Discovering planar disorder in close-packed structures from x-ray diffraction: Beyond the fault model

D. P. Varn,[1,2] G. S. Canright,[2,3] and J. P. Crutchfield[1]

[1]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501*
[2]*Department of Physics and Astronomy, University of Tennessee, Knoxville, Tennessee 37996*
[3]*Telenor Research and Development, 1331 Fornebu, Norway*

We solve a long-standing problem—determining structural information for disordered materials from their diffraction spectra—for the special case of planar disorder in close-packed structures (CPS's). Our solution offers the most complete possible statistical description of the disorder and, from it, we find the minimum effective memory length for stacking sequences in CPS's. We contrast this description with the so-called "fault" model by comparing the structures inferred using both approaches on two previously published zinc sulphide diffraction spectra.

Stacking faults[1,2]—deviations from crystallinity that result when an entire plane of atoms breaks a stacking rule—are common in a broad class of materials known as polytypes.[2,3] Found in metals such as Li, micas, simple inorganic compounds like ZnS and CdI_2, and more complicated materials such as the oxides $Ba_5Nb_4O_{15}$ and Ba_2FeSbO_6, polytypism is defined as the building up of solids from identical *modular layers* (ML's) (Ref. 4) that differ only in their stacking orientation. The Bragg peaks in x-ray diffractograms have long been used to infer the crystal structure of ordered solids. While there has been some success in treating weakly disordered crystals, the general problem of extracting structural information about a disordered crystal from its diffraction spectrum has been an unsolved problem for nearly 70 years.

In this paper, we solve the general problem of inferring crystal structure from diffraction spectra for close-packed structures (CPS's) containing any kind or amount of one-dimensional stacking disorder. Our solution offers the unique, minimal description of the stacking. Our solution is *direct* because we make no assumptions about either the crystal structure or the kind of disorder. It further represents the most complete possible solution to this problem. We illustrate our method on two previously published diffraction spectra for ZnS and compare our results to the best previous analysis of these spectra.

There has been considerable interest in determining crystal structure from diffraction spectra for quite some time (Refs. 2 and 5–12, for example). Previous attempts for CPS's have proceeded as follows. Different kinds of stacking faults based on physically plausible mechanisms were postulated, such as *growth faults*, *deformation faults*, and *layer displacement faults*.[2] By examining the intensity, placement, and broadening of the Bragg peaks it was possible to estimate the kind and amount of disorder present. This approach is necessarily confined to those instances where there is weak faulting of a particular type (or at most a few types) in a single parent crystal. This approach is *indirect*, since one must assume *a priori* both a crystal structure and a faulting mechanism. These efforts met with good success for weakly faulted specimens as such Co.[5] However, for polytypes such as ZnS and SiC, this approach has not been altogether satisfactory.

We refer to this class of approach as the *fault model* (FM), which we define as any model that assumes a parent crystal permeated with one or more kinds of stacking faults. Nearly all previous attempts (Refs. 2,8–10,12, and 13, for example) to infer crystal structure from diffuse scattering for disordered crystals have been applications of the FM. There are several drawbacks to the this approach. Among them are (a) the need to assume a single parent crystalline structure into which stacking defects are introduced, precluding the description of disorder interspersed between distinct crystal structures, (b) the limitation to the case of weak faulting, and (c) the restriction of the quantitative analysis to the effects of faulting on the Bragg peaks only, ignoring the information in the diffuse scattering. Other techniques have attempted to include both the Bragg peaks and diffuse scattering, such as the reverse Monte Carlo procedure (see Ref. 23, for example). This generates structures that give good agreement with experiment but are physically implausible. Therefore, one makes assumptions concerning the disorder so that the procedure is not *a priori*. To our knowledge, reverse Monte Carlo techniques have not yet been applied to planar disorder.

We present a solution that overcomes all of these difficulties. We break our method into three parts.

(i) We note that a polytype is simply described by its *stacking sequence*—the one-dimensional list of successive orientations found as one moves along the stacking direction. We refer to the effective stochastic process induced by scanning the list as the *stacking process*. For CPS's, a stacking sequence is most compactly described by the Hägg notation.[2] One replaces the set $\{A,B,C\}$ of allowed orientations with a binary alphabet $\mathcal{A}=\{0,1\}$: a ML is labeled "1," if it is cyclically related to the preceding ML, or 0, if it is not.

