Microporous and Mesoporous Materials

Characterization and Analysis of Ring Topology of Zeolite Frameworks

--Manuscript Draft--

Manuscript Number:	MICMAT-D-22-01939R1
Article Type:	Full length article
Keywords:	zeolite; topology; rings; vertex symbols
Corresponding Author:	William F. Schneider University of Notre Dame Notre Dame, Indiana UNITED STATES
First Author:	Jerry T. Crum
Order of Authors:	Jerry T. Crum
	Justin R. Crum
	Cameron Taylor
	William F. Schneider
Abstract:	The topology of zeolite frameworks and of associated tetrahedral sites (T-sites) are commonly characterized by their associated rings, typically defined as some set of closed paths or cycles through a framework that cannot be decomposed into shorter cycles. These ring descriptors have been used to identify feasible zeolite topologies, to describe the similarity and differences between zeolites, to identify sites or voids of catalytic relevance, and as machine learning fingerprints. Numerous definitions and algorithms for finding zeolite rings have been proposed and applied throughout the literature. Here we report an analysis of rings and T-sites in a large number of zeolite frameworks using Zeolite Simulation Environment, a Python package that implements an efficient algorithm presented by Goetzke and Klein for finding rings in arbitrary frameworks. We compare the result of a number of common and new ring definitions applied to a large number of common zeolite frameworks. We discover previously unrecognized rings in a number of frameworks. We show that the vertex symbol, a common approach used to characterize T-sites, misses important parts of the stereochemistry around a T-site, and propose an alternative definition. This tool provides an effective platform for characterizing zeolite and T-site structures useful for building models and doing machine learning.

University of Notre Dame
College of Engineering

University of Notre Dame
Department of Chemical and Biomolecular Engineering
250 Nieuwland Science Hall
Notre Dame, Indiana 46556

Telephone: (574) 631-8754
Facsimile: (574) 631-8366
Email: wschneider@nd.edu
https://wfschneidergroup.github.io

January 5, 2023

Heloise de Oliveira Pastore, Dr.
Editor-in-Chief
Microporous and Mesoporous Materials
Email: em@editorialmanager.com

Ms. Ref. No.: MICMAT-D-22-01939
Title: "Characterization and Analysis of Ring Topology of Zeolite Frameworks"

Dear Dr. Pastore,

We would like to thank the reviewers for their constructive comments on our work. Below we reproduce the comments from the reviewers along with detailed responses and a description of corresponding changes made to the manuscript. In the process, we have also made a number of editorial changes to improve the presentation, but not scientific content, of the work. We attach for review only a version of the manuscript with all changes highlighted. We think these changes improve the work and hope that with them the manuscript is acceptable for publication in *Microporous and Mesoporous Materials*.

Sincerely,

William F Schneider

William F. Schneider
Dorini Family Chair and Department Chair, Department of Chemical and Biomolecular Engineering
Executive Editor, *The Journal of Physical Chemistry C*
Chair, Catalysis Science and Technology Division, American Chemical Society

Reviewer: 1

Comment #1:

This contribution is devoted to the development of new topological descriptors for the atomic nets in crystals. All the proposed descriptors are based on the known definition of a ring, which is a cycle without shortcuts. In addition to the widely used vertex symbol, which accounts for the shortest rings meeting in a given net vertex (atom), the authors propose to consider larger rings. Another descriptor is the modified shortcut ring, which strengthens the shortcut condition by requiring the shortcut to decompose the ring into at least one (not two) shorter rings. The third descriptor, shortest path ring, can extend the set of the rings describing a given node with the rings that are associated with the T nodes, which compose a ring containing the given node. The authors conclude that these new descriptors provide additional information about the net and enable one to discriminate better the nodes in the net.

Thanks to the reviewer for the constructive comments, which have helped us improve the manuscript.

No doubt, any new descriptor bears some additional information on the object. The question is if this information helps to solve problems that already existing descriptors cannot solve. In this respect, the following should be taken into account:

The idea to consider not only shortest rings, but all rings up to a given size is not new. In particular, the program ADS of the ToposPro program package (https://topospro.com) computes all rings up to a specified size with the 'All rings' option.

We have added references to ToposPro and ZeoTSites, two pieces of software that implement various ring counting methods, on p. 2:

> *"A number of definitions and enumerations of rings in zeolites have been presented [8–14] and implemented in software tools, including ZeoTSites [14] and ToposPro [15]."*

Comment #2:

Besides rings, an important descriptor is strong ring, which is not a sum of shorter rings. In particular, the 12-ring in CHA mentioned by the authors is not strong; it is the sum of 10 shorter rings: six 4-rings, one 6-ring and three 8-rings. Strong rings have a clear physical meaning: they are sections of channels in a zeolite framework (see Blatov V.A., Delgado-Friedrichs O., O'Keeffe M., Proserpio D. M. Three-periodic nets and tilings: natural tilings for nets. Acta Cryst. 2007, A63, 418-425). What physical meaning does the modified shortcut ring have?

We have added the definition for strong rings to the Introduction (p. 3) where we introduce other ring definitions:

> *"A more restrictive definition for a ring is that of a "strong ring", which is not the sum of a set of smaller rings [9, 16]."*

We introduce the modified shortcut ring in the Introduction (p. 6):

> *"We highlight rings that are found by these conventions but not typically discussed in the literature for a number of frameworks. We propose a modification of the shortcut existing*

conventions. We also show that the vertex symbol, a common approach used to characterize T-sites [17], based on the shortest rings connecting the neighboring oxygen, misses important parts of the stereochemistry around a T-site.

We have rewritten the description of the modified shortcut ring on p. 8 to more clearly motivate the concept:

"The shortest path convention generally finds fewer (and smaller) rings than does an all non-shortcut bearing ring convention. As we show below, these two conventions bracket the set of rings captured in the IZA Structure Database, the all-rings convention generally including even larger rings and the shortest path convention missing large rings. In seeking a convention of intermediate stringency, we identified a convention that would require the shortcut to be shorter than one, rather than both, of the paths connecting two nodes along the ring. This modified shortcut definition is illustrated in Fig. 3(b). By the standard shortcut convention, AFI would be considered to contain a 14-MR formed from the union of Path 1 (blue) and Path 2 (purple) through T1 and T2. However, because T1 and T2 are also connected by Path 3, which, while the same length as Path 1, is shorter than Path 2, the 14-MR does not satisfy the modified shortcut criterion. As we show below, the modified shortcut rule identifies rings of a given zeolite that are closer to those identified in the IZA Structure Database as rings of interest than do either all-ring or shortest-path- ring rules."

We have updated Figure 11 (formerly Figure 12) to refer to "Modified Shortcut Rings" instead of "This Work."

Comment #3:
ADS can also compute extended point symbols and cycle sequences as additional cycle-type net invariants, which help to discriminate nets and nodes topologically. Do we really need more descriptors?

The goal of our work is to present an analysis of ring definitions and their implications for descriptions of zeolites. In doing so we discovered a lack of structural uniqueness in the definition of a vertex symbol that warranted further analysis. As we reported, "ordering" the vertex symbol to assure that T-sites that share the same symbol are stereochemically identical increases the number of unique T-sites only modestly. We do not disagree with the observation that there are other ways of capturing this same information. Nonetheless, vertex symbols are commonly employed in the community, and we believe it is thus of value to highlight and search for the presence of stereochemical differences in vertex symbols. We have rewritten the narrative on p. 19-21 to remove "descriptor" and to focus on the most important insights:

"The conventional definition of a vertex symbol, which lists pairs of rings along opposite edges of the tetrahedron from smallest to largest, only partially captures the stereochemistry around a T-site. As example, Fig. 16 illustrates the rings associated with T3, T9, and T1 or MOR, MON, and EON, respectively, all of which share the same $4{\cdot}5_2{\cdot}5{\cdot}8_2{\cdot}5{\cdot}8_2$ vertex symbol. While MOR T3 and EON T9 are similar, MON T1 differs in the location of the 5_2- and 4-MR edges, as highlighted in the lower panel of Fig. 16. To remove this ambiguity, an alternative approach is to list the rings of the vertex symbol by starting with the largest one and systematically adding the remaining rings following the edges of the tetrahedron such that each ring listed is connected to one of the oxygen

atoms of the next ring listed. This provides a physical meaning to the order of the vertex symbol rings which accounts for differences in the connectivity of those rings around the T-site. Within this algorithm MOR T3 and EON T9 would be labeled as: $8_2 \cdot 8_2 \cdot 5_2 \cdot 5 \cdot 4 \cdot 5$, and MON T1 as: $8_2 \cdot 8_2 \cdot 4 \cdot 5 \cdot 5_2 \cdot 5$. The difference is subtle but highlights the distinct structural difference between the two types of T-sites that is not otherwise captured by a vertex symbol. We computed the ordered vertex symbols for all 1460 T-sites in the IZA Database. The 649 unique vertex symbols in the database increase to 666 unique ordered vertex symbols. Thus, the consideration of stereochemistry only modestly in- creases the space of unique tetrahedral sites. Table 4 lists some common T-site vertex symbols and associated ordered vertex symbols. Complete results are presented in the Supplementary Information."

Comment #4:

The topology of any zeolite net as well as any T node in the net is unambiguously described by vertex symbols and coordination sequences of the nodes. What is the advantage of the proposed descriptors? Do they discriminate all nodes and framework nets in zeolites?

Please see the response to Comment 3.

Comment #5:

One more important issue concerns the completeness of the study. First of all, the authors considered the rings up to size 18. Why? There are zeolites that contain larger rings even if we account for only the shortest rings at a node: IFU - 20-rings; EWT - 21-rings; SYT - 24-rings; ITV - 30-rings. Next, the authors state that they applied their approach to all IZA frameworks. How many frameworks were considered? A complete list with the computed descriptors like in Table 2 should be provided in the Supplementary.

Reviewer 1 makes a terrific point regarding inclusion of results for all frameworks that we collected information on. We have included a zipped file containing the rings found for every T-site in every zeolite framework listed in the IZA Database using each of the ring finding conventions as supplementary information. We have added a Data Statement section to the end of the manuscript to describe the results included in the S.I. on p. 23

"Complete ring finding results are available for download via the Supplementary Information. This folder contains a file for each ring counting convention: all rings, modified shortcut rings, shortest path rings, vertex symbol rings, and ordered vertex symbol rings. These files contain the rings associated with each oxygen atom and T-site in every zeolite framework listed on the IZA Structure Database."

Reference 2 shows the probability of finding rings of various sizes in zeolites, where that probability above 18-MR is not sufficient to have meaningfully enhanced our results. There is also a precedent for limiting the search to 18-MR set by Reference 3.

The four frameworks listed as containing larger rings (IFU, EWT, SYT, and ITV) are all disconnected type frameworks (as shown on the IZA database) meaning not all their T-sites are connected to four oxygen atoms. The focus of our work was not on discerning all of the larger rings in these disconnected frameworks but rather to highlight and compare differences in ring finding conventions. Updated text on p. 9 explains the decision for limiting the search.

"We limit the search to 18-MR and smaller because the differences in ring finding

conventions are captured within the smaller rings, because probabilities of occurrence become small with increasing ring size,[2] and for computational expediency."

Reviewer: 2

Comment #1:

The authors report a detailed study of various methods for identifying and counting rings in known, synthesizable zeolite framework structures. Counting rings in 3D network structures remains a challenging task without a single mathematical definition of correctness. It's also an important task in zeolite science because rings are used to help zeolite scientists quickly understand (1) building blocks that make up zeolites, and (2) channel / window dimensions that constrain molecular transport properties, thus defining a given zeolite's usefulness in catalysis and separations.

The authors suggest that rings can also be helpful for understanding cavity volumes, which are important for characterizing a zeolite's adsorption capacity. I disagree with the idea of generalizing the definition of zeolite rings so they can say something about cavity capacities. There are other, better methods to do this (IZA "largest sphere", simulation methods such as described in https://pubs.acs.org/doi/10.1021/acs.langmuir.7b01682). Also, such rings added to existing ring lists may confuse zeolite scientists by making them think that these bigger ring sizes allow transport of much larger molecules than can actually fit through proper zeolite windows. Thus, I don't agree with some of the authors' motivation for this study.

In detail, the authors have compared ring counts obtained from four approaches: (1) a standard code ("RINGS"), (2) the International Zeolite Association (IZA) vertex symbol method used in the online IZA database, (3) a shortest-path approach (Sastre and Corma), and (4) their own modified version of the shortest-path approach implemented in a new "Zeolite Simulation Environment" (ZSE) python package. The authors have analyzed the nature of discrepancies among these four methods, finding a hierarchy of strictness in defining rings, with IZA vertex symbols being the strictest (and thus leaving out rings that other methods find), shortest-path being less strict, the authors' method being even less strict, and RINGS being the least strict. This finding is important and worthy of publication in MMM, so zeolite scientists can rationalize discrepancies among various definitions of rings.

As such, I think this study may be publishable in MMM after the following detailed list of issues is considered by the authors. Also, the manuscript could use another round of careful proofreading to fix grammatical errors (missing words, incorrect usage of "less" vs. "fewer", etc.):

Author List: First and last author are the same, why is this?

