Modellizzazione della resistenza di diodi a giunzione PN per alte correnti di lavoro

L. Ciucci(*) S. Bruzzesi(*) M. Romagnoli(*) M. Alighieri(*)
B. Tomelleri(*)

2020/11/01

Riassunto: — Studiamo il comportamento di diodi in silicio PN, esplorando la propria curva caratteristica $V - I$ al di fuori dei regimi di lavoro ordinari, tramite campionamenti digitali dei segnali ai capi del componente. Discutiamo la presenza di una componente resistiva del diodo e ne misuriamo l'entità, al fine di arrivare ad un modello teorico in grado giustificare eventuali deviazioni dal modello di Shockley, verso una risposta -ohmica- dovuta alla resistenza della giunzione PN al passaggio di correnti.

PACS 01.40.-d – Education.
PACS 01.50.Pa – Laboratory experiments and apparatus.

1 INTRODUZIONE

L'alta resistività intrinseca al silicio, di cui è composta la giunzione bipolare, comporta la presenza di una sua componente resistiva: questa risulta sempre meno trascurabile agli effetti del passaggio di corrente attraverso il diodo, all'aumentare della tensione ai suoi capi e della sua temperatura. Per poter modellare la componente -ohmica- di un diodo percorso da correnti alte si propone un modello semplice di resistenza parassita in serie, in grado di descriverne gli effetti, verificandone sperimentalmente la validità e i limiti.

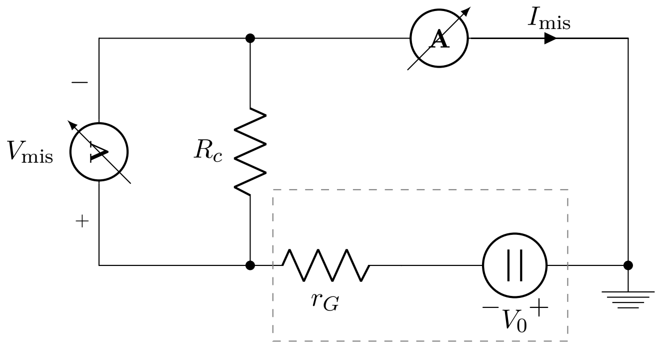
2 METODO E APPARATO SPERIMENTALE

Dovendo lavorare con correnti relativamente alte per il circuito sotto studio, al fine di minimizzare effetti termici-dissipativi secondari ed evitare danni all'apparato, si imprimono sui componenti correnti pulsate o di durate molto brevi, limitando dunque il trasferimento di energia sui componenti.

2.1 *Acquisizione dati*

Si è fatto uso della piattaforma di sviluppo `Teensy 3.2`[1] per il campionamento di segnali, essendo non solo più veloce e capiente in memoria rispetto ad Arduino, ma oltretutto dotato di due ADC, entrambi con risoluzione maggiore nel campionamento analogico, a 12 bit (reali). In particolare, `Teensy` è in grado di effettuare una lettura differenziale su due ingressi analogici e sincronizzata tra i due ADC. Questo ci permette di misurare pressoché -simultaneamente- la differenza di potenziale e intensità di corrente percepite dal diodo, dunque la sua curva caratteristica. Per prima cosa si sono misurati i rispettivi fattori di conversione ξ_V [V/digit] e ξ_I [A/digit] tramite un fit lineare sui campionamenti di tensione e corrente continua generati da un trasformatore di d.d.p $V_0 \approx 4.95$ V. Come illustrato sotto:

