Formatting Instructions For NeurIPS 2022

Tianyu Xie
School of Mathematical Sciences
Peking University
xietianyu@pku.edu.cn

Abstract

The abstract paragraph should be indented ½ inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

1 Simulation Studies

Although the authors of ARTICLE had done experiments to verify the validity of their propoesd methods, we will repeat their experiments in this review. In this section, we aim to evaluate performance of of various estimators discussed earlier in the article using artificial data, following the same settings in ARTICLE.

Data is generated from the following models:

$$\begin{cases} \delta^{\mathrm{L}}(X) = \tanh\left(\alpha^{\mathrm{T}} X\right) \\ \delta^{\mathrm{M}}(X) = \exp\left(\alpha^{\mathrm{T}} X\right) \\ \phi_i(X) = \operatorname{sigmoid}\left(\beta_i^{\mathrm{T}} X\right), i = 1, \ldots, 4 \\ \mathrm{OP}^{\mathrm{CO}}(X) = \exp\left(\eta^{\mathrm{T}} X\right) \\ \Pr(Z = 1 \mid X) = \operatorname{sigmoid}\left(\gamma^{\mathrm{T}} X\right) \end{cases} \tag{1}$$

where we set $\alpha = (0, -1)^{\mathrm{T}}, \beta_i = (-0.4, 0.8)^{\mathrm{T}}, i = 1, \ldots, 4, \eta = (-0.4, 1)^{\mathrm{T}}$ and $\gamma = (0.1, -1)^{\mathrm{T}}$. The covariates X has two dimensions, including an intercept and another covariate generated from Uniform$(-1, 1)$. Our goal is to estimate α and then estimate the local average treatment effect, while other nuisance parameters are estimable but of no interest.

Similar to ARTICLE, we will fit models of the same functional form as in (1), although this is impossible in practical applications. We repeat the simulation studies of

- `mle`: the maximum likelihood estimator in ARTICLE;
- `drw`: the doubly robust estimator with the optimal weighting function in ARTICLE;
- `dru`: the doubly robust estimator with the identity weighting function in ARTICLE;

and cite the reported results of

- `reg.ogburn`: the outcome regression estimator in Ogburn et al. (2015, ğ 3.1);
- `drw.ogburn`: the doubly robust estimator in Ogburn et al. (2015, ğ 3.3) with the optimal weighting function;
- `dru.ogburn`: the doubly robust estimator in Ogburn et al. (2015, ğ 3.2) with the identity weighting function;
- `mle.wang`: the maximum likelihood estimator in Wang & Tchetgen Tchetgen (2018, ğ 4.1);

	LATE		MLATE	
Method	α_0	α_1	α_0	α_1
`mle.bth`	4.55 (0.44)	9.31 (0.93)	13.81 (0.81)	18.33 (1.32)
`mle.opc`	5.86 (0.48)	11.94 (0.95)	16.47 (0.85)	21.11 (1.35)
`mle.psc`	5.86 (0.48)	11.94 (0.95)	16.47 (0.85)	21.11 (1.35)
`mle.bad`	19.15 (0.69)	2.80 (0.62)	41.28 (1.46)	0.53 (1.05)
`dru.bth`	0.95 (0.45)	6.16 (1.04)	3.58 (1.00)	9.75 (1.79)
`dru.opc`	1.21 (0.43)	4.55 (0.97)	0.02 (1.05)	10.78 (1.99)
`dru.psc`	1.21 (0.43)	4.55 (0.97)	0.02 (1.05)	10.78 (1.99)
`dru.bad`	15.25 (0.70)	29.99 (1.80)	25.23 (1.43)	17.30 (2.56)
`drw.bth`	0.29 (1.04)	1.55 (1.20)	3.04 (1.24)	4.75 (1.77)
`drw.opc`	3.71 (1.27)	0.63 (1.45)	0.30 (1.10)	6.39 (1.78)
`drw.psc`	3.71 (1.27)	0.63 (1.45)	0.30 (1.10)	6.39 (1.78)
`drw.bad`	10.97 (0.56)	13.68 (1.19)	29.54 (1.65)	4.18 (2.59)

