HELMHOLTZ MUNICH

Levelling Up Your Scientific Writing

Bastian Rieck

Motivation

Sturgeon's Law

Ninety percent of everything is crap.

Motivation

Sturgeon's Law

Ninety percent of everything is crap.

Coupette's Corollary

Most of our 'best practices' are, in fact, our *worst* practices.

What is this talk about?

1. Some writing tips.

What is this talk about?

1. Some writing tips.
2. Some tooling tips.

What is this talk about?

1. Some writing tips.
2. Some tooling tips.
3. A lot of opinions.

What is this talk about?

1. Some writing tips.
2. Some tooling tips.
3. A lot of opinions.

Caveat

These rules are not carved in stone. You can always deviate from them if you know what you are doing. Use this to develop your own style.

Why care?

- Great science should look and read great.

Why care?

- Great science should look and read great.
- Attention is the reader's currency. Spend it well.

Style over substance?

Modeling the creative process of the mind by prime numbers and a simple proof of the Riemann Hypothesis

Shi Huang, Ph.D.

The Burnham Institute for Medical Research
10901 North Torrey Pines Road
La Jolla, CA 92037, USA

Institute of Biomedical Sciences
Center for Evolutionary Biology
Fudan University
Shanghai 200433, China.

shuangtheman@yahoo.com

Key words: prime numbers, uniqueness, uniformity, creativity, creation, evolution, Darwinism, the Riemann Hypothesis, the Prime Law

Style over substance?

Modeling the creative process of the mind by prime numbers and a simple proof of the Riemann Hypothesis

Shi Huang, Ph.D.

The Burnham Institute for Medical Research
10901 North Torrey Pines Road
La Jolla, CA 92037, USA

Institute of Biomedical Sciences
Center for Evolutionary Biology
Fudan University
Shanghai 200433, China.

shuangtheman@yahoo.com

Key words: prime numbers, uniqueness, uniformity, creativity, creation, evolution, Darwinism, the Riemann Hypothesis, the Prime Law

Proof of the Riemann Hypothesis

Björn Tegetmeyer

11.10.2023

Abstract

The Riemann hypothesis, stating that the real part of all non-trivial zero points of the zeta function must be $\frac{1}{2}$, is one of the most important unproven hypotheses in number theory. In this paper we will prove the Riemann hypothesis by using the integral representation $\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{x-\lfloor x \rfloor}{x^{s+1}}\,\mathrm{d}x$ and solving the integral for the real- and imaginary part of the zeta function.

1 Introduction

In 1859 Bernhard Riemann found one of the most eminent mathematical problems of our time: In his paper "On the Number of Primes Less Than a Given Magnitude" [4] he published the assumption that all non-trivial zero-points of the zeta function extended to the range of complex numbers $\mathbb{C}$ have a real part of $\frac{1}{2}$, noting the demand of a strict proof for this. Ever since David Hilbert in 1900 added this problem to his list of the 23 most important problems of 20th century, mathematicians have been working on finding evidence for Riemanns hypothesis. This paper aims to provide the proof and fill this gap in modern mathematics.

2 Proof of the Riemann Hypothesis

The zeta-function $\zeta(s)$ in the complex range $s \in \mathbb{C}$ for a positive real-part of s can be formulated as integral representation

$$\zeta(s) = \frac{s}{s-1} - s\int_1^\infty \frac{x-\lfloor x \rfloor}{x^{s+1}}\,\mathrm{d}x \tag{1}$$

with $s \in \mathbb{C}$, where s can be expressed by $s = \alpha + ib$; $a, b \in \mathbb{R}$ and $0 < a < 1$ as well as $0 < b$. Be s_0 a zero point of the zeta function. From [3] we know, that the zeta-function is symmetrical in a way that $\zeta(s_0) = 0 \Leftrightarrow \zeta(1-s_0) = 0$ for all zero-points $s_0 \in \mathbb{C}$ (see appendix for details). In accordance to equation 1 we can write $\zeta(1-s)$ as

$$\zeta(1-s) = \frac{1-s}{-s} - (1-s)\int_1^\infty \frac{x-\lfloor x \rfloor}{x^{1-s}}\,\mathrm{d}x \tag{2}$$

