

Compressing random microstructures via stochastic Wang tilings

Jan Novák,^{1,*} Anna Kučerová,^{1,†} and Jan Zeman^{1,2,‡}

¹Department of Mechanics, Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29 Praha, Czech Republic

²Centre of Excellence IT4Innovations, VŠB-TU Ostrava, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic

(Received 5 June 2012; published 22 October 2012)

This Rapid Communication presents a stochastic Wang tiling-based technique to compress or reconstruct disordered microstructures on the basis of given spatial statistics. Unlike the existing approaches based on a single unit cell, it utilizes a finite set of tiles assembled by a stochastic tiling algorithm, thereby allowing to accurately reproduce long-range orientation orders in a computationally efficient manner. Although the basic features of the method are demonstrated for a two-dimensional particulate suspension, the present framework is fully extensible to generic multidimensional media.

DOI: [10.1103/PhysRevE.86.040104](https://doi.org/10.1103/PhysRevE.86.040104)

PACS number(s): 05.20.-y, 81.05.Zx, 61.43.Gt

In 1961, Wang introduced a tiling concept based on square dominoes with colored edges permitting their mutual assembly in a geometrically compatible (hard) manner [1]. Since then, his tiles have been the subject of studies in discrete mathematics [2–4] and have found an extensive use in computer graphics [5], game industry [6], theory of quasicrystals [7], or biology [8]. From the perspective of this Rapid Communication, the appealing feature of Wang tilings is that they can compress and reproduce naturally looking planar patterns or three-dimensional surfaces by employing only a small number of distinct tiles [3,9,10]. Motivated by this observation, we further explore the potential of Wang tiles to represent long-range spatial correlations in disordered microstructures, a problem common to materials science [11], geostatistics [12], or image analysis [13].

In this regard, two closely related applications can be distinguished, namely, the *microstructure reconstruction* [14–16] based on given spatial statistics and *microstructure compression* [17–19] aiming at efficient representation of material structures in multiscale computations [20]. Our focus is on the latter, since these procedures usually have the microstructure reconstruction techniques at heart, hereby covering the common features of both.

To the best of our knowledge, compression algorithms reported to date use a single cell [periodic unit cell (PUC)] that is periodically extended to tile the plane in a deterministic manner [18]. Such structures then inevitably manifest strong long-range correlations with a period of the PUC dimensions. We shall demonstrate that these artifacts can be effectively controlled when utilizing small Wang tile sets [3,10], carefully designed to capture morphological trademarks of compressed media, combined with the fast stochastic Cohen-Shade-Hiller-Deussen (CSHD) tiling algorithm [5] for real-time texture generation. A potential of this approach will be demonstrated for equilibrium two-dimensional particulate suspensions consisting of equisized disks of radius ρ uniformly distributed in a homogeneous matrix (cf. Ref. [21]).

To this goal, consider a two-dimensional microstructured domain $\mathcal{D}$ discretized by a regular square lattice. Each

lattice cell contains specific morphological patterns that are compatible on contiguous boundaries [Fig. 1(b)]. If there are no missing cells in $\mathcal{D}$, the discretization is called a *valid tiling*, and a single cell is referred to as the *Wang tile* [1] [Fig. 1(a)]. The tiles have different codes on their edges [lowercase Greek symbols in Fig. 1(a)], and are not allowed to rotate when tiling a plane. The number of distinct tiles within $\mathcal{D}$ is limited, though arranged in such a fashion that none of them or any of their subsequence periodically repeats. The gathered distinct tiles are referred to as the *tile set* [Fig. 1(a)]. Sets that enable uncountably many, always aperiodic, tilings are called *aperiodic* [3,4]. In real-world applications, the assumption of strict aperiodicity of the tile sets is relaxed to aperiodicity of tilings, ensured, e.g., by the CSHD algorithm [5] introduced next.