Table 1: bias $\times$ 100 (standard error $\times$ 100) of parameter estimation under different scenarios (repeated by us)

	LATE		MLATE	
Method	α_0	α_1	α_0	α_1
`mle.bth`	0.28 (0.35)	-3.5 (0.78)	-0.092 (0.71)	-3.0 (1.2)
`mle.bad`	-20 (0.42)	-15 (0.80)	-48 (1.2)	-18 (2.1)
`drw.bth`	0.55 (0.36)	-4.1 (0.82)	-0.54 (0.77)	-5.6 (1.5)
`drw.psc`	0.060 (0.38)	-5.9 (1.0)	-0.38 (1.2)	-12 (2.7)
`drw.opc`	0.55 (0.36)	-3.9 (0.79)	0.49 (0.75)	-5.3 (1.4)
`drw.bad`	-10 (0.40)	-9.6 (1.1)	-28 (1.4)	25 (3.3)
`dru.bth`	1.3 (0.44)	-5.8 (1.0)	1.8 (0.84)	-8.1 (1.7)
`reg.ogburn.bth`	-5.7 (1.6)	-2.9 (3.1)	1.8 (0.84)	-8.1 (1.7)
`reg.ogburn.bad`	-9.0 (0.25)	100 (0.23)	140 (5.6)	93 (3.6)
`drw.ogburn.bth`	0.10 (0.46)	-4.2 (0.99)	3.2 (1.4)	-13 (2.5)
`dru.ogburn.bth`	1.3 (0.45)	-5.8 (1.1)	1.9 (0.85)	-8.2 (1.7)
`dru.wang.bth`	1.3 (0.45)	-5.8 (1.1)	-	-
`ls.abadie.bth`	-0.19 (0.37)	-4.1 (0.93)	0.42 (0.79)	-11 (1.6)
`ls.abadie.bad`	-23 (0.88)	22 (1.2)	-32 (1.9)	7.7 (3.6)
`mle.crude`	-2.8 (0.10)	60 (0.19)	0.36 (0.25)	51 (0.42)

Table 2: bias $\times$ 100 (standard error $\times$ 100) of parameter estimation under different scenarios (reported by ARTICLE)

- `dru.wang`: the doubly robust estimator in Wang & Tchetgen Tchetgen (2018, ğ 4.4) with the identity weighting function;
- `ls.abadie`: the least squares estimator of Abadie (2003, ğ 4.2.1);
- `mle.crude`: the maximum likelihood estimator of the crude association on the additive or multiplicative scale (Richardson et al., 2017, ğ 2).

We also investigate the performance of these methods in misspecified scenarios. Consider two misspecified covariates: $X^{\dagger}$ includes an intercept and another covariate generated from Uniform$(-1, 1)$ independent of X; and X' includes $(\underbrace{1, \ldots, 1}_{0.5n}, \underbrace{0 \ldots, 0}_{0.5n})^{\mathrm{T}}$, $(\underbrace{0, \ldots, 0}_{0.1n}, \underbrace{1, \ldots, 1}_{0.9n})^{\mathrm{T}}$. We consider the following four misspecified scenarios, while the model of $\theta(X)$ is always correctly specified.

- `bth`: X is used in all nuisance models;
- `psc`: X is used in the instrumental density model, and X' is used in other nuisance models;
- `opc`: $X^{\dagger}$ is used in the instrumental density model, and X is used in other nuisance models;
- `bad`: $X^{\dagger}$ is used in the instrumental density model, and X' is used in other nuisance models.

Results repeated by us and reported by ARTICLE are presented in Table 1 and Table 2 respectively. Firstly, our results is different from ARTICLE, but they are roughly of the same magnitude. The