The manuscript reports the authors correctly. The cover sheet is in error and will be corrected with this resubmission.

Comment #2:

Shortest-rings paragraph in Introduction: Would be helpful to explain why you are reviewing their work. What is the advantage of this approach? Why is it important enough to be discussed? And then, what is the disadvantage of this work, prompting you to engage in the present work. That is the background that lends context to your work, and can help explain the significance of your contribution. Your background is helpful, but not thoroughly stitched together to make a compelling story. Thank you.

We have taken the comments of the reviewer to heart and in response have significantly modified and sharpened the Introduction to more clearly explain the motivations for introducing the various definitions of rings within a framework. The changes are too large to be easily summarized here; rather, we refer the editor and reviewer to the revised manuscript marked with revisions.

Comment #3:

Methods: I wonder what "stereochemistry around a T site" means. It would be good if you had introduced / explained this concept in the introduction, and given some background to show why this is or may be important.

We have added a more specific definition of what the ring stereochemistry is referring to in our introduction to vertex symbols p. 5-6:

> *"Lastly, we consider the vertex symbol [17] characterization of T-sites and show that T-sites described by the same vertex symbol are not necessarily superimposable but rather can have distinct local stereochemistries. We identify and enumerate these cases of constitutional isomerization within T-sites of identical vertex symbol and propose an alternative convention that uniquely takes into account their orientation and connectivity around the T-site."*

We discuss this further in our results section on p. 20 where we provide examples of T-sites in various frameworks which have the same vertex symbol yet are shown visually to be constitutional isomers of each other due to variations in the connectivity of the rings making up the vertex symbol p. 19-21.

> *"The conventional definition of a vertex symbol, which lists pairs of rings along opposite edges of the tetrahedron from smallest to largest, only partially captures the stereochemistry around a T-site. As example, Fig. 16 illustrates the rings associated with T3, T9, and T1 or MOR, MON, and EON, respectively, all of which share the same $4{\cdot}5_2{\cdot}5{\cdot}8_2{\cdot}5{\cdot}8_2$ vertex symbol. While MOR T3 and EON T9 are similar, MON T1 differs in the location of the 52- and 4-MR edges, as highlighted in the lower panel of Fig. 16. To remove this ambiguity, an alternative approach is to list the rings of the vertex symbol by starting with the largest one and systematically adding the remaining rings following the edges of the tetrahedron such that each ring listed is connected to one of the oxygen atoms of the next ring listed. This provides a physical meaning to the order of the vertex symbol rings which accounts for differences in the connectivity of those rings around the T-site. Within this algorithm MOR T3 and EON T9 would be labeled as: $8_2{\cdot}8_2{\cdot}5_2{\cdot}5{\cdot}4{\cdot}5$, and MON T1 as: $8_2{\cdot}8_2{\cdot}4{\cdot}5{\cdot}5_2{\cdot}5$. The difference is subtle but highlights the distinct structural difference between the two types of T-sites that is not otherwise captured by a vertex symbol. We computed the ordered vertex symbols for all 1460 T-sites in the IZA Database. The 649 unique vertex symbols in the database increase to 666 unique ordered vertex symbols. Thus, the consideration of stereochemistry only modestly in- creases the space of unique tetrahedral sites. Table 4 lists some common T-site vertex symbols and associated ordered vertex symbols. Complete results are presented in the Supplementary Information."*

Comment #4:

Methods: awkward construction in "In this work we implement an efficient algorithm that was presented by Goetzke and Klein to find all the rings associated with a T-site that do not contain a shortcut [9] in a Python package called the Zeolite Simulation Environment (ZSE)" Sounds like some rings contain a shortcut in a python package.

The ordering of the sentence in Section 2.1 has been changed to reduce confusion:

> *"In this work we implement the efficient algorithm presented by Goetzke and Klein [9] in a Python package called the Zeolite Simulation Environment (ZSE) [17] to find all the rings associated with a T-site that do not contain a shortcut."*

Comment #5:
Fig. 5: This figure does little to explain how the algorithm used by the authors

This figure is taken from the original work of Goetzke and Klein. We believe it will be useful to those interested in the algorithmic elements of the work and prefer to leave as-is.

Comment #6:
Fig. 6: Is it possible that this example demonstrates the fundamental motivation for this work? If so, it can be moved much earlier in the narrative of this work, like into the Introduction itself.

Figure 6 does indeed represent an important motivation. We have integrated Figure 6 into Figure 3 and incorporated motivation into the revised Introduction:

> *"A comparison of the rings reported by various conventions to those listed in the IZA highlights differences related to fused rings illustrated in Fig. 3b. We propose a modification of the shortcut definition that results in a set of rings closer to those listed in the IZA database than do existing conventions."*

Comment #7:
Fig. 8: "We would argue that this is still a ring that provides an important topological descriptor of CHA because none of the tabulated rings provides information about the size of the CHA cage." This seems like the first time the authors suggest the motivating idea that rings should not only describe zeolite building blocks (for solid state chemists) and window dimensions (for reaction engineers) but also should describe zeolite cavity sizes (for understand sorption capacities). As a motivation, this should be mentioned much earlier in the narrative. Even so, I disagree with this motivation, as discussed above. However, I am just one zeolite scientist, and others may disagree with me. So I suggest we publish this paper in some revised form in MMM and let the field decide if this idea is good or bad.

We have added text to the first paragraph of the Introduction to further motivate practical interest in rings in zeolites:

> *"Rings can be used to describe the channel sizes for understanding shape selectivity in catalysis, zeolite building blocks for solid state chemists, and the sizes of framework cages and windows to provide insights into adsorption properties."*

Comment #8:
Above Sec. 3.2: "METHODS REFERENCE" What does this mean?

Thanks to the reviewer for catching this missing reference. The wording has been replaced with:

> *"as explained in Section 2.4"*

Comment #9:

Below Fig. 10: "drawback to using a ring convention based on connectivity and shortcuts is the exclusion of cycles that don't fit this definition" - very vague, not sure what this means.

The authors agree that this sentence can be reworded to better convey the point. The text in section 3.2 has been changed as shown below:

> *"One drawback to using a ring convention based on connectivity and shortcuts is the exclusion of non-ring cycles that exhibit geometric properties similar to those designated as rings."*

Comment #10:

Below Table 3: "This raises the question, if you want to ascertain chemical or physical properties about a T-site based on its ring count, and differentiate these T-sites from other similar but distinct T-site." I'd respond to this question with another question: Are you sure you should be using rings to distinguish T-sites? That is, using rings (topological structural descriptors) to distinguish complex, 3D geometrical objects like T-sites in their local environments seems like barking up the wrong tree.

We have added text to the Introduction We understand the reviewer's point. , and we have addressed it in the response to Comment #3 from Reviewer 1. It is possible to capture the subtle differences in complex 3D geometrical objects like T-sites, and we have shown examples of doing just that with the ordered vertex symbol. We have added text on p. 2 to illustrate the chemical significance of the rings of a T-site.

> *"Fig. 1 shows a cutout from the CHA framework, highlighting an 8-membered ring (8-MR) and a 4-MR filled in pink and green, respectively. As an illustration of the significance of those rings for a T-site, consider that the T-site labeled 20 is a Brønsted acid site, i.e., the site contains an Al. It will then have associated with it a charge-compensating proton that will bind to one of the four adjacent oxygen atoms. While the T-site is symmetrically identical to all other T-sites in CHA, the proton sites are inequivalent, distinguished by the dimensions of the rings that pass through them. These differences lead to differences in proton energies and vibrational frequenices, a distinction that is experimentally observable in infrared spectra [6]."*

Characterization and Analysis of Ring Topology of Zeolite Frameworks

Jerry T. Crum[a], Justin R. Crum[b], Cameron Taylor[a], William F. Schneider[a,c]

[a]Department of Chemical and Biomoledcular Engineering, University of Notre Dame, 250 Nieuwland Science Hall, Notre Dame, IN 46556, USA
[b]Department of Applied Mathematics, University of Arizona, 617 N Santa Rita Ave, Tucson, AZ 85721, USA
[c]Department of Chemistry and Biochemistry, University of Notre Dame, 251 Nieuwland Science Hall, Notre Dame, IN 46556, USA

Abstract

The topology of zeolite frameworks and of associated tetrahedral sites (T-sites) are commonly characterized by their associated rings, typically defined as some set of closed paths or cycles through a framework that cannot be decomposed into shorter cycles. These ring descriptors have been used to identify feasible zeolite topologies, to describe the similarity and differences between zeolites, to identify sites or voids of catalytic relevance, and as machine learning fingerprints. Numerous definitions and algorithms for finding zeolite rings have been proposed and applied throughout the literature. Here we report an analysis of rings and T-sites in a large number of zeolite frameworks using Zeolite Simulation Environment, a Python package that implements an efficient algorithm presented by Goetzke and Klein for finding rings in arbitrary frameworks. We compare the result of a number of common and new ring definitions applied to a large number of common zeolite frameworks. We discover previously unrecognized rings in a number of frameworks. We show that the vertex symbol, a common approach used to characterize T-sites, misses important parts of the stereochemistry around a T-site, and propose an alternative definition. This tool provides an effective platform for characterizing zeolite and T-site structures useful for building models and doing machine learning.

1. Introduction

Zeolites are three dimensional crystalline structures containing tetrahedral Si or Al atoms connected by oxygen bridges. The International Zeolite Association (IZA) lists over 200 known zeolite frameworks that are described by dimensionality, pore shape and size, and Si/Al ratios [1]. It is natural to characterize differences in zeolites by the size and shapes of the features present in the crystal. Rings are one common type of feature widely reported and used, both to characterize a zeolite crystal and as descriptors of the individual tetrahedral sites (T-sites) of a zeolite. Rings have been used to identify feasible zeolites

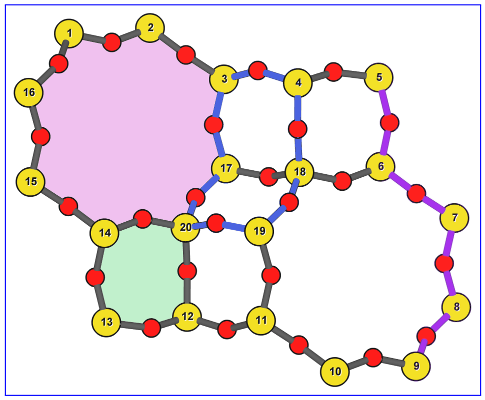


Figure 1: Cutout of the Chabazite framework showing a path (5-6-7-8-9) highlighted with purple bonds, a cycle (3-4-18-19-20-17) highlighted with blue bonds, an 8-MR filled in with pink, and a 4-MR filled in with green. Yellow ~~atoms~~ and red spheres are ~~Si (~~T-sites (typically Si) ~~,~~ and ~~red atoms are~~ oxygen, respectively.

topologies [2], to describe the similarity and differences between zeolites [3, 4], to identify sites or voids of catalytic relevance [5, 6], and as machine learning finger prints of physicochemical properties [7]. Rings can be used to describe the channel sizes for understanding shape selectivity in catalysis, zeolite building blocks for solid state chemists, and the sizes of framework cages and windows to provide insights into adsorption properties. The dimension and number of rings in a framework can be used to characterize entire zeolite frameworks, T-sites that make up these frameworks, or even the oxygen atoms that connect the T-sites. The number and dimension of rings that pass through a T-site can be used to describe and differentiate the local void environment around symmetry distinct T-sites and to differentiate symmetry distinct oxygen atoms.

Fig. 1 shows a cutout from the CHA framework, highlighting an 8-membered ring (8-MR) and a 4-MR filled in pink and green, respectively. As an illustration of the significance of those rings for a T-site, consider that the T-site labeled 20 is a Brønsted acid site, i.e., the site contains an Al. It will then have associated with it a charge-compensating proton that will bind to one of the four adjacent oxygen atoms. While the T-site is symmetrically identical to all other T-sites in CHA, the proton sites are inequivalent, distinguished by the dimensions of the rings that pass through them. These differences lead to differences in proton energies and vibrational frequenices, a distinction that is experimentally observable in infrared spectra [6].

