

**ΑΝΑΠΤΥΞΗ ΔΥΝΑΜΙΚΩΝ ΓΙΑ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΑΔΡΟΠΟΙΗΜΕΝΩΝ
ΜΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΜΕ ΧΡΗΣΗ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ**

**Δ.-Π. Γερακίνη^{1,3}, E. Ricci^{1,2}, Γ. Γιαννακόπουλος², Β. Καρκαλέτσης², Δ.Ν. Θεοδώρου³,
N. Βέργαδου^{1,*}**

¹Ινστιτούτο Νανοεπιστήμης & Νανοτεχνολογίας, Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών
«Δημόκριτος», Αθήνα, Ελλάδα

²Ινστιτούτο Πληροφορικής & Τηλεπικοινωνιών, Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών
«Δημόκριτος», Αθήνα, Ελλάδα

³Σχολή Χημικών Μηχανικών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Ελλάδα

(*n.vergadou@inn.demokritos.gr)

ΠΕΡΙΛΗΨΗ

Η μοριακή προσομοίωση είναι ένα ιδιαίτερα αποτελεσματικό και αξιόπιστο εργαλείο για τη μελέτη μοριακών μηχανισμών, το σχεδιασμό νέων υλικών και την πρόβλεψη των ιδιοτήτων τους. Η μελέτη πολύπλοκων χημικών συστημάτων με μεθόδους μοριακής προσομοίωσης καθιστά αναγκαία την ανάπτυξη ιεραρχικών μεθόδων για την υπολογιστική μελέτη των συστημάτων αυτών σε μεγάλο εύρος κλιμάκων μήκους και χρόνου. Οι μέθοδοι αδροποίησης (coarse-graining, CG) βρίσκονται συχνά στον πυρήνα των μεθοδολογιών αυτών. Η παρούσα εργασία εστιάζεται στην ανάπτυξη και στην εφαρμογή συστηματικής μεθοδολογίας μοριακών προσομοιώσεων ενσωματώνοντας και αλγορίθμους μηχανικής μάθησης (machine learning, ML). Αποτελεσματικές τεχνικές μηχανικής εκμάθησης και πιο συγκεκριμένα βαθιάς μάθησης, με την χρήση συνελκτικών νευρωνικών δικτύων γράφων (Convolutional Graph Neural Networks, CGNN), διερευνώνται και χρησιμοποιούνται για την ανάπτυξη πεδίων δυνάμεων στην αδροποιημένη αναπαράσταση με βάση την ατομιστική πληροφορία. Η χρήση νευρωνικών δικτύων γράφων έχει εμφανιστεί ως πολλά υποσχόμενη για την ανάπτυξη αξιόπιστων ατομιστικών δυναμικών αλληλεπίδρασης με εκπαίδευση πάνω σε κβαντομηχανική πληροφορία. Η χρήση ML για την ανάπτυξη αποτελεσματικών αδροποιημένων δυναμικών έχει ελάχιστα διερευνηθεί και η εφαρμογή της για τη μελέτη οργανικών ρευστών και πολύπλοκων χημικών συστημάτων σε συνεχείς φάσεις (bulk systems) είναι σπάνια. Η μεθοδολογία που ακολουθείται περιλαμβάνει ένα σχήμα που βασίζεται στην αντιστοίχιση δυνάμεων (force-matching scheme) χρησιμοποιώντας το υγρό βενζόλιο ως πρώτο σύστημα ελέγχου. Υιοθετούνται δύο αναπαραστάσεις για την αδροποίηση του μορίου του βενζολίου: μια αναπαράσταση από ένα κέντρο αλληλεπίδρασης και μια αναπαράσταση τριών κέντρων αλληλεπίδρασης. Τα ML CG δυναμικά που εξάγονται χρησιμοποιούνται για τη διεξαγωγή προσομοιώσεων στο αδροποιημένο επίπεδο και οι ιδιότητες που υπολογίζονται από τις προσομοιώσεις αυτές συγκρίνονται με τις αντίστοιχες ατομιστικές προκειμένου να προσδιοριστεί η αποτελεσματικότητα των μοντέλων ML που αναπτύχθηκαν. Η επίδραση των υπερπαραμέτρων, της μορφής της συνάρτησης κόστους και του μεγέθους του δικτύου διερευνήθηκαν διεξοδικά παρέχοντας μια γενική βάση για την αξιόπιστη χρήση μεθόδων CG με χρήση ML στο ευρύ πεδίο της εύπλαστης ύλης σε συνεχή φάση.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Μοντελοποίηση Πολλαπλών Κλιμάκων, Μοριακές Προσομοιώσεις, Αδροποίηση, Μηχανική Μάθηση, Συνελκτικά Νευρωνικά Δίκτυα Γράφων