*Dipartimento di Fisica E. Fermi, Università di Pisa - Pisa, Italy



2.2 *Schema circuitale del sistema*

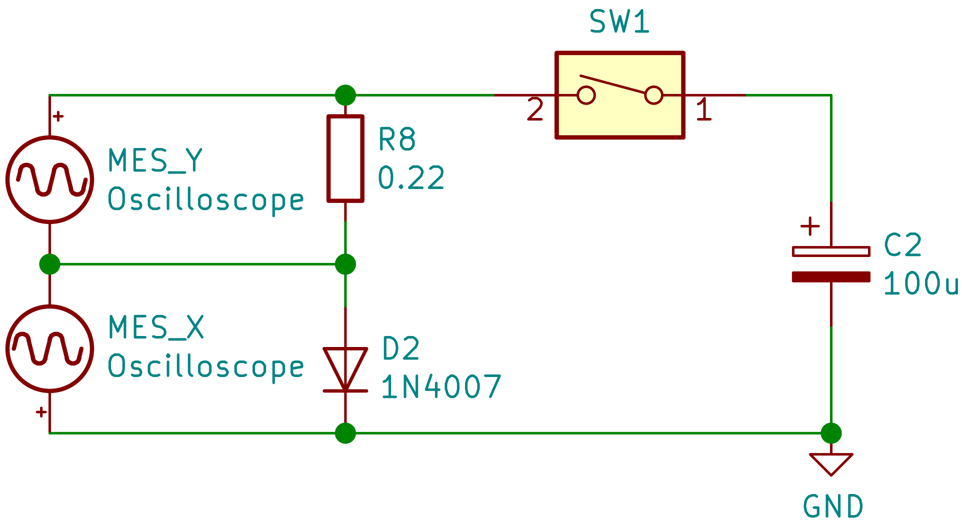


Figura 1: Versione semplice del circuito

Come ultima cautela preliminare, per minimizzare le influenze esterne sulla misura della componente resistiva del diodo, si sono misurate le resistenze dei collegamenti del circuito, in quanto i cavi reali e gli ingressi del diodo possono avere resistenze non trascurabili rispetto a quella opposta dal diodo in regime di conduzione, nell'ordine di qualche ohm.

Tramite il semplice circuito con interruttore in Figura 1 riusciamo a ricostruire sperimentalmente la curva caratteristica V-I del diodo: si lascia scaricare repentinamente il condensatore sulla serie diodo-resistenza chiudendo l'interruttore, dunque dai canali di un oscilloscopio si misurano le le cadute di tensione ai capi del diodo e della resistenza R_P, da cui si ricava l'intensità di corrente che scorre su entrambi i componenti tramite la legge di Ohm.

$$I_d = \frac{\Delta V_{\text{mis}}}{R_P}.$$

Vista la limitata attendibilità del circuito con interruttore azionato manualmente si è costruito una seconda versione del circuito (Fig. 2), in cui figurano:

UN SOTTOCIRCUITO di switch

UN CONDENSATORE di capacità maggiore di un paio di ordini di grandezza.

TEENSY La piattaforma di sviluppo impiegata per caricare (e scaricare) il condensatore a diverse tensioni e per la misura delle tensioni ai capi del diodo e della resistenza.