~~Many researchers have presented methods to define and count the rings present in zeolites [8–14]. Zeolites can naturally be~~ A number of definitions and enumerations of rings in zeolites have been presented [8–14] and implemented in software tools, including ZeoTSites [14] and ToposPro [15]. Zeolites are naturally represented using graph theory, ~~where~~ in which atoms are nodes ~~,~~ and bonds are edges. ~~In some cases researchers will use T-sites as nodes and oxygen atoms as edges, since~~ Because the oxygen atoms are only connected to

Table 1: List of graph based features, their descriptions, and how they apply to frameworks, T-sites, and oxygen atoms.

heightFeature	Description	Application to Frameworks	Application to T-sites	Application to Oxygen Atoms
Node	T-site or oxygen atom	Contains some set of symmetry distinct nodes		
Path	A sequence of nodes that are connected to each other	Traverses through a framework	T-sites and oxygen atoms connected in series forms a path	T-sites and oxygen atoms connected in series forms a path
Cycle	A path that starts and ends at the same node, and does not repeat any other nodes	Contains some set of cycles	Formed by alternating connected T-sites and oxygen atoms	Formed by alternating connected T-sites and oxygen atoms
Rings	A cycle that does not contain a shortcut [8, 9]	Can describe channels of a framework, or other distinct void environments	Can be described by counting the number of rings that pass through it	Can be described by the number of rings that pass through it
Modified Shortcut Rings	A cycle that does not contain a modified shortcut	A subset of the rings will be found in the framework	Used to describe a T-site	Used to describe an oxygen atom
Shortest Path Rings	A ring that is the shortest ring for at one least one set of O-T-O along the cycle	Another subset of the rings of a framework	Used to describe a T-site	Used to describe an oxygen atom
Vertex Symbol Rings	The rings making up the vertex symbol of a T-site	Yet another subset of the rings of the framework	Used to describe a T-site	Not applicable to oxygen atoms

exactly two T-sites~~[9]. Some basic definitions of graphs that will be used going forward are highlighted in a~~ , their inclusion in a graph does not change the graph structure, and thus T-sites can be defined as nodes and the intervening oxygen atoms as edges [9]. The cutout of the chabazite (CHA) framework shown in Fig. 1 ~~, and described in Table 1~~ and Table 1 summarize basic definitions of graphs that will be used here. The size of each of these features can either be defined by the total number of atoms contained in them, or by the number of T-sites contained in them. The latter method is the standard convention, which we will use. A path is a sequence of edges ~~which joins~~ that connects a sequence of nodes in which every edge is distinct (shown with purple highlighted bonds in Fig. 1). A cycle is a path that starts and ends at the same node, and does not repeat any other nodes (shown by blue bonds in Fig. 1). ~~In the most basic form a ring is a cycle that does not contain any shortcuts [8, 9]~~A zeolite crystal contains an infinite number of cycles, but the symmetry of the crystal reduces this to a finite number of symmetry-distrinct rings whose exact identity is subject to definition. Fig. 1 shows an ~~8-membered ring~~ 8-MR and a 4-MR filled in ~~with~~ pink and green, respectively.

In the most basic definition, a ring is any cycle that does not contain any shortcuts [8, 9]. A shortcut is defined as a path connecting two nodes of a cycle that decomposes the cycle into smaller rings [8, 9]~~, or in other words, the~~ ; i.e., a path is a shortcut if it is shorter than both paths between these nodes along the cycle. An example is the path connecting T17 and T18 in Fig. 1, which is a shortcut of the blue cycle. This path contains two T-sites, splitting the 8-MR cycle into two 4-MRs. ~~Many definitions build off this idea that rings do~~

~~not contain shortcuts. In a zeolite, rings are used to define the topology of the framework, such as the channels and void environments that are present. Counting the rings that pass through a T-site can also be used to describe and differentiate the local void environment around symmetry distinct T-sites. Like T-sites, ring counts can be used to differentiate symmetry distinct oxygen atoms.~~ While general, this definition reports, as distinct, ring features of a zeolite framework that do not intuitively relate to local structure, accessibility, or usefully “fingerprint” frameworks. Thus, more restrictive defintions of zeolite rings have been proposed. For instance, a “strong ring” is defined as a ring that is not the sum of a set of smaller rings [9, 16].

~~Cutout of the CHA framework, showing the rings that make up the vertex symbol of the single symmetry distinct T-site. a) Example of the tetrahedron formed by the T-site and four connected oxygen atoms, with labeled edges of the tetrahedron. b) Rings associated with opposite edges E1 and E3. c) Rings associated with opposite edges E5 and E6. d) Ring associated with opposited edges E2 and E4. Rings are colored as: 4-MR (green), 6-MR (blue), and 8-MR (pink).~~

~~Vertex symbols are~~ Another method for pruning the list of rings found in a zeolite frameworks is based off of the concept of a vertex symbol. The vertex symbol is a common way to describe the structure around the T-sites of ~~a zeolite framework, and were~~ zeolite frameworks and was first used to do so in 1997 by O’Keeffe and Hyde ~~in 1997~~ [17]. The vertex symbol of a T-site ~~contains~~ is determined by the shortest rings associated with each of the six edges of the tetrahedron formed by the T-atom and its four ~~bounded~~ bound oxygen atoms. The ~~ring sizes, and their multiplicity, for opposite edges (edges~~ symbol is formed by grouping rings on opposite edges of the tetrahedron ~~that do not connect to the same oxygen atom) are grouped together. These grouped pairs are listed~~ and listing from smallest to largest~~forming the vertex symbol~~. Fig. 2(a) shows ~~an example of~~ as example the tetrahedron formed at the single ~~symmetry distinct~~ symmetry-distinct T-site in the CHA framework. The edges of the tetrahedron are labeled to aid in identification. Fig. 2(b-d) show the rings associated with each opposite pair of edges. The vertex symbol of T1 in CHA is thus ~~able to be determined as~~ $4{\cdot}4{\cdot}4{\cdot}8{\cdot}6{\cdot}8$. For a T-site that contains a multiplicity of rings at one edge, that multiplicity would be represented as a subscript in the vertex symbol. An example would be the vertex symbol of $4{\cdot}6_{\cancel{2}\,2}{\cdot}6{\cdot}6_{\cancel{3}\,3}{\cdot}6_{\cancel{2}\,2}{\cdot}6_{\cancel{3}\,3}$ for T1 in AFI. The rings of a framework are then the union of the rings of all the distinct T-sites, a definition that eliminates some rings that satisfy the no shortcut criterion.

~~Another ring counting convention presented by~~

Sastre and Corma ~~is to count only the shortest path connecting an O-T-O~~ [18] proposed an alternative convetion, in which distinct rings are only those that form the shortest cycle connecting any O−T−O in a framework. ~~With this definition, they can find and count all the rings in a framework that are the shortest path for at least one set of O-T-O along the cycle [18]~~This convention provides an univocal count of the rings that pass through a T-site without the ambiguity of limiting the search to a certain ring size. Using AFI as an

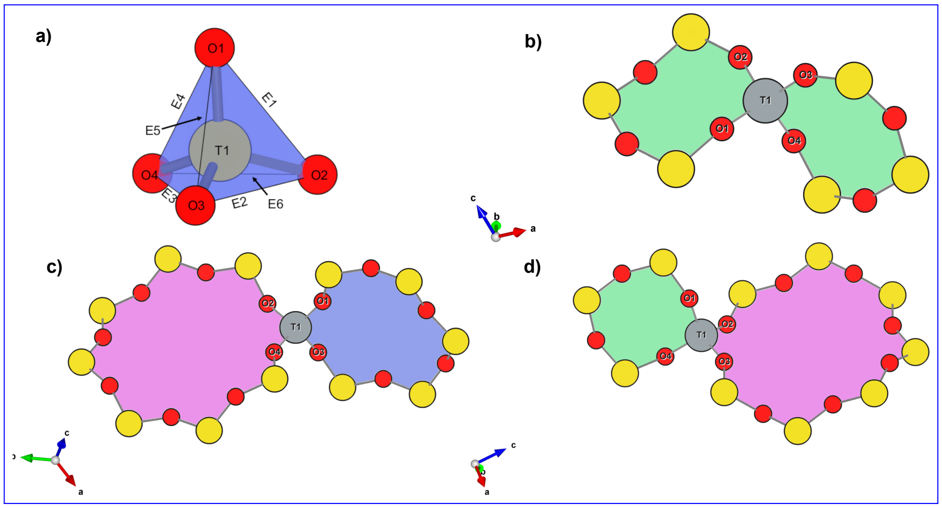


Figure 2: Cutout of the CHA framework, showing the rings that make up the vertex symbol of the single symmetry distinct T-site. a) Example of the tetrahedron formed by the T-site and four connected oxygen atoms, with labeled edges of the tetrahedron. b) Rings associated with opposite edges E1 and E3. c) Rings associated with opposite edges E5 and E6. d) Ring associated with opposited edges E2 and E4. Rings are colored as: 4-MR (green), 6-MR (blue), and 8-MR (pink).

example~~again (Fig. 3), we see~~ , (Fig. 3(a)) shows that both a 12-MR (purple) and a 6-MR (blue) pass through the labeled T1 atom. The 12-MR would not be included in the Sastre and Corma enumeration because for every ~~O-T-O~~ O−T−O along the 12-MR, the shortest path connecting them is not the 12-MR. The difference between this ~~method~~ criterion and the vertex symbol rings is subtle, but with this shortest path convention any ring belonging to the vertex symbol of any T-site in a framework will be included in the ring count for each of the T-sites that ring passes through. Fig. 4 shows a cutout of the TON framework including a 6- (blue) and 10-MR (orange). For T1, only the 6-MR would be counted ~~in~~ within the vertex symbol convention because it is the shortest path connecting O2 and O3. The 10-MR is ~~part of the vertex symbol for T3 because it is the~~ the shortest path connecting O2~~and O14. Since~~ , and since this 10-MR is the shortest path for at least one set of ~~O-T-O along the ring~~O−T−O along the cycle, and passes through T1, it does get counted in the shortest path rings for T1.

~~We can use ring counts to characterize entire zeolite frameworks, T-sites that make up these frameworks, or even the oxygen atoms that connect the T-sites. Since various conventions exist that can reduce the set of rings in a zeolite to more strictly defined properties, the ring counts returned by the various conventions will differ . Differences in ring counts leads to differences in how we might describe the topological environment of a zeolite. Therefore, when using rings to determine the properties of a framework, T-site, or oxygen atom, it is important to know the difference in the conventions, and use one~~

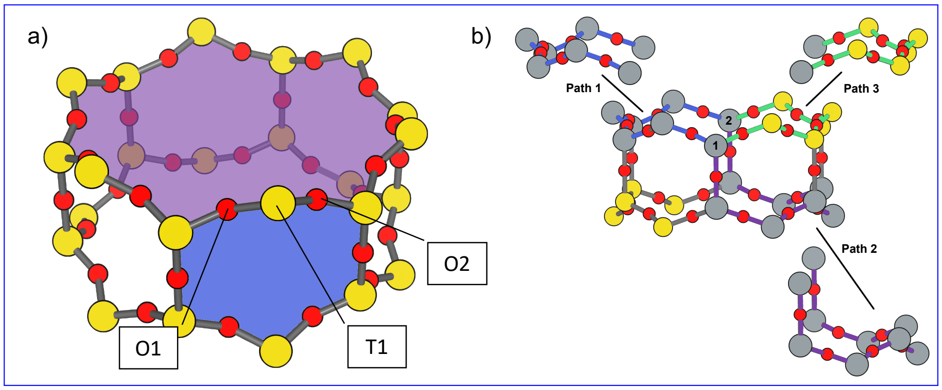


Figure 3: ~~Cutout~~ Cutouts of the 12-MR channel in AFI~~highlighting~~: a) Highlighting a 12-MR in purple, and a 6-MR in blue. The 6-MR is included in the vertex symbol of labeled T1 because it is the shortest path connecting O1 and O2. The 12-MR would not be included in the vertex symbol or shortest path ring list because for each O-T-O along the 12-MR there is a shorter path connecting them. b) A 14-MR is shown as T-sites replaced with aluminum atoms in gray. The two paths connecting Al1 and Al2 that make this 14-MR are highlighted with blue and purple bonds. Path 3 highlighted with green bonds is a modified shortcut connecting Al1 and Al2.

~~that determines the features of interest.~~

Thus, the dimensions and numbers of rings present in a framework will differ depending on precise convention, with potential consequences for the ability to relate properties to rings. Here we present an analysis of rings captured by Goetzke and Klein's efficient ring finding algorithm [9] ~~,~~and compare those rings to the rings found by other previously published ring set reduction conventions. We have implemented all of these ring finding conventions in a Python package called the Zeolite Simulation Environment (ZSE~~,)~~) [19]. We use ZSE to ~~provide an analysis~~ analyze the sets of rings captured by each convention ~~for~~ across the entire set of zeolite frameworks contained ~~on~~ in the IZA Database [1]~~to compare how these sets of rings provide different characterizations of said frameworks~~. We highlight rings that are found by these conventions but not typically discussed in the literature for a number of frameworks. ~~We also show that the vertex symbol, a common approach used to characterize T-sites [17], based on the shortest rings connecting the neighboring oxygen, misses important parts of the stereochemistry around a T-site. Finally, we provide an alternative method for listing~~ A comparison of the rings reported by various conventions to those listed in the IZA highlights differences related to fused rings illustrated in Fig. 3 b. We propose a modification of the shortcut definition that results in a set of rings closer to those listed in the IZA database than do existing conventions. Lastly, we consider the vertex symbol ~~rings that~~ [17] characterization of T-sites and show that T-sites described by the same vertex symbol are not necessarily superimposable but rather can have distinct local stereochemistries. We identify and enumerate these cases of constitutional isomerization within T-sites of identical vertex symbol and propose an alternative convention that uniquely

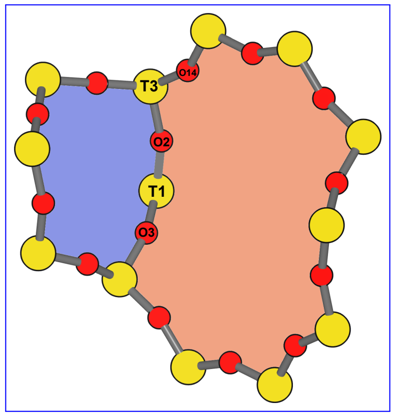


Figure 4: Cutout of the TON framework showing a 6- (blue) and 10-MR (orange). The 10-MR is the shortest path connecting O14-T3-O2, and passes through T1, so it is counted in the shortest path rings for T1.

takes into account their orientation and connectivity around the T-site. ~~.~~

2. Methods

2.1. ~~Finding~~ Rings That ~~do not~~ Do Not Contain Shortcuts

~~In this work we implement an efficient algorithm that was~~ We implemented the efficient algorithm presented by Goetzke and Klein [9] in ZSE [19] to find all the rings associated with a T-site that do not contain a shortcut~~[9] in a Python package called the Zeolite Simulation Environment (ZSE)~~. In ZSE we use the framework put in place by the Atomic Simulation Environment (ASE) [20] to handle routine analysis zeolite crystal structures. All graph theory functions are performed ~~used~~ using the NetworkX Python package [21].