The Riemann hypothesis states, that the real part of s would be $\frac{1}{2}$ for all non-trivial zero-points of zeta (i.e. all zero points of zeta with a positive real part). Furthermore, from [2] we know, that $0 < \Re(s_0) < 1$. Inserting $s_0 = a + ib$ into $\zeta(s)$, using $x^{-1-a-ib} = x^{-1-a}\cos(b\ln(x)) - ix^{-1-a}\sin(b\ln(x))$ and defining $\{x\} := x - \lfloor x \rfloor$ we get the following two equations 3 and 4 out off $\zeta(s)$ and $\zeta(1-s)$:

$$\begin{aligned} \frac{a-1}{(a-1)^2+b^2} - i\frac{b}{(a-1)^2+b^2} &= \int_1^\infty \frac{\{x\}}{x^{a+1}}\,\mathrm{d}x \\ \Leftrightarrow \frac{a-1}{(a-1)^2+b^2} - i\frac{b}{(a-1)^2+b^2} &= \int_1^\infty \frac{\{x\}}{x^{a+1}}(\cos(b\ln(x)) - i\sin(b\ln(x)))\,\mathrm{d}x \end{aligned} \tag{3}$$

and

$$\begin{aligned} \frac{1-a}{-a} &= \int_1^\infty \frac{-\{x\}}{x^{1-a}}\,\mathrm{d}x \\ \Leftrightarrow \frac{1-a}{-a} &= \int_1^\infty \frac{-\{x\}}{x^{1-a}}(\cos(b\ln(x)) + i\sin(b\ln(x)))\,\mathrm{d}x \end{aligned} \tag{4}$$

Thus, we get 4 equations, for the real- and imaginary-part by means of $\zeta(s_0)$ (being called here $\Re_1$ and $\Im_1$) and $\zeta(1-s_0)$ (being called here $\Re_2$ and $\Im_2$):

$$\begin{aligned} \Re_1: \frac{a-1}{(a-1)^2+b^2} &= \int_1^\infty \frac{\{x\}}{x^{a+1}}\cos(b\ln(x))\,\mathrm{d}x \quad \Im_1: -\frac{b}{(a-1)^2+b^2} &= \int_1^\infty \frac{-\{x\}}{x^{a+1}}\sin(b\ln(x))\,\mathrm{d}x \\ \Re_2: \frac{1-a}{-a} &= \int_1^\infty \frac{-\{x\}}{x^{1-a}}\cos(b\ln(x))\,\mathrm{d}x \quad \Im_2: \frac{1-a}{-a} &= \int_1^\infty \frac{-\{x\}}{x^{1-a}}\sin(b\ln(x))\,\mathrm{d}x \end{aligned} \tag{5}$$

Part I: Writing & Plotting

Guiding principles

- Above all, strive for clarity and consistency.

Guiding principles

- Above all, strive for clarity and consistency.
- Omit unnecessary words.

Guiding principles

- Above all, strive for clarity and consistency.
- Omit unnecessary words.
- Use the active voice (because it is clearer).

Guiding principles

- Above all, strive for clarity and consistency.
- Omit unnecessary words.
- Use the active voice (because it is clearer).
- Avoid ‘weasel words’ and absolutes.

Examples

B. Rieck and H. Leitte, ‘Persistent Homology for the Evaluation of Dimensionality Reduction Schemes’, *Computer Graphics Forum* 34.3, 2015, pp. 431–440

- Bad: *The utility of our method is demonstrated using application data from multiple domains and a variety of commonly used DR methods.*

Examples

B. Rieck and H. Leitte, 'Persistent Homology for the Evaluation of Dimensionality Reduction Schemes', *Computer Graphics Forum* 34.3, 2015, pp. 431–440

- Bad: *The utility of our method is demonstrated using application data from multiple domains and a variety of commonly used DR methods.*
- Better: *We demonstrate the utility of our method by analysing data from several scientific domains with a variety of dimensionality reduction methods.*

Examples

B. Rieck and H. Leitte, 'Persistent Homology for the Evaluation of Dimensionality Reduction Schemes', *Computer Graphics Forum* 34.3, 2015, pp. 431–440

- Bad: *The utility of our method is demonstrated using application data from multiple domains and a variety of commonly used DR methods.*
- Better: *We demonstrate the utility of our method by analysing data from several scientific domains with a variety of dimensionality reduction methods.*
- Bad: *For high-dimensional data sets, the calculation is slightly more involved than in the 1-dimensional case.*

