Volume 2 March 2024

JOURNAL of PHYSICS Research and Education

JPRE

DEPARTMENT OF PHYSICS
University of North Bengal
Darjeeling – 734013, West Bengal, India

ISSN: XXXX-XXXX (Print); XXXX-XXXX (Online)

ISSN xxxx-xxxx

JOURNAL *of* Physics Research and Education **Vol. 2** **March, 2024**

CONTENTS

Editorial Note

We are happy to bring out this second volume of the Departmental journal, Journal *of* Physics Research and Education. We apologize for the inordinate delay in bringing out this volume. We aim to publish a new volume regularly in every year. The goal of *JPRE* is to have an amalgamating impact for the diversity of the fields of physics research and physics education. The current volume has a distinctive aspect of this nature contributed together by Ph.D. scholars, Post-doctoral scholars and Professors. I hope this volume will reflect the current state of interdisciplinary research, insightful reviews and will indicate the usefulness of their material to physics education.

Rajat K. Dey
Editor-in-Chief

Editors

Editor-in-Chief
Dr. Rajat Kumar Dey, Associate Professor, Department of Physics, University of North Bengal
Associate and Managing Editor
Dr. Tutul Biswas, Assistant Professor, Department of Physics, University of North Bengal
Assistant Editors
Mr. Lakpa Tamang, Assistant Professor, Department of Physics, University of North Bengal Dr. Prakriti Adhikary, Assistant Professor, Department of Physics, University of North Bengal

Advisory Board

Dr. Arunava Chakraborti, Professor, Presidency University, Kolkata
Dr. Tarun Kanti Ghosh, Professor, Indian Institute of Technology-Kanpur
Dr. Subhasish Chattopadhyay, Professor, Variable Energy Cyclotron Centre, Kolkata
Dr. Kamal Kanti Nandi, Professor, Zel'dovich International Centre for Astrophysics, Ufa, Russia
Dr. Subir Ghosh, Professor, Indian Statistical Institute, Kolkata
Dr. Arunava Bhadra, Director, High Energy & Cosmic Ray Research Centre, University of North Bengal
Dr. Bannanje Sripathi Acharya, Retired Professor, Tata Institute of Fundamental Research, Mumbai
Dr. Subinit Roy, Retired Professor, Saha Institute of Nuclear Physics, Kolkata
Dr. Shyamal Kumar Bhadra, Emeritus Professor, Indian Association for the Cultivation of Science, Kolkata

Editorial Board

Dr. Kamal Kanti Nandi, Professor, Zel'dovich International Centre for Astrophysics, Ufa, Russia
Dr. Arunava Bhadra, Director, High Energy & Cosmic Ray Research Centre, University of North Bengal
Dr. Subinit Roy, Retired Professor, Saha Institute of Nuclear Physics, Kolkata
Dr. Pradip Kumar Mandal, Retired Professor, Department of Physics, University of North Bengal
Dr. Sailojananda Mukherjee, Retired Professor, Department of Physics, University of North Bengal
Dr. Swapan Kumar Ghosal, Retired Professor, Department of Physics, University of North Bengal
Dr. Amitava Mukhopadhyay, Professor, Department of Physics, University of North Bengal
Dr. Malay Kumar Das, Professor, Department of Physics, University of North Bengal
Dr. Bikash Chandra Paul, Professor, Department of Physics, University of North Bengal
Dr. Sripada Haldar, Professor, Department of Physics, University of North Bengal
Dr. Suman Chatterjee, Professor, Department of Physics, University of North Bengal
Dr. Provash Mali, Associate Professor, Department of Physics, University of North Bengal

Articles

J. Phys. Res. Edu., Vol. 2, March 2024

Variable thermoelectric parameters in Si / Ge zNR by electrostrictive application of localized strain

Amretashis Sengupta*

Department of Physics, P. D. Women's College, Jalpaiguri, West Bengal-735101, India

In this work, we propose the tuning of thermoelectric performance of zigzag monolayer Silicene and Germanene nanoribbon (SiNR/GeNR) with a reversible strain engineering method. By a proposed electrostrictive method a tensile or compressive moderate strain of ±2% orthogonal to the transport direction in a short region of the SiNR or GeNR is considered to be applied. A self-consistent density functional tight binding (DFTB) approach is employed for the calculation of the electronic properties of the system, while the vibrational properties are computed with classical molecular dynamics simulations. Electron/phonon transport is computed with the Green's function formalism. With the localized strain application it is observed that electron transmission and current through the Si or GeNR remains largely unaffected while a suppression of the phonon transport and thermal conductance can be achieved. A significant tuning is observed for thermoelectric figure of merit and variations are seen in the Seebeck coefficient and the thermoelectric power factor. Moreover the temperature and doping dependencies of these parameters also showed high degree of tunability with strain. The enhancements in thermoelectric figure of merit by such simple strain ON/OFF mechanism in a CMOS compatible architecture suggest good prospects for nanoscale thermoelectrics.

I. INTRODUCTION

Among the two-dimensional (2D) Dirac materials, Silicene and Germanene have general interest for the semiconductor industry due to their electronic properties and elemental familiarity with the planar CMOS process [1–8]. The prospect of these 2D materials for device and interconnect applications in the next generation of VLSI chips is realistic and also augers well for flexible electronics [5–8]. Moreover, such materials are also being explored for applications in optoelectronics, solar cells and chemical sensors [3–11].

Silicene and Germanene nanoribbons (SiNR and GeNR), have also been successfully synthesized [12] and studied in a number of theoretical and experimental works for their electrical parameters, spin transport, hydrogen sensing and other properties [2–16]. Like graphene nanoribbons these 1D nanostructures can be semi-metallic or metallic depending on their chirality and width [2–4, 17–19].

With the prospect of SiNR and GeNR interconnects and devices, the factor of application specific tai-

loring of material properties is also important. The electronic properties in Si and Ge 2D sheets are found to be altered with external electromagnetic or mechanical stimuli better than graphene due to their buckled honeycomb structure [3, 4, 20–22]. The tuning of electro-thermal properties in nanostructures made of Silicene or Germanene is also been investigated of late [20–30]. This is quite important from the point of view of thermoelectric devices, thermal management in integrated circuits and other applications such as NEMS. An important aspect of tailoring material properties in nanoelectronics application would be the reversibil-

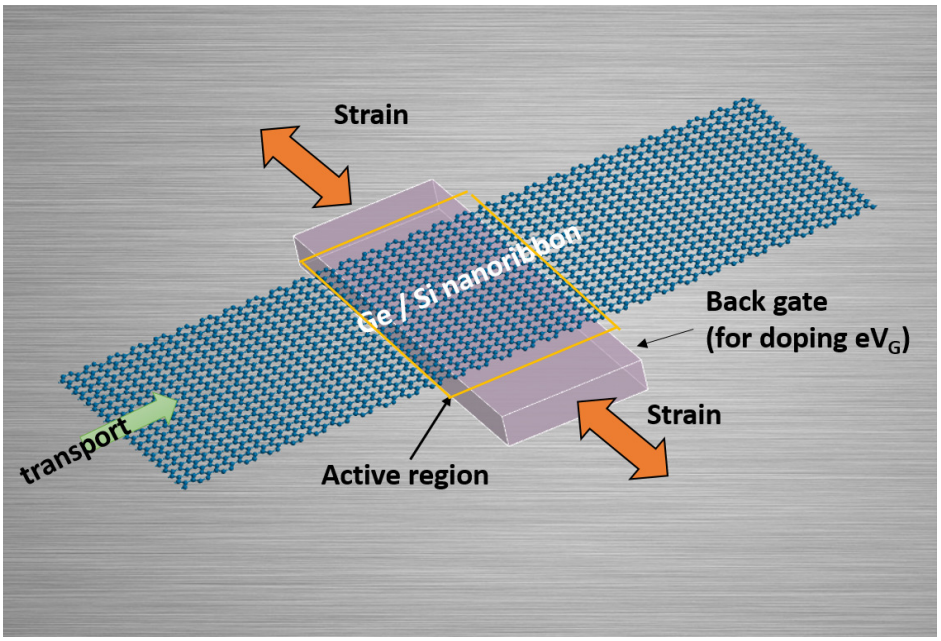

FIG. 1: Schematic of the proposed straining mechanism (not to scale).

ity of the induced enhancement/suppression of certain parameters. While much work has been carried out on the impact of line and point defects and dislocations on the thermoelectric behavior of Dirac materials [17–32] such topological changes are difficult to form in a controlled manner and are not something that can be switched on/off at will. The application of compressive vertical strain by means of electrostriction of a piezoelectric "gate" terminal to switch a bilayer MoS_2 channel has also been proposed by Das [33]. While this can be a highly efficient method to change electrical properties of multi-layered structures with changing the interlayer van der Waals gap, for monolayer systems the application of in-plane strain is more effective for tuning material properties [34] Such a system may be achieved by an in-plane electrostrictive architecture consisting of SiNR and GeNR being subject to a localized strain applied by a piezoelectric finger (or control bar) in a direction orthogonal to carrier transport [33, 34] Depending upon the polarity of the voltage applied on the bar the region of SiNR / GeNR suspended on it can be expanded and contracted. The issue of weak strain transfer in such an layered architecture may be mitigated by encapsulation

of the region of SiNR or GeNR within the piezo-material, such as ZnO or PMN-PT. This gives rise to three possible states of such a system/ device, the perfect state where no bias is applied to the control bar, the state of tensile (positive) strain, and that of compressive (negative) strain. Fig. 1 shows the schematic of the proposed straining mechanism.

Considering a moderate strain of ±2% and the perfect case (0% strain), we compute the electronic and thermo-electric parameters of zigzag SiNR and GeNR devices by means of atomistic simulations. With density functional tight binding (DFTB) formalism [35–37] the electronic properties of the system are evaluated, while the phonon calculations are performed with the frozen phonon approximation [38, 39] with a small-displacement method using classical potentials [38, 40, 41] The electron and phonon transport calculations are done with the Landuer approach within the non-equilibrium Green function (NEGF) framework [38, 42–46]. The thermoelectric parameters were evaluated with the PyQuanTE code [32] We calculate the carrier transport, phonon transmission, spectral currents and the various thermoelectric parameters namely the Peltier and Seebeck coefficients, thermoelectric conductance, thermoelectric figure of merit and the thermoelectric power factor. The variation of thermoelectric parameters with temperature was also studied in detail.

II. METHODOLOGY

We consider$(8, 8)$ zigzag SiNR and GeNR with comparable width of ∼5.5 nm. The proposed device consists of ideal semi-infinite contacts connected to a central region which is about 25 nm long. At the middle of this central region the ±2% strain is considered on a section of 5 nm long 'active' area (indicated in the schematic Fig. 1). The length of the central region ensures sufficient screening distance available on both sides of this active area, so as to properly match with the perfect contacts.

With density functional tight binding (DFTB) formalism the electronic properties of the system is simulated [38]. The tight binding method is chosen for it's high computational efficiency for handling systems with such large number of atoms. We use the Slater-Koster parameters for Si and Ge [38, 47] for our calculations. $1 \times 1 \times 24$ k-points in a Monkhorst-Pack grid [48] are used for sampling the device supercell. The phonon calculations are carried out under the frozen phonon approximation [38, 39] using the classical Stillinger-Weber [40, 41] type classical potential. A central finite difference method, inclusive of the acoustic sum rule with 180 atomic displacements of 0.01 magnitude, was employed for calculation of the symmetric dynamical matrix.

The electron current flowing between the two semi-infinite ideal contacts (1 and 2) is evaluated with

Landauer formula [42–45]

$$I = \frac{q}{h} \int_{-\infty}^{+\infty} dE \, \mathrm{Im}_{el}(E) \left[f_{FD}\left(E - \mu_1, T_1\right) - f_{FD}\left(E - \mu_2, T_2\right) \right], \tag{1}$$

where $f_{FD}(.)$ is the Fermi-Dirac distribution function, μ and T the electrochemical potential and the temperature of the contacts respectively, and Im_{el} the electron transmission function. For a small temperature difference, the electrical conductance derived from Eq. (1) has the following form [32, 38, 46]

$$G_e = \frac{2e^2}{h} \int_{-\infty}^{+\infty} dE \, \mathrm{Im}_{el}(E) \left(-\frac{df_{FD}(E,T)}{dE} \right). \tag{2}$$

The heat current I_Q develops between contacts due to temperature gradient and is related to the thermal conductance κ as [32–38]

$$\kappa = \left. \frac{dI_Q}{dT} \right|_{I=0}. \tag{3}$$

The methodology for electron transport and phonon transport is available in detail in standard literature, such as Refs. [42–46].

The electron thermal conductance can be written as [32]

$$\kappa_e = \frac{2}{k_B T^2 h} \left(L_2 - \frac{L_1^2}{L_0} \right), \tag{4}$$

where the Onsager coefficient L_n is expressed as [32, 49]

$$L_n = \int_{-\infty}^{+\infty} dE (E - E_F)^n \mathrm{Im}_{el}(E) \left(-\frac{df_{FD}(E,T)}{dE} \right). \tag{5}$$

In the low temperature difference limit, the phonon thermal conductance can be defined as

$$\kappa_{ph} = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \hbar\omega \, \mathrm{Im}_{ph}(\omega) \left(\frac{\partial f_{BE}(\omega,T)}{\partial T} \right), \tag{6}$$

where Im_{ph} is the phonon transmission, $f_{BE}(.)$ is the Bose-Einstein distribution function and ω the phonon frequency.

The Seebeck coefficient or thermopower (S) can be expressed as [45–49]

$$S(T) = -\frac{1}{eT} \frac{L_1}{L_0}. \tag{7}$$

The thermoelectric figure of merit ZT is given as [45–49]

$$ZT = \frac{G_e S^2 T}{\kappa_e + \kappa_{ph}}. \tag{8}$$

III. RESULTS AND DISCUSSION

In our calculations the electron-phonon and the phonon-phonon couplings were not taken into account since the scattering length in such Dirac materials greatly exceed the dimensions considered herein [2, 3, 17–32].

The phonon transmission spectra as in Fig. 2(b) shows significant suppression upon the application of ±2% strain in the NR structures. A tensile strain shows slightly less suppression at lower energies as compared to compressive strain for GeNR. For the SiNR both strained conditions show a very similar nature of reduction from the unstrained condition. For GeNR the phonon transmission states exist mostly between 0 to 0.04 eV, for SiNR the range expands to 0 to 0.08 eV. The greater suppression of the phonon transmission against electron transmission upon introduction of non-uniformity in lattice is something that is observed in other 1 dimensional systems as CNTs [32].

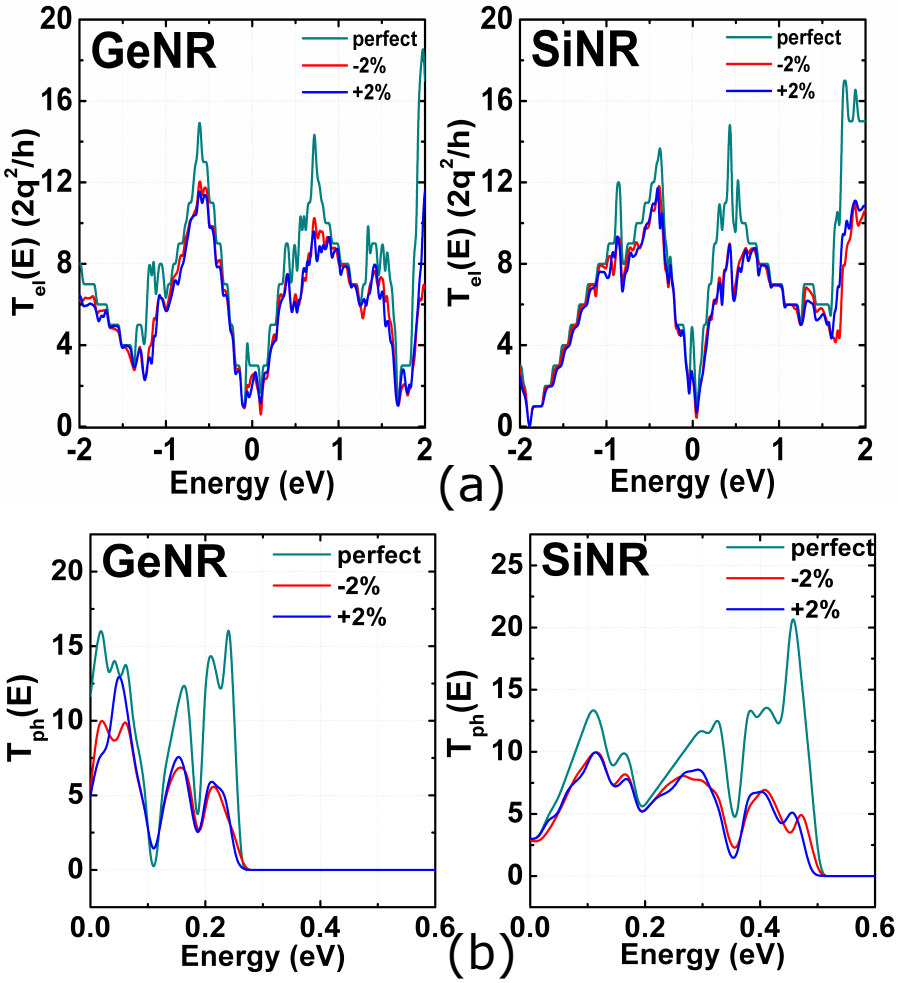

FIG. 2: (a) Electron and (b) phonon transmission spectra of the various devices at equilibrium condition.

In Fig. 3, the conductance of the various conditions of the SiNR and GeNR systems for variation in temperature and Fermi level position is shown. The change in fermi level position is considered resulting from a variation in chemical doping or electrostatic doping by the application of a gate voltage (VG). We designate this Fermi level modulation as eV_G in our plots and refer to them as simply doping. We see that with the increase in temperature the peaks of the conductance reduces in magnitude, at the same time the peaks become more broadened and less strong for the same amount of doping.

In the presence of a symmetry breaking event as strain, the coherence of the transmission states is lost to some degree resulting in the reduction in the conductance. The distinct fluctuations at 100 K smoothen out as the temperature is increased and can be attributed to change in the Fermi distribution of the system. The conductance in both SiNR and GeNR seem vary more with change in temperature, near the maxima or minima than in other positions of the plots. Also the GeNR displays slightly more conductance against SiNR structures for higher amounts of doping.

Coming to the thermal conductivities in Fig. 4, we can see a reduction in the thermal conductivity of the SiNR and GeNR configurations upon the application of both tensile and compressive localized strain. This is consistent with our aim of enhancing thermoelectric qualities of the nanoribbons with strain engineering. For the low temperatures the contribution due to phonons is shown by dashed lines. An interesting observation at this point is the dominance of the electron contribution to the thermal conductivity over the lattice part which is largely due to the metallic nature of zigzag SiNR and GeNR. For these NRs the electronic κ_e and lattice κ_{ph} part of the thermal conductivity are at best equal at low doping levels at 100 K while for 300 K and above the electronic part becomes dominant even for the undoped condition.

The Seebeck coefficients of the SiNR and GeNR shown in Fig. 5, show a minor increase with strain application (both tensile and compressive) for the same level of doping at temperatures in excess of 300 K. In the SiNR it is seen that with increase in negative doping and temperature the plots tend to bunch together more than that in case of GeNR. An oscillatory behavior of S with doping is seen in all cases, with a gradual smoothing of the fluctuations with increase in temperature. Even at temperatures at and above 300 K, the change of sign of the Seebeck coefficient with change in doping sign and also magnitude is present. For instance, for the GeNR a changeover from positive values to negative value of Seebeck coefficient occurs around -0.6 eV for the GeNR. On the positive doping side a changeover of S from the negative to positive value occurs around 0.8 eV for the GeNR. For SiNR these turning point is around -0.4 eV ($S > 0$ to $S < 0$) and +0.4 eV (for $S < 0$ to $S > 0$). This indicates a modulation of the thermal transport behavior between electron dominated (where $S < 0$) and hole dominated (where $S > 0$) regimes in such SiNR and GeNR systems as the doping and temperature changes.

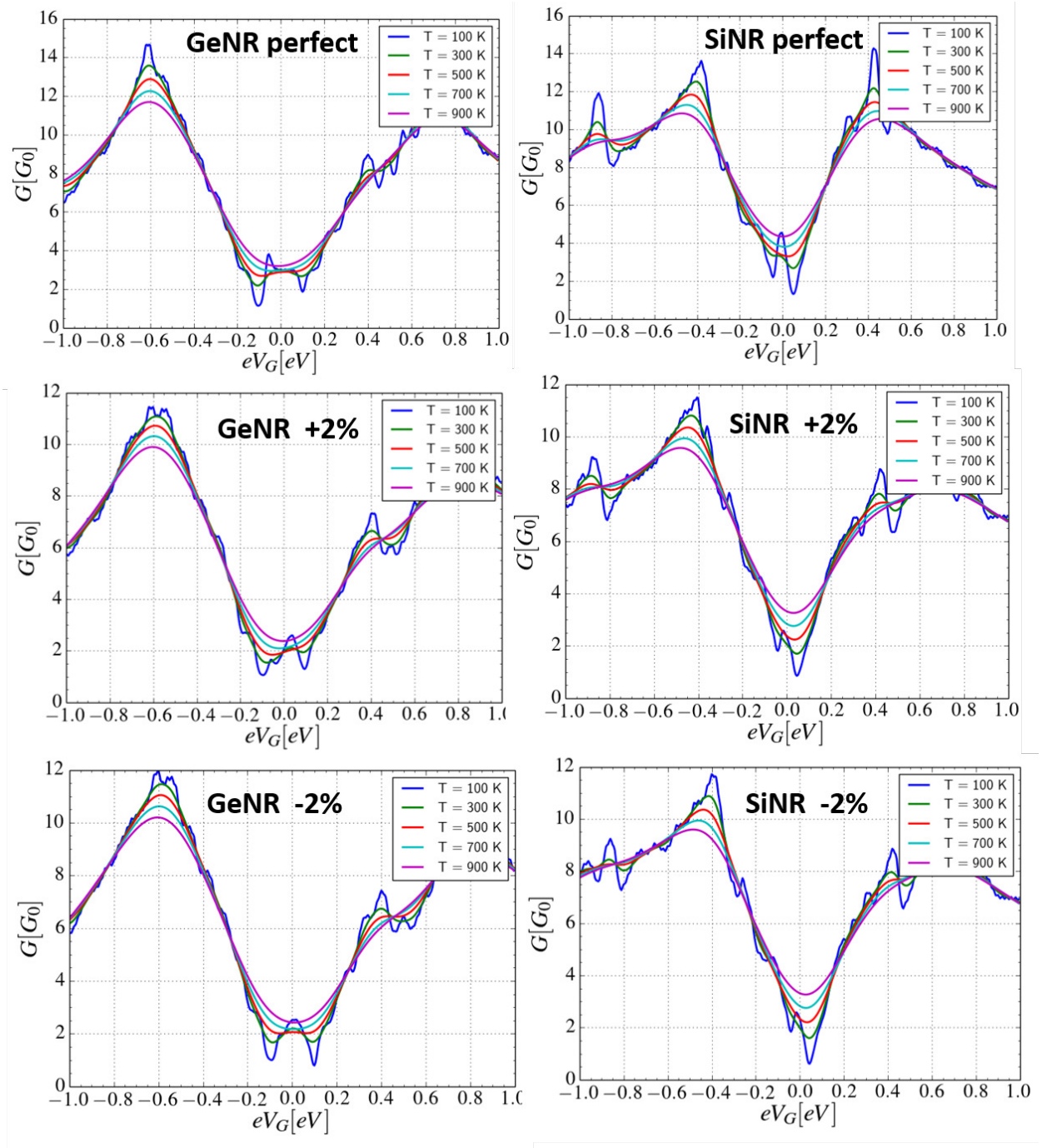

FIG. 3: Variation of conductance (G) of the SiNR and GeNR for different conditions with doping and temperature.

The thermoelectric figure of merit (ZT) plotted in Fig. 6, shows more diverse behavior in terms of change of material and strain than other thermoelectric coefficients presented earlier. The peak values of the figure of merit differ depending on the material, temperature and nature of strain and are usually between 0.12-0.25, which is acceptable considering the single layered one dimensional structures considered in our

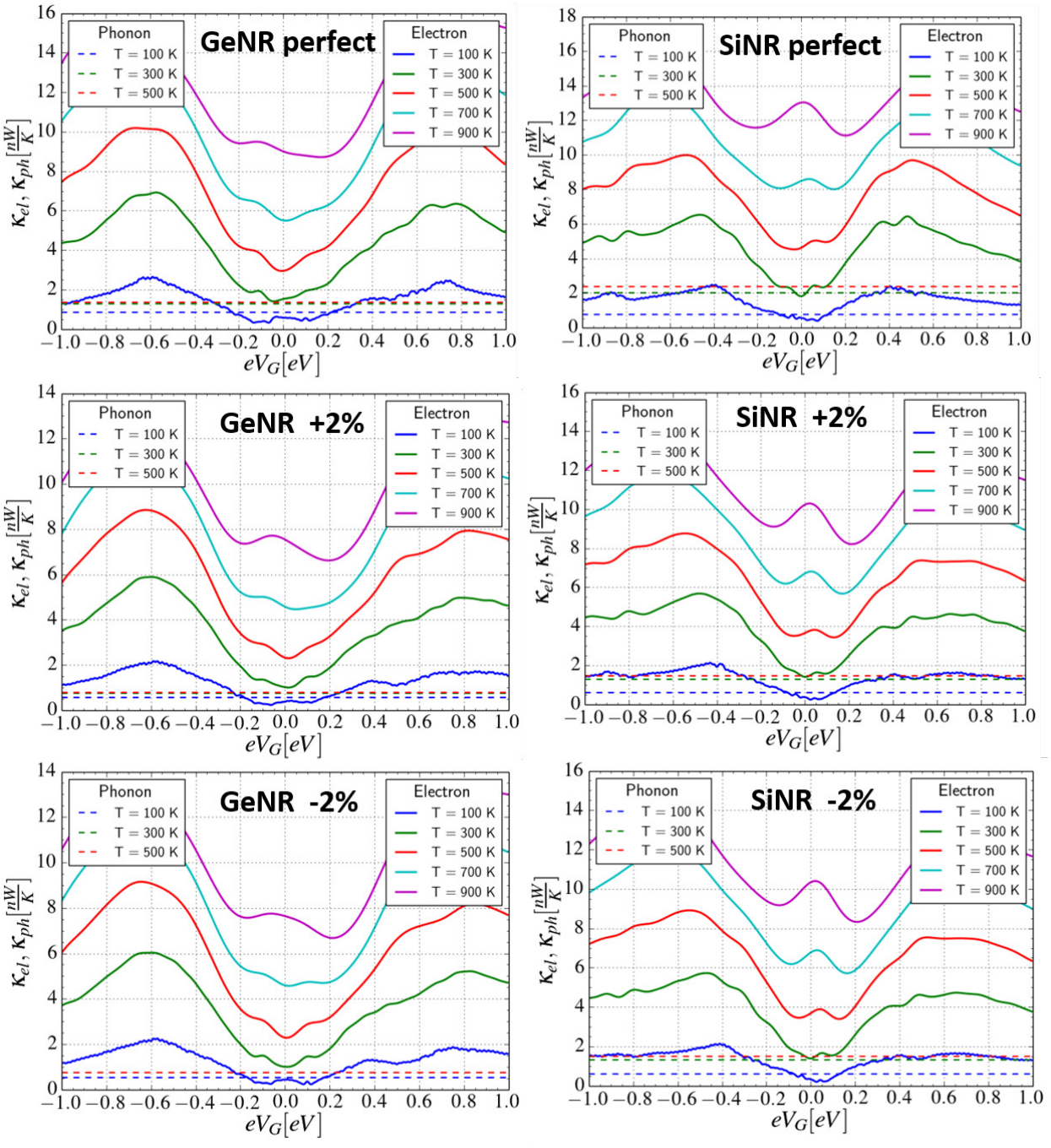

FIG. 4: Variation of thermal conductance ($\kappa_{e,ph}$) of the SiNR and GeNR for different conditions with doping and temperature. Electronic part shown in solid lines, while lattice (phonon) part shown in dotted lines.

studies. It is seen in SiNR that at very low doping levels and low temperature of 100 K, ZT shows values larger than that for temperatures of 300 K and above for similar doping. For SiNR with increase in temperature the ZT plot spreads out more and has a lesser maxima at the same positive doping value, which

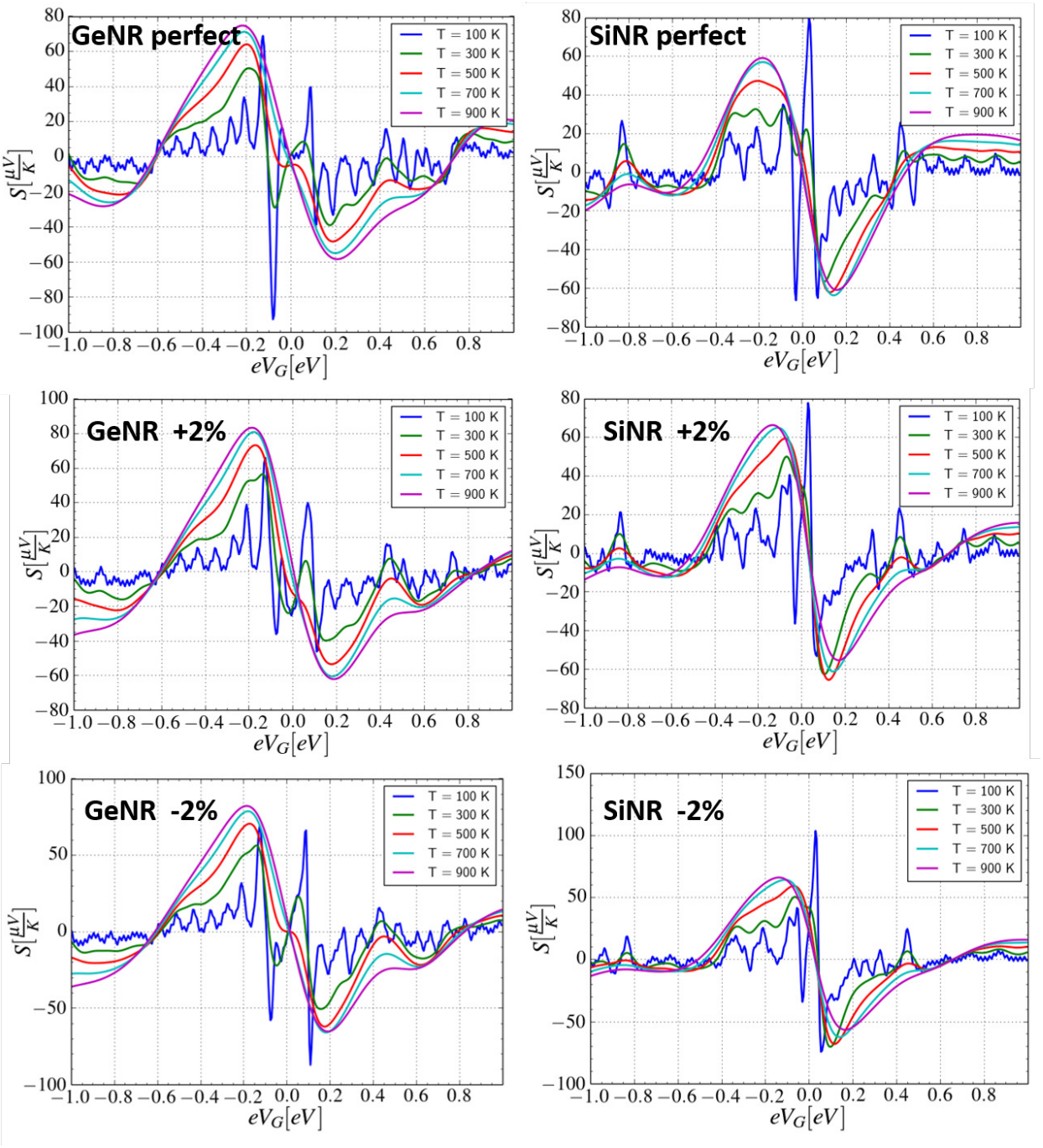

FIG. 5: Variation of Seebeck coefficient (S) of the SiNR and GeNR for different conditions with doping and temperature.

is different from GeNR. With application of strain, the most significant increase in ZT is seen in SiNR, where for temperatures of 300 K and above a $\sim$30-35% increase is seen ZT for minor doping levels with $\pm$2% localized strain. Also this enhancement in SiNR is more for the negative doping than positive doping values. In GeNR at temperatures above 300 K a higher degree of ZT enhancement of about 25% can be

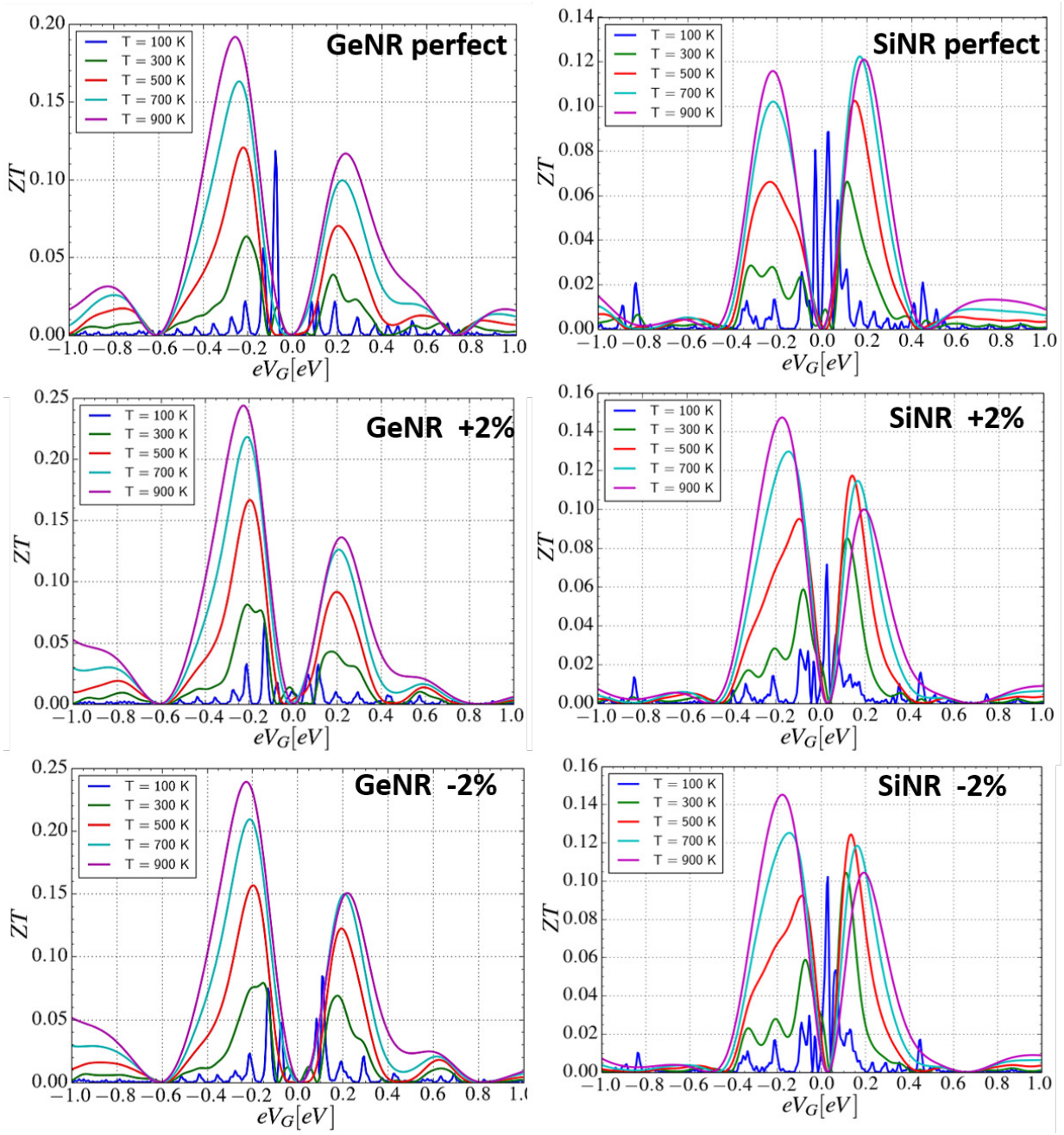

FIG. 6: Variation of the thermoelectric figure of merit (ZT) of the SiNR and GeNR for different conditions with doping and temperature.

seen for $\pm 2\%$ of strain. These results indicate that by applying a moderate strain($\pm 2\%$) on a small section of a 1D SiNR or GeNR, it is possible to significantly tune the thermoelectric behavior of the material up to 35% for SiNR 25% for GeNR. The reversible nature of the induced deformation, makes it a good candidate for tunable thermoelectric devices/ interconnects.

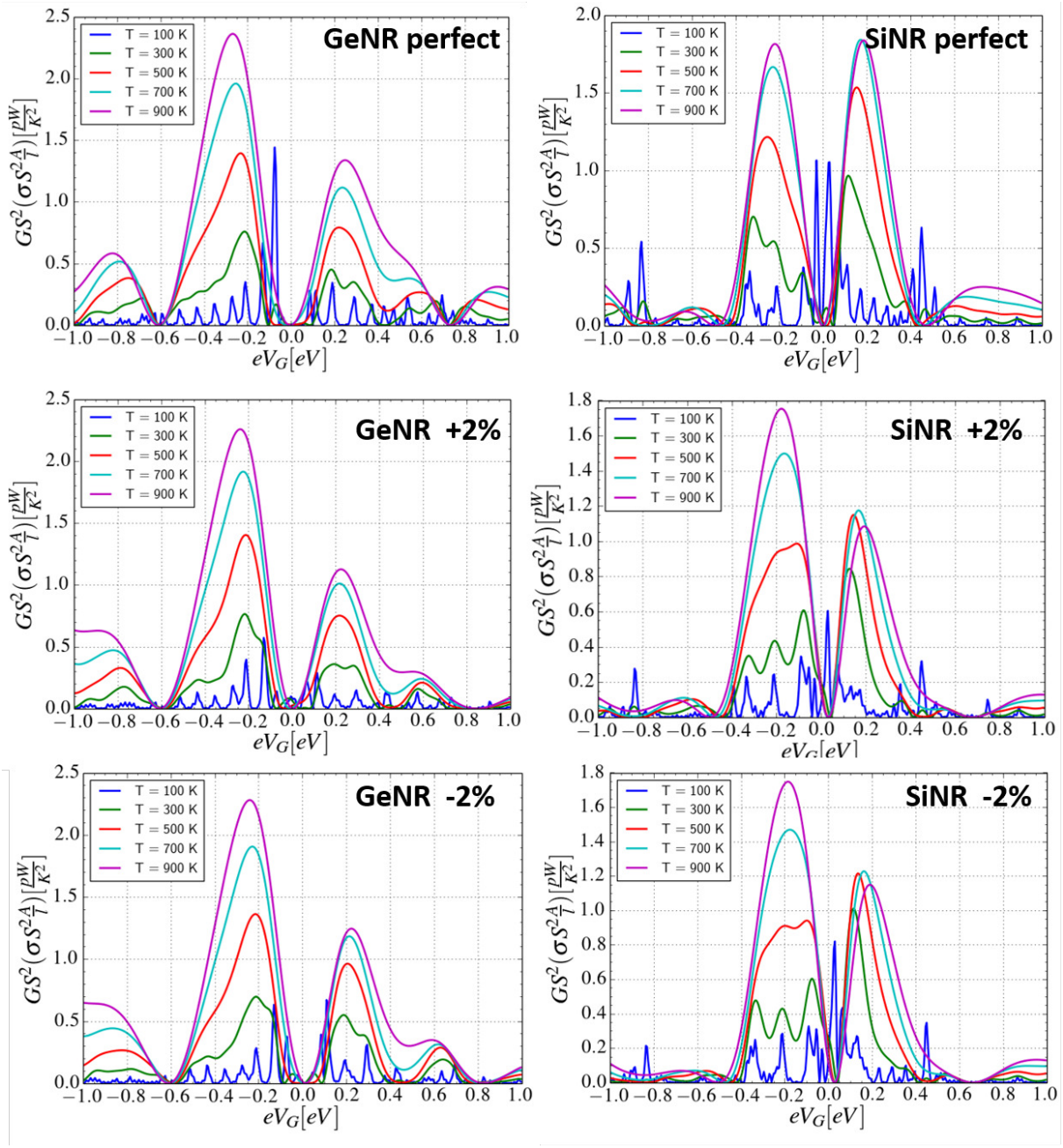

FIG. 7: Variation of the thermoelectric power factor of the SiNR and GeNR for different conditions with doping and temperature.