First, we convert the ASE atoms object into a connectivity matrix which represents every atom across the columns and rows. If two atoms are bound together, their respective entry in the connectivity matrix contains a 1, else a 0. This connectivity matrix is then converted to a NetworkX graph object, and then a distance dictionary using NetworkX built in functions. Then we implement Step 3 from Geotzke and Klein's algorithm [9] summarized here: to find the rings that pass through a T-site, we iteratively search for every size ring between 3-MR and a maximum ring value that is ~~user specified~~user-specified. For this work we set a cutoff of 18-MRs. A schematic showing the evolution of the ring search is shown in Fig. 5. For ring size λ we start at the T-site of interest (labeled 1)~~Fig. 5~~, and search the distance matrix for any T-sites that are $\lambda/2$ (even λ) or (λ-1)/2 (for odd λ) distance from the starting T-site (labeled 2). Next we attempt to create ~~to~~ two distinct paths from $1 \rightarrow 2$ and from $2 \rightarrow 1$ alternating adding a node to each path as indicated by Fig. 5. Each node added to each of the paths must be $\lambda/2$ (even λ) or (λ-1)/2 (odd λ) from the head of the other path. Also each node added to each path needs to be the correct distance from 1 and 2 for the given step respectively. If either of the previous two conditions are not met, a ring cannot be formed of length λ along the given paths, we backtrack and repeat until all possible options have been explored for λ. Then we increase λ and continue until the cutoff ring size is completed.

2.2. ~~Finding~~ Vertex Symbol Rings

Starting from the set of all rings found in Section 2.1, we ~~can~~ prune the ring list to the set of vertex symbol rings. We find the shortest ring in the set that connects each pair of oxygens bound to our initial T-site. It is possible for there to be multiple rings of the same size connecting each oxygen, in which case all the rings of that size are kept.

2.3. ~~Finding~~ Shortest Path Rings

~~Here we~~ We prune the set of all rings from Section 2.1 to a subset of rings the meets the shortest path definition published by Sastre and Corma [18]. For each ring, we iterate over every group of ~~O-T-O~~ O–T–O atoms in the ring, and check ~~if~~ whether this ring is the shortest path connecting the two oxygen atoms. If so, the loop is broke, because the ring need only be the shortest path for one group of ~~O-T-O~~ O–T–O atoms to fit the definition. This is the most time consuming process out of all the ring finding conventions we have implemented.

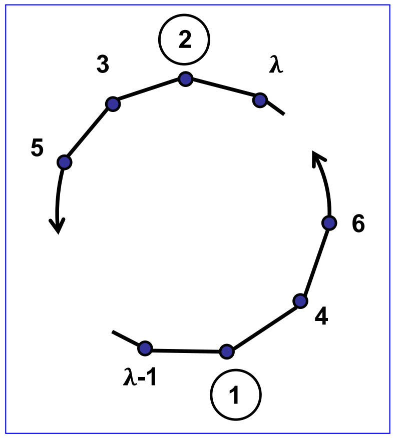


Figure 5: Diagram showing how the ring finding algorithm evolves. Adapted from Goetzke and Klein [9].

2.4. Modified Shortcut Rings

~~In this work we present a modified definition of shortcut to capture a different subset of rings from any of the other ring finding conventions. Traditionally a shortcut is a path connecting two nodes of a cycle that decomposes the cycle into smaller rings [8, 9]. We propose that a shortcut is a path connecting two nodes of a cycle that decomposes the cycle into at least one smaller ring. This definition does not require that the shortcut between two nodes be shorter than both paths connecting those nodes along the cycle. Our shortcut only needs~~ The shortest path convention generally finds fewer (and smaller) rings than does an all non-shortcut bearing ring convention. As we show below, these two conventions bracket the set of rings captured in the IZA Structure Database, the all-rings convention generally including even larger rings and the shortest path convention missing large rings. In seeking a convention of intermediate stringency, we identified a convention that would require the shortcut to be shorter than one, rather than both, of the paths connecting ~~those~~ two nodes along the ~~cycle.~~

~~This definition is explained graphically in **??**, where we present a cutout of the AFI framework showing a portion of the 12-MR channel. There is a~~ ring. This modified shortcut definition is illustrated in Fig. 3(b). By the standard shortcut convention, AFI would be considered to contain a 14-MR ~~that traverses seven T-sites in each of the 12-MRs, through the combination~~ formed from the union of Path 1 (blue) and Path 2 (purple) ~~. With the classical definition of a shortcut, this cycle is considered a ring. However, Path 3 (in green) connecting~~ through T1 and T2. However, because T1 and T2 ~~produces a 12-MR when combined with Path~~ are also connected by Path 3, which, while the same length as Path 1, ~~making Path 3 a modified shortcut . This~~ is shorter than Path 2, the 14-MR ~~would not be counted under our new definition~~does not satisfy the modified shortcut criterion. As we show below, the modified shortcut rule identifies rings of a given zeolite that are closer to those identified in the IZA Structure Database as rings of interest than do either all-ring or shortest-path-ring rules.

~~Cutout of the AFI framework showing two stacked 12-MRs from the channel. A 14-MR is shown as T-sites replaced with aluminum atoms in gray. The two paths making connecting Al 1 and Al2 that make this 14-MR are highlighted with blue and purple bonds. Path 3 highlighted with green bonds is a modified shortcut connecting Al1 and Al2.~~

~~To~~ Algorithmically, to remove rings containing modified shortcuts from the full set of rings, we iterate over every T-site pair of the ring and check for the shortest path connecting them. If that shortest path is shorter than either of the two paths along the ring connecting the two T-sites, we check ~~if~~ whether the combination of this shorter path and the shorter of the two ring paths forms a new smaller ring. If so the iteration is broken, and the ring is removed from the counted set.

2.5. Ordered Vertex Symbols

To add information about the spatial orientation of the rings around a T-site to the vertex symbol, we have developed a method to order the edges in the vertex symbol. We systematically list the rings by following the edges of the tetrahedron such that each ring listed is connected to one of the oxygens of the next ring listed. After removing all the rings that are not a part of the vertex symbol (Section 2.2) we use the following process to order them.

1. List all the possible arrangements of the oxygens bound to the T-site (4! = 24 possible arrangements).
2. Use a predetermined order of edges: [[0,2],[0,1],[1,2],[2,3],[3,0],[1,3]].
   (a) Where each of those values represents the index of the oxygen to use.
3. Find the ring size (and multiplicity) connecting each pair of oxygens in this predetermined order.
4. Make a list of weights, where for each pair of oxygens the weight is the ring size $\times$ multiplicity.
5. Reverse sort the list of all possible oxygen arrangements by the correlating list of weights.
6. Use the first oxygen arrangement coupled with the predetermined edge order to list the rings and multiplicity for each edge.

2.6. Determining All Ring Sizes Contained in a Zeolite Framework

Finally, to determine all the ring sizes exhibited with in a zeolite framework, we take advantage of T-site symmetry. The rings of a framework are made of T-sites, and if two T-sites are symmetrically identical they will have the same set of rings passing through them. Therefore, we only need to find the rings associated with each symmetry distinct ~~T-sites~~ T-site to know all the possible ring sizes within a framework. For example, AFI only contains one symmetry distinct T-site (T1). Using the basic definition of a shortcut, T1 is a part of 4-, 6-, 12-, and 14-MRs when using a cutoff of 18-MR. Every other T-site in the AFI framework is also a T1, thus the only possible ring sizes in AFI are 4-, 6-, 12-, and 14-MRs.

3. Results

3.1. Characterizing Rings in a Zeolite Graph

The IZA Database [1] is a common reference used to identify all the rings in a zeolite framework, however it only lists the rings that define a channel (ex: 12-MR in AFI), or rings associated with the symbol of a T-site. These rings listed by the IZA are referred to as tabulated rings in the literature [3]. In some frameworks, other rings (cycles not containing shortcuts) exist that are not included in the list of tabulated rings. These ~~’untabulated’~~ “untabulated” rings may still provide important topological information about a zeolite framework, or the local void environment around a T-site. Fig. 6 shows counts of frameworks containing each size ring from 3- to 18-MR using the Goetzke algorithm and the listed rings on the IZA database [1]. There are slight differences in the counts up to 6-MRs, but the main divergence takes place ~~as we get to ring sizes >6-MR.~~ at ring sizes greater than six. We limit the search to 18-MR and smaller because the differences in ring finding conventions are captured within the smaller ring sets, because probabilities of occurrence become small with increasing ring size [2], and for computational expediency.

Taking a closer look at some of these untabulated rings ~~,~~ highlights rings not typically listed for some frameworks, but are still relevant to describing their topology. Using CHA as an example, Fig. 7 displays a 12-MR (in purple) that ~~exists in CHA that circumferences the~~ circumferences the CHA cage. This ring is not associated with the vertex symbol of the single symmetry distinct T-site in CHA and does not define a channel. Thus, this ring is not included in the list of tabulated rings. We would argue that this is still a ring that provides an important topological descriptor of CHA because none of the tabulated rings ~~provides~~ provide information about the size of the CHA cage.

Using AFI as another example, we find another type of ring that arises from traversing a pair of stacked rings and is not included in the list of tabulated rings. AFI, like CHA, contains one symmetry distinct T-site. According to the IZA, the AFI framework contains 4-, 6-, and 12-MRs [1]. When we search for rings using the Goetzke algorithm [9], we also find that it contains 14-MRs created by using seven T-sites from two 12-MRs that are separated by a distance of one oxygen (Fig. 8). Rings of this nature are prevalent in many frameworks~~,~~ ; another example can be seen in the bottom right of Fig. 7, where an 8-MR is highlighted traversing the two 6-MRs of the D6R. These types of rings may not be of interest depending on which topological feature one intends to describe. This has led us to create a modified definition of a shortcut as explained in ~~METHODS REFERENCE~~Section 2.4, which excludes these types of rings. The benefit of this new shortcut definition is that larger rings that are missed by the vertex symbol or shortest path rings (i.e., 12-MR in AFI) are still captured,

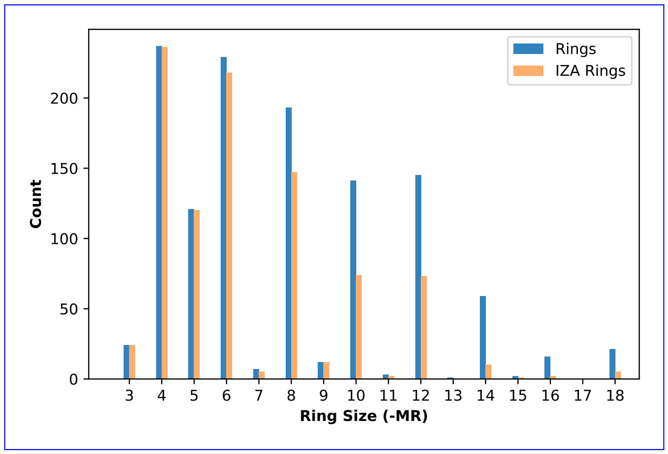


Figure 6: Counts of IZA frameworks containing each size ring between 3- and 18-MR using the Goetzke algorithm and the tabulated rings listed by the IZA [1].

while excluding rings that arise from convolution of ~~staked~~ stacked rings found with the typical shortcut definition.

3.2. Characterizing Frameworks by Rings

With the addition of our modified shortcut definition, we ~~have~~ can compare the results of four ring finding conventions to ~~compare, as well as including~~ the tabulated rings ~~from~~ listed in the IZA Database. Fig. 9 shows how many frameworks contain each size ring found using the various ring counting conventions from 3- to 18-MRs. This plot highlights the differences in the conventions and shows that a topological description of a framework based on rings will depend on the way that you define a ring. In general, a hierarchy of ring sizes found by each convention is: all rings not containing a shortcut >this work >shortest path rings >vertex symbol rings. ~~While the~~ The IZA listed rings ~~includes~~ include the vertex symbol rings, and a selection of general rings [1].

