavior using suction-controlled triaxial tests. They found the dependence of collapse deformation on stress path could be interpreted using the expansion of the yield surface due to wetting or loading. The loading-wetting sequence may influence the collapse deformation (collapse here is not limited to wetting-induced collapse) because of the difference between both elastic and plastic stiffness parameters for changes in net mean stress and that for changes in matric suction. For example, soil specimen that is loaded first and then wetted may yield due to wetting, while the soil specimen wetted first and then loaded may yield due to loading. The difference between stiffness parameters results in the difference between collapse deformations induced under two stress paths. In general, the elastic stiffness parameters associated with changes in net mean stress and in matric suction are assumed to be independent of matric suction and net mean stress, respectively (Alonso et al., 1990; Wheeler and Sivakumar, 1995). Due to this reason, collapse deformation may not be affected by the loading-wetting sequence if the stress state is confined to elastic region. However, many researchers have experimentally proved that the volume change is independent of stress path for natural loess soils, where volume change is not limited to that induced by wetting (Zheng and Zhang, 1989; Jiang et al., 1999).

While Kato and Kawai (2000) tests were limited to single density, Sun et al. (2004, 2007a) performed a series of wetting tests on a compacted clay with different initial densities. They found that the wetting-induced collapse deformations mainly depend on the initial void ratio and net mean stress under which the soil is wetted, irrespective of the matric suction that was initially imposed on the soil. The collapse behavior of compacted clay can be explained using the BBM. Later, Sun et al. (2007b) investigated the effects of initial density and stress states on the wetting-induced collapse of compacted clay. The collapse behavior with respect to initial density from isotropic compression tests is shown in Fig. 15. The collapse behavior using suction-controlled triaxial tests. They found the dependence of collapse deformation on stress path could be interpreted using the expansion of the yield surface due to wetting or loading. The loading-wetting sequence may influence the collapse deformation (collapse here is not limited to wetting-induced collapse) because of the difference between both elastic and plastic stiffness parameters for changes in net mean stress and that for changes in matric suction. For example, soil specimen that is loaded first and then wetted may yield due to wetting, while the soil specimen wetted first and then loaded may yield due to loading. The difference between stiffness parameters results in the difference between collapse deformations induced under two stress paths. In general, the elastic stiffness parameters associated with changes in net mean stress and in matric suction are assumed to be independent of matric suction and net mean stress, respectively (Alonso et al., 1990; Wheeler and Sivakumar, 1995). Due to this reason, collapse deformation may not be affected by the loading-wetting sequence if the stress state is confined to elastic region. However, many researchers have experimentally proved that the volume change is independent of stress path for natural loess soils, where volume change is not limited to that induced by wetting (Zheng and Zhang, 1989; Jiang et al., 1999).

While Kato and Kawai (2000) tests were limited to single density, Sun et al. (2004, 2007a) performed a series of wetting tests on a compacted clay with different initial densities. They found that the wetting-induced collapse deformations mainly depend on the initial void ratio and net mean stress under which the soil is wetted, irrespective of the matric suction that was initially imposed on the soil. The collapse behavior of compacted clay can be explained using the BBM. Later, Sun et al. (2007b) investigated the effects of initial density and stress states on the wetting-induced collapse of compacted clay. The collapse behavior with respect to initial density from isotropic compression tests is shown in Fig. 15. The
volumetric strain and shear strain due to collapse with respect to initial density in triaxial tests are shown in Fig. 16. From Sun et al. (2007b) tests, it is also found that the volume contraction due to wetting is more sensitive to changes in the degree of saturation than changes in the suction. These results suggest the possibility to predict the wetting-induced collapse behavior from the degree of saturation, using the SWCC as a tool.

4.3. Studies conducted on natural loess soils for interpreting their collapse mechanism

The soil mechanics-based approaches for interpreting the collapse behavior discussed above were mostly proposed based on the experimental results of compacted soils, which may show different collapse behaviors from natural loess soils. Hu et al. (2000) highlighted the difference between the volume change behavior of natural and compacted loess soils from the results of a set of oedometer tests on both natural and compacted loess soil specimens. Their test results are shown in Fig. 17, in which the after-collapse curves of initially unsaturated loess soils coincide with the compression line of saturated natural loess soil. These results suggest that wetting-induced collapse deformation is independent of the loading-wetting sequence. This means that both single- and double-oedometer tests would yield similar results about the collapse behavior of natural and compacted loess soils from different stress paths on artificial loess soil specimens. Similar conclusion was arrived at by Jiang et al. (2012a, b) between the mechanical behavior of natural and compacted loess soils. This also illustrates the difference between the mechanical behavior of natural and compacted loess soils. Similar conclusion was arrived at by Jiang et al. (2012a, b) from different stress path tests on artificial loess soil specimens (including single- and double-oedometer tests, strain- and stress-controlled triaxial tests, etc.). They attributed this difference to the different sources of bonding strength for natural loess soils and compacted or remolded loess soils. For this reason, it is implied that models proposed for compacted soils may not be reliable to explain the collapse behavior of natural loess soils although most of these models have been validated for collapsible soils. On the other hand, because of the difficulties associated with obtaining “identical” natural loess soil specimens and the variability in natural soil properties, experimental studies on compacted or remolded soils are valuable for understanding the mechanical behavior of natural loess soils. In addition, artificial loess soil samples which have been experimentally proved to behave similar to natural loess soils compensate this limitation (Hu et al., 2000; Jiang et al., 2012a, b).