Intuitively, the ability of a tile set to control long-range order effects arises from tile and edge code diversities [Fig. 1(a)]. Both factors are related, so that while the number of edge codes n_i^c in the i th spatial direction can be chosen arbitrarily, the number of tiles n^t must satisfy

$$n^t = n^{\text{NW}} \sqrt{n^{\text{cs}}}, \quad (1)$$

where $n^{\text{cs}} = (n_1^c n_2^c)^2$ is the number of tiles in the complete set and $n^{\text{NW}} = 2, \dots, \sqrt{n^{\text{cs}}}$ stands for the optional number of tiles with an identical arrangement of northwestern (NW) edge codes. The complete set of n^{cs} tiles is obtained by permuting the chosen codes c_i . In valid tilings, the southeastern edge codes must match those assigned to NW edges [Fig. 1(b)]. Thus, the tiles in the complete set are sorted according to existing NW edge code combinations and the desired number of tiles in a user defined set is formed by selecting n^{NW} of unique tiles from each group. Such a set is referred to as $Wn^t/n_1^c n_2^c$ in what follows. Notice that the $W1/1-1$ set corresponds to the PUC setting.

In the stochastic tiling algorithm, the index of a new tile to be placed is selected randomly with the uniform probability from an appropriate NW group compatible with the eastern code of the tile previously placed and the southern code of the tile above the one to be placed [edges γ and α in bold adjacent to the shaded cell in Fig. 1(b)]. Aperiodicity of the resulting tiling is guaranteed provided that the random generator never returns a periodic sequence of numbers and that each NW group contains at least two distinct tiles [5].

*novakj@cml.fsv.cvut.cz

†anicka@cml.fsv.cvut.cz

‡zemanj@cml.fsv.cvut.cz

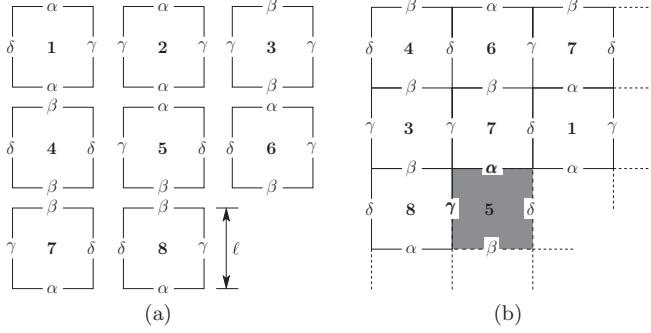


FIG. 1. (a) A Wang tile set W8/2-2 with edge length ℓ and codes $\{\alpha, \beta, \gamma, \delta\}$. (b) Example of an aperiodic valid tiling.

Analogously to the existing works on reconstruction and compression of random media, the tile morphologies are designed by an optimization procedure expressed in terms of suitable statistical descriptors. As our focus is to control long-range artifacts, we limit the exposition to the two-point probability function $S_2(\mathbf{x})$ [11]. For statistically uniform ergodic media, it provides the probability that two arbitrary points from $\mathcal{D}$, separated by $\mathbf{x}$, are simultaneously found in the particle phase. The function satisfies $S_2(\mathbf{0}) = \phi$, where ϕ is the volume fraction of particles, and $S_2(\mathbf{x}) \approx \phi^2$ for $\|\mathbf{x}\| > \lambda$ indicates the absence of long-range orders at the characteristic length λ [Fig. 2(b)].

In the current setting, the Wang tiling compression consists of a set of n^t tiles of the edge length ℓ , in which we distribute n^d disks. The configuration of particles is determined by the parameter vector $[t_d, \xi_{1,d}, \xi_{2,d}]_{d=1}^{n^d}$, where $t_d \in \{1, \dots, n^t\}$ specifies the parent tile index of the d th disk and $\xi_{j,d} \in [0, \ell]$ the local position of the disk at the j th direction. To determine the two-point probability function $\tilde{S}_2$ for a given configuration, we assemble a tiling that covers the domain of the same size as the representative sample $\mathcal{D}$ [Fig. 2(a)]. Notice that such tiling corresponds to a realization of a statistically homogeneous material, since the tiles are selected from NW edge groups with the uniform probability. The proximity of the tile-based morphology to the original sample is quantified by an objective function

$$E = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} [S_2(\mathbf{x}) - \tilde{S}_2(\mathbf{x})]^2 d\mathbf{x}, \quad (2)$$

which can be efficiently evaluated using the fast Fourier transform techniques (e.g., Ref. [11]). The minimization of (2) is carried out by the simulated reannealing method with computational cost similar to existing PUC design strategies [22]. The algorithm ensures that the tiles in the set satisfy the corner constraint [5], requiring that the tile corners are not occupied by a disk, and determines the number of disks n^d and the cell size ℓ such that the local volume fractions associated with edges (gray disks in Fig. 5) and tile interiors (white disks in Fig. 5) are as close to the target value ϕ as possible.