We use the diffraction spectrum to find average correlations between ML's as a function of the number n of separating layers. We make the usual assumption about ML stacking in CPS's,[2] namely that the ML's themselves are undefected, that each ML has the same scattering power, and that the spacing between ML's is independent of the local stacking arrangement; then correlation functions (CF's) $Q_c(n)$ and $Q_a(n)$ (Ref. 14) can be found by Fourier analysis

of the diffraction spectrum.[1] $Q_c(n)$ and $Q_a(n)$ are defined as the probability that any two ML's at separation n are cyclically or anticyclically related, respectively.

(ii) We infer the spatial patterns of ML's that reproduce these CF's by reconstructing an ϵ-machine,[15–17] which describes the minimal effective states of the stacking process. Assume we know the probability $p(\omega)$ of stacking sequences ω. At each ML in a stacking sequence define the "past" $\overleftarrow{\omega}$ as those ML's already seen and the "future" $\overrightarrow{\omega}$ as those yet to be seen: $\omega = \overleftarrow{\omega}\overrightarrow{\omega}$. The effective states of the stacking process then are defined as the *sets* of pasts $\overleftarrow{\omega}$ that lead to statistically equivalent futures:

$$\overleftarrow{\omega}_i \sim \overleftarrow{\omega}_j \text{ if and only if } p(\overrightarrow{\omega}|\overleftarrow{\omega}_i) = p(\overrightarrow{\omega}|\overleftarrow{\omega}_j). \quad (1)$$

These equivalence classes of pasts are the stacking process's *causal states*. Along with their transitions, they comprise the process's ϵ-*machine*—a statistical description of the ensemble of spatial patterns that produces the stacking distribution $p(\omega)$. It has been shown that the ϵ-machine is the minimal-size (as measured by the number of states), optimal predictor of a process, and up to state relabeling, it is the unique such a description.[15–17]

To find the causal states we must first estimate the probability $p(\omega)$ of stacking sequences ω averaged over the sample. Note that, from conservation of probability, $p(u) = p(0u) + p(1u) = p(u0) + p(u1)$, for all $u \in \mathcal{A}^r$, where $\mathcal{A}^r$ is the set of all sequences of length r. Additionally, the probabilities for sequences of the same length are normalized: $\Sigma_{\omega \in \mathcal{A}^{r+1}} p(\omega) = 1$. Together these constraints provide 2^r independent relations among probabilities for the 2^{r+1} possible stacking sequences of length $r+1$.

The other 2^r constraints come from relating CF's to sequence probabilities via

$$Q_\alpha(n) = \sum_{\omega \in \mathcal{A}^n_\alpha} p(\omega), \quad (2)$$

where $\mathcal{A}^n_\alpha$ is that subset of length-n sequences with a cyclic ($\alpha = c$) or an anticyclic ($\alpha = a$) rotation between ML's at separation n. We take as many of these latter relations as necessary to form a complete set of equations. At a fixed r, the set of equations describes the stacking sequence as generated by an rth-order Markov process. At $r = 3$ one encounters the first nonlinearities due to the necessity of using CF's at $n = 5$ to obtain a complete set of equations. We rewrite the probability of sequences of length $n = 5$ in terms of the conditional probabilities of those at $n = 4$, and it is this mapping that is nonlinear. We solve numerically for the stacking sequence probabilities $p(\omega)$ and then find the set of causal states using the equivalence relation, Eq. (1). The causal-state transitions are estimated from the conditional distributions of the next ML orientation given pasts $\overleftarrow{\omega}$ associated with each causal state.

(iii) We begin with the $r = 1$ reconstructed ϵ-machine and use it to generate a sample stacking sequence (here we used length 400 000), and from this we estimate the ϵ-machine's predicted CF's and diffraction spectrum. We then compare the latter to the experimental diffraction spectrum. If there is
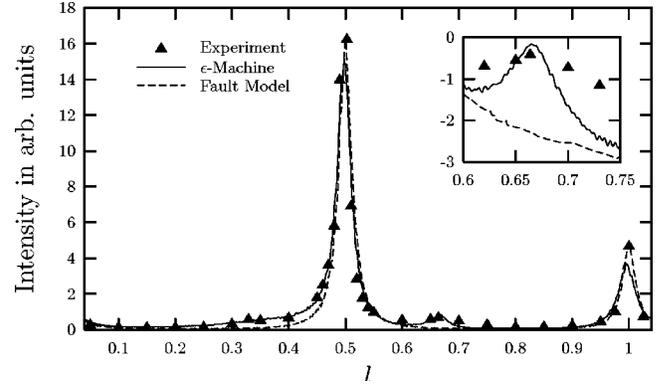
FIG. 1. Comparison of the experimental diffraction spectrum SK134 along the 10.l row (triangles) for a disordered hcp ZnS single crystal (Ref. 2, p. 134) with spectra estimated from the FM with 5% deformation faulting (dashed line) and $r = 3$ ϵ-machine (solid line). The vertical scale in the inset is the logarithmic intensity.

