Machine Learning-Accelerated Computational Solid Mechanics: Application to Linear Elasticity

Rajat Arora[1]

[1]Siemens Technology
rajat.arora9464@gmail.com

Abstract

This work presents a novel physics-informed deep learning based super-resolution framework to reconstruct high-resolution deformation fields from low-resolution counterparts, obtained from coarse mesh simulations or experiments. We leverage the governing equations and boundary conditions of the physical system to train the model without using any high-resolution labeled data. The proposed approach is applied to obtain the super-resolved deformation fields from the low-resolution stress and displacement fields obtained by running simulations on a coarse mesh for a body undergoing linear elastic deformation. We demonstrate that the super-resolved fields match the accuracy of an advanced numerical solver running at 400 times the coarse mesh resolution, while simultaneously satisfying the governing laws. A brief evaluation study comparing the performance of two deep learning based super-resolution architectures is also presented.

1 Introduction

Image super-resolution (SR) is an active area of research in the field of computer science which aims at recovering high-resolution (HR) image from a low-resolution (LR) image. In this work, we focus on exploring the concept of image super-resolution to develop a physics-informed Deep Learning (DL) model to reconstruct HR deformation fields (stress and displacements) from LR fields without requiring any HR labeled data. The LR data could be obtained by running simulations on a coarse mesh or from experiments such as digital image correlation. We also present a brief study that compares two DL architectures and evaluate their suitability for developing physics-informed super-resolution framework. The overall schematic of the proposed physics-informed strategy for super resolution is depicted in Fig. 1. The use of such physics-informed SR framework will allow researchers to solve computationally expensive simulations much faster and enable them to increase accuracy without additional costs.

Recently, several researchers have explored the possibility of using deep learning based super-resolution to reconstruct HR fluid flow fields from LR (possibly noisy) data. The data-driven approaches for reconstructing HR flow field (Fukami, Fukagata, and Taira 2021, 2019; Deng et al. 2019; Bode et al. 2019; Xie et al. 2018) relies on the availability of large amount of HR labeled data. Moreover, the HR output obtained from data-driven approaches may fail to satisfy physics-based constraints because of the lack of any embedded physical constraints in the model itself. Several studies have demonstrated the merits of developing physics-informed DL models for SR in the fluid mechanics community (Esmaeilzadeh et al. 2020; Subramaniam et al. 2020; Sun and Wang 2020; Gao, Sun, and Wang 2021). However, to the best of author's knowledge, developing an effective physics-informed DL model for super-resolution in label-scarce or label-free scenarios for solid mechanics problems has not yet been explored.

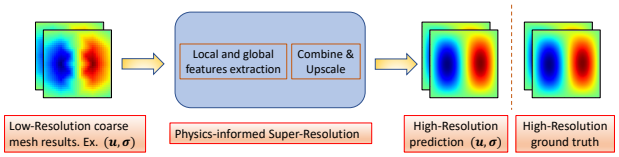


Figure 1: Schematic of the super-resolution framework.

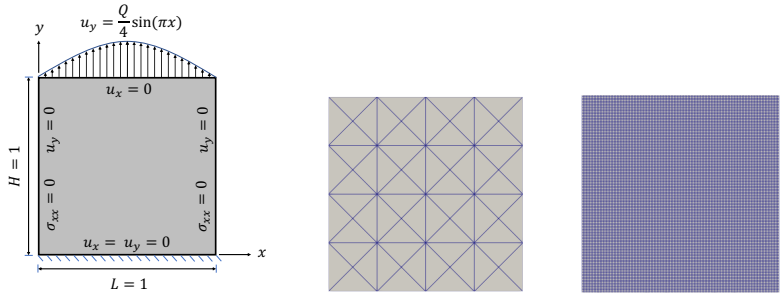
The layout of the rest of this paper is as follows: In Sec. 2, a brief review of the governing equations for modeling elastic deformation in solids is presented. Model architectures and construction of loss function are discussed in Secs. 3 and 4, respectively. Sec 5 presents the findings for the evaluation of the proposed SR-framework after briefly discussing simulation setup and data collection strategy. Conclusions and future opportunities are presented in Sec. 6.