The thermoelectric power factor(PF), shown in Fig. 7, also confirms tuning with application of strain. The response however is a different for SiNR and GeNR structures. Overall there is a rather minor change in case of the PF for the GeNR with negative doping and a slight decrease in the maxima of the power factor, for the positive doping. The SiNR however shows a marked increase in PF with $\pm 2\%$ strain for the low negative doping region, especially at 300 K. For doping levels of -0.2 eV, the PF rises about 3-4 times

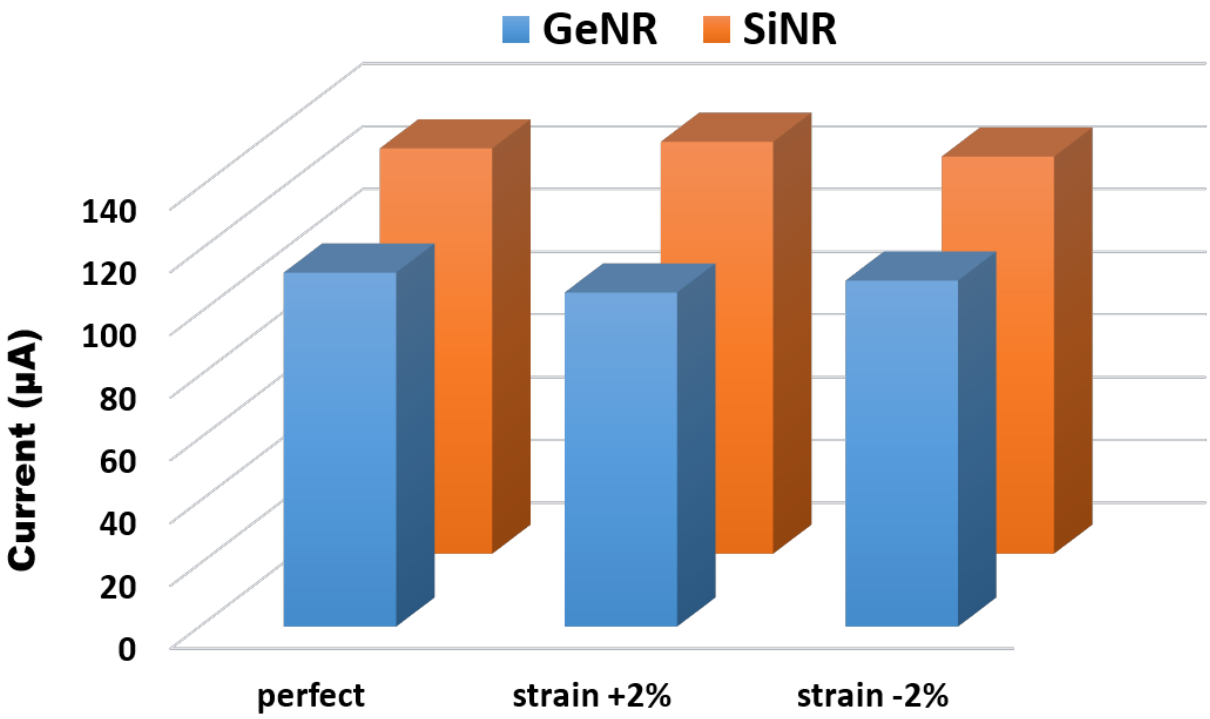

FIG. 8: Variation of the current through the SiNR and GeNR structures at a driving voltage of 1V.

for 300 K operating temperature.

While engineering the thermoelectric properties it is important to reduce thermal conductivity while ensuring minimal effect on the electrical conduction. In order to investigate the effects of strain on the electrical current in the SiNR and GeNR, we carried out non-equilibrium Green function (NEGF) simulations for a moderate bias of 1 V (at 300 K) and evaluated the maximum ballistic current with the two probe Landauer method. The results (Fig. 8) show that the zigzag SiNR and GeNR structures can offer a good driving current in the range of 90-120 μA and the effect of localized strain on the output current is rather small($< 5 - 7\%$). This is quite a positive factor considering the aim in thermoelectric tuning is to have variation in the heat transport properties of a system without much significantly affecting the electronic transport. Since Silicene or Germanene are Dirac materials, through charge pumping or adsorption doping they can be rather easily modulated between n-type or p-type materials [50]. Reports regarding silicene based TEGs have also appeared [51].

IV. CONCLUSION

In this work we explore how a reversible effect as in-plane strain application in a localized manner can be employed to tune the thermoelectric properties of zigzag silicene and germanene nanoribbons. With atomistic simulations we investigate the electron transmission properties and with classical potentials the phonon properties are investigated. The thermoelectric properties are evaluated from the electron and the phonon transmission spectrum calculated under the NEGF formalism. Our studies show a good degree of tuning of electrothermal coefficients are attainable by such strain engineering in SiNR and GeNR systems. A suppression of thermal conductivity while keeping the electronic transport not as much affected was possible to achieve in particular for both the SiNR and the GeNR systems. For GeNR and SiNR devices, 25% and 35%, respectively enhancement is seen in ZT for similar conditions by application of the moderate strain of ±2% in a small portion of the entire NR. These results indicate a good prospect of controlling the electrothermal performance of SiNR and GeNR structures with a simple mechanism and is important from the point of view of nanoscale thermoelectrics in next generation devices and sensors.

ACKNOWLEDGEMENTS

A. S. thanks SERB, India for the SERB Research Scientist Fellowship. A.S. also wishes to thank Dr. Dmitry A. Ryndyk of BCCMS, University of Bremen for valuable comments and discussions.

---

* Electronic address: amretashis@gmail.com

[1] A. Molle, J. Goldberger, M. Houssa, Y. Xu, S.-C. Zhang, and D. Akinwande, Nat. Mater. **16**, 163 (2017).

[2] J. Zhao, H. Liu, Z. Yu, R. Quhe, S. Zhou, Y. Wang, C.C. Liu, H. Zhong, N. Han, J. Lu, Y. Yao, and K. Wu, Prog. Mater. Sci. **83**, 24 (2016).

[3] A. Acun, L. Zhang, P. Bampoulis, M. Farmanbar, A. van Houselt, A.N. Rudenko, M. Lingenfelder, G. Brocks, B. Poelsema, M.I. Katsnelson, and H.J.W. Zandvliet, J. Phys. Condens. Matter **27**, 443002 (2015).

[4] M.E. Davila and G. Le Lay, Sci. Rep. **6**, 20714 (2016).

[5] L. Tao, E. Cinquanta, D. Chiappe, C. Grazianetti, M. Fanciulli, M. Dubey, A. Molle, and D. Akinwande, Nat. Nanotechnol. **10**, 227 (2015).

[6] G. Fiori, F. Bonaccorso, G. Iannaccone, T. Palacios, D. Neumaier, A. Seabaugh, S.K. Banerjee, and L. Colombo, Nat. Nanotechnol. **9**, 768 (2014).

[7] S. Yamacli, J. Nanoparticle Res. **16**, 2576 (2014).

[8] A.H. Bayani, D. Dideban, and N. Moezi, J. Comput. Electron. **15**, 381 (2016).

[9] A.H. Bayani, D. Dideban, M. Vali, and N. Moezi, Semicond. Sci. Technol. **31**, 045009 (2016).

[10] Y. Wang, Z. Ni, Q. Liu, R. Quhe, J. Zheng, and M. Ye. Adv. Funct Mater. **25**, 68 (2015).

[11] S. Wippermann, Y. He, M. Voros, and G. Galli, Appl. Phys. Rev. **3**, 040807 (2016).

[12] M.R. Tchalala, H. Enriquez, A.J. Mayne, A. Kara, G. Dujardin, M.A. Ali, and H. Oughaddou, J. Phys. Conf. Ser. **491**, 12002 (2014).

[13] B. Aufray, A. Kara, S. Vizzini, H. Oughaddou, C. Ĺandri, B. Ealet, and G. Le Lay, Appl. Phys. Lett. **96**, 183102 (2010).

[14] M. Derivaz, D. Dentel, R. Stephan, M.C. Hanf, A. Mehdaoui, P. Sonnet, and C. Pirri, Nano Lett. **15**, 2510 (2015).

[15] P. De Padova, O. Kubo, B. Olivieri, C. Quaresima, T. Nakayama, M. Aono, and G. Le Lay, Nano Lett. **12**, 5500 (2012).

[16] P. De Padova, C. Quaresima, C. Ottaviani, P.M. Sheverdyaeva, P. Moras, C. Carbone, D. Topwal, B. Olivieri, A. Kara, H. Oughaddou, B. Aufray, and G. Le Lay, Appl. Phys. Lett. **96**, 261905 (2010).

[17] M. Houssa, A. Dimoulas, and A. Molle, J. Phys. Condens. Matter **27**, 253002 (2015).

[18] B. Zhou, B. Zhou, Y. Zeng, G. Zhou, and M. Duan, Phys. Lett. A **380**, 282 (2016).

[19] Q. Pang, Y. Zhang, J.-M. Zhang, V. Ji, and K.-W. Xu, Nanoscale **3**, 4330 (2011).

[20] S. Mehdi Aghaei and I. Calizo, J. Appl. Phys. **118**, 104304 (2015).

[21] L. Matthes and F. Bechstedt, Phys. Rev. B **90**, 165431 (2014).

[22] Z. Ni, Q. Liu, K. Tang, J. Zheng, J. Zhou, R. Qin, Z. Gao, D. Yu, and J. Lu, Nano Lett. **12**, 113 (2012).

[23] W. Zhao, Z.X. Guo, Y. Zhang, J.W. Ding, and X.J. Zheng, Solid State Commun. **227**, 1 (2016).

[24] H. Sadeghi, S. Sangtarash, and C.J. Lambert, Sci. Rep. 5, 9514 (2015).

[25] K. Zberecki, M. Wierzbicki, J. Barnaa, and R. Swirkowicz, Phys. Rev. B **88**, 115404 (2013).

[26] K. Yang, S. Cahangirov, A. Cantarero, A. Rubio, and R. D'Agosta, Phys. Rev. B **89**, 125403 (2014).

[27] L. Pan, H.J. Liu, X.J. Tan, H.Y. Lv, J. Shi, X.F. Tang, and G. Zheng, Phys. Chem. Chem. Phys. **14**, 13588 (2012).

[28] M. Hu, X. Zhang, and D. Poulikakos, Phys. Rev. B. **87**, 195417 (2013).

[29] S.M. Aghaei and I. Calizo, Conf. Proc. - IEEE SOUTHEASTCON 2015–June, (2015).

[30] A. Sengupta, Proc. SISPAD 2016, pp. 97-100 (2016).

[31] D. Saha, A. Sengupta, S. Bhattacharya, and S. Mahapatra, J. Comput. Electron. **13**, 862 (2014).

[32] T. Lehmann, D.A. Ryndyk, and G. Cuniberti, Phys. Status Solidi Appl. Mater. Sci. **213**, 591 (2016).

[33] S. Das, Sci. Rep. **6**, 34811 (2016).

[34] R. Roldan, A. Castellanos-Gomez, E. Cappelluti and F. Guinea, J. Physics. Condens. Matter **27**, 313201 (2015).

[35] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kaschner, Phys. Rev. B **51**, 12947 (1995).

[36] G. Seifert, D. Porezag, and T. Frauenheim, Int. J. Quantum Chemistry **58**, 185 (1996).

[37] A. Pecchia and A. Di Carlo, Rep. Prog. Phys. **67**, 1497 (2004).

[38] DFTB+ code available online www.dftb.org.

[39] K. Kaasbjerg, K.S. Thygesen, and K.W. Jacobsen, Phys. Rev. B **85**, 115317 (2012).

[40] F.H. Stillinger and T. A. Weber, Phys. Rev. B, **31**, 5262 (1985).

[41] https://www.tremolo-x.com/.

[42] G. Jha and T. Heine, J. Chem. Theory Comput. **18**, 7, 4472 (2022).

[43] J. Guo, M. Lundstrom, and S. Datta, Appl. Phys. Lett. **80**, 3192 (2002).

[44] S. Datta, Quantum Transport: Atom to Transistor, (Cambridge University Press, NY 2005).

[45] J. Guo, S. Datta, M. Lundstrom, and M. P. Anantram, Int. J. Multiscale Computational Engineering **2**, 257 (2004).

[46] D.A. Ryndyk, Theory of quantum transport at nanoscale, (Springer, Heidelberg, 2016).

[47] P. Vogl, H.P. Hjalmarson, and J.D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).

[48] H. J. Monkhorst, and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[49] T. Markussen, A.-P. Jauho , and M. Brandbyge, Phys. Rev. Lett. **103**, 055502 (2009).

[50] Z. Ni *et al*, Nanoscale **6**, 7609 (2014).

[51] M. M. El Banna, A.H. Phillips, A.S. Bayemi, IEEE Access, **9**, 103564 (2021).

J. Phys. Res. Edu., Vol. 2, March 2024

Recent Efforts Towards Understanding the Early Universe from a Fundamental Quantum Perspective

Sujoy K. Modak*

Facultad de Ciencias - CUICBAS, Universidad de Colima, Colima, C.P. 28045, México and

Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, POB 70-543, 04510, CDMX, México

The observable universe is fundamentally *inhomogeneous* and *anisotropic*. Quantum description of the generation of these inhomogeneities and anisotropies is ill-understood and unsatisfactory. After providing a brief account of the standard approach of the generation of the classical density perturbations starting from the quantum fluctuations of inflaton field, I critically review various assumptions which are crucial for the success of this description, and point out various shortcomings around it. I also discuss the basic ideas and recent works by using an alternative path to overcome those shortcomings which is motivated by the so-called Collapse Model interpretation of quantum mechanics. Inspired by these works, I argue the necessity of constructing a class of manifestly inhomogeneous and anisotropic quantum states after inflation and discuss my recent works which provide one such prescription of building such a state, called the $T-$vacuum, defined in the radiation dominated stage of the early universe.

I. INTRODUCTION

For long our central source of information about the structure and the evolution of the universe were restricted to the observations of redshifts from distant galaxies at the astrophysical scale. In 1965 a nearly isotropic background of microwave radiation was discovered which revolutionized our understanding of the universe due to an access to a large cosmological data. Observations on the distant galaxies provided us an information, even before the observed microwave radiation, that the universe is in an expanding phase. If we have to believe that the expansion of the universe is a feature from the earliest of times, then we could imagine that in the past the matter was hotter and denser than present. In a distant past the temperature was so high that electrons were not bound with atoms and radiation in the form of photons were in collisions with matter and hence the radiation was in thermal equilibrium with matter. The number density of photons in equilibrium with matter at temperature T followed a black-body distribution. As time passed matter become cooler and denser forming nucleus and then atoms but the photons got separated and began a free expansion, but its *spectrum kept the same form*. Usually, it is assumed that there was a time $t = t_L$ where the radiation got suddenly separated from matter, and this L stands for the surface of the "last scattering".

For the expanding universe with scale factor $a(t)$ the frequency ν of freely moving photons at a time $t > t_L$ would have had a frequency $\nu a(t)/a(t_L)$ at the time of decoupling. The effect of cosmic expansion is such that the form of the blackbody spectrum remain unaffected during the free expansion after photons went out of equilibrium with matter just the temperature gets a modification. This modification relates the temperature at any later time after this "photon decoupling" with the temperature at the last scattering, given by $T(t) = T(t_L)a(t_L)/a(t)$. While the instantaneous decoupling of photons is an approximation, it is quite a convincing one as frequent elastic and infrequent inelastic collisions even during the transparent photon era would not modify the blackbody spectrum by much.

George Gamow was one of the first to realize the possible existence such a background spectrum [1] but it was not until 1965 when a detection was made by the radio telescope by Penzias and Wilson [2] which could only detect the radio wavelength of 7.5 cm and a temperature of about $3.5 \pm 1.0\ K$. Subsequently, NASA's COBE and WMAP missions measured this Cosmic Microwave Background (CMB) radiation's average temperature as 2.726 K.

Although this Cosmic Microwave Background (CMB) radiation is highly isotropic there exist several types of relatively small anisotropies and these characteristics of CMB provide some of the most revolutionary information about our universe. The frame of CMB provides a natural frame of reference for studying the universe. Among the various types of anisotropies, that do appear in the CMB temperature distribution over the sky, one is due to the Earth's own motion, and the temperature of CMB varies depending on which direction we look up to. Earth's own motion with respect to the CMB frame depends on the motion of the solar system as well as the rotation of the galaxy itself. According to the WMAP satellite experiment the net velocity of the local group of galaxies relative to the microwave background is 627 ± 22 km/s in a specific direction given by an azimuthal angle $\ell = 276 \pm 3$ degrees and the angle between the galactic plane with the line of sight $b = 30 \pm 3$ degrees [3]. The second type of anisotropy is due to the Sunyaev-Zel'dovich effect [4] which attributes to the scattering of radiation by electrons in the intergalactic space in between the clusters of galaxies along the line of sight. Apart from these secondary anisotropies which are caused by the late universe and appear during the journey of CMB photons to their way to us, there is also primary type of anisotropies that have their origin in the early universe [5]. The primary anisotropies are originated at the Last Scattering Surface (LSS at $z = 1090$) due to various processes, for example, (i) an intrinsic temperature fluctuations in the electron–nucleon–photon at the time of last scattering, (ii) the Doppler effect due to velocity fluctuation in plasma, and (iii) Sachs-Wolfe effect which is redshift/blueshift due to a fluctuating gravitational potential [6].

While it is possible to derive an *empirical formula* relating the anisotropies in CMB with a stochastic fluctuation of a classical field representing the density perturbations, it is of utmost importance to understand

the *fundamental origin* of those empirical formulas of classical, stochastic density perturbations, for which, we need a quantum treatment simply because it is the the large fluctuations of the primordial quantum fields which are responsible for the generation of density perturbations in the first place.

In this article I review the physics behind the origin of the tiny anisotropies at the LSS from a quantum standpoint. The reason behind this is the fact that the in the inflationary universe, and even a short (but unknown) period after inflation, all matter fields were quantum, and the fluctuations of these quantum fields, famously named "inflaton", are the source of these observed angular temperature fluctuations in the CMB. This article is organized as follows: in the next section I discuss the empirical formula that can be constructed using the temperature fluctuations and relating those with the fluctuations of the classical density perturbations. We introduce the famous Harrison-Zel'dovich power spectrum which is what is observed from CMB. Next, in section I provide an elaborate discussion, based on the monograph by Weinberg [3], describing standard quantum description that is most famous in the literature. Here, I will be using a critical stance, strictly reviewing the underlying assumptions, and in the next section I will refute all four popular assumptions, which will lead us to the foundation problems with quantum theory. In section , I shall present new alternative path to understand the quantum description using the so-called Collapse Model interpretations of quantum theory. Finally, in section I will present some new results derived by us in order to understand better quantum fields in radiation dominated universe. This last section is a complementary path to define a set of physically motivated anisotropic quantum vacuum states with an ambition that finally we may be able to point these works down to yet unknown *quantum state* of the early universe at LSS (or before) generating the density perturbations more convincingly.

II. CMB TEMPERATURE FLUCTUATION AND THE POWER SPECTRUM

In the integrated Sachs-Wolfe effect the perturbation to the gravitational potential is a time dependent function $\delta\phi(\mathbf{x})$ where $\mathbf{x}$ is the comoving coordinate. This perturbation slightly changes the frequency of the emitted photon from the point $\mathbf{x}$ on the LSS, and its energy is also slightly shifted. This implies that the temperature obtained by looking in a particular direction $\hat{n}$ is shifted from the averaged value over the whole sky, given by [3]

$$\frac{\Delta T(\hat{n})}{T_0} = \delta\phi(\hat{n}R_{LSS}), \tag{1}$$

where

$$R_{LSS} = \frac{1}{H_0 a(t_0)\sqrt{\Omega_K}} \sinh\left[\sqrt{\Omega_K}\int_{1/(1+z_L)}^{1} \frac{dx}{\sqrt{\Omega_\Lambda x^4 + \Omega_K x^2 + \Omega_M x + \Omega_R}}\right] \tag{2}$$

is the radial coordinates for the LSS, $\Omega_\lambda, \Omega_M, \Omega_R$ are the fractions of energy densities corresponding to vacuum, non-realtivistic and relativistic matters and $\Omega_K = 1 - \Omega_M - \Omega_R - \Omega_K$, $K = 0, \pm 1$ (spatially flat, open/close), $x = a/a_0 = 1/(1+z)$ (z is the redshift factor). The subscript zero belongs to the present value and $z_L = 1090$ is the redshift of the LSS. The perturbation to the gravitational potential also affect the rate of expansion of the universe that also leads to a fluctuation of the temperature given by

$$\frac{\Delta T(\hat{n})}{T_0} = -\frac{2}{3}\delta\phi(\hat{n}R_{LSS}). \tag{3}$$

The total fractional change in temperature is then the sum of Eq. (1) and Eq. (4), and is given by

$$\frac{\Delta T(\hat{n})}{T_0} = \frac{1}{3}\delta\phi(\hat{n}R_{LSS}). \tag{4}$$

which is known as the Sachs-Wolfe effect which is related with the perturbation $\delta\rho$ in the total mass density through the Poisson equation, given by

$$\frac{1}{a^2}\nabla^2\delta\phi(\mathbf{x}) = 4\pi G\delta\rho(\mathbf{x}). \tag{5}$$

This equation, after expressing the Fourier transformation of $\delta\rho(\mathbf{x})$ in terms of $\delta\phi(\mathbf{x})$ yields the correlation function of the density fluctuations

$$\langle\delta\rho(\mathbf{x},t)\delta\rho(\mathbf{x}',t')\rangle = \frac{1}{(4\pi G a(t)a(t'))^2}\int d^3k\ \ \mathcal{P}_\phi(k)e^{i\mathbf{k}.(\mathbf{x}-\mathbf{x}')}, \tag{6}$$

where $\mathcal{P}_\phi(k)$ is the power spectrum which is conventionally expressed as

$$\mathcal{P}_\phi(k) = N_\phi^2 k^n. \tag{7}$$

Observations of the density correlation function has shown that the above expression takes the so-called Harrison-Zel'dovich form [7] with $N_\phi \simeq 10^{-5}$ and $n = 1$.

III. THE STANDARD QUANTUM DESCRIPTION OF THE POWER SPECTRUM

One of the celebrated successes of inflationary cosmological theories lies in the fact that it provides a natural quantum mechanical origin of cosmological fluctuations observed in CMB as described in the last section and in the large scale structure of matter. In this section I shall briefly discuss the standard picture keeping a close attention to various assumptions and hypothesis that are at play. I present here only the relevant part for this article following Weinberg's monograph [3]. My aim is not repeat everything that is already discussed there but to give a much condensed version with a special focus on the basic structure

beneath the complete construction. Let us consider the simplest model of inflation with a single inflaton field $\phi(x)$, given by the action

$$I = \int d^4x\sqrt{-g}\left(\frac{1}{2}g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi)\right), \tag{8}$$

where $V(\varphi)$ is an arbitrary real potential. The first step is to express the scalar field as a sum of unperturbed $\bar{\varphi}(t)$ plus perturbation $\delta\varphi(\mathbf{x},t)$ that depends on both space and time, given by

$$\varphi(\mathbf{x},t) = \bar{\varphi}(t) + \delta\varphi(\mathbf{x},t). \tag{9}$$

Similarly, the metric is given by the unperturbed Friedman-Robertson-Walker metric $\bar{g}_{\mu\nu}(t)$ plus a small perturbation $h_{\mu\nu}(\mathbf{x},t)$,

$$g_{\mu\nu}(\mathbf{x},t) = \bar{g}_{\mu\nu}(t) + h_{\mu\nu}(\mathbf{x},t). \tag{10}$$

The unperturbed Friedman equation (with $K=0$) is

$$H^2 = \frac{8\pi G}{3}\left(\frac{1}{2}\dot{\bar{\varphi}}^2 + V(\bar{\varphi})\right) \tag{11}$$

while the field equation for the unperturbed inflaton field is

$$\ddot{\bar{\varphi}} + 3H\dot{\bar{\varphi}} + V'(\bar{\varphi}) = 0 \tag{12}$$

Now for the perturbations one may consider the Newtonian gauge to select $h_{00} = -2\Psi,\ h_{0i} = 0,\ h_{ij} = -2a^2\delta_{ij}\Psi$. The Einstein equation and energy conservation equations for the perturbations are

$$\dot{\Psi} + H\Psi = 4\pi G\dot{\bar{\varphi}}\delta\varphi \tag{13}$$

$$\delta\ddot{\varphi} + 3H\delta\dot{\varphi} + \partial^2_{\bar{\varphi}}V(\bar{\varphi})\delta\varphi - \frac{1}{a^2}\nabla^2\delta\varphi = -2\Psi\partial_{\bar{\varphi}}V(\bar{\varphi}) + 4\dot{\Psi}\dot{\bar{\varphi}} \tag{14}$$

while the constraint equation can be simplified using the relationship $\dot{H} = -4\pi G\dot{\bar{\varphi}}^2$ and is given by

$$\left(\dot{H} - \frac{1}{a^2}\nabla^2\right)\Psi = 4\pi G\left(-\dot{\bar{\varphi}}\delta\dot{\varphi} + \ddot{\bar{\varphi}}\delta\varphi\right). \tag{15}$$

Once the above set of differential equations Eq. (13)-Eq. (15) describing the perturbations of the metric and the inflaton field are obtained the next step is to find the complete set of field modes which can be used to express the fluctuations as

$$\delta\phi(\mathbf{x},t) = \int d^3k\left[\delta\varphi_k(t)e^{i\mathbf{k}.\mathbf{x}}a_k + \delta\varphi_k^*(t)e^{-i\mathbf{k}.\mathbf{x}}a_k^*\right], \tag{16}$$

$$\Psi(\mathbf{x},t) = \int d^3k\left[\Psi_k(t)e^{i\mathbf{k}.\mathbf{x}}a_k + \Psi_k^*(t)e^{-i\mathbf{k}.\mathbf{x}}a_k^*\right]. \tag{17}$$

At sufficiently early times one has $k/a >> H$ and $k/a >> \partial^2_{\bar{\varphi}} V(\bar{\varphi})$, and the solution of the field equations Eq. (13)-Eq. (15) in their mode form satisfy the initial conditions as $a(t) \to 0$,

$$\lim_{a(t)\to 0} \delta\varphi_k(t) = \frac{1}{(2\pi)^{3/2} a(t)\sqrt{2k}} e^{-ik\int_{t_*}^{t} dt'/a(t')}, \tag{18}$$

$$\lim_{a(t)\to 0} \Psi_k(t) = \frac{4i\pi G\dot{\bar{\varphi}}(t)}{(2\pi)^{3/2} a(t)\sqrt{2k}} e^{-ik\int_{t_*}^{t} dt'/a(t')}, \tag{19}$$

while their complex conjugates are the other set of independent solutions. Also, in the early times the coefficients a_k and a_k^* can be identified with the creation and annihilation operators. Note that, although here the same creation and annihilation operators appear in the scalar field and gravitational perturbations Ψ does not correspond to gravitational radiation, rather it is an auxiliary field and a functional of the inflaton field $\delta\varphi$ in the similar way that in the Coulomb gauge quantization of quantum electrodynamics the time-dependent component of the vector potential is a functional of the charged matter fields. The vacuum expectation values of the paired fields are given by

$$\langle 0|\delta\varphi(\mathbf{x},t)\delta\varphi(\mathbf{y},t)|0\rangle = \int d^3k |\delta\phi_k(t)|^2 e^{i\mathbf{k}.(\mathbf{x}-\mathbf{y})}, \tag{20}$$

$$\langle 0|\Psi(\mathbf{x},t)\Psi(\mathbf{y},t)|0\rangle = \int d^3k |\Psi_k(t)|^2 e^{i\mathbf{k}.(\mathbf{x}-\mathbf{y})}, \tag{21}$$

$$\langle 0|\delta\varphi(\mathbf{x},t)\Psi(\mathbf{y},t)|0\rangle = \int d^3k \delta\phi_k(t)\Psi_k^*(t) e^{i\mathbf{k},(\mathbf{x}-\mathbf{y})}, \tag{22}$$

$$\langle 0|\Psi(\mathbf{x},t)\delta\varphi(\mathbf{y},t)|0\rangle = \int d^3k \Psi_k(t)\delta\phi_k^*(t) e^{i\mathbf{k}.(\mathbf{x}-\mathbf{y})}. \tag{23}$$

As mentioned in the literature [3, 8] that clearly these are quantum averages and not the averages over an ensemble of classical field configurations which is evident from equations Eq. (22) and Eq. (23). These two equations produce complex results for the averages of products of real scalar fields. However, it is critical that they reproduce classical results of density perturbations that will eventually be seen in the CMB map. In order that to happen these averages, although quantum and complex in the beginning of inflation, must lead to classical and real values at the end of inflation for the appropriate field modes which contribute to the primordial inhomogeneities in the classical density and an anisotropy over the CMB sky. This is a challenging task for various reasons one of them is of course to explain satisfactorily how this quantum-to-classical transition might have taken place.

There are three assumptions are made to explain the emergence of classicality:

- Assumption 1: Just as in the measurement of the spin in the laboratory a decoherence will take place making the above-mentioned quantum averages Eq. (20)-Eq. (23) "classical".

- Assumption 2: This quantum-to-classical transition happen when the perturbations exit the Hubble

horizon. These perturbations $\delta\phi_k$ and Ψ_k become classical and are locked into one of the ensembles of classical configurations. Hence can be treated classically.

- Assumption 3: Once the universe become "classical" in the above sense one may use the so-called Ergodic Theorem [3] to interpret averages over ensembles of possible classical universes as averages over the position of the observer in *our universe*.

The first goal of this article is to present some concrete counter-arguments of all the above assumptions that have been made in various recent works and to show that none of the assumptions are satisfactory enough to resolve the issue at hand. In fact, these assumptions are inevitably related with the foundation of quantum theory – the so-called "Measurement Problem" – which till date is unresolved.

But before we delve into such arguments let us first complete the review of the remaining steps leading us to the standard explanation of the Harrison-Zel'dovich power spectrum for the CMB anisotropies.

Since the observational quantities related with the cosmological fluctuations are outside the horizon, after inflation it is not necessary to calculate the full set of equations Eq. (13)-Eq. (15) with the initial conditions Eq. (18) and Eq. (19) for which one also has to select a potential $V(\varphi)$. Rather one focuses on a quantity $\mathcal{R}_k = -\Psi_k + H\delta\varphi/\dot{\bar{\varphi}}$ which is conserved outside the horizon during inflation. The aim is to use this quantity to provide an initial condition for the evolution of perturbations after they re-enter the horizon later but before than the decoupling time. The function $\mathcal{R}(\mathbf{x}, t)$ is related with $\mathcal{R}_k$ much to the same way as perturbations in equations Eq. (16) and Eq. (17) while the quantum average is given by

$$\langle 0|\mathcal{R}(\mathbf{x}, t)\mathcal{R}(\mathbf{y}, t)|0\rangle = \int d^3k\, |R_k(t)|^2\, e^{i\mathbf{k}.(\mathbf{x}-\mathbf{y})}. \tag{24}$$

It is easier to solve $\mathcal{R}_k(t)$ directly than solving for $\delta\varphi_k(t)$ or $\Psi_k(t)$, however, usually a different gauge (than Newtonian) is chosen for a simpler calculation. The vacuum is defined in the early times as $a(t) \to 0$, for the modes satisfying $k/a >> H$ which are essentially same as free massless real scalar field modes in Minkowski spacetime, and it is simply given by $a_k|0\rangle = 0$. The assumption is that the state of the universe during inflation is the vacuum $a_k|0\rangle = 0$ which is known as the Bunch-Davies vacuum [9]. This is the fourth assumption:

- Assumption 4: the state of the universe during inflation is in the Bunch-Davies vacuum $a_k|0\rangle = 0$, in Eq. (20)-Eq. (23) and Eq. (24), which is defined in the early times of inflationary epoch as $a(t) \to 0$.

This Fourier transform of $\mathcal{R}(\mathbf{x}, t)$ satisfy the Mukhanov-Sasaki equation [10]

$$\frac{d^2\mathcal{R}_q}{d\eta^2} + \frac{2}{\eta}\frac{d\mathcal{R}_q}{d\eta} + q^2\mathcal{R}_q = 0, \tag{25}$$

where $\eta = \int \frac{dt}{a(t)}$ is the conformal time. The initial condition, for $a(t) \to 0$, is given by

$$\mathcal{R}_k(t) = -\frac{H(t)}{(2\pi)^{3/2}\sqrt{2k}a(t)\dot{\varphi}} e^{-ik\int \frac{dt'}{a(t')}}. \tag{26}$$

While integrating Eq. (25) one considers the limit from $a = 0$ to beyond the horizon for which $q/a << H$. In this limit it is possible to drop the last term Eq. (25) multiplying q^2 and then there are basically two solutions for $\mathcal{R}_k$. One of those is decaying while the other is a nonzero constant $\mathcal{R}_k = \mathcal{R}_k^0$. This constant value $\mathcal{R}_k^0$ in the super-Horizon limit is used for the analysis of cosmological fluctuations.

The quantity $\mathcal{R}_k^0$ is independent of the nature of inflaton potential in the deep sub-Hubble region while it is constant in the deep super-Hubble scale. However, the potential $V(\varphi)$ plays its role while the perturbations becomes super-Hubble from the sub-Hubble scale as $a(t)$ increases with time – a phenomena usually named as the "Hubble exit" of the perturbations. This feature of the inflationary era, while imprinted on the cosmological perturbations, can be revealed by observing the scalar fluctuations in the later stages of the expanding universe.

This property of Hubble exit in the inflationary universe provide a bound on the maximum number of e-foldings during the inflationary period before the beginning of the next epoch (the radiation dominate era) which, for the slow-roll inflation, is given by

$$\mathcal{N}_0 = \ln\left(\frac{\rho_r^0}{0.037\ h\ \text{eV}}\right), \tag{27}$$

where ρ_r^0 energy density at the beginning of the radiation dominated era. Putting the values of other quantities one can calculate $\mathcal{N}_0 \simeq 68$ which means that we are restricted to explore these 68 e-foldings of inflation by studying the scalar perturbations.

As mentioned above, the particular value of the scalar perturbation $\mathcal{R}_k^0$ does not depend on the particular choice of the potential but on the slow-roll character of the same, and this is helpful to carry out a specific calculation with a potential that obey such a property, for example $V(\varphi) = Ae^{-\lambda\varphi}$ where $A,\ \lambda$ are constants. The final result for $\mathcal{R}_k^0$ in the limit $q/a << H$ can be expressed as [3]

$$\mathcal{R}_k^0 = -i\frac{\sqrt{16\pi G}}{k^{3/2}}\left(\frac{H(t_k)}{8\pi^{3/2}\sqrt{\epsilon}}\right), \tag{28}$$

where t_k is the comoving time for Hubble exit. The power spectral function for this solution is then given by the factor $k^4|\mathcal{R}_k^0|^2$, which is proportional to k, and therefore matches with the Harrison-Zel'dovich spectrum as discussed in the previous section and given by equation Eq. (7). The theoretical restriction on the value of the constant $\epsilon < 1$ and it is defined as $\epsilon = -\frac{\dot{H}}{H^2} = \frac{\lambda^2}{16\pi G}$. Observations, such as WMAP, is able to fix the value of this parameter which is found to be (third year WMAP result) $\epsilon = 0.021 \pm 0.008$ which is in accordance with the theoretical restriction.

This completes our brief review of the standard approach to describe the observed anisotropies based on the cosmological perturbation theory using inflaton field with a general but slowly rolling potential. While it is remarkable that the correct form of the Harrison-Zel'dovich spectrum Eq. (7) is obtained it is also important to review necessary assumptions that are at the core of this approach which we want to perform in the next section.

IV. UNSATISFACTORY ISSUES WITH THE STANDARD APPROACH

In this section I review the four assumptions that are listed in the previous section and discuss if the arguments behind these assumptions are acceptable beyond criticisms and why there is a wide range of disagreement on their acceptability.

- Reviewing assumption 1: in their works Perez, Sahlman and Sudarsky [11, 12] had pointed out several shortcomings that are associated with the standard explanation of the quantum origin of classical density perturbations. One cannot compare the appearance of classical result starting from Eq. (20)-Eq. (23) to a laboratory situation of spin measurement. While in the latter there is an "observer" performing the act of "measurement" by an appropriate use of an "apparatus", *none* of these is true for the situation at hand in the early universe where we are describing the vary fact of structure formation. No physical entity can perform any such act of measurement and therefore assumption 1 described in the previous section cannot be objectively acceptable. As Weinberg mentions in his monograph [3] *"... decoherence cannot occur until expectation value of products of real fields become real..."*, and therefore decoherence do not make the complex quantum averages into real statistical averages by itself. The unitary property of quantum theory does not allow this to happen and we do need something extra such as a spontaneous collapse of the wavefunction which is the topic for the next section.

- Reviewing assumption 2: although it is often argued that when the perturbations exit the Hubble horizon during the inflationary period they suddenly become classical and get locked in that classical state until they reappear inside the Hubble horizon during the radiation dominated era and detected in the LSS. In an important work by Singh, Modak and Padmanabhan [13] a systematic study was performed to review this particular assumption. We used a Schrodinger picture and studied the evolution of real, massive scalar field in a toy universe with three stages of expansions given by the inflationary stage, radiation dominated stage and late time de Sitter stage. Using a notion of the strength of "classicality" as a measure of standard correlation between the field modes and its

momentum via Wigner function, developed in some earlier works [14, 15], we could quantify the level classical behavior for such modes. This model although is not exact to the quantization of both the gravitational and inflaton perturbations at the same time, it is in fact a very good approximation and in line with a semiclassical viewpoint where one only quantize the fields but not the background metric. Our result showed interesting properties – (i) we could show that the degree of classicality is maximum when the field modes exit the Hubble radius during the inflationary period, which is in agreement with the first part of the assumption made in [3], however, we also found (ii) that the second part of the assumption is violated since the degree of classicality oscillates when the same mode *reenters the Hubble scale during the radiation dominated era*. Therefore, they do not get locked in any classical configuration and the hypothesis that those modes are responsible for the classical density perturbations in the LSS is misleading.

- Reviewing assumption 3: the validity of this assumption needs the validity of the first two assumptions which, as we have discussed above, are not acceptable from a more fundamental point of view. We may need new physics in order to establish the first two hypotheses and once this objective is ensured then one may review the "Ergodic Theorem" as described by Weinberg [3].

- Reviewing assumption 4: the assumption that the quantum state of the universe during inflation is given by the Bunch-Davies vacuum has a serious problem which was also pointed out by Perez, Sahlman and Sudarsky [11, 12]. The main objection is the following – one cannot explain the observed anisotropies after inflation starting from quantum averages Eq. (20)-Eq. (23) or Eq. (24) where the quantum state $|0\rangle = |0\rangle_{\rm BD}$, the initial conditions are Eq. (18), Eq. (19) and Eq. (26), while the evolution during the inflationary period is all *unitary*. This is simply because the Bunch-Davies state, by definition is homogeneous and isotropic, and no unitary evolution can break this symmetry and is able to dynamically produce anisotropies which we eventually observed in CMB and associate with the density perturbations on the LSS.

We therefore conclude that none of the assumptions mentioned in the section and used in the literature (as summarized in [3]) *are not* convincing enough and there is a scope of new ideas to eventually reproduce the Harrison-Zel'dovich spectrum in Eq. (7) from other viewpoints.

V. ALTERNATIVE PROPOSALS USING THE OBJECTIVE COLLAPSE MODELS

The quest for satisfactory resolution of the "Measurement Problem" gave birth to several versions of quantum mechanics, and among them are (a) Many World Interpretation (MWI) [16], (b) Bohmian Me-

chanics [17] and (iii) Objective Collapse Models [18–20]. Considerations for addressing the generation of classical density perturbations is a common topic that has been addressed from all of the above three versions except for the MWI, for which this issue the problem does not need any new explanation. However, using the Bohmian approach this problem was addressed to some extent [21, 22]. Finally, there were various attempts and discussions within the scope of collapse model interpretation independently from several groups [23–26]. I shall only discuss the main ideas and developments based on the Collapse Model approaches in this note.

Delicate issues related with the application of quantum theory in the cosmological framework has been stated before [11, 12, 27]. There are several works where authors have studied the emergence of classical density perturbations [28–31] which are essentially identical or very similar to the method outlined in Weinberg's monograph [3], and which suffer from the problems I discussed in the preceding section. The emergence of classicality or dynamically breaking the symmetry of an initial quantum state without measurement is something cannot be addressed within the Copenhagen interpretation where an observer performs the act of "measurement" making the wavefunction collapse and only after this an outcome is achieved. As Roger Penrose had pointed there are two completely distinct processes of evolution in the Copenhagen version of quantum theory – (i) the "unitary" or $U-$process which is unitary and dictates the evolution of a quantum system in isolation, and (ii) the "reduction" or the $R-$process dictates the reduction of quantum superposition to a stochastically chosen outcome, following the Born probability rule, once the system is measured. These two dynamical processes (i.e., unitary vs stochastic) are separated by this vague act of measurement where neither the apparatus, nor the observer or the details of measurement process are included in the theory. Although one may get around these complications in the laboratory by subjecting quantum theory as a prescription to compare the outcome of measurements, a logical extension of this outside laboratory setting is challenging. The absence of any "observer" or "apparatus" in the cosmological framework makes the quantum mechanical "Measurement Problem" [32–34] more explicit than in the laboratory situations [11, 12, 27].

Collapse Models [18–20] provide a mathematically rigorous treatment to unify the $U-$process and R in a single evolution equation and include the effect of apparatus in the mathematical formulation of the quantum theory. The most evolved version of Collapse Models is called Continuous-Spontaneous-Localization (CSL) theory [35], [1]. To understand the CSL theory we need to know following two main equations [35]: first, (i) a modified quantum dynamical evolution, accompanied by the choice of a certain observable $\hat{A}$,

[1] Note that this is an evolved form of the previous discrete versions [36, 37].

which is a (stochastically) modified Schrödinger equation, and whose solution is given by

$$|\psi,t\rangle_w = \hat{\mathcal{T}} e^{-\int_0^t dt' \left[i\hat{H} + \frac{1}{4\lambda}[w(t') - 2\lambda\hat{A}]^2\right]} |\psi,0\rangle, \tag{29}$$

where $\hat{\mathcal{T}}$ is a time-ordering operator and $w(t)$ is a random, white noise type classical function of time. The second equation provides the probability of $w(t)$ via (ii) the Probability Distribution Density $[P(Dw(t))]$ function:

$$P(Dw(t)) \equiv {}_w\langle\psi,t|\psi,t\rangle_w \prod_{t_i=0}^{t} \frac{dw(t_i)}{\sqrt{2\pi\lambda/dt}}. \tag{30}$$

In this way the standard Schrödinger evolution and the changes in the state corresponding to a "measurement" of the observable $\hat{A}$ are unified. For non-relativistic quantum mechanics of a single particle, in all situations (without invoking any measurement device or observer), this theory assumes a spontaneous and continuous reduction characterized by $\hat{A} = \hat{\vec{X}}_\delta$, where $\hat{\vec{X}}_\delta$ is a suitably smeared position operator. This can be generalized to multi-particle systems and thereby everything, including, the apparatuses can treated quantum mechanically. The theory seems to successfully address the "measurement problem". In order to match observational evidences at the quantitative level, the collapse parameter λ must be small enough not to be in conflict with known tests of QM in subatomic physics range. However, it also ensures rapid localization of the "macroscopic objects" and ensuring no superpositions in them. Recently we made a bold claim that Collapse Models can precisely identify the classical-quantum boundary [38]. The originally suggested value for the collapse parameter was $\lambda \sim 10^{-16} sec^{-1}$ but over the decades the CSL parameter space has being tested by several experiments and for the current status of the theory, and its empirical constraints, we refer the reader to consult some review articles [18–20].