	LATE		MLATE	
Method	α_0	α_1	α_0	α_1
`mle.bth`	95.6	95.8	95.4	96.4
`mle.bad`	65.4	91.0	46.3	94.6
`drw.bth`	94.7	95.2	96.9	95.9
`drw.psc`	95.4	95.5	97.6	97.4
`drw.opc`	95.0	95.6	96.1	96.1
`drw.bad`	87.0	95.3	91.8	96.8
`dru.bth`	94.5	94.6	96.3	96.9
`reg.ogburn.bth`	98.0	99.9	99.6	100.0
`reg.ogburn.bad`	75.6	0.1	99.9	86.1
`drw.ogburn.bth`	97.0	98.1	98.5	98.4
`dru.ogburn.bth`	94.5	95.0	96.2	97.2
`dru.wang.bth`	94.4	94.7	-	-
`dru.simple.bth`	94.3	94.8	96.1	96.5
`ls.abadie.bth`	94.8	95.5	96.4	95.6
`ls.abadie.bad`	87.3	94.5	93.1	95.7
`mle.crude`	84.3	0.0	94.0	6.3

Table 3: Coverage probabilities $(\times 100)$ of confidence interval obtained from 500 bootstrap samples in selected scenarios (reported by ARTICLE)

reason for this difference is that we use different optimization methods and the covariates $X, X^\dagger$ are generated randomly. Secondly, in Table 2, the estimators `mle`, `drw`, `dru` generally have small bias and small standard error relative to other methods when all nuisance models are correctly specified, which vefiries the validity and efficiency of the method proposed by ARTICLE. Thirdly, in Table 2, we can also find that the performances of `dru.ogburn.bth`, `dru.wang.bth`, `dru.bth` are similar and less efficient than `drw.bth`, which implies that estimators using the optimal weighting function is more efficient than estimators using identity weighting function. At last, in Table 1, the two doubly robust estimators `drw` and `dru` performs worser than `mle` in the misspecification scenarios, which suggests the weighting function may yields less efficient estimations when model is misspecified. This phenomenon is not stated by ARTICLE, and they only partly reported their results.

Table 3 is the selected coverage probabilities of 95% confidence intervals obtained from the quantile bootstrap based on 500 bootstrap samples, reported by ARTICLE. We do not repeat the experiements of coverage probabilities since it cost too much time. Firstly, `mle` is less efficient in the misspecification scenarios when estimating the intercept α_0; in these scenarios, the doubly robust estimator with optimal weighting function `drw` seems better than `mle`. ARTICLE did not report the results of `dru` and did not explain why. Secondly, `mle`, `drw` and `dru` have higher converage probabilities than other methods, except for `reg.ogburn` and `drw.ogburn`. The authors of ARTICLE guessed it is due to misspecification of variation-dependent models.

2 Experiments on 401(k) Data

To explore the performances of estimators proposed by ARTICLE as well as other compared estimators, we use them to analyze the 401(k) data. According to Wikipedia, in the United States, a 401(k) plan is an employer-sponsored, defined-contribution, personal pension (savings) account, as defined in subsection 401(k) of the U.S. Internal Revenue Code. Collected by Employee Benefit Research Institute (EBRI) and the Investment Company Institute (ICI), 401(k) is the largest, most representative repository of information about individual 401(k) plan participant accounts. To be more specific, we aim to investigate whether the participation in 401(k) will reduce the paticipation in another retirement plan - Individual Retirement Accounts.

There may exists unobserved individual-level confounders such as the consuming willingness which influence the participation in 401(k) and Individual Retirement Accounts, thus directly estimating the treatment effect may yield invalid causal inferences. Therefore, we might wonder a instrumental variable irrelevant to the unobserved individual confounders. Following Ababie (2003), we choose eligibility for 401(k) as an instrument variable. Determined by companies, eligibility is weakly relevant to the individual-level variables. The monotonicity and instrumental variable relevance assumptions holds since only those eligible for 401(k) may participate in it.

ARTICLE argued that eligibility for 401(k) has an impact on participation in Individual Retirement Accounts only through participation in 401(k) plans. However, it seems that this argument is problematic since the employees who is eligible for 401(k) is more likely to be eligible for Individual Retirement Accounts. In our review, we will set this problem aside, and repeat the real data studies in ARTICLE using the dataset provided by Abadie (2003). In the following, we will use '401(k) data' to refer to this dataset.