2 Governing Equations for Elasticity

The governing equations for elasticity problems, in the absence of inertial forces, are given as follows:

$$\begin{aligned} Div\boldsymbol{\sigma} + \boldsymbol{B} &= \boldsymbol{0}, \quad \text{in } \Omega, \\ \boldsymbol{\sigma} = \mathbb{C} : \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} &= \frac{1}{2}\left(\boldsymbol{\nabla u} + (\boldsymbol{\nabla u})^T\right), \\ \boldsymbol{\sigma n} = \boldsymbol{t}_{bc} \ \text{ on } \ \partial\Omega_N \ \text{ and } \ \boldsymbol{u} &= \boldsymbol{u}_{bc} \ \text{ on } \ \partial\Omega_D. \end{aligned} \tag{1}$$

In the above, $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$ denotes the stress and the (linearized) strain in the material. $\boldsymbol{u}$ and $\boldsymbol{B}$ denotes the displacement vector and body force vector (per unit volume), respectively. Ω denotes the volumetric domain, Div denotes the divergence operator, and $\mathbb{C}$ is the fourth order elasticity tensor. $\boldsymbol{t}_{bc}$



(a) (b) (c)

Figure 2: a) Schematic showing the geometry and the applied boundary conditions. b) Coarse triangular mesh with 41 nodes. c) 128×128 fine mesh with 16384 nodes. The LR data is refined by 400 times.

and $\boldsymbol{u}_{bc}$ denote the known traction and displacement vectors on the (non-overlapping) parts of the boundary $\partial\Omega_N$ and $\partial\Omega_D$, respectively. $\boldsymbol{n}$ denotes the unit outward normal on the external boundary $\partial\Omega$. Under two-dimensional plane-strain conditions, the unknown components for displacement vector and stress tensor are (u_x, u_y) and $(\sigma_{xx}, \sigma_{yy}, \sigma_{xy})$, respectively.

3 Model Architectures

We train physics-informed DL framework to approximate the mapping $\Psi : \mathcal{I}^{LR} \rightarrow \mathcal{I}^{HR}$ to reconstruct the HR deformation field ($\mathcal{I}^{HR}$) from the LR ($\mathcal{I}^{LR}$) data. The two architectures evaluated in this study are i) Residual Dense Network (RDN) (Zhang et al. 2018), and ii) FSRCNN (Dong, Loy, and Tang 2016). In this work, we use the following hyper-parameters for the RDN model: number of residual blocks: 2, number of layers in each residual block: 4, growth rate: 32, and number of features: 32. For the FSRCNN model, we use the following hyper-parameters: number of layers: 8, and LR feature dimensions $d = 128$ and $s = 64$. These hyper-parameters also ensure that both the models have (almost) same number of trainable parameters. The inputs to both the models consist of LR data $\{u_x, u_y, \sigma_{xx}, \sigma_{yy}, \sigma_{xy}\}$ obtained by running simulations on a coarse mesh (see Fig. 2) and then evaluating the solution (using underlying interpolating basis functions) on a 32×32 structured grid. The outputs of these models correspond to the HR data on a 128×128 structured grid as shown in Figure 2.

4 Constructing the Loss Function

For the unsupervised model, wherein the HR labeled data is not needed, the total network loss $\mathcal{L}$ is obtained only from the physics-based constraints corresponding to the governing equations and boundary conditions. For the mixed-variable formulation (displacement vector $\boldsymbol{u}$ and stress tensor $\boldsymbol{\sigma}$ as outputs), the total loss function $\mathcal{L}$ is constructed as follows