Examples

B. Rieck and H. Leitte, 'Persistent Homology for the Evaluation of Dimensionality Reduction Schemes', *Computer Graphics Forum* 34.3, 2015, pp. 431–440

- Bad: *The utility of our method is demonstrated using application data from multiple domains and a variety of commonly used DR methods.*
- Better: *We demonstrate the utility of our method by analysing data from several scientific domains with a variety of dimensionality reduction methods.*
- Bad: *For high-dimensional data sets, the calculation is slightly more involved than in the 1-dimensional case.*
- Better: *For high-dimensional data sets, the calculation is ~~slightly~~ more involved than in the 1-dimensional case.*

Examples

B. Rieck and H. Leitte, 'Persistent Homology for the Evaluation of Dimensionality Reduction Schemes', *Computer Graphics Forum* 34.3, 2015, pp. 431–440

- Bad: *The utility of our method is demonstrated using application data from multiple domains and a variety of commonly used DR methods.*
- Better: *We demonstrate the utility of our method by analysing data from several scientific domains with a variety of dimensionality reduction methods.*
- Bad: *For high-dimensional data sets, the calculation is slightly more involved than in the 1-dimensional case.*
- Better: *For high-dimensional data sets, the calculation is ~~slightly~~ more involved than in the 1-dimensional case.*
- Bad: *The "Isomap faces" data set is well-suited for a parameter study of Isomap.*

Examples

B. Rieck and H. Leitte, 'Persistent Homology for the Evaluation of Dimensionality Reduction Schemes', *Computer Graphics Forum* 34.3, 2015, pp. 431–440

- Bad: *The utility of our method is demonstrated using application data from multiple domains and a variety of commonly used DR methods.*
- Better: *We demonstrate the utility of our method by analysing data from several scientific domains with a variety of dimensionality reduction methods.*
- Bad: *For high-dimensional data sets, the calculation is slightly more involved than in the 1-dimensional case.*
- Better: *For high-dimensional data sets, the calculation is ~~slightly~~ more involved than in the 1-dimensional case.*
- Bad: *The "Isomap faces" data set is well-suited for a parameter study of Isomap.*
- Better: *Due to its manifold structure, the "Isomap faces" data set can be used to analyse the hyperparameters of the Isomap algorithm.*

Morals

- Good writing needs a lot of editing.

Morals

- Good writing needs a lot of editing.
- Your first draft will probably not be good—and that is okay!

Morals

- Good writing needs a lot of editing.
- Your first draft will probably not be good—and that is okay!

Morals

- Good writing needs a lot of editing.
- Your first draft will probably not be good—and that is okay!

But also know when to stop:

> *Striving to better, oft we mar what's well.*
>
> *(Shakespeare, King Lear, Act I, Scene IV)*

Graphics

- The same principles apply to graphics.

Graphics

- The same principles apply to graphics.
- Statistical plotting software is not designed by experts in perception research.

Graphics

- The same principles apply to graphics.
- Statistical plotting software is not designed by experts in perception research.
- Tufte's guidelines: *Remove chart junk, maximise data-ink.*

Graphics

- The same principles apply to graphics.
- Statistical plotting software is not designed by experts in perception research.
- Tufte's guidelines: *Remove chart junk, maximise data-ink.*
- Not everything needs to be a figure; consider using tables.

Graphics

- The same principles apply to graphics.
- Statistical plotting software is not designed by experts in perception research.
- Tufte's guidelines: *Remove chart junk, maximise data-ink.*
- Not everything needs to be a figure; consider using tables.
- Strive for a consistent look and feel.

Graphics

- The same principles apply to graphics.
- Statistical plotting software is not designed by experts in perception research.
- Tufte's guidelines: *Remove chart junk, maximise data-ink.*
- Not everything needs to be a figure; consider using tables.
- Strive for a consistent look and feel.
- Hence: Prefer vector graphics over all formats!

Graphics

- The same principles apply to graphics.
- Statistical plotting software is not designed by experts in perception research.
- Tufte's guidelines: *Remove chart junk, maximise data-ink.*
- Not everything needs to be a figure; consider using tables.
- Strive for a consistent look and feel.
- Hence: Prefer vector graphics over all formats!
- Invest some time to learn your tools.