ΕΙΣΑΓΩΓΗ

Η ανάπτυξη καινοτόμων τεχνολογιών που να ανταποκρίνονται στις σύγχρονες, καίριες

περιβαλλοντικές και κοινωνικές απαιτήσεις απαιτεί τον σχεδιασμό νέων υλικών με ελεγχόμενες ιδιότητες γεγονός που καθιστά αναγκαία την κατανόηση σε θεμελιακό επίπεδο των διαφορετικών αλληλεπιδράσεων και των μικροσκοπικών μηχανισμών που καθορίζουν την μακροσκοπική τους συμπεριφορά. Η μοριακή προσομοίωση βασίζεται στις θεμελιώδεις αρχές της στατιστικής μηχανικής και είναι ένα ιδιαίτερα αποτελεσματικό και αξιόπιστο εργαλείο για τη μελέτη μοριακών μηχανισμών, το σχεδιασμό νέων υλικών και την πρόβλεψη των ιδιοτήτων τους. Η μελέτη πολύπλοκων χημικών συστημάτων με μεθόδους μοριακής προσομοίωσης καθιστά αναγκαία την ανάπτυξη και εφαρμογή ιεραρχικών μεθόδων πολλαπλών κλιμάκων^[1-3], για την υπολογιστική μελέτη των συστημάτων αυτών σε μεγάλο εύρος κλιμάκων μήκους και χρόνου. Οι μεθοδολογίες αδροποίησης βρίσκονται στον πυρήνα των μεθόδων πολλαπλών κλιμάκων. Η διαδικασία της αδροποίησης αφορά στην αντικατάσταση ομάδων ατόμων με ισοδύναμα κέντρα αλληλεπίδρασης^[4,5] με τρόπο ώστε να διατηρούνται οι βαθμοί ελευθερίας που είναι σημαντικοί για τους μηχανισμούς και τις ιδιότητες υπό μελέτη. Οι αλληλεπιδράσεις εντός αυτού του χώρου μειωμένων διαστάσεων, μοντελοποιούνται μέσω αδροποιημένων δυναμικών πεδίων τα οποία αποτελούν προσεγγίσεις του δυναμικού μέσης δύναμης (potential of mean force, PMF).

Η μηχανική μάθηση έχει αυξανόμενο αντίκτυπο στις φυσικές επιστήμες, τη μηχανική και την τεχνολογία, συνεισφέροντας στην επίλυση μεγάλου εύρους ερευνητικών προβλημάτων που κυμαίνονται από μηχανισμούς μοριακών αντιδράσεων έως την υψηλής απόδοσης διαλογή λειτουργικών υλικών. Οι τεχνικές ML διερευνώνται τα τελευταία χρόνια για την ανάπτυξη ατομιστικών δυναμικών πεδίων, τα οποία να βασίζονται σε κβαντομηχανικούς υπολογισμούς^[6]. Η χρήση ML για την ανάπτυξη αποτελεσματικών αδροποιημένων δυναμικών έχει ελάχιστα διερευνηθεί και η εφαρμογή της για την μελέτη οργανικών ρευστών πολύπλοκων χημικών συστημάτων σε συνεχείς φάσεις (bulk systems) είναι πολύ περιορισμένη^[7-9].

Στην παρούσα εργασία^[10], χρησιμοποιήθηκαν αρχιτεκτονικές συνελκτικών νευρωνικών δικτύων γράφων^[11] για τη μοντελοποίηση αδροποιημένων δυναμικών εφαρμόζοντας και βελτιστοποιώντας μια μεθοδολογία^[12] που περιλαμβάνει ένα σχήμα που βασίζεται στην μέθοδο αντιστοίχισης δυνάμεων (force-matching)^[4]. Ως συστήματα ελέγχου χρησιμοποιήθηκαν συστήματα υγρού βενζολίου, στα οποία εφαρμόστηκαν δύο διαφορετικές αδροποιήσεις: με ένα και τρία αδροποιημένα κέντρα αλληλεπίδρασης ανά μόριο βενζολίου. Διερευνήθηκε διεξοδικά η επίδραση των τιμών των υπερπαραμέτρων, της μορφής της συνάρτησης κόστους και του μεγέθους της αρχιτεκτονικής του νευρωνικού δικτύου. Τα αδροποιημένα δυναμικά αλληλεπίδρασης που αναπτύχθηκαν με βάση την ML, χρησιμοποιήθηκαν για την πραγματοποίηση προσομοιώσεων μοριακής δυναμικής (MD) των υπό μελέτη συστημάτων στο αδροποιημένο επίπεδο. Στη συνέχεια, οι ιδιότητες που εξήχθησαν κατά την μετεπεξεργασία των δεδομένων συγκρίθηκαν με τις αντίστοιχες ατομιστικές προκειμένου να προσδιοριστεί η αποτελεσματικότητα των ML μοντέλων που αναπτύχθηκαν.