In order to address the problems associated in standard description of the quantum generation of classical density perturbations, [23] Cañate, Pearle and Sudarsky (CPS) used the semiclassical gravity approach, together with collapse of the state vector according to the CSL dynamics, where the semiclassical Einstein equation is given by

$$G_{\mu\nu} = 8\pi\langle\Psi|T_{\mu\nu}|\Psi\rangle. \tag{31}$$

In this approach, the spacetime is treated classically but all matter fields get full quantum treatment. This equation when supplemented by the necessary collapse of the state vector [23] which makes the semiclassical treatment well-defined [2]. It is in fact very appealing that Collapse Models not only useful to address the problem at hand but also cure a fundamental inconsistency in semiclassical gravity at the same time. One

[2] Collapse mechanism is one of the few ways to bypass the criticism on semi-classical gravity by Page and Geiker [39].

major difference of this approach with the standard one described in section is that here *the perturbation of the background metric does not get a quantum treatment*, just the perturbation of the matter field is quantized. In their calculations CPS chose both the field modes $\hat{\delta\phi}_k$ and corresponding momentum conjugate $\hat{\pi}_k$ as collapse generating operators with a choice of collapse parameter not just a constant but a function depending on the momentum of the field modes, chosen to be $\tilde{\lambda} = \lambda/k$ and $\tilde{\lambda} = \lambda k$, respectively. With these choices CPS made a detailed calculation, based on a CSL evolution of the state vector, in which the initial state is the Bunch-Davies vacuum defined in the beginning of inflation, to a stochastically chosen particle excited state with some momentum k, as a final state after inflation, and reproduced the necessary form of the Harrison-Zel'dovich spectrum [Eq. (7)].

In [24] Martin, Vennin and Peter (MVP) made another detailed study of generation of primordial perturbations using Collapse Models. They also used the CSL theory, however, chose the collapse to take place on the eigenstate of the Mukhanov-Sasaki operator $\hat{\mathcal{R}}_k$ rather than the field operator or the field momentum operator. Note that Mukhanov-Sasaki operator is a combination of both the perturbation of the inflaton and perturbation of the background metric. Therefore, strictly speaking MVP approach is not strictly semiclassical where background metric is never quantized, and in this sense this method is different than the method used by CPS in [23]. In their study MVP found two branches of solutions in their analysis, one which reproduces the scale invariant Harrison-Zel'dovich power spectrum given in Eq. (7) and the other which does not. The requirement that the non-scale-invariant part is outside the horizon puts some bounds on the CSL parameters which control the deviation from standard quantum prediction and MVP concluded that in the absence of any amplification mechanism the standard CSL mechanism is not strong enough to reproduce the known power spectrum in expected time frame.

Soon after the above work by MVP, Das *et. al.* revisited the above study [25] and concluded that it is in fact possible to reproduce (a) the observable power spectrum of the superhorizon modes from the appropriate branch, while (b) keeping the non scale invariant branch outside of the horizon, and (c) achieving the above two within the number of e-folds allowed by inflation (unlike MVP in [24]). These are possible by modifying/amplifying the effect of collapse by setting the CSL parameter depending on the field momentum of the Mukhanov-Sasaki operator. This way of amplification of collapse mechanism is in line with what CPS used in their work [23], as well as we used for addressing the information loss during black hole evaporation process [40–44] which is another scenario where Collapse Models are very useful.

While in the above works authors always considered the Bunch-Davies vacuum state as the initial state Bengochea and León performed an interesting study [26] where they considered a more general type of state, known as the Hadamard state which in the semiclassical gravity theories provide finite values for the renormalized energy momentum tensors free from pathological behaviours such as infinite divergences.

They could reproduce the observer angular power spectrum with this new choice of quantum state which puts the overall idea of using the Collapse Models on a firm ground.

More recently Martin and Vennin [45] reported that the CMB constraints on the CSL parameter space, within their model, may invalidate the CSL theory altogether. The conclusions of their study were challenged in [46, 47] with a further reply from the authors [48]. In addition, there are more criticisms on this approach of using collapse models in the context of generation of CMB anisotropies by Kiefer and Vardanyan [49] and by Ashtekar, Corichi, and Kesavan [50] which, in turn, were also got refuted by Berjon, Okon and Sudarsky [51].

Recently, Lechuga and Sudarsky [52] made an exciting proposal resolving the *e*ternal inflation problem which is basically a situation where the inflationary expansion of the early universe cannot be terminated just by assuming the slow-roll condition [53]. The authors show in their work that there exist a valid parameter space in the CSL based proposals which can accommodate (a) the observed anisotropies in the CMB, and (b) avoiding the necessity of eternal inflation.

It is worth mentioning that in the standard accounts of inflation, one obtains predictions for the spectrum of primordial gravity waves (i.e. tensor modes) that are very similar to those for the density perturbations, however the former have yet to be observed. In contrast, the approach described here, predicts primordial gravity waves at a substantially suppressed amplitude [54].

All of the above works, debates, criticisms and counter criticisms make the utilization of the Collapse Models in the cosmological framework a fascinating field of study with a potential of contributing not only to cosmology but also to the foundations of the quantum theory.

VI. BUILDING AN ANISOTROPIC QUANTUM STATE FOR THE EARLY UNIVERSE

One novel feature of quantum field theory (QFT) in a curved background is the possible existence of multiple non-unitarily related quantum vacuum states that are natural state of sets of observers who are non-inertially connected to each other. Even in the Minkowski spacetime if one is to construct a QFT in a non-inertial frame, such as for an observer with constant four momentum, we find ourselves in a similar situation where the standard vacuum in Minkowski (inertial frame) is envisioned as a particle excited state producing particles with blackbody spectrum, which is detected in the accelerating frame. This phenomena is known as Unruh effect [55] and it serves as a cornerstone to understand QFT in a curved space where the background spacetime is other than Minkowski. Other famous examples of particle creation, especially in curved background, are given by the Hawking effect [56, 57] where particles are produced and black holes evaporate itself and in the cosmological models [58–60].

In this section I will give a brief account of my recent works [61–65] which is a complementary line of thought to understand the early universe from a quantum perspective. We believe these studies, in future, will be able to integrate with the efforts mentioned in the last section and may provide a wholesome picture of the generation of classical density perturbations that we observe in the CMB.

The main idea is the following: if one takes semiclassical picture to generate inhomogeneity and anisotropies in the spacetime metric appearing on the left hand side of the semiclassical Einstein equation given in Eq. (31), on the right hand side one must have a quantum state $|\Psi\rangle$ which manifestly breaks the homogeneity and isotropy symmetries. The usual Bunch-Davies vacuum state being invariant under the de Sitter symmetry group it cannot, by simple argument of isometry, be the state $|\Psi\rangle$ at the time of generation of these tiny deviations. One must find another state which is manifestly in-homogeneous (in-H) and an-isotropic (an-I). Of course, there will be numerous conditions applicable for such a state to be well defined. Foremost, it must be a Hadamard state so that the right hand side of Eq. (31) is free of divergences after one renormalize the stress energy tensor. Further, the resultant renormalized value of the stress tensor must satisfy the Wald axioms [66]. One may come up with several examples of such a quantum state and we can only hope that ideally one or at most a class of them could provide correct renormalized values which, via back-reaction on the metric, will generate observed CMB inhomogeneities anisotropies. The main problem, as it stands today, is the fact that there is no such in-H and an-I state was ever built in quantum cosmology which satisfies all of the restrictions required for the well-definedness of semiclassical gravity. We find this to be a glaring omission and have made considerable progress in building one such quantum state [61, 63] and discussed various new results in this respect. Below we provide a quick review on the prescription of building the $T-$vacuum state, as termed in [63], which is an in-H and an-I quantum state and lives in the radiation dominated era of the early universe and therefore it is free from inflationary effect which tends to erase any primordial inhomogeneity and anisotropy.

We shall divide our findings in two separate branches – first (i) the geometric part, where we express the H & I form of the radiation dominated metric into a spherically symmetric, in-H & an-I form, and related geometric results, and second (ii) the field theoretic part, where we discuss the formulation of the $T-$vacuum and related field theoretic results.

A. Novel geometric features of radiation dominated Epoch

Consider the spatially flat FRW metric in comoving coordinates,

$$ds^2 = dt^2 - a^2(t)[dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)], \tag{32}$$

which in cosmological time frame is

$$ds^2 = a^2(\eta)[d\eta^2 - dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)]. \tag{33}$$

where the "cosmological time" $\eta = \int \frac{dt}{a(t)}$. Here the scale factor $a(t)$ has a dimension of length and η is dimensionless, while t has dimension of time. The only thing carry dimension [Eq. (33)] is the conformal factor.

Using the lightcone coordinates $u = \eta - r$ and $v = \eta + r$, Eq. (33) becomes

$$ds^2 = a^2(u,v)\Big[dudv - \frac{(v-u)^2}{4}(d\theta^2 + \sin^2\theta d\phi^2)\Big]. \tag{34}$$

At this point make the following conformal transformation for $a(t) \propto t^{1/2}$, given by [61]

$$U \equiv T - R = \pm\frac{\mathcal{H}e}{2}u^2, \quad V \equiv T + R = \frac{\mathcal{H}e}{2}v^2, \tag{35}$$

where $+$ and $-$ mean $u \geq 0$ and $u \leq 0$, respectively. It is straightforward to show that the above transformation implies that the radiation dominated epoch is given by

$$ds^2 = F_I(T,R)(dT^2 - dR^2) - R^2 d\Omega^2 \tag{36}$$

for $U \geq 0$ ($T \geq R$), and

$$ds^2 = F_{II}(T,R)(dT^2 - dR^2) - T^2 d\Omega^2, \tag{37}$$

for $U \leq 0$ ($T \leq R$) and the functions F_I and F_{II} are $F_I(T,R) = \frac{(\sqrt{T+R}+\sqrt{T-R})^2}{4\sqrt{T^2-R^2}}$ and $F_{II}(T,R) = \frac{(\sqrt{R+T}-\sqrt{R-T})^2}{4\sqrt{R^2-T^2}}$. These new coordinates in (T,R) and (η,r) frames are related by $T = \frac{V+U}{2} = \frac{\mathcal{H}e}{2}(\eta^2 + r^2)$; $R = \frac{V-U}{2} = \mathcal{H}e\eta r$ for region I, and $T = \frac{V+U}{2} = \mathcal{H}e\eta r$; $R = \frac{V-U}{2} = \frac{\mathcal{H}e}{2}(\eta^2 + r^2)$ for region II. The conformal factors $F_I(T,R)$ and $F_{II}(T,R)$ as functions of the Hubble parameter for radiation stage $H = (\frac{\dot{a}}{a})_{RD}$ in the following form [61] $F_I(H,R) = \frac{1}{1-H^2R^2}$ and $F_{II}(H,T) = \frac{1}{H^2T^2-1}$. The line $T = R$ for these observers is the comoving Hubble radius at $R = 1/H$. These metrics in Eq. (36) and Eq. (37) are static up to a dynamical conformal factor and they exhibit a spherical symmetry. It was named as the (T,R) frame [63].

There are few important results derived in the earlier works. For example, the (T,R) frame was backtracked in the inflationary de Sitter epoch and was identified with the de Sitter static patch with a time redefinition [64]. A complete spacetime foliation with Cauchy slices was performed ensuring the well defined initial value problem [63] etc.

B. Novel field theoretic features of radiation dominated epoch

The motivation of constructing a field theory in the (T, R) frame is the following: first we notice that the metrics in the (T, R) frame as appear in Eq. (36) and Eq. (37) are inhomogeneous and anisotropic. The spherically symmetry ensures that we should be able to separate the field equation (such as the KG equation) into the radial and angular parts. The radial part although will be dependent on coordinates T, R, the angular part will provide us some known solution. Given that we are able to decompose the field operator in the positive and negative frequency parts we will be able to define a new vacuum state which will be annihilated by the new annihilation operator multiplying the field modes. This new vacuum state will then also be inhomogeneous and anisotropic and if we can prove the well-definedness of this state we shall have one physically motivated, rigorously constructed anisotropic quantum state in the radiation dominated universe.

In our recent works [63, 65] we showed that it is possible to quantize a free massless scalar field which can be related with the fluctuation of the inflaton field and with no quantization of the background metric fluctuations as necessary in a semiclassical set up. In this process we defined a manifestly in-H and an-I quantum vacuum which we called as the $T-$vacuum [63]. It was shown that the $T-$vacuum creates particles in the frame of cosmological and comoving observers and the distribution of particles is anisotropic in 3+1 dimensions [65]. This is an exciting result since we could see the usefulness of engineering a quantum state that breaks the H & I symmetries. Further, we showed the Hadamard behavior in (1+1) dimensions and a point by point comparison with the Unruh effect [63]. These are definitely encouraging but there remain many more unanswered questions which we want to study in future.

We are now in a position to ask the following questions, (a) could the $T-$vacuum be the yet unknown in-H and an-I state that the universe finds itself after the inflation?, (b) could we reproduce the scale invariant Harrison-Zel'dovich spectrum (7) by using this state?, (c) could we connect it to the studies using Collapse Models as the final state after collapse?, and finally (d) even if these ambitions are refuted could we adjust and build new states in order to match the observation in CMB? To summarize, what we are essentially asking through all these questions is: *what is the precise quantum state of the early universe at the time of the production of density perturbations at the LSS?*

VII. SUMMARY AND CONCLUSION

To summarize, this article provides an up to date information about some efforts related with betting understanding the quantum description of the classical density perturbations from inflationary cosmology

which are famously called the "seeds of cosmic structures" that exist today. We provided some introductory information on the detection of anisotropies in the CMB in the introduction and then discuss the observational quantity given by the Harrison-Zel'dovich power spectrum that is one of the results needs to be reproduced by any valid theory describing the problem at the hand. In this respect, we provide a brief but self contained explanation of the standard inflationary approach of using cosmological perturbations. We critically reviewed various assumption behind the calculations and discuss some of the flaws in details. Then we reviewed strong arguments present in the literature to connect the problem of quantum generation of classical density fluctuations with the foundational problems in quantum theory which cannot be bypassed in the cosmological setting. To this aspect, we reviewed recent works within the Collapse Model interpretations of quantum theory which is one of the most promising paths to overcome various flaws with the standard picture. Finally, we provided brief account of our work on building an anisotropic and physical quantum state in the post-inflationary universe, especially in the radiation dominated era, in order to look for the precise quantum state (if any) of the early universe at the time of the production of density perturbations at the LSS.

ACKNOWLEDGEMENTS

I thank Daniel Sudarsky and Rosa Laura Lechuga for discussions and Dhamar Astilla for reading the manuscript and correcting typos. This research is supported by CONAHCyT Grant CF/140630, Mexico. I acknowledge the Sabbatical Fellowship Program by CONAHCyT for financial support. I also acknowledge ICN-UNAM for hosting my sabbatical stay and providing necessary support for my research.

* Electronic address: smodak@ucol.mx

[1] G. Gamow, Phys. Rev. **70**, 572 (1946).

[2] A. A. Penzias and R. W. Wilson, Astrophys. J. **142**, 419 (1965).

[3] Steven Weinberg, "Cosmology," Oxford University Press (2008).

[4] Ya. B. Zel'dovich, R. A. Sunyaev, Astrophysics Space Science **4**, 301 (1969); Comments Astrophysics and Space Science **2**, 66 (1970); **4**, 173 (1972).

[5] E. D. Reese *e*t al., Astrophysics Journal **533**, 38 (2000).

[6] R. K.Sachs and A. M. Wolfe, Astrophysics Journal, **147**, 73 (1967).

[7] E. R. Harrison, Phys. Rev, **D1**, 2726 (1970); Ya. B. Zel'dovich, MNRAS **160** 1P (1972).

[8] D. H. Lyth and D. Seery, arXiv: astro-ph/0607647

[9] T. S. Bunch and P.C.W. Davies, Proc. Roy. Society Ser. A **360**, 117 (1978).

[10] V. S. Mukhanov, JEPT Lett. **41**, 493 (1986), S. Sasaki, Prog. Theor. Phys. **76**, 1036 (1986).

[11] A. Perez, H. Sahlmann, and D. Sudarsky, Class. Quant. Grav. **23**, 2317 (2006).

[12] D. Sudarsky, Int. J. Mod. Phys. D, **20**, 509 (2011).

[13] S. Singh, S. K. Modak, and T. Padmanabhan, Phys. Rev. D **88**, 125020 (2013).

[14] G. Mahajan and T. Padmanabhan, Gen. Rel. Grav. **40**, 661 (2008).

[15] G. Mahajan and T. Padmanabhan, Gen. Rel. Grav. **40**, 709 (2008).

[16] J. Hartle and T. Hertog, Phys. Rev. D, **95**, 123502 (2017); E. Okon and D. Sudarsky, Found. Phys. **44**, 19 (2014).

[17] R. Tumulka, "Bohmian Mechanics," Chapter 17: The Routledge Companion to Philosophy of Physics (2021).

[18] A. Bassi and G. Ghirardi, Physics Reports, 379, 257, (2003).

[19] A. Bassi, M. Dorato, and H. Ulbricht, Entropy **25**, 645 (2023).

[20] A. Bassi, K. Lochan, S.Satin, T. P. Singh, and H. Ulbricht, Rev. Mod. Phys. **85**, 471 (2013).

[21] S. Goldstein, W. Struyve, and R. Tumulka, arXiv:1508.01017.

[22] N. Pinto-Neto and W. Struyve, arXiv:1801.03353.

[23] P. Cañate, P. Pearle, and D. Sudarsky, Phys. Rev. D **87**, 104024 (2013).

[24] J. Martin, V. Vennin, and P. Peter, Phys. Rev. D **86**, 103524 (2012).

[25] S. Das, K. Lochan, S. Sahu, and T. P. Singh, Phys. Rev. D **88**, 085020 (2013). [erratum: Phys. Rev. D **89**, 109902 (2014).]

[26] G. R. Bengochea and G. León, Phys. Lett. B **774**, 338 (2017).

[27] J. Bell and A. Aspect, "Quantum mechanics for cosmologists," *In Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy,* pages 117–138. Cambridge University Press, (2004).

[28] A. Albrecht, P. Ferreira, M.Joyce, and T. Prokopec, Phys. Rev. D **50**, 4807 (1994).

[29] D. Polarski and A. A. Starobinsky, Classical and Quantum Gravity **13**, 377 (1996).

[30] F. C. Lombardo and D. Lopez Nacir, Phys. Rev. D **72**, 063506, (2005).

[31] C. Kiefer, D. Polarski, Adv. Sci. Lett.**2**, 164 (2009).

[32] J. Bell, Phys. World **3**, 33 (1990).

[33] J. S. Bell, "Speakable and unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy," Cambridge University Press, (1987).

[34] T. Maudlin, Topoi, **14**, 7 (1995).

[35] G. C. Ghirardi, P. Pearle, and A. Rimini, Phys. Rev. A **42**, 78 (1990).

[36] P. M. Pearle, Phys. Rev. D **13**, 857 (1976).

[37] G. C. Ghirardi, A. Rimini and T. Weber, Phys. Rev. D **34**, 470 (1986).

[38] F. Torres, S. K. Modak, and A. Aranda, Found. Phys. **53**, 73 (2023).

[39] D. N. Page and C. D. Geilker, Phys. Rev. Lett. **47**, 979 (1981).

[40] S. K. Modak, L. Ortíz, I. Peña, and D. Sudarsky, Gen. Rel. Grav. **47**, 120 (2015).

[41] S. K. Modak, L. Ortíz, I. Peña, and D. Sudarsky, Phys. Rev. D **91**, 124009 (2015).

[42] D. Bedingham, S. K. Modak, and D. Sudarsky, Phys. Rev. D **94**, 045009 (2016).

[43] S. K. Modak and D. Sudarsky, Fundam. Theor. Phys. **187**, 303 (2017).

[44] S. K. Modak and D. Sudarsky, Eur. Phys. J. C **78**, 556 (2018).

[45] J. Martin and V. Vennin, Phys. Rev. Lett. **124**, 080402 (2020).

[46] G. R. Bengochea, G. Leon, P. Pearle, and D. Sudarsky, arXiv:2006.05313.

[47] G. R. Bengochea, G. León, E. Okon, and D. Sudarsky, Eur. Phys. J. C **80**, 18 (2020).

[48] J. Martin and V. Vennin, Eur. Phys. J. C **81**, 64 (2021).

[49] C. Kiefer and T. Vardanyan, Gen. Rel. Grav., **54**, 30 (2022).

[50] A. Ashtekar, A. Corichi, and A. Kesavan, Phys. Rev. D, **102**, 023512 (2020).

[51] J. Berjon, E. Okon, and D. Sudarsky, Phys. Rev. D, **103**, 043521 (2021).

[52] R. L. Lechuga and D. Sudarsky, JCAP **01**, 038 (2024).

[53] A. Linde, "Eternal inflation," in "Inflationary cosmology", Springer (2010).

[54] G. Leon, L. Kraiselburd, and S. J. Landau, Phys. Rev. D **92**, 083516 (2015); T. Markkanen, S. Rasanen, and P. Wahlman, Phys. Rev. D **91**, 084064 (2015); G. León, A. Majhi, E. Okón, and D. Sudarsky, Phys. Rev. D **96**, 101301(R) (2017); G. Leon, A. Majhi, E. Okón, and D. Sudarsky, Phys. Rev. D **98**, 023512 (2018).

[55] W. G. Unruh, Phys. Rev. D **14**, 870 (1976).

[56] S. W. Hawking, Nature **248**, 30 (1974).

[57] S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).

[58] L. Parker, Phys. Rev. Lett. **21**, 562 (1968).

[59] L. Parker, Phys. Rev. **183**, 1057 (1969).

[60] L. Parker, Phys. Rev. D **3**, 346 (1971); Erratum: [Phys. Rev. D **3**, 2546 (1971)].

[61] S. K. Modak, Phys. Rev. D **97**, 105016 (2018).

[62] S. K. Modak, Int. J. Mod. Phys. D **28**, 1930015 (2019).

[63] S. K. Modak, JHEP **12**, 031 (2020).

[64] J. R. Salazar and S. K. Modak, JHEP **05**, 048 (2022).

[65] D. S. Astilla, S. K. Modak and E. Salazar, arXiv:2312.17129.

[66] R. Wald, "General Relativity", University of Chicago Press (1984).

J. Phys. Res. Edu., Vol. 2, March 2024

A short course on the Su-Schrieffer-Heeger model

Tarun Kanti Ghosh*

Department of Physics, Indian Institute of Technology-Kanpur, Kanpur-208 016, India

We review various topological properties of a dimer Su-Schrieffer-Heeger tight-binding model. Exact analytical expressions of the energy spectrum and corresponding eigenstates for any choice of system parameters are provided. We discuss the system's parity, time-reversal, and chiral symmetries. The system undergoes a topological phase transition while tuning the hopping parameters. The topological phase is associated with the presence of the boundary modes and establishes the bulk-boundary correspondence.

I. INTRODUCTION

One of the successful applications of early quantum mechanics is the band theory of crystallne materials [1]. The superposition of the Bloch states leads to the formation of well separated energy bands. The crystalline materials can be classified into two categories, namely the insulator and the metal, depending on whether the bands are completely or partially filled. Completely filled bands are well separated from the completely empty bands in a band insulator. The energy gap prohibits the flow of the charge carriers in the band insulator. On the other hand, materials having partially filled bands are conductors, where charge carriers can easily flow at the cost of very low energy.

After the phenomenal discovery of the integer quantum Hall effect [2] in early 1980s, the simple concept of the band insulator has changed. In the integer quantum Hall effect, a strong perpendicular magnetic field freezes the electron's motion in the bulk of a two-dimensional (2D) semiconductor but the edge states at the boundary conducts current. Thus, a 2D system subjected to a strong magnetic field behaves like an insulator in the bulk and metallic at the boundaries, by forming a discrete number of edge states. The number of edge states is directly related to the topological invariant of the bulk bands, namely the Chern number.

Over the last few decades, it has been established that the nontrivial topology of the bulk bands gives rise to the conducting edge states in the insulating materials. Such systems are usually called as the topological material/insulator [3–6]. Generally a topological phase transition may take place if a band gap closes and then reappears again as we change the parameter(s) of the system continuously while preserving the internal symmetry. The simplest topological system is an one-dimensional (1D) tight-binding dimer chain with alternating hopping parameters known as the Su-Schrieffer-Heeger (SSH) model [7]. Polyacetylene, a 1D

long polymer having large number of -CH- monomers, can be modeled by the tight-binding Hamiltonian of the SSH model. The SSH model is the simplest model to understand band topology, edge states, and bulk-boundary correspondence in 1D lattice systems. This would help us to understand more topological phenomena in higher dimensions. For example, the zero-energy edge states of the SSH model generalizes into 1D edge states of 2D topological insulators, and surface states of three-dimensional (3D) topological insulators, although the bulk-edge correspondence properties may differ. The topological phase of a finite system is manifested by the presence of exponentially localized zero-energy edge states.

The Zak phase [8] similar to the Berry phase [9] and the winding number have been used as a topological invariant quantity to classify various inversion-symmetric 1D topological systems. The Zak phase is quantized to π or 0 (mod 2π), whereas the winding number is either 1 or 0 [10]. The existence of the edge states and the Zak phase are confirmed in various artificial 1D systems such as cold atomic gases in optical lattices [11], photonic crystals [12, 13] and acoustic crystals [14].

This article is organized as follows. In Sec. II, we discuss some basic symmetry properties and bulk band structure of the SSH model Hamiltonian. The topological phase transition and bulk-boundary correspondence are discussed in Sec. III. In Sec. IV, we summarize the results. In Appendix A, we present alternative Bloch Hamiltonian for the SSH model.

II. SSH MODEL

The SSH model is described by the following nearest-neighbour (NN) tight-binding Hamiltonian

$$H = v\sum_{j=1}^{N}(a_j^\dagger b_j + b_j^\dagger a_j) + w\sum_{j=1}^{N-1}(a_{j+1}^\dagger b_j + b_j^\dagger a_{j+1}), \tag{1}$$

where $a_j^\dagger(a_j)$ and $b_j^\dagger(b_j)$ are the creation (annihilation) operator defined at the sub-lattice sites A and B, respectively in the j-th unit cell (Fig. 1). In the present case the nearest-neighbour (NN) real hopping parameters v and w may be called intra-cell and inter-cell hopping amplitudes, respectively. We will discuss

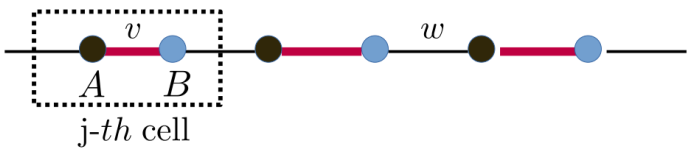

FIG. 1: Schematic diagram of the dimer SSH model. The unit cell of length a is depicted by the dotted rectangle contains two sub-lattices A and B.

some basic properties of the Bloch Hamiltonian of the SSH model. There are mainly two representations of the Bloch Hamiltonian depending on the two distinct choices of the basis states: the periodic Bloch Hamiltonian and the canonical Bloch Hamiltonian. Both the representations are useful and convenient to study certain topological aspects of the systems. We proceed with the periodic Bloch Hamiltonian. The derivation of the canonical Bloch Hamiltonian is provided in Appendix A.

A. Periodic Bloch Hamiltonian

As the system is translationally invariant, we can diagonalize the Hamiltonian using the cell index j. We expand the real space field operators defined at the sites A and B in the j-th cell

$$a_j = \frac{1}{\sqrt{N}} \sum_{k \in BZ} e^{ikaj} a_k, \; b_j = \frac{1}{\sqrt{N}} \sum_{k \in BZ} e^{ikaj} b_k, \tag{2}$$

where k is the crystal wave vector lying within the first Brillouin zone (BZ) ($-\pi$ to π), N is the number of unit cells, a is the length of the unit cell and $a/2$ is the spacing between two successive sites. Using

$$\sum_{j=1}^{N} e^{i(k-k')ja} = N\delta_{k,k'}, \tag{3}$$

the Hamiltonian in Eq. (1) reduces to the following form

$$H = v\sum_k (a_k^\dagger b_k + b_k^\dagger a_k) + w\sum_k (e^{ika} a_k^\dagger b_k + e^{-ika} b_k^\dagger a_k) = \sum_k \psi_k^\dagger \mathcal{H}(k) \psi_k, \tag{4}$$

where the spinor $\psi_k^\dagger = (a_k^\dagger, b_k^\dagger)$. Here, $\mathcal{H}(k)$ is given by

$$\mathcal{H}(k) = \begin{pmatrix} 0 & v + we^{-ika} \\ v + we^{ika} & 0 \end{pmatrix} = \begin{pmatrix} 0 & h_x(k) - ih_y(k) \\ h_x(k) + ih_y(k) & 0 \end{pmatrix} \tag{5}$$

with $h_x(k) = v + w\cos(ka)$ and $h_y(k) = w\sin(ka)$. Note that $\mathcal{H}(k)$ can be expressed in terms of the Pauli matrices (σ_x, σ_y) as

$$\mathcal{H}(k) = \boldsymbol{\sigma} \cdot \boldsymbol{h}(k), \tag{6}$$

where $\boldsymbol{h}(k)$ is $\boldsymbol{h}(k) = h_x(k)\hat{i} + h_y(k)\hat{j}$. The Pauli matrices $\boldsymbol{\sigma}$ do not represent the real spin of the fermions, they act on the sublattice sites (A, B). In real space, the SSH Hamiltonian is a $2N \times 2N$ matrix, whereas it becomes a 2×2 matrix $\mathcal{H}(k)$ for each value of k. This Hamiltonian $\mathcal{H}(k)$ is periodic in k with periodicity $2\pi/a$: $\mathcal{H}(k + 2\pi/a) = \mathcal{H}(k)$ due to the fact that only the cell index j is used to perform the Fourier transformation of the real space Hamiltonian.

Interestingly, the periodic Hamiltonian $\mathcal{H}(k)$ has the same structure as that of a spin-1/2 particle in an in-plane rotating magnetic field. Here, $\boldsymbol{h}(k)$ plays the role of the in-plane rotating magnetic field and the two sublattices A and B play the roles of spin-up and spin-down components. Therefore, the Zak phase due to variation of the parameter k equals half of the solid angle extended by a closed path in the $\boldsymbol{h}$-space.

B. Symmetry properties of the SSH model Hamiltonian

In this sub-section, we shall discuss some of basic symmetries of the Bloch Hamiltonian.

Space-inversion symmetry: The space-inversion symmetry is equivalent to the reflection symmetry. Considering the mid-point between the two sublattices A and B as the center of space-inversion and its operation simply exchanges the two sublattices. Then the space inversion operator in the pseudospin space can be simply written as σ_x. Under the space-inversion operation, the coordinates in phase space transform as $x \to -x$ and $k \to -k$. Further, $h_x(-k) = h_x(k)$ and $h_y(-k) = -h_y(k)$. Therefore, under the space inversion operation

$$\mathcal{P}\mathcal{H}(k)\mathcal{P} = \sigma_x \mathcal{H}(-k)\sigma_x = \mathcal{H}(k). \tag{7}$$

Here we have used $\sigma_x\sigma_y = -\sigma_y\sigma_x$. Thus, the Hamiltonian $\mathcal{H}(k)$ is invariant under the space-inversion with respect to the mid-point of the two sublattices.

Time-reversal symmetry: As the Pauli matrices do not represent the real spin of the charge carrier, the time-reversal operator $\mathcal{T}$ for a spinless fermion is simply $\mathcal{T} = \mathcal{C}$, with $\mathcal{C}$ being the complex conjugation operator, along with $k \to -k$. Note that $\mathcal{C}\sigma_y\mathcal{C} = -\sigma_y$ and the remaining Pauli matrices are even under time-reversal operation. Under the time-reversal operation,

$$\mathcal{T}\mathcal{H}(k)\mathcal{T} = \mathcal{C}\mathcal{H}(-k)\mathcal{C} = \mathcal{H}(k). \tag{8}$$

Therefore, the SSH Hamiltonian is invariant under time-reversal operation.

Chiral symmetry: The chiral symmetry, also known as sublattice symmetry, states that if there is a state u_k with energy $E(k)$ for a given Hamiltonian $H(k)$, there must be a state $\mathcal{S}u_k$ with energy $-E(k)$, with $\mathcal{S}$ being the chiral operator. Mathematically, it can be written as $\mathcal{S}H(k)\mathcal{S} = -H(k)$, or $H(k)\mathcal{S} = -\mathcal{S}H(k)$. The spectrum is symmetric around zero energy of a chiral symmetric Hamiltonian which explains the particle-hole symmetry. In the present case, it can be easily checked that $\sigma_z\mathcal{H}(k)\sigma_z = -\mathcal{H}(k)$. Therefore σ_z is identified as the chiral operator and the system has a chiral symmetry. The chiral symmetry will be lost if σ_z is present in the Hamiltonian $\mathcal{H}(k)$. The absence of σ_z arises due to $h_z(k) = 0$. Therefore, the chiral symmetry enforces the vector $\boldsymbol{h}(k)$ to lie on the equator plane of the Bloch sphere.

The periodic Bloch Hamiltonian $\mathcal{H}(k)$ and its eigenstates (discussed below) for each k are expressed in terms of the unit vector $\hat{\boldsymbol{h}}(k) = \boldsymbol{h}(k)/|\boldsymbol{h}(k)|$ on the equator of the Bloch sphere. The BZ of 1D lattice structure has the shape of a circle since the eigenstates at the two end-points ($k = \pm\pi$) of the BZ are equivalent. At the same time, the tip of the unit vector $\hat{\boldsymbol{h}}(k)$ traces a circle on the equator of the Bloch sphere while k runs across the BZ. The mapping of the circular BZ to the circle on the equator traced by $\hat{\boldsymbol{h}}(k)$ allows us to define a winding number ν of $\boldsymbol{h}(k)$ around the origin. Thus, the winding number ν is protected by the chiral symmetry.

There is an ambiguity of labelling $\nu = 0$ or $\nu = 1$ to trivial phase or non-trivial phase. For an infinite system, we can not determine whether the system is in trivial phase or non-trivial phase based on integer values of ν since the unit cell can be chosen in two different ways. There is no ambiguity if the changes in the winding number is $\Delta\nu = \pm 1$ while going from one phase to another phase. There is no ambiguity for finite chain since the unit cell can be chosen uniquely.

The sublattice symmetry operator can be written in terms of the sub-lattice projection operators P_α of each sub-lattice $\alpha = A, B$ as $\mathcal{S} = P_A - P_B$, where the projection operators are

$$P_A = |A\rangle\langle A| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } P_B = |B\rangle\langle B| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{9}$$

C. Energy bands, eigenstates and their properties

We obtain the energy band structure of the SSH model, by diagonalizing the Hamiltonian $\mathcal{H}(k)$, as

$$E_\pm(k) \;=\; \pm|\boldsymbol{h}(k)| = \pm\sqrt{v^2 + w^2 + 2vw\cos(ka)}. \tag{10}$$

The energy bands are symmetric around zero energy axis since the chiral symmetry is preserved. Further, the energy bands are symmetric in k, $E_\pm(-k) = E_\pm(k)$, due to the space inversion symmetry. The maximum bandwidth at the Γ-point ($k = 0$) is $2(v + w)$. The two energy bands satisfy following relations: (i) $E_+(k) + E_-(k) = 0$ as expected since trace of the Hamiltonian $\mathcal{H}(k)$ is zero and (ii) $E_+^2(k) + E_-^2(k) = 2|\boldsymbol{h}(k)|^2$, describing a circle of radius $\sqrt{2}|\boldsymbol{h}(k)|$ centered at origin.

Figure 2 displays the band dispersion for various parameters. For staggered hopping parameters, $v \neq w$ (see left and right panels of Fig. 2), there is a minimum energy gap $E_g = E_+(k) - E_-(k) = 2|w - v|$ at $k = \pm\pi/a$. On the other hand, the gap E_g closes at $k = \pm\pi/a$ when $v = w$, which is equivalent to a monatomic chain of lattice constant $a/2$ with the metallic dispersion $E(k) = 2t\cos(ka/2)$. Therefore, staggered hopping parameters are required in order to open a gap in the dispersion. Moreover, the gap closing at $v = w$ indicates a topological phase transition.

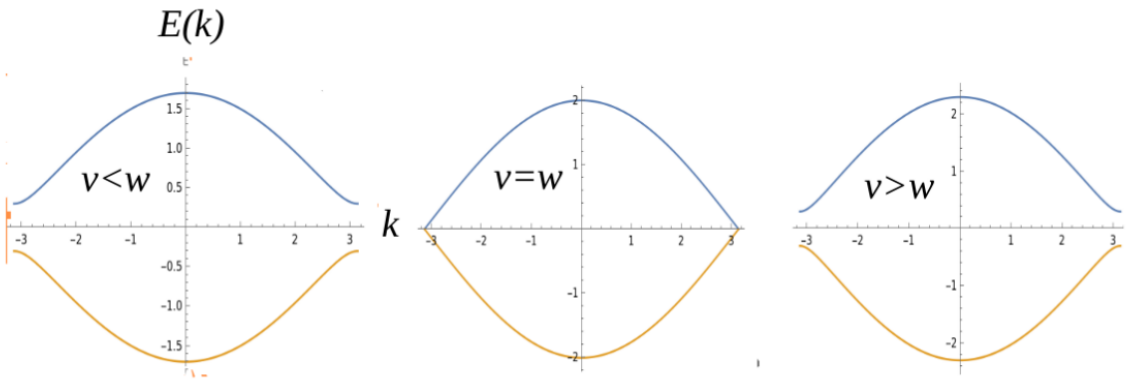

FIG. 2: Energy band dispersion: Left panel: $v = 0.7, w = 1$, Middle panel: $v = w = 1$, Right panel: $v = 1.3, w = 1$.

The eigenspinors of $\mathcal{H}(k)$ located on the equator of the Bloch sphere are given by

$$u_+(k) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ e^{i\gamma_k} \end{pmatrix} \quad \text{and} \quad u_-(k) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -e^{i\gamma_k} \end{pmatrix} \tag{11}$$

with the phase angle $\gamma_k = \arctan(h_y/h_x)$ in the h_x-h_y plane on the equator of the Bloch sphere. Both the eigenspinors are well-defined for all values of k with $v \neq w$. Note that $u_-(k) = \sigma_z u_+(k)$ as expected from the chiral symmetry.

Let us first rewrite the band dispersion $E_\pm(k)$ in terms of $z = \delta/(2t)$ with $t = (v+w)/2$ and $\delta = w - v$ as

$$E_\pm(k) \;=\; \pm 2t\sqrt{1 - (1 - z^2)\sin^2(ka/2)}. \tag{12}$$

This form of $E_\pm(k)$ will be convenient to study the group velocity, density of states and effective mass.

The group velocity can be obtained as

$$v_g = \frac{1}{\hbar}\frac{dE}{dk} \;=\; \frac{at}{\hbar}\frac{(z^2-1)\sin(ka)}{\sqrt{1 - (1 - z^2)\sin^2(ka/2)}}. \tag{13}$$

Total number of states per unit length in one-dimension is $N(k) = k/\pi$. The density of states is then given by

$$\frac{dN(k)}{dE} = \frac{1}{\pi}\frac{dk}{dE}.$$

The density of states of either bands for the spinless fermions can be obtained as

$$D(E) = \frac{1}{\pi}\frac{dk}{dE} = \frac{1}{\pi a}\frac{|E|}{\sqrt{(4t^2 - E^2)(E^2 - \delta^2)}}. \tag{14}$$

The density of states has Van Hove singularity at $E = \delta$ and at $E = 2t$. There is no available density of states in the gap region $-\delta < E < \delta$.

Limiting case: Note that when $\delta = 0$, the lattice is equivalent to a monatomic chain of lattice constant $a/2$ with the dispersion $E = 2t\cos(ka/2)$. The BZ is now $-2\pi/a \leq k \leq 2\pi/a$. The band minimum ($E_{min} = -2t$) occurs at $k = \pm 2\pi/a$ and maximum $E_{max} = 2t$ at $k = 0$.

In this case ($z = 0$), the group velocity becomes $v_g = -(at/\hbar)\sin(ka/2)$. The density of states correctly reproduces to

$$D(E) = \frac{1}{\pi a}\frac{1}{\sqrt{(4t^2 - E^2)}}. \tag{15}$$

III. BULK AND EDGE STATES OF A FINITE SSH MODEL

The SSH model supports the bulk as well as the edge states. In the previous Section, we have discussed properties of the bulk bands for an infinite SSH chain. The difference between the bulk and the edge states is characterized by the spreading of energy eigenstates in the real space. Now we shall consider a finite chain and study both the bulk and the edge states by diagonalizing the Hamiltonian in Eq. 1 in real space. As an example, here we provide the matrix eigenvalue equation, $H\Psi = E\Psi$, for 4 unit cells as given by

$$H\Psi = \begin{pmatrix} 0 & v & 0 & 0 & 0 & 0 & 0 & 0 \\ v & 0 & w & 0 & 0 & 0 & 0 & 0 \\ 0 & w & 0 & v & 0 & 0 & 0 & 0 \\ 0 & 0 & v & 0 & w & 0 & 0 & 0 \\ 0 & 0 & 0 & w & 0 & v & 0 & 0 \\ 0 & 0 & 0 & 0 & v & 0 & w & 0 \\ 0 & 0 & 0 & 0 & 0 & w & 0 & v \\ 0 & 0 & 0 & 0 & 0 & 0 & v & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ a_3 \\ b_3 \\ a_4 \\ b_4 \end{pmatrix} = E \begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \\ a_3 \\ b_3 \\ a_4 \\ b_4 \end{pmatrix}. \tag{16}$$

Here, the eigenvector for four unit cells is $\Psi = (a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4)^T$ with the subscript in (a_j, b_j) denotes the cell index and T denotes the transpose. One can easily extend the matrix eigenvalue equations for an arbitrary N cells and solve the eigenvalue problem for all $2N$ discrete eigenvalues E as well as the corresponding real space eigenvectors.