One drawback to using a ring convention based on connectivity and shortcuts is the exclusion of ~~cycles that don't fit this definition, but still exhibit properties like~~ non-ring cycles that exhibit geometric properties similar to those designated as rings. This is a trade-off between well-defined connectivity rules, and the inclusion of particular void environments that may still have important applications. These shortcut containing cycles can display chemical and/or geometric properties consistent with rings and are of interest to catalysis researchers even though they are not classically considered rings. One example is the 6-membered cycle referred to as the α-6-MR in literature (Fig. 10) and is present in a number of frameworks including but not limited to MOR, FER, MFI, and BEA [22, 23]~~, which~~. This α-6-MR is a potential location for Co~~$^{2+}$~~$^{2+}$ uptake when two Al atoms are 3rd nearest neighbor (NN) in the cycle [23, 24]~~. Similar to Co^{2+}~~, similar to Co^{2+} uptake at 3NN Al atoms in 6-MRs in other frame-

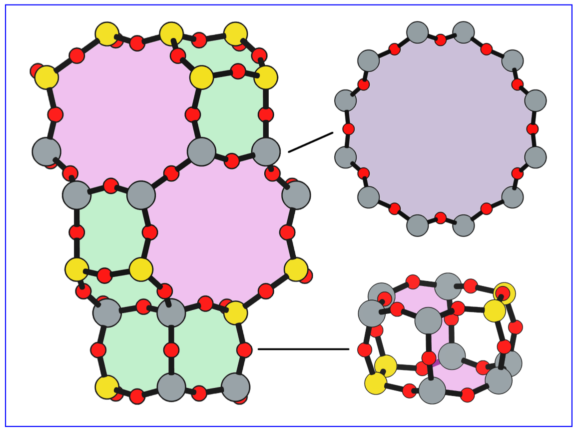


Figure 7: Chabazite cage and double 6-MR (D6R) with highlighted rings: 4-MR in green, 8-MR in pink, and 12-MR in purple. The 8-MR in the D6R and the 12-MR are rings not typically discussed in literature. Si atoms have been replaced with Al atoms to help identify those rings in the overal cage strcuture.

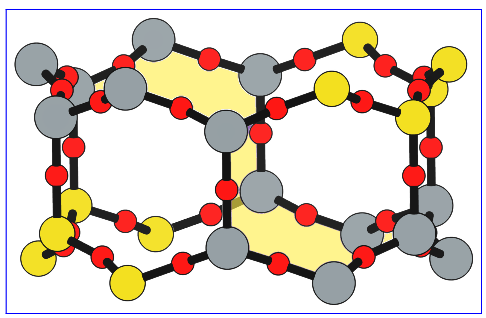


Figure 8: Cutout of the 12-MR channel in AFI with a 14-MR (yellow) traversing seven T-sites of each 12-MR. The T-sites comprising the 14-MR have been replaced with Al for visibility.

works such as CHA [25]. Fig. 10 shows that this particular structure would be considered two 5-MRs using connectivity rules based on a shortcut.

3.3. Characterizing T-sites by Rings

Considering that zeolite frameworks are comprised of one or more symmetry distinct T-sites, it may be of interest to describe those T-sites by the rings that pass through them. Most often the vertex symbol is used to make such a classification [17]. Sastre and Corma also provided characterization of T-sites using the shortest path rings that pass through them [18]. In their work, they presented the ring index, which lists all the rings passing through a T-site from smallest to largest, and a subscript for each size representing it's multiplicity. The rings associated with a T-site can ~~provided~~ provide information about the local void environments around the T-site, and could potentially be correlated

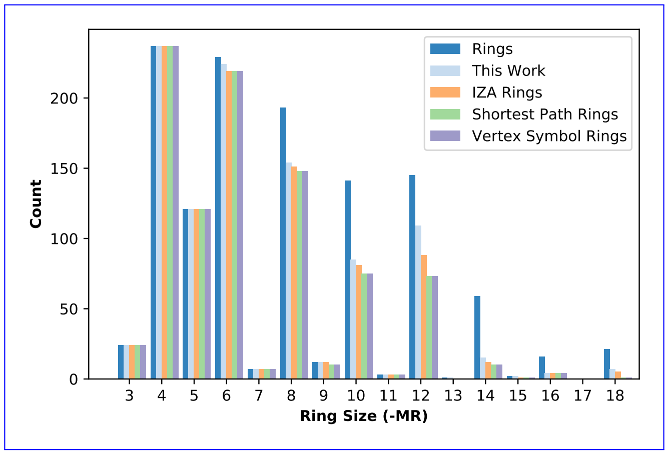


Figure 9: Number of IZA zeolite frameworks containing each size ring, using the various ring counting conventions, as well as the rings listed by the IZA Database [1].

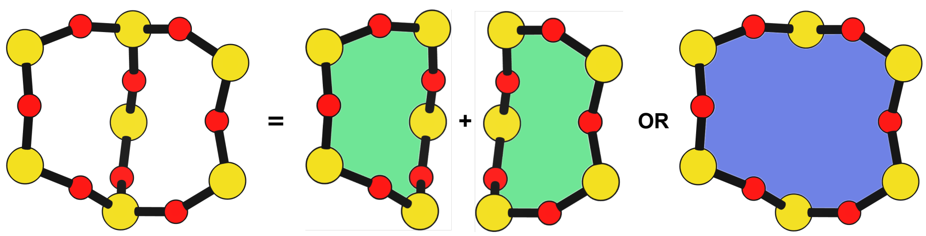


Figure 10: Cutout of MFI framework showing the structure referred to as an α-6-MR in blue, and the two 5-MRs that compose it in green. The 6-membered cycle would not be found by any of the connectivity ring rules outlined in this work.

to other physicochemical properties of the T-site once ~~enough once~~ sufficient data on those physicochemical properties exists.

Take for example the AFI framework, containing one symmetry distinct T-site. AFI contains 4-, 6-, 12-, and 14-MRs. To describe that T-site we can count how many of each of those rings pass through the T-site. We can also prune this list using our modified definition of a shortcut, the shortest path rings definition [18], or the rings contained within the vertex symbol of this T-site [17]. Using the ring index outlined above, each of these conventions will provide a different description of the zeolite (highlighted in Fig. 11):

- Rings: $4{\cdot}6_{\cancel{13}\,13}{\cdot}12{\cdot}14_{\cancel{7}\,7}$
- ~~This work~~Modified Shortcut Rings: $4{\cdot}6_{\cancel{13}\,13}{\cdot}12$
- Shortest Path Rings: $4{\cdot}6_{\cancel{13}\,13}$
- Vertex Symbol Rings: $4{\cdot}6_{\cancel{11}\,11}$

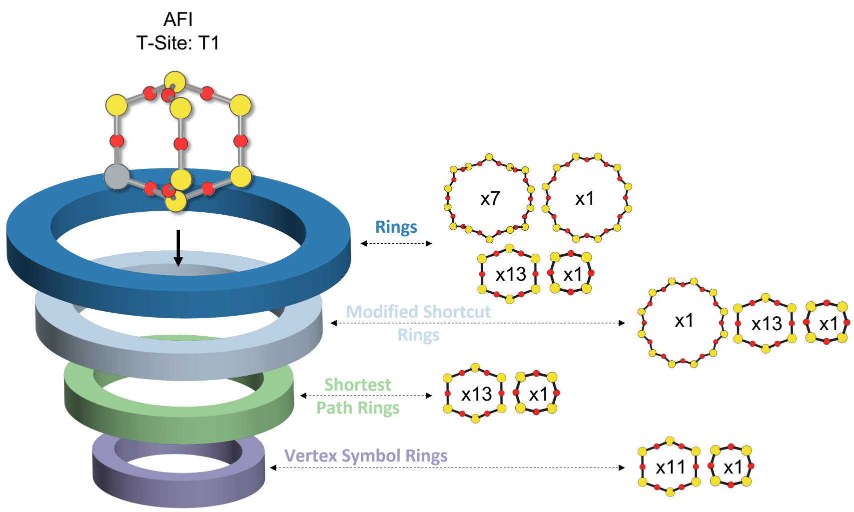


Figure 11: ~~Diagram shwoing the ring~~ Ring counts of each size ring that pass through the single symmetry distinct T-site in AFI for each of the various ring finding conventions.

With an understanding of how we characterize ~~T-site~~ T-sites by counting the rings that pass through them, Table 2 shows the ring index for a selection of T-sites from uninodal (containing only one symmetry distinct T-site) frameworks. This table highlights the differences in ~~rings~~ ring counts found with each convention, and shows that in general as you move from left to right across the table the largest ring size found decreases. The results in the shortest path column were found using ZSE, but agree directly with the results shown by Sastre and Corma [18]. The results in the vertex symbol rings column were also found with ~~zse~~ZSE, and agree directly with the vertex symbols listed on the IZA Database website [1].

Next, we take an in-depth look at the ring counts for a framework with multiple symmetry distinct T-sites, to show how a ring index can provide information about the local environment around a T-site, and help differentiate them. MOZ is a zeolite framework containing 4-, 6-, 8-, 10-, 12-, 14-, and 18-MRs, 6 symmetry distinct T-sites, and two distinct 12-MR channels. Table 3 shows the ring index for each T-site using each ring finding method.

Fig. 12 shows the T-site locations inside a 2-dimensional view of the framework. If you were interested in which T-sites have access to the 12-MR channels, the shortest path rings and vertex symbol rings would only suggest T3 participates in the 12-MR rings. However, ~~all rings and this work~~ the all rings convention and the modified shortcut convention both identify T4 and T6 as participating in the 12-MR channels as highlighted in Fig. 12.

We next used ZSE to find the rings associated with every symmetry distinct

Table 2: Comparison of ring indices for the T-sites in various uninodal zeolite frameworks.

Framework	Rings	This Work	Shortest Path Rings [18]	Vertex Symbol Rings [1]
ABW	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{3}\ 3}\cdot 8_{\cancel{2}2}$
ACO	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{3}\ 3}\cdot 8_{\cancel{6}\ 6}\cdot 10_{\cancel{15}\ 15}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{6}\ 6}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{6}\ 6}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{6}6}$
AFI	$4_{\cancel{1\ 13}\ 1\cdot 13}\cdot 12_{\cancel{1}\ 1}\cdot 14_{\cancel{7}\ 7}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{13}\ 13}\cdot 12_{\cancel{1}\ 1}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{13}\ 13}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{11}11}$
ANA	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{16}\ 16}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{16}\ 16}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{16}\ 16}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{8}8}$
ATO	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{9}\ 9}\cdot 8_{\cancel{8}\ 8}\cdot 12_{\cancel{20}\ 20}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{9}\ 9}\cdot 12_{\cancel{20}\ 20}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{9}\ 9}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{9}9}$
BCT	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{6}\ 6}\cdot 8_{\cancel{20}\ 20}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{6}\ 6}\cdot 8_{\cancel{12}\ 12}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{6}\ 6}$	$4_{\cancel{1}\ 1}\cdot 6_{\cancel{6}6}$
CHA	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{6}\ 6}\cdot 12_{\cancel{1}\ 1}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{2}\ 2}\cdot 12_{\cancel{1}\ 1}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{2}\ 2}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{2}2}$
DFT	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{6}\ 6}\cdot 8_{\cancel{10}\ 10}\cdot 10_{\cancel{10}\ 10}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{6}\ 6}\cdot 8_{\cancel{10}\ 10}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{6}\ 6}\cdot 8_{\cancel{10}\ 10}$	$4_{\cancel{2}\ 2}\cdot 6_{\cancel{4}\ 4}\cdot 8_{\cancel{6}6}$
GIS	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}4}$
GME	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{6}\ 6}\cdot 12_{\cancel{7}\ 7}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{2}\ 2}\cdot 12_{\cancel{1}\ 1}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{2}\ 2}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{1}\ 1}\cdot 8_{\cancel{2}2}$
MER	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}\cdot 10_{\cancel{10}\ 10}\cdot 14_{\cancel{14}\ 14}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}\ 4}$	$4_{\cancel{3}\ 3}\cdot 8_{\cancel{4}4}$
MON	$4_{\cancel{1}\ 1}\cdot 5_{\cancel{5}\ 5}\cdot 8_{\cancel{6}\ 6}$	$4_{\cancel{1}\ 1}\cdot 5_{\cancel{5}\ 5}\cdot 8_{\cancel{6}\ 6}$	$4_{\cancel{1}\ 1}\cdot 5_{\cancel{5}\ 5}\cdot 8_{\cancel{6}\ 6}$	$4_{\cancel{1}\ 1}\cdot 5_{\cancel{4}\ 4}\cdot 8_{\cancel{4}4}$
NPO	$3_{\cancel{1}\ 1}\cdot 6_{\cancel{6}\ 6}\cdot 12_{\cancel{40}\ 40}$	$3_{\cancel{1}\ 1}\cdot 6_{\cancel{6}\ 6}\cdot 12_{\cancel{40}\ 40}$	$3_{\cancel{1}\ 1}\cdot 6_{\cancel{6}\ 6}$	$3_{\cancel{1}\ 1}\cdot 6_{\cancel{6}6}$

Vertex symbols have been represented in ring index format for ease of comparison.