Many researchers who have been involved in the studies on natural loess soils proposed the concept of structural strength, which is associated with bonding and mainly influenced by initial water content. Some researchers suggested the structural strength of natural loess soils as the preconsolidation stress (i.e. yield stress) (e.g. Hu et al., 2000). Fig. 18a shows the variation of structural strength with respect to initial water content for natural loess soils in Shaanxi in China modified after Hu et al. (2000). They attributed collapse to the loss of structural strength and bonding failure as a result of wetting. Hu et al. (2004) proposed a constitutive relationship for relating collapse deformation to water content after wetting based on experimental results, see Eq. (3), in which the two parameters are independent of the loading condition (i.e. confining pressure and stress level) and are required to be determined from triaxial wetting tests. However, most researchers defined the structural strength of natural loess soils as the difference in peak strength between natural and compacted loess soils. Fig. 18b shows the variation of structural strength with respect to initial water content for natural loess soils from different regions in China (Zhang et al., 1994; Dang and Hao, 1998; Liu et al., 2008). The structural strength has been found to be influenced not only by the initial water content but also by the stress applied on the soil, grain size distribution, chemical composition, environmental and climate conditions for loess soil deposition. Fig. 18 highlights that the structural strength of natural loess soils determined from the first method (i.e. preconsolidation stress) is much larger than that determined from the second method (i.e. difference in peak strength between natural and compacted loess soils). Recently, several studies were undertaken to study the relation between structural strength and yield stress or shear strength for natural loess soils (e.g. Dang and Li, 2001; Chen, 2008; Tian et al., 2011; Luo et al., 2014).

\[
\Delta_{c_v} = \alpha_v \Delta_{c_v} (w_s - w_0) + \alpha_v (w - w_0)
\]

where \(w_s\) and \(w_0\) are the saturated and natural water contents, respectively; \(\Delta_{c_v}\) is the final collapse deformation; \(\alpha_v\) is the initial slope of the \(\Delta_{c_v}\) with respect to \(\Delta w\) curve.

Shao et al. (2006, 2010, 2014) put forward a structural parameter, so-called the “structural index”, for structural soils, especially for natural loess soils. They stated that the natural soil structure can change due to disturbance, loading and wetting. Shear strength, unconfined compressive strength or deformation between natural and compacted soil specimens could indicate the bonding effect and particle interlocking arrangement in natural soil structure,
While the properties between natural and saturated natural soil specimens could indicate the effect of soluble salt bonding and capillary tension in natural soil structure. The structural index is defined as the product of strength reduction due to disturbance and that due to wetting:

\[
m_s = \left(\frac{\sigma_1 - \sigma_3}{\sigma_1 - \sigma_3}\right)_y \left(\frac{\sigma_1 - \sigma_3}{\sigma_1 - \sigma_3}\right)_s
\]

where \(m_s\) is the structural index; \(\sigma_1 - \sigma_3\) is the maximum deviatoric stress during shearing process; subscripts \(y\), \(r\) and \(s\) represent natural, compacted and saturated states, respectively. The first component on the right side of the equation is also called load-disturbance sensitivity and the second component is called water-immersion sensitivity. Shao et al. (2006, 2010, 2014) studied the relationships between the structural index and conventional properties of natural loess soils based on test results of natural, compacted and saturated loess soil specimens. They found that soil properties associated with grading, density and humidity (such as Atterberg limits, water content, dry density and void ratio) have significant influence on the structural index (see Fig. 19), while soil with the same structural index may have different values of such conventional properties. For this reason, the structural index together with conventional properties would be more reasonable and accurate for estimating the mechanical behavior of a structural soil. Their work is valuable for interpretation of mechanical behavior, especially strength behavior, of natural loess soils. Jiang et al. (2012b) proposed the concepts of structural damage stress (i.e. SDS, the structural yield stress of saturated soil) and generalized structural damage stress (i.e. GSDS, the structural yield stress of unsaturated soil), both of which could be determined from confined compression tests. Wetting-induced deformation is relatively small when the vertical pressure is less than SDS, and it increases rapidly once the vertical pressure exceeds SDS and decreases after it reaches the peak value of GSDS.