401(k) data consists of 9275 individuals from the Survey of Income and Program Participation of 1991. It has the following 11 variables:

- e401k; *=1 if eligble for 401(k).*
- inc; *annual income*
- marr; *= 1 if married*
- male; *=1 if male respondent*
- age; *age of an individual, in years*
- fsize; *family size*
- nettfa; *net total fin. assets*
- p401k; *=1 if participate in 401(k)*
- pira; *=1 if have Individual Retirement Accounts*
- incsq; inc^2
- agesq; age^2

Following the same setting as ARTICLE, we use the same instrumental variable model to estimate the multiplicative local average treatment effect of 401(k) participation on the participation in IRA. we use e401k as instrumental variable, p401k as treatment variable and pira as response variable. The covariate X include an intercept, inc, inc^2, age, marr and fsize. Since only those eligible for 401(k) will participate in it, i.e. $D(0) = 0$, there is no defiers and always takers. This implies the probability of an individual's being always taker conditional on his being always taker or never taker is zero,i.e. $\phi_2(X) \equiv 0$. In the parameterization of $\Pr(D, Y|Z)$, ϕ_4 is always multiplied by ϕ_2, thus the model of $\Pr(D, Y|Z)$ can be uniquely represented by $\delta^{\mathrm{M}}, \phi_1, \phi_3$ and ${}_{\mathrm{OP}}{}^{\mathrm{CO}}$.

The models for $\delta^{\mathrm{M}}, \phi_1, \phi_3$ and ${}_{\mathrm{OP}}{}^{\mathrm{CO}}$ are

$$\begin{cases} \delta^{\mathrm{M}}(X) = \exp(\alpha' X); \\ \phi_1(X) = \text{sigmoid}(\beta_1' X); \\ \phi_3(X) = \text{sigmoid}(\beta_2' X); \\ {}_{\mathrm{OP}}{}^{\mathrm{CO}}(X) = \exp(\eta^T X) \\ P(Z = 1|X) = \text{sigmoid}(\gamma' X). \end{cases} \tag{2}$$

which are generally in log-linear family and logit-lieanr family. This two family is less flexible compared to machine learning models, such as energy based model. Future works may investigate the performances of these proposed methods with energy based model. We repeat the experiements of methods `mle` and `drw`, and cite the results of methods `drw.ogburn`, `dru.ogburn`, `ls.abadie` and `mle.crude` reported by ARTICLE.

3 Submission of papers to NeurIPS 2022

Please read the instructions below carefully and follow them faithfully.

3.1 Style

Papers to be submitted to NeurIPS 2022 must be prepared according to the instructions presented here. Papers may only be up to **nine** pages long, including figures. Additional pages *containing only acknowledgments and references* are allowed. Papers that exceed the page limit will not be reviewed, or in any other way considered for presentation at the conference.

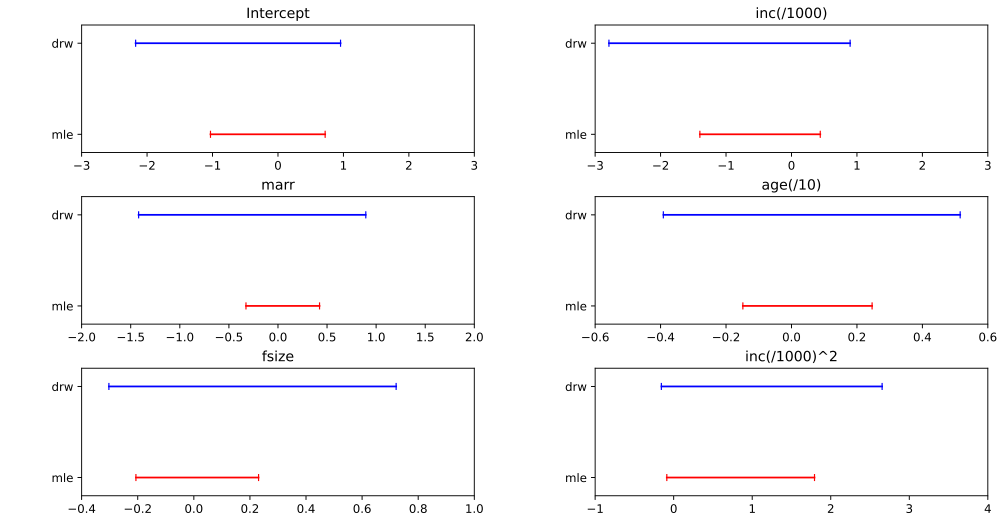

Figure 1: 95% Confidence intervals of estimates of coefficients in δ^{M} using different methods (Repeated by us)