ΜΕΘΟΔΟΛΟΓΙΑ

Το CGNN που εφαρμόζεται στην παρούσα εργασία στηρίζεται στην αρχιτεκτονική του SchNet^[13-15]. Οι κόμβοι του γράφου αναπαριστούν τα σωματίδια και οι ακμές ορίζονται μεταξύ όλων των σωματιδίων, δηλαδή των αδροποιημένων κέντρων αλληλεπίδρασης στη δεδομένη περίπτωση, που βρίσκονται εντός ενός σφαιρικού κελύφους δεδομένης ακτίνας.

Ο καθορισμός των αδροποιημένων δυναμικών αλληλεπίδρασης πραγματοποιείται μέσω της εφαρμογής της μεθόδου αντιστοίχισης δυνάμεων^[16]. Αυτή η μέθοδος λαμβάνει υπόψη την μέση τετραγωνική Ευκλείδεια απόσταση (L2-Squared) μεταξύ των δυνάμεων που προβλέπονται από το μοντέλο, $\mathbf{F}^{CG} \in \mathbb{R}^{3N}$, για τα N αδροποιημένα σωματίδια με συντεταγμένες θέσης, \mathbf{x}_i , και των αδροποιημένων δυνάμεων, όπως αυτές υπολογίζονται για κάθε αδροποιημένο σωματίδιο με βάση τον κανόνα αδροποίησης των ατομιστικών δυνάμεων, $\mathbf{F}^A \in \mathbb{R}^{3v}$, όπου v είναι το πλήθος των ατόμων:

$$\chi^2 = \langle [\mathbf{F}^{CG}(\mathbf{x}_i) - \mathcal{M}(\mathbf{F}^A(\mathbf{r}_j))] \rangle \quad (1)$$

Στην παραπάνω έκφραση το \mathcal{M} είναι ο τελεστής αντιστοίχισης του ατομιστικού χώρου στον αδροποιημένο χώρο, \mathbf{r}_j είναι οι ατομιστικές συντεταγμένες, \mathbf{x}_i είναι οι συντεταγμένες των αδροποιημένων σωματιδίων και οι αγκύλες $\langle \dots \rangle$ αναπαριστούν την μέση τιμή ως προς όλους τους βαθμούς ελευθερίας και το πλήθος των ατομιστικών απεικονίσεων που έχουν δειγματοληπτηθεί.

Η κατάλληλη επιλογή της συνάρτησης κόστους είναι πολύ σημαντική για τη σωστή εκπαίδευση του NN. Στο πλαίσιο της παρούσας εργασίας, διαπιστώθηκε ότι συναρτήσεις κόστους που λαμβάνουν υπόψη μόνο τις δυνάμεις δεν μπορούν να εξασφαλίσουν την εκπαίδευση μοντέλων που να οδηγούν σε θερμοδυναμικά σταθερές προσομοιώσεις και να επιδεικνύουν φυσικά αποδεκτή δυναμική συμπεριφορά και χωρική οργάνωση σε σύγκριση με αυτή της ατομιστικής αναπαράστασης. Ως εκ τούτου, προτείνεται και εφαρμόζεται μια μορφή της συνάρτησης κόστους, στην οποία συμπεριλαμβάνεται και ένας όρος που υπολογίζει την τετραγωνική απόκλιση της ενέργειας που προβλέπει το μοντέλο από έναν επιλεγμένο ενεργειακό όρο της ατομιστικής προσομοίωσης (U_A). Αυτή συνάρτηση κόστους έχει την εξής συναρτησιακή μορφή:

$$L = \frac{1}{3N} \sum_i^N \left(-\nabla_{\mathbf{r}_i} U^{CG} - \mathcal{M}(\mathbf{F}^A(\{\mathbf{x}_k\}))_i \right)^2 + \lambda (U^{CG} - U_A)^2 \quad (2)$$