Shen (1993) and his group were the earliest investigators to interpret the mechanical behavior of natural loess soils using the concepts of elastoplastic breakage mechanics as a tool. They stated that the strength decrease and volume change of loess soils are the results of transition from natural structural soils to remolded soils; during this process, natural structure fails and secondary structure develops. Based on the concepts of breakage mechanics, the binary medium model and the block structural model were proposed particularly for natural loess soils.

As per the studies of Shen (2003), Shen and Deng (2003) and Shen and Hu (2003), natural loess soil structure is regarded as being composed of two elements in the binary medium model, i.e. cemented block and weak zone. The former element is actually the aggregate in natural loess soils that are firmly cemented by bonding materials (i.e. clay platelets and calcium carbonate) and the latter element is the zone among aggregates that has pore space (i.e. inter-aggregate pores) and separate particles (i.e. with no or very small amount of bonds at the contacts). For this reason, when a stress is applied on the soil, both elements are responsible for bearing and transferring the stress (see Eq. (5)), and deformations occur in both elements in response to changes in stress state in the soil. However, since the aggregates are well cemented and the pore space within the aggregates is relatively small that could be neglected, the volume change of natural loess soils due to loading or wetting is predominantly associated with the plastic deformations that occur in weak zones (see Eq. (6)). In the binary medium model, nine parameters are required for modeling the mechanical behavior of natural loess soils due to loading or wetting, which could be determined from compression tests, triaxial tests, etc. on both natural and remolded loess soil samples. The model has been found to be reasonable for estimation of the mechanical behavior (i.e. shear strength and volume change) for natural loess soils. A good agreement between the measured and predicted mechanical behavior using the binary medium model for natural loess soils in Shaanxi in China is presented in Fig. 20. Although it is easy to understand the process of natural loess soil structure failure due to loading and wetting using this model, a large number of tests are required to determine the parameters of the model.
\[
\{\sigma\} = (1 - b)\{\sigma_b\} + b\{\sigma_s\}
\]

\[\varepsilon_v = c_v \ln \frac{\sigma_m(1 + \kappa)}{\sigma_m^0}\]

where \(\{\sigma\}\) is the total stress applied on the soil; \(\{\sigma_b\}\) is the stress born by cemented blocks; \(\{\sigma_s\}\) is the stress born by weak zones; \(b\) is the parameter associated with the compression index of natural loess soils; \(\sigma_m\) is the mean stress; \(\sigma_m^0 = \sigma_m^0/S_T\); \(\sigma_m^0\) is the reference stress at zero volumetric strain, \(S_T\) is the degree of saturation, \(n\) is the parameter of the model; \(\kappa\) is the parameter of the model.

The block structural model treats the natural loess soil structure as an elastic block, in which bonding agents among particles or within aggregates are randomly distributed (Hu, 2000; Hu et al., 2005). When a stress is applied on the soil structure, contacts which have few bonds fail first as the stress exceeds the bonding strength at these contacts, resulting in breakage of primary large soil block into several small blocks. The block theory was originally proposed for rock mass was extended to model the volume change process of natural loess soils. The weakly bonded contacts in loess soils were considered as elements similar to structural surfaces in rock mass. However, almost all the contacts including the ones that are strongly bonded or within the aggregates in loess soils would fail as stress increases to considerably high level. The failure of weak bonding and breakage of the primary soil block were assumed to be elastic, while the subsequent slip of small blocks over each other and further breakage of small blocks were assumed to contribute to plastic deformations. For such a scenario, the stress-strain behavior could be represented using the following equation:

\[
\{\Delta \varepsilon\} = [C]\{\Delta \sigma'\} + A_1 \left\{ \frac{\partial f}{\partial \sigma'} \right\} \Delta f + A_2 \left\{ \frac{\partial g}{\partial \sigma^2} \right\} \Delta g
\]

where \(\{\varepsilon\}\) is the total strain due to loading and wetting; \([C]\) is the elastic flexibility matrix; \(\{\sigma'\}\) and \(\{\sigma^2\}\) are the effective stress and the net stress, respectively; \(f\) and \(g\) are the yield and damage functions, respectively, in which matric suction was involved; \(A_1\) and \(A_2\) are the plastic parameters corresponding to yielding and damage, respectively.