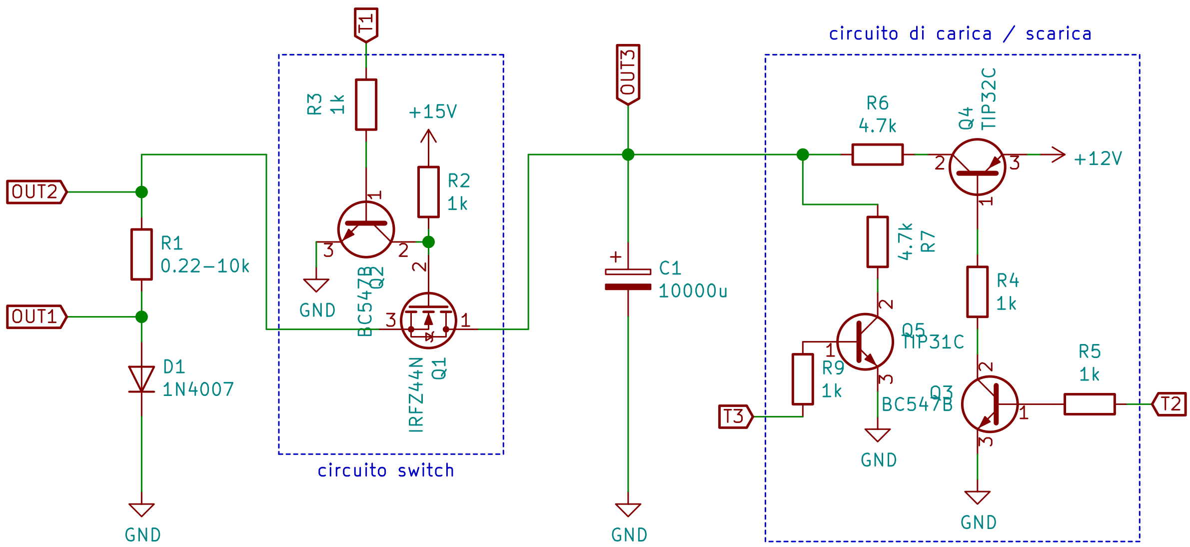


Figura 2: Circuito globale per la gestione del diodo

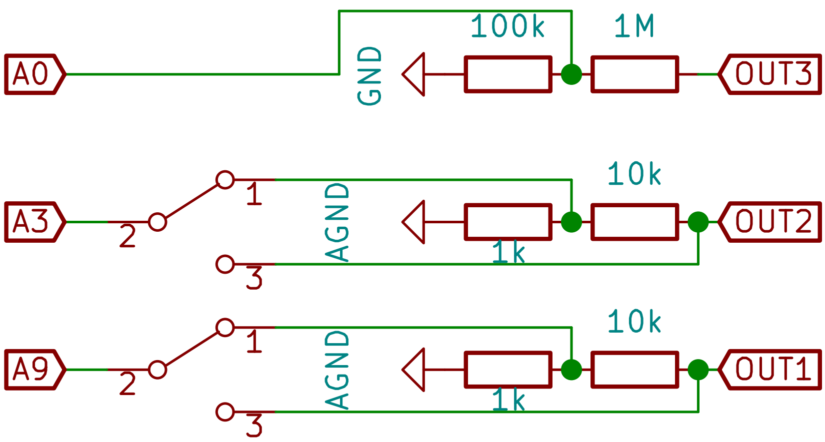


Figura 3: Schema circuitale del sistema di lettura (`Teensy`)

Si sviluppano due casi principali, dipendenti sostanzialmente dal valore della resistenza R_P posta in serie al diodo:

Se lasciamo caricare gradualmente il condensatore, essendo C decisamente maggiore rispetto ai condensatori precedenti si ha un tempo di carica $\tau = RC$ abbastanza prolungato, in cui possiamo campionare contemporaneamente tensione e corrente ai capi del diodo per correnti modeste, nel regime in cui il diodo è "interdetto" ed oppone resistenza al flusso di carica.

Viceversa, nel regime in cui si applichi al diodo una d.d.p. ben al disopra di $V_{\text{thr}} \approx 0.6\text{V}$, che siamo liberi di esplorare variando la tensione di carica di C con il sottocircuito destro, il diodo idealmente lascia passare tutta la corrente impressa sulla giunzione. Allora per caratterizzare la risposta del diodo senza cambiarne drasticamente le caratteristiche (ad esempio per eccessiva agitazione termica) si impiega il circuito di switch per imprimere impulsi di alta corrente e breve durata sulla giunzione PN, di cui misuriamo la curva caratteristica in risposta con i due ADC di `Teensy`.

Se effettivamente il diodo, oltre ad una certa soglia di d.d.p. inizia ad avere componente resistiva sempre più pronunciata, allora riportando le nostre previsioni in carta semilogaritmica ci si aspetterebbe di trovare una retta, entro il regime in cui è valida l'approssimazione di Shockley, ma oltre a questo, una regione in cui la curva caratteristica della giunzione ora mostra una dipendenza apprezzabilmente più lineare/ohmica dell'intensità di corrente dalla ΔV rispetto a prima, a cui corrisponderebbe il grafico "piatto" di un logaritmo.

3 RISULTATI E ANALISI DATI

4 APPENDICE A: FILTRAGGIO DATI

Il sistema di filtraggio di dati implementato nell'eseguibile si compone di 2 fasi: la prima è l'eliminazione di outliers, la seconda consiste nell'eliminazione di dati non significativi.