όπου N ο συνολικός αριθμός των αδροποιημένων σωματιδίων, λ είναι μια υπερπαραμέτρος που ρυθμίζει την κλιμάκωση της συνεισφοράς του ενεργειακού όρου στην συνάρτηση κόστους, $\mathcal{M}(\mathbf{F}^A(\{\mathbf{x}_k\}))_i$ είναι το διάνυσμα δύναμης που ασκείται στο αδροποιημένο κέντρο αλληλεπίδρασης i και το οποίο προκύπτει ως συνισταμένη των δυνάμεων που ασκούνται στα ατομιστικά κέντρα αλληλεπίδρασης που αντιστοιχίστηκαν κατά την διαδικασία αδροποίησης στο ενεργό κέντρο i , $\{\mathbf{x}_k\}$ είναι ένα διάνυσμα θέσης του ατομιστικού συστήματος εντός του χώρου των απεικονίσεων του και $-\nabla_{\mathbf{r}_i} U^{CG} = \mathbf{F}^{CG}(\mathbf{x}_i)$, όπου U^{CG} είναι η ενέργεια που προβλέπεται από το μοντέλο.

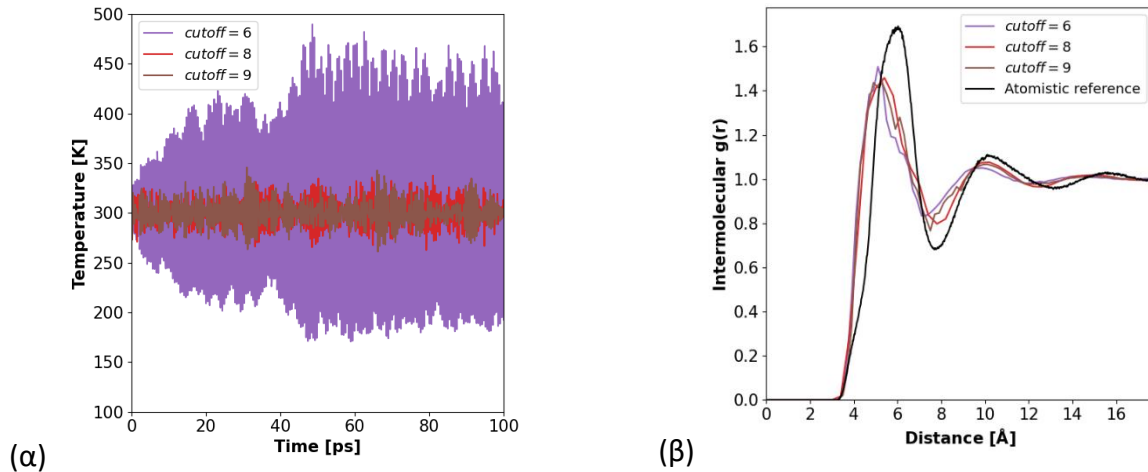
Τα δεδομένα με τα οποία εκπαιδεύτηκε το μοντέλο παράχθηκαν από ατομιστικές προσομοιώσεις MD υγρού βενζολίου στους 300K και σε 1atm, χρησιμοποιώντας περιοδικές οριακές συνθήκες και το δυναμικό αλληλεπίδρασης Amber Cornell^[17]. Τα συστήματα εξισορροπήθηκαν μέσω μιας NPT προσομοίωσης διάρκειας 1ns ακολουθούμενης από μια προσομοίωση NVT 20ns, οι οποίες πραγματοποιήθηκαν με χρήση του LAMMPS^[18]. Μελετήθηκε ένα σύστημα 500 μορίων στην αδροποιημένη αναπαράσταση, αντιστοιχίζοντας κάθε μόριό του σε ένα ισοδύναμο κέντρο αλληλεπίδρασης. Επίσης, αδροποιήθηκε ένα σύστημα 300 μορίων αντιστοιχίζοντας κάθε μόριό του σε τρία ισοδύναμα κέντρα αλληλεπίδρασης, τα οποία συνδέονται μεταξύ τους μέσω ισοδύναμων δεσμών. Από τα τελευταία 10000 βήματα των προσομοιώσεων χρησιμοποιήθηκαν: το 90% ως δεδομένα εκπαίδευσης (training) του μοντέλου και το εναπομένον 10% ως δεδομένα ελέγχου (test). Τα 10000 βήματα διαχωρίστηκαν με τυχαίο τρόπο στις δύο αυτές κατηγορίες.

ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΣΥΖΗΤΗΣΗ

Η μεθοδολογία εφαρμόστηκε για την ανάπτυξη ML μοντέλων για τον προσδιορισμό δυναμικών αλληλεπίδρασης στην αδροποιημένη αναπαράσταση, διάφορες συνθήκες εκπαίδευσης, εστιάζοντας, μεταξύ άλλων: (α) στο μέγεθος της αρχιτεκτονικής του NN, (β) στην υπερπαραμέτρο ακτίνας αποκοπής (*cutoff*), η οποία ορίζει την γειτονιά κάθε αδροποιημένου σωματιδίου, (γ) εξετάσαμε την ικανότητα μεταφοράς των εξαγόμενων μοντέλων χρησιμοποιώντας τα για την πραγματοποίηση CG προσομοιώσεων υπό άλλες θερμοκρασιακές συνθήκες ή και μέγεθος συστήματος από αυτά στα οποία εκπαιδεύτηκαν, (δ) εκπαιδεύσαμε μοντέλα με διάφορους όγκους δεδομένων εκπαίδευσης και (ε) διερευνήσαμε ως συνάρτηση ενεργοποίησης των νευρωνικών επιπέδων την συνάρτηση shifted softplus, πέραν της tanh. Για την αδροποιημένη αναπαράσταση ενός σωματιδίου ανά μόριο του συστήματος βενζολίου, πραγματοποιήθηκαν περισσότερες από 150 δοκιμές, στις οποίες εξετάσαμε διαφορετικούς συνδυασμούς υπερπαραμέτρων διερευνώντας

την επίδραση των διαφορετικών υπερπαραμέτρων στην διαδικασία εκπαίδευσης, στην ευστάθεια των προσομοιώσεων και στην επίδοση των μοντέλων που αναπτύχθηκαν.

Στο Σχήμα 1 παρουσιάζονται αποτελέσματα στα οποία για συγκεκριμένο συνδυασμό υπερπαραμέτρων, εξετάζεται η επίδραση της τιμής της υπερπαραμέτρου cutoff, η οποία ελέγχει το εύρος της τοπικής γειτονιάς κάθε αδροποιημένου σωματιδίου. Από αυτές τις συγκριτικές γραφικές παραστάσεις παρατηρούμε ότι για $cutoff = 6\text{ \AA}$ λαμβάνουμε θερμοδυναμικά ασταθή προσομοίωση. Για μεγαλύτερες τιμές του $cutoff$, η προσομοίωση είναι σταθερή ενώ η θέση της πρώτης κορυφής του διαμοριακού $g(r)$ παραμένει σχεδόν αμετάβλητη.



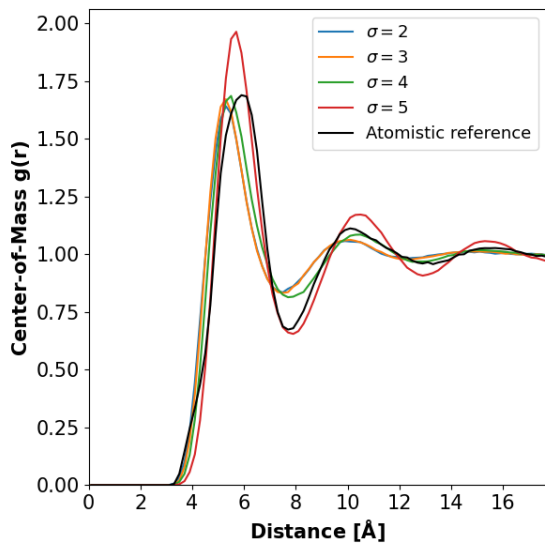
Σχήμα 1. Συγκριτικές γραφικές παραστάσεις (α) της εξέλιξης της θερμοκρασίας και (β) της συνάρτησης ακτινικής κατανομής μεταξύ των κέντρων μάζας των αδροποιημένων μορίων, ($\sigma=5\text{ \AA}$, $n=7$, $\lambda=1$) για διάφορες τιμές του cutoff, $T = 300\text{ K}$, $p = 1\text{ atm}$, για ένα σύστημα βενζολίου 500 μορίων.

Ομοίως, κατά τη μελέτη της αδροποιημένης αναπαράστασης με τρία σωματίδια ανά μόριο του συστήματος βενζολίου, μελετήσαμε διάφορες συνθήκες εκπαίδευσης και συνολικά, πραγματοποιήθηκαν περισσότερες από 200 δοκιμές, στις οποίες εξετάσαμε διαφορετικούς συνδυασμούς. Στην περίπτωση αυτής της αναπαράστασης, εισάγεται μια ακόμη υπερπαραμέτρος, η οποία εκφράζει την δυσκαμψία των δεσμών εντός των αδροποιημένων μορίων.