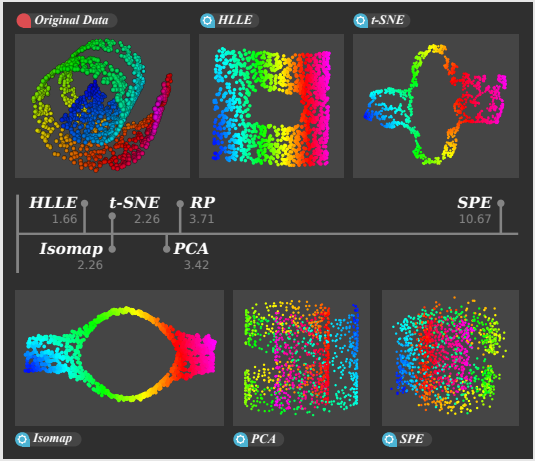
Graphics

- The same principles apply to graphics.
- Statistical plotting software is not designed by experts in perception research.
- Tufte's guidelines: *Remove chart junk, maximise data-ink.*
- Not everything needs to be a figure; consider using tables.
- Strive for a consistent look and feel.
- Hence: Prefer vector graphics over all formats!
- Invest some time to learn your tools.

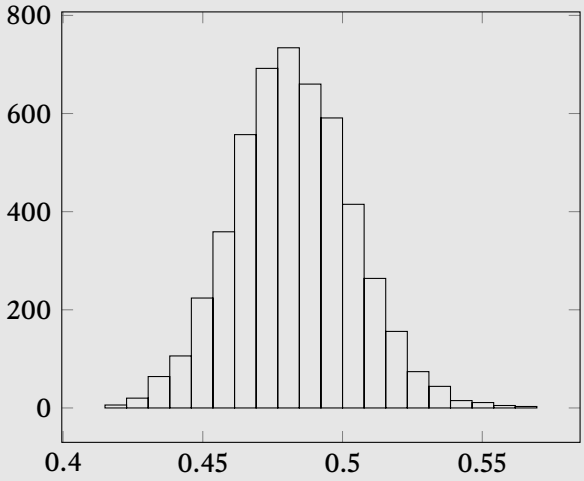
LaTeX packages: `tikz`, `pgfplots`
`Python` packages: `matplotlib`, `seaborn`

Not good!

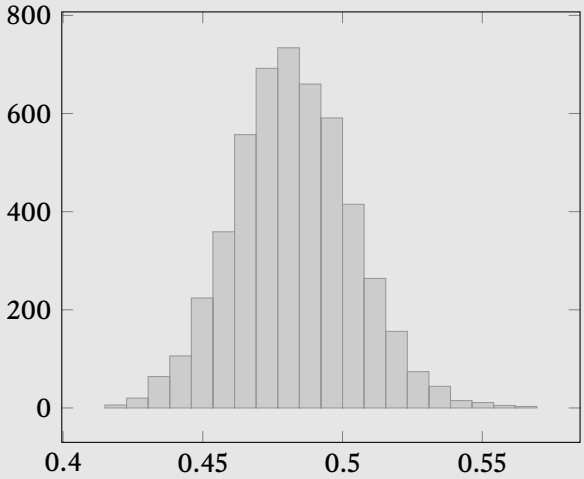
B. Rieck and H. Leitte, ‘Persistent Homology for the Evaluation of Dimensionality Reduction Schemes’, *Computer Graphics Forum* 34.3, 2015, pp. 431–440



