

Some improvements of the ART method for finding transition pathways on potential energy surfaces

E. CANCÈS¹, F. LEGOLL¹, M.-C. MARINICA², K. MINOUKADEH¹, F. WILLAIME²

¹ Ecole Nationale des Ponts et Chaussées & INRIA, France

² CEA, DEN, Service de Recherches de Métallurgie Physique, France

<http://cermics.enpc.fr/~minoukak/>



Introduction

The Activation-Relaxation Technique (ART) [1] is a powerful method for searching saddle points and transition pathways of a given potential energy surface (PES). The zero temperature case of this algorithm is called *ART nouveau* (ART n) method [6, 4], which is the focus of our paper.

Search methods for saddle points and transition paths generally fall into two main categories:

1. **Two local minima of the PES are known** and the path between the two is to be calculated. Examples are the Nudged Elastic Band method and the String method.
2. **Only one local minimum of the PES is known** and the aim is to find a saddle point which lies on the transition path to another local minimum. Examples are the Eigenvector Following method, the Dimer method and the Hybrid Eigenvector Following method [7].

The Activation-Relaxation Technique also falls into this second category of saddle point search methods.

Main results

In our paper [2], we propose:

- Improvements to the ART n method (m-ART n):
 1. the total number of force evaluations per successful saddle point search is decreased by about 50%,
 2. the successful/unsuccessful search ratio is increased by up to 260%.
- A proof of convergence and robustness of a simplified version of the new algorithm

Activation-Relaxation Technique (ART)

Mathematical setting

The PES under consideration is defined as a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ with isolated critical points. For $x \in \mathbb{R}^d$, we denote by $\nabla f(x)$ the gradient of f at the point x and by $H(x) = \nabla^2 f(x)$ the Hessian of f at the point x . For $x \in \mathbb{R}^d$, let $\lambda_1(x) \leq \lambda_2(x) \leq \dots \leq \lambda_d(x)$ be the eigenvalues of $H(x)$ counted with their multiplicity, and let $(v_1(x), \dots, v_d(x))$ be an orthonormal basis of associated eigenvectors.

The main steps

The ART n method defines events by a two-step process:

1. **Activation:** Starting from a local minimum x_0 , the system is pushed up to a saddle point:
 - (a) First, the system is forced out of the convex region of the PES. A constant μ_A and a random local displacement vector Δx are chosen. We then:
 - i. deform the system:

$$x_{k+1} = x_k + \mu_A \Delta x$$
 - ii. relax the system in the hyperplane orthogonal to Δx using several steps of damped molecular dynamics. If, after this minimization in the hyperplane, λ_1 is still positive, repeat (a).
 - (b) Once $\lambda_1(x_k) < \lambda_s$ for some threshold $\lambda_s < 0$, the system is pushed up to a saddle point by:
 - i. taking a step in the direction of the corresponding eigenvector $v_1(x_k)$:

$$x_{k+1} = x_k - \frac{\langle \nabla f(x_k), v_1(x_k) \rangle}{\min(\lambda_1(x_k), -\lambda_c)} v_1(x_k)$$

where $\lambda_c > 0$ is an important parameter preventing the algorithm from becoming unstable in the regions where $|\lambda_1(x_k)|$ is small

- ii. relaxing the system in the hyperplane orthogonal to $v_1(x_k)$ using several steps of damped molecular dynamics. The dynamics is carried on until a maximum number of steps, n_k , is reached, or

$$\max_{i=1, \dots, d} |\nabla_i f(x_k)| < 0.01 \text{ eV/\AA}.$$

The maximum number of iterations n_k is gradually increased along the iterations. If λ_1 is found to be positive, then we start the algorithm again from x_0 with a different deformation Δx .

2. **Relaxation:** Once the criterion in the final Activation stage is satisfied, this configuration is relaxed to a new minimum, using standard optimization algorithms (in our case, a variable step steepest descent method).

Approximating $(\lambda_1(x_k), v_1(x_k))$: The Lanczos method

Contrarily to some other second order methods, such as the one proposed in [3], the ART method does not rely on the complete spectrum of the Hessian, but only on the lowest eigenmode $(\lambda_1(x_k), v_1(x_k))$.