Table 3: Ring indices for each distinct T-site in the MOZ framework using each ring counting convention.

T-Site	Rings	This Work	Shortest Path Rings	Vertex Symbol Rings
T1	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{7}\ 7}\cdot 10_{\cancel{7}\ 7}\cdot 18_{\cancel{5}\ 5}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{3}\ 3}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{3}\ 3}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8$
T2	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{7}\ 7}\cdot 10_{\cancel{7}\ 7}\cdot 14_{\cancel{5}\ 5}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{3}\ 3}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{3}\ 3}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8$
T3	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8_{\cancel{5}\ 5}\cdot 10_{\cancel{4}\ 4}\cdot 12_{\cancel{4}\ 4}\cdot 14_{\cancel{5}\ 5}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8\cdot 12_{\cancel{4}\ 4}$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8$	$4_{\cancel{3}\ 3}\cdot 6_{\cancel{2}\ 2}\cdot 8$
T4	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{6}\ 6}\cdot 10_{\cancel{6}\ 6}\cdot 12\cdot 18_{\cancel{26}\ 26}$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{6}\ 6}\cdot 12$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{6}\ 6}\cdot 12$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{6}\ 6}\cdot 12$
T5	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{7}\ 7}\cdot 10_{\cancel{6}\ 6}\cdot 14_{\cancel{18}\ 18}$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{7}\ 7}$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{7}\ 7}$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{7}7}$
T6	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{3}\ 3}\cdot 10_{\cancel{2}\ 2}\cdot 12_{\cancel{8}\ 8}\cdot 14_{\cancel{18}\ 18}$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{3}\ 3}\cdot 12_{\cancel{8}\ 8}$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{3}\ 3}$	$4_{\cancel{2}\ 2}\cdot 6\cdot 8_{\cancel{3}3}$

Vertex symbols have been represented in ring index format for ease of comparison.

T-site in every framework across the IZA Database using each of the four ring counting conventions. For each T-site we used the rings to generate a ring index, and Fig. 13 shows how many unique ring indices are present when using each of the ring counting conventions. The plot follows intuition with the number of unique ring indices decreasing as we use more restrictive ring counting conventions, because ~~less~~ fewer rings are found ~~and provides~~ providing less room for differentiation. This raises the question, if you want to ascertain chemical or physical properties about a T-site based on its ring count, and differentiate these T-sites from other similar but distinct ~~T-site~~T-sites, which ring counting convention will suffice? The answer will depend on what level of detail is desired. Larger rings can be found with the standard shortcut definition, while rings traversing other stacked rings will be excluded with our modified shortcut definition. The shortest path definition and vertex symbol rings will provide the most localized information about a T-site.

To further compare the ring counting conventions, we show a distribution of the number of T-sites containing each size ring between 3- and 18-MR in Fig. 14 (right). This plot highlights that more T-sites contain larger sized rings when using the basic definition of a shortcut, and at smaller rings sizes (<6-MR) all the ring counting conventions return the same results. To further emphasize ~~this~~ these results, we have provided a cumulative distribution of the same data normalized to the maximum 'rings' value in Fig. 14. At 6-MRs is where we see the cumulative distribution functions deviate from each other. The largest deviation takes place at 12-MRs, and the cumulative distributions start to level

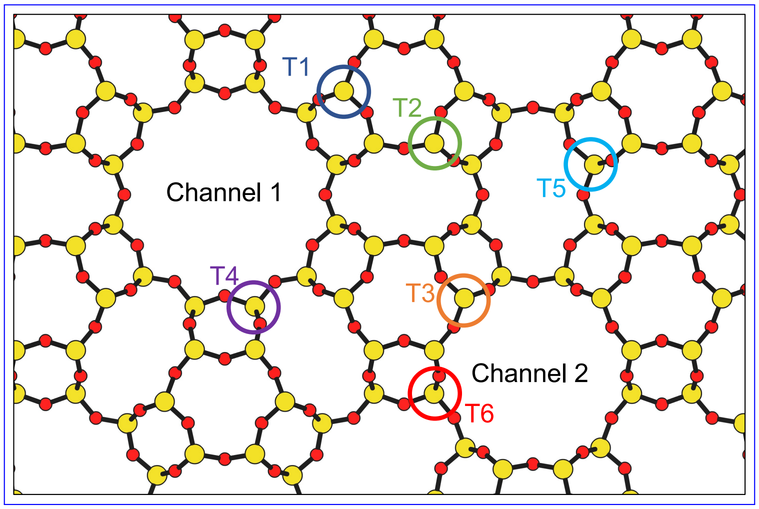


Figure 12: Cutout of the MOZ framework showing two 12-MR channels, with an example of each distinct T-site highlighted. T1: navy, T2: green, T3: orange, T4: purple, T5: blue, and T6: red. As shown, T3, T4, and T6 are all ~~associted~~ associated with the 12-MR channels, while T1, T2, and T5 are not connected to the 12-MR channels.

out at larger ring sizes.

To complete the comparison of ring counting conventions, we have developed a method to determine ~~how similar the ring counts of one convention are to another~~their similarity. We do this by comparing the ring index for a T-site using each convention, where the similarity of the ring indices is scored with Eq. (1). In this equation, sr is the number of similar rings that are found in both counting conventions, and mr is the maximum number of rings found by either convention. For example: the ring index of AFI using the classic shortcut definition and the shortest path definition are $4\cdot 6_{\cancel{13}\,13}\cdot 12\cdot 14_{\cancel{7}\,7}$ and $4\cdot 6_{\cancel{13}\,13}$. The number of similar rings found by both conventions is 14, and the maximum number of rings found by either convention is 22. This would lead to a similarity score of 0.636. We do this for every T-site between two conventions and average the similarity score to get the results in Fig. 15. Down the diagonal each method is compared to itself and clearly has a similarity of 1. The remainder of the table follows intuition, in that the most restrictive ring counting convention (vertex symbol rings) ~~compare~~ compared to the least restrictive convention (rings) has the lowest similarity score. The two most similar ring counting methods are our modified shortcut definition and the shortest path rings.

$$s = \frac{\mathrm{sr}}{\mathrm{mr}} \tag{1}$$

~~One final point we would like to make about T-site characterization, is that the vertex symboland ring indices only provide information about rings, and do not give any information about the spatial orientation of those rings around the T-site. The vertex symbol comes close to accomplishing that, by finding~~

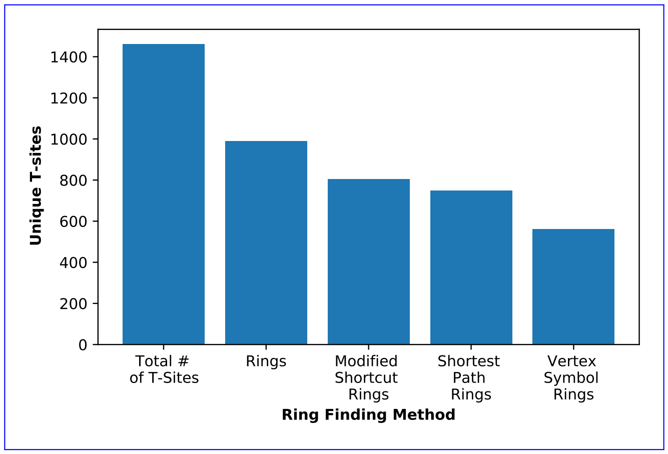


Figure 13: Number of unique T-site ring indices when classified by the rings passing through them using the various ring counting conventions. There are 1460 T-sites across all the frameworks in the IZA Database. As we move from less restrictive to more restrictive (left to right) ring counting conventions, the number of unique ring indices decreases.

~~rings associated with opposite edge pairs~~ The conventional definition of a vertex symbol, which lists pairs of rings along opposite edges of the tetrahedron ~~, and then listing the those pairs~~ from smallest to largest~~. However, the vertex symbol does not capture subtle but distinct differences in the orientation of the rings around the~~ , only partially captures the stereochemistry around a T-site~~that can lead to varying local void environments. For example: MOR~~ . As an example, Fig. 16 illustrates the rings associated with T3, ~~MON~~ T9, and T1 ~~, and EON T9 all have the same vertex symbol of:~~ or MOR, MON, and EON, respectively, all of which share the same $4_2 \cdot 5_2 \cdot 5 \cdot 8_2{}_2 \cdot 5 \cdot 8_2$~~. However, the orientation of those rings around each of those T-sites are not identical. Fig. 16 shows a cutout of each of these frameworks that only includes the atoms that make up the rings of the vertex symbolaround the specified T-sites. We can see that~~ $_2$ vertex symbol. While MOR T3 and EON T9 ~~have the same ring orientation, and that orientation is different from the rings making up the vertex symbol of~~ are similar, MON T1 ~~. The main difference is~~ differs in the location of the $5_{2}{}_{2}$- and 4-MR edges~~. They are highlighted in Fig. 16.~~

~~Cutout of the MOR, EON, and MON frameworks that only shows the rings associated with the vertex symbol of T3, T9, and T1 respectively. The 4-MR (green) and 2× 5-MR (teal) that are in swapped positions are highlighted for emphasis. The 4-MR for MON, and the 25-MRs for MOR and EON are into the plane, and not easily shown.~~

~~The structural differences shown in Fig. 16 that are not able to be captured by the vertex symbol leads us to believe that~~ , as highlighted in the lower panel of Fig. 16. To remove this ambiguity, an alternative approach is to list the rings of the vertex symbol ~~is not a complete descriptor, and there is room to define a new descriptor that takes into consideration ring orientation. This has led~~

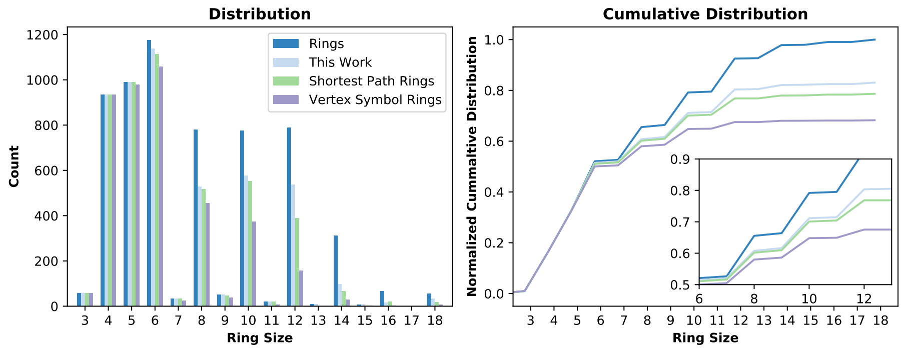


Figure 14: Frequency of T-sites accross all IZA frameworks containing ring sizes between 3- and 18-MR (left), and cumulative distribution of T-sites containing each ring size normalized to the final 'rings' value (right).

~~us to create a new method for listing the rings in~~ by starting with the largest one and systematically adding the remaining rings following the edges of the tetrahedron such that each ring listed is connected to one of the oxygen atoms of the next ring listed. This provides a physical meaning to the order of the vertex symbol ~~that considers the structural connection of the rings described in Section 2.5.~~

~~With this new descriptor~~ rings which accounts for differences in the connectivity of those rings around the T-site. Within this algorithm MOR T3 and EON T9 would be labeled as: $8_{\cancel{2}\,2}\cdot 8_{\cancel{2}\,2}\cdot 5_{\cancel{2}\,2}\cdot 5\cdot 4\cdot 5$, and MON T1 as: $8_{\cancel{2}\,2}\cdot 8_{\cancel{2}\,2}\cdot 4\cdot 5\cdot 5_{\cancel{2}\,2}\cdot 5$. The difference is subtle but highlights the distinct structural difference between the two types of T-sites that is not otherwise captured by a vertex symbol. ~~Stereochemistry of the rings associated with a T-site could influence the chemical properties we care about such as deprotonation energy, T-site substitution energy, or catalytic properties of reactions happening at that T-site.~~

~~We have used this new ordered vertex symbol to characterize~~ We computed the ordered vertex symbols for all 1460 T-sites in the IZA Database. ~~We found that using the standard vertex symbol to characterize T-sites there are~~ The 649 unique vertex symbols ~~present. In contrast, there are~~ in the database increase to 666 unique ordered vertex symbols~~across every T-site. This would imply that not a large amount of vertex symbols contain stereochemical differences. In Table 4 we provide a list of~~. Thus, the consideration of stereochemistry only modestly increases the space of unique tetrahedral sites. Table 4 lists some common T-site vertex symbols ~~, and their representative~~ and associated ordered vertex symbols. Complete results are presented in the Supplementary Information.