An example of optimal approximations of the target microstructure from Fig. 2 in terms of a PUC and the Wang tile set W18/3-3 is shown in Fig. 3. The representations are based on $n^d = 49$ particles and tile sizes $\ell = 24.5\rho$ and $\ell = 7.5\rho$, respectively. Evidently, both heterogeneity patterns carry long-range order effects with the period of ℓ , manifested

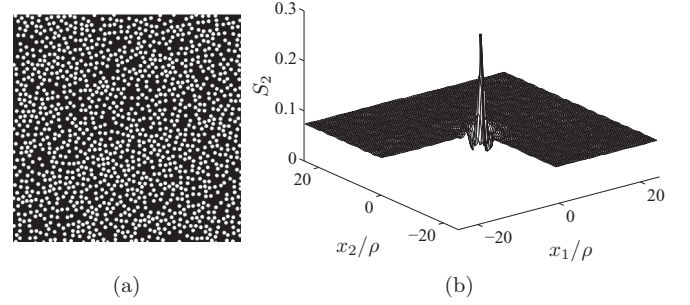


FIG. 2. (a) Reference two-phase medium of size $174\rho \times 174\rho$ formed by equilibrium distribution of 1300 equisized disks of volume fraction 26.8% and (b) its two-point probability function S_2 ; ρ is the disk radius.

as the local peaks

$$\hat{S}_2 = \max_{\mathbf{k} \setminus \{\mathbf{0}\}} \tilde{S}_2(\mathbf{k}\ell) \quad (3)$$

in the two-point probability functions [Figs. 3(b) and 3(d)]. Notice that $\hat{S}_2$ is always equal to ϕ for the PUC approach, whereas the Wang tiles are capable of adjusting these artifacts by the proper morphology design. This is also reflected in visual regularity of the generated suspensions [compare Fig. 2(a) with Figs. 3(a) and 3(c)]. Also observe the locally periodic region in Fig. 3(c), arising from the local character of the CSHD algorithm and from the lowest number of tiles in groups of admissible NW edge code combinations, $n^{\text{NW}} = 2$. Such a phenomenon is thus less likely when increasing the parameter n^{NW} , however, at the expense of increasing set sizes, especially for higher edge code diversities, recall Eq. (1).

The principal features of Wang tile-based compressions are further illustrated in Fig. 5. Instead of relying on a

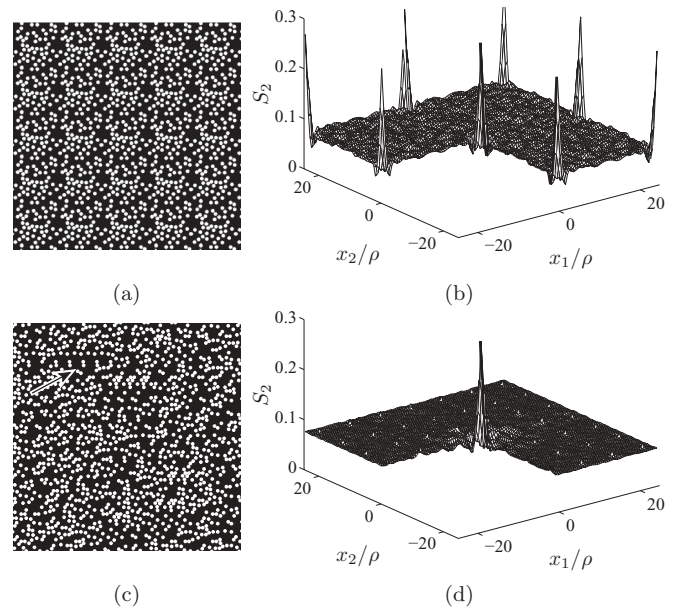


FIG. 3. Optimized microstructures and two point probability functions $\tilde{S}_2$ for PUC (a), (b) and set W18/3-3 (c), (d). The arrow in (c) denotes the periodic region due to the local character of tile placement in the CSHD tiling algorithm.