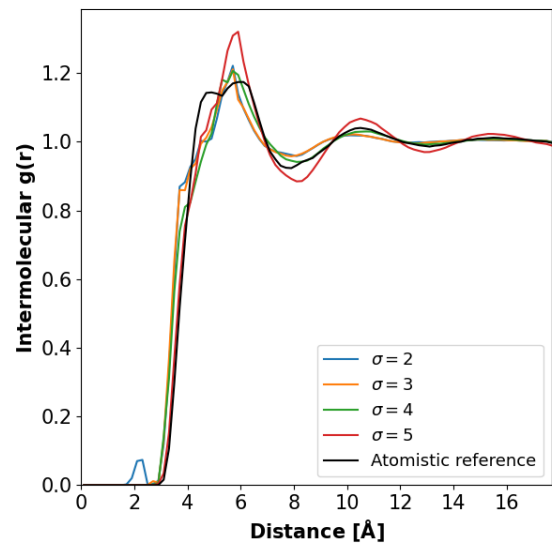
Στο Σχήμα 2 παρουσιάζονται τα αποτελέσματα προσομοιώσεων μοριακής δυναμικής στο αδροποιημένο επίπεδο με συγκεκριμένο συνδυασμό υπερπαραμέτρων και για διάφορες τιμές της υπερπαραμέτρου σ που σχετίζεται με την ακτίνα του αποκλειόμενου όγκου που θεωρείται για κάθε σωματίδιο και συμπεριλαμβάνεται ως πρόσθετος ενεργειακός όρος. Σε αυτές τις συγκριτικές γραφικές παραστάσεις παρατηρείται μια μονότονη συμπεριφορά όσον αφορά το ύψος της πρώτης αφύσικης κορυφής της διαμοριακής συνάρτησης ακτινικής κατανομής, $g(r)$, συναρτήσεως του σ . Πιο συγκεκριμένα, καθώς η τιμή του σ αυξάνει, το ύψος της μη-φυσικά αποδεκτής κορυφής στα $\sim 2.3\text{ \AA}$ μειώνεται και εξαφανίζεται για $\sigma \geq 3$. Την ίδια στιγμή στην γραφική παράσταση της ακτινικής κατανομής των κέντρων μάζας παρατηρείται μια έντονη κορυφή για μεγάλα σ μαζί με μια ελαφρά μετατόπισή της δεξιότερα, προς μεγαλύτερες αποστάσεις, βελτιώνοντας έτσι την συμφωνία με τα ατομικά δεδομένα.

Στο Σχήμα 3 παρουσιάζονται τα αποτελέσματα προσομοιώσεων μοριακής δυναμικής στο αδροποιημένο επίπεδο με χρήση μηχανικής μάθησης στα οποία μεταβάλλεται η υπερπαραμέτρος στάθμισης του ενεργειακού όρου στην συνάρτησης κόστους, λ , για ένα συγκεκριμένο συνδυασμό υπερπαραμέτρων. Σε αυτή την περίπτωση, διαφαίνεται επίσης ότι είναι αναγκαία η χρήση του ενεργειακού όρου εντός της συνάρτησης κόστους προκειμένου να μπορεί να πραγματοποιηθεί μια σταθερή προσομοίωση. Ο συνδυασμός υπερπαραμέτρων που παρουσιάζεται στο Σχήμα 3 (για $\lambda \neq 0$) πέραν των σταθερών προσομοιώσεων παρουσιάζει και καλή συμφωνία όσον αφορά την δομή με την συνάρτηση ακτινικής κατανομής των κέντρων μάζας, ενώ η περίπτωση με $\lambda = 1$ φαίνεται

να είναι σε θέση να αναπαράγει την διπλή κορυφή που παρατηρείται στην διαμοριακή συνάρτηση ακτινικής κατανομής.

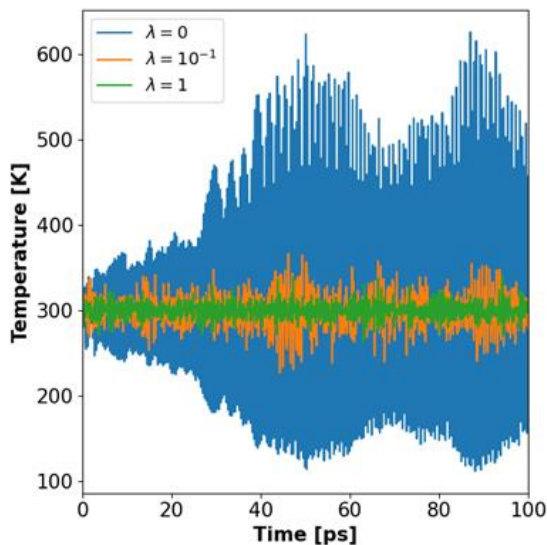


(α)