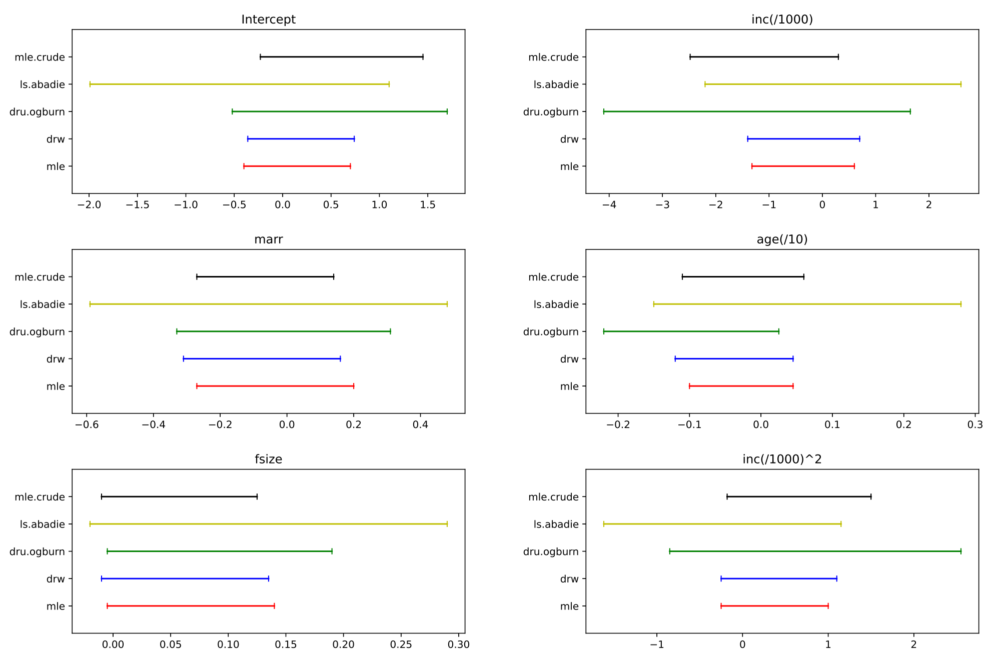

Figure 2: 95% Confidence intervals of estimates of coefficients in δ^{M} using different methods (Reported by ARTICLE)

The margins in 2022 are the same as those in 2007, which allow for $\sim 15\%$ more words in the paper compared to earlier years.

Authors are required to use the NeurIPS LaTeX style files obtainable at the NeurIPS website as indicated below. Please make sure you use the current files and not previous versions. Tweaking the style files may be grounds for rejection.

3.2 Retrieval of style files

The style files for NeurIPS and other conference information are available on the World Wide Web at

`http://www.neurips.cc/`

The file `neurips_2022.pdf` contains these instructions and illustrates the various formatting requirements your NeurIPS paper must satisfy.

The only supported style file for NeurIPS 2022 is `neurips_2022.sty`, rewritten for LaTeX 2_ε. **Previous style files for LaTeX 2.09, Microsoft Word, and RTF are no longer supported!**

The LaTeX style file contains three optional arguments: `final`, which creates a camera-ready copy, `preprint`, which creates a preprint for submission to, e.g., arXiv, and `nonatbib`, which will not load the `natbib` package for you in case of package clash.

Preprint option If you wish to post a preprint of your work online, e.g., on arXiv, using the NeurIPS style, please use the `preprint` option. This will create a nonanonymized version of your work with the text "Preprint. Work in progress." in the footer. This version may be distributed as you see fit. Please **do not** use the `final` option, which should **only** be used for papers accepted to NeurIPS.

At submission time, please omit the `final` and `preprint` options. This will anonymize your submission and add line numbers to aid review. Please do *not* refer to these line numbers in your paper as they will be removed during generation of camera-ready copies.