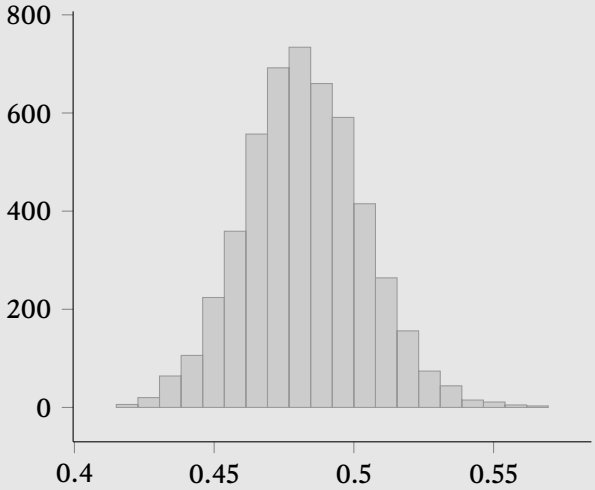
Improving the defaults



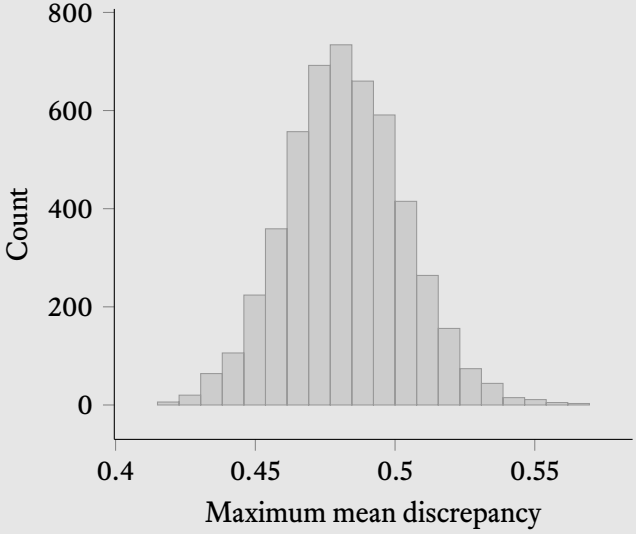
Improving the defaults



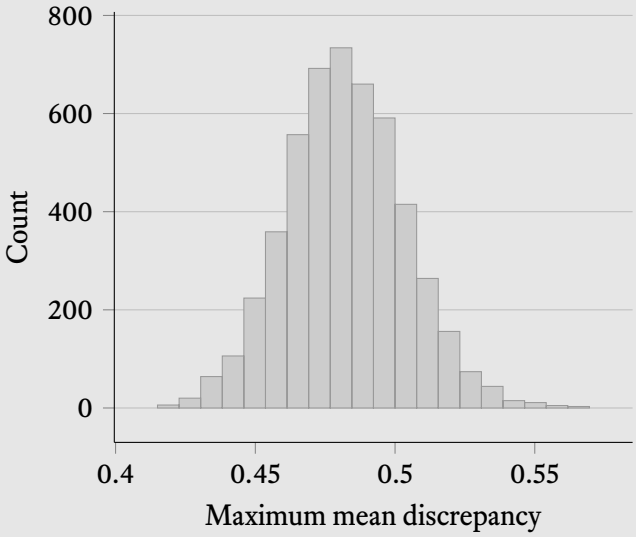
Improving the defaults



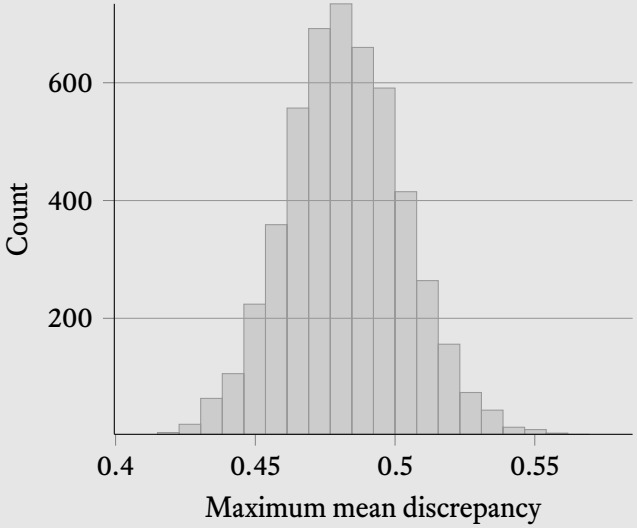
Improving the defaults



Improving the defaults



Improving the defaults



Part II: Tools

Guidelines

- LaTeX sets text with ‘rubber lengths,’ i.e. lengths that can change.

Guidelines

- LaTeX sets text with ‘rubber lengths,’ i.e. lengths that can change.
- Thus, text can *expand* or *contract* to a certain extent.

Guidelines

- LaTeX sets text with ‘rubber lengths,’ i.e. lengths that can change.
- Thus, text can *expand* or *contract* to a certain extent.
- Use this to your advantage when text spills over!

Guidelines

- LaTeX sets text with ‘rubber lengths,’ i.e. lengths that can change.
- Thus, text can *expand* or *contract* to a certain extent.
- Use this to your advantage when text spills over!
- Pay attention to warnings like this one:
  `Underfull \hbox (badness 10000) has occurred while \output is active`

Guidelines

- LaTeX sets text with ‘rubber lengths,’ i.e. lengths that can change.
- Thus, text can *expand* or *contract* to a certain extent.
- Use this to your advantage when text spills over!
- Pay attention to warnings like this one:
  `Underfull \hbox (badness 10000) has occurred while \output is active`
- Sometimes, spacing matters. Use % whenever necessary to signal to LaTeX that no additional space is intended.