Figures 3(a) and 3(b) display the spectrum of a finite SSH model of $N = 12$ and $N = 25$ unit cells considering the open boundary condition, while varying v for fixed $w = 1$. A careful analysis of the dispersion reveals that there are two nearly zero-energy states within the bulk band gap when $v < w$. The eigenvectors of the zero-energy edge states for $v = 0.2$ in the form of symmetric and anti-symmetric pairs (blue dot for $N = 12$) are shown in Figs. 3(c) and 3(d). Only the sub-lattice A at the left edge and the

sub-lattice B at the right edge are occupied. The amplitude of the eigenvectors inside the lattice is simply zero, thus the eigenvectors are exponentially localized.

The amplitude of the eigenvectors begins to spread inside the lattice and the bulk band gap shrinks to zero as v approaches to w. Similar to the previous case, two eigenvectors of the nearly-zero-energy edge states at $v = 0.7$ (maroon dot for $N = 12$) are shown in Figs. 3(e) and 3(f). It shows that amplitude of the eigenvectors become finite inside the lattice, giving rise to splitting in the energy levels. The edge state energies for $N = 12$ case start splitting earlier than that of the $N = 25$ case, due to the finite size effects.

There is no edge states within the bulk gap when $v > w$. The eigenvectors for $v = 2.5$ (red dots for $N = 12$) are shown in Figs. 3(e) and 3(f). The eigenvector profile of the bulk modes looks like a standing wave on the chain.

For $v < w$, only A sub-lattices are occupied at left half of the lattice, whereas B sub-lattices are occupied at right-half of the lattice. On the other hand, both the sub-lattices are occupied across the lattice for $v > w$ case. These zero-energy states localized at the edges are responsible for the topological phase for $v < w$ case. This phase can be characterized by the topological invariants, namely, the Zak phase and the winding number.

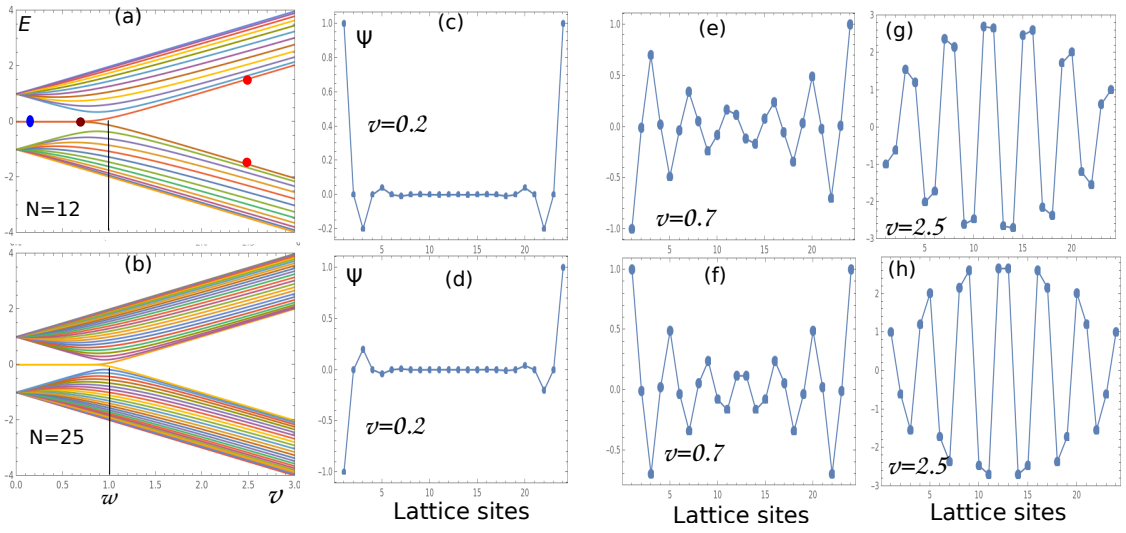

FIG. 3: Bulk bands and edge states of the finite SSH model in the open boundary condition, as a function v for fixed $w = 1$. Second, third and fourth panels show the eigenvectors of the hybridized states for various values of v.

A. Space inversion symmetry and sub-lattice symmetry operator for finite SSH Hamiltonian

Like the space inversion symmetry operator for bulk Hamiltonian, one can also construct the same for the finite SSH Hamiltonian. It is given by I = off-diag(1, 1,.....)$_{2N}$, satisfying the relation $IHI = H$.

The sub-lattice operator for the finite SSH Hamiltonian can be written as $\Sigma = P_A - P_B$, where the sub-lattice projection operators for the finite system are $P_A = \sum_{j=1}^{N} |j, A\rangle\langle j, A|$ and $P_B = \sum_{j=1}^{N} |j, B\rangle\langle j, B|$. For $N = 2$, the sub-lattice symmetry operator can be expressed in matrix form as

$$\Sigma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{17}$$

It can be easily checked that $\Sigma H \Sigma = -H$.

B. Approximate expression of the zero-energy eigenstates

The SSH Hamiltonian is

$$\mathcal{H} = [v + w\cos(ka)]\sigma_x + w\sin(ka)\sigma_y. \tag{18}$$

Expanding $\cos(ka)$ and $\sin(ka)$ around the gap closing points $ka = \pm\pi$ as $ka = \pm\pi + qa$, and keeping upto linear in qa, we get the low-energy Hamiltonian around the gap closing point as

$$\mathcal{H} \simeq (v - w)\sigma_x - wqa\sigma_y. \tag{19}$$

Setting $q \to -i\partial_x$, the Hamiltonian in real space reads as

$$\mathcal{H} = (v - w)\sigma_x + iwa\sigma_y\partial_x. \tag{20}$$

Let us see if there is any zero-energy edge state under certain conditions or not. The zero-energy eigenvalue equation can be written as

$$[(v - w)\sigma_x + iwa\sigma_y\partial_x]\psi(x) = 0, \tag{21}$$

where $\psi(x)$ is the zero-energy eigenstate. Multiplying σ_y, we get

$$\sigma_0\partial_x\psi(x) \;=\; \frac{(v - w)}{wa}\sigma_z\psi(x), \tag{22}$$

where σ_0 is the 2×2 identity matrix.

Consider a finite SSH chain of length $L = Na$ with N number of unit cells and the left end is at $x = 0$. Now, we substitute the following ansatz in the above equation, which will satisfy the boundary condition at $x = L$

$$\psi_r(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \phi_r(x).$$

The first-order differential equation for $\phi_r(x)$ has the solution

$$\phi_r(x) = e^{\frac{(v-w)}{wa}(L-x)}. \tag{23}$$

It clearly shows that the condition, $w > v$, must be satisfied in order to get a normalizable solution. Thus the zero-energy (unnormalized) eigenstate at the right edge is

$$\phi_r(x < L) = e^{-\frac{(w-v)}{wa}(L-x)}, \tag{24}$$

with $w > v$. Similarly, the zero-energy (unnormalized) eigenstate at the left edge can be obtained as

$$\psi_l(x > 0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-\frac{(w-v)}{wa}x}, \tag{25}$$

with $v < w$.

Both the eigenstates $\psi_l(x)$ and $\psi_r(x)$ are exponentially decaying functions with the decay length $\xi_c = wa/(w-v)$. The decay length decreases with the increase of $(w-v)$. Further, the spinor parts of $\psi_l(x)$ and $\psi_r(x)$ show that sublattice A at the left edge and the sublattice B at the right edge support a zero-energy state.

IV. BAND GEOMETRIC QUANTITIES AND BULK TOPOLOGICAL INVARIANT

The Berry connection for $E_-(k)$ band is

$$A_-(k) = i\langle u_-(k)|\partial_k u_-(k)\rangle = \frac{1}{2}\frac{d\gamma_k}{dk}. \tag{26}$$

The Zak phase acquired which is while traveling along a closed path in BZ is given by

$$Z_- = \int_{BZ} A_-(k)dk = \frac{1}{2}\oint_{BZ} d\gamma_k. \tag{27}$$

A care must be taken while calculating the above integral of the phase angle γ_k in $\boldsymbol{h}(k)$-space for two different cases: $v < w$ and $v > w$. Note that $h_x(k) = v + w\cos(ka)$ and $h_y(k) = w\sin(ka)$ parameterize a circle. When we vary k across the first BZ, the tip of $\boldsymbol{h}(k)$ sketches a circle of radius w whose center is

at v in the h_x-h_y plane as shown in Fig. 4. The gap closing points $ka = \pm\pi$ for $v = w$ correspond to the origin of the h_x-h_y plane. When $w > v$, the trajectory of $\boldsymbol{h}$ encloses the origin when k varies across the first BZ. Therefore, $\oint_{BZ} d\gamma_k = 2\pi$ and so the Zak phase is π. When $w < v$, the trajectory of the vector $\boldsymbol{h}$ excludes the origin, which leads to $\oint_{BZ} d\gamma_k = 0$ and so the Zak phase for this path is simply zero. The Zak phase for different cases can be written as

$$Z_- = \begin{cases} \pi & v < w \text{ (topological insulator)} \\ 0 & v > w \text{ (band insulator)} \end{cases} \tag{28}$$

The topology of this kind of closed loop can also be characterized by an integer, the bulk winding number ν. This counts the number of times the loop winds around the origin of the h_x-h_y plane. We have $\nu = 0$ when $w < v$, $\nu = 1$ when $w > v$ and ν is undefined when $w = v$ since the loop passes through the origin.

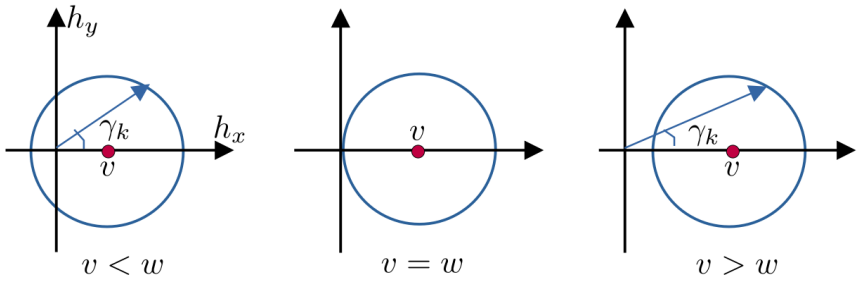

FIG. 4: Sketch of the trajectory of the tip of the $\boldsymbol{h}(k)$ vector while k varies in the first BZ for different choices of the hopping energies v and w.

Electric polarization: The electric polarization is created due to spatially separated charges. The electric polarization at k of a given band can be defined as

$$P_n(k) = e\langle x\rangle = ie\langle\partial_k\rangle. \tag{29}$$

The total polarization is the sum over all the occupied states in a given band

$$P_n = \frac{e}{2\pi}\oint P_n(k)dk = \frac{e}{2\pi}\oint i\langle\partial_k\rangle dk = \frac{e}{2\pi}\oint A(k)dk = \frac{e}{2\pi}Z_n. \tag{30}$$

Therefore, the electric polarization P is given by

$$P = \frac{e}{2\pi}\oint_{BZ} A(k)dk \tag{31}$$

$$= \begin{cases} \frac{e}{2}, & v < w \\ 0 & v > w. \end{cases} . \tag{32}$$

V. SUMMARY

In this article, we have discussed some interesting topological properties of the dimer SSH model. Various symmetries like the parity, the time-reversal and the chiral symmetries of the bulk Hamiltonian are also mentioned. We provide exact analytical expressions of the bulk band structures and corresponding eigenstates. The appearance of the boundary modes is shown by studying a finite SSH chain and establishing the bulk-boundary correspondence.

APPENDIX A: CANONICAL BLOCH HAMILTONIAN

In order to get the canonical Bloch Hamiltonian from the real space Hamiltonian given in Eq. (1), we use the Fourier transformation

$$a_j = \frac{1}{\sqrt{N}} \sum_k e^{ikaj} a_k,\ b_j = \frac{1}{\sqrt{N}} \sum_k e^{ika(j+1/2)} b_k. \tag{33}$$

The real space Hamiltonian H reduces to the form

$$H = v \sum_k (e^{ika/2} a_k^\dagger b_k + e^{-ika/2} b_k^\dagger a_k) + w \sum_k (e^{-ika/2} a_k^\dagger b_k + e^{ika/2} b_k^\dagger a_k) \tag{34}$$

$$= \sum_k \psi_k^\dagger H(k) \psi_k, \tag{35}$$

where $\psi_k^\dagger = (a_k^\dagger, b_k^\dagger)$. Here, the Hamiltonian $H(k)$ is given by

$$H(k) = \begin{pmatrix} 0 & d_x(k) - id_y(k) \\ d_x(k) + id_y(k) & 0 \end{pmatrix} \tag{36}$$

with $d_x(k) = (v + w)\cos(ka/2)$ and $d_y(k) = (w - v)\sin(ka/2)$. The canonical Bloch Hamiltonian can be further written as

$$H(k) = \boldsymbol{\sigma} \cdot \boldsymbol{d}(k), \tag{37}$$

where $\boldsymbol{d}(k) = \hat{i} d_x(k) + \hat{j} d_y(k)$.

Note that the periodicity of the Bloch Hamiltonian has increased by two times: $H(k + 4\pi/a) = H(k)$ because the separation distance between the sites in a given unit cell becomes half of the size of the unit cell. It is noteworthy that the canonical Bloch Hamiltonian $H(k)$ is directly obtained from the periodic Bloch Hamiltonian $\mathcal{H}(k)$ via the unitary transformation: $H(k) = U\mathcal{H}(k)U^\dagger$, where the unitary matrix U is given by $U = \text{diag}(1, e^{ika})$.

As expected, the energy dispersion is given by

$$E_{\pm}(k) \;=\; \pm|\boldsymbol{d}(k)| = \pm\sqrt{v^2 + w^2 + 2vw\cos(ka)}. \tag{38}$$

The eigenspinors for the canonical Bloch Hamiltonian $H(k)$ located on the equator of the Bloch sphere are

$$u_{+}(k) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ e^{i\phi_k} \end{pmatrix}, \qquad u_{-}(k) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -e^{i\phi_k} \end{pmatrix}, \tag{39}$$

where $\phi_k = \arctan[d_y(k)/d_x(k)]$. The eigenspinors are not single-valued and ill-defined at $k = \pm\pi/a$.

---

* Electronic address: tkghosh@iitk.ac.in

[1] N. W. Ashcroft and N. D. Mermin, Solid State Physics.

[2] K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45** 494, (1980).

[3] Topological Insulators: Dirac equation in condensed matter, Springer series in solid state sciences 174.

[4] J. K. Asbóth, L. Oroszlány and A. Pályi, A Short Course on Topological Insulators, arXiv:1509.02295v1.

[5] D. Vanderbilt, Berry Phases in Electronic Structure Theory (Cambridge: Cambridge University Press, 2018).

[6] J. Cayssol and J. N. Fuchs, J. Phys. Mater. **4**, 034007 (2021).

[7] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Solitons in polyacetylene, Phys. Rev. Lett. **42**, 1698 (1979).

[8] J. Zak, Berry's phase for energy bands in solids, Phys. Rev. Lett. **62**, 2747 (1989).

[9] M. V. Berry, Quantal Phase Factors Accompanying Adiabatic Changes, Proc. R. Soc. Lond. A **392**, 45 (1984).

[10] P. Delplace, D. Ullmo, and G. Montambaux, Zak phase and the existence of edge states in graphene, Phys. Rev. B **84**, 195452 (2011).

[11] M. Atala, M. Aidelsburger, J. T. Barreiro, D. Abanin, T. Kitagawa, E. Demler, and I. Bloch, Direct measurement of the zak phase in topological Bloch bands, Nature Physics **9**, 795 (2013).

[12] M. Xiao, Z. Q. Zhang, and C. T. Chan, Surface Impedance and Bulk Band Geometric Phases in One- Dimensional Systems, Phys. Rev. X **4**, 021017 (2014).

[13] W. S. Gao, M. Xiao, C. T. Chan, and W. Y. Tam, Deter- mination of Zak phase by reflection phase in 1D photonic crystals, Opt. Lett. **40**, 5259 (2015).

[14] M. Xiao, G. C. Ma, Z. Y. Yang, P. Sheng, Z. Q. Zhang, and C. T. Chan, Geometric phase and band inversion in periodic acoustic systems, Nat. Phys. **11**, 240 (2015).

J. Phys. Res. Edu., Vol. 2, March 2024

The Exactly Solvable Pöschl-Teller Potential

Parikshit Dutta*

Department of Physics, Asutosh College, 92 Shyama Prasad Mukherjee Road, Kolkata 700026, India

We revisit a well known quantum mechanical problem, the trigonometric Pöschl-Teller potential, which is an exactly solvable one-dimensional problem. The potential appears in many physical systems of study and thus the technique of solution is interesting and important for students and readers to know. We wish to elaborate on this in this article.

I. INTRODUCTION

An interesting problem in one-dimensional quantum mechanics is the trigonometric Pösch-Teller potential. This appears in several interesting scenarios in physics [1–3]. Although it is complicated to look at, it turns out that it is exactly solvable. Here we follow Ref. [4] and elaborate the technique of solution for the purpose of the reader. The form of the potential is given by

$$V(x) = \frac{V_0}{2}\left\{\frac{\chi(\chi-1)}{\sin^2\alpha x} + \frac{\lambda(\lambda-1)}{\cos^2\alpha x}\right\} \tag{1}$$

with $\chi > 1$ and $\lambda > 1$ and $V_0 = \hbar^2\alpha^2/m$. This potential is bounded on both sides by singularities at $x = 0$ and $x = \frac{\pi}{2\alpha}$. As $\chi \to 1$, the singularity becomes less and less pronounced at $x = 0$, until at $\chi = 1$, it vanishes. In that case the potential becomes

$$V(x) = \frac{V_0}{2}\frac{\lambda(\lambda-1)}{\cos^2\alpha x} = \frac{V_0}{8}\left[\frac{1}{\sin^2(\frac{\alpha x}{2}-\frac{\pi}{4})} + \frac{1}{\cos^2(\frac{\alpha x}{2}-\frac{\pi}{4})}\right], \tag{2}$$

which again has the same form, with singularities at $\alpha x = \pm\frac{\pi}{2}$, and $\chi = \lambda$. Thus in the $\chi = 1$ case, the two potential 'holes' between $-\frac{\pi}{2}$ and 0 and 0 and $\frac{\pi}{2}$ unite into one large hole. Thus the potential form remains same with redefinitions $\alpha \to \alpha/2$ and $V_0 \to V_0/4$. More over, the potential is clearly periodic but for solving the Schrodinger equation, it plays no role, since the barriers are impenetrable. Thus we pick a single hole, say the interval $0 \le x \le \frac{\pi}{2\alpha}$ and solve the Schrodinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x). \tag{3}$$

$\psi(x)$ is obviously a bound state since we have infinite barriers at 0 and $\frac{\pi}{2\alpha}$. Eq. (4) can be rewritten as

$$\frac{d^2\psi(x)}{dx^2} - \frac{2m}{\hbar^2}V(x)\psi(x) + \frac{2m}{\hbar^2}E\psi(x) = 0. \tag{4}$$

II. SOLUTION BY HYPERGEOMETRIC SERIES

Making the following change of variables : $y = \sin^2 \alpha x$, Eq. (4) becomes

$$y(1-y)\frac{d^2\tilde{\psi}(y)}{dy^2} + \left(\frac{1}{2} - y\right)\frac{d\tilde{\psi}(y)}{dy} + \frac{1}{4}\left(\frac{\kappa^2}{\alpha^2} - \frac{\chi(\chi-1)}{y} - \frac{\lambda(\lambda-1)}{y}\right)\tilde{\psi}(y) = 0, \tag{5}$$

where $\kappa^2 = \frac{2mE}{\hbar^2}$. Here, we denote $\psi(x)$ as $\tilde{\psi}(y)$. This differential equation has three regular singular points at $y = 0, 1, \infty$ and is very close to the hyper-geometric differential equation

$$z(1-z)f''(z) + [c - (a+b+1)z]f'(z) - abf(z) = 0. \tag{6}$$

To we recast them in the above form, assume a series solution about $y = 0$, $\tilde{\psi}(y) = \sum_{n=0}^{\infty} a_n y^{\mu+n}$. Equating the lowest power of y i.e. $y^{\mu-1}$ we get the indicial equation

$$\mu(\mu-1) + \frac{1}{2}\mu - \frac{1}{4}\chi(\chi-1) = 0. \tag{7}$$

Since $a_0 \neq 0$, we get the two roots of the indicial equation : $\mu = \frac{\chi}{2}, \frac{1-\chi}{2}$. Since $\chi > 1$ the second indicial root leads to a solution of the form

$$\lim_{y\to 0} y^{\frac{1-\chi}{2}} \sum_{n=0}^{\infty} a_n y^n \to \infty. \tag{8}$$

Thus, we choose the first indicial root. We can do the same about the point $y = 1$. Assuming a series solution about $y = 1$, $\tilde{\psi}(y) = \sum_{n=0}^{\infty} \bar{a}_n (1-y)^{\nu+n}$, we get the indicial equation

$$\nu(\nu-1) + \frac{\nu}{2} - \frac{1}{4}\lambda(\lambda-1) = 0. \tag{9}$$

We get the following roots $\nu = \frac{\lambda}{2}, \frac{1-\lambda}{2}$. Again, we choose the first root as the second root leads to singular behavior at $y = 1$. Thus, we can again write

$$\tilde{\psi}(y) = y^{\frac{\chi}{2}}(1-y)^{\frac{\lambda}{2}} f(y). \tag{10}$$

Plugging this into E. (5), we get the equation for $f(y)$

$$y(1-y)f''(y) + [(\frac{1}{2} + \chi) - (1 - \chi + \lambda)y]f'(y) + \frac{1}{4}[\frac{\kappa^2}{\alpha^2} - (\chi+\lambda)^2]f(y) = 0. \tag{11}$$

This is finally in Hypergeometric form with $a + b = \chi + \lambda$, $c = \chi + \frac{1}{2}$, $ab = \frac{1}{4}[(\chi+\lambda)^2 - (\chi+\lambda)^2]$. Since the equation is symmetric in a and b we can choose

$$\begin{aligned} a &= \frac{1}{2}(\chi + \lambda + \frac{\kappa}{\alpha}) \\ b &= \frac{1}{2}(\chi + \lambda - \frac{\kappa}{\alpha}). \end{aligned} \tag{12}$$

The Hypergeometric equation can now be solved using Frobenius method, with the trial solution

$$f(y) = \sum_{n=0}^{\infty} a_n y^{k+n} \tag{13}$$

from which we get the two possible values of $k = 0, 1-c$ and the recurrence relation

$$a_{n+1} = a_n \frac{(k+n+a)(k+n+b)}{(k+n+1)(k+n+c)} \tag{14}$$

which yields the solution

$$f_k(y) = a_0 \sum_{n=0}^{\infty} \frac{(k+a)_n (k+b)_n}{(k+1)_n (k+c)_n} y^{k+n}. \tag{15}$$

Here, the Pochammer symbol is defined as

$$\begin{aligned} (k+w)_n &= (k+w)(k+w+1)\dots(k+w+n-1) \\ &= \frac{\Gamma(k+w+n)}{\Gamma(k+w)}. \end{aligned} \tag{16}$$

Hence, the two solutions are

$$\begin{aligned} f_0(y) &= {}_2F_1(a,b,c;y) \\ f_{1-c}(y) &= y^{1-c}{}_2F_1(a+1-c, b+1-c, 2-c; y). \end{aligned} \tag{17}$$

Now, the full solution in our case is

$$f(y) = c_1 {}_2F_1(a,b,c,y) + c_2 y^{1-c}{}_2F_1(a+1-c, b+1-c, 2-c; y) \tag{18}$$

with $a = \frac{1}{2}(\chi + \lambda + \frac{\kappa}{\alpha})$, $b = \frac{1}{2}(\chi + \lambda - \frac{\kappa}{\alpha})$, $c = \chi + \frac{1}{2}$. Now we want to impose the boundary conditions that the wave function vanishes at $x = 0$ and $x = \frac{\pi}{2\alpha}$ as $V \to \infty$ at those points i. e.

$$\tilde{\psi}(0) = 0, \;\; \tilde{\psi}(1) = 0. \tag{19}$$

Note that

$$\lim_{y\to 0} {}_2F_1(a,b,c;y) \to 1. \tag{20}$$

Thus

$$\lim_{y\to 0} \tilde{\psi}(0) = \lim_{y\to 0} y^{\frac{\chi}{2}} (1-y)^{\frac{\lambda}{2}} f(y) = \lim_{y\to 0} (1-y)^{\frac{\lambda}{2}} \left[c_1 y^{\frac{\chi}{2}} + c_2 y^{\frac{1-\chi}{2}} \right] \to \infty. \tag{21}$$

Since $\chi > 1$, the second term is singular, so we must have $c_2 = 0$. For the case $y \to 1$, it is a bit more subtle. First, we observe the domain convergence of the hypergeometric series using De Alembert's ratio test

$$\lim_{n\to\infty} \left| \frac{(a)_{n+1}(b)_{n+1} y^{n+1}}{(1)_{n+1}(c)_{n+1}} \frac{(1)_n c_n}{(a)_n (b)_n y^n} \right| \to |y|. \tag{22}$$

Thus, the series diverges for $y > 1$. For $y = 1$, since $a_{n+1} \sim a_n$ as $n \to \infty$, the series again diverges. Thus, we can get a solution $\tilde{\psi}(y)$ which vanishes at $y = 1$ if the series truncates. Note that this heuristic argument does not take into account the rate of divergence of the hypergeometric series which plays an overall role required for convergence due to the factor of $(1-y)^{\frac{\lambda}{2}}$ appearing in $\tilde{\psi}(y)$, but it is sufficient for the purpose. Now looking at the Pochammer symbol

$$\begin{aligned}(a)_n &= a(a+1)\dots(a+n-1)\\ (b)_n &= b(b+1)\dots(b+n-1).\end{aligned} \tag{23}$$

If a and b are 0 or negative integers say $-m$, then the series truncates as the $(m+1)^{\text{th}}$ onward coefficient in the hypergeometric series becomes 0. It $a = -m$, then $b = \chi + \lambda + m$ and vice versa. Since the solution is invariant under exchange of a, b both conditions are equivalent. Finally, we can find the energy eigen value as

$$\begin{aligned}a.b &= -m(\chi+\lambda+m) = \frac{1}{4}\Big[(\chi+\lambda)^2 - \frac{\kappa^2}{\alpha^2}\Big]\\ \Rightarrow E_m &= \frac{V_0}{2}(\chi+\lambda+2m)^2.\end{aligned} \tag{24}$$

The corresponding eigenfunction function finally takes the form

$$\begin{aligned}\tilde{\psi}(y) &= c_1 y^{\frac{\chi}{2}}(1-y)^{\frac{\lambda}{2}}{}_2F_1(a,b,c;y)\\ &= c_1 \sin^{\chi}(\alpha x)\cos^{\lambda}(\alpha x){}_2F_1(-m,\chi+\lambda+m,\chi+\frac{1}{2};\sin^2\alpha x).\end{aligned} \tag{25}$$

III. CONCLUSION

Here, we provide a review of the method for solving the trigonometric Pösch-Teller potential, which is an exactly solvable one-dimensional quantum mechanical problem. The primary goal of this review was to elaborate the solution to the readers, on how a complicated looking problem can still be solved exactly. In fact, this potential plays an important role in the study of supersymmetric quantum mechanics as well, and the reason for the solvability could be guessed from the nature of the potential. The interested readers may refer to Ref. [5].

* Electronic address: parikshit.dutta@asutoshcollege.in

[1] A. Contreras-Astorga and V. Hussin, arXiv:1901.04606.

[2] A. Contreras-Astorga and D. J. Fernandez C, Journal of Physics A: Mathematical and Theoretical **41**, 475303 (2008).

[3] M. Hamzavi and A. A. Rajabi, Advances in High Energy Physics **2013**, 196986 (2013).

[4] S. Flugge, *Practical Quantum Mechanics*: Classics in Mathematics, Springer-Verlag Berlin Heidelberg 1971.

[5] A. Khare, Pramana **49**, 41 (1997).

J. Phys. Res. Edu., Vol. 2, March 2024

A brief overview of Nanomechanical Qubits

Chandan Samanta*

Department of Physics, Indian Institute of Science Education and Research Bhopal, Bhopal-462066, Madhya Pradesh, India

Qubits are the building blocks for quantum computers and quantum information processing. However, there is a great deal of dispute over the most ideal types of qubits. The nanomechanical qubit might be one potential addition to the qubit platforms. Here, we briefly outline the underlying physical principle of a nanomechanical qubit, where the mechanical vibrations stores the information.

I. INTRODUCTION

Qubits serve as the fundamental units for quantum computers and quantum information processing. Research and development in this realm are rapidly progressing to determine which system or platform will emerge as the frontrunner. Among the array of possibilities are superconducting Josephson junctions[1, 2], semiconductor qubits [3–5], trapped ions [6–8], topological qubits [9, 10], ultra-cold neutral atoms [11, 12], and even diamond vacancies [13–15]. However, only a select few qubit platforms have thus far showcased the potential for quantum computing, meeting criteria such as high-fidelity controlled gates, effortless qubit-qubit coupling, and robust isolation from environmental interference, ensuring prolonged coherence. Nanomechanical resonators stand as a potential contender within this elite group of platforms. Over the past three decades, we have observed remarkable advancements in micro/nano-electromechanical systems [16–18]. This progress has led to technology that demonstrates exceptional performance across various applications, including sensing, imaging, timing, signal processing, and logic devices, among others. Moreover, these devices provide an intriguing opportunity to explore fundamental physics phenomena that include the quantum effects in macroscopic objects and the coupling of mechanical motion with photons, spins, and electrons.

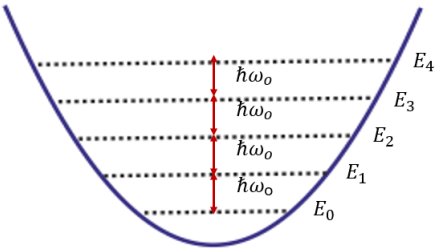

FIG. 1: The quantized energy levels of a simple harmonic oscillator having a frequency ω_0. The energy levels are evenly spaced.

A mechanical resonator is usually modeled as a harmonic oscillator. The energy of a simple harmonic oscillator can be written as $E_n = (n+1/2)\hbar\omega_0$, where $n = 0, 1, 2, ...$, $\hbar$ is the reduced Planck constant, and ω_0 is the resonance frequency of the oscillator. Here, $n = 0$ corresponds to the zero point energy with a corresponding fluctuation in displacement known as zero point displacement fluctuation $x_{\rm zp} = \sqrt{\hbar/(2m\omega_0)}$. This allows a mechanical resonator to be a test-bed for the quantum nature of macroscopic objects. There has been an intense effort to study mechanical resonators in the quantum regime. Several cool-down techniques have been implemented to achieve the ground state of a mechanical resonator. These devices are becoming a crucial new avenue in quantum science and technology. A plethora of proposals suggests their utility in storing, processing, and transducing quantum information. This requires progressively advanced methods to control mechanical motion within the quantum realm. These devices have been employed in achieving quantum ground state [19, 20], quantum squeezing [21–23], backaction-evading measurements [24, 25], entanglement [26, 27], coherent microwave-optical interface [28, 29], and superconducting qubit-mechanical interfaces [19, 30].

A recurrent question has been whether it is possible to realize strong nonlinearities in nanomechanical resonators approaching the quantum ground state. The origin of nonlinearity in case of large displacements is connected to the stress that depends nonlinearly on the displacement of a particular mode. It modifies the Hooks law as $F = -m\omega_0^2 x - \gamma x^3$ where γ is the the weak Duffing (or Kerr) constant and m is effective mass of the the mechanical eigenmode [31, 32]. The conventional devices are perfectly linear in the quantum regime. This means that the energy levels are equidistant. The vibrations in these resonators approaching the quantum ground state are only nonlinear at much larger values of x_{nl}, typically at $x_{nl}/x_{\rm zp} = 10^6$ or above. There have been efforts to introduce nonlinearity in mechanical resonators without much success. The proposition of anharmonicity resulting from proximity to a buckling instability has been made [33, 34]. Nevertheless, implementing such a scheme experimentally poses significant challenges.

Pistolesi *et al* [35] recently introduced a theoretical framework for a mechanical qubit utilizing the

coupling of one of the flexural modes of a carbon nanotube resonator to the charge state of a double-quantum dot defined in the same nanotube. It becomes feasible to generate adequate anharmonicity in the mechanical oscillator, enabling the coupled system to function as a mechanical quantum bit. Nevertheless, it is attainable solely when the device is operated in the ultrastrong coupling regime.

II. THE DEVICE AND THE ORIGIN OF ANHARMONICITY

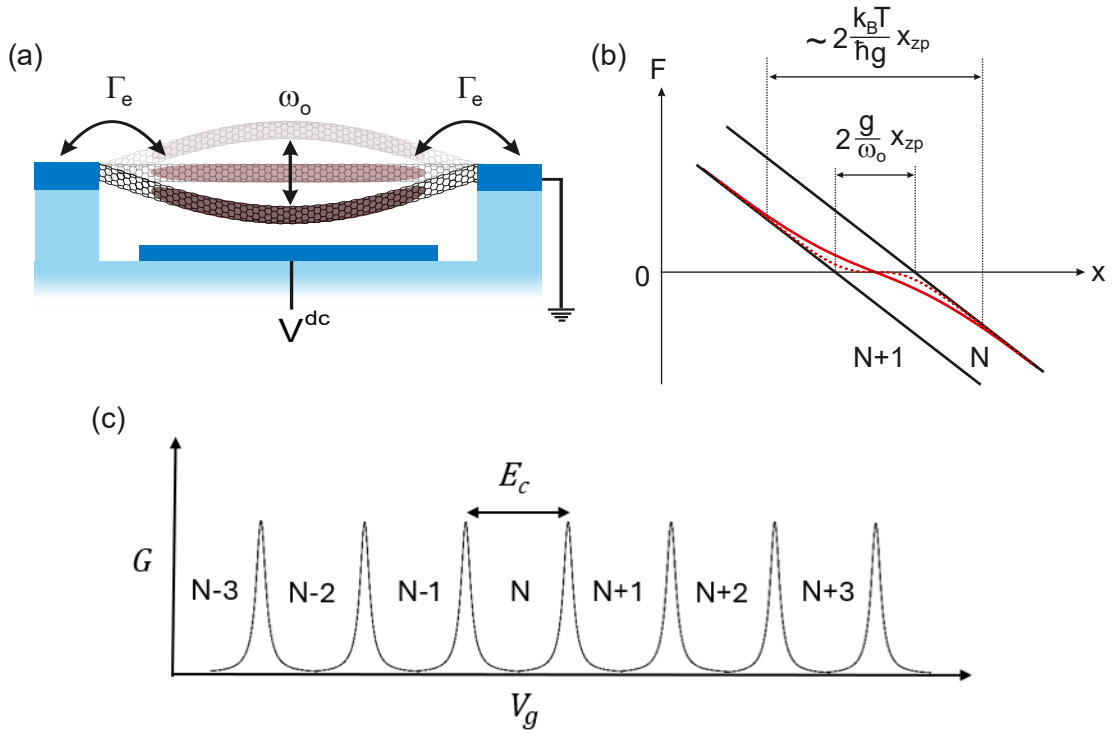

FIG. 2: (a) A schematic diagram that depicts the nanotube oscillating at ω_0. A quantum dot, shown in purple, emerges along the suspended nanotube, with an overall electron tunneling rate $\Gamma_{\rm e}$ to the left and right leads. (b) When the fluctuations between the two electron states are rapid compared to the mechanical motion ($\Gamma_{\rm e} > \omega_0$), the force experienced by the vibrations is an average of the two linear F vs x curves (black straight lines) represents the situation in which the dot is filled with either N or $N+1$ electrons weighted by the Fermi-Dirac distribution. The resulting force becomes nonlinear (shown in red) when the vibration displacement is smaller than $k_{\rm B}Tx_{\rm zp}/(\hbar g)$. The reduced slope at zero displacement is a measure of the decrease in ω_0. As the temperature decreases, nonlinearity increases, as evidenced by the further reduction in slope at zero vibration, as depicted by the dashed red line. The separation between the two curves is $\Delta x = 2(g/\omega_0)x_{\rm zp}$ due to the force generated by the electron tunneling onto the quantum dot. (c) A schematic of the conductance G as a function of the gate voltage.

Recently, there has been an experimental demonstration of a new mechanism to boost the anharmonicity of a carbon nanotube mechanical resonator [36]. This was achieved by coupling the resonator with single electron tunneling (SET) via a quantum dot non-resonantly. Figure 2(a) shows the schematic of a vibrating nanotube at a resonance frequency ω_0. The typical dimensions of the devices range from approximately $1 - 1.5$ μm in length, with a diameter of about 3 nm. The separation distance between the nanotube and the gate electrode is about 150 nm. A highly resistive silicon dioxide substrate with prepatterend source drain and gate electrodes were used to grow nanotube using chemical vapour deposition technique. The nanotube has a narrow band-gap and its electrochemical potential is tunable by underlying gate electrodes. The quantum dot is formed by creating a p-n tunnel junction at the ends of the suspended nanotube by applying a DC gate voltage. Figure 2(c) shows a simple schematic of gate voltage dependence of the conductance in the incoherent tunneling regime. This clearly indicates how a degenerate two-level system fluctuates between two states having N and $N + 1$ electrons. The charging energy $E_c = e^2/C_\Sigma$ for the devices were about 8 meV, where e is electronic charge and C_Σ is the total capacitance of the quantum dot. A significant capacitive coupling was established between the nanotube island and the gate electrode by maintaining a minimal separation distance. This results in C_g being much greater than C_s and C_d, where C_s and C_d represent the capacitances between the nanotube island and the source and drain electrodes, respectively. The capacitance values were estimated by measuring the charge stability diagram of the quantum dot.

The resonator is coupled to the electrons within the dots through capacitive coupling between the nanotube and the gate electrode. This coupling can be described by the Hamiltonian $H = -\hbar g n_e x/x_{\rm zp}$, where g is the electromechanical coupling, and $n_e = 0, 1$ is the additional electron number in the quantum dot. Figure 2(b) describes the origin of nonlinearity in the mechanical resonator. The two black lines correspond to the linear force-displacement curves, when the dot is filled with either N or $N + 1$ electrons, whose slopes are governed by the spring constant $m{\omega_0}^2$. The separation between two curves, $\Delta x = 2(g/\omega_0)x_{\rm zp}$, is caused by the force generated by one electron tunnelling onto the quantum dot. When the fluctuations between the two electron states are rapid compared to the mechanical motion ($\Gamma_{\rm e} > \omega_0$), the force experienced by the vibrations is simply an average of the two black force-displacement lines weighted by the Fermi-Dirac distribution. The resulting force (shown in red) is nonlinear when the vibration displacement is smaller than $\frac{k_B T}{\hbar g} x_{\rm zp}$. The reduced slope at zero vibration displacement is an indication of a decrease in ω_m. Here, T is the temperature, and k_B is the Boltzmann constant. If we decrease the temperature, the slope at zero point displacement is reduced further as shown by dashed red line in Fig. 2(b). The electron fluctuation rate is faster than the bare mechanical frequency i.e. ($\Gamma_e > \omega_0$) in the adiabatic limit. In this

case, the fluctuations result in the nonlinear restoring force as given by

$$F_{\text{eff}} = -\left[m{\omega_0}^2 - \frac{1}{4x_{\text{zp}}^2}\frac{(\hbar g)^2}{k_B T}\right] x - \frac{1}{48x_{\text{zp}}^4}\frac{(\hbar g)^4}{(k_B T)^3}x^3 \tag{1}$$

for $\Gamma_e < k_B T$ and $x \ll 2k_B T/\hbar g$.

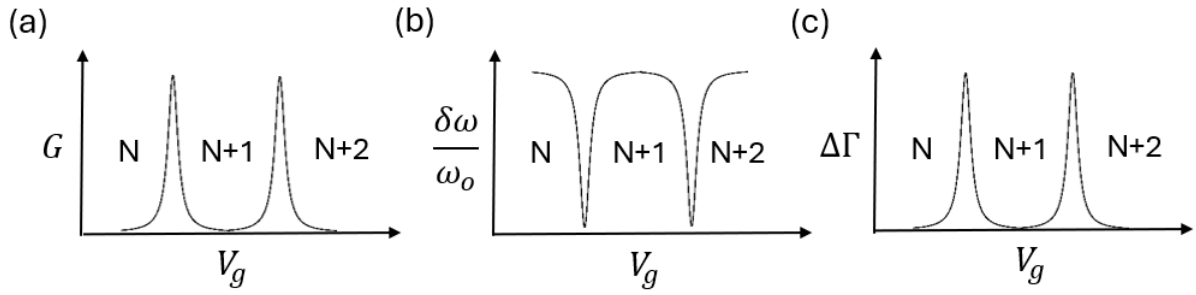

FIG. 3: (a) Schematic of conduction peaks with gate voltage in the SET regime. (b) and (c) represent schematics of change in resonance frequency and mechanical linewidth of the resonator across the conduction peaks, respectively.