On the right hand side of Eq. (7), the first part is due to the elastic deformations, the other two are associated with the plastic deformations due to the slip of small blocks over each other and their breakage, respectively. The model involves six parameters which are influenced by the damage parameter or loading condition. In general, both confined and unconfined compression tests, wetting tests and triaxial undrained tests on natural loess soil specimens are required for determining these parameters. The model was validated for natural loess soil in Shaanxi in China, and a comparison between the measured and predicted volume change behavior is shown in Fig. 21. It is shown that the model is capable of modeling the volume change behavior of natural loess soils under relatively higher confining pressures (i.e. 200 kPa and 300 kPa), while the deformations are overestimated under lower confining pressures (i.e. 50 kPa and 100 kPa). Both the binary medium model and block structural model provide reasonable predictions of the mechanical behavior associated with loading and wetting for natural loess soils; however, they were rarely used in conventional geotechnical engineering practice since cumbersome tests (such as the oedometer tests and triaxial tests) are required for determining the various parameters. For this reason, much simpler models with less parameters or parameters that are easy-to-determine from conventional tests are suggested for future researches for interpreting and predicting the mechanical behavior of collapsible loess soils.

Collapse surface has been proposed particularly for natural loess soils during the last two decades (Zhang and Zheng, 1990; Shen, 1993; Hu et al., 2000, 2004; Jiang et al., 2012a, b), which is similar to the yield surface in elastoplastic models for judging whether collapse or irreversible deformations will occur. Zhang and Zheng (1990) proposed the concept of moistening deformation,
suggesting that significant volume contraction would occur when there is a soil water content increase rather than the soil getting fully saturated. In such a scenario, the deformation is controlled by two external factors, i.e. water and pressure. Therefore, the collapse surface for natural loess soils would be presented using a 3D plot with two abscissas representing the water content and volumetric strain, and the ordinate representing the vertical pressure under which the soil is wetted, as shown in Fig. 22a. This figure clearly shows the deforming development due to loading and wetting. Shen (1993) stated that there is an identifiable collapse surface for natural loess soils. When the stress path is within the collapse surface, deformations are almost elastic that could be modeled using elastic models. When the stress path crosses the collapse surface, the soil would show high compressibility, and soil collapses. In Hu et al. (2000, 2004) tests, since all the compression lines after collapse due to free access to water coincide with the compression line of initially saturated loess soil, they presumed that there would be a collapse surface for natural loess soils. Such a collapse surface could be defined from the compression line of natural saturated loess soils. However, as per Jiang et al. (2012a, b) study results of triaxial wetting tests on artificial loess soil specimens, it can be concluded that there exists a particular collapse surface, which is denoted as the dashed curve as shown in Fig. 22b. The vectors of total wetting-induced strains under different loading conditions are shown in p–q space in Fig. 22b (p is the mean stress; q is the deviatoric stress; \( R_{v} \) is the stress level, which is the ratio of the deviatoric stress to its maximum value or the value when axial strain reaches 10%; \( e_{v} \) is the axial strain; Tw1 to Tw9 represent different stress paths for triaxial tests, as summarized in the inset table in Fig. 22b). The horizontal and vertical components of the vector represent the total wetting-induced volumetric strain \( \Delta e_{v} \) and deviatoric strain \( \Delta e_{d} \), respectively. The proportion of vertical to horizontal components increases with stress level and the surface perpendicular to these vectors is approximately oval in shape. The wetting-induced strain is negligible when a specimen is wetted at a stress point inside the surface, while it becomes significant when the soil is wetted beyond the surface. Such a collapse surface could be defined from the results of wetting tests under different stress levels.