$$\begin{aligned}\mathcal{L} =& \lambda_1 \underbrace{||\boldsymbol{\nabla} \cdot \boldsymbol{\sigma}||}_{\text{PDE}} + \lambda_2 \underbrace{||\boldsymbol{\sigma} - \mathbb{C} : \boldsymbol{\epsilon}||}_{\text{Constitutive law}} \\ &+ \lambda_3 \underbrace{||\boldsymbol{u} - \boldsymbol{u}_{bc}||_{\partial\Omega_U}}_{\text{Dirichlet BC}} + \lambda_4 \underbrace{||\boldsymbol{\sigma}\boldsymbol{n} - \boldsymbol{t}_{bc}||_{\partial\Omega_N}}_{\text{Neumann BC}},\end{aligned} \tag{2}$$

where $||(\cdot)||$ denotes the L^1 norm of the quantity $(\cdot)$. L^1 norm is chosen to make the model robust to noise and outliers in the LR data. The scalar constants λ_1, λ_2, λ_3, and λ_4 are chosen to nondimensionalize the individual loss components. In this work, we choose $\lambda_1 = \frac{H}{\mu}$, $\lambda_2 = \frac{1}{\mu}$, $\lambda_3 = \frac{20}{U_0}$, and $\lambda_4 = \frac{20}{\mu}$, where μ and U_0 represent the shear modulus and the characteristic displacement in the body, respectively. H denotes the height of the body. Relatively larger magnitudes of λ_3 and λ_4 are chosen to assign more weight to the boundary conditions.

Two neural networks based on RDN and FSRCNN architectures are implemented and trained using PyTorch framework (Paszke et al. 2019). The network's total loss $\mathcal{L}$ is minimized by iteratively updating trainable parameters by using Adam optimizer (Kingma and Ba 2015) for around 2000 epochs with learning rate $\eta = 10^{-4}$. We also use `ReduceLROnPlateau` scheduler with the `patience` = 30. L-BFGS algorithm (Zhu et al. 1997) is then used for local fine-tuning of the solution until loss converges. The training is performed using NVIDIA Quadro RTX 8000 graphics card and takes around 8 and 14 hours for RDN and FSRCNN models, respectively. The source code and the dataset used in this research can be found at https://github.com/sairajat/SR_LE upon acceptance of this paper.

5 Results & Discussion

In what follows, we demonstrate the effectiveness of SR framework in reconstructing HR displacement and stress fields from LR input data for linear elastic simulations – which we believe is a first step in demonstrating the strength of machine-learned super-resolution techniques in solid mechanics.

We apply the framework to resolve the stress and displacement fields within an isotropic body undergoing linear elastic deformation. The schematic of the body along with