4.1 *Introduzione*

In generale possiamo immaginare che ad ogni misura sia associata una forma quadratica rappresentata dalla matrice di covarianza della stessa, cioè che durante la misura si commetta un errore statistico normale noto. Possiamo poi immaginare che il misurando abbia un'altra matrice di covarianza. Nel procedimento proposto questo potrebbe essere teoricamente trattato per esteso, tuttavia non si troverebbe una forma chiusa generale per il problema in questione. Supponiamo allora che gli errori su x e y siano indipendenti e che $\sigma_x^2 := \mathrm{Var}(x)$ sia nota a priori e che abbia distribuzione gaussiana per ogni misura. Questo non è in generale vero, ma questa ipotesi può essere trascurata quando la correlazione tra le varianze delle misure su x e y sono indipendenti dai valori assunti dalle x e y stesse e vi sono "molti" dati entro una deviazione lungo x. In questo caso, infatti, la correlazione viene inclusa nella varianza lungo y e, dal teorema del limite centrale si vede che la non-normalità delle distribuzioni è trascurabile. In ogni modo, nei nostri dati x e y risultano ragionevolmente indipendenti, quindi l'assunzione dovrebbe essere lecita.

4.2 *Procedimento*

Supponiamo, per ogni punto, che la distribuzione sia gaussiana secondo una matrice di covarianza diagonale nella base $\{x, y\}$: allora la densità di probabilità che un punto che abbia misurato x si trovi a tale ascissa x_i si ricava banalmente integrando lungo y a x fissata:

$$\mathrm{d}P = \frac{1}{\sigma_{x_i}\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-x_i)^2}{\sigma_{x_i}^2}} \mathrm{d}x.$$

Dunque, ripetendo più volte la stessa misura, la probabilità

$$P\left(\mid x - x_i \mid \leq \varepsilon\right) = \varepsilon G_{x_i}.$$

dove

$$G_{x_i} := \frac{1}{\sigma_{x_i}\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-x_i)^2}{\sigma_{x_i}^2}}.$$

e $\varepsilon > 0$ è piccolo a piacere. Scegliendo allora solo quelle misure x per cui vale $\mid x - x_i \mid \leq \varepsilon$, queste saranno in numero intorno a:

$$N_i := N_{\mathrm{tot}} \frac{G_{x_i}}{\sum_j G_{x_j}} = N_{\mathrm{tot}} w_i.$$

che definisce implicitamente i pesi w_i con cui si mediano le distribuzioni di probabilità gaussiane G_{x_i}. Allora, detto:

$$G_{y_i} := \frac{1}{\sigma_{y_i}\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(\mu_y-y_i)^2}{\sigma_{y_i}^2}}.$$

Per il principio di massima verosimiglianza siamo quindi interessati a massimizzare la quantità:

$$\mathcal{L} = \prod_{i=1}^{n} \prod_{j=1}^{N_i} G_{y_i} = \prod_{i=1}^{n} G_{y_i}^{N_i}.$$

Per la monotonia del logaritmo il problema equivale a massimizzare la quantità:

$$\ln \mathcal{L} = \sum_{i=1}^{n} \ln {G_{y_i}}^{N_{\text{tot}} w_i} = \frac{N_{\text{tot}}}{\sum_{j=1}^{n} G_{x_j}} \sum_{i=1}^{n} G_{x_i} \ln G_{y_i}.$$

Per cui, a meno di costanti risulta:

$$\ln \mathcal{L} - \text{const.} \propto \sum_{i=1}^{n} -G_{x_i} \ln \sigma_y - \frac{1}{2} G_{x_i} \left(\frac{y_i - \mu_y}{\sigma_y} \right)^2 \tag{1}$$

Imponendo la condizione di stazionarietà rispetto a μ_y si ottiene dunque:

$$\mu_y = \sum_{i=1}^{n} y_i w_i \tag{2}$$

Una volta sostituito in (1) quanto appena trovato per μ_y e imponendo la stessa condizione di stazionarietà rispetto a σ_y si ha:

$$\sigma_y^2 = \sum_{i=1}^{n} (y_i - \mu_y)^2 w_i. \tag{3}$$

Infine è possibile ricavare la varianza di μ_y dalla definizione di valore di aspettazione, riconducendola più volte a integrali di gaussiane di altezze e ampiezze diverse:

$$\text{Var}(\mu_y) = \sum_{i=1}^{n} \left[w_i^2 \sigma_y^2 + \left(\frac{y_i}{\sum_{j=1}^{n} w_j} \right)^2 \frac{\frac{e^{-\frac{(x-x_i)^2}{3\sigma_x^2}}}{\sigma_x \sqrt{6\pi}} + \frac{e^{-3\frac{(x-x_i)^2}{4\sigma_x^2}}}{\sigma_x \sqrt{\pi}} + \frac{e^{-\frac{(x-x_i)^2}{\sigma_x^2}}}{\sigma_x \sqrt{2\pi}}}{\sqrt{2\pi}} \right] = \tag{4}$$