Figure 3(b) illustrates the reduction of resonance frequency across the conduction peaks [shown in 3(a)] due to capacitive coupling between the mechanical vibration and the SET. Figure 3(c) shows a schematic of increase in linewidth Γ of the vibration across the conduction peaks [shown in 3(a)] due to incoherent tunneling of electrons between the dots and source/drain electrodes. A large number of experiments have been carried out [37–44], but the decrease in ω_0 has always been around 1% as the value of g was modest. The linear coefficient of the equation gives us

$$\omega_m = \omega_0\sqrt{1 - \frac{\varepsilon_p}{4k_B T}}, \tag{2}$$

where $\varepsilon_p = 2\hbar g^2/\omega_0$ is the polaronic energy. It is also clear that ω_m decreases with decreasing temperature. It was satisfactorily explaining the experimental observation. However, all the previous experiments could not be performed in a regime where $2k_B T \gg \hbar g^2/\omega_0$ due to lower values of g. This regime is referred as ultrastrong coupling regime, which can be realized in a mechanical system not in the ground state ($k_B T > \hbar\omega_0$) if $g > \sqrt{2}\omega_0$. It is also clear that ω_m vanishes when $2k_B T = \hbar g^2/\omega_0$ since the linear part of the restoring force vanishes as illustrated by the blue solid line in Fig. 4.

However, when we operate the device in the ultra-strong coupling regime, the nonlinear part dominates the dynamics of the resonator. The nonlinear force can even result in a vibration potential that is purely quartic in displacement. As a consequence of this nonlinearity, the oscillation period becomes highly dependent on the oscillation amplitude. Thermal fluctuations enable the oscillator to explore various amplitudes, resulting in different resonance frequencies. These fluctuations produce an observed resonance frequency that

is significantly higher than ω_m when averaged as shown in Eq. (2). The dashed red line in Fig. 4 depicts the theoretical prediction of the resonance frequency as a function of ϵ_p/k_BT incorporating both quartic nonlinearity and thermal fluctuations. This result was used to fit the temperature dependence of the resonance frequency at the conductance peak to estimate ϵ_p. The numerical result in Fig. 4 was fitted with the following analytical expression

$$\omega_m = \omega_0 \left[1 + \sum_{n=1}^{5} a_n \left(\frac{\epsilon_p}{k_B T} \right)^n \right] \tag{3}$$

with $a_1 = -0.127655$, $a_2 = 0.010475$, $a_3 = 0.0125029$, $a_4 = -0.00480876$, and $a_5 = 0.000515142$, which is within 0.1% of the numerical result for $0 \leq \frac{\epsilon_p}{k_B T} \leq 4$.

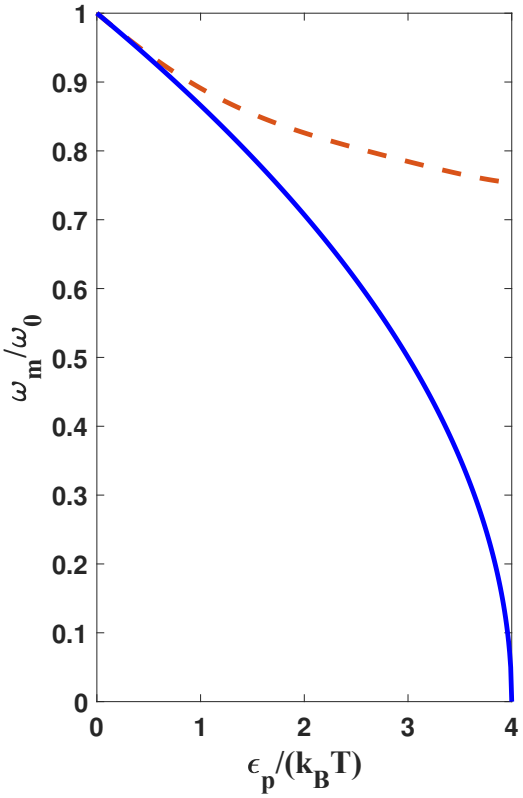

FIG. 4: The dashed red line illustrates the theoretical prediction for the renormalization of resonance frequency as a function of ϵ_p/k_BT, encompassing both quartic nonlinearity and thermal fluctuations. In contrast, the solid blue line represents the renormalization of resonance frequency excluding quartic nonlinearity and thermal fluctuations.

In contrast to earlier experiments, in Ref. [36], a substantial dip of approximately 25% in ω_0 was observed, along with an increase in the mechanical linewidth over 90% when the system was positioned on a conductance peak (Fig. 3), where the electronic two-level system reaches a state of degeneracy. Then, the authors repeated the measurements at different temperatures. The measured reduction in resonance

frequency as a function of temperature was then fitted using Eq. (3) to estimate the value of g. The largest value of the coupling strength g obtained from measurements was $g/2\pi = 0.5$ GHz. That corresponds to $g/\omega_0 = 17$. The value of g was consistent with the estimation $g/2\pi = 0.55$ GHz obtained from independent electron transport measurements. They used the formula $g = e(C'_g/C_\Sigma)V_g^{dc}/\sqrt{2m\hbar\omega_0}$, where m is estimated from driven spectral response measurements [45]. The spatial derivative of the dot-gate capacitance C'_g and the total capacitance C_Σ of the quantum dot were obtained from charge stability diagram of the device in the SET regime. These measurements established that the system was deep in the ultrastrong coupling regime.

III. FUTURE DEVICES AND EXPERIMENTAL PLAN

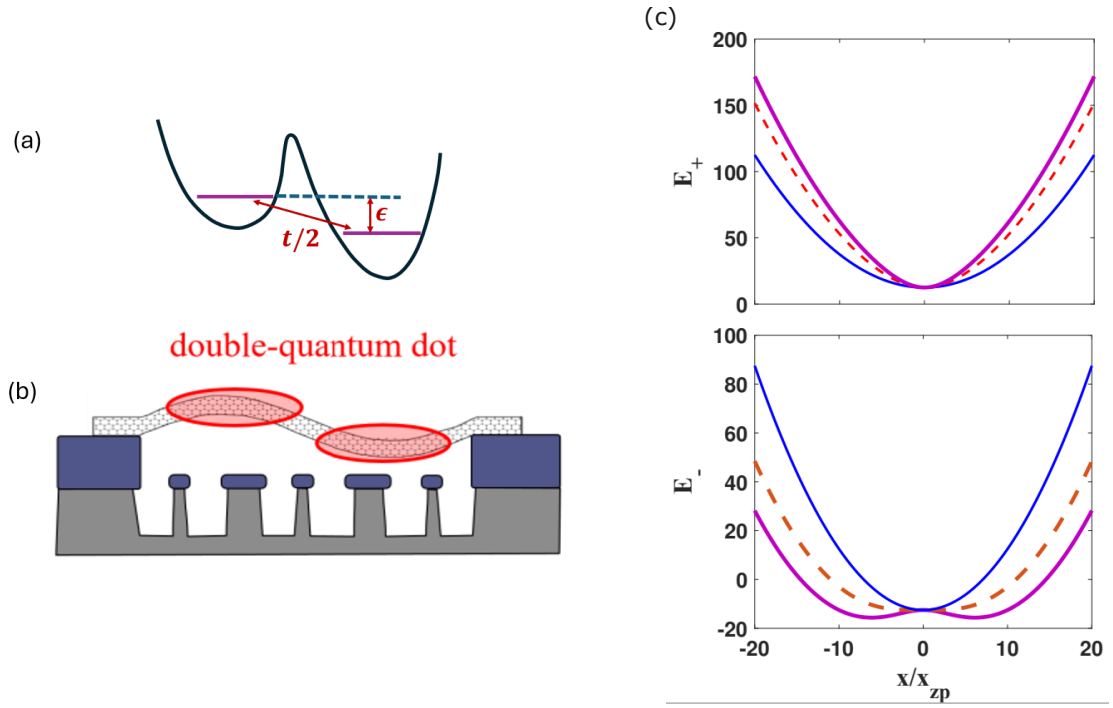

FIG. 5: (a) A schematic of the electronic confinement potential where t is the the hopping amplitude and ϵ is energy difference between the two individual charge states. (b) A schematic of a suspended carbon nanotube which hosts a double quantum dot. It's one-electron charged state is interconnected with the secondary flexural mode. (c) Upper and lower right panels show effective potential for E_+ and E_-, respectively using Eq. (6) for $t/\hbar\omega_0$=20. Here, E_+ and E_- are scaled by $\hbar\omega_0$. Thin solid, dashed and thick solid lines correspond to $4g/\omega_0 = 0, 100$, and 200, respectively.

The experimental validation of the ultrastrong coupling regime and strong anharmonicity represents a significant achievement, as it establishes a foundation for realizing nanomechanical qubits. However, at

low temperatures, the damping rate increases significantly due to the coupling between the resonator and the single quantum dot. The mechanical resonance width in the high temperature limit ($k_BT \gg \hbar g^2/\omega_0$) is given by [35]

$$\Delta\omega = \Gamma_0 + \frac{1}{2}\frac{\hbar g^2}{k_B T}\frac{\omega_0}{\Gamma_e}, \tag{4}$$

where Γ_0 is the damping rate due to other dissipation mechanisms. This charge fluctuation can be minimized by introducing two quantum dots in the resonators. It becomes feasible to define and control a charge qubit embedded in the resonator by adjusting the gate voltages independently for the two quantum dots. Pistolesi *et al* [35] theoretically proposed a possibility of a mechanical qubit by coupling the charge qubit with the second flexural mode in the strong coupling regime as shown in Fig. 5(b).

The Hamiltonian of the system can be described as

$$H = \frac{p^2}{2m} + \frac{m\omega_0^2 x^2}{2} + \frac{\epsilon}{2}\sigma_z + \frac{t}{2}\sigma_x - \hbar g\frac{x}{x_{\rm zp}}\sigma_z \tag{5}$$

where the first two terms correspond to the mechanical mode of frequency ω_0 with effective mass m, displacement x, and momentum p. The second and third terms corresponds to the electronic state of the quantum dots, where the two Pauli matrices σ_z and σ_x represent the dot charge energy splitting and inter-dot charge hopping, respectively. The final term corresponds to the coupling between the vibration and the charge state. This arises physically from the fluctuation of the force exerted on the mechanical mode when the charge transition occurs between the two quantum dots. The magnitude and the polarity of g can be adjusted across a wide range by adjusting the gate voltages. In the semiclassical Born Oppenheimer picture, by diagonalizing the above Hamiltonian, we find the energy eigenvalues as

$$E_\pm(x) = \frac{m\omega_0^2 x^2}{2} \pm \sqrt{\left(\epsilon - 2\hbar g\frac{x}{x_{\rm zp}}\right)^2 + \frac{t^2}{2}}. \tag{6}$$

Considering $\epsilon = 0$ and for small x we can rewrite equation 6 as

$$E_\pm(x) = \pm\frac{t}{2} + \frac{m\omega_0^2}{2}\left(1 \pm \frac{4\hbar g^2}{\omega_0 t}\right)x^2 \mp \frac{4m^2\omega_0^2\hbar^2 g^4}{t^3}x^4. \tag{7}$$

The interaction between the resonator and the double dot results in a renormilization of resonnace frequency and the emergence of quartic and higher-order terms. Figure 5(c) shows that this interaction stiffens the resonant frequency of the upper branch while softens that of the lower one. The quadratic term in Eq. (7) becomes negative if $g > \sqrt{\omega_0 t/4\hbar}$. This results in a double-well potential and bistability akin to what is anticipated for a single quantum dot coupled to a mechanical oscillator. Pistolesi *et al* [35] estimated the quantized energy levels of the system by diagonalizing Eq. (5) numerically. It was shown that for a large value of g, the anharmonicity $a = (\omega_{21} - \omega_{01})/\omega_{01}$ is enough for enabaling quantum control of the qubit

formed by the $|0\rangle$ and $|1\rangle$ state. The anharmonicity about 5% is typically used for a superconducting qubit for full quantum control, where the transition frequencies ω_{01} and ω_{21} are between $|0\rangle$ and $|1\rangle$ states and between $|1\rangle$ and $|2\rangle$, respectively. At very low temperature, the damping originates from the double quantum dots is suppressed exponentially. It is possible to achieve a very high quality factor of the resonator about a few millions. The decoherence of the hybrid system is improved by three oders of magnitude compared to its charge qubit counterpart. It is expected to have sub-kHz decoherence rate in this system. Additionally, mechanical qubits have novel prospects for quantum sensing. Rather than using traditional mechanical resonators to investigate AC forces, it is possible to detect weak DC forces with an exceptional sensitivity down to $\sim 10^{-21} N/\sqrt{Hz}$ [35]. The qubit state can be read out by coupling the system to a microwave superconducting cavity and utilizing a dispersive interaction, similar to the technique used with superconducting qubits.

IV. CONCLUSION

We discussed why a mechanical qubit is not possible to realize with conventional mechanical resonators though it has a remarkable success in preparation of quantum ground state, squeezing, entanglement etc. in a macroscopic objects. We explained the physical origin of the anharmonicity in a carbon nanotube resonator embedded with a quantum dot in detail. Then, we addressed the limitations of a single quantum dot hosted in nanotube the tube. Subsequently, we elaborated the methodology to overcome those limitations by hosting a double quantum dot in the nanotube as prescribed by Pistolesi *et al* [35]. Mechanical qubits may offer new perspectives for quantum sensing and quantum computing.

ACKNOWLEDGEMENTS

C. S. acknowledges the support from Indian Institute of Science Education and Research Bhopal (IISERB).

* Electronic address: csamanta@iiserb.ac.in

[1] M. H. Devoret and R. J. Schoelkopf, Science **339**, 1169 (2013).

[2] F. Arute *et al*, Nature **574**, 505 (2019).

[3] G. Burkard, T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta, Rev. Mod. Phys. **95**, 025003 (2023).

[4] J. Yoneda *et al*, Nature nanotechnology **13**, 102 (2018).

[5] N. Hendrickx, W. Lawrie, L. Petit, A. Sammak, G. Scappucci, and M. Veldhorst, Nature communications **11**, 3478 (2020).

[6] C. D. Bruzewicz, J. Chiaverini, R. McConnell, and J. M. Sage, Applied Physics Reviews **6**, 021314 (2019).

[7] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz, Nature Physics **18**, 1053 (2022).

[8] C. Fluhmann, T. L. Nguyen, M. Marinelli, V. Negnevitsky, K. Mehta, and J. Home, Nature **566**, 513 (2019).

[9] A. Stern and N. H. Lindner, Science **339**, 1179 (2013).

[10] V. Lahtinen and J. Pachos, SciPost Phys. **3**, 021 (2017).

[11] B. Yang, H. Sun, C.-J. Huang, H.-Y. Wang, Y. Deng, H.-N. Dai, Z.-S. Yuan, and J.-W. Pan, Science **369**, 550 (2020).

[12] H.-N. Dai, B. Yang, A. Reingruber, X.-F. Xu, X. Jiang, Y.-A. Chen, Z.-S. Yuan, and J.-W. Pan, Nature Physics **12**, 783 (2016).

[13] M. W. Doherty, N. B. Manson, P. Delaney, F. Jelezko, J. Wrachtrup, and L. C. Hollenberg, Physics Reports **528**, 1 (2013).

[14] F. Jelezko and J. Wrachtrup, physica status solidi (a) **203**, 3207 (2006).

[15] S. Pezzagna and J. Meijer, Applied Physics Reviews **8**, 011308 (2021).

[16] H. G. Craighead, Science **290**, 1532 (2000).

[17] K. Ekinci and M. Roukes, Rev. Sci. Instrum. **76**, 061101 (2005).

[18] M. Roukes, Physics World **14**, 25 (2001).

[19] A. D. O'Connell *et al*, Nature **464**, 697 (2010).

[20] J. D. Teufel, T. Donner, D. Li, J. W. Harlow, M. Allman, K. Cicak, A. J. Sirois, J. D. Whittaker, K. W. Lehnert, and R. W. Simmonds, Nature **475**, 359 (2011).

[21] J. M. Pirkkalainen, E. Damskagg, M. Brandt, F. Massel, and M. A. Sillanpaa, Phys Rev Lett **115**, 243601 (2015).

[22] M. Rossi, D. Mason, J. Chen, Y. Tsaturyan, and A. Schliesser, Nature **563**, 53 (2018).

[23] E. E. Wollman, C. U. Lei, A. J. Weinstein, J. Suh, A. Kronwald, F. Marquardt, A. A. Clerk, and K. C. Schwab, Science **349**, 952 (2015).

[24] J. Suh, A. J. Weinstein, C. U. Lei, E. E. Wollman, S. K. Steinke, P. Meystre, A. A. Clerk, and K. C. Schwab, Science **344**, 1262 (2014).

[25] C. F. Ockeloen-Korppi, E. Damskagg, J. M. Pirkkalainen, A. A. Clerk, M. J. Woolley, and M. A. Sillanpaa , Phys. Rev. Lett. **117**, 140401 (2016).

[26] S. Kotler *et al*, Science **372**, 622 (2021).

[27] E. A. Wollack, A. Y. Cleland, R. G. Gruenke, Z. Wang, P. Arrangoiz-Arriola, and A. H. Safavi-Naeini, Nature **604**, 463 (2022).

[28] R. W. Andrews, R. W. Peterson, T. P. Purdy, K. Cicak, R. W. Simmonds, C. A. Regal, and K. W. Lehnert, Nature physics **10**, 321 (2014).

[29] K. Fang, M. H. Matheny, X. Luan, and O. Painter, Nature Photonics **10**, 489 (2016).

[30] A. Reed *et al*, Nature Physics **13**, 1163 (2017).

[31] R. Lifshitz and M. C. Cross, Reviews of nonlinear dynamics and complexity, **Chapter 1** (2008).

[32] H. W. C. Postma, I. Kozinsky, A. Husain, and M. L. Roukes, Appl. Phys. Lett. **86**, 223105 (2005).

[33] S. Rips and M. J. Hartmann, Phys. Rev. Lett. **110**, 120503 (2013).

[34] S. Rips, I. Wilson-Rae, and M. Hartmann, Phys. Rev. A **89**, 013854 (2014).

[35] F. Pistolesi, A. N. Cleland, and A. Bachtold, Phys. Rev. X **11**, 031027 (2021).

[36] C. Samanta *et al*, Nature Physics **19**, 1340 (2023).

[37] R. G. Knobel and A. N. Cleland, Nature **424**, 291 (2003).

[38] B. Lassagne, Y. Tarakanov, J. Kinaret, D. Garcia-Sanchez, and A. Bachtold, Science **325**, 1107 (2009).

[39] G. A. Steele, A. K. Huttel, B. Witkamp, M. Poot, H. B. Meerwaldt, L. P. Kouwenhoven, and H. S. J. van der Zant, Science **325**, 1103 (2009).

[40] Y. Okazaki, I. Mahboob, K. Onomitsu, S. Sasaki, and H. Yamaguchi, Nature Communications **7**, 11132 (2016).

[41] I. Khivrich, A. A. Clerk, and S. Ilani, Nature Nanotechnology **14**, 161 (2019).

[42] C. Urgell, W. Yang, S. L. De Bonis, C. Samanta, M. J. Esplandiu, Q. Dong, Y. Jin, and A. Bachtold, Nature Physics **16**, 32 (2020).

[43] Y. Wen, N. Ares, F. J. Schupp, T. Pei, G. A. D. Briggs, and E. A. Laird, Nature Physics **16**, 75 (2020).

[44] S. Blien, P. Steger, N. Hittner, R. Graaf, and A. K. Huttel, Nature Communications **11**, 1636 (2020).

[45] C. Samanta, D. Czaplewski, S. De Bonis, C. Moller, R. Tormo-Queralt, C. Miller, Y. Jin, F. Pistolesi, and A. Bachtold, Appl. Phys. Lett. **123**, 203502 (2023).

J. Phys. Res. Edu., Vol. 2, March 2024

Prediction of Cross Port Output in Electro-optic Mach-Zehnder interferometer Using Artificial Neural Network

Dilip Kumar Gayen[1] and Tanay Chattopadhyay[2*]

[1]*Department of Computer Science and Engineering, College of Engineering and Management, Kolaghat, KTPP Township, Purba Medinipur, 721171, W.B, India.*

[2]*Solar division, Bidyut Unnayan Bhaban, WBPDCL, Kolkata, 700106, West Bengal, India.*

Artificial neural networks are essential for the design and analysis of complex algorithms. We have designed a circuit with an electro-optic Mach-Zehnder interferometer (EOMZI) driven by a photodiode. The output from the EOMZI's cross bar port is used. Then, we have developed a photonic artificial neuron based on this setup. The outputs are validated using MATLAB, and the Artificial neural networks successfully generate the sigmoid function corresponding to the EOMZI switch's cross port.

I. INTRODUCTION

Optical electro-optic modulators are critical devices in the field of photonics. They are widely used in optical communication systems. One common configuration of the modulator is the electro-optic Mach-Zehnder interferometer (EOMZI). EOMZIs have high modulation bandwidth, low power consumption, and high performance [1–5]. EOMZIs are used in a variety of photonic applications, including optical fiber communications, optical signal processing, and optical computing. The modulation bandwidth of an EOMZI is the range of frequencies over which it can modulate an optical signal. The power consumption of an EOMZI is the amount of electrical power required to operate it. The performance of an EOMZI is measured by its modulation depth, extinction ratio, and linearity. Artificial neural networks (ANNs) are a type of machine learning algorithm that can be used to solve a wide variety of problems, including those in optics. ANNs are inspired by the human brain, and they are able to learn from data by adjusting their connections over time. The use of ANNs in optics is a rapidly growing field, and there are many exciting new research directions being explored [6–10]. For example, ANNs are being used to develop new optical imaging techniques, to improve the performance of optical communication systems, and to create new optical computing platforms. This study explores the use of a photodetector (PD) sensor to convert light into an electrical signal. The amplified electrical signal is used to drive a Mach-Zehnder interferometer (MZI), which produces an optical output at the cross port. Cross port output is used to design optical NOT gates, 1's

compliment operation in optical computing. An equivalent model of these circuits is also developed using an artificial neural network (ANN). The ANN takes light intensities as inputs and generates optical signals as outputs. The conditions required for the framework to work are derived and validated using the ANN and MATLAB. The output of the ANN is verified to be in agreement with the sigmoid activation function.

II. SCHEMATIC REPRESENTATION OF THE PROPOSED CONFIGURATION

The block diagram of the proposed scheme is shown in Fig. 1(a). The proposed scheme consists of a photodetector (PD) sensor, an amplifier and a Mach-Zehnder interferometer (MZI). The PD sensor converts light into an electrical signal. The electrical signal is amplified and used to drive the MZI. The amplifier is used to amplify the signal from the PD. This is necessary because the signal from the PD is typically very weak. The amplifier increases the signal strength so that it can be used to drive the MZI. The MZI is a type of interferometer that can be used to modulate the phase of an optical signal. The MZI works by splitting the optical signal into two paths and then recombining them. This amplified voltage creates phase shift($\Delta\delta$) of the optical signal propagated through the two arms of MZI switch. When 0, destructive interference occurs at output coupler of MZI and data is transmitted to lower cross port. But when $\Delta\delta \sim 0$, then output at cross port is absent. The EOMZI figure is shown in Fig. 1(b). According to theoretical calculation cross port output can be expressed a $P_{out} = |\cos(\Delta\delta/2)|^2$ [11].

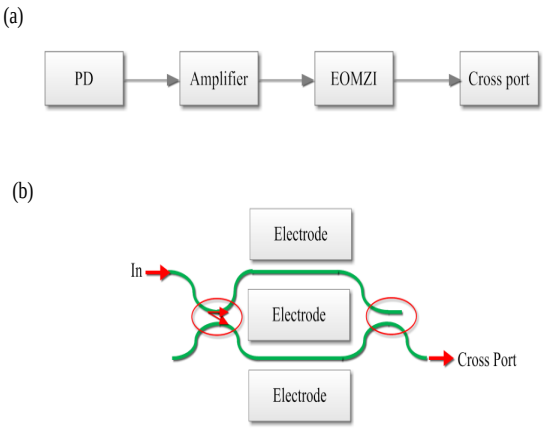

FIG. 1: (a) Block diagram of the proposed scheme, (b) EOMZI layout.

III. PREDICTION USING AN ARTIFICIAL NEURAL NETWORK MODEL

Artificial neural networks (ANNs) are used by many researchers to build output prediction systems [12–14]. ANNs require a large number of input and output data points to be created. The neural network design in Fig. 2 has three levels: the input layer, the hidden layers, and the output layer. Each neuron in the hidden layer has a weight assigned to it. In this architecture, the input layer comprises one neurons, denoted as $a_1^{(0)}$, which represent the light intensity of the system as measured by the photodiode. In our architecture, the hidden layer is composed of six neurons, namely $a_1^{(1)}$ to $a_6^{(1)}$. The weights and activations of these neurons play a crucial role in the neural network's overall performance and output predictions. The output layer consists of one neuron, labeled $a_1^{(2)}$, representing the output optical power at cross port of EOMZI. The data set has been divided into three parts: the test data set, the training data set, and the validation data set. The data set was partitioned into three segments, with 85% allocated for training, 10% for testing, and the remaining 5% for evaluating the network's overall performance.

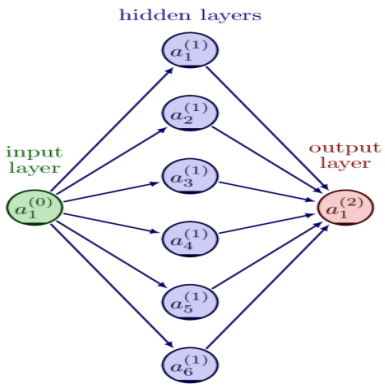

FIG. 2: ANN model layout.

Figure 3 illustrates the Mean Squared Error (MSE) variations for the training, validation, and testing data sets concerning the number of epochs at the output $a_1^{(2)}$. As depicted in the figure, the best validation performance is achieved at epoch 61, with an MSE of 8.034E-06. Beyond epoch 61, the errors begin to increase, leading us to terminate the training process to prevent over fitting.

Figure 4 shows the error values for the training, validation, testing, and overall data sets. The error factor, which is a measure of how close the neural network's output is to the desired output, is closer to 1 for all four data sets. This indicates that the neural network is able to fully associate the input data set with the model data set. The error values for the training, validation, testing, and overall data sets are 0.99999, 0.99997, 0.99998, and 0.99998, respectively. These results show that the designed neural network is performing well. After training the neural network, we tested it with different input values. Fig. 5(a) shows the distribution of

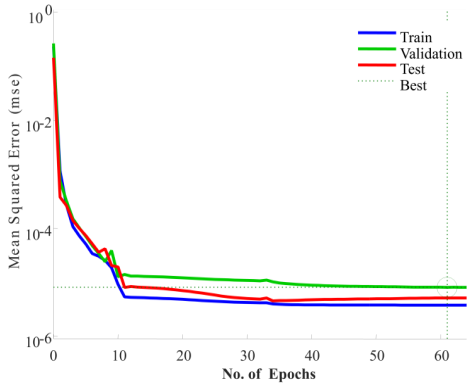

FIG. 3: Error with number of iterations (Epochs) at the output $a_1^{(2)}$.

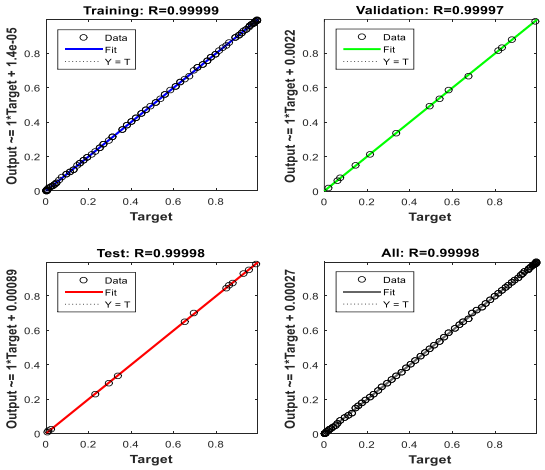

FIG. 4: Variation of error for various data sets at the output $a_1^{(2)}$.

training outputs and target outputs for different data sets: the training data set, the validation data set, and the test data set. Figure 5(b) shows the errors, which are the differences between the training output and target output, for different sets of inputs.

The results show that our constructed neural network accurately predicts the output values. This indicates that the neural network is robust and can generalize to new data sets. We can therefore be confident in using this neural network to predict output values for unknown data sets. The figure also shows that the sigmoid function is accurately modeled at the output terminal ($a_1^{(2)}$). We find the function as

$$f(x) = \left(\frac{e^{i0.022x} + e^{-i0.022x}}{2} \right)^2 . \tag{1}$$

We also plot the Cross port theoretical value with ANN model value in Fig. 6. This further validates the effectiveness of our approach.

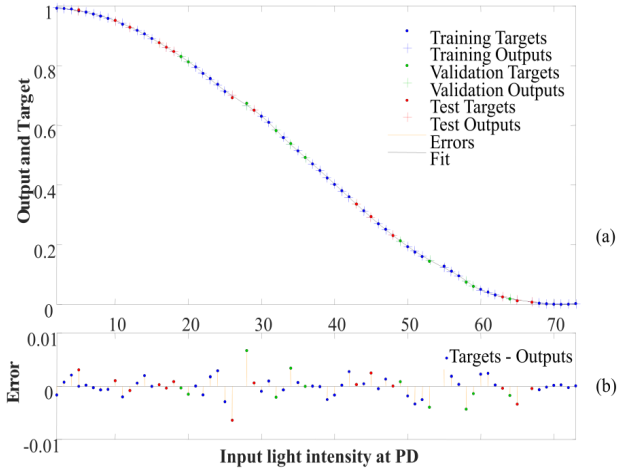

FIG. 5: (a) Displays the distribution of training outputs against input intensity of light at photo diode and target outputs for different data sets at the output $a_1^{(2)}$ and (b) the errors, calculated as the differences between the training outputs and target outputs, are depicted for various sets of inputs at the output $a_1^{(2)}$.

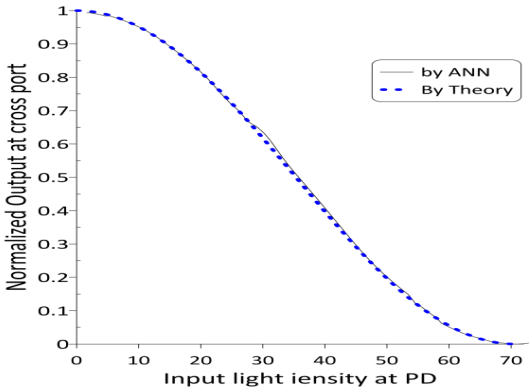

FIG. 6: Theoretical output and ANN plot of EOMZI cross output.

IV. CONCLUSION

This research effectively showcases the utilization of a photodetector (PD) sensor for converting light into an electrical signal, which in turn propels an EOMZI to generate optical outputs at the cross port. Through the implementation of Artificial Neural Networks (ANNs), precise predictions of optical outputs were achieved. The outcomes underscore the ANN's capability to accurately approximate the sigmoid function at the output terminal, reinforcing the credibility and appropriateness of the ANN model for this undertaking. Evaluation of the network's performance using testing data revealed an error factor approaching 1, signifying the successful association of input data with the model data set.

* Electronic address: tanay2222@rediffmail.com

[1] S. Awasthi *et al*, Journal of Computational Electronics **22**, 485 (2023).

[2] I. V. Kuznetsov and A. S. Perin, Journal of Optical Technology **90**, 93 (2023).

[3] H-S. Jung, Journal of the Optical Society of Korea **16**, 47 (2012).

[4] F. Valdez, V. Mere, X. Wang, and S. Mookherjea, Optics Express **31**, 5273 (2023).

[5] L. Lu, L. Zhou, X. Li, and J. Chen, Optics Letters **39**, 1633 (2014).

[6] T. Shan, J. Li, X. Hou, P. Huang, and X. Guo, 5th International Proceedings on Artificial Intelligence Circuits and Systems (AICAS), Hangzhou, China, **1-2**, (2023).

[7] T. Wang *et al*, Nature Photonics **17**, 408 (2023).

[8] Y. Du *et al*, IET Optoelectronics **17**, 1 (2023).

[9] P. Celard *et al*, Neural Computing and Applications **35**, 2291 (2023).

[10] S. R. Mohandes, X. Zhang, and A. Mahdiyar, Neurocomputing **340**, 55 (2019).

[11] C. Reis, T. Chattopadhyay, P. Andre, and A. Teixeira, Applied Optics, **51**, 8693 (2012).

[12] N. P. Juan, and V. N. Valdecantos, Ocean Engineering **259**, 111947 (2022).

[13] O. Guest, and A. E. Martin, Computational Brain & Behavior **6**, 213 (2023).

[14] D. Zhang et al., Artificial Intelligence Review **56**, 10651 (2023).

J. Phys. Res. Edu., Vol. 2, March 2024

Spin and valley dependent transport in a biased dice lattice

Lakpa Tamang*

Department of Physics, University of North Bengal, Raja Rammohunpur-734013, India

We study the spin and valley-dependent transport in a spin-orbit coupled biased dice lattice. We find that the presence of a bias term and the spin-orbit interaction (SOI) give rise to the spin-split energy spectrum. The SOI couples the valley and the spin degrees of freedom, resulting in a spin and valley-resolved Berry curvature. We find a profound variation in the Berry curvature for different spin states around both valleys. The spin and valley Hall conductivities are calculated for various values of the bias term. We find the interplay between the bias term and SOI term leads to a quantum phase transition from a topological insulating phase to a trivial band insulating phase accompanied by the emergence of the valley Hall effect and the suppression of the spin Hall effect.

I. INTRODUCTION

Since the discovery of graphene [1], there has been a growing interest in two-dimensional crystals like silicene [2], MoS_2 [3], etc., with a honeycomb lattice structure for potential applications in new-generation electronic devices [4, 5]. The low-energy electrons in such crystals have an extra degree of freedom called "valley". The valley index is associated with the special points of the Brillouin zone, called the Dirac points, K and K' where the energy bands meet each other. Since there is a significant distance in momentum space that separates the valleys, the intervalley scattering is highly suppressed in the presence of smooth scattering potential, thus making valley index an intrinsic property of low energy carriers. Similar to the application of spin in spintronics, the utilization of the valley index for encoding and controlling information leads to the emergence of a new field of research known as valleytronics [6–9]. The valley Hall effect stands out as an intriguing phenomenon resulting from the manipulation of the valley degree of freedom, where a longitudinal electric field drives carriers in opposite transverse directions based on their valley index. D. Xiao [6] first predicted it theoretically in graphene with broken space-inversion symmetry, and subsequently, it was observed experimentally in MoS_2 [10]. The breaking of inversion symmetry in crystals is necessary to manifest valley-contrasting physics. The SOI in such materials couples the valley index with the spin of the carriers, leading to various spin and valley-coupled phenomena. Over the years, various fascinating spin and valley-dependent phenomena like valley polarization by spin [11], spin-valley locking [12, 13], spin-valley interactions [14, 15], spin-valley polarization [16, 17], spin-valley dependent optical selection rules [18], etc, have been extensively studied.

In this paper, we calculate the spin and valley Hall conductivities to study the spin and valley coupled transport properties of a biased dice lattice. To study the effect of the SOI, we consider an intrinsic SOI of Kane-Mele type [19]. The consideration of such SOI usually results in a spin Hall effect [19–21], marked by the emergence of conducting channels along the boundaries of the sample. Our system of interest is a spin-orbit coupled biased dice lattice [22]. The bias term and the SOI break the space-inversion symmetry of the dice lattice. The Dice lattice is a graphene-like two-dimensional material with T_3 symmetry. Like graphene, electrons in the dice lattice obey the Dirac equation near the K and K' points. Although the zero-field spectrum of dice and graphene lattices appears identical, these systems differ fundamentally because the former supports a zero-energy flat band. In the dice lattice, low-energy excitations follow the Dirac-Weyl Hamiltonian with an enlarged pseudospin $S = 1$. Moreover, a significant disparity arises between the graphene and the dice lattice when look at the closed trajectories of the quasiparticles in momentum space. Unlike the Berry phase of π observed in graphene, the quasiparticles in the dice lattice do not acquire any nontrivial Berry phase. Figure 1 illustrates the geometry of the dice lattice, which includes an additional site positioned at the center of each hexagon in the honeycomb lattice connecting with one of the two inequivalent sites of the honeycomb lattice. Among these sites, two (A and C) are referred to as rim sites, while the third (B) is known as a hub site, having coordination numbers of 3 and 6, respectively. Experimental realization of the dice lattice is feasible in a $SrTiO_3/SrIrO_3/SrTiO_3$ heterostructure [23]. Additionally, it can be replicated in an optical lattice [24] setup by confining ultra cold atoms with the aid of three pairs of oppositely moving laser beams.

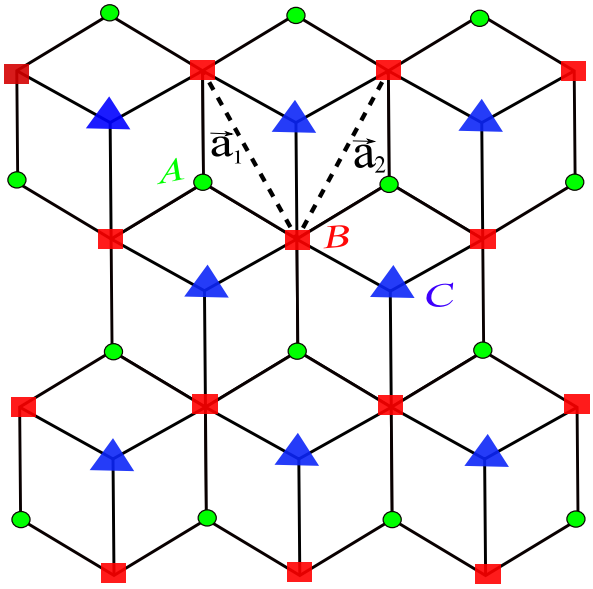

FIG. 1: The lattice geometry of the dice lattice is depicted here, where a_1 and a_2 represent the lattice vectors.

II. MODEL AND METHODS

Allowing the nearest-neighbor (NN) and next nearest-neighbor (NNN) hoppings only, the real space tight-binding Hamiltonian for biased dice lattice can be written as

$$H = t\sum_{\langle ij\rangle}\mathcal{C}_{is}^{\dagger}\mathcal{C}_{js} + t\sum_{\langle jk\rangle}\mathcal{C}_{js}^{\dagger}\mathcal{C}_{ks} + \frac{i\lambda}{3\sqrt{3}}\sum_{\langle\langle ij\rangle\rangle ss'}\mu_{ij}\mathcal{C}_{is}^{\dagger}s_z\mathcal{C}_{js'} + \frac{i\lambda}{3\sqrt{3}}\sum_{\langle\langle jk\rangle\rangle ss'}\mu_{jk}\mathcal{C}_{js}^{\dagger}s_z\mathcal{C}_{ks'} + \sum_{is}\nu_i\Delta_i\mathcal{C}_{is}^{\dagger}\mathcal{C}_{is} \quad (1)$$

where $\mathcal{C}_{is}^{\dagger}(\mathcal{C}_{is})$, $\mathcal{C}_{js}^{\dagger}(\mathcal{C}_{js})$, and $\mathcal{C}_{ks}^{\dagger}(\mathcal{C}_{ks})$ creates (annihilates) an electron of spin polarization s at sites A, B, and C, respectively. The first term in Eq. (1) illustrates the NN hopping between A and B sites, while the next term represents the NN hopping between B and C sites. In both cases the NN hopping strength is t. The third and fourth terms are the Kane-Mele type SOI term arising from the NNN hoppings A-B-A and C-B-C, respectively. Here, λ denotes the strength of the SOI. For clockwise (counter-clockwise) NNN hopping we have $\mu_{ij}, \mu_{jk} = -1\,(+1)$. Here, s, s', and s_z represent the real spin Pauli operators. The final term is the staggered lattice potential or the bias term Δ_i, where ν_i takes $+1$ and -1 for the A and C sites, respectively.

The Hamiltonian in Eq. (1) on projecting into a reciprocal space and expanding it around the Dirac points results in an effective Hamiltonian

$$H(\boldsymbol{k}) = \begin{pmatrix} \Delta - \lambda\tau s & \frac{f_{\boldsymbol{k}}}{\sqrt{2}} & 0 \\ \frac{f_{\boldsymbol{k}}^*}{\sqrt{2}} & 0 & \frac{f_{\boldsymbol{k}}}{\sqrt{2}} \\ 0 & \frac{f_{\boldsymbol{k}}^*}{\sqrt{2}} & -\Delta + \lambda\tau s \end{pmatrix}, \quad (2)$$

where $f_k = \hbar v_F(\tau k_x + ik_y)$ with the Fermi velocity $v_F = 3at/\sqrt{2}$ and $\tau = \pm 1$ is the valley index representing K and K' points. Diagonalizing the Hamiltonian given in Eq. (2), we find the following spin and valley-dependent eigenenergies

$$E_n^{\tau s} = n\sqrt{\epsilon_k^2 + (\Delta - \lambda\tau s)^2}. \quad (3)$$

Here, $n = -1/0/+1$ denotes the valence/flat/conduction band, respectively, and $\epsilon_k = \hbar v_F k$ with $k = \sqrt{(k_x^2 + k_y^2)}$. The presence of the bias term and the SOI term together introduces an energy gap $E_g = |\Delta - \lambda\tau s|$ between the flat band and the conduction band or the valence band at a particular Dirac point.

The corresponding normalized eigenstates are obtained as

$$u_0^{\tau s}(\boldsymbol{k}) = \frac{1}{\sqrt{2}N_0}\begin{pmatrix} \frac{-f_{\boldsymbol{k}}}{(\Delta-\lambda\tau s)} \\ \sqrt{2} \\ \frac{f_{\boldsymbol{k}}^*}{(\Delta-\lambda\tau s)} \end{pmatrix}, u_{\pm1}^{\tau s}(\boldsymbol{k}) = \frac{1}{\sqrt{2}N_{\pm1}}\begin{pmatrix} \frac{f_{\boldsymbol{k}}}{\left(E_{\pm1}^{\tau s}(k)-\Delta+\lambda\tau s\right)} \\ \sqrt{2} \\ \frac{f_{\boldsymbol{k}}^*}{\left(E_{\pm1}^{\tau s}(k)+\Delta-\lambda\tau s\right)} \end{pmatrix}, \tag{4}$$

where $N_0 = \frac{\sqrt{\epsilon_k^2+(\Delta-\lambda\tau s)^2)}}{|\Delta-\lambda\tau s|}$ and $N_{\pm1} = \frac{\sqrt{2}|E_{\pm1}^{\tau s}|}{\epsilon_k}$ are the normalization constants.