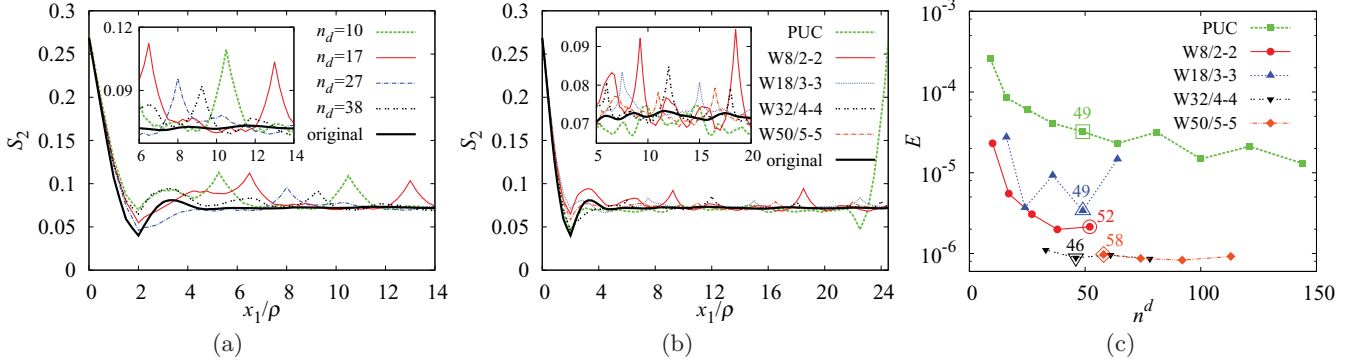


FIG. 4. (Color online) Comparison of $\tilde{S}_2$ in the x_1 - S_2 plane for (a) tile set W8/2-2 with respect to the number of disks n^d involved and (b) different sets. (c) Comparison of different sets in terms of objective function (2) and with respect to the number of disks n^d involved. The curves in (b) are plotted for particular values of n^d highlighted in (c).

single cell containing the complete morphological information [Fig. 5(a)], the tiling-based approach utilizes substantially simpler building blocks [Fig. 5(b)] assembled to comply with edge constraints (gray disks). This, however, restricts the space of admissible disk configurations in Wang tiles compared to the single PUC design.

Now, we are in the position to quantify to which extent is the quality of reconstructed suspensions determined by the tile set diversity and the morphology design itself. This aspect is examined first in Fig. 4 by means of sections of the two-point probabilities $S_2(x_1, 0)$ and the objective function E , revealing that two effects govern the amplitude and period of the local extremes $\tilde{S}_2$. First, for a fixed tile set, increasing the number of disks increases the tile edge dimension ℓ (and thus the period) and slightly decreases the amplitude [Fig. 4(a)]. On the other hand, increasing the number of tiles decreases the period as well as the magnitude of local extremes [Fig. 4(b)]. Also notice that the quality of the W18/3-3 set in terms of the objective function (2) is systematically inferior to W8/2-2 [Fig. 4(c)]. This is caused by an inaccurate representation of disk volume fraction for the former set, which pollutes the shape of $\tilde{S}_2$ statistics [Fig. 4(b)] [23]. It further follows from Fig. 4(c) that increasing the tile set diversity is much more efficient; for sets containing more than 32 tiles, the error is almost independent of the number of disks. This saturation value reflects the inaccurate representation of short-range values of S_2 , caused by the particular form of the objective function (2). If needed, the local details can be incorporated in terms of higher-order statistics or specifically tailored descriptors [14, 15, 24, 25].