The m-ART n method calculates this lowest eigenmode using the Lanczos method with a Krylov basis of small size ($\sim 10 - 20$), where the matrix-vector products are approximated using the first order finite-difference formula:

$$H(x_k) v \approx \frac{1}{\epsilon} (\nabla f(x_k + \epsilon v) - \nabla f(x_k))$$

References

- [1] G.T. BARKEMA AND N. MOUSSEAU, Event-based relaxation of continuous disordered systems, *Phys. Rev. Lett.* **77**, 4358 (1996).
- [2] E. CANCÈS, F. LEGOLL, M.-C. MARINICA, K. MINOUKADEH, F. WILLAIME, *to appear in J. Chem. Phys.*, (2009), [arXiv: 0806.4354](https://arxiv.org/abs/0806.4354).
- [3] C.J. CERJAN AND W.H. MILLER, On finding transition states, *J. Chem. Phys.*, **75**, 2800 (1981).
- [4] R. MALEK AND N. MOUSSEAU, Dynamics of Lennard-Jones clusters: A characterization of the activation-relaxation technique, *Phys. Rev. E*, **62**, 7723 (2000).
- [5] M.-C. MARINICA, F. WILLAIME, AND N. MOUSSEAU, Small interstitial clusters in iron BCC structure, preprint (2008).
- [6] N. MOUSSEAU AND G.T. BARKEMA, Traveling through potential energy surfaces of disordered materials: the activation-relaxation technique, *Phys. Rev. E*, **57**, 2419 (1998).
- [7] L.J. MUNRO AND D.J. WALES, *Defect migration in crystalline silicon*, *Phys. Rev. B* **59**, 3969 (1999).

Numerical results for iron BCC structure

We compare numerical results between the m-ART n algorithm and the standard ART n method in the case of basic defects in α -iron: small self interstitial (SIA) and vacancy (VAC) clusters (1 to 3 defects). The crystal consists of $1024 \pm n$ atoms ($n=1,2,3$). The crystal of α -Fe is modeled by the EAM potential.

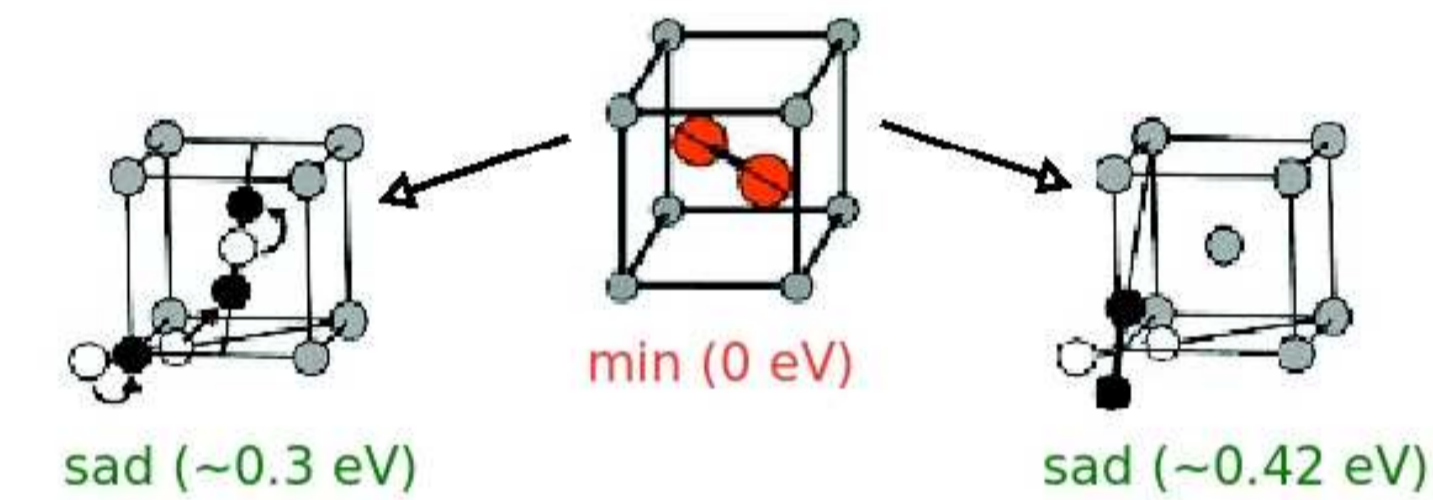
Numerical results

Below we present a comparison of a previous ART n approach [5] and the m-ART n algorithm presented in our paper for interstitial and vacancy defects. A good measure of efficiency of such algorithms is the average number of force evaluations, $\langle f \rangle$, needed to reach a saddle point and the ratio of successful to unsuccessful searches, η . The table below summarizes our results