not sufficient agreement, we increment r and repeat the reconstruction and comparison. The resulting r is called the stacking process's *memory length*, since it is the amount of history (in ML's) one must use to optimally predict the process. We note that the reconstruction algorithm has but a single "free" parameter, namely r. Once an r is selected, the ϵ-machine is completely determined by the experimental data. We find the minimum r that satisfactorily explains the experimental data.

ZnS can be thought to have a CPS with a basis composed of two atoms, zinc and sulphur, with the sulphurs displaced one quarter of a body diagonal (as referred to the conventional unit cell) along the stacking direction.[2] We take an ML to be this zinc-sulphur pair arranged in a hexagonal net.[4] We correct the experimentally obtained diffraction spectrum for the atomic scattering factors, the structure factor, dispersion factors, and the polarization factor.[18]

We now give the results for ϵ-machine reconstruction for two experimental diffraction spectra, SK134 and SK135 from Ref. 2. Let l be a continuous variable that indexes the magnitude of the perpendicular component of the diffracted wave $k = 2\pi l/c$, where c is the spacing between adjacent ML's. We select a unit interval in l on which to analyze each spectrum. Since many diffraction spectra suffer from experimental error,[2] we show elsewhere[19] that there are relations that the CF's must obey for any CPS and that we can use these to select a relatively error-free l interval. The spectra from experiment and ϵ-machine reconstruction are normalized.

The triangles in Fig. 1 show the experimental diffraction spectrum SK134 along the 10.l row for an hexagonally close-packed (hcp) ZnS crystal annealed at 300 °C for 1 h. (For the sake of clarity, here we only show a few selected points from the experimental diffraction spectrum. Experimentalists report the spectrum in increments of $\Delta l = 0.005$. For our analysis, we used this finer mesh.) Sebastian and Krishna[2] attribute the observed disorder to a 5% probability of deformation faulting at each ML. (This is the FM-predicted spectrum given as a dashed line in Fig. 1.) We find
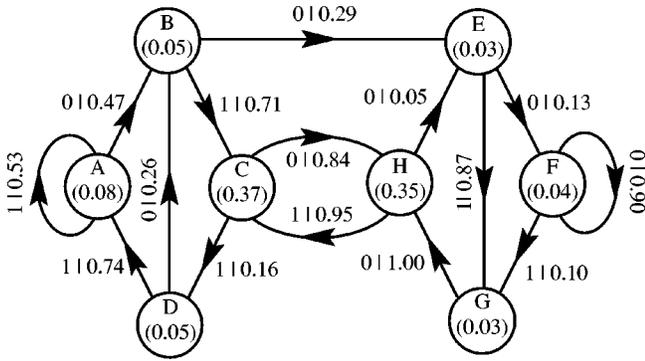
FIG. 2. The recurrent causal states {A–H} of the reconstructed ϵ-machine estimated from the experimental diffraction spectrum SK134 of Fig. 1 with $r=3$. Asymptotic state probabilities are given in parentheses; edge label $s|p$ indicates a transition on symbol s with probability p.

FIG. 3. Comparison of the experimental diffraction spectrum SK135 along the 10.l row (triangles) for a disordered ccp ZnS single crystal [2, p. 135] with the diffraction spectra calculated from the the FM with 12% twinned faulting (dashed line) and $r=3$ ϵ-machine (solid line).

that the smallest-r ϵ-machine that gives adequate agreement (solid line) with experiment is estimated at $r=3$; it is shown in Fig. 2.

It is possible to give an approximate equivalent of this ϵ-machine in terms of the FM, but we stress that this decomposition is *not unique*. We associate each closed, non-self-intersecting loop [called a *simple cycle* or SC (Ref. 20)] in the ϵ-machine with either a crystal structure or a fault. In this way, ϵ-machines directly describe familiar structures in polytypes. For instance, the closed loop between causal states C and H in Fig. 2 implies a stacking sequence . . . 010101 . . . , which is simply the Hägg notation for the hcp structure. One concludes, then, that there is no qualitative difference between what one calls faults and crystal structure. The distinction is, in fact, quantitative and one of convenience—crystal structures have relatively high probabilities, as opposed to the rarer faults. For the most general $r=3$ ϵ-machine, it is known that there are 19 such SC's.[21] Since eight independent CF's are sufficient to specify an $r=3$ ϵ-machine, the problem of decomposing the ϵ-machine into SC's is underdetermined. This conclusion holds for all $r \geq 2$. Therefore, without a fortuitous vanishing of causal states or transitions, the FM, *unlike* the ϵ-machine, is not unique.