The file `neurips_2022.tex` may be used as a "shell" for writing your paper. All you have to do is replace the author, title, abstract, and text of the paper with your own.

The formatting instructions contained in these style files are summarized in Sections 4, 5, and 6 below.

4 General formatting instructions

The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long. The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing (leading) of 11 points. Times New Roman is the preferred typeface throughout, and will be selected for you by default. Paragraphs are separated by ½ line space (5.5 points), with no indentation.

The paper title should be 17 point, initial caps/lower case, bold, centered between two horizontal rules. The top rule should be 4 points thick and the bottom rule should be 1 point thick. Allow ¼ inch space above and below the title to rules. All pages should start at 1 inch (6 picas) from the top of the page.

For the final version, authors' names are set in boldface, and each name is centered above the corresponding address. The lead author's name is to be listed first (left-most), and the co-authors' names (if different address) are set to follow. If there is only one co-author, list both author and co-author side by side.

Please pay special attention to the instructions in Section 6 regarding figures, tables, acknowledgments, and references.

5 Headings: first level

All headings should be lower case (except for first word and proper nouns), flush left, and bold.

First-level headings should be in 12-point type.

5.1 Headings: second level

Second-level headings should be in 10-point type.

5.1.1 Headings: third level

Third-level headings should be in 10-point type.

Paragraphs There is also a `\paragraph` command available, which sets the heading in bold, flush left, and inline with the text, with the heading followed by 1 em of space.

6 Citations, figures, tables, references

These instructions apply to everyone.

6.1 Citations within the text

The `natbib` package will be loaded for you by default. Citations may be author/year or numeric, as long as you maintain internal consistency. As to the format of the references themselves, any style is acceptable as long as it is used consistently.

The documentation for `natbib` may be found at

```
http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf
```

Of note is the command `\citet`, which produces citations appropriate for use in inline text. For example,

```
\citet{hasselmo} investigated\dots
```

produces

> Hasselmo, et al. (1995) investigated. . .

If you wish to load the `natbib` package with options, you may add the following before loading the `neurips_2022` package:

```
\PassOptionsToPackage{options}{natbib}
```

If `natbib` clashes with another package you load, you can add the optional argument `nonatbib` when loading the style file:

```
\usepackage[nonatbib]{neurips_2022}
```

As submission is double blind, refer to your own published work in the third person. That is, use "In the previous work of Jones et al. [4]," not "In our previous work [4]." If you cite your other papers that are not widely available (e.g., a journal paper under review), use anonymous author names in the citation, e.g., an author of the form "A. Anonymous."

6.2 Footnotes

Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number[1] in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).

Note that footnotes are properly typeset *after* punctuation marks.[2]

Figure 3: Sample figure caption.

Table 4: Sample table title

Part		
Name	Description	Size (μm)
Dendrite	Input terminal	$\sim$100
Axon	Output terminal	$\sim$10
Soma	Cell body	up to 10^6

6.3 Figures

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction. The figure number and caption always appear after the figure. Place one line space before the figure caption and one line space after the figure. The figure caption should be lower case (except for first word and proper nouns); figures are numbered consecutively.

You may use color figures. However, it is best for the figure captions and the paper body to be legible if the paper is printed in either black/white or in color.

6.4 Tables

All tables must be centered, neat, clean and legible. The table number and title always appear before the table. See Table 4.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

Note that publication-quality tables *do not contain vertical rules.* We strongly suggest the use of the `booktabs` package, which allows for typesetting high-quality, professional tables:

`https://www.ctan.org/pkg/booktabs`

This package was used to typeset Table 4.

7 Final instructions

Do not change any aspects of the formatting parameters in the style files. In particular, do not modify the width or length of the rectangle the text should fit into, and do not change font sizes (except perhaps in the **References** section; see below). Please note that pages should be numbered.

[1] Sample of the first footnote.
[2] As in this example.

8 Preparing PDF files

Please prepare submission files with paper size "US Letter," and not, for example, "A4."

Fonts were the main cause of problems in the past years. Your PDF file must only contain Type 1 or Embedded TrueType fonts. Here are a few instructions to achieve this.

