
Glauber Dynamics Based Discrete Diffusion for Text Generation

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 Denoising Diffusion Models have achieved impressive performance on various
2 continuous generation tasks like images and waveform generation. In comparison,
3 for discrete domains, in particular for text generation, their performance lags behind
4 the performance of autoregressive language models. This is because existing
5 approaches only consider forward “noising” at the token level without accounting
6 for sentence structure or semantics.

7 In this work, we propose a discrete diffusion model for text generation with the
8 forward process modeled using Glauber dynamics which allows adding sequence-
9 level noise to the input by randomly selecting and replacing its tokens to gradually
10 converge to the stationary distribution corresponding to an energy-based language
11 model.

12 This setup also naturally allows us to model *controlled* text generation directly into
13 the diffusion model. We show improved performance for a variety of benchmark
14 metrics on character-level generation tasks.

15 1 Introduction

16 Large language models (LLMs) are capable of generating high quality text (cite) but typically work
17 by generating text in an auto-regressive fashion which is inherently word-by-word which may not
18 be properly accounting for long range sentence structure in the text generated. Consequently, this
19 leads to the text generated to have errors [Quantify]. On the other hand, for continuous domains like
20 image and waveform generation, denoising diffusion models have achieved impressive performance
21 performing much better than other methods (cite). These models work by considering a forward
22 “noising” Markov process based on a stationary distribution, which is typically taken to be the
23 Ornstein-Uhlenbeck process corresponding to the stationary distribution being the standard Gaussian
24 distribution $d\gamma_n$. Then the process is “reversed” to “steer” the process such that the reversed process
25 converges to another target distribution $f d\gamma_n$ whose density f is absolutely continuous with respect
26 to the Gaussian distribution. This reverse process is typically requires a time-varying drift of the
27 dynamics, referred to as the score function, which depends on f . This score function is typically
28 estimated through a neural network using samples from the target distribution. These diffusions can
29 also be “controlled” to steer the generation towards favorable parts of the search space modelled via
30 some constraints.

31 In contrast to continuous domains, diffusion model based generation for discrete domains, in particular,
32 for text generation, are not yet competitive with auto-regressive Language Models (LMs). One
33 approach to this task tries to embed text into a continuous embedding space and then try to apply
34 continuous diffusion techniques. The text thus generated in the continuous embedding space then needs
35 to be “rounded” to generate the actual discrete text. Relying on a good “learned” embedding space to
36 represent the text and the final “rounding” step to get back to text are two sources of error, amongst

possibly others which may be leading to such models being less competitive than autoregressive models.

Another recently proposed approach is to work in the discrete domain directly and consider discrete denoising diffusion models (cite). Certainly, the important variable here is the forward Markov process, which now corresponds to (discrete or continuous time) Markov Chains, and the corresponding stationary distribution. These recent works have primarily focused on modelling transitions across words independently and the transition probabilities between words to be either uniform or based on modelling distance again via some embedding space, which is then passed via a Gaussian filter to get transition probabilities (cite Austin). However, unlike in the continuous domains, where Gaussian distribution seems to be competitive, in text domains, it seems important to model the right forward process to account for semantic structure and [other stuff Sachin] that are inherent in text. This is inherently hard to achieve if each word is diffused independently without considering how changing a word affects the sentence as a whole.

In this paper,

Please read the instructions below carefully and follow them faithfully. **Important:** This year the checklist will be submitted separately from the main paper in OpenReview, please review it well ahead of the submission deadline: <https://neurips.cc/public/guides/PaperChecklist>.

1.1 Style

Papers to be submitted to NeurIPS 2023 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.

The margins in 2023 are the same as those in previous years.

Authors are required to use the NeurIPS L^AT_EX 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.

1.2 Retrieval of style files

The style files for NeurIPS and other conference information are available on the website at

<http://www.neurips.cc/>

The file `neurips_2023.pdf` contains these instructions and illustrates the various formatting requirements your NeurIPS paper must satisfy.

The only supported style file for NeurIPS 2023 is `neurips_2023.sty`, rewritten for L^AT_EX 2_ε. **Previous style files for L^AT_EX 2.09, Microsoft Word, and RTF are no longer supported!**

The L^AT_EX 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, as long as you do not say which conference it was submitted to. 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_2023.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 2, 3, and 4 below.

85 **2 General formatting instructions**

86 The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long.
87 The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing (leading) of 11 points.
88 Times New Roman is the preferred typeface throughout, and will be selected for you by default.
89 Paragraphs are separated by 1/2 line space (5.5 points), with no indentation.

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

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

98 Please pay special attention to the instructions in Section 4 regarding figures, tables, acknowledgments,
99 and references.

100 **3 Headings: first level**

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

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

103 **3.1 Headings: second level**

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

105 **3.1.1 Headings: third level**

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

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

109 **4 Citations, figures, tables, references**

110 These instructions apply to everyone.

111 **4.1 Citations within the text**

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

115 The documentation for `natbib` may be found at

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

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

119 `\citet{hasselmo}` investigated\dots

120 produces

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

122 If you wish to load the `natbib` package with options, you may add the following before loading the
123 `neurips_2023` package:



Figure 1: Sample figure caption.

124 `\PassOptionsToPackage{options}{natbib}`

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

127 `\usepackage[nonatbib]{neurips_2023}`

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

133 4.2 Footnotes

134 Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number¹
135 in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote
136 with a horizontal rule of 2 inches (12 picas).

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

138 4.3 Figures

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

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

145 4.4 Tables

146 All tables must be centered, neat, clean and legible. The table number and title always appear before
147 the table. See Table 1.

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

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

153 <https://www.ctan.org/pkg/booktabs>

154 This package was used to typeset Table 1.

¹Sample of the first footnote.

²As in this example.

Table 1: Sample table title

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

4.5 Math

Note that display math in bare TeX commands will not create correct line numbers for submission. Please use LaTeX (or AMSTeX) commands for unnumbered display math. (You really shouldn't be using \$\$ anyway; see <https://tex.stackexchange.com/questions/503/why-is-preferable-to> and <https://tex.stackexchange.com/questions/40492/what-are-the-differences-between-align-equation-and-displaymath> for more information.)

4.6 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.

5 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 file uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdf fonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- `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}{\mathbb{R}} %real numbers
\newcommand{\Nat}{\mathbb{N}} %natural numbers
\newcommand{\CC}{\mathbb{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.

5.1 Margins in L^AT_EX

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}
```

191 See Section 4.4 in the graphics bundle documentation ([http://mirrors.ctan.org/macros/](http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf)
192 [latex/required/graphics/grfguide.pdf](http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf))
193 A number of width problems arise when L^AT_EX cannot properly hyphenate a line. Please give LaTeX
194 hyphenation hints using the \- command when necessary.

195 **6 Supplementary Material**

196 Authors may wish to optionally include extra information (complete proofs, additional experiments
197 and plots) in the appendix. All such materials should be part of the supplemental material (submitted
198 separately) and should NOT be included in the main submission.

199 **References**

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

204 [1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In
205 G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp.
206 609–616. Cambridge, MA: MIT Press.

207 [2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the*
208 *GENeral NEural Simulation System*. New York: TELOS/Springer-Verlag.

209 [3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent
210 synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.