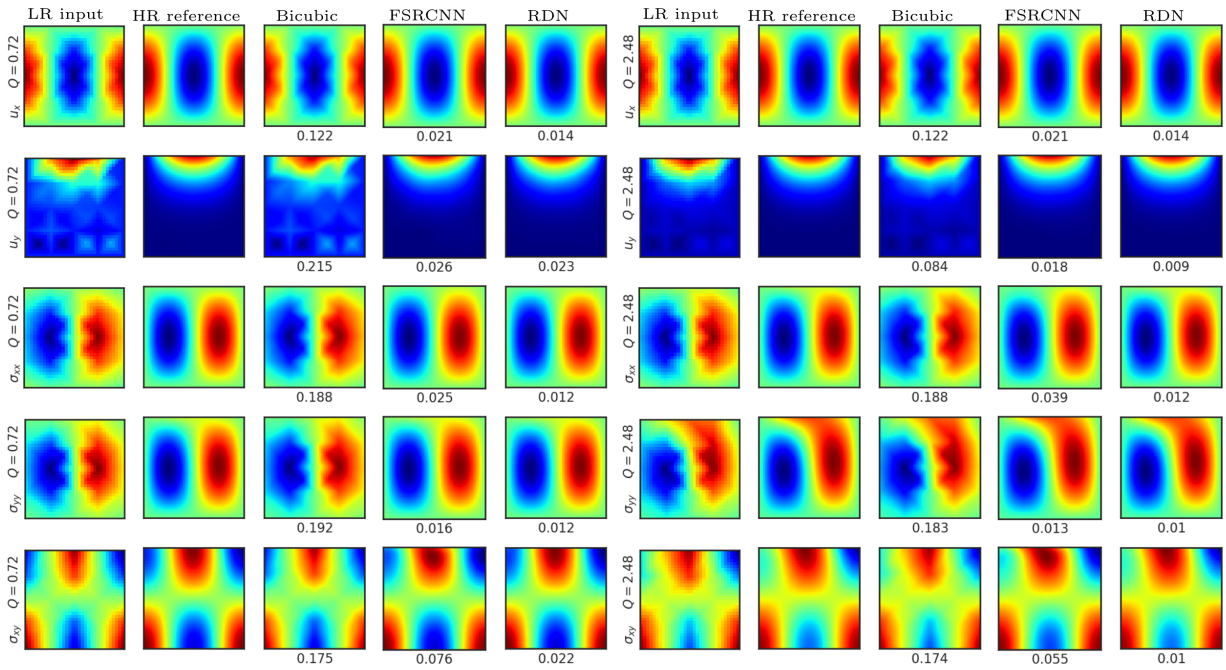


Figure 3: The color contours of displacement vector and stress tensor components in two-dimensional elastic deformation reconstructed with physics-informed super-resolution frameworks. Values below the plots indicate the L_2 error e. In both the blocks, the LR input data, HR ground truth data, bicubic interpolation, FSRCNN output, and the RDN output are plotted from the left to the right.

the boundary conditions is shown in Figure 2. The body is assumed to deform quasi-statically under plane strain conditions with the body force vector $\boldsymbol{B}$ given as

$$\begin{aligned} B_x &= \lambda \left[4\pi^2 \cos(2\pi x)\sin(\pi y) - \pi\cos(\pi x)Qy^3\right] \\ &\quad + \mu \left[9\pi^2 \cos(2\pi x)\sin(\pi y) - \pi\cos(\pi x)Qy^3\right], \\ B_y &= \lambda \left[2\pi^2 \sin(2\pi x)\cos(\pi y) - 3\sin(\pi x)Qy^2\right] \\ &\quad + \mu \left[-6\sin(\pi x)Qy^2 + 2\pi^2 \sin(2\pi x)\cos(\pi y)\right. \\ &\quad \left. + 0.25\pi^2 \sin(\pi x)Qy^4\right]. \end{aligned}$$

In this work, the material constants λ and μ are taken to be 1 and 0.5, respectively. The quantity $Q \in [0, 4]$ affects the boundary conditions (see Fig. 2) and the body force $\boldsymbol{B}$. The characteristic displacement U_0 is taken to 1 (maximum value of u_y on the top boundary). The ground truth data is generated by solving the system of equations (1) on a coarse mesh (shown in Fig. 2) using Finite Element Method for Q regularly sampled at an interval of 0.04. The data is then randomly split in a $80:20$ ratio for training and test purposes.

The framework super-resolves the deformation fields onto the 128×128 mesh, shown in Fig. 2, which is ≈ 400 times finer than the coarse mesh used to obtain the LR data. The HR outputs for each model are obtained by doing a forward propagation through the corresponding trained models. For comparison, along with HR labeled data, we also utilize a simple bicubic interpolation of fields. We note that the HR labeled data is used only for the comparison with the predicted outputs.

Figure 3 presents the results for the reconstructed displacement and stress fields for 2 different values of Q for both the models. We can see that the both the frameworks are able to super-resolve all the deformation fields with great accuracy as the plots show great agreement with the reference HR ground truth data. To qualitatively measure the accuracy, we define a relative error measure as $e = \frac{||\mathcal{I}^{HR} - \mathcal{I}^{LR}||_{L^2}}{||\mathcal{I}^{HR}||_{L^2}}$. The values of e are reported underneath the reconstructed fields obtained using the SR frameworks and the bicubic interpolation. As can be seen, the error is largest for the bicubic interpolated data as compared to both the physics-informed models. This is expected since the interpolated data may not faithfully satisfy the governing laws of the system. We also notice that the error is larger for FSRCNN based model as compared to RDN based model. The reconstructed HR outputs obtained from the RDN based model almost match the accuracy of an advanced numerical solver running at 400 times the coarse mesh resolution. We believe that the better accuracy for the RDN model results from the use of residual connections and smaller kernel sizes during convolution and upsampling operations in its architecture. This validates the concept that a deep-learning based physics-informed SR framework can be used to faithfully reconstruct the fields at a higher resolution while simultaneously satisfying the governing laws. We note that the proposed physics-informed SR strategy can be easily extended to non-rectangular domains (Gao, Sun, and Wang 2020) or account for boundary conditions in a *hard* manner (Rao,

Sun, and Liu 2021).