# defects	$\langle f \rangle$	SIA		VAC	
		ART n	m-ART n	ART n	m-ART n
1	$\langle f \rangle$	462	298	780	291
	η	4.6	4.7	1.8	7.9
2	$\langle f \rangle$	548	328	705	323
	η	4.2	4.4	2.6	7.1
3	$\langle f \rangle$	691	320	667	321
	η	2.6	4.4	2.8	7.4

Table 1: The new algorithm reduces the average number of force evaluations ($\langle f \rangle$) by about 40% and 55% for the self-interstitial atoms (SIA) and vacancies (VAC) case respectively. In the case of SIA, the ratio of successful to unsuccessful searches (η) is almost constant. However, in the case of vacancies, η is increased by over 260%.

Below are two saddle points, found by ART n , corresponding to the minimum energy configuration of the $\langle 110 \rangle$ SIA (mono-interstitial) in an α -iron crystal.



Local convergence analysis

We present a proof of convergence for a simplified prototype of the m-ART n algorithm, and robustness, when the eigenvalue in the direction of the negative curvature is approximately calculated.

Prototypical algorithm

The prototypical algorithm reads

$$x_{k+1} = x_k - \frac{\langle \nabla f(x_k), v_1(x_k) \rangle}{\min(\lambda_1(x_k), -\lambda_c)} v_1(x_k) - \mu_t \Pi_{v_1(x_k)^\perp} \nabla f(x_k)$$

for a suitably chosen constant $\mu_t > 0$. Note that in this simplified algorithm, the relaxation step of the ART method is replaced by a single steepest descent step projected onto the hyperplane $v_1(x_k)^\perp$.

Notation

Let x_* be the exact configuration of the saddle point, i.e. such that $\nabla f(x_*) = 0$ and

$$\lambda_1(x_*) < 0 < \lambda_2(x_*) \leq \dots \leq \lambda_d(x_*).$$

We introduce the notation $v_1^* = v_1(x_*)$, $\lambda_1^* = \lambda_1(x_*)$, $H_* = \nabla^2 f(x_*)$, and finally let

$$\begin{aligned} e_k &= x_k - x_*, & \text{such that } |e_k| \text{ is the distance to the exact solution} \\ z_k &= (x_k - x_*) \cdot v_1^*, & \text{such that } z_k v_1^* \text{ is the projection of } e_k \text{ on } v_1^* \\ y_k &= \Pi_{v_1^*^\perp}(x_k - x_*), & \text{the projection of } e_k \text{ in hyperplane orthogonal to } v_1^* \end{aligned}$$

In order to take into account the fact that $(\lambda_1(x_k), v_1(x_k))$ is only an approximately calculated in practice, we replace this in the prototypical algorithm by the approximations $(\tilde{\lambda}_1(x_k), \tilde{v}_1(x_k))$ defined by:

$$\tilde{v}_1(x_k) = v_1(x_k) + \alpha_k, \quad \tilde{\lambda}_1(x_k) = \frac{\lambda_1(x_k)}{1 + \beta_k},$$

Assumptions

- $|\alpha_k| \ll 1$ and $|\beta_k| \ll 1$ and $\mu_t < 2/\lambda_d(x_*)$
- $\lambda_1(x_*) < -\lambda_c$, $\lambda_1(x_k) \leq -\lambda_c$ and $\tilde{\lambda}_1(x_k) \leq -\lambda_c$ for all k sufficiently large first and second derivatives

Results

Convergence is at least linear:

$$|x_{k+1} - x_*| \leq \gamma |x_k - x_*| + O(|x_k - x_*|^2) + O(|x_k - x_*| |\alpha_k|) + O(|x_k - x_*| |\beta_k|),$$

where $\gamma = \|(I - \mu_t H_*)|_{v_1^*^\perp}\|_2 < 1$. If we assume that the eigenvalues are calculated exactly, i.e. $\alpha_k = 0$ and $\beta_k = 0$ then we obtain

$$z_{k+1} = O(|x_k - x_*|^2) \quad \text{and} \quad y_{k+1} = (1 - \mu_t H_*) y_k + O(|x_k - x_*|^2).$$

That is to say that with exact λ_1 and v_1 , convergence is quadratic in the principal direction of negative curvature and linear in the orthogonal directions.