Common mistake: Incorrect citations

- Both the `natbib` and `biblatex` packages have specific citation commands that communicate the *intent*.

Common mistake: Incorrect citations

- Both the `natbib` and `biblatex` packages have specific citation commands that communicate the *intent*.
- When using `natbib`, use `\citet` for in-text citations and `\citep` for parenthetical citations.

Common mistake: Incorrect citations

- Both the `natbib` and `biblatex` packages have specific citation commands that communicate the *intent*.
- When using `natbib`, use `\citet` for in-text citations and `\citep` for parenthetical citations.
- It looks better and will work consistently across different citation styles.

Common mistake: Incorrect hyphenation and spacing

Persistent homology: *Persistent homology*, the technique used by our method, has already been used to complement standard data analysis methods. Singh et al. [SMC07] showed the importance of studying the behaviour of a given function on the data. Carlsson [Car14] refers to this as *functional persistence*. Sheehy [She14] recently proved that the topological features of distance functions remain stable under projections, implying that the study of functions (and their connectivities) on a data set contains salient information.

Persistent homology: *Persistent homology*, the technique used by our method, has already been used to complement standard data analysis methods. Singh et al. [SMC07] showed the importance of studying the behaviour of a given function on the data. Carlsson [Car14] refers to this as *functional persistence*. Sheehy [She14] recently proved that the topological features of distance functions remain stable under projections, implying that the study of functions (and their connectivities) on a data set contains salient information.

Fix

```
\usepackage[british]{babel}
\usepackage{microtype}
```

Common mistake: Using default `hyperref` colours

Theorem 5 implies that topological distances are generally more discriminative than the distances between the filtration functions. Thus, calculating topological representations of graphs based on a class of functions improves discriminative power. To further understand the expressive power of curvature filtrations, we analyse strongly-regular graphs, which are often used for studying GNN expressivity as they cannot be distinguished by k-WL, the k-dimensional Weisfeiler–Le(h)man test, if $k \leq 3$ [2, 7, 52]. Additionally, we explore how curvature filtrations can count substructures, an important tool for evaluating and comparing expressivity [56]. To the best of our knowledge, ours is the first work to explore discrete curvature and curvature-based filtrations in this context.

Theorem 5 implies that topological distances are generally more discriminative than the distances between the filtration functions. Thus, calculating topological representations of graphs based on a class of functions improves discriminative power. To further understand the expressive power of curvature filtrations, we analyse strongly-regular graphs, which are often used for studying GNN expressivity as they cannot be distinguished by k-WL, the k-dimensional Weisfeiler–Le(h)man test, if $k \leq 3$ [2, 7, 52]. Additionally, we explore how curvature filtrations can count substructures, an important tool for evaluating and comparing expressivity [56]. To the best of our knowledge, ours is the first work to explore discrete curvature and curvature-based filtrations in this context.

Fix

```
\usepackage{hyperref}

% Better than nothing!
\hypersetup{
  colorlinks = true,
  urlcolor   = blue,
  linkcolor  = blue,
  citecolor  = blue,
}
```

Common mistake: Using TEX primitives

- The `\emph` command is aware of its surroundings, as its goal is to *emphasise* text.

Common mistake: Using TEX primitives

- The `\emph` command is aware of its surroundings, as its goal is to *emphasise* text.
- In articles, an `\emph` within an `\emph` will thus be set upright again (try it out when using a `theorem` environment or a definition, for instance).

Common mistake: Using TEX primitives

- The `\emph` command is aware of its surroundings, as its goal is to *emphasise* text.
- In articles, an `\emph` within an `\emph` will thus be set upright again (try it out when using a `theorem` environment or a definition, for instance).
- The `\textit` or `\it` should be avoided because they *enforce* their respective style.

Common mistake: Using TEX primitives

- The `\emph` command is aware of its surroundings, as its goal is to *emphasise* text.
- In articles, an `\emph` within an `\emph` will thus be set upright again (try it out when using a `theorem` environment or a definition, for instance).
- The `\textit` or `\it` should be avoided because they *enforce* their respective style.
- The same goes for `\textbf`, `\bf`, etc.