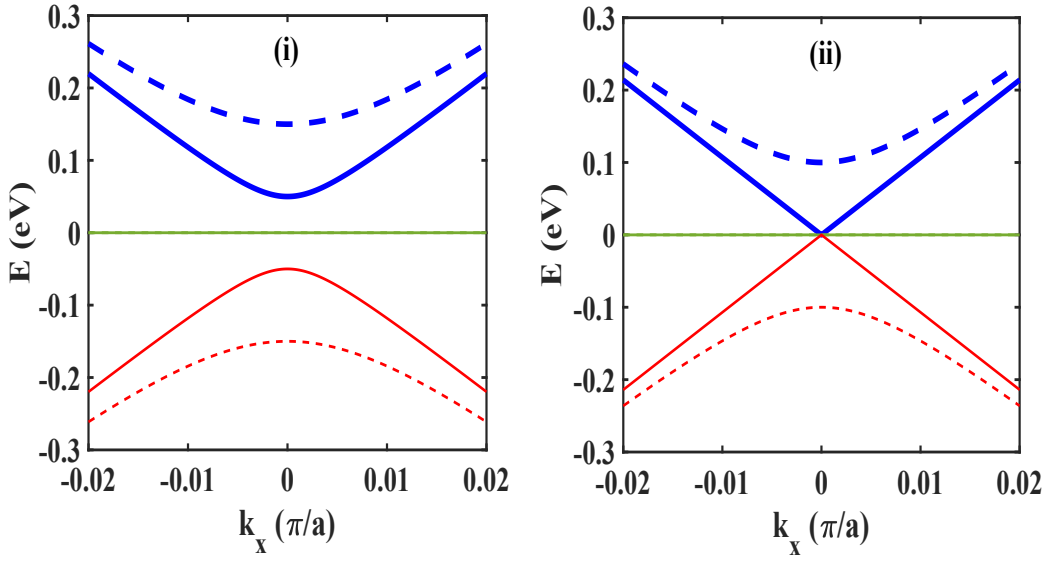

FIG. 2: The energy spectrum of the biased dice lattice for (i) $\Delta \neq \lambda$ and (ii) $\Delta = \lambda$. Thick solid (dashed) line represents the spin-up (spin-down) conduction band (blue in color), thin solid (dashed) line is for the spin-up (spin-down) valence band (red in color) and the Flat green line is the spin degenerate flat band. Here, we have taken $\lambda = 50$ meV.

The broken space-inversion symmetry due to the presence of the bias and the SOI results in a spin and valley-dependent Berry curvature $\Omega_n^{\tau s}(\boldsymbol{k})$. The Berry curvature can be calculated using the relation $\Omega_n^{\tau s}(\boldsymbol{k}) = i\langle\nabla_{\boldsymbol{k}} u_n^{\tau s}(\boldsymbol{k})| \times |\nabla_{\boldsymbol{k}} u_n^{\tau s}(\boldsymbol{k})\rangle$, where $u_n^{\tau s}(\boldsymbol{k})$ is the periodic component of the Bloch function corresponding to the n^{th} energy band. The Berry curvature transforms as $\boldsymbol{\Omega}_n^{\tau s}(\boldsymbol{k}) = \boldsymbol{\Omega}_n^{\tau' s}(\boldsymbol{k})$ and $\boldsymbol{\Omega}_n^{\tau s}(\boldsymbol{k}) = -\boldsymbol{\Omega}_n^{\tau' s'}(\boldsymbol{k})$ under the space-inversion and the time reversal symmetry operations, respectively. Here, $\tau' = -\tau$ and $s' = -s$. Therefore a non-zero Berry curvature demands the breaking of any one of these symmetries.

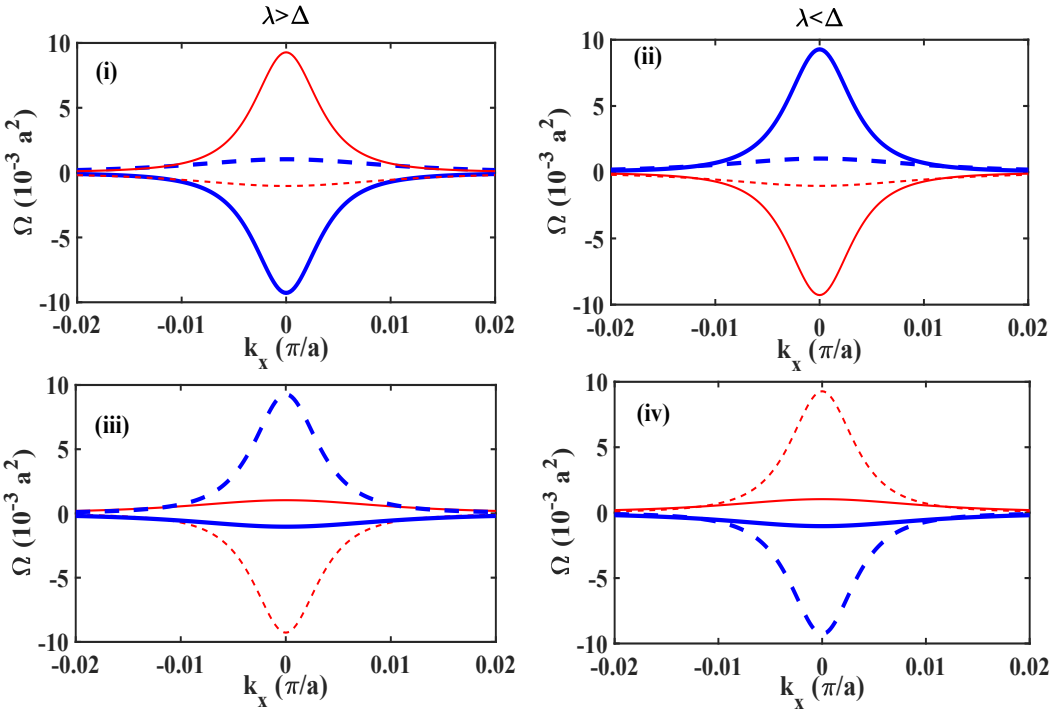

FIG. 3: The variation of Berry curvature around the K valley[(i) and (ii)] and the K' valley[(iii) and (iv)]. The left (right) panel corresponds to $\Delta < \lambda$ ($\Delta > \lambda$). Thick solid (dashed) line indicates the Berry curvature of the spin-up (spin-down) conduction band (blue in color) and thin solid (dashed) line is for the Berry curvature of the spin-up (spin-down) valence band (red in color).

In the presence of an in-plane electric field, $\Omega_n^{\tau s}(\boldsymbol{k})$ imparts an anomalous velocity to an electron perpendicular to its motion, resulting in an intrinsic contribution to the anomalous Hall conductivity [6]

$$\sigma_{xy}^{\tau s} = (e^2/\hbar) \int \frac{d^2\mathbf{k}}{(2\pi)^2} \sum_n f_n(E_n^{\tau s}) \Omega_n^{\tau s}(\boldsymbol{k}), \tag{5}$$

where $f_n(E_n^{\tau s})$ is the Fermi-Dirac distribution function. We define the spin ($\sigma_{xy}^{\text{spin}}$) and valley ($\sigma_{xy}^{\text{valley}}$) Hall conductivities as

$$\sigma_{xy}^{\text{spin}} = \sum_{\tau s} s\sigma_{xy}^{\tau s}, \tag{6}$$

$$\sigma_{xy}^{\text{valley}} = \sum_{\tau s} \tau\sigma_{xy}^{\tau s}. \tag{7}$$

III. RESULTS AND DISCUSSION

The energy spectrum of the biased dice lattice is depicted in Figs. 2 for K valley. There are, in principle, six energy bands, three associated with spin-up and three with spin-down electrons. Here, we treat λ as a fixed parameter and assume that Δ is a freely adjustable parameter. We have plotted the energy spectrum for $\lambda \neq \Delta$ and $\lambda = \Delta$ cases. In Fig. 2(i), when $\lambda \neq \Delta$ ($\lambda = 100$ meV, $\Delta = 50$ meV), we observe a splitting of the conduction and valence bands, resulting in a finite gap between both spin-up and spin-down bands. Conversely, when $\Delta = \lambda = 50$ meV, Fig. 2(ii) demonstrates a gapless Dirac cone for the spin-up electrons and a gapped Dirac cone for the spin-down electrons. This state has been termed a valley-spin-polarized metal (VSPM) [25]. In this case, electrons with only one spin polarization are significant for device applications facilitating spin-polarized transport. However, the flat band is nondispersive and spin degenerate in both cases. Because of the time-reversal symmetry, the spin splitting in the two valleys exhibits an opposite behavior. The role played by spin-up bands in the K valley is mirrored by spin-down bands in the K' valley. Using Eq. (4), we find the Berry curvature for the individual band analytically as

$$\Omega_{\pm}^{\tau s}(\boldsymbol{k}) = \mp\tau \frac{\hbar^2 v_F^2(\Delta - \lambda\tau s)}{[\epsilon_k^2 + (\Delta - \lambda\tau s)^2]^{3/2}} \text{ and } \Omega_0^{\tau s}(\boldsymbol{k}) = 0. \tag{8}$$

The Berry curvature associated with the conduction and valence bands are equal in magnitude but opposite in sign as a consequence of the particle-hole symmetry. However, Berry curvature of the flat band vanishes due to its nondispersive nature. Furthermore, with the inclusion of the bias and the SOI terms, the Berry curvature becomes spin and valley resolved. Figures 3(i) and 3(iii) depict the Berry curvature of conduction and valence bands for $\lambda > \Delta$ around K and K' valleys, respectively. The distribution of Berry curvatures is mainly centered around $\boldsymbol{k}$=0. In the K valley, for spin-up states, the Berry curvature of the conduction band is negative, and that of the valence band is positive. Conversely, its sign reverses with reduced magnitude for the spin-down states. However, in the K' valley, the Berry curvature follows the relation $\boldsymbol{\Omega}_n^{\tau s}(\boldsymbol{k}) = -\boldsymbol{\Omega}_n^{\tau' s'}(\boldsymbol{k})$ as a consequence of the time-reversal symmetry. The variation of the Berry curvature for $\lambda < \Delta$ around K and K' valleys is depicted in Figs. 3(ii) and 3(iv), respectively. In this case, the Berry curvature for spin-up states exhibits an opposite trend, while that for spin-down states remains consistent with the observations made when $\lambda > \Delta$. The pronounced variation observed in the Berry curvature for different spin states around both the valleys emphasizes the potential for manipulating spin and valley characteristics of the Berry curvature.

Using Eqs. (5) and (8), we calculate the analytical expression of Hall conductivity at zero temperature for various positions of the chemical potential (μ). For μ in the band gap, we obtain

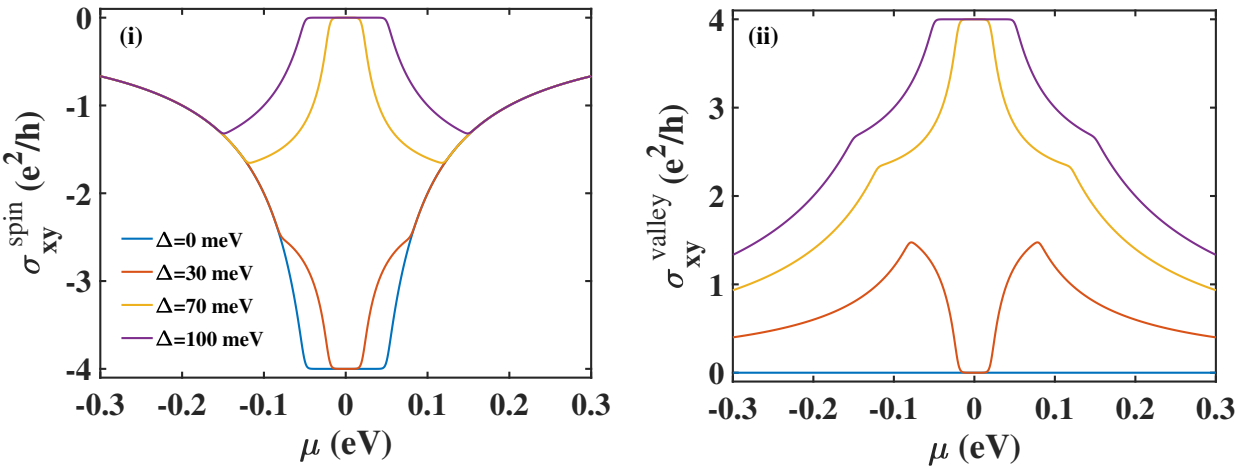

FIG. 4: The variation of (i) Spin and (ii) Valley Hall conductivities with the chemical potential (μ) for different values of Δ at $T = 20$K. Here we have considered $\lambda = 50$ meV.

$$\sigma_{xy}^{\tau s} = \tau \frac{e^2}{h} \frac{(\Delta - \lambda \tau s)}{|\Delta - \lambda \tau s|}. \tag{9}$$

Using Eqs. (6) and (7), we obtain the analytical expressions of $\sigma_{xy}^{\text{spin}}$ and $\sigma_{xy}^{\text{valley}}$ as

$$\sigma_{xy}^{\text{spin/valley}} = 2\frac{e^2}{h}\Big[\text{sgn}(\Delta - \lambda) \mp \text{sgn}(\Delta + \lambda)\Big], \tag{10}$$

where the $-(+)$ sign denotes the spin (valley) Hall conductivity. The interplay between Δ and λ leads to various phases, which can be characterized based on the values of $\sigma_{xy}^{\text{spin}}$ and $\sigma_{xy}^{\text{valley}}$ as

$$\sigma_{xy}^{\text{spin}} = \frac{e^2}{h} \begin{cases} -4 & \Delta < \lambda \\ -2 & \Delta = \lambda \\ 0 & \Delta > \lambda \end{cases}$$

and

$$\sigma_{xy}^{\text{valley}} = \frac{e^2}{h} \begin{cases} 0 & \Delta < \lambda \\ 2 & \Delta = \lambda \\ 4 & \Delta > \lambda. \end{cases}$$

When μ lies in the band gap, both $\sigma_{xy}^{\text{spin}}$ and $\sigma_{xy}^{\text{valley}}$ take a quantized values. For $\Delta < \lambda$, we have a finite $\sigma_{xy}^{\text{spin}}$ with vanishing $\sigma_{xy}^{\text{valley}}$, indicative of a topological insulator or Quantum spin Hall insulator phase [19],

with conductive channels at the boundary of the system. When $\Delta = \lambda$, both $\sigma_{xy}^{\text{spin}}$ and $\sigma_{xy}^{\text{valley}}$ simultaneously acquire non-zero values, indicating a VSPM phase [25]. A trivial insulator phase [6] emerges for $\Delta > \lambda$ with vanishing $\sigma_{xy}^{\text{spin}}$ and a finite $\sigma_{xy}^{\text{valley}}$.

When μ lies in the conduction band, we obtain

$$\sigma_{xy}^{\tau s} = \tau \frac{e^2}{h} \left[\frac{(\Delta - \lambda \tau s)}{E_F} \right], \tag{11}$$

where $E_F = \sqrt{\hbar^2 v_F^2 k_F^2 + (\Delta - \lambda \tau s)^2}$. The corresponding $\sigma_{xy}^{\text{spin}}$ and $\sigma_{xy}^{\text{valley}}$ are calculated as

$$\sigma_{xy}^{\text{spin/valley}} = 2 \frac{e^2}{h} \Big[\frac{(\Delta - \lambda)}{E_F} \mp \frac{(\Delta + \lambda)}{E_F} \Big]. \tag{12}$$

Here the $-(+)$ sign corresponds to the spin (valley) Hall conductivity. Due to the particle-hole symmetry, a similar expression can be obtained when μ lies in the valence band. The variation of $\sigma_{xy}^{\text{spin}}$ as a function of μ for different values of the parameter Δ is depicted in Fig. 4(i). Within a band gap, $\sigma_{xy}^{\text{spin}}$ takes a quantized and finite value as long as $\Delta < \lambda$ and vanishes for $\Delta > \lambda$, indicating a transition from a topological insulator to a trivial insulator phase. Additionally, it is observed that the width of the Hall plateau diminishes as Δ increases, suggesting a reduction in the width of a band gap. The behavior of $\sigma_{xy}^{\text{valley}}$ as a function of μ with increasing values of Δ is shown in Fig. 4(ii). For $\Delta = 0$, $\sigma_{xy}^{\text{valley}}$ vanishes for all values of μ, implying the necessity of a non-zero value of the bias term to break the valley degeneracy and observe the valley Hall effect. As Δ increases, for $\Delta < \lambda$, $\sigma_{xy}^{\text{valley}}$ remains zero within a band gap, while for $\Delta > \lambda$, it takes a quantized and finite value.

IV. CONCLUSION

In summary, we have investigated the spin and valley-dependent transport in the biased dice lattice. The presence of the bias and the SOI terms leads to a spin-split energy band structure. However, the flat band remains spin degenerate. We have obtained analytical expressions of the Berry curvature corresponding to individual bands. We analyzed the variations of the spin and valley Hall conductivities as a function of the chemical potential analytically and numerically. It is observed that the interplay between the bias and the SOI terms leads to a quantum phase transition from a topological insulator phase to a trivial insulator phase. This phase transition can be characterized by the appearance of the valley Hall effect and the suppression of the spin Hall effect, indicating a profound impact of the interplay of the bias and the SOI terms on the transport properties of the system. Our results could provide valuable insights for developing spintronics and valleytronics devices.

ACKNOWLEDGMENTS

The author sincerely acknowledges Dr. Tutul Biswas for his valuable suggestions.

* Electronic address: lakpatamang@nbu.ac.in

[1] K. S. Novoselov *et al*, Science **306**, 666 (2004).

[2] P. Vogt *et al*, Phys. Rev. Lett. **108**, 155501 (2012).

[3] D. Xiao, G-B. Liu, W. Feng, X. Xu, and W. Yao, Phys. Rev. Lett.**108**, 196802 (2012).

[4] K. S. Novoselov *et al*, Proc. Natl. Acad. Sci. **102**, 10 451 (2005).

[5] C. Lee, Q. Li, W. Kalb, X.-Z. Liu, H. Berger, R. W. Carpick, and J. Hone, Science **328**, 76 (2010).

[6] D. Xiao, W. Yao, Q. Niu, Phys. Rev. Lett. **99**, 236809 (2007).

[7] A. Rycerz, J. Tworzydlo, and C. W. J. Beenakker, Nat. Phys. **3**, 172 (2007).

[8] M. Ezawa, Phys. Rev. Lett. **110**, 026603 (2013).

[9] Z. P. Niu and S. Dong, Appl. Phys. Lett. **104**, 202401 (2014).

[10] K. F. Mak, K. L. McGill, J. Park, and P. L. McEuen, Science **344**, 1489 (2014).

[11] O. L. Sanchez, D. Ovchinnikov, S. Misra, A. Allain, and A. Kis, Nano Lett. **16**, 5792 (2016).

[12] L. Bawden *et al*, Nat Commun **7**, 11711 (2016).

[13] J.Y. Liu *et al*, Nat Commun **12**, 4062 (2021).

[14] R. Pisoni *et al*, Phys. Rev. Lett. **121**, 247701 (2018).

[15] N. Stefanidis and I. S. Villadiego, Phys. Rev. B **108**, 235137 (2023).

[16] J-T. Sun *et al*, 2D Mater. **3**, 035026, (2016).

[17] T. Zhou *et al*, npj Quant Mater **3**, 39 (2018).

[18] M. Ezawa, Phys. Rev. B **86**, 161407(R) (2012).

[19] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005).

[20] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).

[21] C-C. Liu, W. Feng, and Y. Yao, Phys. Rev. Lett. **107**, 076802 (2011).

[22] L. Hao, Phys. Rev. B **104**, 195155 (2021).

[23] F. Wang, Y. Ran, Phys.Rev. B. **84**, 241103 (2011).

[24] M. Rizzi, V. Cataudella, and R. Fazio, Phys. Rev. B **73**, 144511 (2006).

[25] M. Ezawa, Phys. Rev. Lett. **109**, 055502 (2012).

J. Phys. Res. Edu., Vol. 2, March 2024

Spectral variability in the orbital profiles of TeV Binary $LSI + 61^o$ 303 in X-Ray window using $XMM - Newton$ Observations

Tamal Sarkar*

HECRRC, University of North Bengal, Siliguri, WB 734 013 India

$LSI + 61^o$ 303 is a high-mass X-ray binary consisting of a low-mass $[M\ (1-4)M_\odot]$ compact object orbiting around an early type B0 Ve star along an eccentric $e = 0.7$ orbit. It along with LS5039 are the only two known gamma-ray binary detected in the TeV band. Despite extensive observations the nature of this source, particularly whether it is a pulsar or a black hole (micro-quasar) system, is not clear. The mechanisms that lead to the multi-wavelength behavior are also uncertain. Recent approach to the study of $LSI + 61^o$ 303 has been to focus on possible correlated variability. There was an early indication that there is a correlation between the X-ray and TeV emission at the time where the latter was measurable. In a more recent observation, however, such a feature has not been found. In this work, we study the spectral variability of $LSI + 61^o$ 303 using four observations between 2001 to 2007 of $XMM - Newton$ available from NASA archive and examine the correlation between low energy range and high energy part.

I. INTRODUCTION

The Be star binary $LSI + 61^o$ 303 is among the few X-ray binaries from which radio and very high-energy gamma-ray emission [1] in TeV energies were observed [2, 3]. Be stars are a heterogeneous collection of stars with B spectral types and emission lines. The radio source GT 0236 +601, is considered to be associated with this Be Star [4], the radio outbursts show a periodicity of about 26.496 days [4] and a further modulation of the outburst phase and outburst peak flux with a period of 1667 +/- 8 days [5]. This indicate that the position of the maximum of radio emission along the orbit with P_{orb} = 26.496 days [4], as well as its intensity, are modulated with a super-orbital period [6] of P_{sup} = 1667 days. Albert *et al* [7] found that the peak flux at TeV energies occurs at orbital phase $\phi_{orb} = 0.65$, while no high energy emission is detected around peri-astron passage $\phi_{orb} = 0.23$. Using VLBA imaging obtained by Dhawan *et al* (2006) over full orbit of LS +61 303 has shown the radio emission to come from angular scales smaller than about about 7 mas (projected size 14 AU at an assumed distance of 2 kpc). Rea *et al* suggest that X-ray emission of $LSI + 61^o$ 303 must necessarily come from the collision between winds or from the internal wind zone shock region instead of the magnetosphere of the supposed pulsation using Deep Chandra observations of TeV binaries [8]. Dhawan *et al* concluded that the radio and TeV emissions from $LSI + 61^o$ 303 are origi-

TABLE I: *Mode specific parameters for EPIC PN camera*

Obs. Mode	Time Resolution	Live Time [%]	MRC diffuse sec^{-1}	MRC point source sec^{-1}
Large window	$48ms$	94.9	$1500ms$	10
Small window	$6ms$	71.0	$12000ms$	100
Timing	$0.03ms$	99.5	N/A	800
Burst	$7\mu s$	3.0	N/A	60000

MRC= Maximum Count Rate

TABLE II: *Observation data considered for analysis with their* ϕ_{orb}

Observation ID	MJD	Exposure	window	ϕ_{orb}	Date
0112430401(X1)	52533	59999	PN Small	0.97	09/16/2002
0207260101(X2)	53397	49972	PN Small	0.60	01/27/2005
0505980901(X3)	54349	09860	PN Large	0.51	09/06/2007
0505981401(X4)	54354	12293	PN large	0.70	09/11/2007

MJD = Modified Julian Day

The orbital phases ϕ_{orb} are computed using $MJD = 43366.275$ and $P_{orb} = 26.496$

nated by the interaction of the wind of a young pulsar with that of the stellar companion [9]. While, Sarkar *et al* based on their studies suggested that particle acceleration sites can be different for events with and without flares, as in Microquasar model, in which the flaming radiation can come from hot spots located above the black hole using RXTE Mission Data. While stationary emissions are due to jets [10]. Kravtsov *et al* in 2020 [11] done studies on the orbital variability and confirm the existence of super-orbital variability using Optical measurements perform with the broad-band BVR polarimeter Dipol-2.

This Be star binary was observed with XMM EPIC PN camera in different modes Table I. Here, in this work, we have considered two modes prime small window as well as prime large window Fig. 1 using medium filter from 2001 to 2007, Some of which are taken as a part of the observational campaign of the system using XMM-Newton [12]. Here, we have considered four observation for our study as mentioned in Table II. Out of four two are in small window mode and two are in large window mode with medium filter. The first observation was made during from 02:23:47.000 hrs of 16th September 2002 to 04:11:28.000 hrs of 16th September 2002 for total 6.426×10^3 sec with the observation ID 0112430401, the second with observation ID 0202760301 was made during from 17:40:46.000 hrs of 27th Jan 2005 to 07:41:04.000 hrs of 27th Jan 2005 for total 50.418×10^3 sec, the third with observation ID 0505980901 was made during from 01:31:25.000 hrs to 05:15:02.000 hrs of 6th September 2007, the fourth with observation ID 0505981401 was made during from 01:27:25.000 hrs to 05:29:20.000 hrs of 11th September 2007.

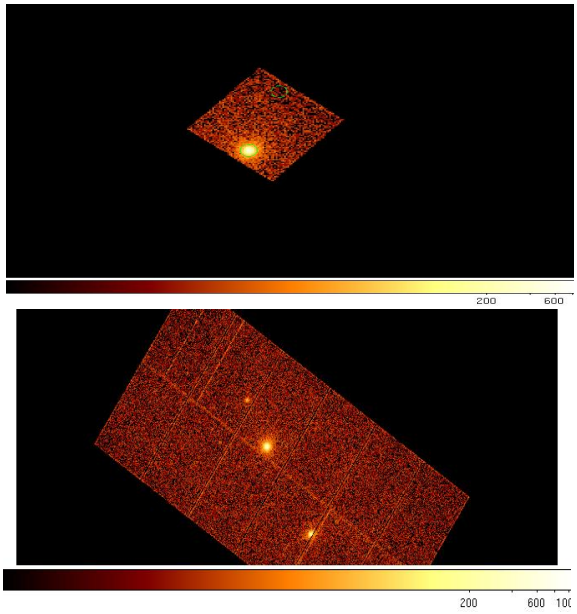

FIG. 1: Spectral Image of LSI $+61^{O}$ 303 in Small Window Mode [Top] and Large window mode [Bottom].

II. DATA REDUCTION, ANALYSIS & OBSERVATIONS

LS I +61 303 was observed with $XMM-Newton$ EPIC cameras for different time intervals in various windows mode with medium filters in between 2001 to 2007. Here, for study purpose, we have considered few suitable observations. Out of the four data sets, two data sets (0207260101, 0505980901) were having flaring regions at the beginning of the light curve. In this present work, we have processed the data giving special attentions to these regions. For data analysis, we have used XMM Science Analysis System. During data reduction, we have used the SAS's task "epchain" for extraction of EPIC PN camera data. The data was further filtered to include only science events and for patterns less than or equal to 4 in the energy range from 0.2-15 keV. Source spectrum was extracted from a circular region of radius 30 arc-sec around the source position using SAO Image Analysis Software ds9. Background PN spectra was extracted from nearby, source free regions having low value of counts, about $.1\%$ of the central region located at approximately the same readout position as that for the source circle.

A. Spectral analysis

To obtain the data for spectral analysis, the extracted regions for the source and background spectra as given above is required. Now, one can obtain the related response files using the task 'especget' or using OGIP Spectral Products of 'xmmselect' in Graphical User Interface (GUI). In GUI, the user can

TABLE III: *Best fit results of Spectral analysis with wabs *(po)*

ID	grouping of energy bin	Powerlaw Index	χ^2_{red}	$Flux^{**}$	$Lumin^{*}$ $10^{+33} ergs.sec^{-1}$
X1	gr 0 4095 5	$1.55^{+}_{-}0.02$	406.49/378	14.139	6.8437
X2	gr 0 4095 5	$1.67^{+}_{-}0.01$	438.32/378	08.622	4.1707
X3	gr 0 4095 5	$1.72^{+}_{-}0.02$	448.92/378	08.111	3.9234
X4	gr 0 4095 5	$1.67^{+}_{-}0.01$	400.57/378	08.302	4.0158
X1	gr min 35	$1.48^{+}_{-}0.02$	215.82/217	15.058	7.2838
X2	gr min 35	$1.66^{+}_{-}0.02$	855.90/812	08.588	4.1542
X3	gr min 35	$1.67^{+}_{-}0.02$	265.88/260	08.458	4.0914
X4	gr min 35	$1.64^{+}_{-}0.01$	410.22/387	08.455	4.0901

$N_H = 0.5 \times 10^{22}/cm^2$

* Luminosity is calculated for assume distance of 2 kpc.

** Flux measured in unit of $10^{-12} ergs.cm^{-2}.sec^{-1}$

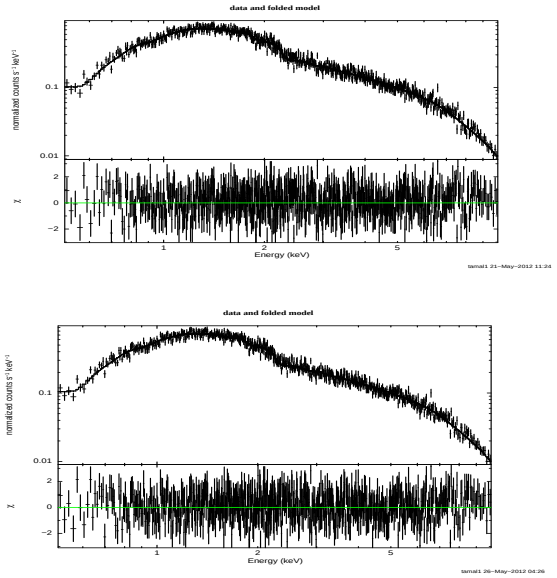

FIG. 2: Spectrum showing the normalized counts per second per keV Vs Energy in keV for model wabs*(powerlaw) for observation ID 0207260101 with Chi square for Flare region [TOP] and Flare free region[BOTTOM].

interactively define the source and background regions.

The GUI mode is provides more flexibility compare to command line in extraction of the spectra. Accordingly we have used the GUI. After using the above task four output files were obtained. They are source spectrum file; background spectrum file, ancillary response file and redistribution matrix file. The next job is to analysis the spectra which can be done using the standard package like xspec for X-ray spectral analysis.

TABLE IV: *Best fit results of Spectral analysis with wabs *(po) allowing N_H to vary*

ID	N_H	Powerlaw	χ^2_{red}	$Flux^*$	$Lumin^{**}$
	$10^{22}/cm^2$	Index			$10^{+33}ergs.sec^{-1}$
$X1^{@}$	0.598	$1.67^{+}_{-}0.02$	371.68/380	13.880	6.8437
$X2^{@}$	0.512	$1.69^{+}_{-}0.02$	431.62/377	08.587	4.1707
$X3^{@}$	0.500	$1.72^{+}_{-}0.02$	448.92/377	08.111	3.9234
$X4^{@}$	0.516	$1.69^{+}_{-}0.02$	398.22/377	08.272	4.0158
$X1^{@@}$	0.562	$1.55^{+}_{-}0.02$	204.35/217	14.872	7.2838
$X2^{@@}$	0.513	$1.67^{+}_{-}0.02$	853.84/809	08.564	4.1542
$X3^{@@}$	0.492	$1.65^{+}_{-}0.02$	263.82/259	08.477	4.0914
$X4^{@@}$	0.494	$1.63^{+}_{-}0.01$	410.15/384	08.456	4.0901

* Flux measured in unit of $10^{-12}ergs.cm^{-2}.sec^{-1}$

** Luminosity is calculated for assume distance of 2 kpc and $N_H = 0.5 \times 10^{22}/cm^2$.

$^{@}$ Obtained using gr 0 4095 5

$^{@@}$Obtained using gr min 35

Before going for such an analysis, the created spectrum was grouped (binned) depending on the available signal to noise ratio in the data and the science the user wants to perform. The FTOOLS task grppha allows regrouping of the source spectrum based on different criteria. In our case, we have divided the spectrum with flare into two parts i.e one with flare and other with no flare and done the analysis for an energy range of 0.5-10.0 keV Fig. 2. Initially, the spectrum was fitted with a powerlaw(po). Then, other models such as the black body, the gauss(ga) and the multicolor disc black body (diskbb) were tried. It is found that the model multiplicative absorption model "wabs" was used along-with a powerlaw wabs*(po) works well for the source for all data set to fit the spectrum. The best fit parameters along with χ^2_{red} are given in Table III and Table IV.

B. Timing analysis

Data for timing analysis were reduced using the XMM Newton Scientific Analysis Software package. The EPIC pn background subtracted light curves in two energy ranges (0.5 to 2.0 keV and 2.0 to 10.0 keV) are reported and the Light curves are plotted for all the data sets along with Hardness ratio and cross correlation as shown in Fig. 3 and Fig. 4, respectively.

The observed X-Ray datas also shows the presence of flaring events which was reported in past also one such report is reported by Smith *et al* [13].

In the Fig. 4, the cross correlation for flaring region and flare free regions are shown and it indicates a

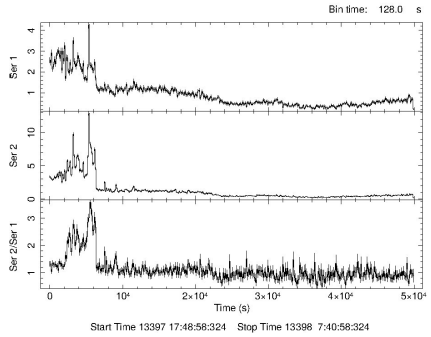

FIG. 3: $LSI + 61^o$ 303 light curve and hardness ratio for observation ID 0207260101.

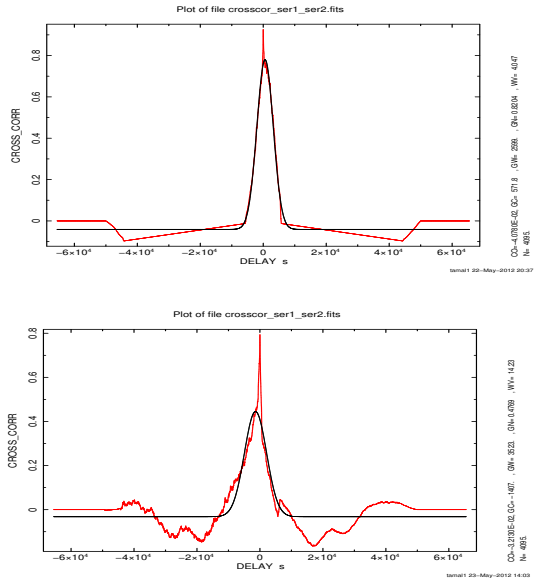

FIG. 4: $LSI+61^o$ 303 cross-correlation for observation ID 0207260101 flare [TOP]and flare free regions[BOTTOM].

clear difference in pattern and same has been discuss in the next section.

III. DISCUSSION

We found that there is no appreciable change in the average flux for a particular orbital phase band.The spectra for different observation period exhibit power law (without break) feature along with observation. The power index varies with orbital phase; the spectrum is harder at brighter stage. However, the difference in spectral index between low and high emission phase decreases if the absorbing material is considered as variable rather than fixed as done in most of the analysis. We notice a flaring feature at the beginning of two observations . Such a feature was previously noticed in one of the BeppoSAX observation [12]. The flaring and increase of flux in observation ID X1 with $\Phi_{orb} = 0.97$ may be explained using two-accretion

peak model with Inverse Compton [14] , as per which the accretion rate might develop two peaks along the eccentric orbit . It may be noted that $LSI + 61\ 303$ is a X-Ray binary with an eccentric orbit $e = 0.7$. As per two-accretion peak model, the rate of accretion i.e $\frac{dM}{dt} \propto \rho_w / v_{rel}^3$, where ρ_w = density of the straitified wind of the B0 Ve star and v_{rel} = Relative velocity between the wind and the orbiting accretor. The accretion rate may peak depending upon ρ_w and v_{rel} i.e when the density is highest or v_{rel} is lowest.

IV. SUMMARY

The X-Ray flux (0.5-10 KeV) is found to be highest and spectral Power-law Index (Γ) is lowest for $\Phi_{orb} = 0.97$ with the two grouping scheme as shown in Table III and Table IV, respectively. It is found that the best fit results of this spectral analysis with wabs*(po) allowing N_H to vary (see Table IV). With grouping scheme "$gr0\ 4095\ 5$", it is observed that there is clear variation of Γ , flux and N_H with orbital phase ϕ_{orb}. The Γ has highest value whereas N_H and flux has lowest at $\phi_{orb} = 0.5$ for observation ID X3. However, if we considered the grouping scheme "gr min 35", there is no such variation of Γ , but, clear variation of N_H is seen. Under these grouping scheme, the variation or decay in $N_H \approx 0.562\text{-}0.492 = .07 \times 10^{22} cm^{-2}$.Taking the mean density of the disk to be $n_0 \approx 10^9 cm^{-3}$ [6], we get the size of the clumps $s = N_H / n_0 \approx 7 \times 10^{11} cm$ which is comparable with the radius of the Be Star if we assume the the presences of clumps of highly or partially ionised hydrogen. With this simple model, the variation of N_H with ϕ_{orb} as observed can be explained on the basis of presence of clumps. In the present work, we have mentioned about the presence of flaring in the Observation ID X1 and X3. Fig. 3 indicates one such event. During such flaring behavior, the hardness ratio is found to increase as in Fig. 3, implying short bursts of high energy radiation as shown in Fig. 4. It indicates that lagging of soft X-Rays during flaring and lagging of hard X-Rays in the flare free event which may be due to the different sites of particle acceleration for flaring and flare free events.

* Electronic address: ta.sa.nbu@hotmail.com

[1] G. Dubus, A&A, **456**, 801 (2006).
[2] M. Chernyakova *et al*, MNRAS, **372**, 1585 (2006).
[3] D. B. Kieda, VERITAS Collaboration, 36th ICRC2019.
[4] P. C. Gregory and A. R. Taylor, Nature, **272**, 704 (1978).
[5] P. C. Gregory, ApJ, **575**, 427 (2002).
[6] M. Chernyakova *et al*, MNRAS, **470**, 1718 (2017).

[7] Albert, J. *et al*, ApJ, **674**, 1037 (2008).

[8] N. Rea, *et al*, MNRAS, **405**, 2206 (2010).

[9] V. Dhawan, A. Mioduszewski, and M. Rupen, VI Microquasar Workshop: Microquasars and Beyond, 52.1 (2006).

[10] T. Sarkar, *et al*, RAA, **16**, 104 (2016).

[11] V. Kravstov *et al*, A&A, **643**, A170 (2020).

[12] Sidoli, L.Pelliizoni, *et al*, A&A **459**, 901 (2006).

[13] A. Smith *et al*, ApJ, **693**, 1621 (2009).

[14] M. Massi *et al*, MNRAS, **498**, 3592 (2020).

J. Phys. Res. Edu., Vol. 2, March 2024

Na_2O Doped Solid State Battery Electrolytes: Preparation and Electrical Characterization

Rakesh Ram and Sanjib Bhattacharya*

Composite Materials Research Laboratory, UGC-HRDC (Physics),

University of North Bengal, District: Darjeeling-734013, West Bengal India.

Na_2O doped glassy electrolytes have been developed to explore their electrical transport behaviour at ambient temperature. The electrical conductivity (mixed conduction) has been studied in the light of both frequency as well as temperature. In the low frequency range, the diffusional motion of Na^+ ions is expected to be the possible reason for a flat conductivity. It is also noted that the conductivity shows dispersion near the high frequency regime. The DC conductivity (σ_{dc}) and hopping frequency have been extracted from the best fitted plots of experimental data. The thermally active nature of electrical conductivity data have been studied to establish the dynamics of charge carriers via hopping conduction in sodium oxide glassy systems. Negligible small difference in pathways in the I-V characteristics in both the directions should make a sense for their applications for new generation battery-electrolyte.

I. INTRODUCTION

In the recent days, the availability of lithium in the earth's crust is decreasing, which is leading the researches for alternative ways [1–3]. One of the possible best candidates for this purpose is sodium ion batteries (NIBs) [1–3]. Owing to the abundant storage and low cost preparation techniques, sodium ion batteries serve best to replace lithium ion batteries [4–6]. Here, Na_2O doped glassy electrolytes have been developed which serves as the electrolyte for NIBs in terms of exhibiting certain important features such as in electrochemical stability, sensors etc. [4–6]. Na_2O doped glassy electrolytes in solid form, are preferred over liquid electrolyte due to their higher stability and ease of access, which may lead us to develop Na_2O doped solid electrolytes with higher conductivity at ambient temperatures [6]. Study of various characteristics of such systems has been done such as AC conductivity [7, 8]. Estimation of hopping frequency and conduction pathways of charge carriers at different temperatures have been performed using Almond-West formalism [7, 8] which is given by

$$\sigma(\omega) = \sigma_{dc}\left[1 + \left(\frac{\omega}{\omega_H}\right)^n\right], \tag{1}$$

where, σ_{dc} is the low frequency conductivity, ω-H is the hopping frequency of the charge carriers, n is the power-law exponent and ω is the frequency of applied electric field. So, the present work has been executed

for the benefit of the research community not only for application point of view but also for academic interest in line with the following objectives: (i) Easy development of Na_2O doped solid electrolytes at low cost (ii) To improve the electrical conductivity of the present system at various temperatures. (iii) Establishment of relation between the ion migration process and the dielectric properties in the system. (iv) Conduction mechanisms were studied in comparison with other works. (v) Hopping frequency along with other parameters were studied and analyzed to check the efficiency of as prepared electrolytes to be used for NIBs.