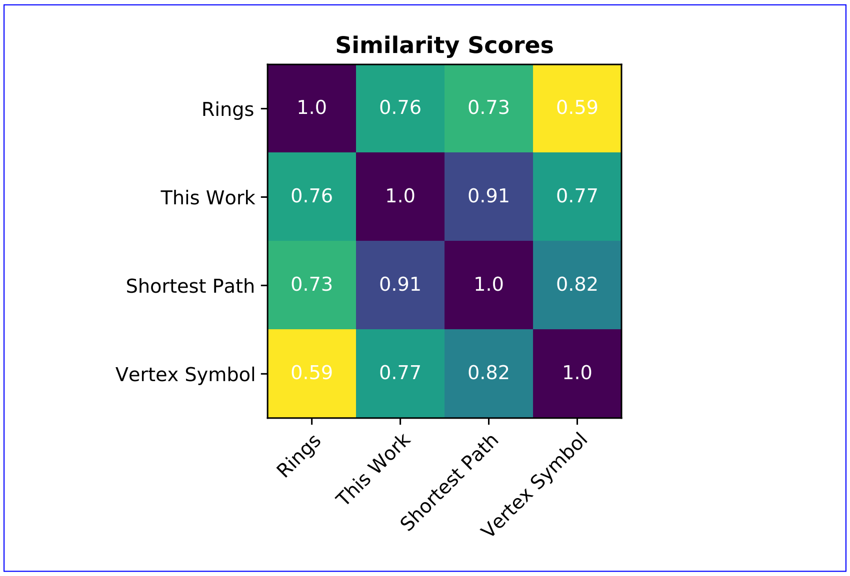


Figure 15: Heat map showing the similarity score for four ring counting methods. Similarity score of 1 means identical set of rings returned, while a similarity of 0 would mean no matching rings are returned.

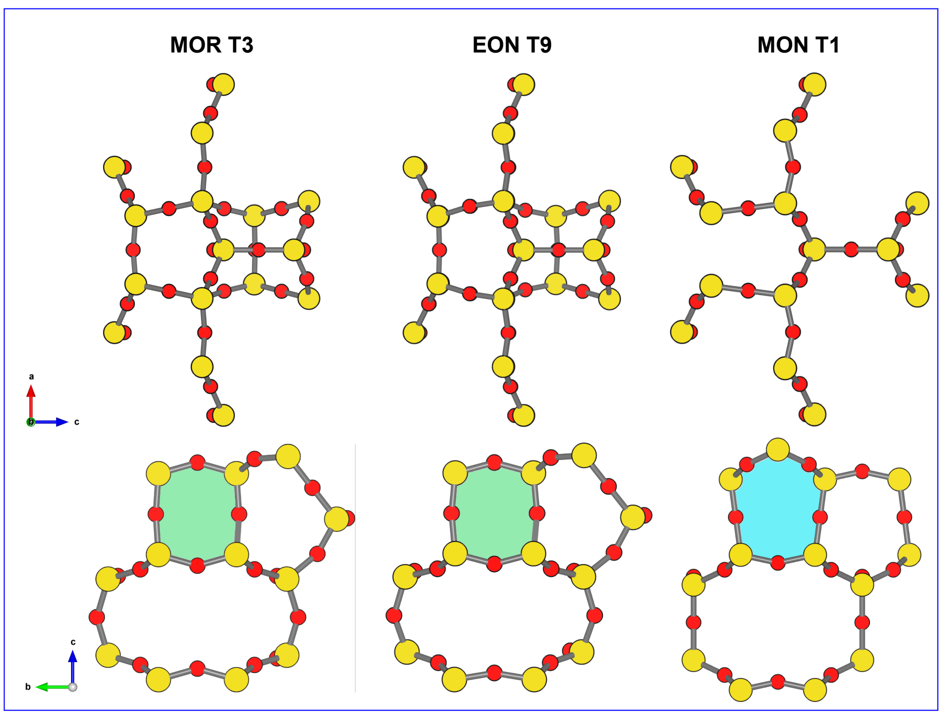


Figure 16: Cutout of the MOR, EON, and MON frameworks that only shows the rings associated with the vertex symbol of T3, T9, and T1 respectively. The 4-MR (green) and 2× 5-MR (teal) that are in swapped positions are highlighted for emphasis. The 4-MR for MON, and the 25-MRs for MOR and EON are into the plane, and not easily shown.

Table 4: List of three vertex symbols, the T-sites associated with them, and the representative ordered vertex symbol for those T-sites.

Vertex Symbol	Framework T-site	Ordered Vertex Symbol
$5\cdot5\cdot5\cdot5_{\cancel{2}2}\cdot5\cdot10$	EWS T3	$10\cdot5_{\cancel{2}2}\cdot5\cdot5\cdot5\cdot5$
	ITN T9	$10\cdot5_{\cancel{2}2}\cdot5\cdot5\cdot5\cdot5$
	ITN T21	$5_{\cancel{2}2}\cdot10\cdot5\cdot5\cdot5\cdot5$
	OKO T2	$10\cdot5_{\cancel{2}2}\cdot5\cdot5\cdot5\cdot5$
	OKO T5	$5_{\cancel{2}2}\cdot10\cdot5\cdot5\cdot5\cdot5$
	PCS T2	$10\cdot5_{\cancel{2}2}\cdot5\cdot5\cdot5\cdot5$
	PCS T3	$5_{\cancel{2}2}\cdot10\cdot5\cdot5\cdot5\cdot5$
	SFS T10	$5_{\cancel{2}2}\cdot10\cdot5\cdot5\cdot5\cdot5$
	SFV T3	$5_{\cancel{2}2}\cdot10\cdot5\cdot5\cdot5\cdot5$
	SFV T7	$5_{\cancel{2}2}\cdot10\cdot5\cdot5\cdot5\cdot5$
	TUN T10	$10\cdot5_{\cancel{2}2}\cdot5\cdot5\cdot5\cdot5$
$4\cdot5_{\cancel{2}2}\cdot5\cdot8\cdot5\cdot8$	DAC T3	$5_{\cancel{2}2}\cdot8\cdot8\cdot5\cdot5\cdot4$
	DAC T4	$5_{\cancel{2}2}\cdot8\cdot8\cdot5\cdot5\cdot4$
	EON T10	$5_{\cancel{2}2}\cdot8\cdot8\cdot5\cdot5\cdot4$
	EPI T1	$5_{\cancel{2}2}\cdot8\cdot8\cdot5\cdot5\cdot4$
	MOR T4	$5_{\cancel{2}2}\cdot8\cdot8\cdot5\cdot5\cdot4$
	RSN T4	$5_{\cancel{2}2}\cdot8\cdot5\cdot5\cdot8\cdot4$
	VNI T2	$5_{\cancel{2}2}\cdot8\cdot5\cdot5\cdot8\cdot4$
	VSV T2	$5_{\cancel{2}2}\cdot8\cdot5\cdot5\cdot8\cdot4$
	YFI T9	$5_{\cancel{2}2}\cdot8\cdot8\cdot5\cdot5\cdot4$
$4\cdot6\cdot4\cdot6\cdot6\cdot8$	ATN T1	$8\cdot6\cdot4\cdot4\cdot6\cdot6$
	JSN T3	$8\cdot6\cdot6\cdot4\cdot4\cdot6$
	PON T1	$8\cdot6\cdot4\cdot4\cdot6\cdot6$
	SAS T1	$8\cdot6\cdot4\cdot4\cdot6\cdot6$
	ZON T3	$8\cdot6\cdot6\cdot4\cdot4\cdot6$

4. Conclusions

Rings of a graph are well defined; here we identified all rings up to 18-MR in every zeolite framework listed on the IZA Structure Database [1]. We find that the commonly reported ring sizes in literature and on the IZA website leave out many rings that fit the classical definition of a cycle that ~~does~~ do not contain a shortcut. To completely describe the topology of a zeolite these rings are required, however there are often cases where someone might want to consider only a subset of rings of interest.

We have shown a comparison of three ~~different~~ existing conventions used to count rings, and highlighted the differences in rings that are found by each convention. The classic definition of a ring identifies the largest set of ring sizes across all the zeolite frameworks, while the shortest path ring and vertex symbol rings only identify smaller ring sizes. We have provided a modified definition of a shortcut that ~~when used to find rings still~~ finds larger rings defining channel openings, but excludes rings that are able to be decomposed into at least one smaller ring. It is important to understand the difference of ring sizes and types found by each convention when discussing the rings of a zeolite framework. A disadvantage to using purely connectivity based definitions of rings is the exclusion of cycles in a framework that behave physicochemically like a ring but contain a shortcut. We have displayed an example case of these geometric rings, and in the future it would be beneficial to develop a computation method of identifying these cycles.

~~This same methodology was used to~~ We describe T-sites in zeolite frameworks by counting all of the rings that pass through the T-site using each of the ring counting conventions described. When using ring finding conventions that find larger rings, we see more diversity in the descriptions of T-sites, which can aid researchers who want to identify similar T-sites across multiple frameworks. We have also shown that the vertex symbol used to describe T-sites leaves out subtle but distinct stereochemical differences in the spatial orientation of the rings around a T-site. To address this shortcoming we have provided a new method for ordering the rings of a vertex symbol that takes into consideration the ring stereochemistry and is able to identify differences in T-sites that have the same vertex symbol. In the future, correlating physicochemical properties of T-sites to the ring descriptors identified with each ring counting convention can help identify sets of frameworks with desired T-site properties.

5. Data Statement

Complete ring finding results are available for download via the Supplementary Information. This folder contains a file for each ring counting convention: all rings, modified shortcut rings, shortest path rings, vertex symbol rings, and ordered vertex symbol rings. These files contain the rings associated with each oxygen atom and T-site in every zeolite framework listed on the IZA Structure Database.

6. Acknowledgments

We acknowledge financial support provided by the National Science Foundation under Cooperative Agreement No. EEC-1647722, which is an Engineering Research Center for the Innovative and Strategic Transformation of Alkane Resources. JTC thanks the Arthur J. Schmitt Foundation for financial aid in the form of a PhD fellowship. We thank Dr. Christian Baerlocher for the numerous discussions about the topology of zeolites, and methodologies used by the IZA Structure Database. We thank Dr. German Sastre for providing a copy of zeoTsites to compare results, and for the helpful conversions about use cases and zeolite topology. This research was supported in part by the Notre Dame center for Research Computing through access to high performance computing clusters.

References

[1] C. Baerlocher, L. McCusker, Database of Zeolite Structures. URL `http://www.iza-structure.org/databases/`

[2] X. Li, M. W. Deem, Why Zeolites Have So Few Seven-Membered Rings, The Journal of Physical Chemistry C 118 (29) (2014) 15835–15839.

doi:10.1021/jp504143r.
URL https://pubs.acs.org/doi/10.1021/jp504143r

[3] R. A. Curtis, M. W. Deem, A Statistical Mechanics Study of Ring Size, Ring Shape, and the Relation to Pores Found in Zeolites, Journal of Physical Chemistry B 107 (2003) 8612–8620.

[4] D. Bermúdez, G. Sastre, Calculation of pore diameters in zeolites, Theoretical Chemistry Accounts 136 (10) (2017) 116. doi:10.1007/s00214-017-2143-6.
URL https://doi.org/10.1007/s00214-017-2143-6

[5] S. Li, H. Li, R. Gounder, A. Debellis, I. B. Müller, S. Prasad, A. Moini, W. F. Schneider, First-Principles Comparison of Proton and Divalent Copper Cation Exchange Energy Landscapes in SSZ-13 Zeolite, The Journal of Physical Chemistry C 122 (41) (2018) 23564–23573. doi:10.1021/acs.jpcc.8b07213.
URL https://pubs.acs.org/doi/10.1021/acs.jpcc.8b07213

[6] P. M. Kester, J. T. Crum, S. Li, W. F. Schneider, R. Gounder, Effects of Brønsted acid site proximity in chabazite zeolites on OH infrared spectra and protolytic propane cracking kinetics, Journal of Catalysis 395 (2021) 210–226. doi:10.1016/j.jcat.2020.12.038.
URL https://linkinghub.elsevier.com/retrieve/pii/S0021951721000191

[7] B. A. Helfrecht, R. Semino, G. Pireddu, S. M. Auerbach, M. Ceriotti, A new kind of atlas of zeolite building blocks, The Journal of Chemical Physics 151 (15) (2019) 154112. doi:10.1063/1.5119751.
URL http://aip.scitation.org/doi/10.1063/1.5119751

[8] L. Guttman, Ring structure of the crystalline and amorphous forms of silicon dioxide, Journal of Non-Crystalline Solids 116 (2-3) (1990) 145–147. doi:10.1016/0022-3093(90)90686-G.
URL https://linkinghub.elsevier.com/retrieve/pii/002230939090686G

[9] K. Goetzke, H.-J. Klein, Properties and efficient algorithmic determination of different classes of rings in finite and infinite polyhedral networks, Journal of Non-Crystalline Solids 127 (2) (1991) 215–220. doi:10.1016/0022-3093(91)90145-V.
URL https://linkinghub.elsevier.com/retrieve/pii/002230939190145V