6 Conclusion

In summary, we successfully trained and evaluated two physics-informed super-resolution frameworks based on Residual Dense Network (Zhang et al. 2018) and FSRCNN (Dong, Loy, and Tang 2016) architectures to super-resolve the deformation fields in a body undergoing elastic deformation. Among the two deep learning architectures evaluated in this work, we show that the framework based on RDN is more accurate and matches the accuracy of an advanced numerical solver running at 400 times the coarse mesh resolution (see Figs. 2 and 3). The approach is successfully able to learn high-resolution spatial variation of displacement and stress fields from their low-resolution counterparts for the linear elastic case discussed. These advantages are possible due to the combined effect of two rapidly evolving research areas - Physics informed neural networks (Raissi, Perdikaris, and Karniadakis 2017, 2019) and computer vision (Voulodimos et al. 2018). We emphasize that the current work focuses on the demonstration of feasibility of the concept while the assessment of potential computational advantages, including the extension to hyperelastic deformation, is deferred to future research.

The approach exemplifies how machine-learning can be leveraged to conduct such mechanical calculations for materials with complex constitutive response (eg. dislocation mediated plastic deformation and fracture modeling (Arora, Zhang, and Acharya 2020; Nielsen and Niordson 2019; Niordson and Tvergaard 2019; Arora and Acharya 2020a; Yingjun et al. 2016; Arora 2019; Kuroda and Tvergaard 2008; Lynggaard, Nielsen, and Niordson 2019; Arora and Acharya 2020b; Borden et al. 2014)) to reduce the computational complexity and accelerate scientific discovery and engineering design.

Acknowledgments

This work was conceptualized during the author's time at Carnegie Mellon University (CMU). The author thank Ankit Shrivastava, Ph.D. candidate at CMU, for useful discussions and comments on the manuscript.

References

Arora, R. 2019. *Computational Approximation of Mesoscale Field Dislocation Mechanics at Finite Deformation*. Ph.D. thesis, Carnegie Mellon University.

Arora, R.; and Acharya, A. 2020a. Dislocation pattern formation in finite deformation crystal plasticity. *International Journal of Solids and Structures*, 184: 114–135.

Arora, R.; and Acharya, A. 2020b. A unification of finite deformation J2 Von-Mises plasticity and quantitative dislocation mechanics. *Journal of the Mechanics and Physics of Solids*, 143: 104050.

Arora, R.; Zhang, X.; and Acharya, A. 2020. Finite element approximation of finite deformation dislocation mechanics. *Computer Methods in Applied Mechanics and Engineering*, 367: 113076.

Bode, M.; Gauding, M.; Lian, Z.; Denker, D.; Davidovic, M.; Kleinheinz, K.; Jitsev, J.; and Pitsch, H. 2019. Using physics-informed super-resolution generative adversarial networks for subgrid modeling in turbulent reactive flows. *arXiv preprint arXiv:1911.11380*.

Borden, M. J.; Hughes, T. J.; Landis, C. M.; and Verhoosel, C. V. 2014. A higher-order phase-field model for brittle fracture: Formulation and analysis within the isogeometric analysis framework. *Computer Methods in Applied Mechanics and Engineering*, 273: 100–118.

Deng, Z.; He, C.; Liu, Y.; and Kim, K. C. 2019. Super-resolution reconstruction of turbulent velocity fields using a generative adversarial network-based artificial intelligence framework. *Physics of Fluids*, 31(12): 125111.