II. EXPERIMENTAL PROCEDURE

Glassy system, $x$$Na_2O$-$(1-x)$(0.2ZnO-0.8CdO) for x = 0.1 and 0.2 have been developed using melt quenching route. The powder of Na_2CO_3 was heated at 500 ℃ in a furnace to obtain Na_2O. The powders of Na_2O, ZnO, and CdO were properly mixed in suitable stoichiometric ratio in an alumina crucible after measuring them in required amount by an electro-mechanical analytical weight balance (Dhona, 200D). The mixture was placed in an electric furnace at temperature 1100 ℃ to 1150 ℃ and was kept for half an hour. After the heat-treatment, the powdered mixture was used to make pallets of thicknesses 1 - 2 mm using a pelletizer. Graphite paste was used to paint both sides of the samples for making electrodes of the sample to conduct electrical and dielectric measurements. Usually, silver paste or gold coat was used as electrode on both sides of the as-prepared samples. But, in the present sodium glassy electrolytes, this conductive layer may develop solid electrolyte interface of blocking current collector, which may produce large polarization effect. To avoid this issue, graphite paste was used in order to produce defective non-blocking carbon current collector with multivalency defects induced homogeneous solid electrolyte interface. This type of conductive graphite paste was expected to reduce charge polarization effect at the interface. A dedicated LCR meter (HIOKI, model no. 3532-50) with good precision was employed to collect electrical and dielectric measurements data [C (capacitance), G (conductance) and tan(dielectric loss tangent)] at numerous temperature at different frequencies ranging from 42 Hz to 5 MHz. Also, a Keithley made electrometer (model No. 6514/E) was employed to measure DC conductivity and current-voltage data for the present samples. Measurements of I-V characteristics of the present samples were performed on a Keithley made electrometer (model No. 6514/E) at room temperature. Electrometer (model No. 6514/E) was used to apply the voltage across the sample and to measure the current through the sample as it had an in- built capability of output independent voltage source of 50 V. The indigenously designed sample holder with copper electrodes (two probe method) was used in the circuitry for bulk pellets were used with graphite paste contact in circuitry.

III. RESULTS AND DISCUSSIONS

A. DC Conductivity

The DC conductivity of the sample was plotted with different concentration of the sample along with the activation energy. These data were analyzed using Arrhenius [9] nature of DC conductivity as

$$\sigma_{dc} = \sigma_0 \, e^{-E_\sigma / k_B T}, \tag{2}$$

where σ_0 is the pre-exponential factor of the DC conductivity, E_σ is the activation energy corresponding to DC conductivity, k_B is the Boltzmann constant, and T is the temperature. The DC conductivity at a specific temperature (393 K) and the corresponding activation energy are illustrated based on composition in Figs. 1(a) and 1(b), respectively. Analysis of Fig. 1(a) reveals a gradual decrease. This behavior is attributed to the presence of various defects [10] in the system, as indicated in the literature. The existence of defects is likely a result of non-bridging oxygen [11, 12], which plays a role in their development.

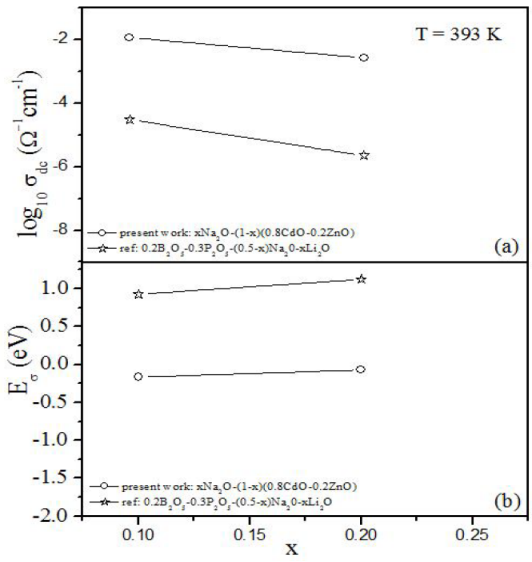

FIG. 1: (a) Comparison of DC conductivity with compositions with others' work. (b) Variation of activation energy corresponding to DC conductivity with compositions along with others work.

According to existing literature [1, 11, 12], bridging oxygen in the network structure is expected to participate in covalent bonding, playing a crucial role in forming Na-O bonding [1, 11, 12]. Consequently, bridging oxygen emerges as a key factor in the ionic DC conductivity process by facilitating the movement of Na^+ ions within the glassy matrix. Conversely, some oxygen ions do not engage in the network structure [1, 11, 12], and it is proposed that these oxygens contribute to non-bridging configurations. These oxygen atoms are believed to give rise to non-bridging oxygen through permanently broken oxygen bonds [1, 11,

12], as discussed in the literature. The oxygen atoms that are not bridging are assumed to be the trapping centers or the defects that are not localized which are formed by modifiers like Na_2O. The presence of such defects is thought to hinder the movement of Na^+ ions, consequently leading to an increase in the activation energy associated with DC conductivity [1, 11, 12].

The glassy system doped with Li_2O and Na_2O, as reported in previous work, exhibits lower DC conductivity up to $x = 0.2$ compared to the current system. Other studies [13] incorporating B_2O_3 and P_2O_5 as formers have demonstrated the stability of their systems. In contrast, our study employs a network former comprising CdO and ZnO. The presence of CdO is anticipated to introduce various defects in the glassy matrices [14], while ZnO serves as a stabilizer [15]. The nature of the DC conductivity in our system suggests that adjusting the CdO and ZnO ratio can alter its electrical properties [16]. Moreover, the activation energy (E_σ) associated with DC conductivity in our system is lower than that reported in other studies. This lower E_σ value indicates the absence of specific defect states related to grain boundaries, which typically result in a higher potential barrier [16]. In our context, "particular defect states" refer to defects that may arise between two charged states (D^+ and D^-) due to a dangling-bond-like structure [1]. The absence of these defect states in the higher potential barrier region, as evidenced by smaller E_σ values, suggests the exemption of correlated motion of polaron hopping (bi-polaron) via Coulombic interaction [1] in the present system. The lack of grain boundaries is further supported by the absence of additional arcs in the semicircular isotherms depicted in Fig. 1(a), confirming a much higher DC conductivity in our system compared to the Na_2O-Li_2O system. This finding underscores the potential suitability of our system as an electrolyte for upcoming materials for the batteries.

B. Study of Alternating Current Conductivity

At different temperatures the conductivity spectra has been scrutinized, and a particular spectrum for $x = 0.2$ is presented in Fig. 2(a) across. Examination of Fig. 2(a) reveals that AC conductivity spectra exhibit both thermal activation and frequency dependence [17, 18]. In the lower frequency range, the spectra display flat isotherms, akin to the conduction related to DC part. Alteration in electrical conduction patterns initiates at the hopping or leap frequency of charge-carriers [17], and after this frequency, distribution becomes evident. It was anticipated as the movement of charge-carriers was interrelated as well as confined within the high frequency range [19]. The Na_2O-doped system demonstrates a higher conductivity level of approximately $10^{-2}\ \Omega^{-1}\ \mathrm{cm}^{-1}$, suggesting its potential as a promising electrolyte for next-generation battery materials. The thermal activation and frequency dependence observed in the AC conductivity spectra further emphasize the dynamic and versatile nature of the electrical behavior in the studied system.

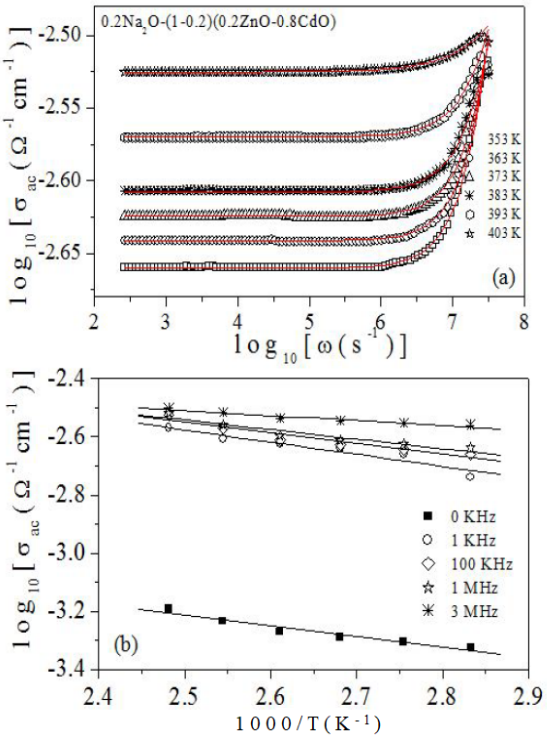

FIG. 2: (a) AC conductivity spectra for $x = 0.2$ at different temperatures. (b) Reciprocal temperature dependence of AC conductivity at different fixed frequencies for a particular composition.

The reciprocal temperature dependence of AC conductivity has been examined at various frequencies to gain insights into the thermally activated nature of the present system at a specific frequency [18]. Figure 2(b) illustrates one such plot for $x = 0.2$ across different frequencies. The plot indicates a broad dispersion of AC conductivity near the low-temperature regime. Interestingly, with increasing temperature, all the conductivity spectra converge and collapse. It is worth noting that as the frequency rises, there is an accompanying increase in AC conductivity. This behavior underscores the dynamic relationship between temperature, frequency, and conductivity in the studied system, revealing its intricate thermally activated characteristics.

The data obtained in Fig. 2(a) were meticulously examined by Eq. (1), in order to acquire comprehensive insights for the conduction spectra which rely on frequency. A rigorous fitting procedure has been employed to ensure the optimal alignment between the experimental data and the theoretical model [19], as defined by Eq. (1). This approach aims to establish the maximum dependency between the experimental observations and the theoretical framework encapsulated in Eq. (1). The fitting process is undertaken with the goal of attaining a robust and accurate representation of the underlying relationship between the variables involved. The fitting process has yielded essential parameters, including the hopping frequency (ω-H), DC

conductivity (σ_{dc}), and frequency exponent (n). The estimation and interpretation of these fitting parameters contribute to a deeper understanding of the underlying mechanisms governing the conductivity in the studied system, facilitating meaningful comparisons and predictions. The values of E_H (the activation energy corresponding to the hopping frequency) and ω_H estimated at 393 K are plotted in Fig. 3(a). The plot reveals a similar trend to that observed in DC conductivity.

The power law model [17], as outlined by Sidebottom, offers the parameter n, which is linked to the dimensionality of conduction pathways [20]. According to Sidebottom's proposition, a value of n around 0.67 signifies the silver (Ag^+) ions tridimensional movement in $MoIO_4$ (iodo-molybdate) glassy systems [21], showcasing independence of temperatures as well as concentration. A more in-depth investigation indicates that an average value of n around 2.72 can be regarded as the fractal dimensionality of quasi-crystals [22].

In Fig. 3(b), the average values of n associated with different compositions are presented. This representation allows for a visual exploration of the dimensionality characteristics across the studied compositions. The observed trends in n provide valuable insights into the conduction pathways and the complex nature of charge carrier motion within the system. Figure 3(b) shows that the value of n can be sometimes higher than 1 (for $x = 0.2$) and sometimes lower than 1 (for $x = 0.1$). With a lower concentration of Na_2O the Na^+ ion motion is dominated by electron or polaron hopping. As the concentration of Na_2O increases the conduction is again dominated by Na^+ ion motion over electron or polaron hopping.

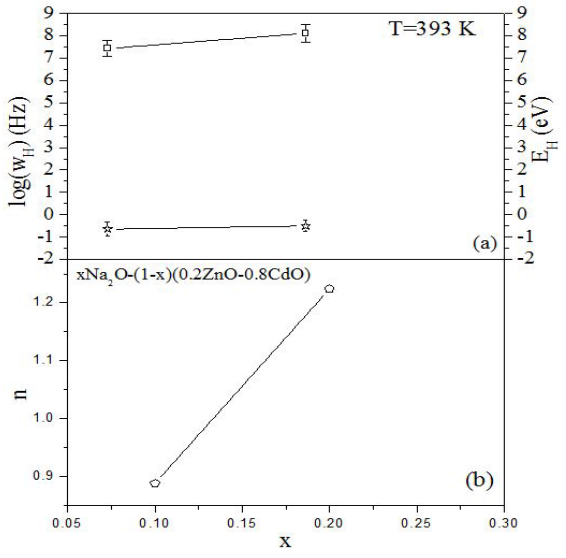

FIG. 3: (a) Hopping frequency vs concentration at a particular temperature T=393 K. (b) Frequency exponent(n) with compositions.

The activation energy depicted in Fig. 1 (b) demonstrates an increase with increasing x, suggesting a

potential shift in conduction mechanisms. This phenomenon could be attributed to the conduction of Na^+ ions encountering a larger barrier potential. Another plausible explanation could involve the existence of an increasing number of oxygens that are not bridge oxygens for concentrations ranging from 0.10 to 0.2. This presence may prompt the release of more electrons from ZnO. The additional electrons from ZnO will be quickly liberated due to its presence, given the expectation that Na^+ is likely to form bonds with O^{-2}.This line of reasoning leads to the consideration of electron/polaron hopping as the predominant conduction mechanism in samples with lower Na_2O content.

Conversely, with an increase in Na^+ ion content, more bridging oxygens are expected to associate with Na^+, suggesting Na^+-O^{-2} releasing additional sodium ions. Consequently, conductivity is anticipated causing domination due to the movement of Na^+ ions with increasing value of compositions. In summary, the magnitude of frequency exponent might be linked to the polaron-to-ion transferring mechanism as the quantity of Na_2O varies from lower to higher concentrations. The nature of activation energy for electrical conductivity is crucial in overcoming the respective potential barriers associated with these conduction mechanisms. This discussion provides insights into the dynamic interplay between composition, conduction mechanisms, and activation energy in the studied system.

By manipulating the electric field, one can anticipate the thermal activation of charge carriers as the determining factor for electrical transport in disordered solids [23, 24]. Jonscher's universal power law is often considered appropriate when analyzing the AC conductivity data of amorphous semiconductors [23, 24], expressed as

$$\sigma(\omega) = \sigma_0 + A\omega^S. \tag{3}$$

Here, σ_0 is the DC conductivity, A is the coefficient, and S is the power of frequency. Researchers [25–27] have explored the AC conductivity spectra in the higher frequency range, as illustrated in Fig. 4(a) for ($x = 0.2$) at different temperatures. The graph indicates a linear increase in AC conductivity with frequency in the high-frequency region, reflecting the nature of thermal activation. Equation (3) is employed to fit the AC conductivity values from Fig. 4(a), revealing a best-fit lines. The S-values are calculated from the slopes of these best-fit lines. Additionally, Fig. 4(a) suggests that as the frequency rises, and so does the AC conductivity. It is inferred from the figure that different conduction mechanisms with varying slopes S may underlie the conductivities at low and high temperature at various frequency ranges.

The other samples prepared in the similar manner also exhibit a same type of AC conduction process within higher frequency regime. In the lower frequency region the VRH (variable range hopping) model [28–30] is applied for the probable conductivity process. Drawing from the defect model developed by Ingram *et al* [31], the characterization of the effectiveness of conductivity is accurately performed in the

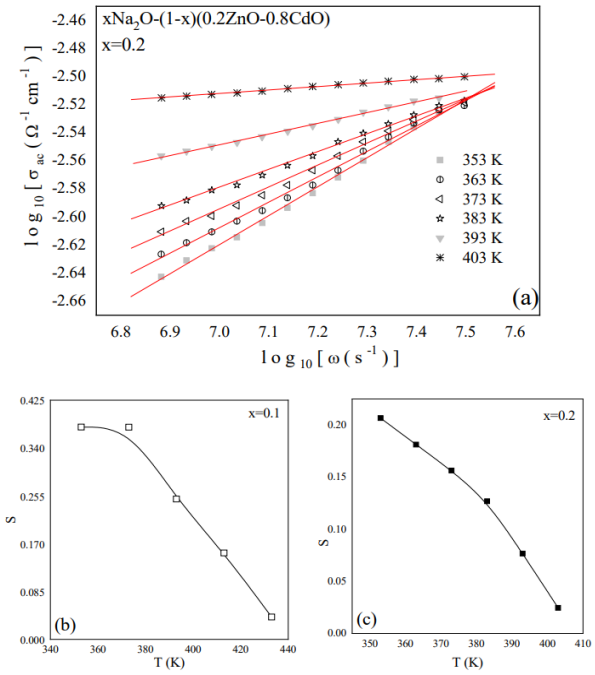

FIG. 4: (a) High conducitivity spectra for x=0.2 at various temperatures, solid straight lines indicates the best fitted straight line fits. (b) S-T plots for x=0.1 and (c) x=0.2.

high-temperature regions. The current researches [32] proposes that in the lower frequency region the polaron hopping aided by phonons is caused by the characteristic Ag-S vibration, whereas in the higher frequency regions it is caused by the unequal elongating vibrational modes of Zn-Te-Se bonding forming primary portion of the conductivity process.

Hence, it is common practice to explore the concept of conduction process within higher frequencies in relation with complete temperature. Approximate S-values depicted in Figs. 4(a) and 4(b) for (x = 0.1 and 0.2), respectively, providing insights into their conduction behavior. The trend reveals a decrease in S with increasing temperature for all samples, as illustrated in Figs. 4(a) and 4(b). These experimental findings, when fitted with an appropriate model, offer crucial information about electrical conduction in the current system. To resolve this problem several attempts [8] were made, and CBH (correlated barrier hopping) model [8, 26] has been proven to be suitable for the study of conduction mechanism.

The CBH model [26, 33, 34] predicts that the paired hopping paths of charge carriers (polarons) involve current transfer between localized sites at the Fermi level in a system exhibiting negatively dependence on temperature. According to this model, the expression for S is given by [33]

$$S = 1 - \frac{6k_BT}{W_m + k_B(T - T_0)\ln(\omega\tau_0)}. \tag{4}$$

Here, W_m is the maximum barrier height, k_B is the Boltzmann constant, T is the absolute temperature, T_0 is the temperature at which S becomes unity, and τ_0 is the relaxation period. The bold curves in Figs. 4(b) and 4(c) represent the best-fit curves using Eq. (4), with all fitting parameters estimated. For AC conductivity, the CBH model is given by

$$\sigma_{ac} = \frac{1}{24} n_0 \pi^3 N(E_F)^2 \varepsilon \varepsilon_0 \omega R_{H\omega}^6. \tag{5}$$

Here, n_0 is the number of polarons involved in the hopping process, $N(E_F)$ is the concentration of pair states, $R_{H\omega}$ is the hopping distance at frequency ω. At this moment the bi-polar hopping mechanism is likely to be primarily conduction process.

C. Current-Voltage Characteristics

To assess the suitability of the current system as a battery electrolyte on the scale of charge transport, the researchers investigated the current–voltage characteristics of prepared samples. An example of such characteristics is presented in Fig. 5 for $x = 0.2$ in both the forward and backward directions. Figure 5 clearly shows that the system under study displays a smooth straight curve having a consistent slope and there is no variation in back and forth ways. This observation clearly suggests Debye-like relaxation [18, 35–38] and dynamic resistance under a narrow range of applied voltage.

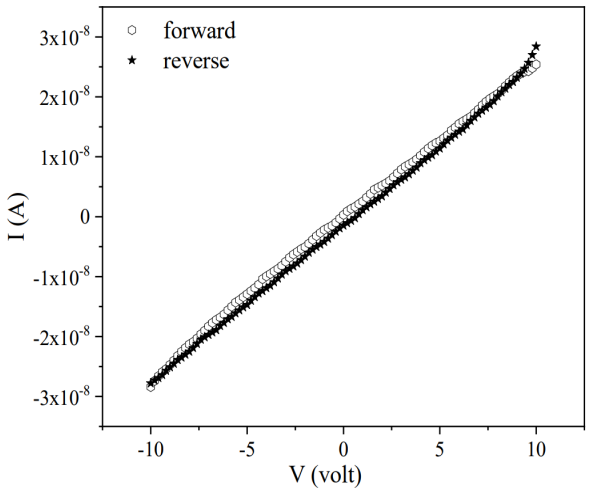

FIG. 5: I-V characteristics for $x = 0.2$.

The content of ZnO within the system is used to stabilize the dynamic resistance distribution in this context as an impact of phonon vibration. Robert Hill and D.A. Dissado proposed that, the arrangement of relaxation time [18, 38] in the non-Debye process is linked to a δ-function value, indicating a highly rapid

process. In contrast, δ-function value associated with a slow process characterized by a rapid decline in relaxation time indicates a Debye-like relaxation process. It is an indication of mixed type charge carriers stable conduction process as anticipated by the slower declination process of Debye-like relaxation mechanism. As observed from current-voltage characteristics in Fig. 5 the nearly zero gap in the conduction paths in both the forward and backward direction the glassy system doped with Na_2O is very much suitable as an electrolyte for upcoming era of batteries.

D. Relaxation process in the present system

The electrical relaxation characteristics of the investigated system were analyzed through the real (M') and imaginary (M'') components of the complex electrical modulus framework [39], employing the following expressions

$$M^*(\omega) = \frac{1}{\varepsilon^*(\omega)} = \frac{\varepsilon'(\omega)}{[\varepsilon'(\omega)]^2 + [\varepsilon''(\omega)]^2} + i\frac{\varepsilon''(\omega)}{[\varepsilon'(\omega)]^2 + [\varepsilon''(\omega)]^2} = M'(\omega) + iM''(\omega). \tag{6}$$

Figure 6(a) shows the plot of electric relaxation (M'') versus frequency for the samples $x = 0.1$ and $x = 0.2$ at a particular temperature (393 K). it is observed that M'' spectra exhibit a single relaxation peak where the height of the peak was increasing with higher value of x. the peak of the plots corresponds to the maximum value of frequency (ω_m) which is utilized to find the value of relaxation time (τ) the relation $\tau = 1/\omega_m$.

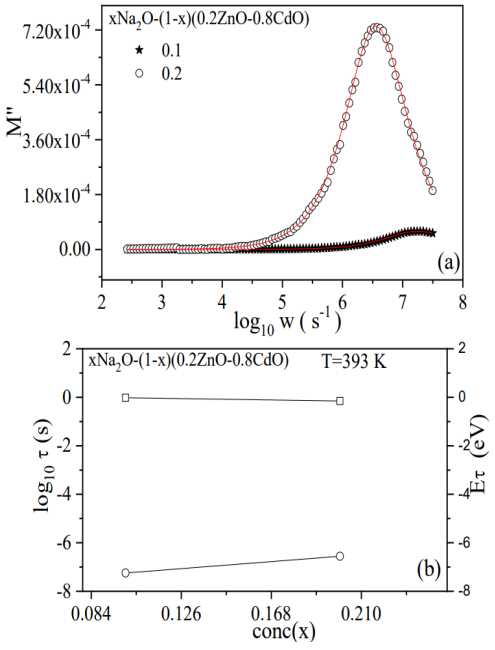

FIG. 6: (a) Modulus spectra for x=0.1 and x=0.2 (b) relaxation times and activation energy corresponding to relaxation times with compositions.

The required estimation in the relaxation process is identified as Debye-type relaxation for the current glassy system, as indicated by the relation $\omega_m \tau = 1$, where τ represents the relaxation time. The estimated relaxation times (τ) demonstrate a thermally activated nature, expressed as [18, 38]

$$\tau = \tau_0 \, e^{-E_\tau / k_B T}. \tag{7}$$

Here, τ_0 is the pre-exponential factor, k_B is the Boltzmann constant, and T is the absolute temperature. Equation (7) is used for least square fitting of the experimental data which yield relaxation times energy barrier E_τ. In Fig. 6(b), the relaxation times (τ) at a fixed temperature and their associated activation energies (E_τ) are presented. Notably, Fig. 6(b) shows that τ increases whereas E_τ decrease with varying compositions.

For compositions with lower Na_2O content, a short-time relaxation process is presumed, likely associated with the conduction of polarons. The overall relaxation mechanism within the current material under study is elucidated when it is associated with the lowest oscillations of sodium ions in compositions featuring higher quantity of Na_2O.Consideration is given to all types of lattice vibrations during relaxation mechanism for dielectrics. The values of structural parameters (β) [37, 38] against composition is given in Table I. The β values also distinctly signify a transition in the relaxation process, shifting from polaron hopping to ionic conduction [1].

x	$\beta(\pm 0.01)$
0.1	0.793
0.2	0.887

TABLE I: Structural parameters with compositions.

IV. CONCLUSION

In this study, glassy electrolytes of the composition $x$$Na_2O$-$(1-x)$(0.2ZnO-0.8CdO) were synthesized for x values of 0.1, 0.2, through the melt quenching route. These Na^+ conducting glassy electrolytes were investigated for its possible utilization as material for batteries, specifically studying its conductive properties of electricity at room temperature, as well as educational purpose. The analysis included the study of AC conductivity and relaxation behavior. The power law model, also known as the Almond-West formalism, yielded the frequency exponent (n), providing insights into the dimensionality of conduction pathways. For x values of 0.20, the observed n values were occasionally higher than 1, indicating the

potential presence of polaron-ion conductivity in the system. By the application of CBH-model the conductivity mechanism of electrons was recognized, elucidating the pathways of polarons hopping in pairs through the transfer of current among specific locations. Three-dimensional Na^+ motion was identified as the dominant charge carrier for samples with a higher Na_2O content. For the more content of Na_2O, the tridimensional sodium ions (Na^+) movement was identified as the dominant one. The composite conductive system demonstrated significantly greater direct current conduction process relative to Na_2O-Li_2O system, confirming the suitability of the present system as electrolytes for new-generation battery materials. The I-V characteristics exhibited a negligible difference in displacement pathways in both directions, suggesting the system's potential as a promising candidate for new-generation battery electrolytes, as a permanent deviation was absent. Na_2O-doped glassy systems were found to offer advantages over Li_2O-doped systems due to their highly conducting nature and reduced leakage, making them promising for improved battery materials. Additionally, the abundance of Na^+ in nature enhances their ease of use compared to Li^+.

ACKNOWLEDGMENTS

Financial assistance for this study from the DST-CRG (Department of Science and Technology, Govt. of India) via Sanction No.CRG/2018/000464 is gratefully acknowledged.

* Electronic address: ddirhrdc@nbu.ac.in

[1] B Karmakar, K. Rademann, and A. Stepanov, Glass nanocomposites: synthesis, properties and applications, Elsevier Publication, 2016.

[2] J. Y. Hwang, S. T. Myung, and Y. K. Sun, Chem. Soc. Rev., **46**, 3529 (2017).

[3] L. P. Wang, L. Yu, X. Wang, M. Srinivasan, Z. J. Xu, Journal of Materials Chemistry A, **3**, 9353 (2015).

[4] H. Aono, and Y. Sadaoka, Chemistry Letters, **29**, 34 (2000).

[5] I. A. Sokolov, V. N. Naraev, I. V. Murin, A. A.Pronkin, and A. V. Naraev. Russian journal of applied chemistry, **75**, 1240 (2002).

[6] Y. Zhao et al., Bulletin of Materials Science, **43**, 1 (2020).

[7] R. Cheruku, L. Vijayan, and G. Govindaraj, Materials Science and Engineering: B, **177**, 771 (2012).

[8] M. M. Ahmad, Physical Review B, **72**, 174303 (2005).

[9] L. K. Sudha, R. Sukumar, and K. Uma Rao, Int. J. Mater. Mech. Manuf. **2**, 96 (2014).

[10] E. D. Cubuk, S. S. Schoenholz, E. Kaxiras, and A. J. Liu, J. Phys. Chem. B **120**, 6139 (2016).

[11] F. A. Abdel-Wahab, A. M. Fayad, M. Abdel-Baki, and H. AbdelMaksoud, J. Non-Cryst. Solids **500**, 84 (2018).

[12] J. F. Stebbins and Z. Xu, Nature **390**, 60 (1997).

[13] J. S. Ashwajeet, T. Sankarappa, R. Ramanna, T. Sujatha, and A. M. Awasthi, AIP Conf. Proc. **1675**, 020017 (2015).

[14] A. Karmakar and A. Ghosh, J. Nanopart. Res. **13**, 2989 (2011).

[15] S. Bhattacharya and A. Ghosh, Appl. Phys. Lett. **88**, 133122 (2006).

[16] S. Shawuti, M. M. Can, M. A. Gulgun, S. Kaneko, and T. Endo, Compos. Part B: Eng. **147**, 252 (2018).

[17] S. Bag, P. Das, and B. Behera, J. Theor. Appl. Phys. **11**, 13 (2017).

[18] Y. B. Saddeek, A. A. El-Maaref, M. G. Moustafa, M. M. El-Okr, and A. A. Showahy, J. Mater. Sci.: Mater. Electron. **29**, 9994 (2018).

[19] V. V. Korotyeyev, V. A. Kochelap, V. V. Kaliuzhnyi, and A. E. Belyaev, Appl. Phys. Lett. **120**, 252103 (2022).

[20] M. Novy, H. Avila-Paredes, S. Kim, and S. Sen, J. Chem. Phys. **143**, 241104 (2015).

[21] D. L. Sidebottom, Phys. Rev. Lett. **83**, 983 (1999).

[22] L. Mancini, C. Janot, L. Loreto, R. Farinato, J. Gastaldi, and J. Baruchel, Philos. Mag. Lett. **78**, 159 (1998).

[23] C. Tsonos, Curr. Appl. Phys. **19**, 491 (2019).

[24] A. Dhahri, E. Dhahri, and E. K. Hlil, RSC Adv. **8**, 9103 (2018).

[25] D. Das and S. Samanta, ACS Appl. Electr. Mater. **3**, 1634 (2021).

[26] H. Bouaamlat, N. Hadi, N. Belghiti, H. Sadki, M. NaciriBennani, F. Abdi, M. Abarkan, T.-d. Lamcharfi, and M. Bouachrine, Adv. Mater. Sci. Eng. 2020, 1 (2020).

[27] A. Langar, N. Sdiri, H. Elhouichet, and M. Ferid, Results Phys. **7**, 1022 (2017).

[28] B. I. Shklovskii, A. L. Efros, B. I. Shklovskii, and A. L. Efros, Electronic Properties of Doped Semiconductors, **45**, 202 (1984).

[29] R. Plugaru, T. Sandu, and N. Plugaru, Results Phys. **2**, 190 (2012).

[30] N. T. T. Hong, V. S. Zakhvalinskii, T. T. Pham, N. T. Dang, E. A. Pilyuk, G. V. Rodriguez, and T. V. Vu, Mater. Res. Express **6**, 055915 (2019).

[31] B. J. Ingram, B. J. Harder, N. W. Hrabe, T. O. Mason, and K. R. Poeppelmeier, Chem. Mater. **16**, 5623 (2004).

[32] A. Chamuah, K. Bhattacharya, C. K. Ghosh, and S. Bhattacharya, Mater Today: Proc. **66**, 3218 (2022).

[33] Ajili, B. Louati, and K. Guidara, J. Mater. Sci.: Mater. Electron. **29**, 8649 (2018).

[34] M. Singh, K. L. Bhatia, N. Kishore, P. Singh, and R. S. Kundu, J. Non-Cryst. Sol. **180**, 251 (1995).

[35] F. Tian and Y. Ohki, IEEE Trans. Dielectrics Electr. Insul. **21**, 929 (2014).

[36] M. Rajnak, B. Dolnik, J. Kurimsky, R. Cimbala, P. Kopcansky, and M. Timko, J. Chem. Phys. **146**, 014704 (2017).

[37] C. R. Mariappan, G. Govindaraj, S. V. Rathan, and G. V. Prakash, Mater. Sci. Eng: B **123**, 63 (2005).

[38] J. Liu, C. G. Duan, W. G. Yin, W. N. Mei, R. W. Smith, and J. R. Hardy, J. Chem. Phys. **119**, 2812 (2003).

[39] S. N. Mohamed, M.K. Halimah, R. H. Y. Subban, A. K. Yahya, Physica B: Condensed Matter, **602**, 412480 (2021).

J. Phys. Res. Edu., Vol. 2, March 2024

Flexible Piezoelectric Energy Harvester Based on UV Light Emitting Ce^{3+}-Complex-P(VDF-HFP) Composite Films

Priti Sundar Barman and Prakriti Adhikary*

Department of Physics, University of North Bengal,

Raja Rammohunpur, District: Darjeeling-734013, West Bengal, India

In this study, we report on the preparation of a polymer composite film with enhanced ($\sim 99\%$) electroactive phases β- and γ-phase) based on P(VDF-HFP) copolymer and Ce^{3+}-complex. Furthermore, its utilization in piezoelectric-based flexible energy harvesting (PFEH) device fabrication has been studied, where the electrostatic interactions between the surface charges of Ce^{3+}-complex and -CH_2-/-CF_2- molecular dipoles of P(VDF-HFP) via H-bonding co-operate to stabilize the electroactive phases and enhance its piezoelectric properties. PFEH generates $\sim$ 3 V of open circuit voltage and 0.16 μA short-circuit current under the external pressure impacting amplitude of 14.20 kPa. Moreover, it can successfully charge up capacitor by repeating finger impact which indicates its potency as an efficient energy-harvesting device. Besides this, the composite film exhibits an intense photoluminescence in the UV-region that might be very promising in the area of high-performance, energy-saving, flexible, solid-state UV light emitters and fabrication of hybrid multifunctional energy harvester where mechano-luminescence phenomenon might be possible to include.

I. INTRODUCTION

Nowadays, nanomaterial-based pressure and strain sensors are the most rapidly growing area of research, as a consequence piezoelectric-based energy harvesters (namely nanogenerators) are connected synergistically with the field of energy harvesting, a recent topic with great importance. Nanogenerators are becoming the simple and easiest implementable key energy harvesting that converts mechanical energy into electrical energy from nano-piezoelectric materials as initially proposed by Wang and Song using ZnO nanowire arrays [1]. Therefore, several inorganic piezoelectric materials, such as PZT, $BaTiO_3$, $ZnSnO_3$, CdS, ZnS, etc., have been potentially utilized as power sources in device forms [2–6]. However, often brittleness, toxicity, and poor biocompatibility become the limitations in inorganic materials-based piezoelectric device fabrication and applicability. In contrast, semi-crystalline and long-chain polymers, such as PVDF (polyvinylidene fluoride) and its copolymers P(VDF-HFP) (poly(vinydilene fluoride-co-hexafluoropropylene)) and P(VDF-TrFE) (poly(vinilydene fluoride-co-trifluororthylene) from the fluoropolymer family, might be the best alternate of inorganic counterparts due to its light weight, flexibility, stretchability, cost-effectiveness

and most importantly bio-compatibility [6–9]. In this study, copolymer P(VDF-HFP) has been selected as the host material due to the presence of strongly polarizable C-F bonds and its possibility to generate electroactive crystalline phases spontaneously in the PVDF counterpart. Moreover, P(VDF-HFP) exhibits good mechanical properties with excellent film-forming capability, chemical stability, and an unusual piezoelectric response, i.e.,$|d_{31}/d_{33}| > 1$, which is not a common feature in other piezo-polymers [9–12]. Generally, P(VDF-HFP) has mainly four crystalline conformations, viz., α, β, γ, and δ, signified by the stereo-chemical representation of the structure with alternating s-trans and s-gauche carbon-carbon bond such as $TGT\bar{G}$, $TTTT$, $T_3GT_3\bar{G}$, and $TGT\bar{G}$ (polar) (T = trans, G = gauche $+$, $\bar{G}$ = gauche $-$), respectively [9–12]. Among them β- and γ-phases are primarily the focus of interest due to their valuable piezo-, pyro-, and ferroelectric properties [13]. For instance, there are several processes for nucleating β/γ-phases, such as mechanical stretching, under an electric field, and by controlling pressure have been adopted however, the main bottleneck of these techniques is not cost-effective and industrial viable in terms of large-scale device production. In this work, we focus on the best alternative method where the addition of selective external filler may lead high yield of β/γ- phases controlling the solvent casting parameters, such as concentration of filler and P(VDF-HFP) in solvent and temperature.

Thus, we report on a novel, flexible electro-active (more than 99% electro-active phase) Ce^{3+}-complex and P(VDF-HFP) composite film prepared by a simple solution casting method where electrical poling is possible to avoid in order to get the piezoelectric behaviour [4, 5, 7]. This composite film-made PFEH possesses superior piezoelectric performance (output voltage of 3 V and 0.16 μA short-circuit current by repetitive finger touch motion) under external pressure stimuli that indicate it is very efficient in mechanical energy harvesting-based applications. Ce^{3+}-complex not only nucleates piezoelectric β-phase in the composite film via electrostatic (like H-bonding) interaction it also has the capability of UV light emitting that provides a new possibility in optoelectronic device fabrications.

II. EXPERIMENTAL SECTION

A. Film preparation

The P(VDF-HFP) copolymer pellets (Sigma-Aldrich (Mw$\sim$ 400,000)) were immersed in N, N-dimethyl formamide (DMF, Merk, India) solvent to prepare 6 wt% (w/v) P(VDF-HFP) solution by continuous stirring at room temperature ($\sim$ 30 ℃ for 12 hours to make sure that the pellets are completely dissolved in DMF. Then 0.1 gm of cerium salt ($(NH_4)_4Ce(SO_4)_4 \cdot 2H_2O$) (Alpha-Asser) was added with P(VDF-HFP) solution and stirred for 10 days at room temperature till the colour of the resulting solution changed from orange

to milky white. A reference solution of pure P(VDF-HFP) was also been prepared for comparison. The solutions were casted on clean glass substrates followed by drying at 90 ℃ for 6 hours and peel-off for the characterizations and PFEH fabrication. Finally, two sets of films were prepared, (i) HFP (where no additive is present) and (ii) Ce-HFP (where Ce-salt is utilized).

B. Characterization

X-ray diffractometer (Bruker, D8 Advance) with CuK_α radiation of wavelength 0.15405 nm was used for crystallographic investigations. The characteristic vibrational modes were analyzed by FT-IR by using Bruker Tensor II spectrometer for crystalline phase identification with 4 cm^{-1} of spectral resolution. The in-situ thermal FT-IR was performed in transmission mode with a thermal ramp of $\sim$ 1 ℃/min. The surface morphology of composites was studied using a field emission scanning electron microscope (FE-SEM), INSPECT F50, operated at 20 kV accelerated voltage. Photoluminescence (PL) spectra were recorded with a Horiba (iHR320) luminescence spectrometer with an excitation wavelength (λ_{exc}) of 270 nm.

C. Fabrication of PFEH

Devices were fabricated by simply pasting carbon tape as electrodes (area $\sim 15 \times 30$ mm^2) onto both sides of the composite films (thickness $\sim$ 0.3 mm $\pm$ 10 μm) and lastly encapsulating with polydimethylsiloxane (PDMS) layers to protect the surface from physical damage, Piezoelectric responses were recorded in terms of open-circuit output voltage using a digital storage oscilloscope (Agilent, DSO3102A). All the measurements were carried out at room temperature.

III. RESULTS AND DISCUSSIONS

The XRD results reveal that HFP film exhibits three major peaks originating at 19.8° (1 1 0), 18.3° (0 2 0), and 17.6° (0 2 1) as shown in Fig. 1 which are characteristic of the non-polar α-phase. Whereas, Ce-HFP film possesses a diffraction peak at 20.7° that is attributable to overlapping lattice reflections of (1 1 0) and (2 0 0) planes owing to the existence of polar β-crystalline phase along with trace amount of crystalline γ-phase that is confirmed by additional three peaks originated at 20.4°, 19.1°, and 18.4° arises for (1 1 0), (0 0 2), and (0 2 0) planes respectively [9, 10, 12].

Therefore, all the α-phase characteristic peaks are translated to electro-active phases in Ce-HFP film as Ce-salt filler takes a significant role in crystallographic phase transformation ($\alpha \rightarrow \beta/\gamma$) in P(VDF-HFP) polymer matrix, which is further affirmed with Fourier Transform Infrared (FT-IR) bands in the subsequent

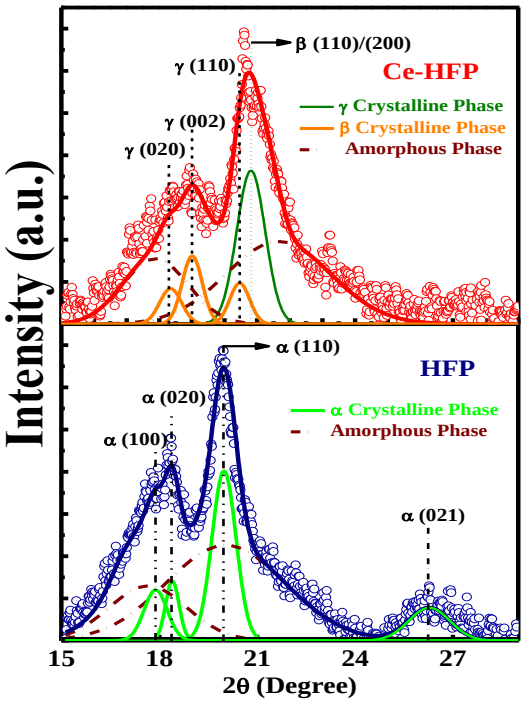

FIG. 1: XRD peak and its curve deconvolution from HFP and Ce^{3+}-HFP films respectively, where the open circles are experimental data points and solid lines correspond to the best-fitted curves.

section. The broad diffraction peak at 20.7° develops for the overlapping of the polar β-phase at 20.8° (1 1 0)/(2 0 0) and semi-polar γ-phase at 20.4° (1 1 0), which demands a curve deconvolution for describing amorphous and crystalline phases and further computing their crystallite sizes. The overall degree of crystallinity (χ_c) is calculated from the relation,

$$\chi_c = \frac{\sum A_{cr}}{\sum A_{cr} + \sum A_{amr}}, \tag{1}$$

where $\sum A_{cr}$ and $\sum A_{amr}$ represent the integrated areas for the crystalline peaks and amorphous halos of XRD pattern respectively. The degree crystallinity(χ_c) diminishes in the Ce-HFP film (see, Table I) which is a very common behaviour of a reinforced polymer composites and suitable for the enhancement of the dielectric properties also. The average crystallite size (D) is obtained according to Debye–Scherrer formula,

$$D = \frac{k\lambda}{B\cos\theta}, \tag{2}$$

where B is the full width at half maximum (FWHM) of the diffraction peak in radians and λ is the X-ray wavelength (i.e., 0.154 nm), k = 0.89, a constant. The crystallite sizes are tabulated in Table. I which reveals that β-crystallite size is relatively greater than the γ-crystals. In Ce-HFP film the hydrogen bonding interaction between the Ce-salt filler and co-polymer matrix possess all-trans conformations in layer-by-

layer configuration and large content of polar β-phase is achieved which will deliver better piezo-response as it is directly proportionate to the β-crystallinity.