[10] D. S. Franzblau, Computation of ring statistics for network models of solids, Physical Review B 44 (10) (1991) 4925–4930. doi:10.1103/PhysRevB.44.4925.
URL https://link.aps.org/doi/10.1103/PhysRevB.44.4925

[11] X. Yuan, A. Cormack, Efficient algorithm for primitive ring statistics in topological networks, Computational Materials Science 24 (3) (2002) 343–360. doi:10.1016/S0927-0256(01)00256-7.
URL https://linkinghub.elsevier.com/retrieve/pii/S0927025601002567

[12] F. Wooten, Structure, odd lines and topological entropy of disorder of amorphous silicon, Acta Crystallographica Section A: Foundations of Crystallography 58 (4) (2002) 346–351, number: 4 Publisher: International Union of Crystallography. doi:10.1107/S0108767302006669.
URL http://scripts.iucr.org/cgi-bin/paper?bk0109

[13] S. Le Roux, P. Jund, Ring statistics analysis of topological networks: New approach and application to amorphous GeS2 and SiO2 systems, Computational Materials Science 49 (1) (2010) 70–83. doi:10.1016/j.commatsci.2010.04.023.
URL https://linkinghub.elsevier.com/retrieve/pii/S0927025610002363

[14] G. Sastre, J. D. Gale, ZeoTsites: a code for topological and crystallographic tetrahedral sites analysis in zeolites and zeotypes, Microporous and Mesoporous Materials 43 (1) (2001) 27–40. doi:10.1016/S1387-1811(00)00344-9.
URL https://linkinghub.elsevier.com/retrieve/pii/S1387181100003449

[15] V. A. Blatov, A. P. Shevchenko, D. M. Proserpio, Applied Topological Analysis of Crystal Structures with the Program Package ToposPro, Crystal Growth & Design 14 (7) (2014) 3576–3586. doi:10.1021/cg500498k.
URL https://pubs.acs.org/doi/10.1021/cg500498k

[16] V. A. Blatov, O. Delgado-Friedrichs, M. O'Keeffe, D. M. Proserpio, Three-periodic nets and tilings: natural tilings for nets, Acta Crystallographica Section A Foundations of Crystallography 63 (5) (2007) 418–425. doi:10.1107/S0108767307038287.
URL https://scripts.iucr.org/cgi-bin/paper?S0108767307038287

[17] M. O'Keeffe, S. Hyde, Vertex symbols for zeolite nets, Zeolites 19 (5-6) (1997) 370–374. doi:10.1016/S0144-2449(97)00133-4.
URL https://linkinghub.elsevier.com/retrieve/pii/S0144244997001334

[18] G. Sastre, A. Corma, Topological Descriptor for Oxygens in Zeolites. Analysis of Ring Counting in Tetracoordinated Nets, The Journal of Physical Chemistry C 113 (16) (2009) 6398–6405. doi:10.1021/jp8100128.
URL https://pubs.acs.org/doi/10.1021/jp8100128

[19] J. Crum, jtcrum/zse: ZSE Stable Version (Sep. 2022). doi:10.5281/ZENODO.7080433.
URL https://zenodo.org/record/7080433

[20] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, Journal of Physics: Condensed Matter 29 (27) (2017) 273002. doi:10.1088/1361-648X/aa680e. URL `http://stacks.iop.org/0953-8984/29/i=27/a=273002?key=crossref.20f9751653d872507bf6c0cc5737032c`

[21] A. A. Hagberg, D. A. Schult, P. J. Swart, Exploring network structure, dynamics, and function using NetworkX, in: Proceedings of the 7th Python in Science Conference, Pasadena, CA USA, 2008, pp. 11–15.

[22] J. Dědeček, Z. Sobalík, B. Wichterlová, Siting and distribution of framework aluminium atoms in silicon-rich zeolites and impact on catalysis, Catalysis Reviews 54 (2) (2012) 135–223. doi:10.1080/01614940.2012.632662.
URL `http://www.tandfonline.com/doi/abs/10.1080/01614940.2012.632662`

[23] M. Bernauer, E. Tabor, V. Pashkova, D. Kaucký, Z. Sobalík, B. Wichterlová, J. Dedecek, Proton proximity – New key parameter controlling adsorption, desorption and activity in propene oligomerization over H-ZSM-5 zeolites, Journal of Catalysis 344 (2016) 157–172. doi:10.1016/j.jcat.2016.09.025.
URL `https://linkinghub.elsevier.com/retrieve/pii/S002195171630197X`

[24] C. T. Nimlos, A. J. Hoffman, Y. Gul Hur, B. Jin Lee, J. R. Di Iorio, D. Hibbitts, R. Gounder, Experimental and Theoretical Assessments of Aluminum Proximity in MFI Zeolites and its Alteration by Organic and Inorganic Structure-Directing Agents, Chemistry of Materials (2020). doi:https://doi.org/10.1021/acs.chemmater.0c03154.

[25] J. R. Di Iorio, S. Li, C. B. Jones, C. T. Nimlos, Y. Wang, E. Kunkes, V. Vattipalli, S. Prasad, A. Moini, W. F. Schneider, R. Gounder, Cooperative and Competitive Occlusion of Organic and Inorganic Structure-Directing Agents within Chabazite Zeolites Influences Their Aluminum Arrangement, Journal of the American Chemical Society 142 (10) (2020) 4807–4819. doi:10.1021/jacs.9b13817.
URL `https://pubs.acs.org/doi/10.1021/jacs.9b13817`

Dr. German Sastre, Instituto de Tecnología Química UPV-CSIC, gsastre@itq.upv.es
Expert in computational analysis of zeolite rings, member of the (ref. 14)

Prof. François-Xavier Coudert, Chimie ParisTech, PSL University, CNRS, Institut de Recherche de Chimie Paris, fx.coudert@chimieparistech.psl.eu
Expert in machine learning applied to zeolites

Prof. Scott M. Auerbach, Department of Chemistry and Department of Chemical Engineering, University of Massachusetts Amherst, auerbach@umass.edu
Expert in zeolite simulation

Prof. Christian Baerlocher, Department of Materials, ETH Zürich, christian.baerlocher@mat.ethz.ch
Knowledgeable about all things zeolites, created the Database of Zeolite Structures, and member of the International Zeolite Association Structure Commission (ref. 1)

Prof. Michael M. J. Treacy, Department of Physics, Arizona State University, teacy@asu.edu
Expert in computationally deriving new zeolites structures and developed a database of hypothetical zeolite structures

Prof. Michael Deem, Physics and Astronomy, Rice University, mwdeem@rice.edu
Pioneer in models of OSDA-zeolite interactions (refs. 2,3)

Prof. Julian Gale, School of Molecular and Life Sciences, Curtin University, j.gale@curtin.edu.au
Expert in zeolite simulation (ref. 14)

Characterization and Analysis of Ring Topology of Zeolite Frameworks

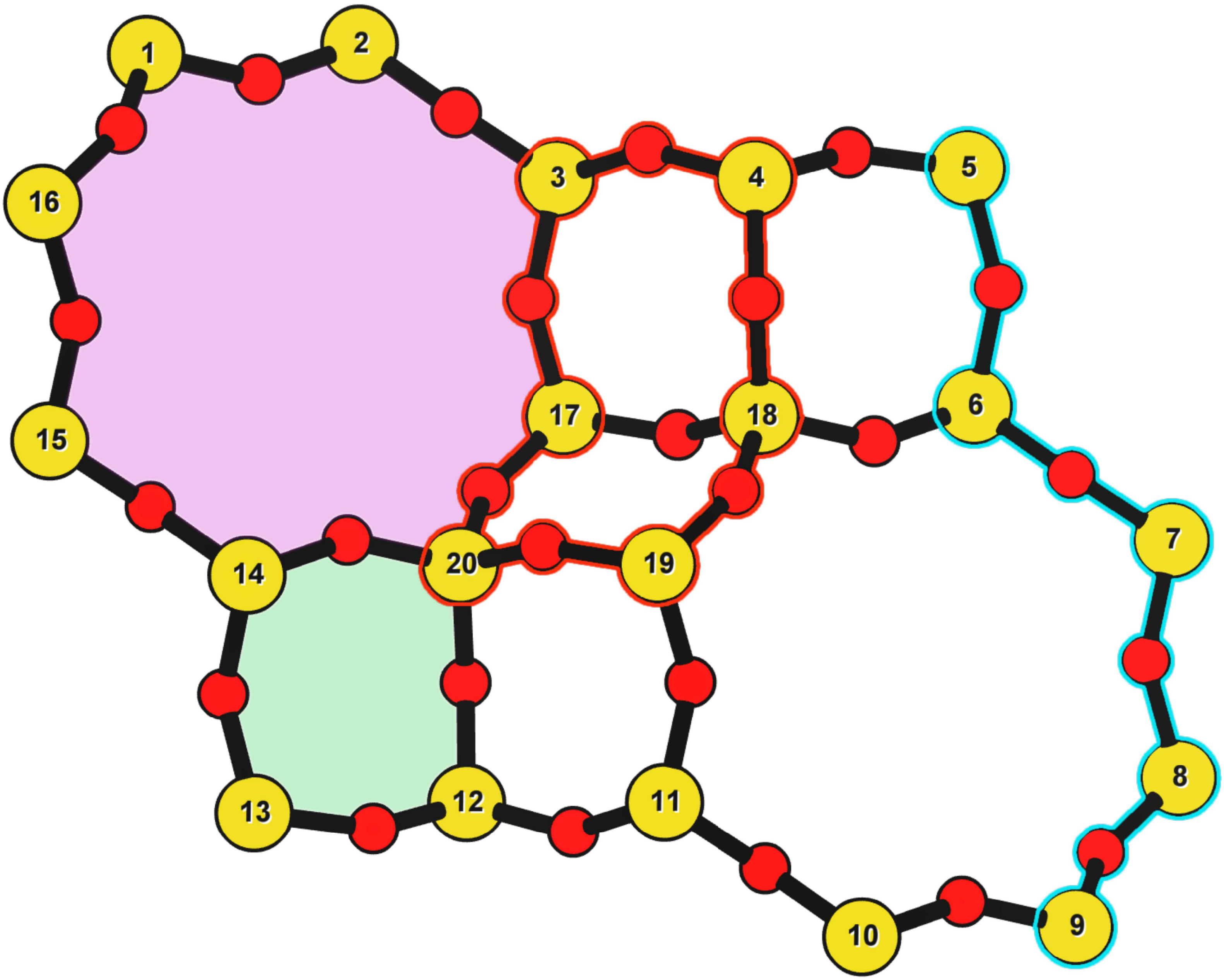
William F. Schneider, Jerry T. Crum[a], Justin R. Crum[b], Cameron Taylor[a], William F. Schneider[a,c]

[a]*Department of Chemical and Biomoledcular Engineering, University of Notre Dame, 250 Nieuwland Science Hall, Notre Dame, IN 46556, USA*
[b]*Department of Applied Mathematics, University of Arizona, 617 N Santa Rita Ave, Tucson, AZ 85721, USA*
[c]*Department of Chemistry and Biochemistry, University of Notre Dame, 251 Nieuwland Science Hall, Notre Dame, IN 46556, USA*

1. Highlights

1. Classes of zeolite rings identified
2. Algorithm and software to identify and enumerate rings developed
3. Software used to analyze ring patterns across all zeolites

Graphical Abstract



Click here to access/download
Supplementary Material
all_rings.txt

Click here to access/download
Supplementary Material
modified_shortcut_rings.txt

Click here to access/download
Supplementary Material
vertex_symbol_rings.txt

Declaration of Interest Statement (Please upload in Word format)

Declaration of interests

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

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

```
This is pdfTeX, Version 3.141592653-2.6-1.40.24 (TeX Live 2022)
(preloaded format=pdflatex 2022.10.28)  11 JAN 2023 00:52
entering extended mode
 restricted \write18 enabled.
 %&-line parsing enabled.
**manuscript.tex
(./manuscript.tex
LaTeX2e <2022-06-01> patch level 5
L3 programming layer <2022-10-26> (c:/TeXLive/texmf-
local/tex/latex/aries/elsar
ticle.cls
Document Class: elsarticle 2008/10/09, 1.0.2: Elsevier Science
\@bls=\dimen138

! LaTeX Error: File `article.cls' not found.

Type X to quit or <RETURN> to proceed,
or enter new name. (Default extension: cls)

Enter file name:
! Emergency stop.
<read *>

l.80 \RequirePackage
                    {graphicx}^^M
*** (cannot \read from terminal in nonstop modes)


Here is how much of TeX's memory you used:
 84 strings out of 475066
 1378 string characters out of 5782772
 341620 words of memory out of 5000000
 21586 multiletter control sequences out of 15000+600000
 469259 words of font info for 28 fonts, out of 8000000 for 9000
 1141 hyphenation exceptions out of 8191
 34i,0n,50p,144b,17s stack positions out of
10000i,1000n,20000p,200000b,200000s
!  ==> Fatal error occurred, no output PDF file produced!
```