It is now clear that the local extremes can be attributed to a limited number of tiles used in a repetitive, although random, fashion. Actually, two components repeat when

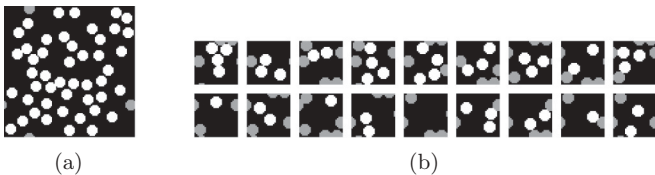


FIG. 5. Building blocks of microstructure compression based on (a) PUC and (b) tile set W18/3-3 with 49 disks assigned to tile edges (gray) and interiors (white).

tiling the plane: tile edges and interiors. To study the local artifacts analytically, we consider user defined sets with tiles selected so that their edges incorporate each code c_i at least once. Assuming that tiles and edges repeat independently, the maximum local extremes can be estimated as

$$\hat{S}_2^p \approx \frac{\phi^t}{n^t} [\phi + (n^t - 1)\phi^2] + \max_i \left\{ \frac{\phi^e}{n_i^e} [\phi + (n_i^e - 1)\phi^2] \right\}, \quad (4)$$

where $\phi^t = (\ell - 4\rho)^2/\ell^2$ and $\phi^e = 1 - \phi^t$ denote the volume fractions of tile interiors (occupied by white disks in Fig. 5) and edges (occupied by gray disks in Fig. 5), respectively [26].

In Fig. 6, we compare the actual values of $\tilde{S}_2$ with theoretical predictions (4) for several values of ϕ^e . Apart from the limit cases, $\phi^e \in \{0, 1\}$, $\hat{S}_2^p$ was also explored for $\phi^e = 0.2$ (average value from all considered tile sets). We observe that an almost exact match is obtained for the lower bound with $\phi^e = 0$ [red solid curve in Fig. 6(a)], demonstrating that the long-range artifacts are carried mainly by the tile interiors. This is rather surprising, since all considered tile sets satisfy $n^t \gg n_1^e = n_2^e$, so that edges repeat more often than the tile interiors. Moreover, the magnitude of spatial artifacts converges rapidly to the limit value ϕ^2 . Altogether, this leads us to the conclusion that artifacts due to the discrete nature of Wang tilings can be almost eliminated by a proper morphology optimization.

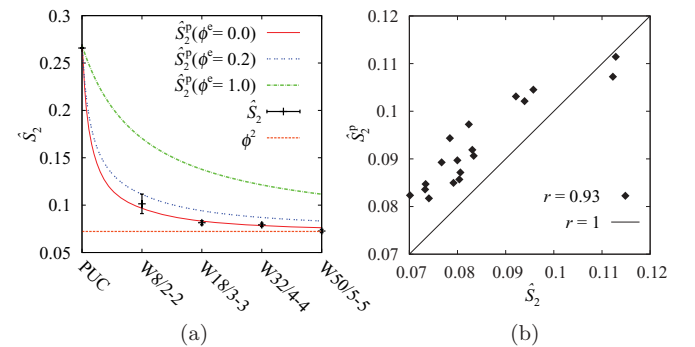


FIG. 6. (Color online) (a) Dependence of local extremes $\hat{S}_2$ on the particular tile set. (b) Correlation of local peaks obtained from two-point probabilities of optimized microstructures and their predictions given by Eq. (4); r in (b) denotes the Pearson correlation coefficient.

Also note that the accuracy of the estimate (4) appears to be reasonable, both for the average value of $\phi^e = 0.2$ [blue double dotted curve in Fig. 6(a)], as well as for values corresponding to individual tile sets [Fig. 6(b)]. It may thus serve as a basis for the *a priori* selection of the tile set parameters n_i^e and n^l .

Summary. A compression and reconstruction technique based on Wang tilings has been proposed and applied to two-dimensional microstructures of disordered particulate media. The technique is extensible to generic three-dimensional microstructures, adopting the frameworks of Wang cubes [27,28] and image synthesis [5]; it substantially generalizes the periodic unit cell concept by making use of multiple tiles instead of a single cell, preserves long-range spatial features,

and is computationally efficient. A formula for estimates of long-range order spatial artifacts has also been proposed and verified for the studied material system.