Sample	χ_c(%)	D_α	D_β	D_γ
HFP	48	9	–	–
Ce-HFP	34	–	12	8

TABLE I: Degree of crystallinity(χ_c), sizes of α-, β- and γ-crystallites of the composite films (D_α, D_β, D_γ) calculated from the XRD curve deconvolution pattern.

In the FT-IR spectra in Fig. 2(a) evidence that the vibration bands that appeared at 612, 763, 970, and 1211 cm^{-1} confirm the predominant non-polar α-phase in HFP film. In contrast in Ce-HFP film, all α-phases characteristic peaks completely disappeared, whereas two prominent peaks at 1276 cm^{-1} and 1231 cm^{-1} are illustrated, which are the signatures of β-phase ($TTTT$ conformation) and γ-phase ($TTTG$ conformation) respectively [9, 10, 12]. As β- and γ-phases resemble each other structurally common TTT chain conformation that possess some common vibration bands such as 840 and 510 cm^{-1} which exhibit the dual character of β and γ-phases [9, 10]. It indicates that the inclusion of cerium salt takes a crucial role in the crystallographic transformation, i.e., $\alpha \rightarrow \beta/\gamma$ phase in P(VDF-HFP). Two additional well-resolved bands at 640 and 602 cm^{-1} shown in Fig. 2(a) and another arrow-marked broad band at 1041 cm^{-1} correspond to Ce-O stretching vibrations (ν_{Ce-O}) that indicates that cerium salt resembled as Ce^{3+} complex in DMF [6]. The absorption intensity of the 840 cm^{-1} band is assigned to quantify the relative proportion of electroactive phases fraction (F_{EA}), by using the following equation,

$$F_{EA} = \frac{A_{EA}}{A_{EA} + \frac{K_{840}}{K_{763}} A_{NEA}}, \tag{3}$$

where A_{EA} and A_{NEA} are the absorbance at 840 cm^{-1} and 763 cm^{-1}, respectively, and K_{840} (7.7×104 cm^2 mol^{-1}) and K_{763} (6.1×104 cm^2 mol^{-1}) are the absorption coefficients at the respective wavenumbers. 99% of electroactive phases are achieved in Ce-HFP film by utilizing 1 wt% of Ce^{3+}-salt filler in P(VDF-HFP) matrix [Fig. 2(c)]. After that, the β- and γ-phases are quantified individually by curve deconvolution of 840 cm^{-1} band as shown in Fig. 2(d) by using the following relations,

$$F(\beta) = F_{EA}\left(\frac{A_\beta}{A_\beta + A_\gamma}\right) \times 100\% \tag{4}$$

and

$$F(\gamma) = F_{EA}\left(\frac{A_\gamma}{A_\beta + A_\gamma}\right) \times 100\%, \tag{5}$$

where, A_β and A_γ are the area under the deconvoluted bands of the β- and γ-phases, respectively centred at 840 cm^{-1}. The individual values of $F(\beta)$ and $F(\gamma)$ are incurred 85% and 15% respectively, which portrayed that, there present a large content of β-phase in Ce-HFP film compared to γ-phase. The shifting of -CH_2 asymmetric (ν_{as}) and symmetric (ν_s) stretching vibration bands [inset of Fig. 2(b)] in Ce-HFP film towards lower frequency regions reveal the confirmation of interfacial interaction from the positively charged Ce^{3+}-complex and -CF_2 dipoles in P(VDF-HFP) matrix. Additionally, a broad band observed at 3342-3140 cm^{-1} [Fig. 2(b)] region, originated due to the -OH stretching vibrations (ν_{O-H}) at the interface between Ce^{3+}-complex and P(VDF-HFP) copolymer matrix. Actually, the -HSO_4 ligand from Ce^{3+}-complex creates hydrogen bonding (O-H$\cdots$F-C) with F-ion of P(VDF-HFP) chain that may encapsulate the Ce^{3+}-complex [Fig. 2(e)], which provided a higher yield of electroactive β-phase in hydrated salt utilized P(VDF-HFP) film [9, 10].

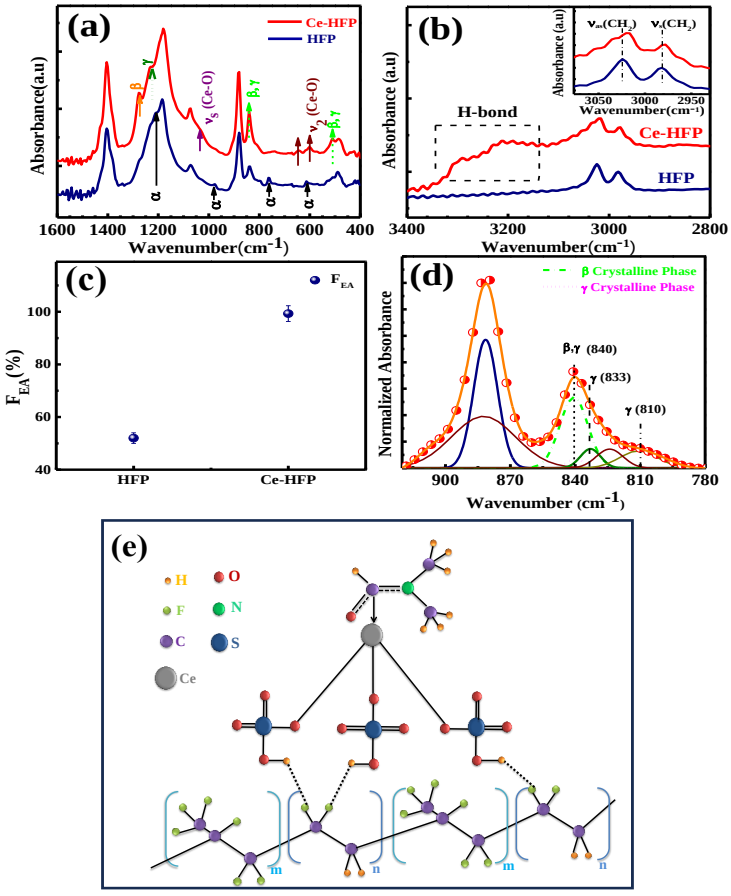

FIG. 2: FT-IR spectra of HFP and Ce-HFP films in different frequency regions, (a) 1600-400 cm^{-1} and (b) 3400-2800 cm^{-1}. (c) Relative % of FEA for the HFP and Ce-HFP films. (d) Deconvolution curve of FT-IR spectra of Ce-HFP film in frequency 920-780 cm^{-1}, the dots are experimental data and the solid lines correspond the best fitted data and, (e) Schematic drawing of Ce-complex coordinates to DMF and its interaction with P(VDF-HFP) via H-bonding.

The in-situ temperature-dependent FT-IR of Ce-HFP film was also executed to feel the thermal behaviour of the crystalline polymorphs (β- and γ-phases), shown in Figs. 3(a) and 3(b). The absorbance intensities of β-(β_{1275}), γ-(γ_{1231}) and α-(α_{763}) phase vs. temperature is represented in Fig. 3(c) that affirms the lower melting behaviour of the β-phase (144 ℃) comparable to α-phase (184 ℃) and γ-phase (199 ℃) [14]. Noteworthy that piezoelectric properties in P(VDF-HFP) is primarily governed by the electro-active β-phase, therefore in-situ thermal FT-IR study provides a clear idea about piezoelectric based device operation region which is around 144 ℃. It is worth noting that the improvement in the thermal stability has prime importance for widening the device operating temperature regions, since the electro-active phases are mainly responsible for the piezoelectric origin.

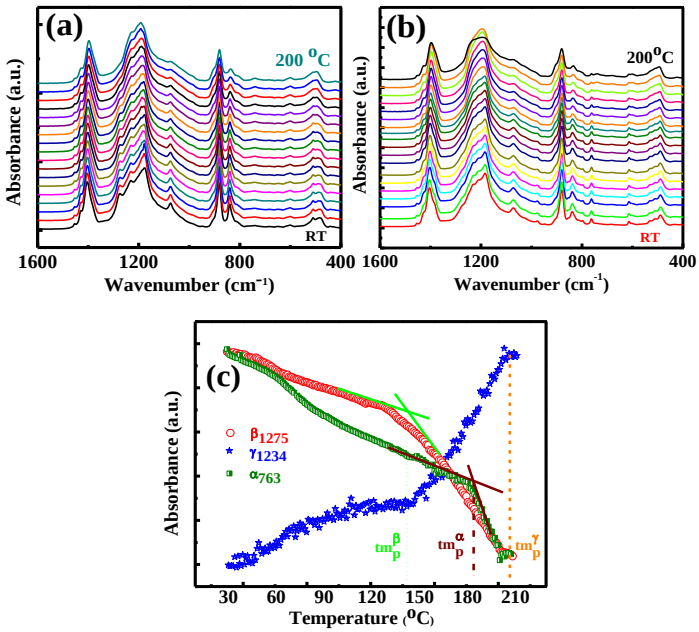

FIG. 3: In-situ thermal FT-IR spectra of (a) Ce-HFP and (b) HFP film with thermal ramp $\sim$ 1 ℃/min, where the spectra in between room temperature 30 ℃ to 200 ℃ are shown in 10 ℃ C interval for clarity. (c) The absorbance intensities of β-phase at 1275 cm^{-1}, γ-phase at 1234 cm^{-1} from Ce-HFP film and α-phase at 763 cm^{-1} from HFP film for representing their peak melting temperatures via in-situ thermal FT-IR study.

The FE-SEM images represented in Fig. 4(a) (HFP film) and Fig. 4(b) (Ce-HFP film) illustrated that the doping of Ce^{3+}-complex in P(VDF-HFP) copolymer matrix improves the polymer surface morphology as it embedded in copolymer matrix. A large number of particles with rod shapes but in different sizes are formed in Ce-HFP film as observed from the micrograph. The Inset of Fig. 4(b) shows the digital photograph in foldable condition that demonstrates the flexibility of the corresponding film. On the other hand, the HFP film shows the predominant α-phase, evidenced by α-spherulitic fibril growth which is consistent with the FT-IR spectra. So, in Ce^{3+}-complex utilized HFP film the α-phase characteristic is

diminished which signifies that the doping can hinder the growth of the non-polar phase and favor the formation of the electroactive phases.

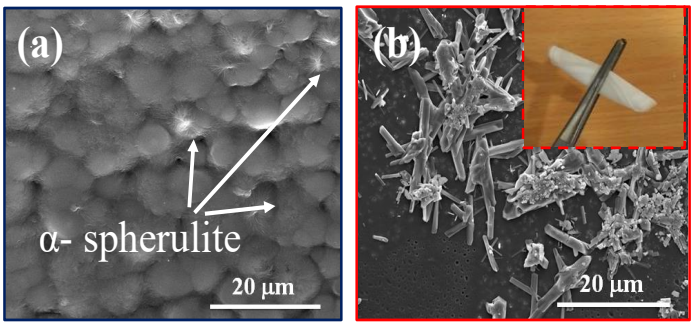

FIG. 4: FE-SEM images of the (a) HFP and (b) Ce-HFP films, inset of (b) represent the digital photograph of the composite film to demonstrate the flexibility.

Piezoelectric output signals from PFEH were performed under repetitive finger touch and release motion with applied stress amplitude of 14.20 kPa and it delivers up to $\sim$ 3 V of open circuit voltage as shown in Fig. 5(a). To affirm the output voltage from PFEH is monitored by the piezoelectric properties, the polarity test was performed simply by reversing the electrode connections with the top and bottom and the insets of Fig. 5(a) represent the reverse and forward circuit diagrams respectively. It is important to note that due to the different strain rate at touch and release response the output signals are not uniform. The enhanced output response of PFEH is due to the change in dipole moment (i.e. $-CH_2-/-CF_2-$) of P(VDF-HFP) itself with the applied stress along with change of initial dipole moment in non-centrosymmetric Ce^{3+}-complex encapsulate with P(VDF-HFP) via electrostatic interaction/H-bonding. The deformation of non-centrosymmetric crystal structure of Ce^{3+}-complex in Ce-HFP film causes remarkable improvement in the piezoelectric response by inter play of its deformed structure to stable one and vice versa with the application of applied stress [6]. Additionally, the hydrogen bonding (O-H$\cdots$F-C) interaction not only promote the nucleation of piezoelectric β-phase, but also improves the polarization ordering by increasing the stability of the molecular chain that leads to a better alignment of dipoles ($-CH_2-/-CF_2-$) due to self-poling phenomenon [9, 12]. Herein the dipoles are oriented in a particular direction that favour a stress induced polarization and as a result the external electrical poling is possible to avoid. The dependency of output voltage and short-circuit current signals from PFEH as a function of different load resistances (RL) were evaluated and it is found that the output voltage through the outer resistor gradually rises up to certain extend (i.e., $\sim$ 3 V) at theoretically infinite resistance, in contrast the corresponding current signal decreases with increasing the load resistance [Fig. 5(b)], which corresponds to the principle of open circuit condition. The instantaneous power reaches maximum of $\sim$ 0.07 μW [Fig. 5(c)] at a resistance of 10 MΩ. The capacitor

charging ability of the PFEH was also tested with rectifier bridge circuit [inset of Fig. 5(d)], and the result shows in Fig. 5(d) its ability to successfully charge up 1 μF capacitor within very short interval of time ($\sim$ 2 min.) by repetitive finger touch and release motion. Therefore, the capacitor charging performance of the PFEH demonstrates that it might be applicable as alternative choices towards the area of self-powered energy harvesting devices.

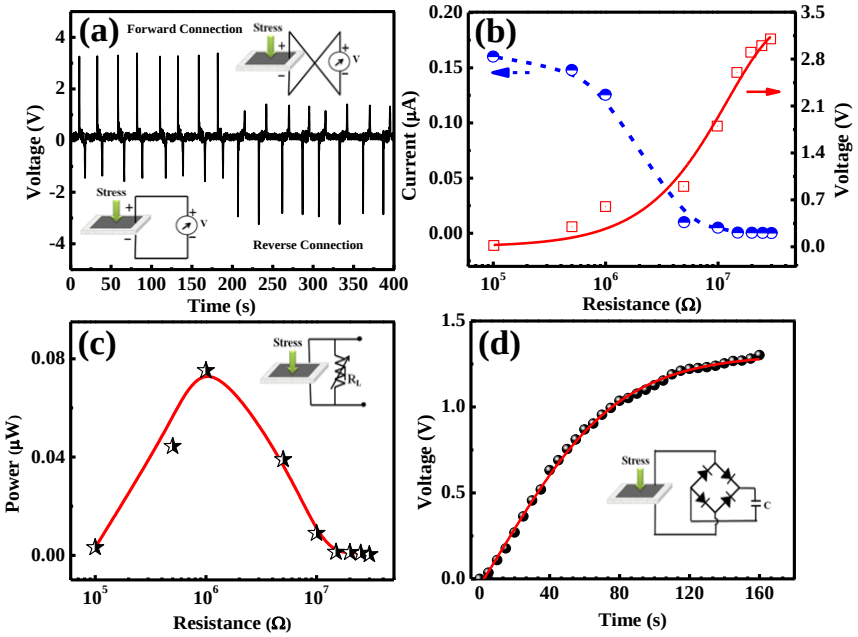

FIG. 5: (a) Piezoelectric output voltage generated in forward connection (left half) and reverse connection (right half). (b) Variation of output voltage and current signals as a function of variable resistance under finger touch. (c) Effective output power with the equivalent circuit diagram and, (d) Capacitor charging response from PFEH by repetitive finger tapping (applied pressure amplitude $\sim$ 14.20 kPa) where inset of (d) represents the corresponding circuit diagram used to charge up 1 μF capacitor.

The room temperature PL spectrum of the Ce-HFP film is recorded by using an excitation wavelength at $\lambda_{\text{exc}} = 270$ nm as shown in Fig. 6(a). Ce^{3+} ions have only one electron in its $4f$ shell that can be excited to the $5d$ orbitals upon UV light irradiation. The excited $5d$ state of Ce^{3+} splits into two spin-orbit components, depicted as $^2D_{5/2}$ and $^2D_{3/2}$ as shown in Fig. 6(b). When Ce^{3+} ions are excited at 270 nm, the transitions from the ground $^2F_{7/2}$ state to excited $5d$ states is occurred. As a consequence, emission shows two bands due to the double character of the $^4f_1$ ground state (spin-orbit components are $^2F_{5/2}$ and $^2F_{7/2}$ states), which are attributed to the $^2D \rightarrow {}^2F_{5/2}$ (321 nm) and $^2D \rightarrow {}^2F_{7/2}$ (344 nm) transitions. These transitions are distinctly expressed by curve deconvolution process ensuing two Gaussian peaks with maxima at 321 and 344 nm, respectively [6, 15]. Thus Ce^{3+}-complex doped P(VDF-HFP) film promises a new type of photonic application apart from its piezoelectric activity, that extends its importance as potential

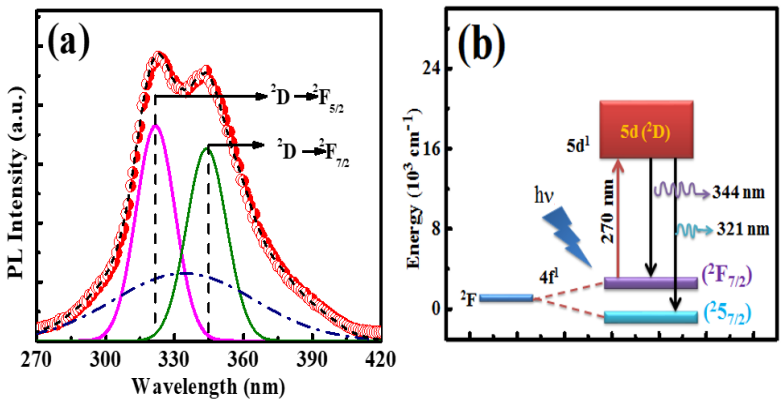

FIG. 6: (a) Room temperature PL spectra of the Ce-HFP film. (b) Energy level diagram for explaining the light emission from Ce^{3+}- complex with excitation at $\lambda_{exc} = 270$ nm.

multifunction film for versatile sensor developments.

IV. CONCLUSIONS

In conclusion, more than 99% of electro-active β- and γ-phases are achieved in the P(VDF-HFP) polymer matrix by introducing Ce^{3+}-complex filler. The electrostatic interaction between Ce^{3+}-complex and CF_2 groups of host polymer plays an important role in enhancing piezoelectric properties via H-bonding. PFEH can deliver a maximum open circuit voltage of $\sim$ 3 V and a short-circuit current of $\sim$ 0.16 μA by repetitive finger touch motions. The charging capability of the capacitor reflects its potential to be utilized as a piezoelectric-based energy harvesting device. Additionally, Ce-HFP displays an intense photoluminescence in the UV-region that might have potential application in the area of future optoelectronics.

ACKNOWLEDGEMENTS

The authors are greatly thankful to the Department of Physics, North Bengal University, for providing financial support and research facilities. Priti Sundar Barman thankfully acknowledges Govt. of West Bengal for providing Swami Vivekananda Merit-cum-Means Scholarship (WB-SVMCM; Applicant ID: WBP211629445105).

* Electronic address: prakriti@nbu.ac.in

[1] Z. L. Wang and J. Song, Science **312**, 242 (2006).

[2] X. Chen, S. Xu, N. Yao, and Y. Shi, Nano Lett. **10**, 2133 (2010).

[3] K. I. Park, S. Xu, Y. Liu, G. T. Hwang, S. J. L. Kang, Z. L. Wang, and K. J. Lee, Nano Lett. **10**, 4939 (2010).

[4] L. Wu, Z. Jin, Y. Liu, H. Ning, X. Liu, Alamusi, and N. Hu, Nanotechnology Reviews **11**, 1386 (2022).

[5] L. Wu, M. Jing, Y. Liu, H. Ning, X. Liu, and S. Liu, Compos Part B. **164**, 703 (2019).

[6] K. Maity, S. Garain, K. Henkel, D. Schmeiber, and D. Mandal, Self-powered human-health, ACS Appl Polym Mater. **2**, 862 (2020).

[7] V. Bhavanasi, V. Kumar, K. Parida, J. Wang, and P. S. Lee, ACS Appl. Mater. Interfaces **8**, 521 (2016).

[8] D. Mandal, S.Yoon, and K. J. Kim, Macromol. Rapid Commun. **32**, 831 (2011).

[9] P. Adhikary, S. Garain, S. Ram, and D. Mandal, J. Polym. Sci. B: Polym. Phys. **54**, 2335 (2016).

[10] P. Adhikary, S. Garain, and D. Mandal, Phys. Chem. Chem. Phys. **17**, 7275 (2015).

[11] P. S. Barman, S. Garain, and P. Adhikary, Phys. Chem. Chem. Phys. **25**, 30583 (2023).

[12] S. K. Ghosh, T. K. Sinha, B. Mahanty, and D. Mandal, Energy Technol. **3**, 1190 (2015).

[13] P. Costa, J. Nunes-Pereira, N. Pereira, N. Castro, S. Gonçalves, and S. Lanceros-Mendez, Energy Technol. **7**, 1800852 (2019).

[14] P. Adhikary and D. Mondal, Phys. Chem. Chem. Phys. **19**, 17789 (2017).

[15] A. Kaminska, A. Duzynska, M. Berkowski, S. Trushkin, and A. Suchocki, Phys. Rev. B. **85**, 155111 (2012).

J. Phys. Res. Edu., Vol. 2, March 2024

A method of estimating the EAS cores of Monte Carlo showers for the GRAPES-3 experiment

Animesh Basak*

Department of Physics, University of North Bengal, Siliguri, WB, 734 013, India

The procedure of estimating the different extensive air shower (EAS) parameters is inherently linked to the accurate estimation of the cosmic-ray EAS cores. In EAS data analyses, the core of an EAS is estimated simultaneously with other crucial EAS parameters like shower size, shower age, etc. by fitting the lateral density data (LDD) of either the EAS charged secondaries or purely electrons with some suitably chosen lateral density function employing the maximum likelihood method. The present analysis estimates EAS cores using the LDD of electrons that fall on the scattered array detectors from the simulated EASs initiated by proton and iron primaries. Considering a densely packed detector array, including configurations akin to GRAPES-3, the research employs a straightforward weight average method (WAM) for the EAS core estimation. The findings reveal that around 95.5% of simulated showers exhibit EAS cores within a deviation range of approximately 1 m to 3 m from the actual cores of the CORSIKA Monte Carlo showers initiated by proton and iron primaries.

I. INTRODUCTION

Cosmic rays (CRs) are the high-energy particles originating from sources beyond our solar system, continually bombard the Earth's atmosphere. When these CRs collide with atmospheric nuclei, they trigger a cascade of secondary particles, resulting in what is known as an extensive air shower (EAS). Formation of an EAS is a complex phenomenon involving the interaction of CRs with the atmosphere, producing a myriad of secondary particles, including electrons, positrons, muons, photons, and hadrons. These particles propagate through the atmosphere, creating a shower-like pattern extending several square kilometers. One remarkable aspect of EASs is their immense energy. Primary CRs can possess energies ranging from 10^9 to 10^{20} eV, far surpassing the energies achievable in particle accelerators on Earth. As the primary CR particle collides with atmospheric nuclei, its energy is distributed among the secondary particles, leading to the development of the EAS. The study of EASs provides valuable insights into the properties and origins of primary CRs and the fundamental processes governing high-energy particle interactions in the cosmos. By analyzing the characteristics of air showers, crucial information about EASs can be inferred, such as the energy spectrum, composition, and arrival directions of CRs, shedding light on their astrophysical sources and acceleration mechanisms. Statistically, EASs exhibit a power-law distribution in terms of their energy spectrum, with

higher-energy events being less frequent than lower-energy ones. Some interesting characteristics of the distribution, like - the knee, the ankle and the GZK cutoff of the primary particle energy spectrum, reflect the underlying physics of CR interactions and are a vital feature analyzed by researchers to understand the nature of CR accelerators and the mechanisms responsible for their acceleration to such extreme energies. Furthermore, the statistical properties of EASs, including their lateral spread, lateral/radial particle density, and the depth of shower maximum development in the atmosphere, offer valuable constraints for theoretical models and simulations of CR propagation and interaction processes.

Detecting and analyzing EASs require sophisticated instruments and experimental setups, such as those employed in facilities like the GRAPES-3 experiment [1, 2]. The precision of measurements pertaining to EAS parameters rests upon the meticulous estimation of the EAS shower cores. Various methodologies are being explored to attain a more accurate core estimation. In the case of densely configured EAS arrays like GRAPES-3, employing the stated weighted average method (WAM) [3] through simulation is anticipated to streamline the proceedure while ensuring accuracy in delineating the EAS core.

The plan of the paper is as follows. A brief description of the GRAPES-3 experiment is given in Section II. The air shower simulations for the study is presented in Section III. Section IV describes briefly the EAS core estimation formulae. In Section V, the main results of the work with the necessary discussion is given. The paper ends with a conclusion in Section VI.

II. THE GRAPES-3 EXPERIMENTAL CONDITIONS

The GRAPES-3 (Gamma Ray Astronomy at PeV EnergieS; phase-3) [2] experiment is located in the mountain valley of Ooty, Tamil Nadu in southern India (11.4^o N, 76.7^o E, 2200 m a.s.l.). The air shower detector-array of the GRAPES-3 facility is a sprawling network of detectors designed to capture the cascades of secondary particles generated when CRs interact with the Earth's atmosphere. These detectors are deployed in a symmetric hexagonal geometry with different (X_i , Y_i , Z_i ; for i^{th} detector) coordinates. The meticulously designed EAS array consisting of 395 scintillators [1], each 1 m^2 in area with inter-detector separation of only 8 m (Fig. 1). The second major component of the GRAPES-3 experiment is the 560 m^2 GRAPES-3 muon detector that contains 16 tracking modules (each 35 m^2 in area and energy threshold of 1 GeV for vertical muons) [2], which provides reliable measurement of the muon size (Fig. 1).

FIG. 1: The GRAPES-3 detector array with a varying elevation, showing muon detectors module in the left and data acquisition center in the middle of the figure. The figure on the right represents the schematic layout of the detector array.

III. EAS SIMULATION

The simulation of EAS events involves a coupling between two distinct hadronic interaction models within the CORSIKA Monte Carlo (MC) program version 7.690 framework [4]. Specifically, the high-energy regime (above 80 GeV/n) is governed by the QGSJet 01 version 1c model [5], while the UrQMD model [6] characterizes the lower energy section (below 80 GeV/n). For the electromagnetic component, the EGS4 program library is employed. The EAS events are simulated to mirror the geographical location of the GRAPES-3 experiment. The kinetic energy thresholds for muons and electrons are set at 0.3 and 0.003 GeV. About 20000 showers are simulated for each of the proton and iron primaries, spanning primary energy ranges between 8×10^{14} eV and 3×10^{15} eV, with zenith angles constrained below 25°. In addition to this comprehensive data set, a numerous number of additional showers are simulated for each mentioned primary type, encompassing various specific primary energy values and zenith angles in order to broaden our quest.

IV. ESTIMATION OF THE EAS CORE

By default, CORSIKA simulates the core of each EAS at the centre of the detector plane (0,0). However, in reality, EAS cores are randomly distributed across the detector array. To reflect this reality in our shower simulations, we deliberately employ random numbers to distribute the EAS cores arbitrarily within the red hexagon [1] (refer to Fig. 1), which closely mimics the real-world scenarios. These cores are represented by coordinates (x_r, y_r). Subsequently, the core position of each shower is estimated as (x_e, y_e) by calculating the weighted mean of the first nineteen selected detectors with the highest-density of electron counts (ρ_i)

$$(x_e, y_e) = (\frac{\sum X_i \rho_i^2}{\sum \rho_i^2}, \frac{\sum Y_i \rho_i^2}{\sum \rho_i^2}); i = 1, ..19. \quad (1)$$

The core position of each shower is estimated considering *plane detector* configuration (i.e. taking $Z_i = 0$ for each detector) as well as for *actual detector* configuration (i.e. taking $Z_i \neq 0$ and the modified coordinate of each detector for varying elevation is $(X_i^Z, Y_i^Z) = (X_i + Z_i \tan\Theta \cos\phi, Y_i + Z_i \tan\Theta \sin\phi)$, where ϕ is the polar angle of detector plane, and Θ is the zenith angle of an EAS).

The error in the core estimation is determined as follows:

$$\Delta r = \sqrt{(x_r - x_e)^2 + (y_r - y_e)^2}. \quad (2)$$

V. RESULTS AND DISCUSSIONS

The WAM estimates the EAS core quite accurately for hexagonal symmetric density distribution for the detector configuration of the GRAPES-3 experiment centre. Here, hexagonal symmetric density distribution represents a scenario for the energy deposition of an EAS in the detector array in such a manner that there is a highest density detector surrounded by six next highest-density detectors inside the first ring, which are surrounded by twelve next highest-density detectors inside the second ring, and the weighted average of these nineteen density detectors represents the estimated core of the EAS. The CR showers with higher zenith angle and lower primary energy inhibit the hexagonal symmetric distribution of density of EAS electrons, thereby reducing the accuracy of the estimated core.

The primary challenges encountered in employing the WAM for the core analysis dwell around fluctuations in densities observed at the detectors, and the considerable spacing between them. Remarkably, these separations pose a significant issue with low-energy EASs, where the number of active detectors is minimal. An extensive hindrance of hexagonal symmetry in density distribution can be observed in Fig. 2 for EASs initiated by the proton with energy 10^{14} eV, and the problem rises with increasing zenith angle, thereby worsening the difficulty in estimating the EAS cores. Additionally, the highest density detectors are dispersed in such a manner that EASs exhibit multiple core structures, rendering the WAM ineffective in accurately pinpointing core locations using the mentioned nineteen highest density detectors particularly for inclined EASs initiated by low-energy primaries.

To mitigate the impact of density fluctuations, especially at lower energies, we adopted a cluster-based approach. In this method, we select the highest-density detector within a designated cluster for each shower event. Subsequently, the remaining eighteen detectors are chosen within a 25 m radial distance from this maximum density detector. The core location is then estimated utilizing the WAM, based on these chosen

nineteen highest density detectors. Through this approach, we find that for low-energy proton showers ($4 \times 10^{12}\text{eV} \leq E \leq 1 \times 10^{13}\text{eV}$), we can accurately estimate the core location for 95% of events with an uncertainty of ± 5 m. In the energy range ($3 \times 10^{14}\text{eV} \leq E \leq 7.5 \times 10^{14}\text{eV}$), we found 95% events within ± 4 m for the actual detector array. The value of Δr (in m) within which 95.5% of events are present for p and Fe initiated showers, considering actual array configuration at two different average energies (0.3 PeV and 0.5 PeV) is shown in the Table I.

TABLE I: Δr (in m) for 95.5% events considering actual detector array

PCR	Proton		Iron	
$E(PeV)$	0.3	0.5	0.3	0.5
$\Theta = 0^o$	2.27	1.58	4.52	3.47
5^o	2.32	1.88	4.62	3.67
15^o	2.67	2.27	4.62	3.57
25^o	2.97	2.62	4.67	4.32

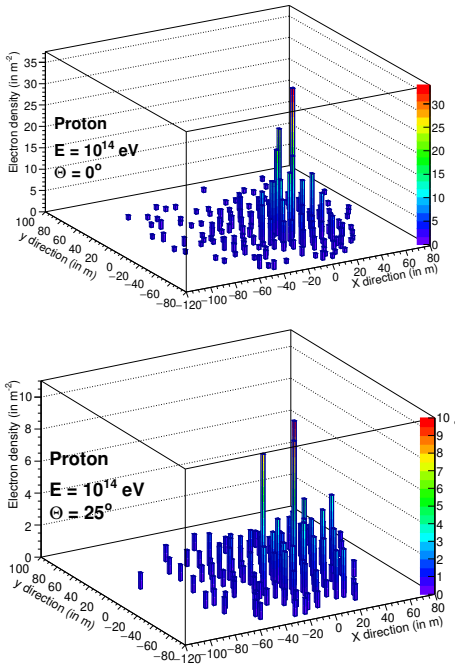

FIG. 2: Density plot for proton-initiated showers with E = 10^{14} eV and $\Theta = 0^o$ (*top*), and $\Theta = 25^o$(*bottom*).

The density fluctuations around the estimated core is minimal for high-energy showers, as illustrated in Fig. 3. Additionally, there is a noticeable contrast in density between the highest-density detector and other detectors registering hits. The cases involving high-energy showers, the strategic expansion of the cluster boundary is undertaken to envelop the whole of the detector array. This deliberate extension ensures

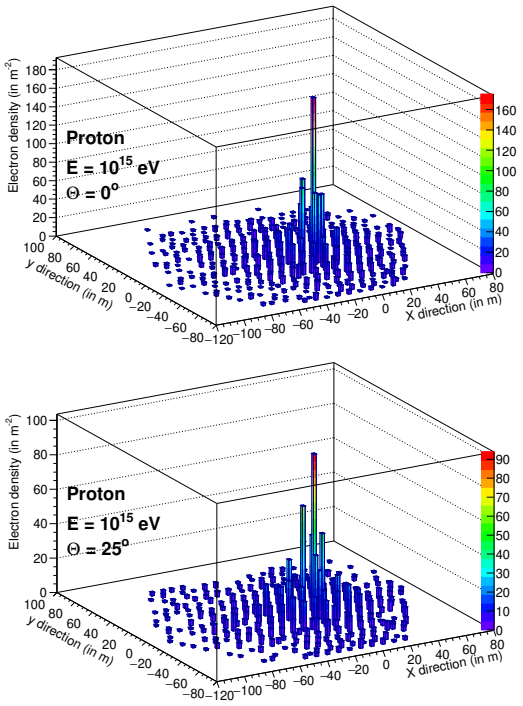

FIG. 3: Density plot for proton showers with E = 10^{15} eV and $\Theta = 0^o$ (*top*), and $\Theta = 25^o$(*bottom*).

the comprehensive selection of the 19 detectors exhibiting the highest-density. By broadening the scope of the cluster boundary to encapsulate the entire detector array, we guarantee the inclusive selection of these critical detectors, thereby focussing the pivotal attributes of the high energy shower event. Table II presents the Δr values (in m), which delineates the radial distance within which 95.5% of events occur for both p- and Fe-initiated showers. This analysis is conducted for both the plane detector array and the actual array, accounting for a geometric correction for the z-coordinate of each detectors. The Δr values are computed at three distinct average energies.

Finally, the methodology is applied to a data set comprising the entire 20000 showers initiated by each of proton and iron primaries, with primary energies falling within proximity to the knee region (8×10^{14}eV $-$ 3×10^{15}eV) and Θ not exceeding 25°. Illustrated in Fig. 4 is the population distribution of Δr for both proton and iron showers, considering the plane detector configuration of the GRAPES-3 array. Notably, the mean of this distribution stands at 0.074 m for proton and 1.094 m for iron showers. Conversely, when accounting for the actual detector configuration of the GRAPES-3 array, the mean values of the population distribution of Δr escalate to 1.094 m for proton and 1.385 m for iron showers as illustrated in Fig. 5.

TABLE II: Δr (in m) for different conditions.

PCR		Proton			Iron		
$E(PeV)$		0.8	1	3	0.8	1	3
Plane array (Δr)	$\Theta = 0^o$	1.33	1.23	1.08	2.42	1.93	1.08
	5^o	1.33	1.28	1.18	2.52	1.93	1.13
	15^o	1.58	1.33	1.28	2.72	2.82	1.28
	25^o	1.83	1.48	1.28	3.72	3.77	1.58
Actual array (Δr)	0^o	1.33	1.38	1.18	2.62	2.23	1.28
	5^o	1.43	1.38	1.23	2.57	2.32	1.28
	15^o	2.47	2.32	1.83	3.87	3.02	2.13
	25^o	3.82	3.48	3.32	4.87	4.12	3.57

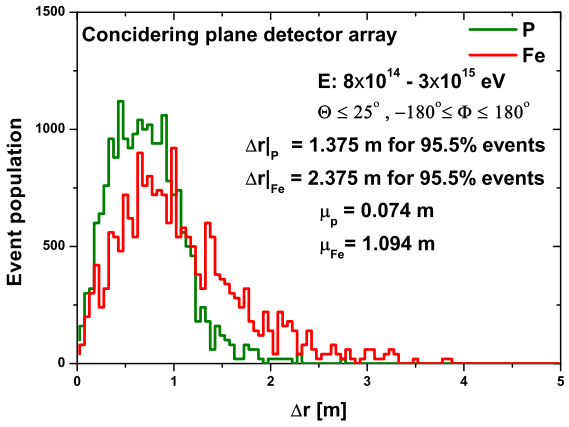

FIG. 4: Frequency distribution of Δr for the EASs produced by proton and iron primaries while considering plane array configuration of detectors.

VI. CONCLUSION

The analysis explored a methodology for accurately estimating the cores of EASs for a detector array like the GRAPES-3 experiment. We observed that the WAM proves effective, particularly in scenarios of high-energy showers with dense detector arrays. Initially, we encountered issues with density fluctuations and asymmetric density distributions over the whole detector array, prompting the exploration of novel approaches, such as cluster-based methods, to improve core estimation accuracy, especially in low-energy showers. As the energy of the primary particle increases, EASs exhibit larger shower sizes accompanied by an expanded lateral spread in density distribution, encompassing the entire configuration of the detector

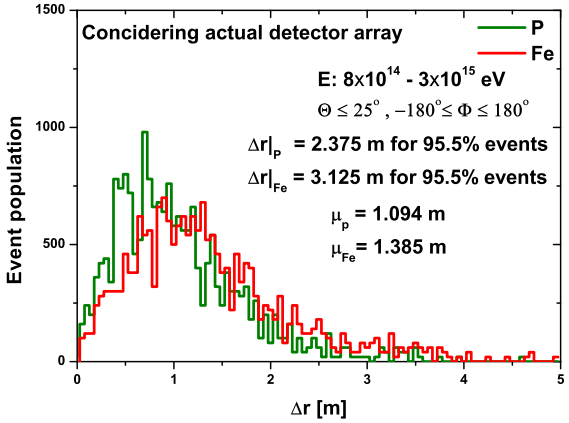

FIG. 5: Frequency distribution of Δr for the EASs produced by proton and iron primaries while considering actual array configuration of detectors.

array. Consequently, the cluster boundary is widened to encompass the entirety of the detector array, ensuring comprehensive coverage of the shower event. Notably, our findings emphasized the significance of geometric considerations and detector array configurations to accurately estimating the EAS core, which is crucial for reconstructing the showers and asociated EAS parameters.

Our investigation has used MC data generated by the CORSIKA code of version 7.690. It has been observed that the value of Δr is larger for an iron shower compared to a proton shower The iron nuclei have a larger interaction cross-section compared to protons. Consequently, the iron nuclei are more likely to interact with atmospheric particles, leading to more frequent collisions and interactions. Due to their intense atmospheric interactions, iron nuclei lose energy more rapidly as they travel through the atmosphere. The higher mass and energy of iron nuclei result in more energetic collisions with atmospheric particles, leading to a more significant number of secondary particle production. Consequently, the larger number of secondary particles in iron-initiated showers contributes to a higher attenuation effect as they propagate through the atmosphere. Protons, being lighter particles, tend to interact less frequently and penetrate deeper into the atmosphere before initiating the showering process compared to heavier nuclei like iron. These factors are responsible for smaller shower sizes for iron EASs than proton-initiated showers. Again, lower-energy primaries initiate fewer interactions and consequently produce fewer secondary particles, reducing shower size. The EASs with reduced shower sizes result in a sparser distribution of detector hits. This reduced the density of detectors due to EAS particles which may lead to more significant uncertainty in determining the precise location of the shower core, as fewer data points may be available for analysis.

The asymmetric lateral density distribution observed in the detector array for inclined extensive air showers primarily emanates from the unequal attenuation of secondary cosmic ray particles for different polar

angles in the shower front plane while propagating through the atmosphere, causing an unequal dispersion of secondary particles. Moreover, asymmetries in particle generation, atmospheric absorption, scattering processes, and the Earth's geomagnetic field together contribute to this asymmetry, resulting in enhanced uncertainty in EAS core estimation.

ACKNOWLEDGMENTS

A. Basak wants to express his sincere gratitude to Dr. R. K. Dey and Dr. A. Bhadra for their invaluable support and guidance throughout preparing this paper. Their expertise, encouragement, and constructive feedback have been instrumental in shaping the quality and direction of this research.

* Electronic address: animesh21@nbu.ac.in

[1] S. K. Gupta, *et al*, Nuclear Phys. B **196**, 153 (2009).

[2] H. Tanaka, *et al.*, Nuclear Phys. B **175**, 280 (2008).

[3] A. Basak, *et al*, DAE Symp. on Nucl. Phys. (Proc.) **65** (2021).

[4] D. Heck, *et al*, FZKA report-6019 ed. FZK Karlsruhe (1998).

[5] N. N. Kalmykov, S. S. Ostapchenko, and A. I. Pavlov, Nucl. Phys. B (Proc. Suppl.) **52**, 17 (1997).

[6] M. Bleicher *et al*, J. Phys. G: Nucl. Part. Phys. **25**, 1859 (1999).