- You should directly generate PDF files using `pdflatex`.
- You can check which fonts a PDF files uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdffonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- The IEEE has recommendations for generating PDF files whose fonts are also acceptable for NeurIPS. Please see `http://www.emfield.org/icuwb2010/downloads/IEEE-PDF-SpecV32.pdf`
- `xfig` "patterned" shapes are implemented with bitmap fonts. Use "solid" shapes instead.
- The `\bbold` package almost always uses bitmap fonts. You should use the equivalent AMS Fonts:

  ```
  \usepackage{amsfonts}
  ```

  followed by, e.g., `\mathbb{R}`, `\mathbb{N}`, or `\mathbb{C}` for $\mathbb{R}$, $\mathbb{N}$ or $\mathbb{C}$. You can also use the following workaround for reals, natural and complex:

  ```
  \newcommand{\RR}{I\!\!R} %real numbers
  \newcommand{\Nat}{I\!\!N} %natural numbers
  \newcommand{\CC}{I\!\!\!\!C} %complex numbers
  ```

  Note that `amsfonts` is automatically loaded by the `amssymb` package.

If your file contains type 3 fonts or non embedded TrueType fonts, we will ask you to fix it.

8.1 Margins in LaTeX

Most of the margin problems come from figures positioned by hand using `\special` or other commands. We suggest using the command `\includegraphics` from the `graphicx` package. Always specify the figure width as a multiple of the line width as in the example below:

```
\usepackage[pdftex]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.pdf}
```

See Section 4.4 in the graphics bundle documentation (`http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf`)

A number of width problems arise when LaTeX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command when necessary.

Acknowledgments and Disclosure of Funding

Use unnumbered first level headings for the acknowledgments. All acknowledgments go at the end of the paper before the list of references. Moreover, you are required to declare funding (financial activities supporting the submitted work) and competing interests (related financial activities outside the submitted work). More information about this disclosure can be found at: `https://neurips.cc/Conferences/2022/PaperInformation/FundingDisclosure`.

Do **not** include this section in the anonymized submission, only in the final paper. You can use the `ack` environment provided in the style file to autmoatically hide this section in the anonymized submission.

References

References follow the acknowledgments. Use unnumbered first-level heading for the references. Any choice of citation style is acceptable as long as you are consistent. It is permissible to reduce the font size to `small` (9 point) when listing the references. Note that the Reference section does not count towards the page limit.

[1] Wang, L., Zhang, Y., Richardson, T. S., & Robins, J. M. (2021). Estimation of local treatment effects under the binary instrumental variable model. *Biometrika*, 108(4), 881-894.

[2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System.* New York: TELOS/Springer–Verlag.

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.

Checklist

The checklist follows the references. Please read the checklist guidelines carefully for information on how to answer these questions. For each question, change the default **[TODO]** to [Yes] , [No] , or [N/A] . You are strongly encouraged to include a **justification to your answer**, either by referencing the appropriate section of your paper or providing a brief inline description. For example:

- Did you include the license to the code and datasets? [Yes] See Section 4.
- Did you include the license to the code and datasets? [No] The code and the data are proprietary.
- Did you include the license to the code and datasets? [N/A]

Please do not modify the questions and only use the provided macros for your answers. Note that the Checklist section does not count towards the page limit. In your paper, please delete this instructions block and only keep the Checklist section heading above along with the questions/answers below.

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? **[TODO]**
   (b) Did you describe the limitations of your work? **[TODO]**
   (c) Did you discuss any potential negative societal impacts of your work? **[TODO]**
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? **[TODO]**
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? **[TODO]**
   (b) Did you include complete proofs of all theoretical results? **[TODO]**
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? **[TODO]**
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? **[TODO]**
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? **[TODO]**
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? **[TODO]**
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? **[TODO]**
   (b) Did you mention the license of the assets? **[TODO]**

(c) Did you include any new assets either in the supplemental material or as a URL? **[TODO]**

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? **[TODO]**

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? **[TODO]**

5. If you used crowdsourcing or conducted research with human subjects...

   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? **[TODO]**

   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? **[TODO]**

   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? **[TODO]**

A Appendix

Optionally include extra information (complete proofs, additional experiments and plots) in the appendix. This section will often be part of the supplemental material.