$$= \sum_{i=1}^{n} \left[w_i^2 \sigma_y^2 + \left(\frac{y_i}{\sum_{j=1}^{n} w_j} \right)^2 \frac{e^{-\frac{(x-x_i)^2}{3\sigma_x^2}} + \sqrt{3} \left(e^{-\frac{(x-x_i)^2}{\sigma_x^2}} - \sqrt{2} e^{-3\frac{(x-x_i)^2}{4\sigma_x^2}} \right)}{2\sqrt{3}\pi\sigma_x^2} \right] \tag{5}$$

Riassumendo:
Nella (2) prendiamo una media dei campionamenti intorno ad un' ascissa x in esame, pesata sulla distanza che gli x_i hanno da questa. Effettivamente quello che stiamo facendo è un stima di densità di kernel, per cui consideriamo i punti come -sfocati- da un *blur gaussiano*; dove però nel nostro caso riscaliamo la stima in base al valore assunto da y. Lo stesso ragionamento vale per σ_y^2, si ha una stima della varianza dei dati la variare di y, pesata sulla distanza dai valori studiati. Dunque $\mu_y \pm \sigma_y$ ci dà una stima della distribuzione dei nostri dati.

4.3 *Var* (μ_y)

Mentre σ_y rappresenta la distribuzione dei dati intorno al valor medio μ_y, $\text{Var}(\mu_y)$ ci dà un'idea dell'incertezza che attribuiamo alla miglior stima di y. Questo ci è utile per determinare la convergenza della stima in funzione dei dati acquisiti. Infatti: più la densità dei dati è grande rispetto alla deviazione standard σ_x, più la stima del valore centrale risulta precisa. Graficamente la banda di confidenza è più ristretta dove si concentrano i dati, viceversa tende ad allargarsi dove i dati sono sparsi, a distanze paragonabili a σ_x. Dalla seconda somma nell'espressione (5) segue allora che la stima del valore centrale è statisticamente significativa solo quando si media su un intervallo campionato con almeno qualche punto ogni deviazione σ_x, altrimenti $\text{Var}(\mu_y)$ tende a $+\infty$ come $\sim e^{x^2}$ in assenza di dati, dove non è possibile stabilire con precisione il valore di μ_y.

4.4 *Filtro outliers*

La parte più semplice nel filtraggio dati consiste nello scartare tutti quei punti che distano da μ_y più di una soglia arbitraria k di deviazioni standard dalla media σ_x, ovvero quegli y_i per cui, indipendentemente dal modello di fit vale $|y_i - \mu_y| \geq k\sigma_y$.

4.5 *Filtro dati non significativi*

Supponiamo di avere 2 set di dati fatti con diverse resistenze, il primo con una resistenza bassa, il secondo con una alta: Il primo set esplorerà la regione ad alta corrente, mentre il secondo la regione di basse correnti. Però il primo insieme conterrà, in generale, anche campionamenti delle zone basse, ma su queste fornirà dei valori meno significativi: Esponiamo dunque il criterio sviluppato per ridurre l'influenza di questi punti meno significativi sulla ricerca dei parametri di best-fit e sulla rappresentazione finale dei dati.
Per capire se in un certo punto i dati di A sono significativi, calcoliamo la misura di significatività che abbiamo sviluppato in (5): $\mathrm{Var}(\mu_y)$ di A e di B. Perciò se $\mathrm{Var}(\mu_y)$ di A è maggiore di $q\mathrm{Var}(\mu_y)$ di B, con q arbitrario, questo indica che i dati di A ci stanno dando meno informazioni rispetto a quelli di B. A questo punto è sufficiente controllare tutti i punti scorrendo su tutte le combinazioni di set per eliminare i dati non significativi, che rendono meno immediata l'interpretazione il grafico.