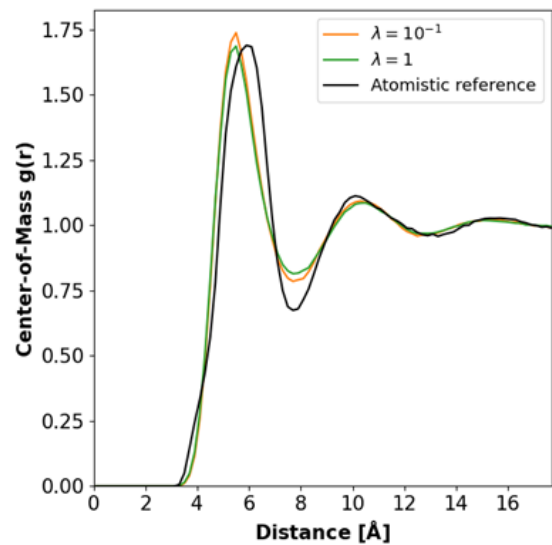


(β)

Σχήμα 2. Συγκριτικές γραφικές παραστάσεις (α) της κατανομής ακτινικής απόστασης μεταξύ των κέντρων μάζας των αδροποιημένων μορίων της και (β) της διαμοριακής συνάρτησης ακτινικής κατανομής μεταξύ των διαμοριακών αδροποιημένων σωματιδίων, ($cutoff=6\text{\AA}$, $n=6$, $\lambda=1$, $k_b=1000$) για διάφορες τιμές του σ , $T = 300\text{K}$, $p = 1\text{atm}$, για ένα σύστημα βενζολίου 300 μορίων.



(a)



(b)

Σχήμα 3. Συγκριτικές γραφικές παραστάσεις (α) της εξέλιξης της θερμοκρασίας και (β) της συνάρτησης ακτινικής κατανομής μεταξύ των κέντρων μάζας των αδροποιημένων μορίων, ($cutoff=6\text{\AA}$, $\sigma=4\text{\AA}$, $n=6$, $k_b=1000$) για διάφορες τιμές του λ , $T = 300\text{K}$, $p = 1\text{atm}$, για ένα σύστημα βενζολίου 300 μορίων.

Στην παρούσα εργασία αναγνωρίστηκαν αρκετές ανοικτές προκλήσεις κατά την προσομοίωση συστημάτων συνεχούς φάσης, στο αδροποιημένο επίπεδο, με χρήση δυναμικών αλληλεπίδρασης που έχουν μοντελοποιηθεί με χρήση ML, συνεισφέροντας έτσι στην βαθύτερη κατανόηση αυτού του καινοτόμου πεδίου. Αντίθετα με πολλές συμβατικές εφαρμογές της ML, το γεγονός της σύγκλισης της συνάρτησης κόστους σε μια τιμή δεν διαπιστώθηκε να αποτελεί αξιόπιστο μέτρο αξιολόγησης της απόδοσης των εκπαιδευμένων μοντέλων κατά την προσομοίωση στο αδροποιημένο επίπεδο. Παράλληλα, διαφάνηκε η ανάγκη χρήσης μιας συνάρτησης κόστους με συναρτησιακή μορφή πολλών όρων, αναδεικνύοντας την ανάγκη ανάπτυξης κατάλληλων στρατηγικών εκπαίδευσης. Μελλοντικά βήματα της μελέτη περιλαμβάνουν την επέκταση της

τρέχουσας μεθόδου ώστε να συνυπολογίζει κατευθυντικά (directional) χαρακτηριστικά κατά τον υπολογισμό του τοπικού περιβάλλοντος. Οι προαναφερθείσες τροποποιήσεις της μεθοδολογίας είναι καίριας σημασίας στο πλαίσιο της προσπάθειας ανάπτυξης αδροποιημένων δυναμικών αλληλεπίδρασης που να διαθέτουν πιστότητα (accuracy), συνέπεια (consistency), δυνατότητα μεταφοράς (transferability), την ικανότητά να ενσωματώνουν συσχετίσεις υψηλότερης τάξης, καθώς και την ικανότητα να μπορούν εν δυνάμει να ενταχθούν σε ένα γενικευμένο πλαίσιο αδροποίησης^[19], διαβλέποντας ως τελικό στόχο την εφαρμογή τους σε ιεραρχικές προσομοιώσεις πολλαπλών κλιμάκων για συστήματα μαλακής ύλης στη συνεχή φάση.