Dong, C.; Loy, C. C.; and Tang, X. 2016. Accelerating the super-resolution convolutional neural network. In *European conference on computer vision*, 391–407. Springer.

Esmaeilzadeh, S.; Azizzadenesheli, K.; Kashinath, K.; Mustafa, M.; Tchelepi, H. A.; Marcus, P.; Prabhat, M.; Anandkumar, A.; et al. 2020. MeshfreeFlowNet: a physics-constrained deep continuous space-time super-resolution framework. In *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, 1–15. IEEE.

Fukami, K.; Fukagata, K.; and Taira, K. 2019. Super-resolution analysis with machine learning for low-resolution flow data. In *11th International Symposium on Turbulence and Shear Flow Phenomena, TSFP 2019*.

Fukami, K.; Fukagata, K.; and Taira, K. 2021. Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows. *Journal of Fluid Mechanics*, 909.

Gao, H.; Sun, L.; and Wang, J.-X. 2020. PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parametric PDEs on irregular domain. *arXiv e-prints*, arXiv–2004.

Gao, H.; Sun, L.; and Wang, J.-X. 2021. Super-resolution and denoising of fluid flow using physics-informed convolutional neural networks without high-resolution labels. *Physics of Fluids*, 33(7): 073603.

Kingma, D. P.; and Ba, J. 2015. Adam: A Method for Stochastic Optimization. *ICLR*.

Kuroda, M.; and Tvergaard, V. 2008. A finite deformation theory of higher-order gradient crystal plasticity. *Journal of the Mechanics and Physics of Solids*, 56(8): 2573–2584.

Lynggaard, J.; Nielsen, K. L.; and Niordson, C. F. 2019. Finite strain analysis of size effects in wedge indentation into a Face-Centered Cubic (FCC) single crystal. *European Journal of Mechanics / A Solids*, 76: 193–207.

Nielsen, K. L.; and Niordson, C. F. 2019. A finite strain FE-Implementation of the Fleck-Willis gradient theory: Rate-independent versus visco-plastic formulation. *European Journal of Mechanics / A Solids*, 75: 389–398.

Niordson, C. F.; and Tvergaard, V. 2019. A homogenized model for size-effects in porous metals. *Journal of the Mechanics and Physics of Solids*, 123: 222–233.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga,

L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; and Chintala, S. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Wallach, H.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 32*, 8024–8035. Curran Associates, Inc.

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2017. Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. *arXiv preprint arXiv:1711.10561*.

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378: 686–707.

Rao, C.; Sun, H.; and Liu, Y. 2021. Physics-Informed Deep Learning for Computational Elastodynamics without Labeled Data. *Journal of Engineering Mechanics*, 147(8): 04021043.

Subramaniam, A.; Wong, M. L.; Borker, R. D.; Nimmagadda, S.; and Lele, S. K. 2020. Turbulence enrichment using physics-informed generative adversarial networks. *arXiv preprint arXiv:2003.01907*.

Sun, L.; and Wang, J.-X. 2020. Physics-constrained bayesian neural network for fluid flow reconstruction with sparse and noisy data. *Theoretical and Applied Mechanics Letters*, 10(3): 161–169.

Voulodimos, A.; Doulamis, N.; Doulamis, A.; and Protopapadakis, E. 2018. Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, 2018.

Xie, Y.; Franz, E.; Chu, M.; and Thuerey, N. 2018. tempogan: A temporally coherent, volumetric gan for super-resolution fluid flow. *ACM Transactions on Graphics (TOG)*, 37(4): 1–15.

Yingjun, G.; Zhirong, L.; Lilin, H.; Hong, M.; Chuanggao, H.; and Kui, L. 2016. Phase field crystal study of nano-crack growth and branch in materials. *Modelling and Simulation in Materials Science and Engineering*, 24(5): 055010.

Zhang, Y.; Tian, Y.; Kong, Y.; Zhong, B.; and Fu, Y. 2018. Residual dense network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2472–2481.

Zhu, C.; Byrd, R. H.; Lu, P.; and Nocedal, J. 1997. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on mathematical software (TOMS)*, 23(4): 550–560.