5. Summary and conclusions

The collapse triggering mechanism associated with wetting received much concern from various soils researchers during the past six decades all over the world. The discussions and contributions on collapse behavior from the literature are summarized in this paper under three categories, i.e. traditional, microstructure and soil mechanics-based approaches, with special reference to loess soil deposits. Traditional approaches have been found to be unsatisfactory to universally explain the collapse behavior for all loess soils. These approaches however were valuable to better understand collapsibility of local loess soils from simple tests. Soil microstructure is widely acknowledged to play an important role in controlling the mechanical behavior of loess soils. The microstructure of loess soils can be analyzed in terms of four factors, i.e. particle pattern, contact relation, pore form and bonding material. Among these factors, pore form and bonding material are suggested as the two dominant factors that have more influence on the collapse behavior. During the last two decades, a large number of researchers have devoted to classifying soil pores and exploring the water stability of possible bonding materials in order to predict the collapsibility of a loess soil from its microstructure. However, the microstructure approach lacks basic quantitative descriptor, in spite of significant advances have been made in image processing techniques during the last decade for quantitative analysis of the soil microstructure. Approaches extending the mechanics of unsaturated soils, not only explain the collapse as a result of changes in stress state variables, but provide a precise way to predict volume change due to collapse by suitable constitutive relations. Elastoplastic models define the yield surface for unsaturated soils, which clearly divide elastic and plastic deformations of an unsaturated soil. These models can also be used as tools to explain the collapse phenomenon as soil yields or the stress path crosses the yield surface due to either loading or wetting. Approaches based on the concepts of elastoplastic breakage mechanics (i.e. the binary medium model and the block structural model) have also been validated to provide reasonable prediction of collapse behavior for natural loess soils. However, there are some limitations for the application of these models in conventional practice mainly because of the difficulty in determining the parameters required for the models from time- and resource-consuming experimental tests. For these reasons, much simpler models with less parameters or parameters that are easy-to-determine from conventional tests are suggested for future research studies to better understand the mechanical behavior of natural loess soils. Such studies would be valuable for conventional geotechnical engineering practice applications.

Conflict of interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant
financial support for this work that could have influenced its outcome.

Acknowledgments

The first author gratefully acknowledges her appreciation to the Chinese Scholarship Council, which funded her Joint PhD research program. The second author thanks the support from Natural Sciences and Engineering Research Council of Canada (NSERC) for his research programs. The third author thanks the Chinese Ministry of Science and Technology for supporting his research program (grant No. 2014CB744701).

References


Terzaghi K. Theoretical soil mechanics. New York, USA: John Wiley and Sons; 1943.


P Morgan received her M.Sc. degree in Geological Engineering in 2013 from Chang’an University, X’An, China. In the same year, Li started her Ph.D. program in the same university. In September 2014, she was awarded a scholarship under the Chinese State Scholarship Fund for supporting her to study in Civil Engineering University of Ottawa, Canada, for 2 years based on her excellent academic performance. P Morgan’s Ph.D. program focuses on developing a simple model for estimating the collapse behavior of loess soils upon wetting based on the concept of unsaturated soil mechanics. Through her thesis, she would like to make a significant contribution to the knowledge of collapsible loess soils and problems that can be applied to engineering practice. Ping Li has authored or co-authored 6 journal papers and 3 international conference papers to date.

Sai Vanapalli is presently a professor of the Department of Civil Engineering, University of Ottawa, Ottawa, Canada. He received his Ph.D. in 1994 from University of Saskatchewan, Saskatoon, Canada. Since then, his research focus has been directed towards the development of simple techniques to interpret and predict the engineering behavior of unsaturated soils. These studies are used in the design of geotechnical structures such as pavements, shallow and deep foundations, slopes, soil nails and retaining walls. More recently, Sai Vanapalli has been developing simple techniques working with his graduate students to estimate and model (i) the stiffness (i.e. resilient modulus $M_r$, maximum shear modulus $G_{max}$, and modulus of elasticity $E$) of unsaturated soils, (ii) the multidi- mensional heave in expansive soils and its influence on geotechnical structures including slope stability, (iii) the mechanics of unsaturated soils during the design of pavements, shallow and deep foundations, and (iv) the influence of dynamic loads and earthquakes on the unsaturated soil behavior. Dr. Sai Vanapalli to-date has authored or co-authored 220 research articles that are disseminated in several journals and conferences, and has delivered 13 keynote lectures in international conferences. He received Sternac Award for service contributions from the Canadian Geotechnical Society in 2010 and Clinksi Award in 2015 for his outstanding research contributions from the University of Ottawa.

Tongli Li obtained M.Eng. Degree from Changchun University of Earth Science, Changchun, China, and Ph.D. degree from X’An College of Geology, X’An, China. He is presently a professor of Geological Engineering, Changchun University, X’An, China, where he used to be the Head of the Department and the Vice-dean of the School of Geological Engineering and Geomatics. Li has been involved in geological research, consultation and education for more than 25 years. His research interests focus on geotechnical environment and geo-disaster in Chinese loess regions. He is authored and co-authored more than 80 academic articles and sits on the Editorial Board of Geoenvironmental Disasters. He is a member of IAEG and the Vice-President of International Consortium on Geo-disaster Reduction. He has consulted for many key projects of urban buildings, highways, hydro-power stations, oil-gas transfer lines and nuclear electric plants in China. In recent years, he has been supported by two National Natural Science Funds and one National Basic Research Program to relate the permeability of loess soils to geo-disasters in loess regions based on the concepts of unsaturated soil mechanics. The research studies that he and his collaborators undertaken are widely used for both interpreting and estimating the geo-disasters associated with loess soils.