5 APPENDICE B: METODO DI FIT

Partiamo dall'osservazione che la tensione ai capi del diodo può essere scomposta in ΔV_r e ΔV_d come illustrato in figura:

dove il resistore a sinistra indica proprio la resistenza di un diodo reale, mentre il diodo alla sua destra rappresenta una giunzione bipolare PN, ossia il diodo ideale descritto dall'equazione di Shockley. Dunque ci aspettiamo che la tensione V campionata con `Teensy` ai capi del diodo rispetti la legge:

$$V = \Delta V = \Delta V_d + \Delta V_r = \eta V_T = \ln \frac{I + I_0}{I_0} + RI \tag{6}$$

All'interno dello script `fit_file.py`, voltages rappresenta il vettore delle ascisse dei punti campionati, mentre currents rappresenta l'analogo per le ordinate. Le incertezze associate a questi due all'interno dello script prendono i nomi `voltageErrs` e `currentErrs`. Ingenuamente potremmo essere tentati di effettuare una sorta di fit "inverso" (delle x in funzione delle y). Tuttavia una tale operazione non è affatto banale, a causa dei valori negativi/nulli che debitamente si riscontreranno all'interno del logaritmo nella legge (6), che possono facilmente portare al fallimento dell'algoritmo di fit. É per questo che si è dovuto definire il modello "inverso" nella forma di una legge di natura approssimata per ricorsione, grazie al metodo di Newton. Al livello di implementazione si è quindi dichiarata una funzione:

```
def curr(V, I0, nVt, R):
    v = V
    for i in range(Nstep):
        a = deriv_errFun(v, I0, nVt, R)
        v = v - errFun(v, V, I0, nVt, R) /a
    return (V - v)/R
```

Questa verrà in seguito usata come modello per la corrente in funzione della tensione ai capi del diodo reale nella routine di fit per minimi quadrati: implementato in `Python` mediante la funzione *curve_fit* dal modulo `optimize` interno alla libreria `Scipy`[2]. In pratica questa funzione cerca di -linearizzare- la relazione tra I e V nell'intorno di un determinato V. Come sappiamo infatti dal metodo di Newton, detta $f(x)$ una funzione tale che $f(x) = 0$ e, dato un valore iniziale, tale che $f(x[0]) = \alpha$ generico, se procediamo con un numero di iterazioni sufficienti, sappiamo che la relazione di ricorsione che segue ci fornirà un risultato soddisfacente per la nostra x:

$$x[0] = \alpha \tag{7}$$

$$x[N+1] = x[N] - \frac{f(x[N])}{f'(x[N])} \tag{8}$$

dove N indica il livello di ricorsione raggiunto nell'approssimazione di x. Nel nostro caso x sarà quella tensione v per cui vale l'identità:

$$I(V) := I_0^{\frac{\Delta V}{\eta V_T}} = \frac{v - V}{R} \tag{9}$$

che identifica un tratto lineare caratterizzato da una retta con coefficiente angolare negativo pari a $-\frac{1}{R}$ ed un intercetta di $I = v/R$ sull'asse delle y, che come è facile dedurre, altro non è che la retta di carico del diodo. Facendo così stiamo quindi cercando il punto di lavoro v per cui la corrente può essere scritta come $\frac{v-V}{R}$ (corrente di lavoro), come si evince dal return nella funzione di fit `curr(V, parametri liberi)`. Sulla base di ciò è facile dedurre che le altre due funzioni che vengono richiamate dentro `curr` possono essere scritte nella forma:

```
def sck(V, I0, nVt):
    return I0*(pylab.exp(V/nVt) - 1)

def errFun(V, V0, I0, nVt, R):
    return sck(V, I0, nVt) + (V - V0)/R

def deriv_errFun(V, I0, nVt, R):
    return I0 / nVt * pylab.exp(V/nVt) + 1./R
```

dove `errFun` e `deriv_errFun` sono, rispettivamente, la relazione da minimizzare specificata dalla legge (9) e la propria derivata. Quindi il fit è stato effettuato secondo le consuete modalità, adottando in primo luogo `curr` come funzione di fit e `currentErrs` come errori sui punti (gli errori sulle y). In seguito, sono stati propagati anche gli errori `voltageErrs` relativi alle x sulle y secondo la consueta modalità.

RIFERIMENTI BIBLIOGRAFICI

[1] P. Stoffregen. Teensy 3.2 development board. Sherwood, Oregon, USA. [Online]. Available: https://www.pjrc.com/teensy/teensy31.html

[2] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and S. . . Contributors, "SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python," *arXiv e-prints*, p. arXiv:1907.10121, Jul 2019.