Common mistake: Using TEX primitives

- The `\emph` command is aware of its surroundings, as its goal is to *emphasise* text.
- In articles, an `\emph` within an `\emph` will thus be set upright again (try it out when using a `theorem` environment or a definition, for instance).
- The `\textit` or `\it` should be avoided because they *enforce* their respective style.
- The same goes for `\textbf`, `\bf`, etc.
- Their usage should always be *deliberate*.

Common mistake: Widows and orphans

- Look out for single words in a new line.

Common mistake: Widows and orphans

- Look out for single words in a new line.
- Rewriting can often help LaTeX set the text better.

Common mistake: Widows and orphans

- Look out for single words in a new line.
- Rewriting can often help LaTeX set the text better.

Common mistake: Widows and orphans

- Look out for single words in a new line.
- Rewriting can often help LaTeX set the text better.

perturbations. The persistence diagrams change in a non-differentiable manner during the training phase. However, for any given update step, a diagram is robust to infinitesimal changes of its entries (Cohen-Steiner et al., 2007). As a consequence, our topological loss is differentiable for each update step during training. We make our code publicly available[4]

for any given update step, a diagram is robust to infinitesimal changes of its entries (Cohen-Steiner et al., 2007). Thus, our topological loss is differentiable for each update step during training. We make our code publicly available.[4]

Better typography: Tables

```
\usepackage{booktabs}

[...]

\begin{tabular}{ll}
  \toprule
  \emph{User} & \emph{Operating System}\\
  \midrule
  Alice       & Arch Linux\\
  Bob         & Bodhi Linux\\
  Charlie     & CentOS\\
  \bottomrule
\end{tabular}
```

User	*Operating System*
Alice	Arch Linux
Bob	Bodhi Linux
Charlie	CentOS

There is often *no need* for a lot of horizontal and vertical rules!

Better typography: Maths

```
\usepackage{amsmath}
\usepackage{amssymb}

\DeclareMathOperator{\diameter}{diam}

[...]

Thus, $\diameter(X) \leq \pi$.
```

Thus, $\operatorname{diam}(X) \leq \pi$. (Good)
Thus, $diam(X) \leq \pi$. (Bad)

In general, the documentation of `amsmath` is well worth a read!

Better typography: References

```
\usepackage[capitalize]{cleveref}

[...]

See \cref{fig:Overview} for an example.
```

We propose a topology-aware loss term based on concepts from topological machine learning and optimal transport. The loss term works on the level of individual volumes, leveraging a valid metric between topological descriptors, while remaining efficiently computable. Owing to its generic nature, the loss can be easily integrated into existing architectures; see Fig. 1 for an overview.

The documentation of `cleveref` offers more opportunities for styling the output of the references (including abbreviations and link placement).

Better typography: Spaces

```
\usepackage{xspace}

[...]

\newcommand{\ourmethod}  {\textsc{Presto}\xspace}
\newcommand{\theirmethod}{\textsc{Mapper}}

[...]

\theirmethod scales worse than \ourmethod.\\
\ourmethod scales better than \theirmethod.
```

Missing usage of `\xspace` can introduce *subtle errors*:

MAPPERscales worse than PRESTO.
PRESTO scales better than MAPPER.

Better typography: Subcaptions

```
\usepackage{subcaption}

[...]

\subcaptionbox{TUM\label{sfig:TUM}}{%
  \includegraphics{Logos/TUM}
}% Do not add any additional space.
\subcaptionbox{Helmholtz\label{sfig:Helmholtz}}{%
  \includegraphics{Logos/HMGU}
}

\cref{sfig:TUM,sfig:Helmholtz} show \dots
```

(a) TUM (b) Helmholtz

Figure 1: Two organisations

Figures 1a and 1b show …

Warning: Do *not* use the deprecated `subfigure` package unless the style file enforces it.

More resources

- Robert Bringhurst, *The Elements of Typographic Style*
- Benjamin Dreyer, *Dreyer's English: An Utterly Correct Guide to Clarity and Style*
- William Strunk, Jr. and E. B. White, *The Elements of Style*
- Edward Tufte, *The Visual Display of Quantitative Information*

Read not to contradict and confute; nor to believe and take for granted; nor to find talk and discourse; but to weigh and consider. *(Francis Bacon)*

Parting words

And whatever your labors and aspirations, in the noisy confusion of life, keep peace in your soul. [...] Strive to be happy. *(Max Ehrmann, 'Desiderata')*