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα εργασία υποστηρίχθηκε εν μέρει από την δράση Marie Sklodowska-Curie (No 101030668) του προγράμματος Horizon 2020 έρευνας και καινοτομίας της Ευρωπαϊκής Ένωσης και πραγματοποιήθηκε με χρήση υπολογιστικού χρόνου που παραχωρήθηκε από το Εθνικό Δίκτυο Υποδομών, Καινοτομίας και Έρευνας (ΕΔΥΤΕ Α.Ε. – GRNET S.A.) στην Εθνική υποδομή HPC – ARIS – στο πλαίσιο των έργων MULTIPOLS (ID:011032), MULTIPOLS II (ID: 013019) και ML-SOFT (ID: 015020).

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Theodorou, DN. (2007). *Chem. Eng. Sci.* 62, 5697–5714.
- [2] Gooneie, A, Schuschnigg, S, Holzer, C. (2017). *Polymers (Basel)*. 9, 16.
- [3] Vergadou, N, Theodorou, DN. (2019). *Membranes (Basel)*. 9, 98.
- [4] Noid, W, Chu, J-W, Ayton, G, Krishna, V, Izvekov, S, Voth, G, Das, A, Andersen, H. (2008). *J. Chem. Phys.* 128, 244114.
- [5] Brini, E, Algaer, EA, Ganguly, P, Li, C, Rodríguez-Roperro, F, van der Vegt, NFA. (2013). *Soft Matter* 9, 2108–2119.
- [6] Unke, OT, Chmiela, S, Sauceda, HE, Gastegger, M, Poltavsky, I, Schütt, KT, Tkatchenko, A, Müller, K-R. (2021). *Chem. Rev.* 121, 10142–10186.
- [7] Ricci, E, Vergadou, N. (2023). *J. Phys. Chem. B* 127, 2302–2322.
- [8] Ruza, J, Wang, W, Schwalbe-Koda, D, Axelrod, S, Harris, WH, Gómez-Bombarelli, R. (2020). *J. Chem. Phys.* 153, 164501.
- [9] Husic, BE, Charron, NE, Lemm, D, Wang, J, Pérez, A, Majewski, M, Krämer, A, Chen, Y, Olsson, S, De Fabritiis, G, Noé, F, Clementi, C. (2020). *J. Chem. Phys.* 153.
- [10] Gerakinis, D-P, Ricci, E, Giannakopoulos, G, Karkaletsis, V, Theodorou, DN, Vergadou, N. (2024, January). doi: 10.5281/zenodo.10501037
- [11] Batzner, S, Musaelian, A, Sun, L, Geiger, M, Mailoa, JP, Kornbluth, M, Molinari, N, Smidt, TE, Kozinsky, B. (2022). *Nat. Commun.* 13, 2453.
- [12] Ricci, E, Giannakopoulos, G, Karkaletsis, V, Theodorou, DN, Vergadou, N. (2022)., in *Proceedings of the 12th Hellenic Conference on Artificial Intelligence*, pp 1–6. ACM, New York, NY, USA. doi: 10.1145/3549737.3549793
- [13] Schütt, KT, Kindermans, P-J, Sauceda, HE, Chmiela, S, Tkatchenko, A, Müller, K-R. (2017)., in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp 992–1002. Curran Associates Inc., Red Hook, NY, USA.
- [14] Chmiela, S, Tkatchenko, A, Sauceda, HE, Poltavsky, I, Schütt, KT, Müller, K-R. (2017). *Sci. Adv.* 3.
- [15] Schütt, KT, Kessel, P, Gastegger, M, Nicoli, KA, Tkatchenko, A, Müller, K-R. (2019). *J. Chem. Theory Comput.* 15, 448–455.
- [16] Ercolessi, F, Adams, JB. (1994). *Europhys. Lett.* 26, 583–588.
- [17] Cornell, WD, Cieplak, P, Bayly, CI, Gould, IR, Merz, KM, Ferguson, DM, Spellmeyer, DC, Fox, T, Caldwell, JW, Kollman, PA. (1995). *J. Am. Chem. Soc.* 117, 5179–5197.
- [18] Plimpton, S. (1995). *J. Comput. Phys.* 117, 1–19.
- [19] Nasikas, D, Ricci, E, Giannakopoulos, G, Karkaletsis, V, Theodorou, DN, Vergadou, N. (2022)., in *Proceedings of the 12th Hellenic Conference on Artificial Intelligence*, pp 1–8. ACM, New York, NY, USA. doi: 10.1145/3549737.3549792