For the sake of comparison with the previous best FM analysis, we decompose the ϵ-machine in Fig. 2 into SC's with the assumption that faults corresponding SC's of length 7 or greater are not present. We define the *fault density* as the sum of the weights of the arcs forming the fault.[19] We can then assign a fault-density distribution for SK134 (second column) as follows and compare it to that of Ref. 2 (third column):

hcp	64%	83%
ccp	8%	0%
deformation fault	16%	17%
growth fault	6%	0%
layer displacement fault	6%	0%

where ccp stands for cubic close packed. Thus, the ϵ-machine description of the crystal differs significantly from
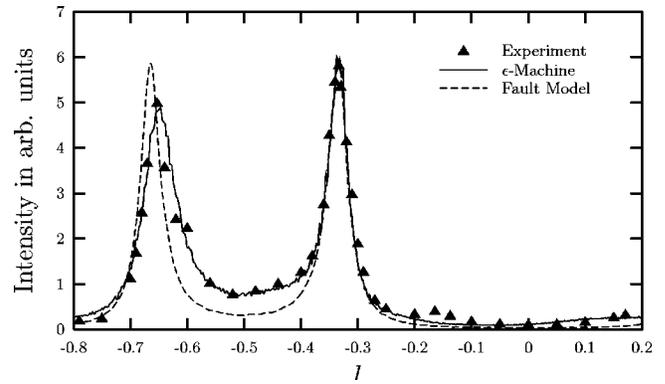
that of Sebastian and Krishna.[2] While we both find qualitatively that deformation faulting is important, we also detect ccp structures, as well as growth faults and layer displacement faults. Overall, ϵ-machine analysis finds a much more disordered crystal. This is borne out when comparing the FM and ϵ-machine diffraction spectra. Figure 1 shows that, while both agree reasonably well with experiment at the broadened peaks at $l=0.5$ and 1, the ϵ-machine is in better agreement along the shoulders of the Bragg peaks, as well as at the rise in broadband intensity at $l \approx 0.67$ (inset in Fig. 1).

Figure 3 plots the experimental diffraction spectrum along the 10.l row (triangles) for a hcp ZnS crystal annealed at $500 \, ^\circ \text{C}$ for 1 h. Sebastian and Krishna[2] find this to be a disordered, twinned-ccp crystal with a twin-fault probability of 12%, calculated from the observed half-widths of the peaks. The diffraction spectrum for such a faulting mechanism is shown in Fig. 3 (dashed line). Only the peak at $l= -0.33$ was used to find the faulting mechanism, and one sees that the FM reproduces it well. However, the second peak at $l=-0.67$ is poorly represented, as is the diffuse scattering between the two peaks. This demonstrates the pitfalls in simply fitting an FM to a single Bragg peak, ignoring the information contained in other peaks and in the diffuse scattering. We also note that the small rise in diffracted intensity at l
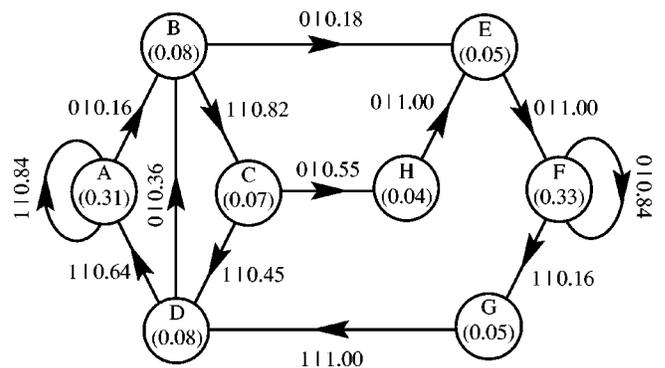
FIG. 4. Recurrent states of the reconstructed ϵ-machine for the experimental diffraction data SK135 of Fig. 3 using $r=3$.