The authors acknowledge the Czech Science Foundation endowment under Grants No. 105/12/0331 (J.N.) and No. 105/11/0411 (A.K. and J.Z.). The work by J.Z. was also partially supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070). In addition, we would like to thank Jiří Šejnoha and Milan Jirásek of CTU in Prague and anonymous referees for helpful comments on the manuscript.

-
- [1] H. Wang, *Bell Syst. Tech. J.* **40**, 1 (1961).
 - [2] R. M. Robinson, *Invent. Math.* **12**, 177 (1971).
 - [3] B. Grünbaum and G. Shephard, *Tilings and Patterns* (Freeman, New York, 1986).
 - [4] K. Culik, *Discrete Math.* **160**, 245 (1996).
 - [5] M. Cohen, J. Shade, S. Hiller, and O. Deussen, *ACM Trans. Graphics* **22**, 287 (2003).
 - [6] E. Demaine and M. Demaine, *Graphs Combinatorics* **23**, 195 (2007).
 - [7] D. Aristoff and C. Radin, *J. Phys. A: Math. Theor.* **44**, 255001 (2011).
 - [8] E. Winfree, F. Liu, L. Wenzler, and N. Seeman, *Nature (London)* **394**, 539 (1998).
 - [9] A. Glassner, *IEEE Comput. Graphics Appl.* **24**, 86 (2004).
 - [10] A. Lagae and P. Dutré, *Comput. Graphics Forum* **27**, 114 (2008).
 - [11] S. Torquato, *Random Heterogeneous Materials* (Springer, New York, 2002).
 - [12] J. Chilés and P. Delfiner, *Geostatistics* (Wiley, New York, 1999).
 - [13] J. Serra, *Image Analysis and Mathematical Morphology* (Academic, London, 1982).
 - [14] C. L. Y. Yeong and S. Torquato, *Phys. Rev. E* **57**, 495 (1998).
 - [15] P. Čapek, V. Hejtmánek, L. Brábec, A. Zikánová, and M. Kočířík, *Transp. Porous Media* **76**, 179 (2009).
 - [16] M. A. Davis, S. D. C. Walsh, and M. O. Saar, *Phys. Rev. E* **83**, 026706 (2011).
 - [17] G. L. Povirk, *Acta Metall. Mater.* **43**, 3199 (1995).
 - [18] J. Zeman and M. Šejnoha, *Modell. Simul. Mater. Sci. Eng.* **15**, S325 (2007).
 - [19] H. Lee, M. Brandyberry, A. Tudor, and K. Matouš, *Phys. Rev. E* **80**, 061301 (2009).
 - [20] D. T. Fullwood, S. R. Niezgodá, B. L. Adams, and S. R. Kalidindi, *Prog. Mater. Sci.* **55**, 477 (2010).
 - [21] M. D. Rintoul and S. Torquato, *J. Colloid Interface Sci.* **186**, 467 (1997).
 - [22] J. Novák, A. Kučerová, and J. Zeman, arXiv:1110.4183.
 - [23] The best resulting disk volume fraction for W18/3-3 and 24 disks was by 1.3% higher than the prescribed value. The remaining sets, however, never resulted in a worse scatter than 0.2%.
 - [24] Y. Jiao, F. Stillinger, and S. Torquato, *Proc. Natl. Acad. Sci. USA* **106**, 17634 (2009).
 - [25] R. Piasecki and A. Plastino, *Physica A* **389**, 397 (2010).
 - [26] Observe that the estimate (4) contains contributions from tile interiors and edges. In addition, the tile interior part arises from two complementary events. If the two adjacent tiles are identical, the probability of simultaneously locating two disks distant by ℓ amounts to $\phi^l \phi / n^l$. Otherwise, we consider the disks in both tiles as independent, which gives rise to the term $\phi^l (1 - 1/n^l) \phi^2$. The contribution of repeated tile edges related to i th spatial direction is established analogously by estimating the probability of simultaneously matching two edges distant by ℓ as $1/n_i^e$.
 - [27] K. Culik and J. Kari, *J. Universal Comput. Sci.* **1**, 675 (1995).
 - [28] A. Lu, D. S. Ebert, W. Qiao, M. Kraus, and B. Mora, *ACM Trans. Graphics* (2007), doi:10.1145/1243980.1243985.