≈ -0.16 is likewise missed by the FM. The ϵ-machine spectrum (solid line) also misses this rise, but otherwise is in excellent agreement with the experiment. Figure 4 shows the reconstructed ϵ-machine obtained at $r=3$. The large probabilities for causal states A and F and their large self-loop transition probabilities, associated with stacking sequences $\dots 1111 \dots$ and $\dots 0000 \dots$, indicate that this is a twinned-ccp crystal. The missing H→C causal-state transition—and thus the resulting absence of the $\dots 0101 \dots$ stacking—implies that the original hcp structure has been eliminated during the annealing.

In conclusion, we have solved the problem of discovering and describing planar disorder and structure in CPS's from their diffraction spectra. We have demonstrated that the FM—the reigning paradigm for understanding and describing planar disorder in crystals—is necessarily inadequate, both in conception and practice. A simple examination of the effects of faulting on the Bragg peaks is insufficient to properly detect the disorder present. In contrast, (a) ϵ-machines provide the most general description possible of structure in one dimension (1D). (b) No assumptions about the crystal or fault structure need be made. (c) Any amount or kind of planar disorder can be treated. (d) More than one crystal structure may be present, as we found in the first example. (e) All the information in the diffraction spectrum—both Bragg like and diffuse scattering—is used to generate the model. Moreover, (f) we quantified the memory length for disordered 1D systems; for the ZnS samples considered, it was 3 ML's. Thus, we find that the memory length in *disordered* structures for ZnS (as for long-period *ordered* structures) clearly extends beyond the calculated range (1 ML) of interlayer interaction.[22] Additionally, (g) we show elsewhere[19] that given the coupling constants between ML's,[22] we can determine the average stacking-fault energy for a disordered crystal. It is expected that other physical parameters will be amenable to calculation directly from the ϵ-machine. And finally, (h) considering the ubiquity of experimentally accessible power spectra in physics, our technique has implications far beyond its present application in polytypism.

[1] A. Guinier, *X-Ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies* (Freeman, San Francisco, 1963).

[2] M. T. Sebastian and P. Krishna, *Random, Non-Random and Periodic Faulting in Crystals* (Gordon and Breach, New York, 1994).

[3] G.C. Trigunayat, Solid State Ionics **48**, 3 (1991).

[4] D.P. Varn and G.S. Canright, Acta Crystallogr., Sect. A: Found. Crystallogr. **57**, 4 (2001).

[5] A.J.C. Wilson, Proc. R. Soc. London, Ser. A **180**, 277 (1942).

[6] S. Hendricks and E. Teller, J. Chem. Phys. **10**, 147 (1942).

[7] H. Jagodzinski, Acta Crystallogr. **2**, 201 (1949); **2**, 208 (1949).

[8] R. Gevers, Acta Crystallogr. **7**, 337 (1954); **7**, 492 (1954).

[9] J.M. Cowley, Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. **A32**, 83 (1976); **A32**, 88 (1976).

[10] D. Pandey, S. Lele, and P. Krishna, Proc. R. Soc. London, Ser. A **369**, 435 (1980); **369**, 451 (1980); **369**, 463 (1980).

[11] R. Berliner and S. Werner, Phys. Rev. B **34**, 3586 (1986).

[12] J.B. Gosk, Cryst. Res. Technol. **35**, 101 (2000).

[13] A notable exception is Jagodzinski's disorder model [H. Jagodzinski, Acta Crystallogr. **2**, 208 (1949)]. He introduces a two-parameter model based on the assumption that the hexagonal close-packed and the cubic close-packed are the parent phases. This approach is a constrained special case of an $r=2$ ϵ-machine.

[14] J. Yi and G.S. Canright, Phys. Rev. B **53**, 5198 (1996).

[15] J.P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).

[16] J.P. Crutchfield and D.P. Feldman, Phys. Rev. E **55**, R1239 (1997).

[17] C.R. Shalizi and J.P. Crutchfield, J. Stat. Phys. **104**, 819 (2001).

[18] *International Tables for Crystallography*, 3rd revised ed. edited by T. Hahn, A. J. C. Wilson, and U. Shmueli, (Kluwer Academic, Dordrecht, 1992).

[19] D. P. Varn, G. S. Canright, and J. P. Crutchfield (unpublished).

[20] G.S. Canright and G. Watson, J. Stat. Phys. **84**, 1095 (1996).

[21] M. Teubner, Physica A **169**, 407 (1990).

[22] G.E. Engel and R.J. Needs, J. Phys.: Condens. Matter **2**, 367 (1990).

[23] T. Proffen and T.R. Welberry, Phase Transitions **67**, 373 (1998).